Artificial Intelligence in Diagnostic Radiology: Where Do We Stand, Challenges, and Opportunities

Ahmed W. Moawad, MD,† David T. Fuentes, PhD,‡ Mohamed G. ElBanan, MD,§ Ahmed S. Shalaby, MD,* Jeffrey Guccione, MD,|| Serageldin Kamel, MD,¶ Corey T. Jensen, MD,* and Khaled M. Elsayes, MD**

Abstract: Artificial intelligence (AI) is the most revolutionizing development in the health care industry in the current decade, with diagnostic imaging having the greatest share in such development. Machine learning and deep learning (DL) are subclasses of AI that show breakthrough performance in image analysis. They have become the state of the art in the field of image classification and recognition. Machine learning deals with the extraction of the important characteristic features from images, whereas DL uses neural networks to solve such problems with better performance. In this review, we discuss the current applications of machine learning and DL in the field of diagnostic radiology.

Deep learning applications can be divided into medical imaging analysis and applications beyond analysis. In the field of medical imaging analysis, deep convolutional neural networks are used for image classification, lesion detection, and segmentation. Also used are recurrent neural networks when extracting information from electronic medical records and to augment the use of convolutional neural networks in the field of image classification. Generative adversarial networks have been explicitly used in generating high-resolution computed tomography and magnetic resonance images and to map computed tomography images from the corresponding magnetic resonance imaging. Beyond image analysis, DL can be used for quality control, workflow organization, and reporting.

In this article, we review the most current AI models used in medical imaging research, providing a brief explanation of the various models described in the literature within the past 5 years. Emphasis is placed on the various DL models, as they are the most state-of-art in imaging analysis.

Key Words: artificial intelligence, machine learning, neural networks, convolutional neural network, recurrent neural network, generative adversarial networks, quality control, workflow organization

(*J Comput Assist Tomogr* 2022;46: 78–90)

The rapid evolution of artificial intelligence (AI) is one of the most important technological developments of the current decade, affecting nearly all aspects of life. In health care, and in particular medical imaging, a large number of AI applications have been proposed, studied, and implemented. Computer-aided diagnosis (CAD), for example, represents the most widely known application of AI in the medical imaging domain. Implementation of clinical applications of CAD in medical imaging dates back to the early 1960s, primarily in the field of breast imaging. Nonetheless, its roles have been limited to the detection of lesions and supporting the interpretations of radiologists.[1,2] Currently, however, we are witnessing a shift in this paradigm with the development of new and robust algorithms for AI and machine learning (ML).

Artificial intelligence, at its core, can be defined simply as the field of sciences that investigates methods to "make machines intelligent," with the goal of developing machines that can perform tasks that historically need human intelligence.[3] This includes learning from data and human language as well as emotions. Machine learning is a subdomain of the vast reach of AI that deals with "learning from data without being explicitly programmed."[4] A widely accepted definition proposed by Mitchell[5] in 1997 describes ML as the field in which "a computer program is said to learn from experience E with respect to some class of tasks T and performance measure P if its performance at tasks in T, as measured by P, improves with experience E."

Deep learning (DL) is a subclass of ML that has demonstrated breakthrough performance in fields of image classification and object recognition as well as natural language processing. Its robust applications in medical imaging have largely been aided by the development of supercomputers and leaps in data availability. The relationships between AI, ML, and DL are summarized in Figure 1.

DL, A NEW PARADIGM IN AI

Machine learning is the category of computer learning from data without explicitly being programmed. This learning ultimately leads to the creation of predictive models that can be used to make decisions on new data that the model has not encountered before. Traditional ML starts with the identification of important characteristic features that are believed to be helpful in solving the problem of interest, a process called "feature engineering." These features are built into a predictive model that uses the given data to try to solve the problem. Two popular predictive models

From the *Department of Abdominal Imaging, The University of Texas MD Anderson Cancer Center, Houston, TX; †Department of Diagnostic and Interventional Radiology, Mercy Catholic Medical Center, Darby, PA; ‡Department of Imaging Physics, Division of Diagnostic Imaging, The University of Texas MD Anderson Cancer Center, Houston, TX; §Department of Diagnostic and Interventional Radiology, Yale New Haven Health, Bridgeport Hospital, CT; ||Department of Diagnostic and Interventional Imaging, The University of Texas Health Sciences Center at Houston, Houston, TX; and ¶Clinical Neurosciences Imaging Center, Yale University School of Medicine, New Haven, CT.

Received for publication May 19, 2021; accepted August 17, 2021.

Correspondence to: Khaled E. Elsayes, MD, Department of Radiology, The University of Texas MD Anderson Cancer Center, 1400 Pressler St, Houston, TX 77030 (e-mail: kmelsayes@mdanderson.org).

Ahmed Moawad ORCID 0000-0001-6860-1513
David Fuentes ORCID 0000-0002-2572-6962
Mohamed Elbanan ORCID 0000-0002-9517-5764
Ahmed Shalaby
Jeffrey Guccione ORCID 0000-0001-8628-7062
Serageldin Kamel ORCID 0000-0002-0046-4337
Corey Jensen ORCID 0000-0001-6540-8961
Khaled Elsayes ORCID 0000-0003-0184-0717

The authors received no financial support for the research, authorship, and/or publication of this article.

The authors declare no conflict of interest.

Author Contributions: A.W.M.: design of the work, interpretation of data for the work, drafting the article, final approval of publication; D.T.F.: substantial contributions to the article conception, critical revision of the draft, final approval of publication; M.G.E.: design the work, critical revision of the draft, interpretation of data for the work, final approval of publication; A.S.S.: design of the work, critical revision of the draft, final approval of publication; J.G.: drafting the work, critical revision, final approval of publication; S.K.: design of the work, critical revision of the draft, final approval of publication; C.T.J.: substantial contributions to the article conception, critical revision of the draft, final approval of publication; K.M.E.: substantial contributions to the article conception, critical revision of the draft, final approval of publication.

DOI: 10.1097/RCT.0000000000001247

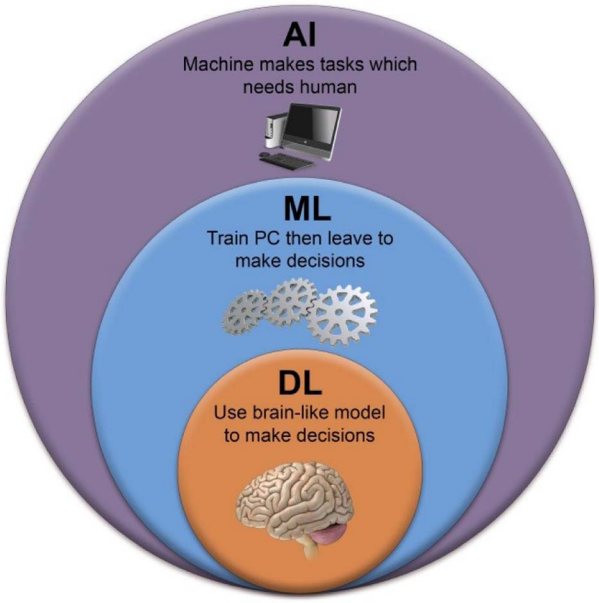

FIGURE 1. A chart demonstrates the relationship between AI, ML, and DL. Deep learning is considered a subbranch of ML which depends on neural networks for implementation, whereas ML is a branch of AI that enables the machine to make decisions. PC, personal computer. Figure 1 can be viewed online in color at www.jcat.org.

are support-vector machines and random decision forests. One problem with the ML approach is that the characteristic features might not be well known beforehand, or at least not comprehensively understood. This is particularly applicable in medical imaging.

Predictive models in DL are based on building blocks called "perceptrons" or artificial neurons, which are interconnected, together with "activation functions," to create what is known as a neural network (Fig. 2). Multiple hidden layers of these neural networks comprise the architecture of DL. With the vast computational power enabled by modern graphical processing units and the readily available digital data ready for analysis, DL algorithms learn directly from data to automatically extract important characteristic features and use this information to solve a problem.

Differentiating benign from malignant tumors is an example of an object classification task that demonstrates the difference between the traditional ML approach and current DL. If there is a training set of benign tumors and malignant tumors, each of them has specific feature vectors. For example, benign tumors are usually smaller and homogenous with well-defined borders, whereas malignant tumors are usually larger with a heterogeneous appearance and ill-defined borders that often invade the surrounding tissue. The traditional ML approach extracts these feature vectors from the image and selects the optimal features as input to the learning algorithm to train the statistical model. The computer then uses this trained model to classify masses on new images as either benign or malignant. This contrasts with DL where the neural network is trained with labeled images (either benign or malignant) as an input; through multiple intermediate layers, the neural network extracts the features learned from the images during the training process then predicts the outcome via a fully automated algorithm[6] (Fig. 3).

Although the DL approach is more structurally complex and computationally demanding, it produces much greater classification accuracy and better performance than the traditional ML approach. Thus, DL and its neural network architecture have become the dominant robust approach to AI in medical imaging applications.[6,7]

Deep learning models differ according to the architecture of their neural networks, mathematical operations to be performed, and the task of interest to be solved by the model.

We reviewed these scientific databases (Scopus, Web of Science, PubMed, IEEE Xplore, and ScienceDirect) with keywords including deep learning, neural networks, convolutional neural networks. We limited our search results to original research performed during the last 5 years to avoid redundancy, then we filtered the search results according to the task (either classification, segmentation, or object detection) and neural network used. We selected the most cited articles and neural network used in the field of diagnostic radiology. There were thousands of published articles using different technologies in imaging analysis. We selected some of them to ensure a comprehensive overview of all technologies used. We will introduce 3 of the most important neural network architectures being used in radiology along with examples of their application in different medical imaging tasks. These include convolutional neural networks (CNNs), recurrent neural networks (RNNs), and generative adversarial networks (GANs).

Convolutional Neural Networks

Convolutional neural networks were famously introduced by LeCun et al[8] in the 1990s and based on the early 1960s works of Hubel and Wiesel[9] on how the visual cortex functions. A basic CNN consists of 3 main types of layers: convolution, a pooling layer, and a fully connected network/layer. Convolution is a specialized pattern of mathematical operation that is used to extract certain aspects of the picture (eg, detection of vertical lines or enhancement of edges). The operation is done using a small array of numbers, usually a 3×3 matrix, called the filter (or kernel), and results in different feature maps, such as a map for vertical edges and a map for horizontal edges. This enables the machine to extract low-level patterns from the image, which is then fed into a pooling layer to enable down-sampling without loss of the main feature of the image itself. Variable numbers of convolution and pooling layers are usually followed by the fully connected layer where every node is connected to the following layer, which will be organized according to the architecture used (Fig. 4).[10] Further

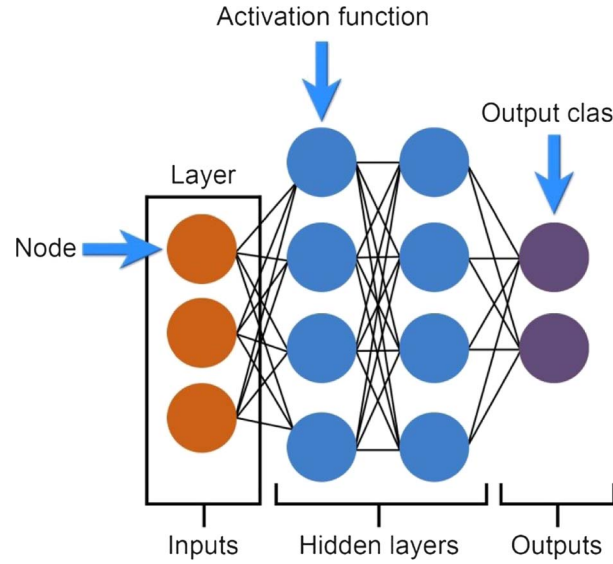

FIGURE 2. Schematic diagram of an artificial neural network, which shows the components of neural network (input, output, and hidden layers). Each layer consists of nodes with certain activation functions. Figure 2 can be viewed online in color at www.jcat.org.

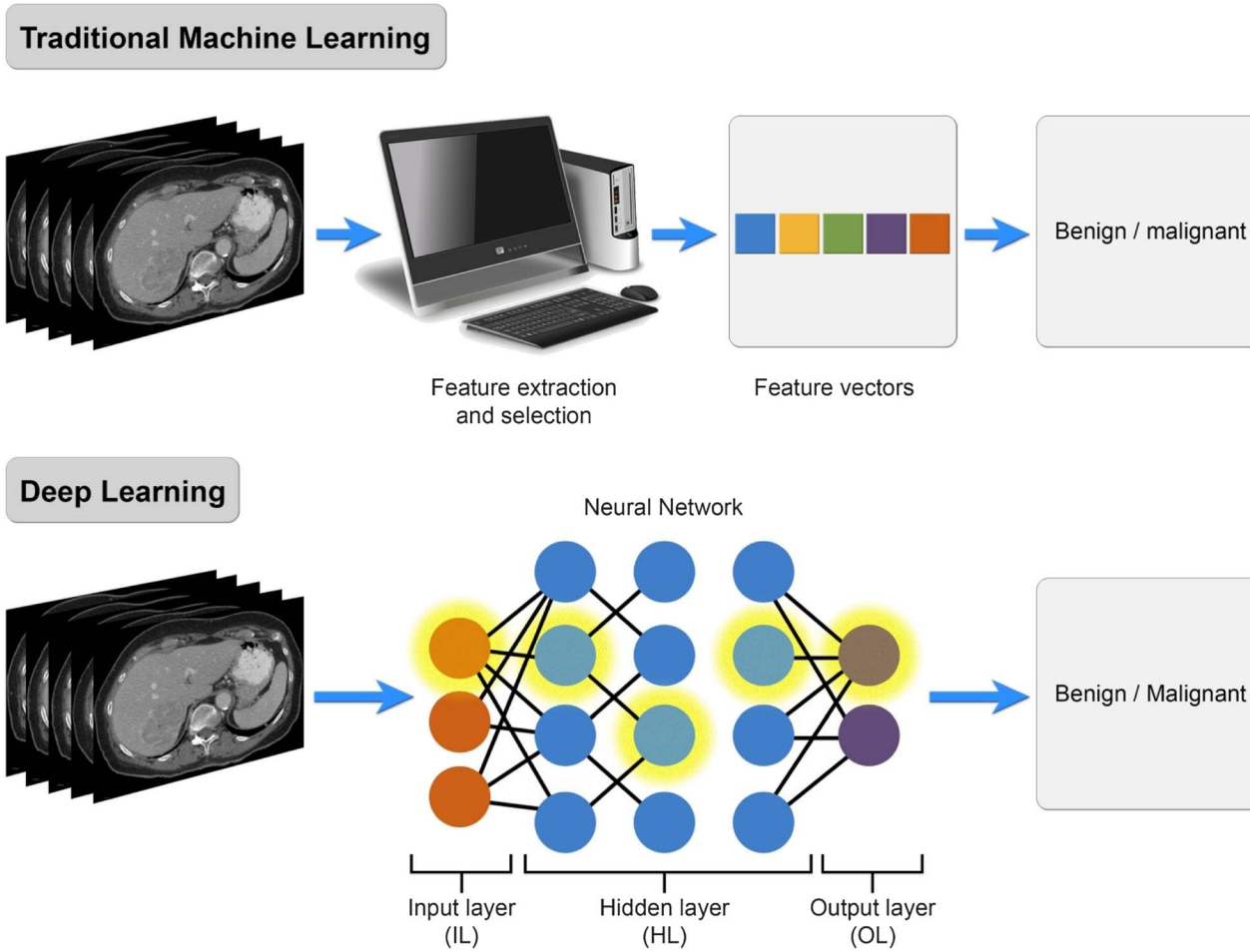

FIGURE 3. Schematic diagram showing the difference between traditional ML and DL approaches for tumor classification. In the ML approach, imaging features are extracted and selected, and then the features are incorporated into selection algorithms. With the DL approach, neural networks receive images and automatically classify the tumor as either benign or malignant. Figure 3 can be viewed online in color at www.jcat.org.

research into the updating and integration of CNN with other classifiers in image analysis has produced deeper, more complex, and more robust neural networks that have surpassed human performance in image classification and object detection tasks.[11] This has been largely aided by the recent advances in parallel computing and readily available labeled data sets for model training. There are several architectures of CNN with different parameters, filters, and number of layers. They all shared the same idea of convolution explained previously and are commonly used architectures in the medical imaging field.

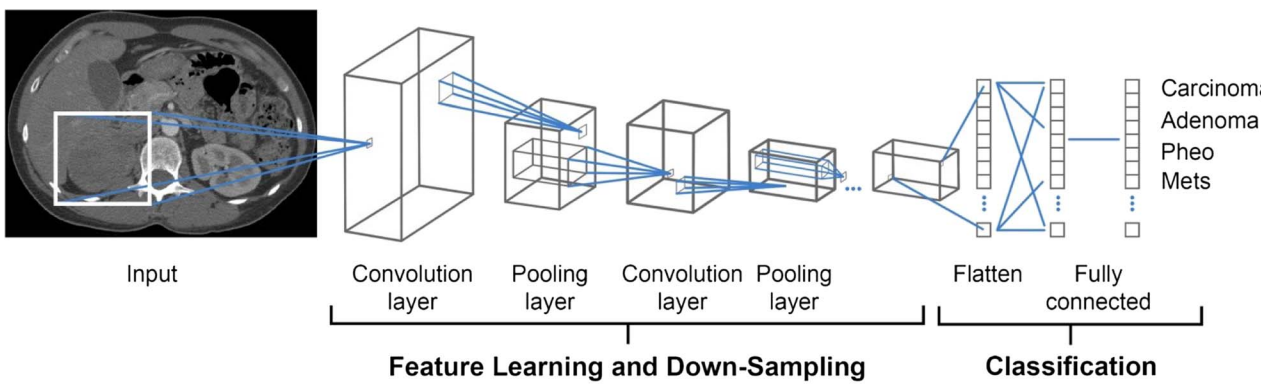

FIGURE 4. Schematic diagram of the classification task using a CNN. Notice the general components of the CNN: a convolution layer, a pooling layer, and a fully connected layer. The output of such NN is the classification of the nature of the tumor. Figure 4 can be viewed online in color at www.jcat.org.

- LeNet: LeNet, built in 1998, is the oldest CNN, but its architecture became the backbone of the following descendants. LeNet consists of 2 convolutional layers and 2 average pooling layers (which select average values of all pixels). The output of these 4 layers serves as an input to 2 successive fully connected layers. The final layer comprises 10 classifier outputs to determine hand-written digits from 0 to 9. Notably, LeNet failed in larger object classes in computational efficiency and accuracy.[8]
- AlexNet: This follows the same architecture as LeNet but with 5 convolutional layers, instead of 2, and 2 max-pooling layers. This was followed by 2 fully connected layers. AlexNet has ~60 million parameters to handle, so it is split into 2 pipelines, which are trained separately yet simultaneously.[12]
  - oVisualization of different layers outputs in AlexNet enables fine-tuning of its hyperparameters to improve the results. This modified network was called ZFNet, which is the same as AlexNet with simpler fine-tuning.[13]

- VGGNet/OxfordNet: There are several configurations of this architecture, depending on the number of layers. 11, 13, 16, and 19 layers exist, including 3 fully connected layers. Despite its simplistic architecture compared with others, VGGNet is very computationally demanding, requiring 130 to 140 million parameters to handle. Conversely, it is therefore a good fit for large classifiers with thousands of output classes.[11]
- GoogleNet: This is dependent on inception modules, in which all convolutional layers are implemented simultaneously and the result of all of them will be contained in the output layer. Building multiple blocks of the inception module results in a very deep network, which retains its accuracy, despite fewer computations.[14]
- DenseNet: As the name describes, DenseNet is a densely packed CNN where layers receive information from all the previous ones while avoiding the presence of redundant layers and unnecessary calculations. This results in dense blocks, which form the backbone of DenseNet.[15]
- ResNet: Residual networks (ResNet) depend on the idea of skip connections, where the output of the first layers skips over a couple of layers. This solves the "vanishing gradient" problem, which is a premature stoppage of model learning after several computations due to shrinkage of the gradient (loss function) to zero, preventing weights from being updated.[16]
  - oSqueeze-and-excitation (SE) networks were recently developed to balance the weights of each convolution layer, rather than constant weight in the previous architectures. Adding SE network to ResNet (SE-ResNet) has yielded the best-known architecture for imaging classification, with a 25% lower error rate than ResNet and a very low computational cost.[17]

Recurrent Neural Networks

Recurrent neural networks are a type of neural network that is organized into successive self-looping nodes. The sequential nodes are identical but executed at different time points; thus, they can be used as a sort of memory.[18] This type of neural network has been investigated in various tasks in radiology report processing, which depends on sentence context and sequential information.[19]

Recurrent neural networks are characterized by the "recurrent" state in which the output from a layer will be an input to the following one, resulting in a final output that is dependent on all the previous inputs. This contrasts with a traditional neural network where all the inputs and outputs are independent of each other and are dependent on the weight of each layer. The hidden layer(s) in RNNs are responsible for this dependence, which acts as "memory" to remember all inputs that were previously calculated. Training of an RNN is complex, yet the proposed output by the network will be compared with the actual output, and the difference is propagated backward to learn the network in a back-propagation manner. In addition, RNN requires fewer computations and complex calculations as parameters because each input is the same along the whole process. Recurrent neural network has been applied in the field of radiology mainly in information extraction from electronic medical records and narrative radiology reports. There are 3 variants of RNNs, described hereinafter.

- Long short-term memory (LSTM): This is considered the most significant variant, as it solves the problem of long-term dependencies. Traditional RNNs cannot efficiently handle large amounts of inputs, such as predicting the last word in a long sentence. LSTM can handle this efficiently using the cell state, which is gated units that can control which information can be kept and what to discard.[20]
- Bidirectional RNNs: These not only learn from the past to predict the future but can adjust and update the past based on the future. For example, in a speech recognition task, usually, the network will depend on the following words to fix the first one according to the context. This bidirectional learning usually occurred using hidden layers of opposite directions.[21]

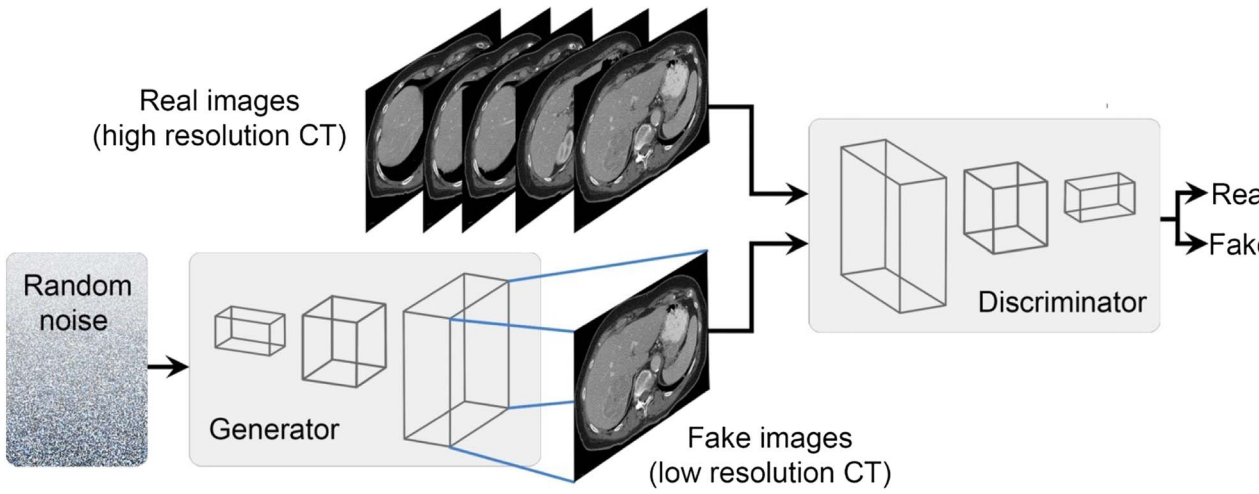

FIGURE 5. Schematic diagram of a GAN, showing the generator and discriminator divisions of the NN. Figure 5 can be viewed online in color at www.jcat.org.

- Deep RNNs: These contain multiple hidden layers, which were found to explore more details than shallow RNNs, but ultimately require more training data.

Generative Adversarial Networks

The GAN has been described as the most interesting achievement in ML algorithms in the last decade. Generative adversarial networks, first introduced by Goodfellow and colleagues[22] in 2014, have come to be regarded as the most robust generative algorithm for generating images from random noise, closely mimicking real images. A classic GAN is composed of 2 separate networks: (*a*) a generative network, which converts noise vectors into an image, and (b) a discriminator, which compares the produced image (fake) with the original image one to classify it as fake or not. Figure 5 shows a simplified demonstration of GAN networks.

This is a dynamic process, as both networks improve their performance until the generated image is nearly identical to the original one.[22] An important application of GAN is the generation

TABLE 1. Summary of Studies Using DL Classifiers in Diagnostic Radiology

Author (Year)	Training Set	Architecture	Classification Task	Performance (ROC/AUC)	Comments
Ma et al (2017)[36]	4782 cases (15,000 images)	VGG	Thyroid nodules classified into benign or malignant using ultrasonography	0.89	Two CNNs were used to extract both low- and high-level features.
KV et al (2019)[37]	389 images, augmentation	VGG-19	Grading of the glioma (grades I–IV) in T2WI	0.982	
Song et al (2018)[38]	444 images, augmentation	Modified VGG	Classification of prostate MRI lesions to either benign or malignant	0.944	Layer numbers were reduced because of simpler task than ILSCV.
Prevedello et al (2017)[39]	385 cases (2006 images)	GoogleNet	Classification of hemorrhage, mass effect, hydrocephalus, and acute infarct in noncontrast CT	0.81–0.91	
Kim et al (2018)[40]	1.389 image, augmentation	GoogleNet	Classification of fractures in lateral plain wrist radiographs	0.95	
Taylor et al (2018)[41]	1990 images	VGG-19, GoogleNet	Classification of plain AP chest radiographs as either containing a moderate/large pneumothorax or as normal	0.96	Testing of this model was done on ChestX-ray NIH data set.
Chung et al (2018)[42]	1702 images, augmentation	ResNet	Classification of fracture types in the proximal humerus using AP shoulder plain radiographs	0.96	CNNs outperform general physicians and specialized orthopedists.
Schwyzer et al (2018)[43]	100 patient (3543 images)	ResNet	Classification of lung nodules in PET-CT into benign or malignant	0.989	They used activation heat maps for detection of nodules.
Lee et al (2019)[44]	160 case (786 images)	ResNet, GoogleNet	Cervical lymph node classification in neck CT scans as either metastatic or benign, in patients with thyroid cancer	0.90–0.95	They used class activation maps for lymph node detection.
Lakhani et al (2017)[45]	857 image, augmentation	Ensemble of AlexNet and GoogleNet	Classification of plain AP chest radiograph as either TB or not	0.99	
Dalmis et al (2019)[46]	576 lesion	DenseNet	Breast lesion classification into benign or malignant in breast MRI (DWI, T2WI, ultrafast dynamic phase)	0.852	They combined traditional ML classifiers with CNN.
Corrhea et al (2018)[47]	60 patients (1611 images)	Fully connected neural neurons (100 hidden layers)	Lung lesion classification in children to either pneumonia or normal in lung ultrasonography		The ultrasound images are divided into multiple vectors after image processing.
Rodriguez-Ruiz et al (2019)[48]	>9000 image		Assisting radiologists in breast lesion classification into benign and malignant using mammography	0.89	The system used is Transpara, which is trained on >9000 mammograms.
Hamm et al (2019)[49]	434 cases, augmentation		Hepatic lesion classification into simple cyst, cavernous hemangioma, FNH, HCC, ICC, or CRC metastasis using MRI	0.91	They also used this model to classify these lesions into their LI-RADS categories.

AP indicates anteroposterior; CRC, colorectal carcinoma; DWI, diffusion-weighted images; FNH, focal nodular hyperplasia; HCC, hepatocellular carcinoma; ICC, intrahepatic cholangiocarcinoma; LI-RADS, Liver Imaging Reporting and Data System; PET, positron emission tomography; ROC/AUC, receiver operator characteristic curve/area under the curve; T2WI, T2-weighted images; TB, tuberculosis.

of high-resolution images from lower-resolution images, known as superresolution GAN.[23]

Generative adversarial network is used in the field of radiology to enhance image resolution without upgrading machines or increasing the radiation dose of computed tomography (CT). Low-dose CT is usually used for high-radiation-dose imaging protocols, such as coronary CT angiography, for patients who undergo multiple follow-up CT scans, and for children and neonates. Image noise is inherent to that approach, with resultant reductions of spatial and temporal resolution, which negatively affect image interpretation. Many studies have examined the use of GAN to generate images with higher resolution from noisy image inputs, supporting its usefulness.[24,25] Another strategy for lowering the radiation dose, especially for patients who must undergo multiple follow-up studies, is to generate CT images from corresponding magnetic resonance images by nonlinear mapping models using GAN algorithms.[26]

Generative adversarial network research is also being carried out with magnetic resonance imaging (MRI) because improvements in MRI resolution require longer scan time and increased costs, as well as reductions in spatial coverage and signal-to-noise ratio. Recent studies have used GAN to generate MRI with superresolution, with 4× better resolution and 6× faster algorithms than other DL models.[27]

Combining a GAN with a CNN has been used to improve the accuracy of classification and segmentation models by the generation of nonrealistic, yet very similar, images for training sets and manual segmentation.[28–30]

USES OF AI IN MEDICAL IMAGES ANALYSIS

Image Classification

Image classification is the assignment of an image into a certain category, for example, to classify a lung nodule as benign or malignant.[31] Artificial intelligence has been used to perform this task efficiently owing to its ability to extract relevant features and perform as needed for the task of image classification using a large preclassified (ie, labeled) training data set. Neural networks, especially CNNs, have demonstrated tremendous potential in image analysis because it outperformed all other image classification algorithms in the ImageNet Large Scale Visual Recognition Challenge since 2012.[12]

Some challenges in applying CNN in medical image classification is the huge amount of data required to train the network. Because of regulations regarding institutional rules and copyright, high-quality radiological images are often not available for public use.[32] In addition, manual image classification (labeling) is a time-consuming process requiring a dedicated, experienced radiologist.

This scarcity of labeled medical image data has the potential to halt the progression of DL algorithms in medical image analysis. However, one approach that has been used to overcome this limitation is "transfer learning."[33] *Transfer learning* uses a pretrained neural network that was trained for image classification on, for example, nonmedical images, modifying certain layers of the neural network architecture to fit medical images and the objects of interest. The concept behind the use of transfer learning in image classification is that different visual tasks—whether natural or medical images—share similar levels of processing, such as edge detection and capturing lines. *Data augmentation* is another approach to expand a limited data set while maintaining a high classification performance. Image processing tools like flipping, rotation, and reorientation can be used to enlarge imaging sets.[34]

Applying the same algorithms that were pretrained in huge nonmedical classification databases, like ImageNet, has been criticized by many investigators, given differences in data set size and class number (which is much larger in ImageNet than for available medical imaging data).[35] The outstanding performance of CNNs in ImageNet Large Scale Visual Recognition Challenge, a former classification competition led by ImageNet, prompted the scientific community to use their architectures in the classification of medical images. Table 1 summarizes the most recent and important studies for the use of DL in medical image classification.

The problem of radiological images scarcity has led to the initiation of development of a large-scale radiological imaging database imitating ImageNet, with highly labeled radiological images (either 2D or 3D) that can be used to develop better algorithms for detection, segmentation, and classification. Examples include MR-Net for magnetic resonance images of the knee, MURA for musculoskeletal radiographs, and ChestX-ray for chest radiographs.

Anomaly/Lesion Detection

Object detection means identification of the presence or absence of an object of interest as well as its localization within an image. Unlike classification tasks, object detection is at its early stages and computer vision cannot yet compete with the performance of an actual human. Common objects in context are a large-scale challenge for object detection and segmentation, and it was developed by Microsoft to enhance and motivate contributors for object detection algorithms innovation. The challenge aims to develop faster and more robust neural networks to supersede human performance in the next few years for real-time object detection.[50] The output of these algorithms is traditionally the original image with a localization object (eg, bounding box surrounding the identified object, or heat maps poorly delineate it; Fig. 6). The more recent algorithms that have been used in medical imaging are further discussed hereinafter.

- Region-based algorithms: Regional proposed networks (RPNs) are the backbone of these algorithms, which proposes regions in an input image to feed to the classifier network.

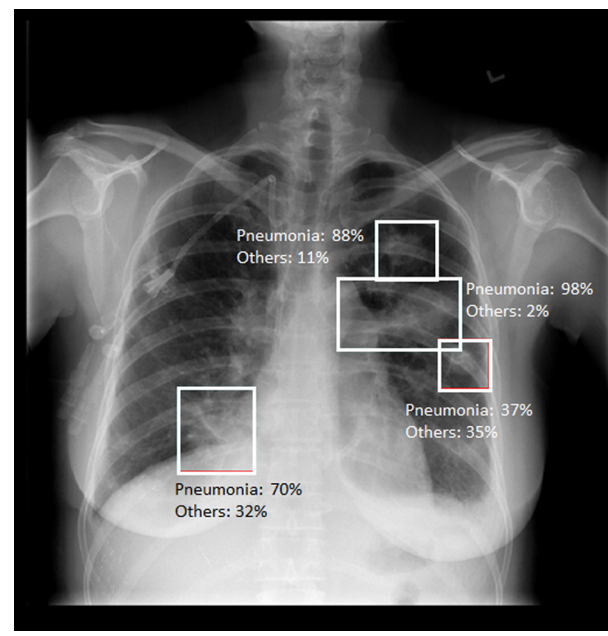

FIGURE 6. Chest x-ray showing bilateral patchy opacities. A CNN is used to detect such lesions in bounding box and then classify these lesions as either pneumonia or others with the probability of each class.

oFaster Regional–CNN (R-CNN): This is the strongest update of R-CNN, which uses CNNs for classification. In Faster R-CNN, the input image is converted into a feature map by using CNNs, based on the various architectures discussed previously. The feature map used by RPNs is to propose different bounding boxes, which then are classified by a classification R-CNN. This functions to label that bounding box into its proper class by using, and adjusting, the coordinates of the bounding box to better fit the object. The main advantage of faster R-CNN is its test-time speed, which is far less than with R-CNN, hence the name.[51]
oMask R-CNN: This is an extension of faster R-CNN with the same architecture, but an additional binary mask classifier used to classify each pixel in the bounding box as related to the object or not. Therefore, Mask R-CNN segments the object at the pixel level. The binary classifier input is the proposed regions suggested by RPN.[52]

- Single-shot detectors: Regional based algorithms work on 2 levels, first to propose regions and then classify them. This makes this method a little bit slower. Single-shot detectors are much different from the previous ones because they look at the whole input image at once. Using CNNs, it then divides the image into grids and draws appropriate bounding boxes with the probability of an object within the bounding. The final output layer then selects the highest probability bounding box to report. The problem in these algorithms is the imbalance between the

TABLE 2. Summary of Studies Using DL for Object Detection in Diagnostic Radiology

Author (Year)	Training Set	Architecture	Object Detection Task	Performance (ROC/AUC)
Gan et al (2019)[60]	2040 images, augmentation	Faster R-CNN for object detection, then GoogleNet for classification.	Detection of fractures in AP wrist radiographs	0.96
Tuzoff et al (2019)[61]	1352 image	Faster R-CNN based on VGG-16 architecture	Detection and numbering of teeth in panoramic teeth radiography	0.9945*
Wang et al (2020)[62]	192 patients, augmentation (11,658 images)	Faster C-NN and VGG-16	Detection of circumferential resection margins in rectal carcinoma MRI (T2WI)	0.953
Liu et al (2017)[63]	120 patients	Faster R-CNN based on ZFNet architecture	Localization of colitis using bounding box and heat map in contrast CT	0.984
Nasrullah et al (2019)[64]	1018 images	Faster R-CNN and U-net (LIDC-IDRI data set)	Pulmonary nodule localization and classification in CT	0.94
Lindsey et al (2018)[65]	135,845 image	Modified U-net architecture	Detection of fractures in plain radiographs using heat maps	0.99
Chang et al (2018)[66]	10,159 case	Mask R-CNN	Detection of hemorrhage in unenhanced CT brain using bounding boxes	0.989
Filice et al (2019)[67]	10,902 image	Mask R-CNN	Detection and localization of chest abnormalities (pneumothorax and chest tubes) in AP chest radiographs	0.87†
Fritz et al (2020)[68]	18,520 studies	GoogleNet	Localizing the meniscal tear using heat maps in MRI	0.961
Kooi et al (2017)[56]	12,138 patients, augmentation	OxfordNet	Detection of breast lesions in mammography	0.93
Kim et al (2018)[69]	26,631 case	ResNet	Localization of breast lesion using specific score and heat map using mammography	0.906
Pan et al (2019)[57]	25,684 case	Ensemble of RetinaNet and R-CNN (RSNA Pneumonia detection challenge)	Localization of pneumonia in AP chest plain radiographs	0.254‡
Wang et al (2019)[71]	276 patients (5007 images)	YOLO v2 integrated with ResNet and pertained on VOC data sets)	Detection and localization of thyroid nodule during ultrasonography	0.90
Cao et a. (2017)[74]	515 patients	SSMD using either ZFNet or VGG-16	Localization of breast lesion using bounding box in US	0.96
Li et al (2017)[72]	888 patients	SSDM and FPN (LUNA grand challenge data set)	Localization of lung nodules in axial CT scans	0.892

*The reported performance metrics is precision for tooth detection.

†The reported performance metrics is sensitivity of pneumothorax detection.

‡The reported performance is the RSNA score metric. This study has the highest ranking until the time of manuscript writing. The score metrics combines true positive, false positive, and false negative and threshold them to obtain intersection over union, which is the agreement between ground truth and network segmentation.

AP indicates anteroposterior; FPN, feature pyramidal network; R-CNN, regional CNN; ROC/AUC, receiver operator characteristic curve/area under the curve; RSNA, Radiological Society of North America; SSMD, single-shot multibox detector; T2WI, T2-weighted images; YOLO, You Only Look Once network.

TABLE 3. Summary of Studies Using DL for Segmentation Task in Diagnostic Radiology

Author (Year)	Training Set	Architecture Used	Object Segmentation	Results (DSC)
Yao et al (2018)[74]	48 cases, data augmentation	Modified U-net with added dilated CNN	Segmentation of the hematoma region in CT	0.62
Park et al (2019)[75]	611 examinations, data augmentation	Modified U-net architecture (HeadXNet)	Pixel-level segmentation of brain aneurysm in CT angiography	0.932*
Blanc-Durand et al (2018)[76]	26 patients, data augmentation	3D U-net architecture	Segmentation of gliomas in PET-CT	0.8231
Liu et al (2020)[77]	590 patients	U-Net	Segmentation of pulmonary embolism and calculate its volume in CT pulmonary angiography	0.926*
Li et al (2018)[78]	LiTs data set (131 patients)	Hybrid Dense U-Net	Hepatic tumor segmentation using CT scans	0.72
Milletari et al (2016)[79]	50 cases, data augmentation	V-Net	Segmentation of prostate in prostatic MRI	0.869

*The results in these studies were recorded in the form of accuracy, not DSC.
DSC indicates dice similarity coefficient.

background (most of the bounding boxes are negative and do not contain objects) and the object. This problem is called class imbalance, and it lowers the performance of these networks.

o You Only Look Once network (YOLO): The architecture of YOLO is traditionally based on GoogleNet, without the inception module but with just 2 fully connected neurons at the output layer, making it fast. YOLO has 24 convolutional layers followed by 2 fully convolutional networks (FCNs) with just one reduction layer, this makes it far faster than Faster R-CNN. YOLO can process clips of 91 frames per second with higher performance than Faster R-CNN, which can only process clips of 17 frame per seconds with lower accuracy level. As a result, when used in medical images, it is commonly used for the detection of objects in ultrasound. The main drawback of YOLO is its inability to detect small objects, so it is less commonly used in cross-sectional imaging.[53]

o Single-shot multibox detector (SSMD): Similar to YOLO, SSMD depends on the whole image and its architecture relies on VGG-16—which is VGG neural network configuration with 16 layers—with the addition of multiple CNNs instead of the last layer in VGG.[54]

- RetinaNet: This algorithm solves the problem of the relative slowness of regional networks and the class imbalance by using a feature pyramidal network (FPN). This FPN can predict objects on multiple resolutions, which extract rich information from the image at all levels while remaining fast. RetinaNet uses the ResNet architecture to extract feature maps from the input image, which then uses FPN to detect objects on multiple resolutions. RetinaNet is one of the most recent advances in object detection algorithms.[55]

Automatic detection of breast and pulmonary lesions using neural networks has been among the most heavily examined. Such tools will support radiologists with a "second opinion," replace and outperform the currently used CAD, and reduce the time needed scanning millions of screening mammograms and chest x-rays.[45,56–59] Table 2 summarizes the most important recent articles that have used object detection algorithms in medical imaging.

Image Segmentation

Image segmentation means dividing the image into different regions, such as organs, lesions, or subvolumes in medical images. It is a vital step in quantitative image analysis for either feature extraction or classification.[73] Segmentation tasks are usually more difficult than classification and object detection, as the computer needs to classify each pixel and label it. There are several algorithms that are important in biomedical image segmentation. Table 3 summarizes important recent articles using imaging segmentation with DL.

- U-Net: This is the most famous neural network and has a proven high level of accuracy in medical image segmentation. U-Net is a CNN organized in a U-shape and developed by the computer science department at the University of Freiburg, Germany. This elegant neural network is composed of 2 parts: an encoding limb, consisting of different convolution layers leading to feature extraction, and a decoding limb, consisting of different

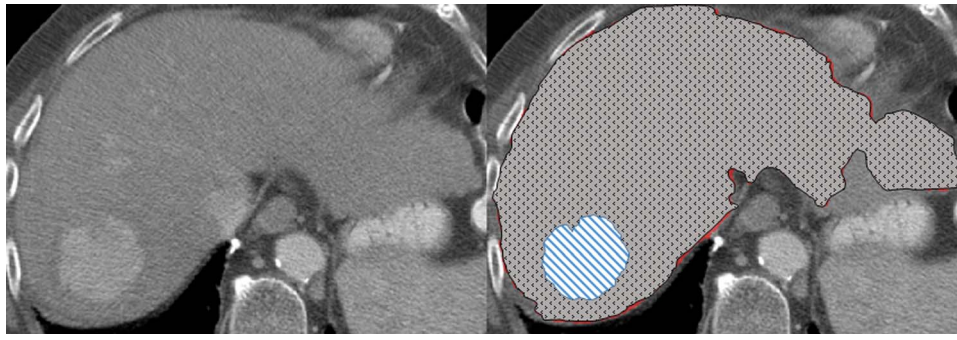

FIGURE 7. A, Contrast-enhanced CT scan of the liver showing hepatic tumor in the right liver. B, U-net was used to segment both the liver (dotted pattern) and the tumor (linear pattern), which can be further used for volumetric assessment to monitor treatment. Figure 7 can be viewed online in color at www.jcat.org.

layers of reverse-convolution or deconvolution. The U-Net architecture implements skip connections at multiple levels between the encoding and decoding limbs that store the spatial information of the image parts. This is in addition to the feature extraction performed using CNNs. The result is a map containing both features and spatial information (Fig. 7). One advantage of U-Net is its strong performance in segmentation using relatively small training data sets. However, one of the current recognized limitations is its inability to segment 3D medical images.[80]

oVariants of U-Net were developed to accept 3D images as an input. This includes V-Net, 3D U-Net, and hybrid Dense U-Net, which all have the same architecture of U-net (encoding and decoding limbs) but contain 3D convolutional filters rather than the 2D filters in U-net. These variants can be used in the segmentation of MRI/CT volume batches, such as in prostate MRI or liver CT. However, they are very demanding on computational and memory resources.[78,79]

- Fully convolutional network (FCN): This uses the same concept of U-net but with a different architecture. FCNs depend on greatly downsizing the input image to produce a very small feature map at the pixel level. It then uses deconvolution (DeConv, Dilated CNN) to apply these feature maps on a larger scale before fusing the product to produce the final output.[81]

TABLE 4. Summary of Studies Using DL in Diagnostic Radiology Other Than Imaging Analysis

Author (Year)	Training Set	Architecture Used	Outcome
Radiology reports manipulation			
Chen et al (2018)[82]	2500 reports	CNN obtained from GloVe*	Extracting PE findings from CT chest free-text reports
Jnawali et al (2019)[83]	12,852 reports	LSTM with one CNN	Labeling the radiological reports that contain either ICH or not
Gale et al (2018)[84]	41,032 images	CNN with DenseNet architecture and LSTM for reporting	Classifying and generating a report for a pelvic plain radiograph with hip fracture
Wang et al (2019)[85]	9240 patients	Classifier CNNs for detection and classification of pulmonary nodules embedded in intelligent system	Intelligent layout for structured reports, classified images. The visualized layout consists of structured reports with annotated images and classified.
Wang et al (2018)[86]	15,472 report	Ensemble of CNN (ResNet architecture) and LSTM	Detection and reporting of chest x-ray findings using specific model
Ben-Cohen et al (2019)[87]	23 PET scans, augmentation	A combination of FCN (U-net architecture) and GAN	Synthetize metabolic PET image from FCN contrast CT
Han et al (2017)[88]	15 patients (2400 training images), augmentation	3D Deep CNN (pretrained on VGG) based on U-net architecture	Synthetize MRI (with better soft tissue delineation) from CT
Xiang et al (2018)[89]	38 MRI scans, augmentation	Deep embedding CNN	Synthetize CT from T1WI MRI
Decision support			
Sepandi et al (2018)[90]	599 cases	Artificial neural (3 layers) network	Predict the risk of breast carcinoma with ANN using 23 variables
Fischer et al (2020)[91]	232 scans	Ensemble of CNN (for detection of calcium) and LSTM (to keep tracking of calcium along the coronaries)	Prediction of calcium score using CT coronary angiography
Buda et al (2019)[92]	1139 patients (1278 images)	Faster R-CNN (based on ResNet architecture)	Deciding whether a thyroid nodule should be biopsied or not
Spasov et al (2018)[93]	ADNI database (308 patients)	3D CNN with GoogleNet architecture	Prediction of Alzheimer disease and clinical evolution through imaging and clinical data
Shen et al (2019)[94]	114 patients, data augmentation (1562 images)	Deep CNN based on inception network	Prediction of outcome in patient with cervical cancer
Morshid et al (2019)[95]	105 patients	Cascaded CNN	Prediction of HCC response to TACE from preprocedural CT scans.
Quality control			
Esses et al (2018)[96]	351 cases (29,790 images)	Deep CNN	Classify the MRI images according to its quality
Lee et al (2018)[97]	5258 examination s	CNN with pretrained word embedding model (Word2vec)†	Protocol determination of MRI
Gurbani et al (2018)[98]	6774 images (7 patients)	Deep CNN	Classification of MR spectroscopy quality as either good or bad

*GloVe is a type of converting words to mathematical vectors (vectorizations) to be easily used in ML algorithms. The embedding is done using word-word co-occurrence matrix.[99]

†Word2vec uses a 2-layer CNN to reconstruct the context of words it used to map each word to other words.[100]

ANN indicates artificial neural network; GloVe, global vectors for word representation; ICH, intracranial hemorrhage; PE, pulmonary embolism; T1WI, T1-weighted images.

AI: BEYOND IMAGE ANALYSIS

Artificial intelligence and DL have applications in radiology other than image analysis. Many of these are still in the research phase, and incorporation of these applications into clinical practice needs extensive validation. The following sections highlight these applications with Table 4 emphasizing some of these efforts.

Quality Control

Choosing an appropriate imaging study protocol is a common quality assurance issue. Selecting the appropriate study protocol based on an individual patient's needs can be a time-consuming process, requiring an experienced radiologist who is aware of various protocols and familiarity with the patient's clinical situation through a review of the medical records.[101] A combination of ML, for natural language interface classification, and DL (eg, CNN) models has been investigated in selecting an appropriate study protocol based on the patient's presenting signs and symptoms and previous imaging reports; the study demonstrated a 93% to 94% accuracy using both of these models.[102]

Another aspect of quality assurance, assessing image quality, is usually performed by the physicist, who checks multiple benchmarks, including the adequacy of exposure, penetration, and artifacts. This quality revision has been explored using CNNs in both CT and MRI scanning and was found to be able to be performed instantaneously after completion of the scan, eliminating the need for patient rescheduling or manual revision by a physician.[96,103–105]

Workflow Organization

A major potential use for AI algorithms in radiology is the optimization of workflow and patient scheduling. Missed appointments can lead to a delay in patient care and increased costs. Certain factors have been demonstrated to be associated with missed appointments, including the previous behavior of the patient, which can be a clue to predict the likelihood of a future missed appointment. Regression models and other ML algorithms have been shown to hold the potential for reducing missed appointments while developing solutions to the problem that can arise from these missed appointments.[106,107]

Optimization of staff scheduling is another difficult task, especially in large radiology practices; multiple factors must be considered in scheduling, such as examination complexity, day of the week, and imaging modality. Inappropriate staff scheduling can lead to radiologist dissatisfaction. Artificial intelligence can be used to ensure a fair distribution of work among staff without affecting patient care.[101,108]

Follow-up of incidental findings on imaging studies is another important potential application of AI. Missing follow-up studies can lead to poor outcomes, malpractice claims, and increased costs. In 2 studies, follow-up imaging for incidental pulmonary nodules and indeterminate lesions was not completed for 71% and 44% of cases, respectively.[109,110] Artificial intelligence algorithms can be used to extract information from radiology reports that suggest the need for follow-up, which has been previously investigated in incidental pulmonary nodules.[111]

Radiology Reporting

There are many other potential applications for DL in radiology reporting. Besides the speech recognition systems now in use, DL algorithms can provide real-time alerts to radiologists to detect reporting errors such as lesion laterality and patient sex.[106] Moreover, information retrieval from medical records by these algorithms can be automated, obviating the need for manual searching.[112]

Using AI can help in the formulation of radiology reports, making different versions for patients, pharmacists, and insurance companies that contain the same information but use simpler language understandable by patients and professionals who need the information in the report but do not have the same level of knowledge or training. Another potential use is optimizing the report according to the specialty of the clinician who is using the report; for example, a report could be generated for a surgeon with details on lesion margins, extensions, and other information specific to surgery. This capability is more efficient for the care team.

CONCLUSIONS

Recent AI techniques and DL act as a breakthrough in the field of imaging analysis. Artificial intelligence–based approaches will aid both radiologists and clinicians but will certainly not replace them. Instead, these approaches will likely be used for consultation and decision support rather than decision making. Regardless, radiologists must be aware of these technologies and their role in the medical field. Although the high accuracy, robustness, speed, and utility of AI algorithms in the field of medical imaging have been demonstrated in research studies, most of the algorithms, particularly DL algorithms, are still in the testing phase and have not yet been scaled up or incorporated into clinical practice. The widespread incorporation of recent AI technologies into clinical practice faces many challenges. First, these algorithms must be trained on very large data sets, which in many cases is not feasible. Moreover, study protocols must be standardized across all institutions before the incorporation of DL algorithms. Improving algorithm accuracy and performance is a complex, difficult problem. Finally, proposed models for solving radiological problems must be externally validated using data from different vendors and institutions before its incorporation into clinical practice.

ACKNOWLEDGMENTS

The authors would like to acknowledge Scientific Publications Service, Research Medical Library, MD Anderson Cancer Center for their help in editing this article.

REFERENCES

1. Shiraishi J, Li Q, Appelbaum D, et al. Computer-aided diagnosis and artificial intelligence in clinical imaging. *Semin Nucl Med*. 2011;41:449–462.
2. Doi K. Computer-aided diagnosis in medical imaging: historical review, current status and future potential. *Comput Med Imaging Graph*. 2007;31:198–211.
3. Turing AM. Computing machinery and intelligence. *Mind*. 1950;59:433–460.
4. Samuel AL. Some studies in machine learning using the game of checkers. *IBM J Res Dev*. 1959;3:210–229.
5. Mitchell T. *Machine Learning*. New York, NY: McGraw Hill; 1997.
6. van Ginneken B. Fifty years of computer analysis in chest imaging: rule-based, machine learning, deep learning. *Radiol Phys Technol*. 2017;10:23–32.
7. Loussaief S, Abdelkrim A, editors. Deep learning vs. bag of features in machine learning for image classification. 2018 International Conference on Advanced Systems and Electric Technologies (IC_ASET); March 22–25, 2018. Hammamet, Tunisia: IEEE; 2018.
8. LeCun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition. *Proc IEEE*. 1998;86:2278–2324.
9. Hubel DH, Wiesel TN. Receptive fields of single neurones in the cat's striate cortex. *J Physiol*. 1959;148:574–591.

10. Yamashita R, Nishio M, Do RKG, et al. Convolutional neural networks: an overview and application in radiology. *Insights Imaging*. 2018;9: 611–629.
11. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. *arXiv*. 2014. Preprint arXiv:14091556.
12. Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. *Adv Neural Inform Proc Syst*. 2012;25: 1097–1105.
13. Zeiler MD, Fergus R. Visualizing and understanding convolutional networks. In: Fleet D, Pajdla T, Schiele B, et al, eds. *European Conference on Computer Vision*. Cham, Switzerland: Springer; 2014.
14. Szegedy C, Liu W, Jia Y, et al. Going deeper with convolutions. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Boston, MA: IEEE; 2015:1–9.
15. Huang G, Liu Z, Van Der Maaten L, et al. Densely connected convolutional networks. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu, HI: IEEE; 2017: 4700–4708.
16. He K, Zhang X, Ren S, et al. Deep residual learning for image recognition. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas, NV: IEEE; 2016:770–778.
17. Hu J, Shen L, Sun G. Squeeze-and-excitation networks. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Salt Lake, UT: IEEE; 2018:7132–7141.
18. Rumelhart DE, Hinton GE, Williams RJ. Learning representations by back-propagating errors. *Nature*. 1986;323:533–536.
19. Friedman C, Alderson PO, Austin JH, et al. A general natural-language text processor for clinical radiology. *J Am Med Inform Assoc*. 1994;1: 161–174.
20. Sundermeyer M, Schlüter R, Ney H. LSTM neural networks for language modeling. In: *Thirteenth Annual Conference of the International Speech Communication Association*. Portland, OR: IEEE; 2012:9161–9172.
21. Berglund M, Raiko T, Honkala M, et al. Bidirectional recurrent neural networks as generative models. *Adv Neural Inform Process Syst*. 2015; 856–864.
22. Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets. *Adv Neural Inform Process Syst*. 2014;27:2672–2680.
23. Ledig C, Theis L, Huszár F, et al. Photo-realistic single image super-resolution using a generative adversarial network. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu, HI: IEEE; 2017.
24. 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.
25. Liu Z, Bicer T, Kettimuthu R, et al. TomoGAN: low-dose x-ray tomography with generative adversarial networks. *arXiv*. 2019. Preprint arXiv:190207582.
26. Nie D, Trullo R, Lian J, et al. *Medical Image Synthesis with Context-Aware Generative Adversarial Networks*. Cham, Switzerland: Springer International Publishing; 2017.
27. Chen Y, Shi F, Christodoulou AG, et al. *Efficient and Accurate MRI Super-Resolution Using a Generative Adversarial Network and 3D Multi-level Densely Connected Network*. Cham, Switzerland: Springer International Publishing; 2018.
28. Singh VK, Romani S, Rashwan HA, et al. *Conditional Generative Adversarial and Convolutional Networks for X-ray Breast Mass Segmentation and Shape Classification*. Cham, Switzerland: Springer International Publishing; 2018.
29. Moeskops P, Veta M, Lafarge MW, et al. *Adversarial Training and Dilated Convolutions for Brain MRI Segmentation*. Cham, Switzerland: Springer International Publishing; 2017.
30. Frid-Adar M, Diamant I, Klang E, et al. GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification. *Neurocomputing*. 2018;321:321–331.
31. Mazurowski MA, Buda M, Saha A, et al. Deep learning in radiology: an overview of the concepts and a survey of the state of the art with focus on MRI. *J Magn Reson Imaging*. 2019;49:939–954.
32. Chen C, Qin C, Qiu H, et al. Deep learning for cardiac image segmentation: a review. *Front Cardiovasc Med*. 2020;7:25.
33. Sebastian Thrun LP. *Learning to Learn*. Boston, MA: Springer; 1998.
34. Perez L, Wang J. The effectiveness of data augmentation in image classification using deep learning. *arXiv*. 2017; Preprint arXiv: 171204621.
35. Raghu M, Zhang C, Kleinberg J, et al. Transfusion: understanding transfer learning with applications to medical imaging. *arXiv*. 2019; Preprint arXiv:190207208.
36. Ma J, Wu F, Jiang T, et al. Cascade convolutional neural networks for automatic detection of thyroid nodules in ultrasound images. *Med Phys*. 2017;44:1678–1691.
37. Ahammed Muneer KV, Rajendran V, Paul Joseph K. Glioma tumor grade identification using artificial intelligent techniques. *J Med Syst*. 2019; 43:113.
38. Song Y, Zhang YD, Yan X, et al. Computer-aided diagnosis of prostate cancer using a deep convolutional neural network from multiparametric MRI. *J Magn Reson Imaging*. 2018;48:1570–1577.
39. Prevedello LM, Erdal BS, Ryu JL, et al. Automated critical test findings identification and online notification system using artificial intelligence in imaging. *Radiology*. 2017;285:923–931.
40. Kim DH, MacKinnon T. Artificial intelligence in fracture detection: transfer learning from deep convolutional neural networks. *Clin Radiol*. 2018;73:439–445.
41. Taylor AG, Mielke C, Mongan J. Automated detection of moderate and large pneumothorax on frontal chest x-rays using deep convolutional neural networks: a retrospective study. *PLoS Med*. 2018;15:e1002697.
42. Chung SW, Han SS, Lee JW, et al. Automated detection and classification of the proximal humerus fracture by using deep learning algorithm. *Acta Orthop*. 2018;89:468–473.
43. Schwyzer M, Ferraro DA, Muehlematter UJ, et al. Automated detection of lung cancer at ultralow dose PET/CT by deep neural networks—initial results. *Lung Cancer*. 2018;126:170–173.
44. Lee JH, Ha EJ, Kim JH. Application of deep learning to the diagnosis of cervical lymph node metastasis from thyroid cancer with CT. *Eur Radiol*. 2019;29:5452–5457.
45. Lakhani P, Sundaram B. Deep learning at chest radiography: automated classification of pulmonary tuberculosis by using convolutional neural networks. *Radiology*. 2017;284:574–582.
46. Dalmis MU, Gubern-Mérida A, Vreemann S, et al. Artificial intelligence–based classification of breast lesions imaged with a multiparametric breast MRI protocol with ultrafast DCE-MRI, T2, and DWI. *Invest Radiol*. 2019;54:325–332.
47. Correa M, Zimic M, Barrientos F, et al. Automatic classification of pediatric pneumonia based on lung ultrasound pattern recognition. *PLoS One*. 2018;13:e0206410.
48. Rodríguez-Ruiz A, Krupinski E, Mordang J-J, et al. Detection of breast cancer with mammography: effect of an artificial intelligence support system. *Radiology*. 2019;290:305–314.
49. Hamm CA, Wang CJ, Savic LJ, et al. Deep learning for liver tumor diagnosis part I: development of a convolutional neural network classifier for multi-phasic MRI. *Eur Radiol*. 2019;29:3338–3347.
50. Lin T-Y, Maire M, Belongie S, et al. Microsoft COCO: common objects in context. In: *European Conference on Computer Vision*. Cham, Switzerland: Springer; 2014.

51. Ren S, He K, Girshick R, et al. Faster R-CNN: towards real-time object detection with region proposal networks. IEEE Trans Pattern Anal Mach Intell. 2017:39:1137–1149.
52. He K, Gkioxari G, Dollár P, et al. Mask R-CNN. In: *Proceedings of the IEEE International Conference on Computer Vision*. Venice, Italy: IEEE; 2017:2980–2988.
53. Redmon J, Divvala S, Girshick R, et al. You only look once: unified, real-time object detection. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas, NV: IEEE; 2016: 779–788.
54. Liu W, Anguelov D, Erhan D, et al. SSD: Single shot multibox detector. In: Leibe B, Matas J, Sebe N, eds. *European Conference on Computer Vision*. Cham, Switzerland: Springer; 2016.
55. Lin T-Y, Goyal P, Girshick R, et al. Focal loss for dense object detection. In: *Proceedings of the IEEE International Conference on Computer Vision*. Venice, Italy: IEEE; 2017:2980–2988.
56. Kooi T, Litjens G, van Ginneken B, et al. Large scale deep learning for computer aided detection of mammographic lesions. *Med Image Anal*. 2017;35:303–312.
57. Pan I, Cadrin-Chênevert A, Cheng PM. Tackling the Radiological Society of North America pneumonia detection challenge. *AJR Am J Roentgenol*. 2019;213:568–574.
58. Fan W, Jiang H, Ma L, et al. A modified faster R-CNN method to improve the performance of the pulmonary nodule detection. In: Jiang X, Hwang J-N, eds. *Tenth International Conference on Digital Image Processing*. SPIE; 2018.
59. Chiao J-Y, Chen K-Y, Liao KY-K, et al. Detection and classification the breast tumors using mask R-CNN on sonograms. *Medicine (Baltimore)*. 2019;98:e15200.
60. Gan K, Xu D, Lin Y, et al. Artificial intelligence detection of distal radius fractures: a comparison between the convolutional neural network and professional assessments. *Acta Orthop*. 2019;90:394–400.
61. Tuzoff DV, Tuzova LN, Bornstein MM, et al. Tooth detection and numbering in panoramic radiographs using convolutional neural networks. *Dentomaxillofac Radiol*. 2019;48:20180051.
62. Wang D, Xu J, Zhang Z, et al. Evaluation of rectal cancer circumferential resection margin using faster region-based convolutional neural network in high-resolution magnetic resonance images. *Dis Colon Rectum*. 2020; 63:143–151.
63. Liu J, Wang D, Lu L, et al. Detection and diagnosis of colitis on computed tomography using deep convolutional neural networks. *Med Phys*. 2017; 44:4630–4642.
64. Nasrullah N, Sang J, Alam MS, et al. Automated lung nodule detection and classification using deep learning combined with multiple strategies. *Sensors (Basel)*. 2019;19:3722.
65. Lindsey R, Daluiski A, Chopra S, et al. Deep neural network improves fracture detection by clinicians. *Proc Natl Acad Sci U S A*. 2018;115: 11591–11596.
66. Chang PD, Kuoy E, Grinband J, et al. Hybrid 3D/2D convolutional neural network for hemorrhage evaluation on head CT. *AJNR Am J Neuroradiol*. 2018;39:1609–1616.
67. Filice RW, Stein A, Wu CC, et al. Crowdsourcing pneumothorax annotations using machine learning annotations on the NIH chest x-ray dataset. *J Digit Imaging*. 2019;1–7.
68. Fritz B, Marbach G, Civardi F, et al. Deep convolutional neural network–based detection of meniscus tears: comparison with radiologists and surgery as standard of reference. *Skelet Radiol*. 2020;1–11.
69. Kim E-K, Kim H-E, Han K, et al. Applying data-driven imaging biomarker in mammography for breast cancer screening: preliminary study. *Sci Rep*. 2018;8:2762.
70. Wang L, Yang S, Yang S, et al. Automatic thyroid nodule recognition and diagnosis in ultrasound imaging with the YOLOv2 neural network. *World J Surg Oncol*. 2019;17:1–9.
71. Cao Z, Duan L, Yang G, et al, eds. Breast tumor detection in ultrasound images using deep learning. In: *International Workshop on Patch-based Techniques in Medical Imaging*. Cham, Switzerland: Springer; 2017.
72. Li N, Liu H, Qiu B, et al. Detection and attention: diagnosing pulmonary lung cancer from CT by imitating physicians. *arXiv*. 2017; preprint arXiv: 1712.05114.
73. Cao C, Liu F, Tan H, et al. Deep learning and its applications in biomedicine. *Genomics Proteomics Bioinformatics*. 2018;16:17–32.
74. Yao H, Williamson C, Soroushmehr R, Gryak J, Najarian K, editors. Hematoma segmentation using dilated convolutional neural network. Paper presented at: 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC); July 17–21, 2018; Honolulu, HI.
75. Park A, Chute C, Rajpurkar P, et al. Deep learning–assisted diagnosis of cerebral aneurysms using the HeadXNet model. *JAMA Netw Open*. 2019; 2:e195600-e.
76. Blanc-Durand P, Van Der Gucht A, Schaefer N, et al. Automatic lesion detection and segmentation of 18F-FET PET in gliomas: a full 3D U-Net convolutional neural network study. *PLoS One*. 2018;13:e0195798.
77. Liu W, Liu M, Guo X, et al. Evaluation of acute pulmonary embolism and clot burden on CTPA with deep learning. *Eur Radiol*. 2020;30:3567–3575.
78. Li X, Chen H, Qi X, et al. H-DenseUNet: hybrid densely connected UNet for liver and tumor segmentation from CT volumes. *IEEE Trans Med Imaging*. 2018;37:2663–2674.
79. Milletari F, Navab N, Ahmadi S. V-Net: fully convolutional neural networks for volumetric medical image segmentation. Cornell University Library. June 15, 2016. https://arxiv.org/abs/1606.04797. Accessed January 12, 2022.
80. Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. In: Navab N, Hornegger J, Wells W, eds. *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Cham, Switzerland: Springer; 2015.
81. Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation. *IEEE Trans Pattern Anal Mach Intell*. 2015;39:640–651.
82. Chen MC, Ball RL, Yang L, et al. Deep learning to classify radiology free-text reports. *Radiology*. 2018;286:845–852.
83. Jnawali K, Arbabshirani MR, Ulloa AE, Rao N, Patel AA, editors. Automatic classification of radiological report for intracranial hemorrhage. Paper presented at: 2019 IEEE 13th International Conference on Semantic Computing (ICSC); January 30–February 1, 2019; Newport Beach, CA.
84. Gale W, Oakden-Rayner L, Carneiro G, et al. Producing radiologist-quality reports for interpretable artificial intelligence. *arXiv*. 2018; preprint arXiv:180600340.
85. Wang Y, Yan F, Lu X, et al. IILS: intelligent imaging layout system for automatic imaging report standardization and intra-interdisciplinary clinical workflow optimization. *EBioMedicine*. 2019;44:162–181.
86. Wang X, Peng Y, Lu L, Lu Z, Summers RM, editors. Tienet: text-image embedding network for common thorax disease classification and reporting in chest x-rays. Paper presented at: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; 18–23 June 2018; Salt Lake City, UT.
87. Ben-Cohen A, Klang E, Raskin SP, et al. Cross-modality synthesis from CT to PET using FCN and GAN networks for improved automated lesion detection. *Eng Appl Artif Intell*. 2019;78:186–194.
88. Han X. MR-based synthetic CT generation using a deep convolutional neural network method. *Med Phys*. 2017;44:1408–1419.

89. Xiang L, Wang Q, Nie D, et al. Deep embedding convolutional neural network for synthesizing CT image from T1-Weighted MR image. *Med Image Anal*. 2018;47:31–44.

90. Sepandi M, Taghdir M, Rezaianzadeh A, et al. Assessing breast cancer risk with an artificial neural network. *Asian Pac J Cancer Prev*. 2018;19: 1017–1019.

91. Fischer AM, Eid M, De Cecco CN, et al. Accuracy of an artificial intelligence deep learning algorithm implementing a recurrent neural network with long short-term memory for the automated detection of calcified plaques from coronary computed tomography angiography. *J Thorac Imaging*. 2020;35(suppl 1):S49–S57.

92. Buda M, Wildman-Tobriner B, Hoang JK, et al. Management of thyroid nodules seen on US images: deep learning may match performance of radiologists. *Radiology*. 2019;292:695–701.

93. Spasov SE, Passamonti L, Duggento A, Liò P, Toschi N, editors. A multi-modal convolutional neural network framework for the prediction of Alzheimer's disease. Paper presented at: 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC); July 18–21, 2018; Honolulu, HI.

94. Shen W-C, Chen S-W, Wu K-C, et al. Prediction of local relapse and distant metastasis in patients with definitive chemoradiotherapy-treated cervical cancer by deep learning from [18 F]-fluorodeoxyglucose positron emission tomography/computed tomography. *Eur Radiol*. 2019;29: 6741–6749.

95. Morshid A, Elsayes KM, Khalaf AM, et al. A machine learning model to predict hepatocellular carcinoma response to transcatheter arterial chemoembolization. *Radiol Artif Intell*. 2019;1:e180021.

96. Esses SJ, Lu X, Zhao T, et al. Automated image quality evaluation of T2-weighted liver MRI utilizing deep learning architecture. *J Magn Reson Imaging*. 2018;47:723–728.

97. Lee YH. Efficiency improvement in a busy radiology practice: determination of musculoskeletal magnetic resonance imaging protocol using deep-learning convolutional neural networks. *J Digit Imaging*. 2018;31:604–610.

98. Gurbani SS, Schreibmann E, Maudsley AA, et al. A convolutional neural network to filter artifacts in spectroscopic MRI. *Magn Reson Med*. 2018; 80:1765–1775.

99. Pennington J, Socher R, Manning CD, editors. Glove: global vectors for word representation. Paper presented at: Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP); October 2014; Doha, Qatar.

100. Mikolov T, Chen K, Corrado G, et al. Efficient estimation of word representations in vector space. *arXiv*. 2013; preprint arXiv:13013781.

101. Lakhani P, Prater AB, Hutson RK, et al. Machine learning in radiology: applications beyond image interpretation. *J Am Coll Radiol*. 2018;15: 350–359.

102. Sohn JH, Trivedi H, Mesterhazy J, et al. Development and validation of machine learning based natural language classifiers to automatically assign MRI abdomen/pelvis protocols from free-text clinical indications. Abstract presented at: Society of Imaging Informatics in Medicine, Annual Meeting; June 3, 2017; Pittsburgh, PA.

103. Lee JH, Grant BR, Chung JH, et al. Assessment of diagnostic image quality of computed tomography (CT) images of the lung using deep learning. In: *Proc. SPIE 10573, Medical Imaging 2018: Physics of Medical Imaging*. Houston, TX: SPIE; 2018. Pittsburgh, PA.

104. Hagiwara A, Otsuka Y, Hori M, et al. Improving the quality of synthetic FLAIR images with deep learning using a conditional generative adversarial network for pixel-by-pixel image translation. *Am J Neuroradiol*. 2019;40:224–230.

105. Su B, Liu Y, Jiang Y, et al. Bone induced artifacts elimination using two-step convolutional neural network. In: Matej S, Metzler SD, eds. *15th International Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine*. Philadelphia, PA: International Society for Optics and Photonics; 2019.

106. Morey JM, Haney NM, Kim W. Applications of AI Beyond Image Interpretation. In: Ranschaert E, Morozov S, Algra P, eds. *Artificial Intelligence in Medical Imaging*. Cham, Switzerland: Springer; 2019: 129–143.

107. Kurasawa H, Hayashi K, Fujino A, et al. Machine-learning-based prediction of a missed scheduled clinical appointment by patients with diabetes. *J Diabetes Sci Technol*. 2016;10:730–736.

108. Boroumand G, Dave JK, Roth CG. Shedding light on the off-hours coverage gap in radiology: improving turnaround times and critical results reporting. 2017.

109. Blagev DP, Lloyd JF, Conner K, et al. Follow-up of incidental pulmonary nodules and the radiology report. *J Am Coll Radiol*. 2014; 11:378–383.

110. Cook TS, Lalevic D, Sloan C, et al. Implementation of an automated radiology recommendation-tracking engine for abdominal imaging findings of possible cancer. *J Am Coll Radiol*. 2017;14:629–636.

111. Lacson R, Desai S, Landman A, et al. Impact of a health information technology intervention on the follow-up management of pulmonary nodules. *J Digit Imaging*. 2018;31:19–25.

112. Pons E, Braun LM, Hunink MM, et al. Natural language processing in radiology: a systematic review. *Radiology*. 2016;279:329–343.