JACC: CARDIOVASCULAR IMAGING VOL. 12, NO. 8, 2019

PUBLISHED BY ELSEVIER

iREVIEW

STATE-OF-THE-ART REVIEW

State-of-the-Art Deep Learning in Cardiovascular Image Analysis

Geert Litjens, PhD,[a] Francesco Ciompi, PhD,[a] Jelmer M. Wolterink, PhD,[b] Bob D. de Vos, PhD,[b] Tim Leiner, MD, PhD,[d] Jonas Teuwen, PhD,[c,e] Ivana Išgum, PhD[b]

ABSTRACT

Cardiovascular imaging is going to change substantially in the next decade, fueled by the deep learning revolution. For medical professionals, it is important to keep track of these developments to ensure that deep learning can have meaningful impact on clinical practice. This review aims to be a stepping stone in this process. The general concepts underlying most successful deep learning algorithms are explained, and an overview of the state-of-the-art deep learning in cardiovascular imaging is provided. This review discusses >80 papers, covering modalities ranging from cardiac magnetic resonance, computed tomography, and single-photon emission computed tomography, to intravascular optical coherence tomography and echocardiography. Many different machines learning algorithms were used throughout these papers, with the most common being convolutional neural networks. Recent algorithms such as generative adversarial models were also used. The potential implications of deep learning algorithms on clinical practice, now and in the near future, are discussed. (J Am Coll Cardiol Img 2019;12:1549-65)

Deep learning techniques have recently revolutionized many fields, from computer vision and natural image classification to speech recognition and language processing (1,2). In addition, deep learning methods have started to have a profound influence on medical imaging (3,4). This is exemplified by several key papers that have illustrated deep learning methods that are performing on par with experienced physicians for specific tasks, such as in disease classification in dermatology (5), diagnosing diabetic retinopathy in ophthalmology (6,7), and finding metastases in lymph

From the [a]Department of Pathology, Radboud University Medical Center, Nijmegen, the Netherlands; [b]Image Sciences Institute, University Medical Center Utrecht, Utrecht, the Netherlands; [c]Department of Radiology, Radboud University Medical Center, Nijmegen, the Netherlands; [d]Department of Radiology, University Medical Center Utrecht, Utrecht, the Netherlands; and the [e]Department of Radiation Oncology, Netherlands Cancer Institute, Amsterdam, the Netherlands. Dr. Litjens is supported by grants from the Dutch Cancer Society (KUN 2015-7970), from Netherlands Organization for Scientific Research (NWO) (project number 016.186.152), and from Stichting IT Projecten (project PATHOLOGIE 2); and has received research funding from Philips Digital Pathology Solutions; and has been a consultant for Novartis. Dr. Ciompi has received research grants from the European Union's Horizon 2020 research and innovation programme under grant agreement number 825292 (ExaMode), and from the Dutch Cancer Society (project number 11917). Dr. Leiner is supported by research grants from NWO/Foundation for Technological Sciences (number 12726 and number P15-26) with industrial participation (Pie Medical Imaging, 3Mensio Medical Imaging, and Philips Healthcare); research grants from the Netherlands Organization for Health Research and Development (FSCAD, number 104003009); industrial research grants from Pie Medical Imaging; and has received grant support from and is a member of the Speakers Bureau for Philips Healthcare and Bayer. Dr. Išgum is supported by research grants from NWO/Foundation for Technological Sciences (number 12726 and number 15-26) with industrial participation (Pie Medical Imaging, 3Mensio Medical Imaging, Philips Healthcare); research grants from the Netherlands Organization for Health Research and Development (FSCAD, number 104003009); research grant from Dutch Cancer Society (UU 2015-7947); and industrial research grants from Pie Medical Imaging; and is the founder and a shareholder of Quantib U. All other authors have reported that they have no relationships relevant to the contents of this paper to disclose.

Manuscript received November 1, 2018; revised manuscript received May 13, 2019, accepted June 13, 2019.

ISSN 1936-878X/$36.00 https://doi.org/10.1016/j.jcmg.2019.06.009

ABBREVIATIONS AND ACRONYMS

3D = 3 dimensional

ACDC = Automatic Cardiac Diagnosis Challenge

ANN = artificial neural network

CCTA = coronary compute tomography angiography

CMR = cardiac magnetic resonance

CNN = convolutional neural network

CT = computed tomography

ED = end-diastolic

ES = end-systolic

FCN = fully convolutional neural network

GAN = generative adversarial network

LVEF = left ventricular ejection fraction

RNN = recurrent neural network

nodes in pathology (8). These developments are not limited to the research setting, because deep learning-based algorithms and software are making their way into the clinic, with the first applications now receiving Food and Drug Administration approval.

Because of the large number of cardiac images that are routinely acquired with a wide range of modalities (9), there has been a surge in publications applying deep learning in the cardiac domain (**Figure 1**). In this review, we aim to provide an introduction to basic deep learning concepts and their possibilities, and offer an overview of the state of the art in deep learning for cardiovascular image analysis.

Several other reviews have already discussed the impact of machine learning or deep learning on (parts of) cardiology and cardiovascular imaging (4,10-14). The work by Slomka et al. (12) and Litjens et al. (4) are most similar to the current survey. Slomka et al. (12) reviewed cardiovascular imaging techniques in general and discussed approaches to automation, in which deep learning played an important part. However, only a few studies that applied deep learning to cardiovascular imaging were covered by Slomka et al. (12). Litjens et al. (4) reviewed applications of deep learning to medical imaging in general, but only covered studies that appeared before early 2017, whereas many works on cardiovascular imaging appeared after that date.

This review includes >80 original research papers on applications of deep learning to cardiovascular image analysis. This includes 15 papers from our previous survey (4) and an update with >60 papers that appeared after the publication of Litjens et al. (4). These papers were selected by querying PubMed for publications between January 1, 2017, and January 1, 2019, containing "deep learning" OR "convolutional" OR "machine learning", which resulted in 4,327 papers. Papers that focused on cardiovascular image analysis were manually selected based on the title and abstract for inclusion in this survey.

The survey is structured as follows. History of Cardiovascular Image Analysis provides a historical background on cardiovascular image analysis before the introduction of deep learning methods. Basis of Deep Learning introduces some key concepts and techniques in deep learning that have driven recent developments. In Applications, we review recent deep learning publications for anatomical, functional, and intraoperative cardiovascular imaging across computed tomography (CT), cardiac magnetic resonance (CMR), single-photon emission computed tomography, ultrasound, and other modalities. Finally, the Discussion section contains an overview on the current and future implications of deep learning on cardiovascular image analysis in clinical practice.

HISTORY OF CARDIOVASCULAR IMAGE ANALYSIS

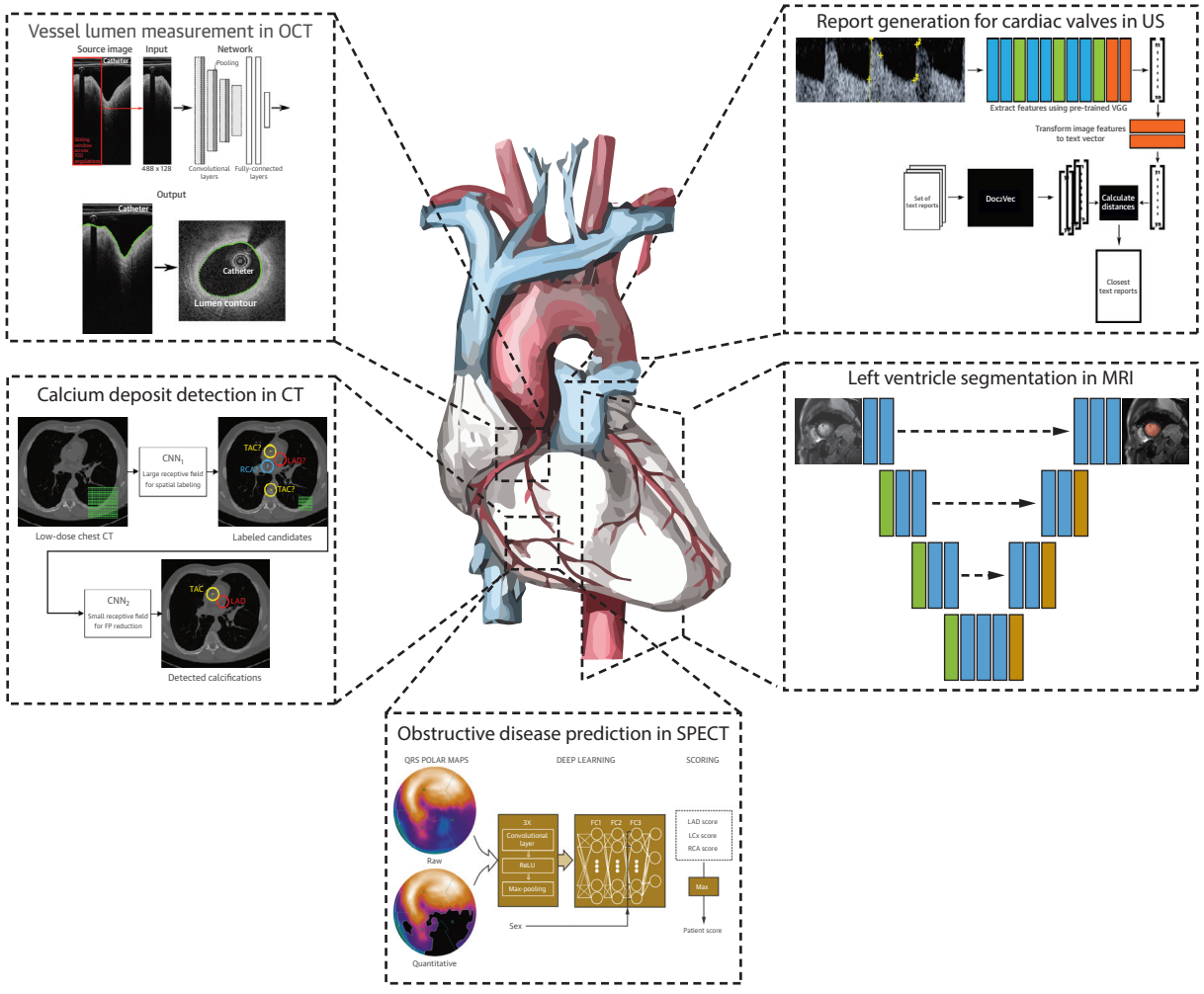
Before the advent of deep learning, a wealth of techniques had been developed to extract clinically relevant information from cardiovascular images. Early algorithms typically required significant manual tuning to transform an input image into the desired output (15).

More modern algorithms included a higher level of automation. Some algorithms directly used image intensities to drive an automated algorithm (e.g., level sets for segmentation) (16,17) or minimum cost path computation for vessel centerline extraction (18). Other methods extracted handcrafted features from cardiovascular images and then fed those to a statistical classifier, such as a support vector machine (19), to perform some form of disease prediction. These features generally described characteristics such as texture and shape (20-22). Recently, research moved in a new direction, typically referred to as radiomics (23), in which hundreds or even thousands of highly abstract features were automatically extracted from medical images. Subsequently, statistical classifiers have to figure out which features are relevant for the task at hand.

One key commonality between these approaches is that the features are designed by humans and subsequently fed to a prediction model. In contrast, deep learning algorithms simultaneously learn relevant features and the prediction model from input image to desired outcome, which is often referred to as end-to-end learning.

Nonetheless, many successful studies were performed with traditional methods. Most focused on segmentation of anatomical structures in cardiac images for visualization or quantitative analysis (24-28). For example, automatic contouring of the left ventricle and determination of the left ventricular ejection fraction (LVEF) in short-axis cine MR images and 3-dimensional (3D) echocardiography has long been a topic of research (24,26). In coronary CT angiography (CCTA), analysis of coronary arteries for the detection of luminal stenosis and atherosclerotic plaque requires the development of vessel detection (27) and lumen segmentation methods (28). In the last couple of years, machine learning approaches have also explored direct prediction of patient outcome

FIGURE 1 Applications of Deep Learning in Cardiovascular Imaging



A simple illustration of the variety of applications deep learning–based methods have in cardiovascular imaging. We highlight 5 examples here across 5 different modalities and anatomies. **(Top left)** vessel lumen measurement (adapted from Yong et al. [58]), **(top right)** report generation (created based on Moradi et al. [59]), **(middle right)** left ventricle segmentation, **(middle left)** obstructive disease prediction (adapted with permission from Betancur et al. [50]), and **(bottom)** calcium deposit detection (adapted with permission from Lessmann et al. [46]). CMR = cardiac magnetic resonance; CNN = convolutional neural network; CT = computed tomography; LCx = left circumflex; LAD = left anterior descending; OCT = optical coherence tomography; RCA = right coronary artery; SPECT = single-photon emission computer tomography; TAC = thoracic aortic calcification; US = ultrasound.

based on parameters extracted from images obtained via nuclear cardiology (29).

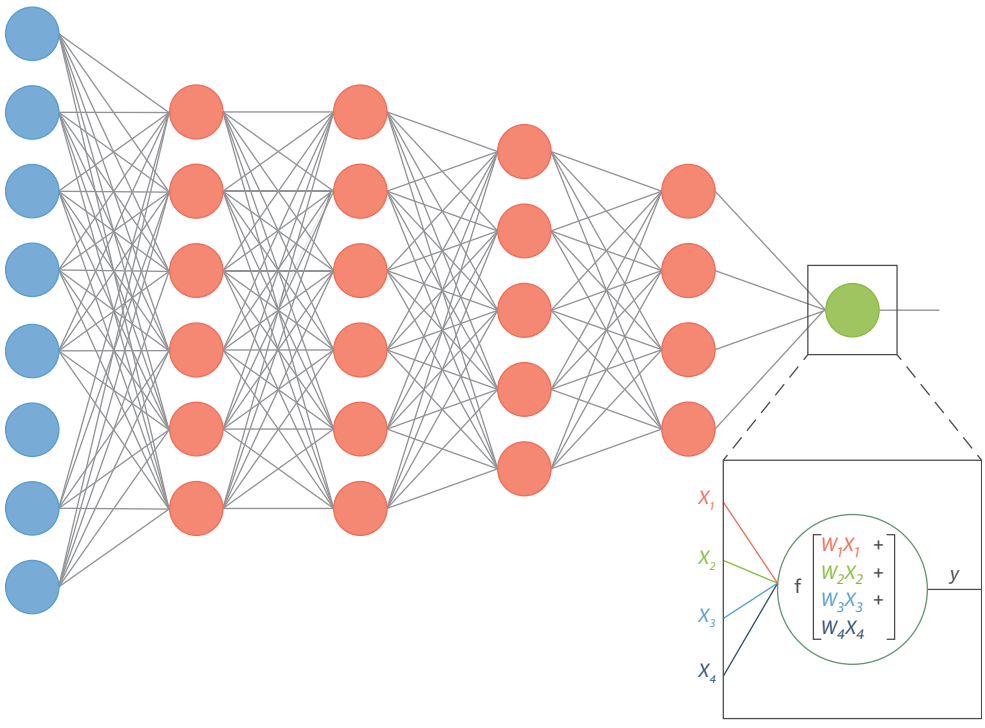
To encourage machine learning researchers to work on cardiovascular imaging and to provide frameworks for comparison of different methods, challenges with standardized datasets and evaluation criteria have been set up. These include challenges for right ventricle segmentation in cardiac MR (30), automatic contouring of the left ventricle in 3D echocardiography (31), coronary centerline extraction (32), stenosis detection in CCTA (33), and coronary artery calcium scoring in cardiac CT (34), among others.

Although many conventional techniques showed promising results in challenging settings, they were often not robust yet for a routine clinical environment. Many have now been superseded, either in performance or in efficiency, by deep learning–based methods, which will be discussed in the Discussion section.

BASICS OF DEEP LEARNING

At the heart of all currently popular deep learning methods are artificial neural networks (ANNs). In this

FIGURE 2 Schematic Drawing of an Artificial Neural Network With 5 Layers



The input and output layers in this drawing are indicated in **blue** and **green**; the hidden layers are in **pink**. The output node y has been highlighted to show explicitly the mathematical operation that occurs in every neuron. Each input x_i to a neuron has its own weight w_i, indicated here with separate colors. Subsequently, a nonlinear function f is applied to the result.

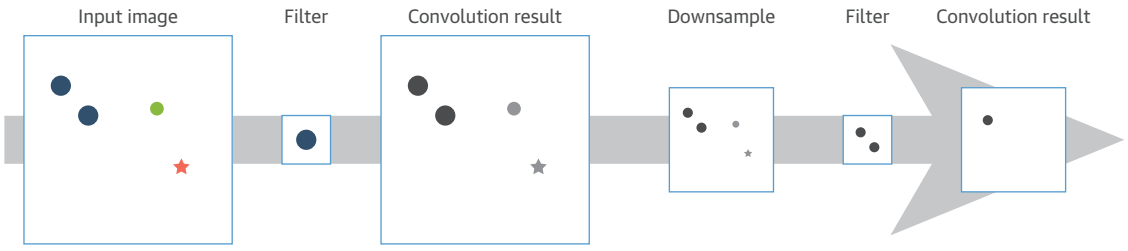
section, we introduce the basics of these networks and the higher level concepts that are important to be able to understand and assess the quality of publications that apply deep learning to cardiovascular image analysis. For a more thorough review of the history and technical aspects of deep learning, we recommend the review by LeCun et al. (2).

ARTIFICIAL NEURONS AND NEURAL NETWORKS. An ANN consists of a large number of artificial neurons. A simple example of an ANN is shown in Figure 2. A single artificial neuron can have multiple inputs that are combined into a single output. Each input is multiplied by a coefficient, typically referred to as a weight, and all multiplied inputs are summed. Often an extra coefficient, the bias, is added to this summation. Subsequently, the resultant value is passed through a nonlinear function.

The strength of ANNs lies in the fact that by combining many neurons in layers, highly nonlinear relationships between the input and the desired output can be modeled, which is typically needed to obtain accurate predictions. Although ANNs have been around since the late 1950s (35) and even saw successful application in medical imaging in 1995 (36), they never achieved the popularity they have now. One of the key reasons, next to methodological developments, was the lack of computational power, which forced researchers to only use networks that were only a few layers deep. Generally, the relationship between a medical image and the desired output is too complex to capture efficiently with such shallow ANNs. Over the past decade, these issues have been resolved because of the increasing availability of affordable computational power. Currently, researchers can easily build ANNSs of tens or even hundreds of layers. This has enabled direct application of neural networks to medical images without previous feature extraction.

DIFFERENT NETWORK STRUCTURES FOR DIFFERENT TASKS. Depending on the task that is being solved (e.g., image classification, image enhancement), a specific neural network structure, or architecture, can be chosen that is optimally suited for that specific task. In this case, the architecture refers to how many layers and neurons are in the neural network and how they are connected.

FIGURE 3 Schematic Drawing of Convolutional Layers in a Neural Network



Schematic drawing of 2 subsequent convolution layers with 1 downsampling (or pooling) layer. This network tries to identify locations of 2 diagonally-oriented **blue dots**. By splitting this across 2 layers, this problem can be solved with 2 small filters. In the first layer, we look for **blue dots**. The filter is applied to every position in the image and shows high responses **(dark grey)** where the image is most similar to the filter. In the second layer, we look for positive detections from the first layer in a diagonal orientation. By stacking many convolutional layers in a neural network, complex patterns can be identified and detected.

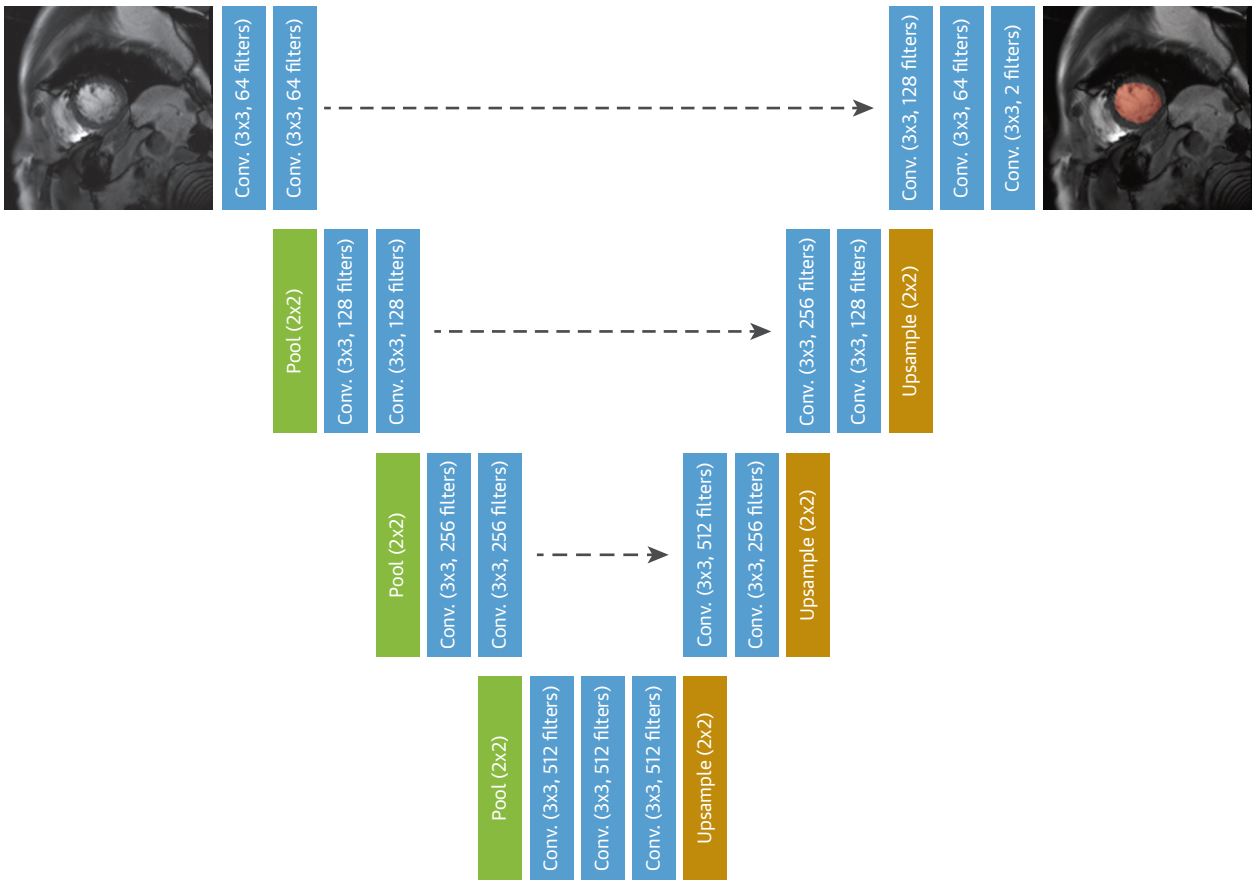
The most common architecture in image analysis is the convolutional neural network (CNN). Compared with standard ANNs, CNNs dramatically reduce the number of coefficients (weights) in the network by sharing weights at each location in the image. This leads to the networks applying a convolution operation, hence, the name convolutional neural network. A simple illustration of a convolution is shown in **Figure 3**. Generally, convolutional neural networks consist of both convolutional and pooling layers. Pooling layers do not have coefficients or weights, but downsample the data, typically by a factor of 2. This reduces the resolution of the image but increases the field of view of subsequent layers, which help CNNs incorporate more contextual information. This is also shown via a simple example in **Figure 3**. Most convolutional networks for image classification or regression also contain ≥1 fully connected layers at the end of the network. These layers are the same as in regular ANNs, as shown in **Figure 2**, and are intended to summarize all feature information into a single prediction. An example of a typical CNN architecture is shown in **Figure 4** (37). Alternatively, the architecture of CNNs can be adapted to perform image segmentation. For example, use cases in cardiovascular imaging are left or right ventricle segmentation. Such architectures, which take an image as input and directly predict an image-sized segmentation, are called fully convolutional networks (FCNs). An example of such an architecture is shown in **Figure 5**, and is called U-net, due to its U-like shape (38).

A second type of neural network architecture used in medical imaging is the recurrent neural network (RNN). RNNs are often used for sequential data, such as electrocardiography, text, or cine-MRI. An RNN takes the first instance of a sequence, makes a prediction, and then takes its own output in combination with the next instance of the sequence for subsequent predictions. This is shown in **Figure 6**.

FIGURE 4 Schematic Drawing of a 15-Layer Convolutional Neural Network

A convolutional neural network for calcium scoring in chest CT (46) , with each layer represented by a color-coded rectangle. Convolutional neural networks, especially those developed over the past few years, contain thousands of neurons. The representation shown in **Figure 2** would become extremely cluttered. Thus, in published works, block diagram representation, such as presented here, are often encountered. Abbreviation as in **Figure 1**.

FIGURE 5 Schematic Drawing of a U-Net



U-net for left ventricle segmentation in MRI, such as used by Tao et al. (55). In addition to the downsampling path, which is similar to the network shown in Figure 4, this type of architecture contains an upsampling path that incorporates both low- and high-resolution features. The **dashed arrows** indicate a concatenation of features. Abbreviation as in Figure 1.

SUPERVISED AND UNSUPERVISED LEARNING. The coefficients (or weights) in neural networks are typically initialized randomly, and their values need to be modified so the networks give a meaningful output. This modification process is what is typically referred to as "training" the neural network. This process is in essence the same as fitting a regression model to your data. Most of the studies included in this review trained deep neural networks using supervised learning, that is, during training, the neural network had access to both the input (e.g., medical image) and the reference (e.g., diagnostic label). This is referred to as a labeled dataset. In the supervised setting, neural networks are optimized by providing a network with example input images, predicting the output, calculating the error, and then computing how much each coefficient contributed to this error. Subsequently, all coefficients of the neural network are slightly updated to minimize the error. This is repeated many times until the error does not significantly decrease anymore.

In contrast, in unsupervised learning, the reference is not available during training, only the input data (e.g. medical images) are available. The data are then referred to as unlabeled data. Although this setting is less common in medical imaging, it is gaining in popularity. The main reason is that medical image data are widely available (e.g., in Picture Archiving and Communication Systems [PACS]), but accurate labels are less easy to obtain and generally require extensive expert input. By using unsupervised learning, we can still leverage this wealth of unlabeled data. We will briefly discuss one of the most popular unsupervised deep learning algorithms: generative adversarial networks (GANs).

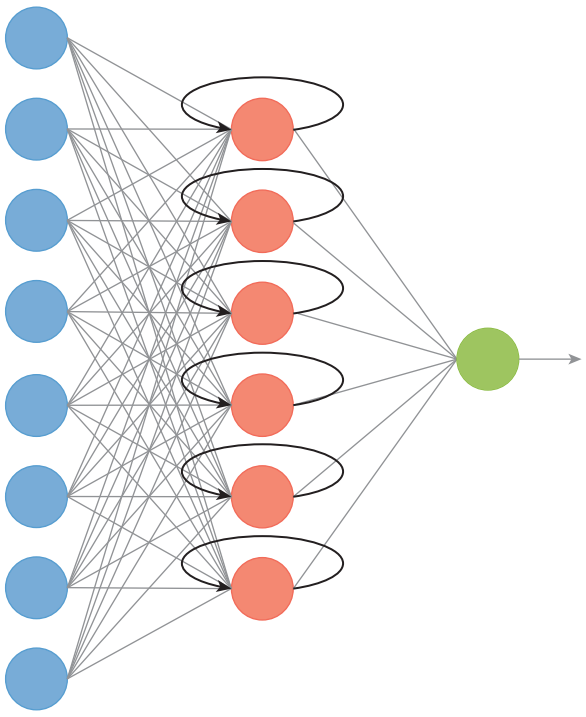
GANs are, often, used for image enhancement (e.g., noise reduction) or generation (e.g., conversion of MRI to CT). GANs consist of 2 networks that have opposing goals. Typically, 1 network acts as a generator (e.g., generating a high-dose CT image from a noisy, low-dose CT image), and the other network acts as a discriminator, assessing whether the generated image is realistic (**Figure 7**). This allows the generator to be optimized without needing exact labels (e.g., voxel-level correspondence between the CT and MRI). This has also been extended to allow the use of completely unpaired data (e.g., CTs and MRIs from different patients) via cycle GANs (39). A brief summary of all of the mentioned algorithms is provided in **Table 1**.

PRACTICAL CONSIDERATION IN TRAINING DEEP NEURAL NETWORKS. Although deep learning has shown incredible results across many fields, it remains a tricky technique to master. Given the same dataset, different researchers can obtain widely varying results. The main cause is that many decisions have to be made when training ANNs: how to pre-process the data; which network architecture to select; and how to optimize the coefficients of the network.

Leveraging task-specific knowledge often helps in optimizing deep learning models. A simple example is dealing with nonuniformity in signal intensities of MRI scanners from different manufacturers. If 1 is unaware of this effect, a network could be trained that performs well on data from 1 MRI machine, but poorly on data from another. A different example can be taken from the Kaggle Data Science Bowl from 2015 on LVEF estimation from cardiac cine-MRI (40). Some teams tried to predict the LVEF based on the image data end-to-end, but the top 3 teams realized that it was better to first segment the left ventricle using a neural network and then use this segmentation to calculate the LVEF.

When deciding on the architecture of a deep neural network, many choices need to be made. Important considerations are the type of network (e.g., ANN, CNN, RNN), the network depth (i.e., the number of layers), or the number of coefficients (i.e., number of nodes in an ANN, number of convolutional filters in a CNN). An example of typical network choices for specific tasks is shown in the **Central Illustration**. Typically, these choices are referred to as hyperparameters. There are no theoretical guarantees or proofs for many of these hyperparameters, and as such, they need to be determined empirically. However, some guidance can be obtained from domain experts or derived from available data. For example, network depth roughly correlates with the field of view (i.e., the number of pixels in each dimension the network can use in its prediction); the deeper the network, the more context it can use. However, deeper networks generally have more coefficients and thus need more training data to be adequately optimized.

FIGURE 6 Schematic Drawing of Recurrent Neural Network

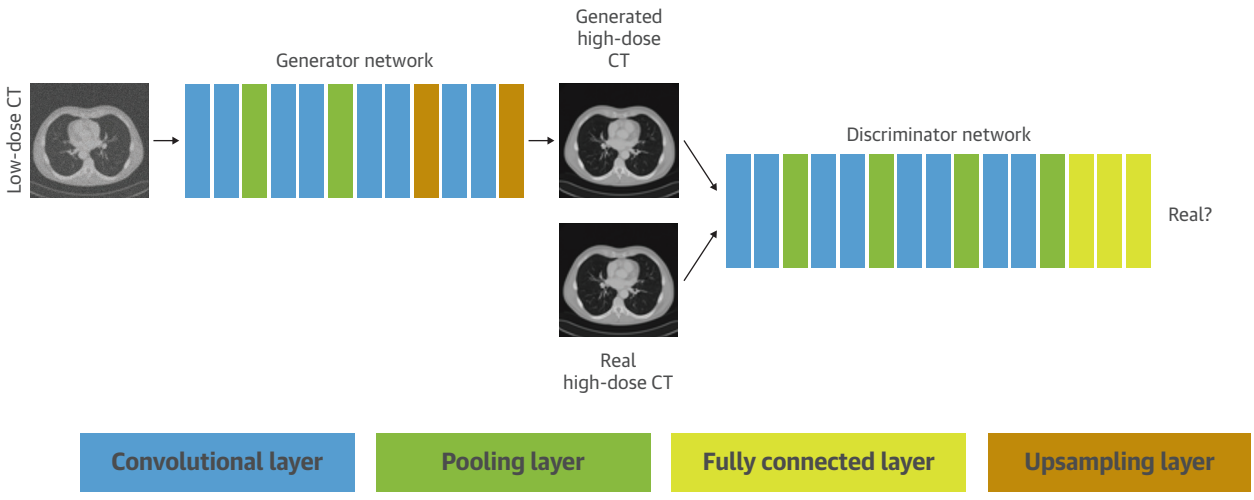


Schematic drawing of a 2-layer recurrent neural network. The **curved arrows** highlight the feedback loop in these networks, which makes them ideal for sequence data, such as electrocardiography or cine-MRI. Another use case is the tracking of vessel centerlines, as done by Wu et al. (56). Abbreviation as in **Figure 1**.

In clinical practice, some modes of variation exist with respect to data that are not present in the dataset. For example, CT images might be acquired with a different dose depending on the equipment used and the specific center. By adding different amounts of noise to the training dataset, the deep neural network can be made more robust to differences in dose without acquiring new data. This strategy is called data augmentation and is an effective method to encode certain variations within an algorithm, improving generalization to new data.

PRACTICAL CONSIDERATION IN EVALUATING DEEP NEURAL NETWORKS. An often underestimated part of algorithm development is performing a meaningful, robust evaluation (41–43). Ideally, the evaluation should result in metric estimates (e.g.,

FIGURE 7 Schematic Drawing of a Generative Adversarial Network



Generative adversarial model for generation of high-dose CT images. The generator network has to convert a low-dose, noisy CT image to a high-dose CT. The discriminator network has to differentiate between real high-dose CT images and generated images. An example is presented in Wolterink et al. (47).

diagnostic accuracy) that are clinically meaningful and generalize to real-world practice. As in all areas of medicine, the strongest evidence for single studies can be obtained via appropriately powered randomized controlled trials. However, most deep learning algorithms are not at that stage of development yet and are evaluated using retrospective cohort or case–control studies. All caveats that apply to these studies in general medical research also apply to studies using deep learning. Aspects that contribute to the likelihood of a dataset being representative are, for example, multicenter and multivendor sources.

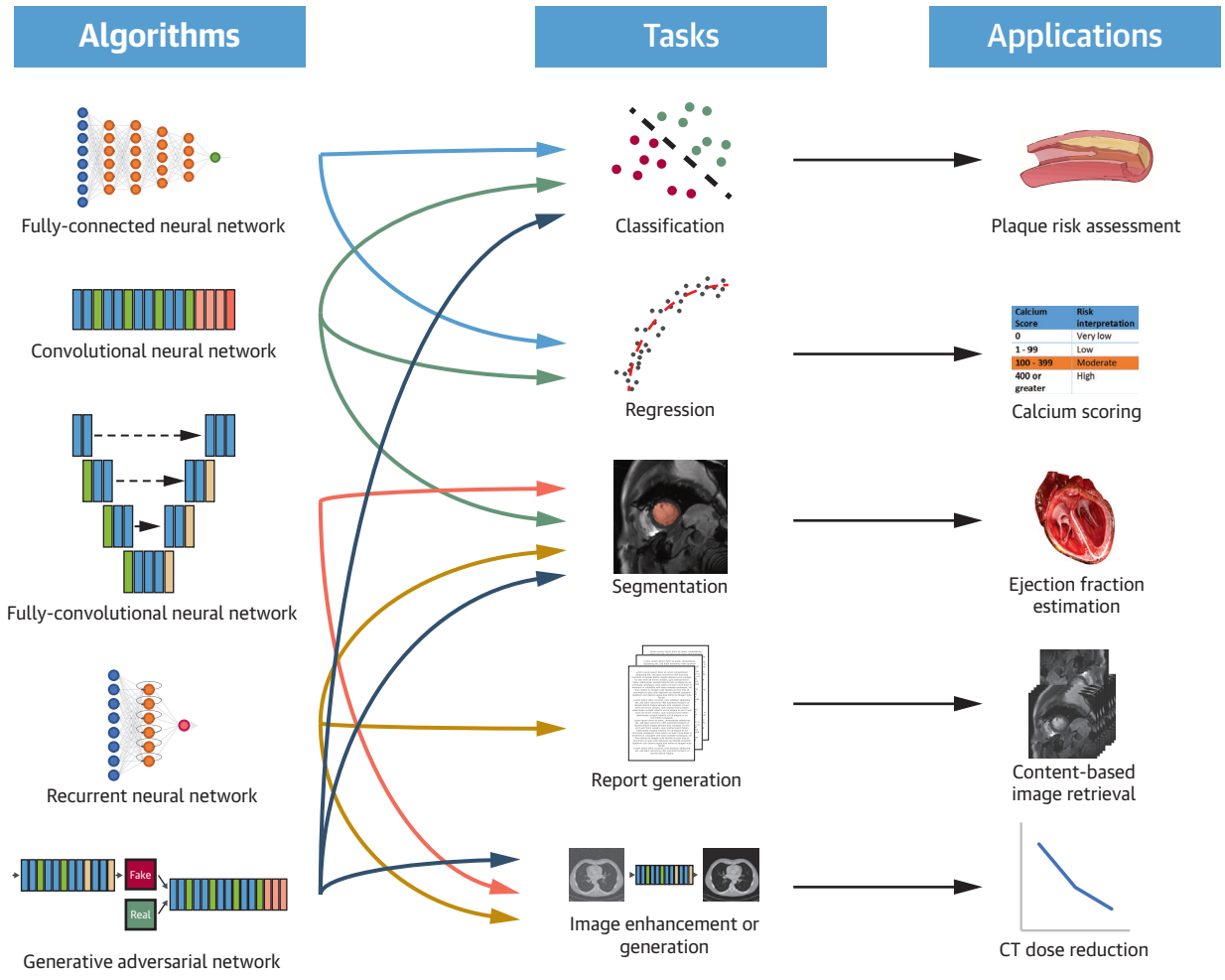
One challenge that is relatively specific to deep neural networks is that they are extremely powerful models that can memorize the training data if trained long enough, resulting in near-perfect

TABLE 1 Different Deep Learning Algorithms and Potential Applications

Method	Description	Example Applications
Fully connected neural network	A neural network in which each node in each layer is connected to all nodes in the subsequent layer (Figure 2).	These networks most often predict outcomes based on unstructured data (e.g., clinical parameters like blood pressure, heart rate, age).
Convolutional neural network	A neural network in which each node is connected to nodes in the subsequent layer in such a way that a convolution operation (Figure 3) is performed. This reduces the number of parameters in the network significantly and allows these networks to identify features independent of their position in an image. Typically, some fully connected layers are added at the end of the network to summarize all information into a single prediction (Figure 4).	These are the most common networks in cardiovascular imaging. Use-cases include calcium scoring of chest or cardiac CT (46) and quality prediction in echocardiograms (54), among many others.
Fully convolutional neural network/ U-net	Fully convolutional neural networks do not contain fully connected layers and often output complete images, for example, segmentation masks. The most common fully convolutional network in cardiovascular imaging is U-net (Figure 5).	Fully convolutional networks find most application in segmentation tasks, such as left/right ventricle segmentation (55).
Recurrent neural network	Recurrent neural networks feed their own output back as input (Figure 6), making them ideally suited for sequence data where a point in the sequence can be combined with information from previous points. Recurrent neural networks are flexible in structure and can also be combine with convolutional layers.	Wide range of use cases, for example, labeling of the coronary artery tree (56) or end-systolic and diastolic frame prediction (57).
Generative adversarial network	Consists of a generator and a discriminator, typically convolutional neural networks (Figure 7). The generator generates an image, and the discriminator predicts whether it is a real or generated image. By optimizing both in conjunction, the generator learns how to generate realistic images.	Most applications are in image enhancement and image generation, for example CT denoising (47).

CT = computed tomography.

CENTRAL ILLUSTRATION **Flowchart of Imaging Modalities, Algorithms, and Potential Applications**



Litjens, G. et al. J Am Coll Cardiol Img. 2019;12(8):1549-65.

This flowchart highlights how certain applications can be realized by using a specific algorithm. The **arrows** indicate for which application an algorithm is typically used. Note that this does not mean that, for example, a fully-connected network cannot be used for segmentation, but it is not the most appropriate choice.

accuracy. However, in general, these networks then perform poorly on new data. This phenomenon is referred to as overfitting. This can be prevented by splitting data into 3 distinct datasets: training, validation (or tuning), and testing (or holdout). The training set is used to optimize the coefficients of the network, whereas the validation or tuning dataset is used to identify overfitting and to optimize hyperparameters. The testing set should not be touched during algorithm development and only be used to obtain the performance metrics. Ideally, the testing set should contain cases from different centers or vendors to increase the likelihood of algorithm robustness in routine practice.

Finally, when setting up an evaluation study, it is also important to make sure relevant metrics are used, such that they relate to meaningful clinical parameters. For example, when optimizing neural networks for segmentation of the left ventricle, the voxel-level accuracy might not be the most useful criterion; some form of volume estimation error might be more suitable. These metrics can also be used for network training so optimization of the most meaningful metric is done.

APPLICATIONS

In this section, we provide an overview of the current state of the art in the application of deep learning

TABLE 2 Applications of Deep Learning in Cardiac Ultrasound

First Author (Ref. #)	Summary	Data			Performance	
		Dev.	Test	CV (folds)	Metric Value	Compared Against
Abdi et al. (Supplemental Ref. 12)	Quality assessment of echocardiograms; CNNs predict the quality of echo series	5,532 (f)	1,384 (f)	N	0.71 MAE	EA
Carneiro et al. (Supplemental Ref. 3)	Left ventricular segmentation; DBN embedded in system using landmarks and nonrigid registration	12	2	N	0.86 MAD	EA
Carneiro et al. (Supplemental Ref. 4)	Left ventricle tracking; extension of (3) for tracking	15	5	N	0.94 MAD	EA
Chen et al. (Supplemental Ref. 2)	Left ventricle segmentation in 4 different 2D views; uses transfer learning	34,361 (f)	8,533 (f)	N	0.864–0.961 DC	EA
Dezaki et al. (Supplemental Ref. 11)	Prediction of end-systolic and end-diastolic frames with a combination of RNN and CNN	2,470	617	N	0.98 R^2	EA
Diller et al. (Supplemental Ref. 8)	Disease classification using CNN and systemic ventricle segmentation using U-Net	159	40	N	0.794–0.881 DC	EA
Dong et al. (Supplemental Ref. 6)	Segmentation of left ventricle in 3D echocardiography using a 2D FCN and deformable models to refine coarse segmentation	12	3	N	0.88–0.90 MDC	EA
Ghesu et al. (Supplemental Ref. 7)	3D aortic valve detection and segmentation; uses shallow and deeper sparse networks	715	150	N	0.90 mm SE	EA
Jun et al. (Supplemental Ref. 9)	Classification of IVUS frames as normal vs. TCFA using CNN	100	*	N	0.911 AUC	EA after OCT/IVUS registration
Moradi et al. (Supplemental Ref. 1)	Automatic generation of text descriptions for Doppler US images of cardiac valves using doc2vec and a combination of models based on CNNs	10,253/496/226 (t)	48	N	96% Acc (valve classification)	EA
Nascimento et al. (Supplemental Ref. 5)	Left ventricular segmentation; DBN applied to patches steers multi-atlas segmentation process	14	14	14	4 pixel AD	EA
Ostvik et al. (Supplemental Ref. 54)	Classification of cardiac view into 7 classes using a real-time model based on CNN	500	500	10	98% Acc	MA
Pereira et al. (Supplemental Ref. 55)	Detection of coarctation in neonates; SVM classify features extracted with stacked denoising autoencoders	150	100	N	12.9% ER	LRC
Yu et al. (Supplemental Ref. 56)	Left ventricular segmentation; multiscale CNNs exploit temporal information	10 (s)	41 (s)	N	0.95 DC	EA
Zhang et al. (Supplemental Ref. 10)	Disease classification, cardiac chamber segmentation and viewpoints classification in echocardiograms using CNNs	14,035 (e)	8,666 (e) (m)	N	0.85–0.93 AUC	EA

If multiple metrics were reported in the paper, the most common 1 is reported to facilitate comparison.

2D = 2 dimensional; Acc = accuracy; AD = average distance; AUC = area under the curve; CNN = convolutional neural network; CV = cross-validation; DBN = deep belief network; DC = dice coefficient; Dev. = development; EA = expert annotation; (e) = echocardiograms; ER = error rate; (f) = frames; FCN = fully connected neural network; IVUS = intravascular ultrasound; LRC = labels from routine care; (m) = mixed splits for different tasks; MA = manual annotation; MAD = mean absolute distance; MAE = mean absolute error; MDC = modified dice coefficient; OCT = optical coherence tomography; RNN = recurrent neural network; (s) = sequences; SE = segmentation error; SVM = support vector machine; (t) = different types of data (e.g., text; images); TCFA = thin-cap fibroatheroma.

algorithms to cardiovascular imaging. This section is organized by imaging modality, with most papers focusing on ultrasound, CT, and MRI. Other modalities are jointly discussed.

ULTRASOUND. Applications of deep learning for cardiovascular ultrasound cover detection, classification, segmentation, tracking, and report generation (**Table 2**). Moradi et al. (Supplemental Ref. 1) presented an application in which a plausible electronic medical record was picked from a database based on a given echocardiography study. A deep model was trained to learn the relationship between the electronic medical record and the echocardiography images. Subsequently, this model was used to identify the most closely related electronic medical record from the database, given an unseen echocardiography image.

Most deep learning ultrasound applications focus on detection and segmentation. Chen et al. (Supplemental Ref. 2) segmented the left ventricle in 5 different 2D views (apical, 2-, 3-, 4-, and 5-chamber) with a CNN. They showed that using multiview CNNs improved segmentation accuracy for each specific view. Segmentation of the left ventricle of the heart was also performed by Carneiro et al. (3) in a series of studies. This first study (Supplemental Ref. 3) used an ANN model to predict landmarks for defining the segmentation contour. This approach was extended in Carneiro and Nascimento (Supplemental Ref. 4) for tracking of the left ventricle. This approach was then combined with multi-atlas registration (Supplemental Ref. 5) for improved left ventricle segmentation. Dong et al. (Supplemental Ref. 6) and Ghesu et al. (Supplemental Ref. 7) also combined deep learning with traditional methods, specifically, deformable and shape models, to segment the left ventricle and the aortic valve, respectively. Ghesu et al. (Supplemental

TABLE 3 Applications of Deep Learning in Cardiac CT

First Author (Ref. #)	Summary	Data			Performance	
		Dev.	Test	CV (folds)	Metric Value	Compared Against
Coenen et al. (Supplemental Ref. 17)	FFR prediction in cardiac CCTA; clinical evaluation of method by Manniesing et al. (16)	(s)	351	N	78% Acc	FFR
Commandeur et al. (Supplemental Ref. 13)	Adipose tissue segmentation in noncontrast CT; 2 subsequent CNNs	1,638	1,638	10	0.719 - 0.822 DC	EA
Gülsün et al. (Supplemental Ref. 57)	Coronary centerline extraction in CCTA; CNN classifies paths as correct or leakages	90	20	N	90.8%-92.7%	MA
Green et al. (Supplemental Ref. 21)	Noise and artifact reduction in low-dose CCTA; CNN for per voxel prediction of routine-dose HU values	(x)	45	N	41.47 PSNR	FD-CCTA
van Hamersvelt et al. (Supplemental Ref. 19)	Identification of patients with functionally significant stenosis in CCTA; clinical evaluation of Gandhi et al. (14)	101	101	10	0.76 AUC	FFR <0.8 or XAS >90%
Itu et al. (Supplemental Ref. 16)	FFR prediction in CCTA; ANN classification of segments based on handcrafted features	(s)	87	N	83.2% Acc	FFR
Jin et al. (Supplemental Ref. 58)	Left atrial appendage segmentation in CCTA; CNN with conditional random field	150	150	5	0.95 DC	MA
Kang et al. (Supplemental Ref. 59)	Noise and artifact reduction in low-dose CCTA; 2D cycle-consistent GAN (CycleGAN)	50	50	N	12.3 SNR	FD-CCTA
Lessmann et al. (Supplemental Ref. 20)	Calcium scoring in low-dose chest CT; 2 CNNs for voxel classification	1,181	506	N	0.91 LWK	EA
López-Lineares et al. (Supplemental Ref. 60)	Segmentation of abdominal aortic thrombus in CTA; CNN for ROI localization and CNN for segmentation	13	13	4	0.82 DC	EA
Lossau et al. (Supplemental Ref. 23)	Detection of motion artifacts in coronary CCTA; CNN for classification of coronary cross-sectional images	13	4	N	99.3% Acc	SD
Mannil et al. (Supplemental Ref. 61)	Identification of MI in noncontrast low-dose CT; texture feature classification with ANN	58	29	N	0.78 AUC	RCP
Moradi et al. (Supplemental Ref. 62)	Labeling of 2D slices from cardiac CT exams; comparison with handcrafted features	75	75	5	0.92 Acc	EA
Trullo et al. (Supplemental Ref. 63)	Organ segmentation in noncontrast CT; CNN with conditional random field	30	30	6	0.90 DC for heart	MA
de Vos et al. (Supplemental Ref. 64)	Anatomical structure localization in CT, including the heart; CNN for 2D slice classification	200	200	N	0.97 F1	EA
Wolterink et al. (Supplemental Ref. 18)	Coronary calcium detection in gated CTA; compares 3D CNN with multistream 2D CNNs	100	100	N	0.872 ICC CACMS	EA
Wolterink et al. (Supplemental Ref. 22)	Noise and artifact reduction in low-dose CT; 3D GAN	28	28	2	43 PSNR	RDCT
Wolterink et al. (Supplemental Ref. 65)	Coronary centerline extraction in CCTA; 3D CNN predicts vessel orientation and radius to guide iterative tracker	8	24	N	93.7 OV	EA
Wu et al. (Supplemental Ref. 15)	Coronary artery tree segment labeling; bi-directional LSTM in tree graph representation	436	436	10	0.87 F1	EA
Zhou et al. (Supplemental Ref. 66)	Organ segmentation in noncontrast CT; 3 2D FCNs and majority voting	228	12	N	0.82 IOU of the heart	EA
Zreik et al. (Supplemental Ref. 14)	Identification of patients with functionally significant stenosis in CCTA; auto-encoder and CNN	126	126	10	0.74 AUC	FFR
Zreik et al. (Supplemental Ref. 67)	Detection and categorization of coronary plaque and stenosis in CCTA; 3D recurrent CNN	98	65	N	0.77 Acc	EA

If multiple metrics were reported in the paper, the most common 1 is reported to facilitate comparison.

ANN = artificial neural network; CACMS = coronary artery calcification mass score; CCTA = cardiac computed tomographic angiography; CT = computed tomography; FD-CCTA = full dose cardiac CT angiography; FFR = invasive fractional flow reserve; GAN = generative adversarial network; ICC = intraclass correlation coefficient; LSTM = long short-term memory; LWK = linear weighted kappa; MI = myocardial infarction; N = no cross-validation; OV = overlap; PSNR = peak signal-to-noise ratio; RCP = regular clinical protocol; ROI = region of interest; (s) = synthetic data; SD = synthetic data; (x) = not mentioned in the paper; XAS = x-ray angiography stenosis; other abbreviations as in Table 2.

Ref. 7) evaluated their method on a large dataset of 869 patients and showed an improvement in segmentation quality of >40% over the previous state-of-the-art method.

Some studies also cover direct disease classification from echocardiography, such as Diller et al. (Supplemental Ref. 8), Jun et al. (Supplemental Ref. 9), and Zhang et al. (Supplemental Ref. 10). Remarkably, Zhang et al. (Supplemental Ref. 10) used >14,000 echocardiograms for development and validation of the entire system, which also included view classification and segmentation.

Deep learning has also been applied to time point and view classification in sequences of ultrasound images. Dezaki et al. (Supplemental Ref. 11) presented a method to predict the end systolic (ES) as well as

TABLE 4 Applications of Deep Learning in Cardiac MRI

		Data				Performance	
First Author (Ref. #)	Summary	Dev.	Test	CV (folds)	Challenges	Metric Value	Compared Against
Avendi et al. (Supplemental Ref. 27)	LV volume segmentation; sequential application of a CNN and an AE	30	15	N	su	0.94 DC	EA
Avendi et al. (Supplemental Ref. 28)	Application of (27) for the right ventricle	16	32	N	r	0.83 DC	EA
Bai et al. (Supplemental Ref. 24)	Comparison of between deep learning and human experts for cardiac segmentation	4,275	600	N		0.94 DC	EA
Bratt et al. (Supplemental Ref. 35)	Segmentation of aortic valve borders in phase-contrast MRI for flow quantification	200	270	N		0.85 ml DF	EA
Bernard et al. (Supplemental Ref. 29)	Overview of segmentation and diagnosis methods used in the Automatic Cardiac Diagnosis Challenge	100	50	N	a	COP	EA
Chakravarty et al. (Supplemental Ref. 68)	Mean curvature velocity prediction for level sets; application of an RNN	10	20	N	st	0.92 DC	EA
Chen et al. (Supplemental Ref. 69)	Patch-based regression of contour coordinates of the right ventricle	145	145	145		0.84 DC	EA
Dou et al. (34)	Segmentation of blood pool and myocardium on congenital heart disease images; CNN with deep supervision	10	10	N	h	0.93 DC	EA
Du et al. (Supplemental Ref. 70)	Prediction of contours with an FCNN. A sequential LSTM-based network predicts descriptive quantities from the contours.	145	*	N		0.87 DC	EA
Emad et al. (Supplemental Ref. 39)	Bounding box localization of LV in short-axis MRI slices; multiscale CNN.	19	14	N		0.99 Acc	MA
Fahmy et al. (Supplemental Ref. 37)	Scar tissue segmentation in short-axis CMR	833	208	N		0.57 DC	EA
Hauptmann et al. (Supplemental Ref. 71)	Residual CNN for artifact removal in undersampled k-space of highly accelerated radial CMR of patients with congenital heart disease	250	25	N		4.2-4.6 IQLS	NR
Khened et al. (Supplemental Ref. 72)	FCNN for multiscale segmentation of short-axis CMR and disease classification	1,400	590	N	a, st, k	100% Acc	EA
Kong et al. (Supplemental Ref. 41)	Prediction of ED and ES frames in cardiac cine MRI; CNN with LSTM units	420	420	4		0.4 AFD	EA
Liao et al. (Supplemental Ref. 33)	Joint segmentation and volume regression of LV volume; multitask CNN for voxel classification and full image regression	700	440	N	k	0.0106 CRPS	EAS
Luo et al. (Supplemental Ref. 32)	Direct quantification of ejection fraction fusing predictions; multiview CNN-based regression	700	440	N	k	9.6 ml RMSE	EAS
Moccia et al. (Supplemental Ref. 36)	Segmenting scar tissue in CMR late gadolinium enhancement; FCNN	30	30	30		0.88 DC	EA
Molaei et al. (Supplemental Ref. 73)	LV segmentation; regular CNN	33	*	N		0.97 Acc	MA
Ngo et al. (Supplemental Ref. 74)	Contour prediction of left ventricle; DBN (i.e., ANN) features for level sets	30	15	N	su	0.88 DC	EA
Schlemper et al. (Supplemental Ref. 43)	Reconstruction of undersampled dynamic cardiac MRI; cascade of CNNs	10	(m)	N		0.04 RMSE	FSKS
Tan et al. (Supplemental Ref. 75)	Semi-automatic approach for contour prediction of LVs; CNN-based edge distance prediction	800	540	N	l, k	0.0124 CPRS	EAS
Tan et al. (Supplemental Ref. 76)	Fully automatic approach of (75); uses separate CNNs exploiting long- and short-axis reconstructions	800	540	N	l, k	0.0122 CRPS	EAS
Tao et al. (Supplemental Ref. 77)	Impact of multicenter and multivendor data variability on LV segmentation performance; U-Net.	400	196	N		0.95 DC	EA
Vigneault et al. (Supplemental Ref. 30)	Segmentation of 5 cardiac structures; sequential application of U-Net for orientation normalization and segmentation	63/100	63/100	3/5	a	0.95 DC	EA
Winther et al. (Supplemental Ref. 78)	Evaluation of v-Net architecture for segmentation of cardiac structures	253	1,031	N	k, l, r	0.95 DC	EA
Xu et al. (Supplemental Ref. 38)	Segmentation of infarcted regions in LV without contrast agents; LSTM-RNN.	165	165	10		0.90 DC	EA
Yang et al. (Supplemental Ref. 79)	Voxel-based segmentation used as inputs to obtain contours	10/100	10/100	10/N	t	0.7 DC	EA
Yang et al. (Supplemental Ref. 26)	CNN-based label fusion for multi-atlas registration for LV myocardium segmentation	83/30	83/15	5/N		0.83 DC	EA
Yang et al. (Supplemental Ref. 42)	Detecting ED and ES time frames in free breathing MRI	20	5	N		76.5% Acc	EA
Zhang et al. (Supplemental Ref. 40)	Automatic assessment for LV coverage; CNN.	5,065	5,065	10		4.6% ER	EA
Zheng et al. (Supplemental Ref. 80)	Segmentation of left and right ventricle structures by propagating adjacent segmentations; U-Net.	3,078	902		a, su, r	0.86 DC (on a)	EA
Zotti et al. (Supplemental Ref. 81)	Segmentation with left and right ventricle shape priors; U-Net.	100	95		a, su	0.91 DC	EA

Performance is reported for a single organ/part. If multiple metrics were reported in the paper, the most common 1 is reported to facilitate comparison. *Splits not specified.

a = MICCAI Automated Cardiac Diagnosis Challenge (ACDC) (2017); AFD = average frame difference; CMR = cardiac magnetic resonance; OP = challenge overview paper; CRPS = continuous ranked probability score; DF = difference in flow; ED = end diastolic; ES = end systolic; FSKS = fully-sampled K-space; h = MICCAI Workshop on Whole-Heart and Great Vessel Segmentation from 3D Cardiovascular MRI in Congenital Heart Disease (HSVMR) (2016); IQLS = image quality on Likert scale; k = Kaggle (2015); l = MICCAI Statistical Atlases and Computational Modeling of the Heart (STACOM) Left Ventricle Segmentation Challenge (LVSC) (2011); LV = left ventricular; (m) = mixed splits for different tasks; MICCAI = Medical Image Computing and Computer Assisted Intervention; N = no cross-validation; NR = no reference; r = MICCAI Right Ventricle Segmentation Challenge (RVSC) (2012); RMSE = root mean square error; sa = MICCAI Challenge Workshop on Segmentation: Algorithms, Theory and Applications (SATA) (2013); st = MICCAI STACOM Challenges (2013); su = MICCAI STACOM LVSC/Sunnybrook (2009); other abbreviations as in Table 2.

TABLE 5 Applications of Deep Learning in Other Modalities

First Author (Ref. #)	Modality	Summary	Data			Performance	
			Dev.	Test	CV (folds)	Metric Value	Compared Against
Abdolmanafi et al. (Supplemental Ref. 49)	OCT	Classification of plaques in intravascular OCT	33	33	33	95%- 99% Acc	EA
Aviles-Rivero et al. (Supplemental Ref. 53)	Video	Prediction of heart motion in intraoperative video using restricted Boltzman machines	1,573 (i)	*	*	0.071 RMSE	PD
Betancur et al. (Supplemental Ref. 44)	SPECT	Convolutional neural networks for prediction of obstructive disease from myocardial perfusion imaging	1,638	1,638	10	0.8 AUC	ICA
Betancur et al. (Supplemental Ref. 45)	SPECT	Extension of the previous paper, now combining supine and semi-upright images for the same patient	1,160	1,160	4 (LOCO)	0.81 AUC	ICA
Breiniger et al. (Supplemental Ref. 47)	X-ray	Segmentation of stents in intraoperative x-ray image using convolutional neural networks	36 (i)	27 (i)	N	0.990-0.996 AUC	MA
Gessert et al. (Supplemental Ref. 50)	OCT	Plaque detection in intravascular OCT using CNNs	40	9	N	0.88 F1	EA
Kolluru et al. (Supplemental Ref. 51)	OCT	Plaque classification in intravascular OCT using CNNs	48	48	10	77.7%-85.3% Acc	EA
Sadda et al. (Supplemental Ref. 46)	X-ray	Real-time denoising of angiography frames using convolutional auto-encoders	16 (v)	4 (v)	N	0.97–0.99 SSIM	CIAN
Toth et al. (Supplemental Ref. 48)	X-ray	Registration of cardiac models from CT to intraoperative x-ray using convolutional neural networks	702	119	N	2.92 mm RE	KT
Yong et al. (Supplemental Ref. 52)	OCT	Segmentation and measurement of the vessel lumen using convolutional neural networks in intravascular OCT	45 (i)	19 i)	N	0.985 DC	MA

In case multiple metrics were reported, the most common 1 is presented. *Splits not specified.

CIAN = clean images with added noise; (i) = images; ICA = invasive coronary angiography; KT = known transformation; LOCO = leave-one-center-out; PD = phantom data; RE = registration error; SPECT = single-photon emission tomography; SSIM = structured similarity; (v) = video; other abbreviations as in Tables 2 and 4.

end diastolic (ED) frame in echocardiograms using a combination of RNNs and CNNs. They developed and validated their approach with 3,087 patients.

Abdi et al. (Supplemental Ref. 12) built a system for quality assessment of echocardiograms. Based on the opinion of 1 expert, who rated approximately 7,000 echograms, a regression CNN was trained to predict the quality of echo series.

CT. Deep learning was used for several applications in CT (full listing in **Table 3**). Most focused on localization and segmentation of anatomical structures for various clinical purposes in both noncontrast and coronary CT angiography (CCTA). Some examples are highlighted here. Commandeur et al. (Supplemental Ref. 13) segmented epicardial and thoracic adipose tissue in noncontrast cardiac CT scans using 2 CNNs. The first CNN segmented adipose tissue, whereas the second CNN found the pericardium. They evaluated their method in a large cohort of 1,638 patients.

The presence of a contrast agent in CCTA allows more precise segmentation of cardiac structures (e.g., volume quantification). Zreik et al. (Supplemental Ref. 14) segmented the left ventricular myocardium using a multiscale patch-based CNN, in which each voxel was classified based on 3 input patches with different scales.

Wu et al. (Supplemental Ref. 15) showed how a long short-term memory RNN could be used to automatically label segments in the coronary artery tree. Such coronary tree models could be used to identify plaque and stenosis. For example, Itu et al. (Supplemental Ref. 16) used machine learning in volumetric models of the coronary artery tree to identify reduced fractional flow reserve. Coenen et al. (Supplemental Ref. 17) showed in a clinical analysis that this method obtained accuracy levels comparable to those obtained with computational fluid dynamics-based CT fractional flow reserve. Alternatively, atherosclerosis could be detected without coronary tree extraction. Wolterink et al. (Supplemental Ref. 18) used 2 CNNs to identify image voxels representing coronary artery calcification in CCTA. The first CNN processed the full CCTA image to identify voxels that were likely to be coronary artery calcification, whereas the second CNN only classified the voxels selected by the first CNN. Zreik et al. (Supplemental Ref. 14) used deep learning–based texture features to identify patients with functionally significant stenosis. A clinical evaluation showed that this information was complementary to CCTA-derived coronary stenosis measurements (Supplemental Ref. 19). As in CCTA, deep learning could be used to detect atherosclerosis in noncontrast CT. Lessmann et al. (Supplemental Ref. 20) used 2 CNNs that progressively excluded noncalcified voxels to identify coronary artery calcification and aortic calcifications in chest CT. This system was trained and evaluated in a large cohort of 1,687 patients.

Not all studies focus on localization, segmentation, and classification. Because CT images are typically

acquired with a low radiation dose, and image quality may be affected by cardiac motion, there is a need for noise and artifact reduction. Green et al. (Supplemental Ref. 21) trained a CNN to predict routine-dose CCTA images based on low-dose CCTA images. Alternatively, Wolterink et al. (Supplemental Ref. 22) used a GAN to translate low-dose CT images into routine-dose CT images, and Kang et al. (Supplemental Ref. 9) trained a GAN with cycle-consistency (CycleGAN) to reduce noise in CCTA images. Lossau et al. (Supplemental Ref. 23) trained a CNN to detect and quantify coronary artery motion in CCTA images. The CNN was trained with samples in which motion was simulated.

MRI. As in CT, deep learning applications in MRI have mainly focused on segmentation (**Table 4**), specifically, cardiac structures in short-axis reconstructions for quantification of cardiac function. State-of-the-art deep learning segmentation methods are on par with human expert performance (Supplemental Ref. 24). Before the advent of deep learning, multi-atlas registration and deformable shape models were popular techniques for segmentation. Some initial studies, such as the those by Yang et al. (Supplemental Refs. 25,26) and Avendi et al. (Supplemental Refs. 27,28) sought to combine deep learning with multi-atlas registration and deformable models, respectively. Although the combination of more traditional techniques with deep learning was an interesting avenue of research, most state-of-the-art algorithms were based purely on CNNs. This was shown by Bernard et al. (Supplemental Ref. 29), who provided an overview of cardiac MRI analysis methods that participated in the Automatic Cardiac Diagnosis Challenge (ACDC). The challenge consisted of 2 tasks. First, automatic segmentation of the left ventricular cavity, the left ventricular myocardium, and the right ventricle in cine cardiac MRI. Second, participants were required to automatically classify patients into 5 types of pathology. Almost all participants in this challenge used CNNs, and algorithms based on 2D and 3D U-nets (**Figure 5**) performed best. Recent work by Vigneault et al. (Supplemental Ref. 30) built on these results and further improved left and right ventricle cavity segmentation.

ACDC was not the only public grand challenge on image analysis in cardiovascular MRI. The biggest, and perhaps the most well-known, was the Kaggle Data Science Bowl of 2015 (Supplemental Ref. 31), which had the goal of assessing LVEF. A plethora of methods were developed for this challenge, mostly segmentation-based (**Table 4**).

Luo et al. (Supplemental Ref. 32) used a regression-based method using multiview CNNs to estimate LVEF directly. The method used long- and short-axis images as input for a CNN that directly predicted the EF and left ventricular volumes at ED and ED. Liao et al. (Supplemental Ref. 33) proposed a hybrid approach, similar to the design by Dou et al. (Supplemental Ref. 34), that performed direct left ventricular volume quantification and segmentation at the same time, without using segmentation masks for algorithm training.

Not every study that applied deep learning to cardiovascular MRI covered segmentation of the cardiac ventricles and atria. Bratt et al. (Supplemental Ref. 35) showed that CNN-based segmentation was a faster and feasible alternative compared with manual delineation of the aortic annulus in phase contrast CMR. Some focused on segmenting scar tissue in the myocardium (Supplemental Refs. 36–38). Other investigators focused on quality control, localization, or reconstruction (Supplemental Refs. 39,40). For example, Kong et al. (Supplemental Ref. 41) and Yang et al. (Supplemental Ref. 42) used deep learning to detect ES and ED time points in cardiac cine-MRI and free-breathing MRI, respectively. Schlemper et al. (Supplemental Ref. 43) showed that a cascade of CNNs could outperform compressed sensing–based reconstruction in undersampled dynamic cardiac MRI.

OTHER MODALITIES. A complete overview of applications of deep learning in other modalities can be found in **Table 5**.

Betancur et al. (Supplemental Ref. 44,45) were the first to apply deep learning algorithms to nuclear cardiology in a pair of studies. Specifically, they used a CNN to predict obstructive disease in myocardial perfusion imaging with single-photon emission computed tomography. A key strength of their studies was the evaluation on a substantial cohort of >1,600 patients from different medical centers.

Next to nuclear cardiology, deep learning has also seen use in intraoperative x-ray. Several different applications were identified: real-time denoising of x-rays (Supplemental Ref. 46), stent segmentation (Supplemental Ref. 47), and registration of 3D preoperative cardiac models to 2D intraoperative x-ray fluoroscopy (Supplemental Ref. 48).

Recently, several studies on applications of deep learning to intravascular optical coherence tomography were published. Three papers focused on plaque detection and classification with CNNs (Supplemental Refs. 49–51). One study covered vessel lumen segmentation using CNNs, which might allow automatic assessment of stenosis (Supplemental Ref. 52). One paper detailed a deep learning method to assess heart motion from an intraoperative video

(Supplemental Ref. 53). These motion estimates could be used to guide robot-assisted surgery, removing the need for mechanical stabilizers in surgeries (e.g., off-pump coronary artery bypass grafting).

DISCUSSION

Applications of deep learning can now be found covering almost all aspects of cardiovascular imaging, from echocardiography to intraoperative fluoroscopy. Most applications can automatically determine relevant clinical parameters from cardiovascular images that could be used for diagnosis or prognosis. For example, machine-learning based segmentation of the left ventricular cavity in MRI can be used to compute EFs and identify patients with heart failure (44). Similarly, automatic identification of calcified voxels in the coronary arteries can be used to determine the coronary calcium score and identify patients at risk of cardiovascular events in CT scans (45,46).

Applications of deep learning are not limited to quantification, they can also be found in enhancing image quality (e.g., noise reduction [47], super-resolution [48]) or merging of different imaging modalities (49). This shows the breadth of possible tasks deep learning can be involved in to streamline the diagnostic workflow or improve quantification of important biomarkers.

Deep learning could also play a role in directly predicting prognosis for patients based on imaging data. Currently, this has not yet been explored, and applications have been limited to mostly estimating biomarkers (e.g., the amount of epicardial fat and the amount of coronary calcium). This is understandable, because directly predicting prognosis requires large datasets to cover all kinds of confounding factors, such as lifestyle, age, and genetic traits. Such datasets are not widely available and challenging to collect. This is a key limitation of currently published applications of deep learning in cardiovascular imaging. Most published studies cover preliminary applications without extensive validation on large, multicenter datasets. Thus, the generalization performance of these studies cannot be guaranteed. Some positive exceptions include the results based on the Kaggle Data Science Bowl and the studies by Betancur et al. (50), Lessmann et al. (46), and Zhang et al. (51), which all cover large cohorts.

We also noted a significant heterogeneity in algorithm evaluation metrics. Different investigators uses different metrics on different proprietary datasets, which made it close to impossible to fairly compare algorithms based on the data in the papers alone. Challenges, such as the Medical Image Computing and Computer Assisted Intervention ACDC Challenge (44), or the Kaggle Data Science Bowl, could play an important role because they offer standardized evaluations on the same dataset, which does allow meaningful comparison of algorithms. More high-profile, well-organized challenges could help accelerate the adoption of deep learning in clinical practice.

A different limitation is that integration of imaging data with other sources of clinical information has not been extensively studied in the context of deep learning. In practice, cardiologists have more information at their disposal for diagnostic conclusions than just the image data. Such information could similarly be analyzed using deep learning. Chamaria et al. (52) showed that integration from multiple sources of data using Bayesian network analysis could be beneficial. Choi et al. (53) showed that this was also feasible using deep learning by applying an RNN to electronic health record data. This allowed prediction of incident heart failure.

Although most publications surveyed in this review focused on MRI, this was not necessarily reflective of clinical usefulness or impact. The work on MRI has immensely benefited from the availability of several large, high-quality public datasets, in contrast to, for example, echocardiography or nuclear cardiology.

There is an important role for cardiologists and other physicians in guiding and applying deep learning research. Although several important contributions on deep learning in cardiovascular imaging were done by groups without a medical background (e.g., the winner of the Kaggle Data Science Bowl on EF prediction), clinical experts can identify the relevant areas for automation and computerized quantification. In addition, they will be the end-users of the developed algorithms and can provide guidance for system developers on how to best integrate algorithms into the clinical workflow to achieve maximum impact. Cardiologists play an important role in making sure data collection and annotation are performed adequately and can help ensure that patient rights are treated in an ethical manner.

Although the introduction of deep learning in cardiovascular imaging is having a big impact, from both the perspectives of research and industry, it is important to realize that deep learning is not a panacea. Although algorithms based on deep learning have shown human-level or even superhuman-level performance in certain medical tasks, the scope of these tasks is relatively narrow. We are nowhere close to artificial general intelligence (AGI), which is often defined as a machine that can

HIGHLIGHTS

- Deep learning has revolutionized computer vision and is now seeing application in cardiovascular imaging.
- This paper provides a thorough overview of the state of the art across applications and modalities for clinicians.
- Clinicians should guide the applications of deep learning to have the most meaningful clinical impact.

perform any task a human could do. To reach this level, many experts believe deep learning will only be part of the solution. The debate now mostly centers on how big this part will be. However, we do not need AGI for machine learning to have impact on cardiovascular imaging. Most, if not all, diagnostic images are acquired in a standardized way, to answer a diagnostic question of relatively narrow scope. Deep learning thrives in such scenarios.

As evident from this survey, deep learning and artificial intelligence have pervaded many areas of cardiovascular imaging. Many algorithms are still in the early research phases, but the first are now going through Food and Drug Administration approval and entering clinical practice. As such, it is important to perceive artificial intelligence and machine learning not as a threat, but as tool to be used to improve diagnosis, prognosis, and treatment for patients. However, it is also paramount to be wary of unsubstantiated claims that are often levied on deep learning–based image analysis tools, especially because at the moment, many are evaluated on single-center datasets of limited size. We hope this survey has, in part, contributed to increased awareness of the types of algorithms available for cardiovascular image analysis and understanding of the aspects that are important for successful clinical application.

ADDRESS FOR CORRESPONDENCE: Dr. Geert Litjens, Huispost 824, Grooteplein-Zuid 10, 6525GA Nijmegen, the Netherlands. E-mail: geert.litjens@radboudumc.nl.

REFERENCES

1. Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. Pereira F, Burges CJC, Bottou L, Weinberger KQ, editors. In Advances in Neural Information Processing Systems, Volume 25; 2012; p 1097–105

2. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature 2005;521:436–44.

3. Greenspan H, Van Ginneken B, Summers RM. Guest editorial deep learning in medical imaging: overview and future promise of an exciting new technique. IEEE Trans Med Imaging 2016;35:1153–9.

4. Litjens G, Kooi T, Bejnordi BE, et al. A survey on deep learning in medical image analysis. Med Image Anal 2017;42:60–88.

5. Esteva A, Kuprel B, Novoa RA, et al. Dermatologist-level classification of skin cancer with deep neural networks. Nature 2017;542:115–8.

6. Gulshan V, Peng L, Coram M, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA 2016;316:2402–10.

7. De Fauw J, Ledsam JR, Romera-Paredes B, et al. Clinically applicable deep learning for diagnosis and referral in retinal disease. Nat Med 2018;24:1342–50.

8. Ehteshami Bejnordi B, Veta M, van Diest P, et al. Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer. JAMA 2017;318:2199–210.

9. Levin DC, Parker L, Halpern EJ, Rao VM. Coronary CT angiography: reversal of earlier utilization trends. J Am Coll Radiol 2019;16:147–55.

10. Johnson KW, Torres Soto J, Glicksberg BS, et al. Artificial intelligence in cardiology. J Am Coll Cardiol 2018;71:2668–79.

11. Henglin M, Stein G, Hushcha PV, et al. Machine learning approaches in cardiovascular imaging. Circ Cardiovasc Imaging 2017;10:1–9.

12. Slomka PJ, Dey D, Sitek A, Motwani M, Berman DS, Germano G. Cardiac imaging: working towards fully-automated machine analysis and interpretation. Exp Rev Med Devices 2017;14:197–212.

13. Shameer K, Johnson KW, Glicksberg BS, Dudley JT, Sengupta PP. Machine learning in cardiovascular medicine: are we there yet? Heart 2018;104:1156–64.

14. Gandhi S, Mosleh W, Shen J, Chow C-M. Automation, machine learning, and artificial intelligence in echocardiography: a brave new world. Echocardiography 2018;35:1402–18.

15. Polonsky T, McClelland R, Jorgensen N, et al. Coronary artery calcium score and risk classification for coronary heart disease prediction. JAMA 2010;303:1610–6.

16. Manniesing R, Velthuis BK, van Leeuwen MS, et al. Level set based cerebral vasculature segmentation and diameter quantification in CT angiography. Med Image Anal 2006;10:200–14.

17. Feng C, Zhang S, Zhao D, Li C. Simultaneous extraction of endocardial and epicardial contours of the left ventricle by distance regularized level sets. Med Phys 2016;43:2741–55.

18. Metz CT, Schaap M, Weustink AC, Mollet NR, van Walsum T, Niessen WJ. Coronary centerline extraction from CT coronary angiography images using a minimum cost path approach. Med Phys 2009;36:5568–79.

19. Cortes C, Vapnik V. Support-vector networks. Machine Learning 1995;20:273–97.

20. Zheng Y, Barbu A, Georgescu B, Scheuering M, Comaniciu D. Four-chamber heart modeling and automatic segmentation for 3-D cardiac CT volumes using marginal space learning and steerable features. IEEE Trans Med Imaging 2008;27:1668–81.

21. Criminisi A, Shotton J, Bucciarelli S. Decision forests with long-range spatial context for organ localization in CT volumes. In Proc. MICCAI Workshop on Probabilistic Models for Medical Image Analysis. 2009;69–80.

22. Wolterink JM, Leiner T, Takx RA, Viergever MA, Išgum I. Automatic coronary calcium scoring in non-contrast-enhanced ECG-triggered cardiac CT with ambiguity detection. IEEE Trans Med Imaging 2015;34:1867–78.

23. Kolossváry M, Kellermayer M, Merkely B, Maurovich-Horvat P. Cardiac computed tomography radiomics: a comprehensive review on radiomic techniques. J Thorac Imaging 2018;33:26–34.

24. Petitjean C, Dacher J-N. A review of segmentation methods in short axis cardiac MR images. Med Image Anal 2011;15:169–84.

25. Yilmaz P, Wallecan K, Kristanto W, Aben J-P, Moelker A. Evaluation of a semi-automatic right ventricle segmentation method on short-axis MR images. J Digit Imaging 2018;31:670–9.

26. Thavendiranathan P, Liu S, Verhaert D, et al. Feasibility, accuracy, and reproducibility of real-time full-volume 3D transthoracic echocardiography to measure LV volumes and systolic function: a fully automated endocardial contouring algorithm in sinus rhythm and atrial fibrillation. J Am Coll Cardiol Img 2012;5:239–51.

27. Frangi AF, Niessen WJ, Vincken KL, Viergever MA. Multiscale vessel enhancement filtering. Med Image Comput Assist Interv 1998;1496:130–7.

28. Lesage D, Angelini ED, Bloch I, Funka-Lea G. A review of 3D vessel lumen segmentation techniques: Models, features and extraction schemes. Med Image Anal 2009;13:819–45.

29. Betancur J, Otaki Y, Motwani M, et al. Prognostic value of combined clinical and myocardial perfusion imaging data using machine learning. J Am Coll Cardiol Img 2018;11:1000–9.

30. Petitjean C, Zuluaga MA, Bai W, et al. Right ventricle segmentation from cardiac MRI: a collation study. Med Image Anal 2015;19:187–202.

31. Bernard O, Bosch JG, Heyde B, et al. Standardized evaluation system for left ventricular segmentation algorithms in 3D echocardiography. IEEE Trans Med Imaging 2016;35:967–77.

32. Schaap M, Metz CT, van Walsum T, et al. Standardized evaluation methodology and reference database for evaluating coronary artery centerline extraction algorithms. Med Image Anal 2009;13:701–14.

33. Kirisli H, Schaap M, Metz C, et al. Standardized evaluation framework for evaluating coronary artery stenosis detection, stenosis quantification and lumen segmentation algorithms in computed tomography angiography. Med Image Anal 2013;17:859–76.

34. Wolterink JM, Leiner T, De Vos BD, et al. An evaluation of automatic coronary artery calcium scoring methods with cardiac CT using the orCaScore framework. Med Phys 2016;43:2361–73.

35. Fukushima K. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. Biol Cybern 1980;36:193–202.

36. Lo S-C, Lou S-L, Lin J-S, et al. Artificial convolution neural network techniques and applications for lung nodule detection. IEEE Trans Med Imaging 1995;14:711–8.

37. Cano-Espinosa C, González G, Washko GR, Cazorla M, Estépar RSJ. Automated Agatston score computation in non-ECG gated CT scans using deep learning. Proc SPIE Int Soc Opt Eng 2018;10574:1–9.

38. Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation. Med Image Comput Assist Interv 2015;9351:234–41.

39. Zhu J.-Y, Park T, Isola P, Efros AA. Unpaired image-to-image translation using cycle-consistent adversarial networks. Presented at: IEEE International Conference on Computer Vision; October 22–29, 2017; Venice, Italy.

40. Kaggle. Second Annual Science Bowl. Transforming how we diagnose heart disease. Available at: https://www.kaggle.com/c/second-annual-data-science-bowl/data. Accessed February 1, 2019.

41. Yu L, Guo Y, Wang Y, Yu J, Chen P. Segmentation of fetal left ventricle in echocardiographic sequences based on dynamic convolutional neural networks. IEEE Trans Biomed Eng 2017;64:1886–95.

42. Emad O, Yassine IA, Fahmy AS. Automatic localization of the left ventricle in cardiac MRI images using deep learning. IEEE Eng Med Biol Soc 2015:683–6.

43. Zreik M, Lessmann N, van Hamersvelt RW, et al. Deep learning analysis of the myocardium in coronary CT angiography for identification of patients with functionally significant coronary artery stenosis. Med Image Anal 2018;44:72–85.

44. Bernard O, Lalande A, Zotti C, et al. Deep learning techniques for automatic MRI cardiac multi-structures segmentation and diagnosis: is the problem solved? IEEE Trans Med Imaging 2018;37:2514–25.

45. Wolterink JM, Leiner T, de Vos BD, et al. Automatic coronary artery calcium scoring in cardiac CT angiography using paired convolutional neural networks. Med Image Anal 2016;34:123–36.

46. Lessmann N, van Ginneken B, Zreik M, et al. Automatic calcium scoring in low-dose chest CT using deep neural networks with dilated convolutions. IEEE Trans Med Imaging 2018;37:615–25.

47. Wolterink JM, Leiner T, Viergever MA, Išgum I. Generative adversarial networks for noise reduction in low-dose CT. IEEE Trans Med Imaging 2017;36:2536–45.

48. Oktay O, Bai W, Lee M, et al. Multi-input cardiac image super-resolution using convolutional neural networks. Med Image Comput Assist Interv 2016;9902:246–54.

49. Toth D, Miao S, Kurzendorfer T, et al. 3D/2D model-to-image registration by imitation learning for cardiac procedures. Int J Comput Assist Radiol Surg 2018;13:1141–9.

50. Betancur J, Commandeur F, Motlagh M, et al. Deep learning for prediction of obstructive disease from fast myocardial perfusion SPECT: a multicenter study. J Am Coll Cardiol 2018;11:1654–63.

51. Zhang J, Gajjala S, Agrawal P, et al. Fully automated echocardiogram interpretation in clinical practice. Circulation 2018;138:1623–35.

52. Chamaria S, Johnson KW, Vengrenyuk Y, et al. Intracoronary imaging, cholesterol efflux, and transcriptomics after intensive statin treatment in diabetes. Sci Rep 2017;7:7001.

53. Choi E, Schuetz A, Stewart WF, Sun J. Using recurrent neural network models for early detection of heart failure onset. J Am Med Inform Assoc 2016;24:361–70.

54. Abdi AH, Luong C, Tsang T, et al. Automatic quality assessment of echocardiograms using convolutional neural networks: feasibility on the apical four-chamber view. IEEE Trans Med Imaging 2017;36:1221–30.

55. Tao Q, Yan W, Wang Y, et al. Deep learning-based method for fully automatic quantification of left ventricle function from cine MR images: a multivendor, multicenter study. Radiology 2019;290:81–8.

56. Wu D, Wang X, Bai J, et al. Automated anatomical labeling of coronary arteries via bidirectional tree LSTMS. Int J Comput Assis Radiol Surg 2019;14:271–80.

57. Dezaki FT, Liao Z, Luong C, et al. Cardiac phase detection in echocardiograms with densely gated recurrent neural networks and global extrema loss. IEEE Trans Medical Imaging 2018 Dec 24 [E-pub ahead of print].

58. Yong YL, Tan LK, McLaughlin RA, Chee KH, Liew YM. Linear-regression convolutional neural network for fully automated coronary lumen segmentation in intravascular optical coherence tomography. J Biomed Opt 2017;22:1–9.

59. Moradi M, Guo Y, Gur Y, Negahdar M, Syeda-Mahmood T. A cross-modality neural network transform for semi-automatic medical image annotation. Med Image Comput Assist Interv 2016;9901:300–7.

KEY WORDS artificial intelligence, cardiovascular imaging, deep learning

APPENDIX For supplemental references, please see the online version of this paper.