Deep Learning in Cardiology

Paschalis Bizopoulos, and Dimitrios Koutsouris

Abstract—The medical field is creating large amount of data that physicians are unable to comprehend and use efficiently. Moreover, rule-based expert systems are inefficient in solving complicated medical tasks or creating insights using big data. Deep learning has emerged as a more accurate and effective technology in a wide range of medical problems such as diagnosis, prediction and intervention. It is a representation learning method that consists of layers that transform the data non-linearly, thus revealing hierarchical relationships and structures. In this review we survey deep learning application papers that use structured data, signal and imaging modalities from cardiology. We discuss the advantages and limitations of applying deep learning in cardiology that also apply in medicine in general, while proposing certain directions as the most viable for clinical use.

Index Terms—deep learning, neural networks, cardiology, cardiovascular disease

I. INTRODUCTION

CARDIOVASCULAR Diseases (CVDs), are the leading cause of death worldwide accounting for 30% of deaths in 2014 in United States [1], 45% of deaths in Europe and they are estimated to cost €210 billion per year just for the European Union [2]. Physicians are still taught to diagnose CVDs based on medical history, biomarkers, simple scores, and physical examinations of individual patients which they interpret according to their personal clinical experience. Then, they match each patient to the traditional taxonomy of medical conditions according to a subjective interpretation of the medical literature. This kind of procedure is increasingly proven error-prone and inefficient. Moreover, cardiovascular technologies constantly increase their capacity to capture large quantities of data, making the work of physicians more demanding. Therefore, automating medical procedures is required for increasing the quality of health of patients and decreasing the cost of healthcare systems.

The need for automating medical procedures ranges from diagnosis to treatment and in cases where there is lack of healthcare service from physicians. Previous efforts include rule-based expert systems, designed to imitate the procedure that medical experts follow when solving medical tasks or creating insights. These systems are proven to be inefficient because they require significant feature engineering and domain knowledge to achieve adequate accuracy and they are hard to scale in presence of new data.

Machine learning is a set of artificial intelligence (AI) methods that allows computers learn a task using data, instead of being explicitly programmed. It has emerged as an effective way of using and combining biomarkers, imaging, aggregated

P. Bizopoulos and D. Koutsouris are with Biomedical Engineering Laboratory, School of Electrical and Computer Engineering, National Technical University of Athens, Greece e-mail: paschalis.bizopoulos@protonmail.com, dkoutsou@biomed.ntua.gr

clinical research from literature, and physician's notes from Electronic Health Records (EHRs) to increase the accuracy of a wide range of medical tasks. Medical procedures using machine learning are evolving from art to data-driven science, bringing insight from population-level data to the medical condition of the individual patient.

Deep learning, and its application on neural networks Deep Neural Networks (DNNs), is a set of machine learning methods that consist of multiple stacked layers and use data to capture hierarchical levels of abstraction. Deep learning has emerged due to the increase of computational power of Graphical Processing Units (GPUs) and the availability of big data and has proven to be a robust solution for tasks such as image classification [3], image segmentation [4], natural language processing [5], speech recognition [6] and genomics [7].

Advantages of DNNs against traditional machine learning techniques include that they require less domain knowledge for the problem they are trying to solve and they are easier to scale because increase in accuracy is usually achieved either by increasing the training dataset or the capacity of the network. Shallow learning models such as decision trees and Support Vector Machines (SVMs) are 'inefficient'; meaning that they require a large number of computations during training/inference, large number of observations for achieving generalizability and significant human labour to specify prior knowledge in the model [8].

In this review we present deep learning applications in structured data, signal and imaging modalities from cardiology, related to heart and vessel structures. The literature phrase search is the combined presence of each one of the cardiology terms indicated by (∗) in Table I with each one of the deep learning terms related to architecture, indicated by (+) in Table II using Google Scholar[1], Pubmed[2] and Scopus[3]. Results are then curated to match the selection criteria of the review and summarized according to two main axis: neural network architecture and the type of data that was used for training/validation/testing. Evaluations are reported for areas that used a consistent set of metrics with the same unaltered database with the same research question. Papers that do not provide information on the neural network architecture or papers that duplicate methods of previous work or preliminary papers are excluded from the review. When *multiple* is reported in the Results column in Tables IV, V, VI, VII, VIII, IX, they are reported in the main text where suitable. Additionally, for CNN architectures the use of term *layer* implies 'convolutional layer' for the sake of brevity.

TABLE I
CARDIOLOGY ACRONYMS

Acronym[a]	Meaning
ACDC	Automated Cardiac Diagnosis Challenge
ACS*	Acute Coronary Syndrome
AF*	Atrial Fibrillation
BIH	Beth Israel Hospital
BP	Blood Pressure
CAC	Coronary Artery Calcification
CAD*	Coronary Artery Disease
CHF*	Congestive Heart Failure
CT*	Computerized Tomography
CVD*	Cardiovascular Disease
DBP	Diastolic Blood Pressure
DWI*	Diffusion Weighted Imaging
ECG*	Electrocardiogram
EHR*	Electronic Health Record
FECG*	Fetal ECG
HF*	Heart Failure
HT	Hemorrhagic Transformation
HVSMR	Heart & Vessel Segmentation from 3D MRI
ICD	International Classification of Diseases
IVUS*	Intravascular Ultrasound
LV	Left Ventricle
MA	Microaneurysm
MI*	Myocardial Infarction
MMWHS	Multi-Modality Whole Heart Segmentation Challenge
MRA*	Magnetic Resonance Angiography
MRI*	Magnetic Resonance Imaging
MRP*	Magnetic Resonance Perfusion
OCT*	Optical Coherence Tomography
PCG*	Phonocardiogram
PPG*	Pulsatile Photoplethysmography
RV	Right Ventricle
SBP	Systolic Blood Pressure
SLO*	Scanning Laser Ophthalmoscopy
STACOM	Statistical Atlases & Computational Modeling of the Heart

[a](∗) denotes the term was used in the literature phrase search in combination with those from Table II.

TABLE II
DEEP LEARNING ACRONYMS

Acronym[a]	Meaning
AE+	Autoencoder
AUC	Area Under Curve
AI+	Artificial Intelligence
CNN+	Convolutional Neural Network
CRF	Conditional Random Field
DBN+	Deep Belief Network
DNN+	Deep Neural Network
FCN+	Fully Convolutional Network
FNN+	Fully Connected Network
GAN+	Generative Adversarial Network
GRU+	Gated Recurrent Unit
LSTM+	Long-Short Term Memory
MFCC	Mel-Frequency Cepstral Coefficient
MICCAI	Medical Image Computing & Computer-Assisted Intervention
PCA	Principal Component Analysis
RBM+	Restricted Boltzmann Machine
RF	Random Forest
RNN+	Recurrent Neural Network
ROI	Region of Interest
SAE+	Stacked Autoencoder
SATA	Segmentation Algorithms, Theory and Applications
SDAE+	Stacked Denoised Autoencoder
SSAE+	Stacked Sparse Autoencoder
SVM	Support Vector Machine
VGG	Visual Geometry Group
WT	Wavelet Transform

[a](+) denotes the term was used in the literature phrase search in combination with those from Table I.

At Section II, we present the fundamental concepts of neural networks and deep learning along with an overview of the architectures that have been used in cardiology. Then at Sections III, IV and V we present the deep learning applications using structured data, signal and imaging modalities from cardiology respectively. Table III provides an overview of the publicly available cardiology databases that have been used with deep learning. Finally at VI section we present the specific advantages and limitations of deep learning applications in cardiology and we conclude proposing certain directions for making deep learning applicable to clinical use.

II. NEURAL NETWORKS

A. Theory overview

Neural networks is a set of machine learning techniques initially inspired by the brain but without a primary aim to simulate it. They are function approximation methods where the input $\mathbf{x}$ is text, image, sound, generic signal, 3D volume, video (or a combination of these) and the output $\mathbf{y}$ is from the same set as $\mathbf{x}$ but with a more informative content. In mathematical terms the objective of a neural network is to find the set of parameters θ (weights $\mathbf{w}$ and biases $\mathbf{b}$):

$$f(\mathbf{x}; \theta) = \hat{\mathbf{y}} \tag{1}$$

where f is a predefined function and $\hat{\mathbf{y}}$ is the prediction. The constraint for θ is to have an as low as possible result for a cost function $J(\theta)$ between $\mathbf{y}$ and $\hat{\mathbf{y}}$.

The basic unit of neural networks is the perceptron depicted in Fig.1 which was first published by Rosenblatt [9] in 1958. It consists of a set of connections with the vector $\mathbf{x} = [x_1, \ldots, x_j, \ldots, x_n]$ as its input, while the strength of the connections, also called weights, $\mathbf{w} = [w_1, \ldots, w_j, \ldots, w_n]$ and the bias b, are learned iteratively from $\mathbf{x}$ and $\mathbf{y}$. The node's decision to fire a signal α to the next neuron or to the output is determined by the activation function ϕ, the weighted sum of $\mathbf{w}$ and $\mathbf{x}$ and the bias b, in the following way:

$$\alpha = \phi(\sum_j w_j x_j + b) \tag{2}$$

The weights $\mathbf{w}$ and biases $\mathbf{b}$ in DNNs are adjusted iteratively using gradient descent based optimization algorithms and backpropagation [10] which calculates the gradient of the cost function with respect to the parameters $\nabla_\theta J(\theta)$ [11]. Evaluating the generalization of a neural network requires splitting the initial dataset $D = (\mathbf{x}, \mathbf{y})$ into three non-overlapping datasets:

$$D_{train} = (\mathbf{x}_{train}, \mathbf{y}_{train}) \tag{3}$$

$$D_{validation} = (\mathbf{x}_{validation}, \mathbf{y}_{validation}) \tag{4}$$

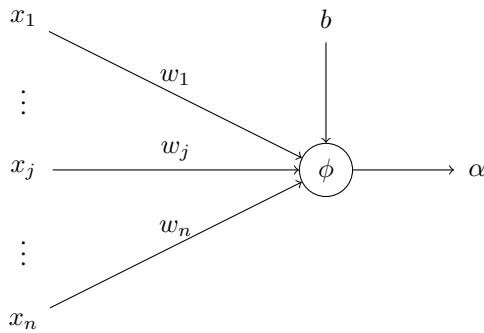$$D_{test} = (\mathbf{x}_{test}, \mathbf{y}_{test}) \tag{5}$$

Fig. 1. Perceptron. It consists of a set of connections $x_{1,...,n}$ as its input, the weights $w_{1,...,n}$, the bias b, the activation function ϕ and the output α.

D_{train} is used for adjusting $\mathbf{w}$ and $\mathbf{b}$ that minimize the chosen cost function while $D_{validation}$ is used for choosing the hyperparameters of the network. D_{test} is used to evaluate the generalization of the network and it should ideally originate from different machines/patients/organizations depending on the research question that is targeted.

The performance of DNNs has improved using the Rectified Linear Unit (ReLU) as an activation function compared to the logistic sigmoid and hyperbolic tangent [12]. The activation function of the last layer depends on the nature of research question to be answered (e.g. softmax for classification, sigmoid for regression).

The cost functions that are used in neural networks also depend on the task to be solved. Cross entropy quantifies the difference between the true and predicted probability distributions and is usually chosen for detection and classification problems. The Area Under the receiver operating Curve (AUC) represents the probability that a random pair of normal and abnormal pixels/signals/images will be correctly classified [13] and is used in binary segmentation problems. Dice coefficient [14] is a measure of similarity used in segmentation problems and its values range between zero (total mismatch) to unity (perfect match).

B. Architectures overview

Fully Connected Networks (FNNs) are networks that consist of multiple perceptrons stacked in width and depth, meaning that every unit in each layer is connected to every unit in the layers immediately before and after. Although it has been proven [15] that one layer FNNs with sufficient number of hidden units are universal function approximators, they are not computationally efficient for fitting complex functions. Deep Belief Networks (DBNs) [16] are stacked Restricted Boltzmann Machines (RBMs) where each layer encodes statistical dependencies among the units in the previous layer; they are trained to maximize the likelihood of the training data.

Convolutional Neural Networks (CNNs), as shown in Fig. 2, consist of a convolutional part where hierarchical feature extraction takes place (low-level features such as edges and corners and high-level features such as parts of objects) and a fully connected part for classification or regression, depending on the nature of the output y. Convolutional layers are much better feature optimizers that utilize the local relationships in the data, while fully connected layers are good classifiers, thus they are used as the last layers of a CNN. Additionally, convolutional layers create feature maps using shared weights that have a fixed number of parameters in contrast with fully connected layers, making them much faster. VGG [17] is a simple CNN architecture that utilizes small convolutional filters (3×3) and performance is increased by increasing the depth of the network. GoogleNet [18] is another CNN-like architecture that makes use of the inception module. The inception module uses multiple convolutional layers in parallel from which the result is the concatenated, thus allowing the network to learn multiple level features. ResNet [19] is a CNN-like architecture that formulates layers as learning residual functions with reference to the layer inputs, allowing training of much deeper networks.

Autoencoders (AEs) are neural networks that are trained with the objective to copy the input x to the output in such a way that they encode useful properties of the data. It usually consists of an encoding part that downsamples the input down to a linear feature and a decoding part that upsamples to the original dimensions. A common AE architecture is Stacked Denoised AE (SDAE) that has an objective to reconstruct the clean input from an artificially corrupted version of the input [20] which prevents the model from learning trivial solutions. Another AE-like architecture is u-net [4], which is of special interest to the biomedical community since it was first applied on segmentation of biomedical images. U-net introduced skip connections that connect the layers of the encoder with corresponding ones from the decoder.

Recurrent Neural Networks (RNNs) are networks that consist of feedback loops and in contrast to previously defined architectures they can use their internal state to process the input. Vanilla RNNs have the vanishing gradients problem and for that reason Long-Short Term Memory (LSTM) [21] was proposed as a solution to storing information over extended time. Gated Recurrent Unit (GRU) [22] was later proposed as a simpler alternative to LSTM.

III. DEEP LEARNING USING STRUCTURED DATA

Structured data mainly include EHRs and exist in an organized form based on data fields typically held in relational databases. A summary of deep learning applications using structured data is shown in Table IV.

RNNs have been used for cardiovascular disease diagnosis using structured data. In [66] the authors predict the BP during surgery and length of stay after the surgery using LSTM. They performed experiments on a dataset of 12036 surgeries that contain information on intraoperative signals (body temperature, respiratory rate, heart rate, DBP, SBP, fraction of inspired O_2 and end-tidal CO_2), achieving better results than KNN and SVM baselines. Choi et al. [67] trained a GRU with longitudinal EHR data, detecting relations among time-stamped events (disease diagnosis, medication orders, etc.) using an observation window. They diagnose Heart Failure (HF) achieving AUC 0.777 for a 12 month window and 0.883 for 18 month window, higher than the MLP, SVM and
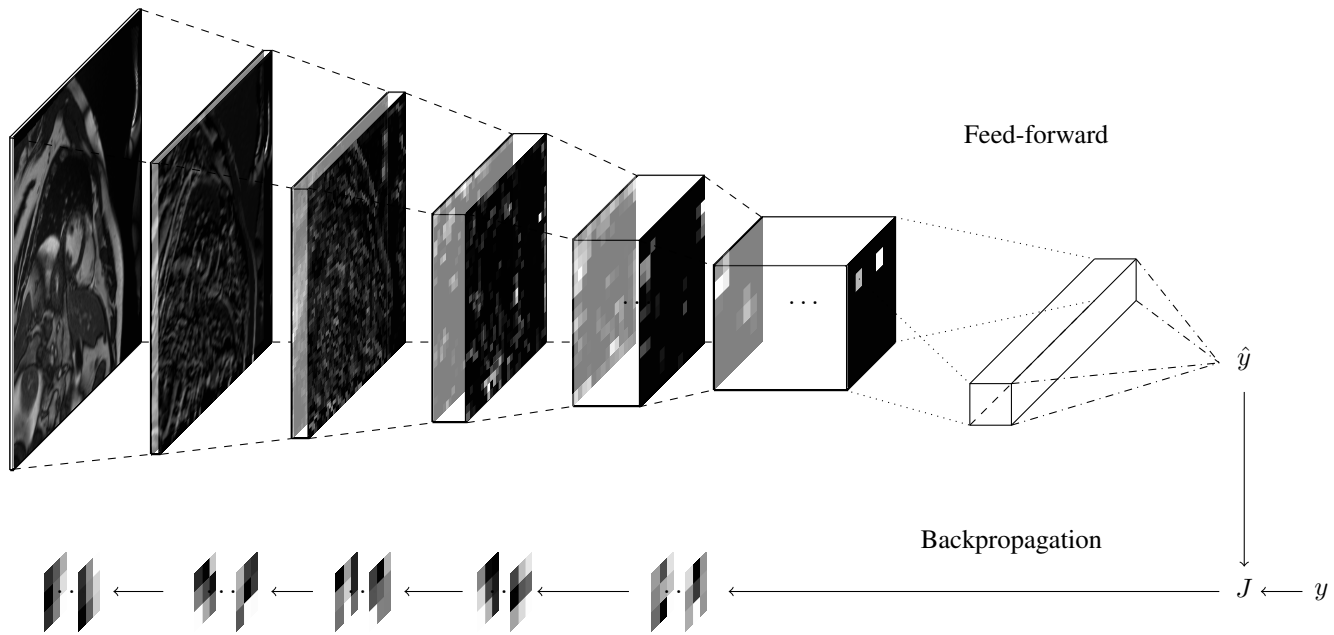
Fig. 2. A Convolutional Neural Network that calculates the LV area ($\hat{y}$) from an MRI image (x). The pyramidoid structure on top denotes the flow of the feed-forward calculations starting from input image x through the set of feature maps depicted as 3D rectangulars to the output $\hat{y}$. The height and width of the set of feature maps is proportional to the height and width of the feature maps while the depth is proportional to the number of the feature maps. The arrows at the bottom denote the flow of the backpropagation starting after the calculation of the loss using the cost function J, the original output y and the predicted output $\hat{y}$. This loss is backpropagated through the filters of the network adjusting their weight. Dashed lines denote a 2D convolutional layer with ReLU and Max-Pooling (which also reduces the height and width of the feature maps), the dotted line denotes the fully connected layer and the dash dotted lines at the end denote the sigmoid layer. For visualization purposes only a few of the feature maps and filters are shown, and they are also not in scale.

KNN baselines. Purushotham et al. [68] compared the super learner (ensemble of shallow machine learning algorithms) [69] with FNN, RNN and a multimodal deep learning model proposed by the authors on the MIMIC database. The proposed framework uses FNN and GRU for handling non-temporal and temporal features respectively, thus learning their shared latent representations for prediction. The results show that deep learning methods consistently outperform the super learner in the majority of the prediction tasks of the MIMIC (predictions of in-hospital mortality AUC 0.873, short-term mortality AUC 0.871, long-term mortality AUC 0.87 and ICD-9 code AUC 0.777). Kim et al. [70] created two medical history prediction models using attention networks and evaluated them on 50000 hypertension patients. They showed that using a bi-directional GRU-based model provides better discriminative capability than the convolutional-based model which has shorter training time with competitive accuracy.

AEs have been used for cardiovascular disease diagnosis using structured data. Hsiao et al. [71] trained an AE and a softmax layer for risk analysis of four categories of cardiovascular diseases. The input included demographics, ICD-9 codes from outpatient records and concentration pollutants, meteorological parameters from environmental records. Huang et al. [72] trained a SDAE using an EHR dataset of 3464 patients to predict ACS. The SDAE has two regularization constraints that makes reconstructed feature representations contain more risk information thus capturing characteristics of patients at similar risk levels, and preserving the discriminating information across different risk levels. Then, they append a softmax layer, which is tailored to the clinical risk prediction problem.

DBNs have also been used in combination with structured data besides RNNs and AEs. In [73] the authors first performed a statistical analysis of a dataset with 4244 records to find variables related to cardiovascular disease from demographics and lifestyle data (age, gender, cholesterol, high-density lipoprotein, SBP, DBP, smoking, diabetes). Then, they developed a DBN model for predicting cardiovascular diseases (hypertension, hyperlipidemia, Myocardial Infarction (MI), angina pectoris). They compared their model with Naive Bayes, Logistic Regression, SVM, RF and a baseline DBN achieving better results.

According to the literature, RNNs are widely used in cardiology structured data because they are capable in finding optimal temporal features better than other deep/machine learning methods. On the other hand, applications in this area are relatively few and this is mainly because there is a small number of public databases in this area which prevents further evaluation and comparison of different architectures on these datasets. Additionally, structured databases by their design contain less information for an individual patient and focus more on groups of patients, making them more suitable for epidemiologic studies rather than cardiology.

IV. DEEP LEARNING USING SIGNALS

Signal modalities include time-series such as Electrocardiograms (ECGs), Phonocardiograms (PCGs), oscillometric and wearable data. One reason that traditional machine learning

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/RBME.2018.2885714, IEEE Reviews in Biomedical Engineering

5

TABLE III
CARDIOLOGY PUBLIC DATABASES

Database Name[a]	Acronym	Patients	Task
Structured Databases			
Medical Information Mart for Intensive Care III [23]	MIMIC	38597	53423 hospital admissions for ICD-9 and mortality prediction
KNHANES-VI [24]	KNH	8108	epidemiology tasks with demographics, blood tests, lifestyle
Signal Databases (all ECG besides [25])			
IEEE-TBME PPG Respiratory Rate Benchmark Dataset [25]	PPGDB	42	respiratory rate estimation using PPG
Creighton University Ventricular Tachyarrhythmia [26]	CREI	35	ventricular tachyarrhythmia detection
MIT-BIH Atrial Fibrillation Database [27]	AFDB	25	AF prediction
BIH Deaconess Medical Center CHF Database [28]	CHFDB	15	CHF classification
St.Petersburg Institute of Cardiological Technics [29]	INDB	32	QRS detection and ECG beat classification
Long-Term Atrial Fibrillation Database [30]	LTAFDB	84	QRS detection and ECG beat classification
Long-Term ST Database [31]	LTSTDB	80	ST beat detection and classification
MIT-BIH Arrhythmia Database [32]	MITDB	47	arrhythmia detection
MIT-BIH Noise Stress Test Database [33]	NSTDB	12	used for noise resilience tests of models
MIT-BIH Normal Sinus Rhythm Database [29]	NSRDB	18	arrhythmia detection
MIT-BIH Normal Sinus Rhythm RR Interval Database [34]	NSR2DB	54	ECG beat classification
Fantasia Database [35]	FAN	40	ECG beat classification
AF Classification short single lead ECG Physionet 2017 [32]	PHY17	—[b]	ECG beat classification (12186 single lead records)
Physionet 2016 Challenge [36]	PHY16	1297	heart sound classification using PCG (3126 records)
Physikalisch-Technische Bundesanstalt ECG Database [37]	PTBDB	268	cardiovascular disease diagnosis
QT Database [38]	QTDB	105	QT beat detection and classification
MIT-BIH Supraventricular Arrhythmia Database [39]	SVDB	78	supraventricular arrhythmia detection
Non-Invasive Fetal ECG Physionet Challenge Dataset [40]	PHY13	447	measurement of fetal HR, RR interval and QT
DeepQ Arrhythmia Database [41][c]	DeepQ	299	ECG beat classification(897 records)
MRI Databases			
MICCAI 2009 Sunnybrook [42]	SUN09	45	LV segmentation
MICCAI 2011 Left Ventricle Segmentation STACOM [43]	STA11	200	LV segmentation
MICCAI 2012 Right Ventricle Segmentation Challenge [44]	RV12	48	RV segmentation
MICCAI 2013 SATA [45]	SAT13	—[b]	LV segmentation
MICCAI 2016 HVSMR [46]	HVS16	20	whole heart segmentation
MICCAI 2017 ACDC [47]	AC17	150	LV/RV segmentation
York University Database [48]	YUDB	33	LV segmentation
Data Science Bowl Cardiac Challenge Data [49]	DS16	1140	LV volume estimation after systole and diastole
Retina Databases (all Fundus besides [50])			
Digital Retinal Images for Vessel Extraction [51]	DRIVE	40	vessel segmentation in retina
Structured Analysis of the Retina [52]	STARE	20	vessel segmentation in retina
Child Heart and Health Study in England Database [53]	CHDB	14	blood vessel segmentation in retina
High Resolution Fundus [54]	HRF	45	vessel segmentation in retina
Kaggle Retinopathy Detection Challenge 2015 [55]	KR15	—[b]	diabetic retinopathy classification
TeleOptha [56]	e-optha	381	MA and hemorrhage detection
Messidor [57]	Messidor	1200	diabetic retinopathy diagnosis
Messidor2 [57]	Messidor2	874	diabetic retinopathy diagnosis
Diaretdb1 [58]	DIA	89	MA and hemorrhage detection
Retinopathy Online Challenge [59]	ROC	100	MA detection
IOSTAR [50]	IOSTAR	30	vessel segmentation in retina using SLO
RC-SLO [50]	RC-SLO	40	vessel segmentation in retina using SLO
Other Imaging Databases			
MICCAI 2011 Lumen+External Elastic Laminae [60]	IV11	32	lumen and external contour segmentation in IVUS
UK Biobank [61]	UKBDB	—[b]	multiple imaging databases
Coronary Artery Stenoses Detection and Quantification [62]	CASDQ	48	cardiac CT angiography for coronary artery stenoses
Multimodal Databases			
VORTAL [63]	VORTAL	45	respiratory rate estimation with ECG and PCG
Left Atrium Segmentation Challenge STACOM 2013 [64]	STA13	30	left atrium segmentation with MRI, CT
MICCAI MMWHS 2017 [65]	MM17	60	120 images for whole heart segmentation with MRI, CT

[a]URLs for each database are provided in the reference section.
[b]The number of patients was not reported.
[c]Authors mention that they plan to make the database publicly available.

TABLE IV
DEEP LEARNING APPLICATIONS USING STRUCTURED DATA

Reference	Method	Application/Notes[a]	Result[b]
Gopalswamy 2017 [66]	LSTM	predict BP and length of stay using multiple vital signals (private)	73.1%
Choi 2016 [67]	GRU	predict initial diagnosis of HF using a GRU with an observation window (private)	0.883[b]
Purushotham 2018 [68]	FNN, GRU	prediction tasks of MIMIC using a FNN and a GRU-based network (MIMIC)	*multiple*
Kim 2017 [70]	GRU, CNN	predict onset of high risky vascular disease using a bidirectional GRU and a 1D CNN (private)	*multiple*
Hsiao 2016 [71]	AE	analyze CVD risk using an AE and a softmax on outpatient and meteorological dataset (private)	—[b]
Kim 2017 [73]	DBN	predict cardiovascular risk using a DBN (KNH)	83.9%
Huang 2018 [72]	SDAE	predict ACS risk using a SDAE with two regularization constrains and a softmax (private)	73.0%

[a]In parenthesis the databases used.

[b]There is a wide variability in results reporting. All results are accuracies besides [67] which report AUC and [71] which is a statistical study.

have worked sufficiently well in this area in previous years is because of the use of handcrafted and carefully designed features by experts such as statistical measures from the ECG beats and the RR interval [74]. Deep learning can improve results when the annotations are noisy or when it is difficult to manually create a model. A summary of deep learning applications using signals is shown in Tables V and VI.

A. Electrocardiogram

ECG is the method of measuring the electrical potentials of the heart to diagnose heart related problems [102]. It is non-invasive, easy to acquire and provides a useful proxy for disease diagnosis. It has mainly been used for arrhythmia detection utilizing the large number of publicly available ECG databases as shown in Table III.

1) Arrhythmia detection with MITDB: CNNs have been used for arrhythmia detection with MITDB. Zubair et al. [75] detected the R-peak using a non-linear transformation and formed a beat segment around it. Then, they used the segments to train a three *layer* 1D CNN with variable learning rate depending on the mean square error and achieved better results than previous state-of-the-art. Li et al. [76] used WT to remove high frequency noise and baseline drift and biorthogonal spline wavelet for detecting the R-peak. Then, they created and resampled segments around the R-peak before feeding them to a two *layer* 1D CNN. In their article Kiranyaz et al. [77] trained patient-specific CNNs that can be used to classify long ECG data stream or for real-time ECG monitoring and early alert system on a wearable device. The CNN consisted of three *layers* of an adaptive implementation of 1D convolution layers. They achieved 99% and 97.6% in classifying ventricular and supraventricular ectopic beats respectively. In [78] the authors used mean removal for dc removal, moving average filter for high frequency removal, derivative-based filter for baseline wander removal and a comb filter for power line noise removal. They detected QRS with Pan-Tompkins algorithm [103], extracted segments using samples after the R-peak and converted them to $256 \times 256 \times 3$ binary images. The images were then fed to an AlexNet feature extractor trained on ImageNet and then to Principal Component Analysis (PCA). They achieved high accuracy classifying three types of arrhythmias of MITDB.

AEs have also been used for arrhythmia detection with MITDB. In their article Luo et al. [79] utilized quality assessment to remove low quality heartbeats, two median filters for removing power line noise, high-frequency noise and baseline drift. Then, they used a derivative-based algorithm to detect R-peaks and time windows to segment each heartbeat. Modified frequency slice WT was used to calculate the spectrogram of each heartbeat and a SDAE for extracting features from the spectrogram. Then, they created a classifier for four arrhythmias from the encoder of the SDAE and a softmax, achieving an overall accuracy of 97.5%. In [80] the authors denoised the signals with a low-pass, a bandstop and a median filter. They detected R-peaks using Pan-Tomkins algorithm and segmented/resampled the heartbeats. Features were extracted from the heartbeat signal using a SDAE, and a FNN was used to classify the heartbeats in 16 types of arrhythmia. Comparable performance with previous methods based on feature engineering was achieved. Yang et al. [81] normalized the ECG and then fed it to a Stacked Sparse AE (SSAE) which they fine-tuned. They classify on six types of arrhythmia achieving accuracy of 99.5% while also demonstrating the noise resilience of their method with artificially added noise.

DBNs have also been used for this task besides CNNs and AEs. Wu et al. [82] used median filters to remove baseline wander, a low-pass filter to remove power-line and high frequency noise. They detected R-peaks using ecgpuwave software from Physionet and segmented and resampled ECG beats. Two types of RBMs, were trained for feature extraction from ECG for arrhythmia detection. They achieved 99.5% accuracy on five classes of MITDB.

2) Arrhythmia detection with other databases: CNNs have been used for arrhythmia detection using other databases besides solely on MITDB. In [83] the authors created a two *layer* CNN using the DeepQ [41] and MITDB to classify four arrhythmias types. The signals are heavily preprocessed with denoising filters (median, high-pass, low-pass, outlier removal) and they are segmented to 0.6 seconds around the R-peak. Then, they are fed to the CNN along with the RR interval for training. The authors also employ an active learning method to achieve personalized results and improved precision, achieving high sensitivity and positive predictivity in both datasets. Rajpurkar et al. [84] created an ECG wearable dataset that contains the largest number of unique patients (30000) than

TABLE V
DEEP LEARNING APPLICATIONS USING ECG

Reference	Method	Application/Notes[a]	Accuracy[b]
arrhythmia detection (MITDB)			
Zubair 2016 [75]	CNN	non-linear transform for R-peak detection and a 1D CNN with a variable learning rate	92.7%
Li 2017 [76]	CNN	WT for denoising and R-peak detection and a two *layer* 1D CNN	97.5%
Kiranyaz 2016 [77]	CNN	patient-specific CNN using adaptive 1D convolutional layers	99%,97.6%[b]
Isin 2017 [78]	CNN	denoising filters, Pan-Tomkins, AlexNet for feature extraction and PCA for classification	92.0%[b]
Luo 2017 [79]	SDAE	denoising filters, derivative-based R-peak detection, WT, SDAE and softmax	97.5%
Jiang 2017 [80]	SDAE	denoising filters, Pan-Tomkins, SDAE and FNN	97.99%
Yang 2017 [81]	SSAE	normalize ECG, fine-tuned SSAE	99.45%
Wu 2016 [82]	DBN	denoising filters, ecgpuwave, two types of RBMs	99.5%[b]
arrhythmia detection			
Wu 2018 [83]	CNN	active learning and a two *layer* CNN fed with ECG and RR interval (MITDB, DeepQ)	*multiple*
Rajpurkar 2017 [84]	CNN	34-layer CNN (private wearable dataset)	80%[b]
Acharya 2017 [85]	CNN	four *layer* CNN (AFDB, MITDB, Creighton)	92.5%
Schwab 2017 [86]	RNN	ensemble of RNNs with an attention mechanism (PHY17)	79%
AF detection			
Yao 2017 [87]	CNN	multiscale CNN (AFDB, LTAFDB, private)	98.18%
Xia 2018 [88]	CNN	CNN with spectrograms from short time fourier transform or stationary WT (AFDB)	98.29%
Andersen 2018 [89]	CNN, LSTM	RR intervals with a CNN-LSTM network (MITDB, AFDB, NSRDB)	87.40%
Xiong 2015 [90]	AE	scale-adaptive thresholding WT and a denoising AE (MITDB, NSTDB)	∼18.7[b]
Taji 2017 [91]	DBN	false alarm reduction during AF detection in noisy ECG signals (AFDB, NSTDB)	87%[b]
Other tasks			
Xiao 2018 [92]	CNN	classify ST events from ECG using transfer learning on Inception v3 (LTSTDB)	0.867[b]
Rahhal 2016 [93]	SDAE	SDAE with sparsity constrain and softmax (MITDB, INDB, SVDB)	>99%[b]
Abrishami 2018 [94]	Multiple	compared a FNN, a CNN and a CNN with dropout for ECG wave localization (QTDB)	96.2%
Wu 2016 [95]	SAE	detect and classify MI using a SAE and multi-scale discrete WT (PTBDB)	∼99%[b]
Reasat 2017 [96]	Inception	detect MI using Inception block for each ECG lead (PTBDB)	84.54%
Zhong 2018 [97]	CNN	three *layer* CNN for classifying fetal ECG segments (PHY13)	77.85%
Other tasks (private databases)			
Ripoll 2016 [98]	RBM	identify abnormal ECG using pretrained RBMs	85.52%
Jin 2017 [99]	CNN	identify abnormal ECG using lead-CNN and rule inference	86.22%
Liu 2018 [100]	Multiple	compared Inception and a 1D CNN for premature ventricular contraction in ECG	88.5%
Hwang 2018 [101]	CNN, RNN	detect stress with an one convolutional layer with dropout and two RNNs on	87.39%[b]

[a]In parenthesis the databases used by paper or by papers in subsection.

[b]There is a wide variability in results reporting. The results of [77] is for ventricular/supraventricular ectopic beats, [78] is for three types of arrhythmias, [82] is for five types of arrhythmias, [84] report precision, [90] report SNR and multiple results depending on added noise, the result of [91] is without noise in a noise resilience study, [92] report AUC, [93] report multiple accuracies for supraventricular/ventricular ectopic beats, [95] report sensitivity and specificity, [101] report results for two cases.

previous datasets and used it to train a 34-layer residual-based CNN. Their model detects a wide range of arrhythmias a total of 14 output classes, outperforming the average cardiologist. In their article Acharya et al. [85] trained a four *layer* CNN on AFDB, MITDB and CREI, to classify between normal, AF, atrial flutter and ventricular fibrillation. Without detecting the QRS they achieved comparable performance with previous state-of-the-art methods that were based on R-peak detection and feature engineering. The same authors have also trained the previous CNN architecture for identifying shockable and non-shockable ventricular arrhythmias [104], identify CAD patients with FAN and INDB [105], classify CHF with CHFDB, NSTDB, FAN [106] and also tested its noise resistance with WT denoising [107].

An application of RNNs in this area is from Schwab et al.

[86] that built an ensemble of RNNs that distinguishes between normal sinus rhythms, AF, other types of arrhythmia and noisy signals. They introduced a task formulation that segments ECG into heartbeats to reduce the number of time steps per sequence. They also extended the RNNs with an attention mechanism that enables them to reason which heartbeats the RNNs focus on to make their decisions and achieved comparable to state-of-the-art performance using fewer parameters than previous methods.

3) AF detection: CNNs have been used for AF detection. Yao et al. [87] extracted instant heart rate sequence, which is fed to an end-to-end multi-scale CNN that outputs the AF detection result, achieving better results than previous methods in terms of accuracy. Xia et al. [88] compared two CNNs, with three and two *layers*, that were fed with spectrograms of

signals from AFDB using Short-Term Fourier Transform and stationary WT respectively. Their experiments concluded that the use of stationary WT achieves a slightly better accuracy for this task.

Besides CNNs other architectures have been used for AF detection. Andersen et al. [89] converted ECG signals from AFDB, into RR intervals to classify them for AF detection. Then, they segmented the RR intervals to 30 samples each and fed them to a network with two *layers* followed by a pooling layer and a LSTM layer with 100 units. The method was validated on MITDB and NSRDB achieving an accuracy that indicates its generalizability. In [90] the authors added noise signals from the NSTDB to the MITDB and then used scale-adaptive thresholding WT to remove most of the noise and a denoised AE to remove the residual noise. Their experiments indicated that increasing the number of training data to 1000 the signal-to-noise ratio increases dramatically after denoising. Taji et al. [91] trained a DBN to classify acceptable from unacceptable ECG segments to reduce the false alarm rate caused by poor quality ECG during AF detection. Eight different levels of ECG quality are provided by contaminating ECG with motion artifact from the NSTDB for validation. With an SNR of $-20dB$ in the ECG signal their method achieved an increase of 22% in accuracy compared to the baseline model.

4) Other tasks with public databases: ECG beat classification was also performed by a number of studies using public databases. In [92] the authors finetuned a Inception v3 trained on ImageNet, using signals from LTSTDB for classifying ST events. The training samples were over 500000 segments of ST and non-ST ECG signals with ten second duration that were converted to images. They achieve comparable performance with previous complex rule-defined methods. Rahhal et al. [93] trained a SDAEs with sparsity constraint and a softmax for ECG beat classification. At each iteration the expert annotates the most uncertain ECG beats in the test set, which are then used for training, while the output of the network assigns the confidence measures to each test beat. Experiments performed on MITDB, INDB, SVDB indicate the robustness and computational efficiency of the method. In [94] the authors trained three separate architectures to identify the P-QRS-T waves in ECG with QTDB. They compared a two layer FNN, a two *layer* CNN and a two *layer* CNN with dropout with the second one achieving the best results.

ECG has also been used for MI detection and classification. In their article Wu et al. [95] detected and classified MI with PTBDB. They used multi-scale discrete WT to facilitate the extraction of MI features at specific frequency resolutions and softmax regression to build a multi-class classifier based on the learned features. Their validation experiments show that their method performed better than previous methods in terms of sensitivity and specificity. PTBDB was also used by Reasat et al. [96] to train an inception-based CNN. Each ECG lead is fed to an inception block, followed by concatenation, global average pooling and a softmax. The authors compared their method with a previous state-of-the-art method that uses SWT, demonstrating better results.

Fetal QRS complexes were identified by a three *layer* CNN

with dropout by Zhong et al. [97] with PHY13. First, the bad quality signals are discarded using sample entropy and then normalized segments with duration of 100ms are fed to the CNN for training. The authors compared their method with KNN, Naive Bayes and SVM achieving significantly better results.

5) Other tasks with private databases: Abnormal ECG detection was studied by a number of papers. Ripoll et al. [98] used RBM-based pretrained models with ECGs from 1390 patients to assess whether a patient from ambulatory care or emergency should be referred to a cardiology service. They compared their model with KNN, SVM, extreme learning machines and an expert system achieving better results in accuracy and specificity. In [99] the authors train a model that classifies normal and abnormal subjects with 193690 ECG records of 10 to 20 seconds. Their model consisted of two parallel parts; the statistical learning and a rule inference. In statistical learning the ECGs are preprocessed using bandpass and lowpass filters, then fed to two parallel lead-CNNs and finally Bayesian fusion is employed to combine the probability outputs. In rule inference, the R-peak positions in the ECG record are detected and four disease rules are used for analysis. Finally, they utilize bias-average to determine the result.

Other tasks include premature ventricular contraction classification and stress detection. Liu et al. [100] used a single lead balanced dataset of 2400 normal and premature ventricular contraction ECGs from Children's Hospital of Shanghai for training. Two separate models were trained using the waveform images. The first one was a two *layer* CNN with dropout and the second an Inception v3 trained on Imagenet. Another three models were trained using the signals as 1D. The first model was a FNN with dropout, the second a three *layer* 1D CNN and the third a 2D CNN same as the first but trained with a stacked version of the signal (also trained with data augmentation). Experiments by the authors showed that the three *layer* 1D CNN created better and more stable results. In [101] the authors trained a network with an one convolutional layer with dropout followed by two RNNs to identify stress using short-term ECG data. They showed that their network achieved the best results compared with traditional machine learning methods and baseline DNNs.

6) Overall view on deep learning using ECG: Many deep learning methods have used ECG to train deep learning models utilizing the large number of databases that exist for that modality. It is evident from the literature that most deep learning methods (mostly CNNs and SDAEs) in this area consist of three parts: filtering for denoising, R-peak detection for beat segmentation and the neural network for feature extraction. Another popular set of methods is the conversion of ECGs to images, to utilize the wide range of architectures and pretrained models that have already been built for imaging modalities. This was done using spectrogram techniques [79], [88] and conversion to binary image [78], [92], [100].

B. Phonocardiogram with Physionet 2016 Challenge

Physionet/Computing in Cardiology (Cinc) Challenge 2016 (PHY16) was a competition for classification of normal/abnormal heart sound recordings. The training set consists

of five databases (A through E) that contain 3126 Phonocardiograms (PCGs), lasting from 5 seconds to 120 seconds.

Most of the methods convert PCGs to images using spectrogram techniques. Rubin et al. [108] used a logistic regression hidden semi-markov model for segmenting the start of each heartbeat which then were transformed into spectrograms using Mel-Frequency Cepstral Coefficients (MFCCs). Each spectrogram was classified into normal or abnormal using a two *layer* CNN which had a modified loss function that maximizes sensitivity and specificity, along with a regularization parameter. The final classification of the signal was the average probability of all segment probabilities. They achieved an overall score of 83.99% placing eighth at PHY16 challenge. Kucharski et al. [109] used an eight second spectrogram on the segments, before feeding them to a five *layer* CNN with dropout. Their method achieved 99.1% sensitivity and 91.6% specificity which are comparable to state-of-the-art methods on the task. Dominguez et al. [110] segmented the signals and preprocessed them using the neuromorphic auditory sensor [120] to decompose the audio information into frequency bands. Then, they calculated the spectrograms which were fed to a modified version of AlexNet. Their model achieved accuracy of 94.16% a significant improvement compared with the winning model of PHY16. In [111] the authors used Adaboost which was fed with spectrogram features from PCG and a CNN which was trained using cardiac cycles decomposed into four frequency bands. Finally, the outputs of the Adaboost and the CNN were combined to produce the final classification result using a simple decision rule. The overall accuracy was 89%, placing this method first in the official face of PHY16.

Models that did not convert the PCGs to spectrograms seemed to have lesser performance. Ryu et al. [112] applied Window-sinc Hamming filter for denoising, scaled the signal and used a constant window for segmentation. They trained a four *layer* 1D CNN using the segments and the final classification was the average of all segment probabilities. An overall accuracy of 79.5% was achieved in the official phase of PHY16.

Phonocardiograms have also been used for tasks such as S1 and S2 heart sound recognition by Chen et al. [113]. They converted heart sound signals into a sequence of MFCCs and then applied K-means to cluster the MFCC features into two groups to refine their representation and discriminative capability. The features are then fed to a DBN to perform S1 and S2 classification. The authors compared their method with KNN, Gaussian mixture models, logistic regression and SVM performing the best results.

According to the literature, CNNs are the majority of neural network architectures used for solving tasks with PCG. Moreover, just like in ECG, many deep learning methods converted the signals to images using spectrogram techniques [108]–[111], [115], [116].

C. Other signals

1) Oscillometric data: Oscillometric data are used for estimating SBP and DBP which are the haemodynamic pressures exerted within the arterial system during systole and diastole respectively [121].

DBNs have been used for SBP and DBP estimation. In their article Lee et al. [114] used bootstrap-aggregation to create ensemble parameters and then employed Adaboost to estimate SBP and DBP. Then, they used bootstrap and Monte-Carlo in order to determine the confidence intervals based on the target BP, which was estimated using the DBN ensemble regression estimator. This modification greatly improved the BP estimation over the baseline DBN model. Similar work has been done on this task by the same authors in [122]–[124].

Oscillometric data have also been used by Pan et al. [115] for assessing the variation of Korotkoff sounds. The beats were used to create windows centered on the oscillometric pulse peaks that were then extracted. A spectrogram was obtained from each beat, and all beats between the manually determined SBPs and DBPs were labeled as Korotkoff. A three *layer* CNN was then used to analyze consistency in sound patterns that were associated with Korotkoff sounds. According to the authors this was the first study performed for this task providing evidence that it is difficult to identify Korotkoff sounds at systole and diastole.

2) Data from wearable devices: Wearable devices, which impose restrictions on size, power and memory consumption for models, have also been used to collect cardiology data for training deep learning models for AF detection.

Shashikumar et al. [116] captured ECG, Pulsatile Photoplethysmographic (PPG) and accelerometry data from 98 subjects using a wrist-worn device and derived the spectrogram using continuous WT. They trained a five *layer* CNN in a sequence of short windows with movement artifacts and its output was combined with features calculated based on beat-to-beat variability and the signal quality index. An accuracy of 91.8% in AF detection was achieved by the method and in combination with its computational efficiency it is promising for real world deployment. Gotlibovych et al. [117] trained an one *layer* CNN network followed by a LSTM using 180h of PPG wearable data to detect AF. Use of the LSTM layer allows the network to learn variable-length correlations in contrast with the fixed length of the convolutional layer. Poh et al. [118] created a large database of PPG (over 180000 signals from 3373 patients) including data from MIMIC to classify four rhythms: sinus, noise, ectopic and AF. A densely connected CNN with six blocks and a growth rate of six was used for classification that was fed with 17 second segments that have been denoised using a bandpass filter. Results were obtained using an independent dataset of 3039 PPG achieving better results than previous methods that were based on handcrafted features.

Besides AF detection, wearable data have been used to search for optimal cardiovascular disease predictors. In [119] the authors trained a semi-supervised, multi-task bi-directional LSTM on data from 14011 users of the Cardiogram app for detecting diabetes, high cholesterol, high BP, and sleep apnoea. Their results indicate that the heart's response to physical activity is a salient biomarker for predicting the onset of a disease and can be captured using deep learning.

TABLE VI
DEEP LEARNING APPLICATIONS USING PCG AND OTHER SIGNALS

Reference	Method	Application/Notes[a]	Accuracy[b]
PCG/Physionet 2016 Challenge			
Rubin 2017 [108]	CNN	logistic regression hidden-semi markov, MFCCs and a two *layer* CNN	83.99%
Kucharski 2017 [109]	CNN	spectrogram and five *layer* CNN with dropout	91.6%[b]
Dominguez 2018 [110]	CNN	spectrogram and modified AlexNet	94.16%
Potes 2016 [111]	CNN	ensemble of Adaboost and CNN and outputs combined with decision rule	86.02%
Ryu 2016 [112]	CNN	denoising filters and four *layer* CNN	79.5%
Chen 2017 [113]	DBN	recognize S1 and S2 heart sounds using MFCCs, K-means and DBN (private)	91%
Other signals			
Lee 2017 [114]	DBN	estimate BP using bootstrap-aggregation, Monte-Carlo and DBN with oscillometry data	*multiple*
Pan 2017 [115]	CNN	assess Korotkoff sounds using a three *layer* CNN with oscillometry data	*multiple*[b]
Shashikumar 2017 [116]	CNN	detect AF using ECG, photoplethysmography, accelerometry with WT and a CNN	91.8%
Gotlibovych 2018 [117]	CNN, LSTM	detect AF using PPG from wearable and a LSTM-based CNN	>99%[b]
Poh 2018 [118]	CNN	detect four rhythms on PPG using a densely CNN (MIMIC, VORTAL, PPGDB)	87.5%[b]
Ballinger 2018 [119]	LSTM	predict diabetes, high cholesterol, high BP, and sleep apnoea using sensor data and LSTM	0.845[b]

[a]In parenthesis the databases used by paper or by papers in subsection. In the 'PCG/Physionet 2016 Challenge' subtable all papers use PHY besides [113], and in the 'Other signals' subtable all papers use private databases besides [118].

[b]There is a wide variability in results reporting. [109] report specificity, [115] report results for SBP and DBP, [117] report sensitivity, specificity, [118] report positive predictive value, [119] report AUC for diabetest, results are also reported for high cholesterol sleep apnea and high BP.

V. DEEP LEARNING USING IMAGING MODALITIES

Imaging modalities that have found use in cardiology include Magnetic Resonance Imaging (MRI), Fundus Photography, Computerized Tomography (CT), Echocardiography, Optical Coherence Tomography (OCT), Intravascular Ultrasound (IVUS), and others. Deep learning has been mostly successful in this area, mainly due to architectures that make use of convolutional layers and network depth. A summary of deep learning applications using images are shown in Tables VII, VIII, and IX.

A. Magnetic resonance imaging

MRI is based on the interaction between a system of atomic nuclei and an external magnetic field providing a picture of the interior of a physical object [125]. The main uses of MRI include Left Ventricle (LV), Right Ventricle (RV) and whole heart segmentation.

1) Left ventricle segmentation: CNNs were used for LV segmentation with MRI. Tan et al. [126] used a CNN to localize LV endocardium and another CNN to determine the endocardial radius using STA11 and SUN09 for training and evaluation respectively. Without filtering out apical slices and using deformable models they achieve comparable performance with previous state-of-the-art methods. In [127] the authors trained a five *layer* CNN using MRI from SUN09 challenge. They trained their model using SGD and RMSprop with the former achieving a better Dice of 92%.

CNNs combined with RNNs were also used. In [128] the authors created a recurrent u-net that learns image representations from a stack of 2D slices and has the ability to leverage inter-slice spatial dependencies through internal memory units. It combines anatomical detection and segmentation into a single end-to-end architecture, achieving comparable results with other non end-to-end methods, outperforming the baselines DBN, recurrent DBN and FCN in terms of Dice.

Other papers combined deep learning methods with level set for LV segmentation. Rupprecht et al. [129] trained a class-specific four *layer* CNN which predicts a vector pointing from the respective point on the evolving contour towards the closest point on the boundary of the object of interest. These predictions formed a vector field which was then used for evolving the contour using the Sobolev active contour framework. Anh et al. [130] created a non-rigid segmentation method based on the distance regularized level set method that was initialized and constrained by the results of a structured inference using a DBN. Avendi et al. [131] used CNN to detect the LV chamber and then utilized stacked AEs to infer the shape of the LV. The result was then incorporated into deformable models to improve the accuracy and robustness of the segmentation.

Atlas-based methods have also been used for this task. Yang et al. [132] created an end-to-end deep fusion network by concatenating a feature extraction network and a non-local patch-based label fusion network. The learned features are further utilized in defining a similarity measure for MRI atlas selection. They compared their method with majority voting, patch-based label fusion, multi-atlas patch match and SVM with augmented features achieving superior results in terms of accuracy. Luo et al. [133] adopted a LV atlas mapping method to achieve accurate localization using MRI data from DS16. Then, a three *layer* CNN was trained for predicting the LV volume, achieving comparable results with the winners of the challenge in terms of root mean square of end-diastole and end-systole volumes.

Regression-based methods have been used for localizing the LV before segmenting it. Yang et al. [134] first locate LV in the full image using a regression CNN and then segment it within

TABLE VII
DEEP LEARNING APPLICATIONS USING MRI

Reference	Method	Application/Notes[a]	Dice[b]
LV segmentation			
Tan 2016 [126]	CNN	CNN for localization and CNN for delineation of endocardial border (SUN09, STA11)	88%
Romaguera 2017 [127]	CNN	five *layer* CNN with SGD and RMSprop (SUN09)	92%
Poudel 2016 [128]	u-net, RNN	combine u-net and RNN (SUN09, private)	*multiple*
Rupprecht 2016 [129]	CNN	combine four *layer* CNN with Sobolev (STA11, non-medical)	85%
Ngo 2014 [130]	DBN	combine DBN with level set (SUN09)	88%
Avendi 2016 [131]	CNN, AE	CNN for chamber detection, AEs for shape inference and deformable models (SUN09)	96.69%
Yang 2016 [132]	CNN	feature extraction network and a non-local patch-based label fusion network (SAT13)	81.6%
Luo 2016 [133]	CNN	a LV atlas mapping method and a three *layer* CNN (DS16)	4.98%*
Yang 2017 [134]	CNN, u-net	localization with regression CNN and segmentation with u-net (YUDB, SUN09)	91%, 93%§
Tan 2017 [135]	CNN	regression CNN (STA11, DS16)	*multiple*
Curiale 2017 [136]	u-net	use residual u-net (SUN09)	90%
Liao 2017 [137]	CNN	local binary pattern for localization and hypercolumns FCN for segmentation (DS16)	4.69%*
Emad 2015 [138]	CNN	LV localization using CNN and pyramid of scales (YUDB)	98.66%$^{-}$
LV/RV segmentation			
Zotti 2017 [139]	u-net	u-net variant with a multi-resolution conv-deconv grid architecture (AC17)	90%
Patravali 2017 [140]	u-net	2D/3D u-net trained (AC17)	*multiple*
Isensee 2017 [141]	u-net	ensemble of u-net, regularized multi-layer perceptrons and a RF classifier (AC17)	*multiple*
Tran 2016 [142]	CNN	four *layer* FCN (SUN09, STA11)	92%, 96%$^{+}$
Bai 2017 [143]	CNN	VGG-16 and DeepLab architecture with use of CRF for refined results (UKBDB)	90.3%
Lieman 2017 [144]	u-net	extension of ENet [145] with skip connections (private)	*multiple*
Winther 2017 [146]	u-net	ν-net is a u-net variant (DS16, SUN09, RV12, private)	*multiple*
Du 2018 [147]	DBN	DAISY features and regression DBN using 2900 images (private)	91.6%, 94.1%$^{+}$
Giannakidis 2016 [148]	CNN	RV segmentation using 3D multi-scale CNN with two pathways (private)	82.81%
Whole heart segmentation			
Wolterink 2016 [149]	CNN	dilated CNN with orthogonal patches (HVS16)	80%, 93%
Li 2016 [150]	CNN	deeply supervised 3D FCN with dilations (HVS16)	69.5%
Yu 2017 [151]	CNN	deeply supervised 3D FCN constructed in a self-similar fractal scheme (HVS16)	*multiple*
Payer [152]	CNN	FCN for localization and another FCN for segmentation (MM17)	90.7%, 87%$^{\#}$
Mortazi 2017 [153]	CNN	multi-planar FCN (MM17)	90%, 85%$^{\#}$
Yang [154]	CNN	deeply supervised 3D FCN trained with transfer learning (MM17)	84.3%, 77.8%$^{\#}$
Others applications			
Yang 2017 [155]	SSAE	atrial fibrosis segmentation using multi-atlas propagation, SSAE and softmax (private)	82%
Zhang 2016 [156]	CNN	missing apical and basal identification using two CNNs with four *layers* (UKBDB)	*multiple*
Kong 2016 [157]	CNN, RNN	CNN for spatial information and RNN for temporal information to identify frames (private)	*multiple*
Yang 2017 [158]	CNN	CNN to identify end-diastole and end-systole frames from LV (STA11, private)	76.5%$^{-}$
Xu 2017 [159]	Multiple	MI detection using Fast R-CNN for heart localization, LSTM and SAE (private)	94.3%$^{-}$
Xue 2018 [160]	CNN, LSTM	CNN, two parallel LSTMs and a Bayesian framework for full LV quantification (private)	*multiple*
Zhen 2016 [161]	RBM	multi-scale convolutional RBM and RF for bi-ventricular volume estimation (private)	3.87%*
Biffi 2016 [162]	CNN	identify hypertrophic cardiomyopathy using a variational AE (AC17, private)	90%$^{-}$
Oktay 2016 [163]	CNN	image super resolution using residual CNN (private)	*multiple*

^{a}In parenthesis the databases used.

^{b}($\S$) denotes for 'for each database', ($*$) denotes mean square error for EF ($+$) denotes 'for endocardial and epicardial', ($-$) denotes accuracy, ($\#$) denotes 'for CT and MRI'

the cropped region of interest using a u-net based architecture. They demonstrate that their model achieves high accuracy with computational performance during inference.

Various other methods were used. Tan et al. [135] parameterize all short axis slices and phases of the LV segmentation task in terms of the radial distances between the LV center-point and the endocardial and epicardial contours in polar space. Then, they train a CNN regression on STA11 to infer these parameters and test the generalizability of the method on DS16 with good results. In [136] the authors used Jaccard distance as optimization objective function, integrating a residual learning strategy, and introducing a batch normalization layer

to train a u-net. It is shown in the paper that this configuration performed better than other simpler u-nets in terms of Dice. In their article Liao et al. [137] detected the Region of Interest (ROI) containing LV chambers and then used hypercolumns FCN to segment LV in the ROI. The 2-D segmentation results were integrated across different images to estimate the volume. The model was trained alternately on LV segmentation and volume estimation, placing fourth in the test set of DS16. Emad et al. [138] localize the LV using a CNN and a pyramid of scales analysis to take into account different sizes of the heart with the YUDB. They achieve good results but with a significant computation cost (10 seconds per image during

inference).

2) LV/RV segmentation: A dataset used for LV/RV segmentation was the MICCAI 2017 ACDC Challenge (AC17) that contains MRI images from 150 patients divided into five groups (normal, previous MI, dilated cardiomyopathy, hypertrophic cardiomyopathy, abnormal RV). Zotti et al. [139] used a model that includes a cardiac center-of-mass regression module which allows shape prior registration and a loss function tailored to the cardiac anatomy. Features are learned with a multi-resolution conv-deconv 'grid' architecture which is an extension of u-net. This model compared with vanilla conv-deconv and u-net performs better by an average of 5% in terms of Dice. Patravali et al. [140] trained a model based on u-net using Dice combined with cross entropy as a metric for LV/RV and myocardium segmentation. The model was designed to accept a stack of image slices as input channels and the output is predicted for the middle slice. Based on experiments they conducted, it was concluded that three input slices were optimal as an input for the model, instead of one or five. Isensee et al. [141] used an ensemble of a 2D and a 3D u-net for segmentation of the LV/RV cavity and the LV myocardium on each time instance of the cardiac cycle. Information was extracted from the segmented time-series in form of features that reflect diagnostic clinical procedures for the purposes of the classification task. Based on these features they then train an ensemble of regularized multi-layer perceptrons and a RF classifier to predict the pathological target class. Their model ranked first in the ACDC challenge.

Various other datasets have also been used for this task with CNNs. In [143] the authors created a semi-supervised learning method, in which a segmentation network for LV/RV and myocardium was trained from labeled and unlabeled data. The network architecture was adapted from VGG-16, similar to the DeepLab architecture [164] while the final segmentation was refined using a Conditional Random Field (CRF). The authors show that the introduction of unlabelled data improves segmentation performance when the training set is small. In [148] the authors adopt a 3D multi-scale CNN to identify pixels that belong to the RV. The network has two convolutional pathways and their inputs are centered at the same image location, but the second segment is extracted from a down-sampled version of the image. The results obtained were better than the previous state-of-the-art although the latter were based on feature engineering and trained on less variable datasets.

FCNs have also been used for LV/RV segmentation. In their article Tran et al. [142] trained a four *layer* FCN model for LV/RV segmentation on SUN09, STA11. They compared previous state-of-the-art methods along with two initializations of their model: a fine-tuned version of their model using STA11 and a Xavier initialized model with the former performing best in almost all tasks.

FCNs with skip connections and u-net have also been used for this task. Lieman et al. [144] created a FCN architecture with skip connections named FastVentricle based on ENet [145] which is faster and runs with less memory than previous ventricular segmentation architectures achieving high clinical accuracy. In [146] the authors introduce ν-net which is a u-net variant for segmentation of LV/RV endocardium and epicardium using DS16, SUN09 and RV12 datasets. This method performed better than the expert cardiologist in this study, especially for RV segmentation.

Some methods were based on regression models. In their article Du et al. [147] created a regression segmentation framework to delineate boundaries of LV/RV. First, DAISY features are extracted and then a point-based representation method was employed to depict the boundaries. Finally, the DAISY features were used as input and the boundary points as labels to train the regression model based on DBN. The model performance is evaluated using different features than DAISY (GIST, pyramid histogram of oriented gradients) and also compared with support vector regression and other traditional methods (graph cuts, active contours, level set) achieving better results.

3) Whole heart segmentation: MICCAI 2016 HVSMR (HVS16) was used for whole heart segmentation that contains MRI images from 20 patients. Wolterink et al. [149] trained a ten *layer* CNN with increasing levels of dilation for segmenting the myocardium and blood pool in axial, sagittal and coronal image slices. They also employ deep supervision [165] to alleviate the vanishing gradients problem and improve the training efficiency of their network using a small dataset. Experiments performed with and without dilations on this architecture indicated the usefulness of this configuration. In their article Li et al. [150] start with a 3D FCN for voxel-wise labeling and then introduce dilated convolutional layers into the baseline model to expand its receptive field. Then, they employ deep-supervised pathways to accelerate training and exploit multi-scale information. According to the authors the model demonstrates good segmentation accuracy combined with low computational cost. Yu et al. [151] created a 3D FCN fractal network for whole heart and great vessel volume-to-volume segmentation. By recursively applying a single expansion rule, they construct the network in a self-similar fractal scheme combining hierarchical clues for accurate segmentation. They also achieve good results with low computational cost (12 seconds per volume).

Another database used for whole heart segmentation was MM17 which contains 120 multimodal images from cardiac MRI/CT. Payer et al. [152] method is based on two FCN for multi-label whole heart localization and segmentation. At first, the localization CNN finds the center of the bounding box around all heart structures, such that the segmentation CNN can focus on this region. Trained in an end-to-end manner, the segmentation CNN transforms intermediate label predictions to positions of other labels. Therefore, the network learns from the relative positions among labels and focuses on anatomically feasible configurations. The model was compared with u-net achieving superior results, especially in the MRI dataset. Mortazi et al. [153] trained a multi-planar CNN with an adaptive fusion strategy for segmenting seven substructures of the heart. They designed three CNNs, one for each plane, with the same architectural configuration and trained them for voxel-wise labeling. Their experiments conclude that their model delineates cardiac structures with high accuracy and efficiency. In [154] the authors used FCN, and couple it

with 3D operators, transfer learning and a deep supervision mechanism to distill 3D contextual information and solve potential difficulties in training. A hybrid loss was used that guides the training procedure to balance classes, and preserve boundary details. According to their experiments, using the hybrid loss achieves better results than using only Dice.

4) Other tasks: Deep learning papers has also been used for detection of other cardiac structures with MRI. Yang et al. [155] created a multi-atlas propagation method to derive the anatomical structure of the left atrium myocardium and pulmonary veins. This was followed by a unsupervised trained SSAE with a softmax for atrial fibrosis segmentation using 20 scans from AF patients. In their article Zhang et al. [156] try to detect missing apical and basal slices. They test the presence of typical basal and apical patterns at the bottom and top slices of the dataset and train two CNNs to construct a set of discriminative features. Their experiments showed that the model with four *layers* performed better than the baselines SAE and Deep Boltzmann Machines.

Other medical tasks in MRI were also studied such as detection of end-diastole end-systole frames. Kong et al. [157] created a temporal regression network pretrained on ImageNet by integrating a CNN with a RNN, to identify end-diastole and end-systole frames from MRI sequences. The CNN encodes the spatial information of a cardiac sequence, and the RNN decodes the temporal information. They also designed a loss function to constrain the structure of predicted labels. The model achieves better average frame difference than the previous methods. In their article Yang et al. [158] used a CNN to identify end-diastole and end-systole frames from the LV, achieving an overall accuracy of 76.5%.

There were also methods that tried to quantify various cardiovascular features. In [159] the authors detect the area, position and shape of the MI using a model that consists of three *layers*; first, the heart localization layer is a Fast R-CNN [166] which crops the ROI sequences including the LV; second, the motion statistical layers, which build a time-series architecture to capture the local motion features generated by LSTM-RNN and the global motion features generated by deep optical flows from the ROI sequence; third, the fully connected discriminate layers, which use SAE to further learn the features from the previous layer and a softmax classifier. Xue et al. [160] trained an end-to-end deep multitask relationship learning framework on MRI images from 145 subjects with 20 frames each for full LV quantification. It consists of a three *layer* CNN that extracts cardiac representations, then two parallel LSTM-based RNNs for modeling the temporal dynamics of cardiac sequences. Finally, there is a Bayesian framework capable of learning multitask relationships and a softmax classifier for classification. Extensive comparisons with the state-of-the-art show the effectiveness of this method in terms of mean absolute error. In [161] the authors created an unsupervised cardiac image representation learning method using multi-scale convolutional RBM and a direct bi-ventricular volume estimation using RF. They compared their model with a Bayesian model, a feature based model, level sets and graph cut achieving better results in terms of correlation coefficient for LV/RV volumes and estimation error of EF.

Other methods were also created to detect hypertrophic cardiomyopathy or increase the resolution of MRI. Biffi et al. [162] trained a variational AE to identify hypertrophic cardiomyopathy subjects tested on a multi-center balanced dataset of 1365 patients and AC17. They also demonstrate that the network is able to visualize and quantify the learned pathology-specific remodeling patterns in the original input space of the images, thus increasing the interpretability of the model. In [163] the authors created an image super-resolution method based on a residual CNN that allows the use of input data acquired from different viewing planes for improved performance. They compared it with other interpolation methods (linear, spline, multi-atlas patch match, shallow CNN, CNN) achieving better results in terms of PSNR. Authors from the same group proposed a training strategy [167] that incorporates anatomical prior knowledge into CNNs through a regularization model, by encouraging it to follow the anatomy via learned non-linear representations of the shape.

5) Overall view on deep learning using MRI: There is a wide range of architectures that have been applied in MRI. Most predominantly CNNs and u-nets are used solely or in combination with RNNs, AEs, or ensembles. The problem is that most of them are not end-to-end; they rely on preprocessing, handcrafted features, active contours, level set and other non-differentiable methods, thus partially losing the ability to scale on the presence of new data. The main target of this area should be to create end-to-end models even if that means less accuracy in the short-term; more efficient architectures could close the gap in the future.

An interesting finding regarding whole heart segmentation was done in [168] where the authors investigated the suitability of state-of-the-art 2D, 3D CNN architectures, and modifications of them. They find that processing the images in a slice-by-slice fashion using 2D networks was beneficial due to the large slice thickness. However, the choice of the network architecture plays a minor role.

B. Fundus photography

Fundus photography is a clinical tool for evaluating retinopathy progress in patients where the image intensity represents the amount of reflected light of a specific waveband [169]. One of the most widely used databases in fundus is DRIVE which contains 40 images and their corresponding vessel mask annotations.

1) Vessel segmentation: CNNs have been used for vessel segmentation in fundus imaging. In [170] the authors first used histogram equalization and Gaussian filtering to reduce noise. A three *layer* CNN was then used as a feature extractor and a RF as the classifier. According to experiments done by the authors the best performance was achieved by a winner-takes-all ensemble, compared with an average, weighted and median ensemble. Zhou et al. [171] applied image preprocessing to eliminate the strong edges around the field of view and normalize the luminosity and contrast inside it. Then, they trained a CNN to generate features for linear models and applied filters to enhance thin vessels, reducing the intensity difference between thin and wide vessels. A dense CRF was then adapted

TABLE VIII
DEEP LEARNING APPLICATIONS USING FUNDUS PHOTOGRAPHY

Reference	Method	Application/Notes[a]	AUC[b]
Vessel segmentation			
Wang 2015 [170]	CNN, RF	three *layer* CNN combined with ensemble RF (DRIVE, STARE)	0.9475
Zhou 2017 [171]	CNN, CRF	CNN to extract features and CRF for final result (DRIVE, STARE, CHDB)	0.7942
Chen 2017 [172]	CNN	artificial data, FCN (DRIVE, STARE)	0.9516
Maji 2016 [173]	CNN	12 CNNs ensemble with three *layers* (DRIVE)	0.9283
Fu 2016 [174]	CNN, CRF	CNN and CRF (DRIVE, STARE)	94.70%*
Wu 2016 [175]	CNN	vessel segmentation and branch detection using CNN and PCA (DRIVE)	0.9701
Li 2016 [176]	SDAE	FNN and SDAE (DRIVE, STARE, CHDB)	0.9738
Lahiri 2016 [177]	SDAE	ensemble of two level of sparsely trained SDAE (DRIVE)	95.30%*
Oliveira 2017 [178]	u-net	data augmentation and u-net (DRIVE)	0.9768
Leopold 2017 [179]	CNN	CNN as a multi-channel classifier and Gabor filters (DRIVE)	94.78%*
Leopold 2017 [180]	AE	fully residual AE with gated streams based on u-net (DRIVE, STARE, CHDB)	0.8268
Mo 2017 [181]	CNN	auxiliary classifiers and transfer learning (DRIVE, STARE, CHDB)	0.9782
Melinscak 2015 [182]	CNN	four *layer* CNN (DRIVE)	0.9749
Sengur 2017 [183]	CNN	two *layer* CNN with dropout (DRIVE)	0.9674
Meyer 2017 [184]	u-net	vessel segmentation using u-net on SLO (IOSTAR, RC-SLO)	0.9771
Microaneurysm and hemorrhage detection			
Haloi 2015 [185]	CNN	MA detection using CNN with dropout and maxout activation (ROC, Messidor, DIA)	0.98
Giancardo 2017 [186]	u-net	MA detection using internal representation of trained u-net (DRIVE, Messidor)	*multiple*
Orlando 2018 [187]	CNN	MA and hemorrhage detection using handcrafted features and a CNN (DIA, e-optha, Messidor)	*multiple*
van Grinsven 2017 [188]	CNN	hemorrhage detection with selective data sampling using a five *layer* CNN (KR15, Messidor)	*multiple*
Other applications			
Girard 2017 [189]	CNN	artery/vein classification using CNN and likelihood score propagation (DRIVE, Messidor)	*multiple*
Welikala 2017 [190]	CNN	artery/vein classification using three *layer* CNN (UKBDB)	82.26%*
Pratt 2017 [191]	ResNet	bifurcation/crossing classification using ResNet 18 (DRIVE, IOSTAR)	*multiple*
Poplin 2017 [192]	Inception	cardiovascular risk factors prediction (UKBDB, private)	*multiple*

[a]In parenthesis the databases used.
[b](*) denotes accuracy.

to achieve the final retinal vessel segmentation, by taking the discriminative features for unary potentials and the thin-vessel enhanced image for pairwise potentials. Amongst their results, in which they demonstrate better accuracy than most state-of-the-art methods, they also provide evidence in favor of using the RGB information of the fundus instead of just the green channel. Chen [172] designed a set of rules to generate artificial training samples with prior knowledge and without manual labeling. They train a FCN with a concatenation layer that allows high level perception guide the work in lower levels and evaluate their model on DRIVE and STARE databases, achieving comparable results with other methods that use real labeling. In [173] the authors trained a 12 CNNs ensemble with three *layers* each on the DRIVE database, where during inference the responses of the CNNs are averaged to form the final segmentation. They demonstrate that their ensemble achieves higher maximum average accuracy than previous methods. Fu et al. [174] train a CNN on DRIVE and STARE databases to generate the vessel probability maps and then they employed a fully connected CRF to combine the discriminative vessel probability maps and long-range interactions between pixels. In [175] the authors used a CNN to learn the features and a PCA-based nearest neighbor search utilized to estimate the local structure distribution. Besides demonstrating good results they argue that it is important for CNN to incorporate

information regarding the tree structure in terms of accuracy.

AEs were used for vessel segmentation. Li et al. [176] trained a FNN and a denoising AE with DRIVE, STARE and CHDB databases. They argue that the learnt features of their model are more reliable to pathology, noise and different imaging conditions, because the learning process exploits the characteristics of vessels in all training images. In [177] the authors employed unsupervised hierarchical feature learning using a two level ensemble of sparsely trained SDAE. The training level ensures decoupling and the ensemble level ensures architectural revision. They show that ensemble training of AEs fosters diversity in learning dictionary of visual kernels for vessel segmentation. Softmax classifier was then used for fine-tuning each AE and strategies are explored for two level fusion of ensemble members.

Other architectures were also used for vessel segmentation. In their article Oliveira et al. [178] trained a u-net with DRIVE demonstrating good results and presenting evidence of the benefits of data augmentation on the training data using elastic transformations. Leopold et al. [179] investigated the use of a CNN as a multi-channel classifier and explore the use of Gabor filters to boost the accuracy of the method described in [193]. They applied the mean of a series of Gabor filters with varying frequencies and sigma values to the output of the network to determine whether a pixel represents a vessel or not. Besides

finding that the optimal filters vary between channels, the authors also state the 'need' of enforcing the networks to align with human perception, in the context of manual labeling, even if that requires downsampling information, which would otherwise reduce the computational cost. The same authors [180] created PixelBNN which is a fully residual AE with gated streams. It is more than eight times faster than the previous state-of-the-art methods at test time and performed well, considering a significant reduction in information from resizing images during preprocessing. In their article Mo et al. [181] used deep supervision with auxiliary classifiers in intermediate layers of the network, to improve the discriminative capability of features in lower layers of the deep network and guide backpropagation to overcome vanishing gradient. Moreover, transfer learning was used to overcome the issue of insufficient medical training data.

2) Microaneurysm and hemorrhage detection: Haloi [185] trained a three *layer* CNN with dropout and maxout activation function for MA detection. Experiments on ROC and DIA demonstrated state-of-the-art results. In [186] the authors created a model that learns a general descriptor of the vasculature morphology using the internal representation of a u-net variation. Then, they tested the vasculature embeddings on a similar image retrieval task according to vasculature and on a diabetic retinopathy classification task, where they show how the vasculature embeddings improve the classification of a method based on MA detection. In [187] the authors combined augmented features learned by a CNN with handcrafted features. This ensemble vector of descriptors was then used to identify MA and hemorrhage candidates using a RF classifier. Their analysis using t-SNE demonstrates that CNN features have fine-grained characteristics such as the orientation of the lesion while handcrafted features are able to discriminate low contrast lesions such as hemorrhages. In [188] the authors trained a five *layer* CNN to detect hemorrhage using 6679 images from DS16 and Messidor databases. They applied selective data sampling on a CNN which increased the speed of the training by dynamically selecting misclassified negative samples during training. Weights are assigned to the training samples and informative samples are included in the next training iteration.

3) Other tasks: Fundus has also been used for artery/vein classification. In their article Girard et al. [189] trained a four *layer* CNN that classifies vessel pixels into arteries/veins using rotational data augmentation. A graph was then constructed from the retinal vascular network where the nodes are defined as the vessel branches and each edge gets associated to a cost that evaluates whether the two branches should have the same label. The CNN classification was propagated through the minimum spanning tree of the graph. Experiments demonstrated the effectiveness of the method especially in the presence of occlusions. Welikala et al. [190] trained and evaluated a three *layer* CNN using centerline pixels derived from retinal images. Amongst their experiments they found that rotational and scaling data augmentations did not help increase accuracy, attributing it to interpolation altering pixel intensities which is problematic due to the sensitivity of CNN to pixel distribution patterns.

There are also other uses such as bifurcation/crossing identification. Pratt et al. [191] trained a ResNet18 to identify small patches which include either bifurcation or crossing. Another ResNet18 was trained on patches that have been classified to have bifurcations and crossings to distinguish the type of vessel junction located. Similar work on this problem has been done by the same authors [194] using a CNN.

An important result in the area of Cardiology using fundus photography is from Poplin et al. [192] who used an Inception v3 to predict cardiovascular risk factors (age, gender, smoking status, HbA1c, SBP) and major cardiac events. Their models used distinct aspects of the anatomy to generate each prediction, such as the optic disc or the blood vessels, as it was demonstrated using the soft attention technique. Most results were significantly better than previously thought possible with fundus photography (>70% AUC).

4) Overall view on deep learning using fundus: Regarding the use of architectures there is a clear preference for CNNs especially in vessel segmentation while an interesting approach from some publications is the use of CRFs as a postprocessing step for vessel segmentation refinement [171], [174]. The fact that there are many publicly available databases and that the DRIVE database is predominantly used in most of the literature makes this field easier to compare and validate new architectures. Moreover the non-invasive nature of fundus and its recent use as a tool to estimate cardiovascular risk predictors makes it a promising modality of increased usefulness in the field of cardiology.

C. Computerized tomography

Computerized Tomography (CT) is a non-invasive method for the detection of obstructive artery disease. Some of the areas that deep learning was applied with CT include coronary artery calcium score assessment, localization and segmentation of cardiac areas.

Deep learning was used for coronary calcium detection with CT. Lessman et al. [195] method for coronary calcium scoring utilizes three independently trained CNNs to estimate a bounding box around the heart, in which connected components above a Hounsfield unit threshold are considered candidates for CACs. Classification of extracted voxels was performed by feeding two-dimensional patches from three orthogonal planes into three concurrent CNNs to separate them from other high intensity lesions. Patients were assigned to one of five standard cardiovascular risk categories based on the Agatston score. Authors from the same group created a method [232] for the detection of calcifications in low-dose chest CT using a CNN for anatomical location and another CNN for calcification detection. In [196] the authors compared a four block u-net and a five block DenseNet for calculating the Agatston score using over 1000 images from a chest CT database. The authors heavily preprocessed the images using thresholding, connected component analysis and morphological operations for lungs, trachea and carina detection. Their experiments showed that DenseNet performed better in terms of accuracy. Cano et al. [197] trained a three *layer* 3D regression CNN that computes the Agatston score

TABLE IX
DEEP LEARNING APPLICATIONS USING CT, ECHOCARDIOGRAPHY, OCT AND OTHER IMAGING MODALITIES

Reference	Method	Application/Notes[a]
CT		
Lessman 2016 [195]	CNN	detect coronary calcium using three independently trained CNNs
Shadmi 2018 [196]	DenseNet	compared DenseNet and u-net for detecting coronary calcium
Cano 2018 [197]	CNN	3D regression CNN for calculation of the Agatston score
Wolterink 2016 [198]	CNN	detect coronary calcium using three CNNs for localization and two CNNs for detection
Santini 2017 [199]	CNN	coronary calcium detection using a seven *layer* CNN on image patches
Lopez 2017 [200]	CNN	thrombus volume characterization using a 2D CNN and postprocessing
Hong 2016 [201]	DBN	detection, segmentation, classification of abdominal aortic aneurysm using DBN and image patches
Liu 2017 [202]	CNN	left atrium segmentation using a twelve *layer* CNN and active shape model (STA13)
de Vos 2016 [203]	CNN	3D localization of anatomical structures using three CNNs, one for each orthogonal plane
Moradi 2016 [204]	CNN	detection of position for a given CT slice using a pretrained VGGnet, handcrafted features and SVM
Zheng 2015 [205]	Multiple	carotid artery bifurcation detection using multi-layer perceptrons and probabilistic boosting-tree
Montoya 2018 [206]	ResNet	3D reconstruction of cerebral angiogram using a 30 layer ResNet
Zreik 2018 [207]	CNN, AE	identify coronary artery stenosis using CNN for LV segmentation and an AE, SVM for classification
Commandeur 2018 [208]	CNN	quantification of epicardial and thoracic adipose tissue from non-contrast CT
Gulsun 2016 [209]	CNN	extract coronary centerline using optimal path from computed flow field and a CNN for refinement
Echocardiography		
Carneiro 2012 [210]	DBN	LV segmentation by decoupling rigin and non-rigid detections using DBN on 480 images
Nascimento 2016 [211]	DBN	LV segmentation using manifold learning and a DBN
Chen 2016 [212]	CNN	LV segmentation using multi-domain regularized FCN and transfer learning
Madani 2018 [213]	CNN	transthoracic echocardiogram view classification using a six *layer* CNN
Silva 2018 [214]	CNN	ejection fraction classification using a residual 3D CNN and transthoracic echocardiogram images
Gao 2017 [215]	CNN	viewpoint classification by fusing two CNNs with seven *layers* each
Abdi 2017 [216]	CNN, LSTM	assess quality score using convolutional and recurrent layers
Ghesu 2016 [217]	CNN	aortic valve segmentation using 2891 3D transesophageal echocardiogram images
Perrin 2017 [218]	CNN	congenital heart disease classification using a CNN trained in pairwise fashion
Moradi 2016 [219]	VGGnet, doc2vec	produce semantic descriptors for images
OCT		
Roy 2016 [220]	AE	tissue characterization using a distribution preserving AE
Yong 2017 [221]	CNN	lumen segmentation using a linear-regression CNN with four *layers*
Xu 2017 [222]	CNN	presence of fibroatheroma using features extracted from previous architectures and SVM
Abdolmanafi 2017 [223]	CNN	intima, media segmentation using a pretrained AlexNet and comparing various classifiers
Other imaging modalities		
Lekadir 2017 [224]	CNN	carotid plaque characterization using four *layer* CNN on Ultrasound
Tajbakhsh 2017 [225]	CNN	carotid intima media thickness video interpretation using two CNNs with two *layers* on Ultrasound
Tom 2017 [226]	GAN	IVUS image generation using two GANs (IV11)
Wang 2017 [227]	CNN	breast arterial calcification using a ten *layer* CNN on mammograms
Liu 2017 [228]	CNN	CAC detection using CNNs on 1768 X-Rays
Pavoni 2017 [229]	CNN	denoising of percutaneous transluminal coronary angioplasty images using four *layer* CNN
Nirschl 2018 [230]	CNN	trained a patch-based six *layer* CNN for identifying heart failure in endomyocardial biopsy images
Betancur 2018 [231]	CNN	trained a three *layer* CNN for obstructive CAD prediction from myocardial perfusion imaging

[a]Results from these imaging modalities are not reported in this review because they were highly variable in terms of the research question they were trying to solve and highly inconsistent in respect with the use of metrics. Additionally all papers use private databases besides [202], [226].

using 5973 non-ECG gated CT images achieving a Pearson correlation of 0.932. In [198] the authors created a method to identify and quantify CAC without a need for coronary artery extraction. The bounding box detection around the heart method employs three CNNs, where each detects the heart in the axial, sagittal and coronal plane. Another pair of CNNs were used to detect CAC. The first CNN identifies CAC-like voxels, thereby discarding the majority of non-CAC-like voxels such as lung and fatty tissue. The identified CAC-like voxels are further classified by the second CNN in the pair, which distinguishes between CAC and CAC-like negatives. Although the CNNs share architecture, given that they have different tasks they do not share weights. They achieve a

Pearson correlation of 0.95, comparable with previous state-of-the-art. Santini et al. [199] trained a seven *layer* CNN using patches for the segmentation and classification of coronary lesions in CT images. They trained, validated and tested their network on 45, 18 and 56 CT volumes respectively achieving a Pearson correlation of 0.983.

CT has been used for segmentation of various cardiac areas. Lopez et al. [200] trained a 2D CNN for aortic thrombus volume assessment from pre-operatively and post-operatively segmentations using rotating and mirroring augmentation. Postprocessing includes Gaussian filtering and k-means clustering. In their article Hong et al. [201] trained a DBN using image patches for the detection, segmentation and severity

classification of Abdominal Aortic Aneurysm region in CT images. Liu et al. [202] used an FCN with twelve *layers* for left atrium segmentation in 3D CT volumes and then refined the segmentation results of the FCN with an active shape model achieving a Dice of 93%.

CT has also been used for localization of cardiac areas. In [203] the authors created a method to detect anatomical ROIs (heart, aortic arch, and descending aorta) in 2D image slices from chest CT in order to localize them in 3D. Every ROI was identified using a combination of three CNNs, each analyzing one orthogonal image plane. While a single CNN predicted the presence of a specific ROI in the given plane, the combination of their results provided a 3D bounding box around it. In their article Moradi et al. [204] address the problem of detection of vertical position for a given cardiac CT slice. They divide the body area depicted in chest CT into nine semantic categories each representing an area most relevant to the study of a disease. Using a set of handcrafted image features together with features derived from a pretrained VGGnet with five *layers*, they build a classification scheme to map a given CT slice to the relevant level. Each feature group was used to train a separate SVM classifier and predicted labels are then combined in a linear model, also learned from training data.

Deep learning was used with CT from other regions besides the heart. Zheng et al. [205] created a method for 3D detection in volumetric data which they quantitatively evaluated for carotid artery bifurcation detection in CT. An one hidden layer network was used for the initial testing of all voxels to obtain a small number of candidates, followed by a more accurate classification with a deep network. The learned image features are further combined with Haar wavelet features to increase the detection accuracy. Montoya et al. [206] trained a 30 layer ResNet to generate 3D cerebral angiograms from contrast-enhanced images using three tissue types (vasculature, bone and soft tissue). They created the annotations using thresholding and connected components in 3D space, having a combined dataset of 13790 images.

CT has also been used for other tasks. Zreik et al. [207] created a method to identify patients with coronary artery stenoses from the LV myocardium in rest CT. They used a multi-scale CNN to segment the LV myocardium and then encoded it using an unsupervised convolutional AE. Thereafter, the final classification is done using an SVM classifier based on the extracted and clustered encodings. Similar work has been done by the same authors in [233] which they use three CNNs to detect a bounding box around the LV and perform LV voxel classification within the bounding box. Commandeur et al. [208] used a combination of two deep networks to quantify epicardial and thoracic apidose tissue in CT from 250 patients with 55 slices per patient on average. The first network is a six *layer* CNN that detects the slice located within heart limits, and segments the thoracic and epicardial-paracardial masks. The second network is a five *layer* CNN that detects the pericardium line from the CT scan in cylindrical coordinates. Then a statistical shape model regularization along with thresholding and median filtering provide the final segmentations. Gulsun et al. [209] created a method for the extraction of blood vessel centerlines in CT.

First, optimal paths in a computed flow field are found and then a CNN classifier is used for removing extraneous paths in the detected centerlines. The method was enhanced using a model-based detection of coronary specific territories and main branches to constrain the search space.

D. Echocardiography

Echocardiography is an imaging modality that depicts the heart area using ultrasound waves. Uses of deep learning in echocardiography mainly include LV segmentation and quality score assessment amongst others.

DBNs have been used for LV segmentation in echocardiography. In [210] the authors created a method that decouples the rigid and nonrigid detections with a DBN that models the appearance of the LV demonstrating that it is more robust than level sets and deformable templates. Nascimento et al. [211] used manifold learning that partitions the data into patches that each one proposes a segmentation of the LV. The fusion of the patches was done by a DBN multi-classifier that assigns a weight for each patch. In that way the method does not rely on a single segmentation and the training process produces robust appearance models without the need of large training sets. In [212] the authors used a multi-domain regularized FCN and transfer learning. They compare their method with simpler FCN architectures and a state-of-the-art method demonstrating better results.

Echocardiography has also been used for viewpoint classification. Madani et al. [213] trained a six *layer* CNN to classify between 15 views (12 video and 3 still) of transthoracic echocardiogram images, achieving better results than certified echocardographers. In [214] the authors created a residual 3D CNN for ejetion fraction classification from transthoracic echocardiogram images. They used 8715 exams each one with 30 sequential frames of the apical 4 chamber to train and test their method achieving preliminary results. Gao et al. [215] incorporated spatial and temporal information sustained by the video images of the moving heart by fusing two CNNs with seven *layers* each. The acceleration measurement at each point was calculated using dense optical flow method to represent temporal motion information. Subsequently, the fusion of the CNNs was done using linear integrations of the vectors of their outputs. Comparisons were made with previous hand-engineering approaches demonstrating superior results.

Quality score assessment and other tasks were also targeted using echocardiography. In [216] the authors created a method for reducing operator variability in data acquisition by computing an echo quality score for real-time feedback. The model consisted of convolutional layers to extract features from the input echo cine and recurrent layers to use the sequential information in the echo cine loop. Ghesu et al. [217] method for object detection and segmentation in the context of volumetric image parsing, is done solving anatomical pose estimation and boundary delineation. For this task they introduce marginal space deep learning which provides high run-time performance by learning classifiers in clustered, high-probability regions in spaces of gradually increasing dimensionality. Given the object localization, they propose a

combined deep learning active shape model to estimate the non-rigid object boundary. In their article Perrin et al. [218] trained and evaluated AlexNet with 59151 echo frames in a pairwise fashion to classify between five pediatric populations with congenital heart disease. Moradi et al. [219] created a method based on VGGnet and doc2vec [234] to produce semantic descriptors for images which can be used as weakly labeled instances or corrected by medical experts. Their model was able to identify 91% of diseases instances and 77% of disease severity modifiers from Doppler images of cardiac valves.

E. Optical coherence tomography

Optical Coherence Tomography (OCT) is an intravascular imaging modality that provides cross-sectional images of arteries with high resolution and reproducible quantitative measurements of the coronary geometry in the clinical setting [235].

In their article Roy et al. [220] characterized tissue in OCT by learning the multi-scale statistical distribution model of the data with a distribution preserving AE. The learning rule of the network introduces a scale importance parameter associated with error backpropagation. Compared with three baseline pretrained AEs with cross entropy it achieves better performance in terms of accuracy in plaque/normal pixel detection. Yong et al. [221] created a linear-regression CNN with four *layers* to segment vessel lumen, parameterized in terms of radial distances from the catheter centroid in polar space. The high accuracy of this method along with its computational efficiency (40.6ms/image) suggest the potential of being used in the real clinical environment. In [222] the authors compared the discriminative capability of deep features extracted from each one of AlexNet, GoogleNet, VGG-16 and VGG-19 to identify fibroatheroma. Data augmentation was applied on a dataset of OCT images for each classification scheme and linear SVM was conducted to classify normal and fibroatheroma images. Results indicate that VGG-19 is better in identifying images that contain fibroatheroma. Abdolmanafi et al. [223] classify tissue in OCT using a pretrained AlexNet as feature extractor and compare the predictions of three classifiers, CNN, RF, and SVM, with the first one achieving the best results.

F. Other imaging modalities

Intravascular Ultrasound (IVUS) uses ultraminiaturized transducers mounted on modified intracoronary catheters to provide radial anatomic imaging of intracoronary calcification and plaque formation [236]. Lekadir et al. [224] used a patched-based four *layer* CNN for characterization of plaque composition in carotid ultrasound images. Experiments done by the authors showed that the model achieved better pixel-based accuracy than single-scale and multi-scale SVMs. In [225] the authors automated the entire process of carotid intima media thickness video interpretation. They trained a two *layer* CNN with two outputs for frame selection and a two *layer* CNN with three outputs for ROI localization and intima-media thickness measurements. This model performs much

better than a previous handcrafted method by the same authors, which they justify in CNN's capability to learn the appearance of QRS and ROI instead of relying on a thresholded R-peak amplitude and curvature. Tom et al. [226] created a Generative Adversarial Network (GAN) based method for fast simulation of realistic IVUS. Stage 0 simulation was performed using pseudo B-mode IVUS simulator and yielded speckle mapping of a digitally defined phantom. Stage I refined the mappings to preserve tissue specific speckle intensities using a GAN with four residual blocks and Stage II GAN generated high resolution images with patho-realistic speckle profiles.

Other cardiology related applications with deep learning use modalities such as mammograms, X-Ray, percutaneous transluminal angioplasty, biopsy images and myocardial perfusion imaging. In their article Wang et al. [227] apply a pixelwise, patch-based procedure for breast arterial calcification detection in mammograms using a ten *layer* CNN and morphologic operation for post-processing. The authors used 840 images and their experiments resulted in a model that achieved a coefficient of determination of 96.2%. Liu et al. [228] trained CNNs using 1768 X-Ray images with corresponding diagnostic reports. The average diagnostic accuracies of the models reached to a maximum of 0.89 as the depth increased reaching eight layers; after that the increase in accuracy was limited. In [229] the authors created a method to denoise low dose percutaneous transluminal coronary angioplasty images. They tested mean squared error and structural similarity based loss functions on two patch-based CNNs with four *layers* and compared them for different types and levels of noise. Nirschl et al. [230] used endomyocardial biopsy images from 209 patients to train and test a patch-based six *layer* CNN to identify heart failure. Rotational data augmentation was also used while the outputs of the CNN on each patch were averaged to obtain the image-level probability. This model demonstrated better results than AlexNet, Inception and ResNet50. In [231] the authors trained a three *layer* CNN for the prediction of obstructive CAD stress myocardial perfusion from 1638 patients and compared it with total perfusion deficit. The model computes a probability per vessel during training, while in testing the maximum probabilities per artery are used per patient score. Results show that this method outperforms the total perfusion deficit in the per vessel and per patient prediction tasks.

G. Overall view on deep learning using CT, echocardiography, OCT and other imaging modalities

Papers using these imaging modalities were highly variable in terms of the research question they were trying to solve and highly inconsistent in respect with the use of metrics for the results they reported. These imaging modalities also lack behind in having publicly available database thus limiting opportunities for new architectures to be tested by groups that do not have an immediate clinical partner. On the other hand there is relatively high uniformity regarding the use of architectures with CNN most widely used, especially the pretrained top performing architectures from the ImageNet competition (AlexNet, VGG, GoogleNet, ResNet).

TABLE X
REVIEWS OF DEEP LEARNING APPLICATIONS IN CARDIOLOGY

Reference	Application/Notes
Mayer 2015 [237]	Big data in cardiology changes how insights are discovered
Austin 2016 [238]	overview of big data, its benefits, potential pitfalls and future impact in cardiology
Greenspan 2016 [239]	lesion detection, segmentation and shape modeling
Miotto 2017 [240]	imaging, EHR, genome and wearable data and needs for increasing interpretability
Krittanawong 2017 [241]	studies on image recognition technology which predict better than physicians
Litjens 2017 [242]	image classification, object detection, segmentation and registration
Qayyum 2017 [243]	CNN-based methods in image segmentation, classification, diagnosis and image retrieval
Hengling 2017 [244]	impact that machine learning will have on the future of cardiovascular imaging
Blair 2017 [245]	advances in neuroimaging with MRI on small vessel disease
Slomka 2017 [246]	nuclear cardiology, CT angiography, Echocardiography, MRI
Carneiro 2017 [247]	mammography, cardiovascular and microscopy imaging
Johnson 2018 [248]	AI in cardiology describing predictive modeling concepts, common algorithms and use of deep learning
Jiang 2017 [249]	AI applications in stroke detection, diagnosis, treatment, outcome prediction and prognosis evaluation
Lee 2017 [250]	AI in stroke imaging focused in technical principles and clinical applications
Loh 2017 [251]	heart disease diagnosis and management within the context of rural healthcare
Krittanawong 2017 [252]	cardiovascular clinical care and role in facilitating precision cardiovascular medicine
Gomez 2018 [253]	recent advances in automation and quantitative analysis in nuclear cardiology
Shameer 2018 [254]	promises and limitations of implementing machine learning in cardiovascular medicine
Shrestha 2018 [255]	machine learning applications in nuclear cardiology
Kikuchi 2018 [256]	application of AI in nuclear cardiology and the problem of limited number of data
Awan 2018 [257]	machine learning applications in heart failure diagnosis, classification, readmission prediction and medication adherence
Faust 2018 [258]	deep learning application in physiological data including ECG

VI. DISCUSSION AND FUTURE DIRECTIONS

It is evident from the literature that deep learning methods will replace rule-based expert systems and traditional machine learning based on feature engineering. In [257] the authors argue that deep learning is better in visualizing complex patterns hidden in high dimensional medical data. Krittanawong et al. [252] argue that the increasing availability of automated real-time AI tools in EHRs will reduce the need for scoring systems such as the Framingham risk. In [241] the authors argue that AI predictive analytics and personalized clinical support for medical risk identification are superior to human cognitive capacities. Moreover, AI may facilitate communication between physicians and patients by decreasing processing times and therefore increasing the quality of patient care. Loh et al. [251] argue that deep learning and mobile technologies would expedite the proliferation of healthcare services to those in impoverished regions which in turn leads to further decline of disease rates. Mayer et al. [237] state that big data promises to change cardiology through an increase in the data gathered but its impact goes beyond improving existing methods such as changing how insights are discovered.

Deep learning requires large training datasets to achieve high quality results [3]. This is especially difficult with medical data, considering that the labeling procedure of medical data is costly because it requires manual labor from medical experts. Moreover most of medical data belong to the normal cases instead to abnormal, making them highly unbalanced. Other challenges of applying deep learning in medicine that previous literature has identified are data standardization/availability/dimensionality/volume/quality issues, difficulty in acquiring the corresponding annotations and noise in annotations [239], [240], [242], [246]. More specifically, in [245] the authors note that deep learning applications on small vessel disease have been developed using only a

few representative datasets and they need to be evaluated in large multi-center datasets. Kikuchi et al. [256] mention that compared with CT and MRI, nuclear cardiology imaging modalities have limited number of images per patient and only specific number of organs are depicted. Liebeskind [259] states that machine learning methods are tested on selective and homogeneous clinical data but generalizability would occur using heterogeneous and complex data. The example of ischemic stroke is mentioned as an example of an heterogeneous and complex disease where the occlusion of middle cerebral artery can lead to divergent imaging patterns. In [253] the authors conclude that additional data that validate these applications in multi-center uncontrolled clinical settings are required before implementation in routine clinical use. The impact of these tools on decision-making, down-stream utilization of resources, cost, and value-based practice also needs to be investigated. Moreover the present literature demonstrated that there is an unbalanced distribution of publicly available datasets among different imaging modalities in cardiology (e.g. no public dataset available for OCT in contrast with MRI).

Previous literature states that problems related to data can be solved using data augmentation, open collaboration between research organizations and increase in funding. Hengling et al. [244] argue that substantial investments will be required to create high quality annotated databases which are essential for the success of supervised deep learning methods. In [238] the authors argue that the continued success of this field depends on sustained technological advancements in information technology and computer architecture as well as collaboration and open exchange of data between physicians and other stakeholders. Lee et al. [250] conclude that international cooperation is required for constructing a high quality multimodal big dataset for stroke imaging. Another solution to better exploit big medical data in cardiology is to apply unsupervised learning methods, which do not require annotations. The

present review demonstrated that unsupervised learning is not thoroughly used since the majority of the methods in all modalities are supervised.

Regarding the problem of lack of interpretability as it is indicated by Hinton [260] it is generally infeasible to interpret nonlinear features of deep networks because their meaning depends on complex interactions with uninterpreted features from other layers. Additionally these models are stochastic, meaning that every time a network fits the same data but with different initial weights different features are learned. More specifically, in an extensive review [231] of whether the problem of LV/RV segmentation is solved the authors state that although the classification aspect of the problem achieves near perfect results the use of a 'diagnostic black box' can not be integrated in the clinical practice. Miotto et al. [240] mention interpretability as one of the main challenges facing the clinical application of deep learning to healthcare. In [250] the authors note that the black-box nature of AI methods like deep learning is against the concept of evidence-based medicine and it raises legal and ethical issues in using them in clinical practice. This lack of interpretability is the main reason that medical experts resist using these models and there are also legal restrictions regarding the medical use of the non-interpretable applications [246]. On the other hand, any model can be placed in a 'human-machine decision effort' axis [261] including statistical ones that medical experts rely on for everyday clinical decision making. For example, human decisions such as choosing which variables to include in the model, the relationship of dependent and independent variables and variable transformations, move the algorithm to the human decision axis, thus making it more interpretable but in the same time more error-prone.

Regarding the solution to the interpretability problem when new methods are necessary researchers should prefer making simpler deep learning methods (end-to-end and non-ensembles) to increase their clinical applicability, even if that means reduced reported accuracy. There are also arguments against creating new methods but instead focus on validating the existing ones. In [262] the authors conclude that there is an excess of models predicting incident CVD in the general population. Most of the models usefulness is unclear due to errors in methodology and lack of external validation studies. Instead of developing new CVD risk prediction models future research should focus on validating and comparing existing models and investigate whether they can be improved.

A popular method used for interpretable models is attention networks [263]. Attention networks are inspired by the ability of human vision to focus in a certain point with high resolution while perceiving the surroundings with low resolution and then adjusting the focal point. They have been used by a number of publications in cardiology in medical history prediction [70], ECG beat classification [86] and CVD prediction using fundus [192]. Another simpler tool for interpretability is saliency maps [264] that uses the gradient of the output with respect to the input which intuitively shows the regions that most contribute toward the output.

Besides solving the data and interpretability problems, researchers in cardiology could utilize the already established deep learning architectures that have not been widely applied in cardiology such as capsule networks. Capsule networks [265] are deep neural networks that require less training data than CNNs and its layers capture the 'pose' of features thus making their inner-workings more interpretable and closer to the human way of perception. However an important constraint they currently have which limits them from achieving wider use, is the high computational cost compared to CNNs due to the 'routing by agreement' algorithm. Amongst their recent uses in medicine include brain tumor classification [266] and breast cancer classification [267]. Capsule networks have not been used in cardiology data yet.

Another underused deep learning architecture in cardiology is GANs [268] that consist of a generator that creates fake images from noise and a discriminator that is responsible of differentiating between fake images from the generator and real images. Both networks try to optimize a loss in a zero-sum game resulting in a generator that produces realistic images. GANs have only been used for simulating patho-realistic IVUS images [226] and the cardiology field has a lot to gain from using this kind of models, especially in the absense of high quality annotated data.

Researchers could also utilize CRFs which are graphical models that capture context-aware information and are able to incorporate higher order statistics, which traditional deep learning methods are unable to do. CRFs have been jointly trained with CNNs and have been used in depth estimation in endoscopy [269] and liver segmentation in CT [270]. There are also cardiology applications that used CRFs with deep learning as a segmentation refinement step in fundus photography [171], [174], and in LV/RV [143]. Multimodal deep learning [271] can also be used to improve diagnostic outcomes e.g. the possibility of combining fMRI and ECG data. Dedicated databases must be created in order to increase research in this area since according to the current review there are only three cardiology databases with multimodal data. In addition to the previous databases MIMIC-III has also been used for multimodal deep learning by [68] for predicting in-hospital, short/long-term mortality and ICD-9 code predictions.

Conclusions

With each technological advance, cardiology and medicine in general is becoming human independent and closer to an automated AI driven field. AI will not only reach the point where it uses real-time physical scans to detect diseases, but it will also interpret ambiguous conditions, precisely phenotype complex diseases and take medical decisions. However, a complete theoretical understanding of deep learning is not yet available and a critical understanding of the strengths and limitations of its inner workings is vital for the field to gain its place in everyday clinical use. Successful application of AI in the medical field relies on achieving interpretable models and big datasets.

Acknowledgements

REFERENCES

[1] E. J. Benjamin, M. J. Blaha, S. E. Chiuve, M. Cushman, S. R. Das, R. Deo, S. D. de Ferranti, J. Floyd, M. Fornage, C. Gillespie *et al.*, "Heart disease and stroke statistics—2017 update: a report from the american heart association," *Circulation*, vol. 135, no. 10, pp. e146–e603, 2017.

[2] E. Wilkins, L. Wilson, K. Wickramasinghe, P. Bhatnagar, J. Leal, R. Luengo-Fernandez, R. Burns, M. Rayner, and N. Townsend, "European cardiovascular disease statistics 2017," 2017.

[3] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097–1105.

[4] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2015, pp. 234–241.

[5] R. Collobert and J. Weston, "A unified architecture for natural language processing: Deep neural networks with multitask learning," in *Proceedings of the 25th international conference on Machine learning*. ACM, 2008, pp. 160–167.

[6] A. Graves, A.-r. Mohamed, and G. Hinton, "Speech recognition with deep recurrent neural networks," in *Acoustics, speech and signal processing (icassp), 2013 ieee international conference on*. IEEE, 2013, pp. 6645–6649.

[7] B. Alipanahi, A. Delong, M. T. Weirauch, and B. J. Frey, "Predicting the sequence specificities of dna-and rna-binding proteins by deep learning," *Nature biotechnology*, vol. 33, no. 8, p. 831, 2015.

[8] Y. Bengio, Y. LeCun *et al.*, "Scaling learning algorithms towards ai," *Large-scale kernel machines*, vol. 34, no. 5, pp. 1–41, 2007.

[9] F. Rosenblatt, "The perceptron: a probabilistic model for information storage and organization in the brain." *Psychological review*, vol. 65, no. 6, p. 386, 1958.

[10] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *nature*, vol. 323, no. 6088, p. 533, 1986.

[11] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT press Cambridge, 2016, vol. 1.

[12] X. Glorot, A. Bordes, and Y. Bengio, "Deep sparse rectifier neural networks," in *Proceedings of the fourteenth international conference on artificial intelligence and statistics*, 2011, pp. 315–323.

[13] J. A. Hanley and B. J. McNeil, "The meaning and use of the area under a receiver operating characteristic (roc) curve." *Radiology*, vol. 143, no. 1, pp. 29–36, 1982.

[14] L. R. Dice, "Measures of the amount of ecologic association between species," *Ecology*, vol. 26, no. 3, pp. 297–302, 1945.

[15] K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators," *Neural networks*, vol. 2, no. 5, pp. 359–366, 1989.

[16] G. E. Hinton, S. Osindero, and Y.-W. Teh, "A fast learning algorithm for deep belief nets," *Neural computation*, vol. 18, no. 7, pp. 1527–1554, 2006.

[17] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[18] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 1–9.

[19] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

[20] P. Vincent, H. Larochelle, I. Lajoie, Y. Bengio, and P.-A. Manzagol, "Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion," *Journal of Machine Learning Research*, vol. 11, no. Dec, pp. 3371–3408, 2010.

[21] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[22] K. Cho, B. Van Merriënboer, D. Bahdanau, and Y. Bengio, "On the properties of neural machine translation: Encoder-decoder approaches," *arXiv preprint arXiv:1409.1259*, 2014.

[23] A. E. Johnson, T. J. Pollard, L. Shen, H. L. Li-wei, M. Feng, M. Ghassemi, B. Moody, P. Szolovits, L. A. Celi, and R. G. Mark, "Mimic-iii, a freely accessible critical care database," *Scientific data*, vol. 3, p. 160035, 2016. [Online]. Available: https://mimic.physionet.org

[24] S. Kweon, Y. Kim, M.-j. Jang, Y. Kim, K. Kim, S. Choi, C. Chun, Y.-H. Khang, and K. Oh, "Data resource profile: the korea national health and nutrition examination survey (knhanes)," *International journal of epidemiology*, vol. 43, no. 1, pp. 69–77, 2014. [Online]. Available: http://knhanes.cdc.go.kr

[25] W. Karlen, S. Raman, J. M. Ansermino, and G. A. Dumont, "Multiparameter respiratory rate estimation from the photoplethysmogram," *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 7, pp. 1946–1953, 2013. [Online]. Available: http://www.capnobase.org/database/pulse-oximeter-ieee-tbme-benchmark/

[26] F. Nolle, F. Badura, J. Catlett, R. Bowser, and M. Sketch, "Crei-gard, a new concept in computerized arrhythmia monitoring systems," *Computers in Cardiology*, vol. 13, pp. 515–518, 1986. [Online]. Available: https://physionet.org/physiobank/database/cudb/

[27] G. Moody, "A new method for detecting atrial fibrillation using rr intervals," *Computers in Cardiology*, pp. 227–230, 1983. [Online]. Available: https://physionet.org/physiobank/database/afdb/

[28] D. S. Baim, W. S. Colucci, E. S. Monrad, H. S. Smith, R. F. Wright, A. Lanoue, D. F. Gauthier, B. J. Ransil, W. Grossman, and E. Braunwald, "Survival of patients with severe congestive heart failure treated with oral milrinone," *Journal of the American College of Cardiology*, vol. 7, no. 3, pp. 661–670, 1986. [Online]. Available: https://www.physionet.org/physiobank/database/chfdb/

[29] A. L. Goldberger, L. A. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley, "Physiobank, physiotoolkit, and physionet," *Circulation*, vol. 101, no. 23, pp. e215–e220, 2000. [Online]. Available: https://www.physionet.org/pn3/incartdb/

[30] S. Petrutiu, A. V. Sahakian, and S. Swiryn, "Abrupt changes in fibrillatory wave characteristics at the termination of paroxysmal atrial fibrillation in humans," *Europace*, vol. 9, no. 7, pp. 466–470, 2007. [Online]. Available: https://physionet.org/pn3/ltafdb/

[31] F. Jager, A. Taddei, G. B. Moody, M. Emdin, G. Antolič, R. Dorn, A. Smrdel, C. Marchesi, and R. G. Mark, "Long-term st database: a reference for the development and evaluation of automated ischaemia detectors and for the study of the dynamics of myocardial ischaemia," *Medical and Biological Engineering and Computing*, vol. 41, no. 2, pp. 172–182, 2003. [Online]. Available: https://www.physionet.org/physiobank/database/ltstdb/

[32] G. B. Moody and R. G. Mark, "The impact of the mit-bih arrhythmia database," *IEEE Engineering in Medicine and Biology Magazine*, vol. 20, no. 3, pp. 45–50, 2001. [Online]. Available: https://physionet.org/challenge/2017/

[33] G. B. Moody, W. Muldrow, and R. G. Mark, "A noise stress test for arrhythmia detectors," *Computers in cardiology*, vol. 11, no. 3, pp. 381–384, 1984. [Online]. Available: https://physionet.org/physiobank/database/nstdb/

[34] R. L. Goldsmith, J. T. Bigger, R. C. Steinman, and J. L. Fleiss, "Comparison of 24-hour parasympathetic activity in endurance-trained and untrained young men," *Journal of the American College of Cardiology*, vol. 20, no. 3, pp. 552–558, 1992. [Online]. Available: https://www.physionet.org/physiobank/database/nsr2db/

[35] N. Iyengar, C. Peng, R. Morin, A. L. Goldberger, and L. A. Lipsitz, "Age-related alterations in the fractal scaling of cardiac interbeat interval dynamics," *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, vol. 271, no. 4, pp. R1078–R1084, 1996. [Online]. Available: https://www.physionet.org/physiobank/database/fantasia/

[36] C. Liu, D. Springer, Q. Li, B. Moody, R. A. Juan, F. J. Chorro, F. Castells, J. M. Roig, I. Silva, A. E. Johnson *et al.*, "An open access database for the evaluation of heart sound algorithms," *Physiological Measurement*, vol. 37, no. 12, p. 2181, 2016. [Online]. Available: https://physionet.org/pn3/challenge/2016/

[37] R. Bousseljot, D. Kreiseler, and A. Schnabel, "Nutzung der ekg-signaldatenbank cardiodat der ptb über das internet," *Biomedizinische Technik/Biomedical Engineering*, vol. 40, no. s1, pp. 317–318, 1995.

[38] P. Laguna, R. G. Mark, A. Goldberg, and G. B. Moody, "A database for evaluation of algorithms for measurement of qt and other waveform intervals in the ecg," in *Computers in cardiology 1997*. IEEE, 1997, pp. 673–676. [Online]. Available: https://physionet.org/physiobank/database/qtdb/

[39] S. D. Greenwald, R. S. Patil, and R. G. Mark, "Improved detection and classification of arrhythmias in noise-corrupted electrocardiograms using contextual information," in *Computers in Cardiology 1990, Proceedings*. IEEE, 1990, pp. 461–464. [Online]. Available: https://physionet.org/physiobank/database/svdb/

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/RBME.2018.2885714, IEEE Reviews in Biomedical Engineering

22

[40] I. Silva, J. Behar, R. Sameni, T. Zhu, J. Oster, G. D. Clifford, and G. B. Moody, "Noninvasive fetal ecg: the physionet/computing in cardiology challenge 2013," in *Computing in Cardiology Conference (CinC), 2013*. IEEE, 2013, pp. 149–152. [Online]. Available: https://physionet.org/challenge/2013/

[41] M.-H. Wu and E. Y. Chang, "Deepq arrhythmia database: A large-scale dataset for arrhythmia detector evaluation," in *Proceedings of the 2nd International Workshop on Multimedia for Personal Health and Health Care*. ACM, 2017, pp. 77–80.

[42] P. Radau, Y. Lu, K. Connelly, G. Paul, A. Dick, and G. Wright, "Evaluation framework for algorithms segmenting short axis cardiac MRI," in *The MIDAS Journal - Cardiac MR Left Ventricle Segmentation Challenge*, 2009. [Online]. Available: http://www.cardiacatlas.org/studies/sunnybrook-cardiac-data/

[43] C. G. Fonseca, M. Backhaus, D. A. Bluemke, R. D. Britten, J. D. Chung, B. R. Cowan, I. D. Dinov, J. P. Finn, P. J. Hunter, A. H. Kadish *et al.*, "The cardiac atlas project - an imaging database for computational modeling and statistical atlases of the heart," *Bioinformatics*, vol. 27, no. 16, pp. 2288–2295, 2011. [Online]. Available: http://www.cardiacatlas.org/challenges/lv-segmentation-challenge/

[44] C. Petitjean, M. A. Zuluaga, W. Bai, J.-N. Dacher, D. Grosgeorge, J. Caudron, S. Ruan, I. B. Ayed, M. J. Cardoso, H.-C. Chen *et al.*, "Right ventricle segmentation from cardiac mri: a collation study," *Medical image analysis*, vol. 19, no. 1, pp. 187–202, 2015. [Online]. Available: http://www.litislab.fr/?projet=1rvsc

[45] A. Asman, A. Akhondi-Asl, H. Wang, N. Tustison, B. Avants, S. K. Warfield, and B. Landman, "Miccai 2013 segmentation algorithms, theory and applications (sata) challenge results summary," in *MIC-CAI Challenge Workshop on Segmentation: Algorithms, Theory and Applications (SATA)*, 2013.

[46] D. F. Pace, A. V. Dalca, T. Geva, A. J. Powell, M. H. Moghari, and P. Golland, "Interactive whole-heart segmentation in congenital heart disease," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2015, pp. 80–88. [Online]. Available: http://segchd.csail.mit.edu/data.html

[47] O. Bernard, A. Lalande, C. Zotti, F. Cervenansky, X. Yang, P.-A. Heng, I. Cetin, K. Lekadir, O. Camara, M. A. G. Ballester *et al.*, "Deep learning techniques for automatic mri cardiac multi-structures segmentation and diagnosis: Is the problem solved?" *IEEE Transactions on Medical Imaging*, 2018.

[48] A. Andreopoulos and J. K. Tsotsos, "Efficient and generalizable statistical models of shape and appearance for analysis of cardiac mri," *Medical Image Analysis*, vol. 12, no. 3, pp. 335–357, 2008. [Online]. Available: http://www.cse.yorku.ca/~mridataset/

[49] "Data science bowl cardiac challenge data," 2016. [Online]. Available: https://www.kaggle.com/c/second-annual-data-science-bowl

[50] J. Zhang, B. Dashtbozorg, E. Bekkers, J. P. Pluim, R. Duits, and B. M. ter Haar Romeny, "Robust retinal vessel segmentation via locally adaptive derivative frames in orientation scores," *IEEE transactions on medical imaging*, vol. 35, no. 12, pp. 2631–2644, 2016.

[51] J. Staal, M. D. Abràmoff, M. Niemeijer, M. A. Viergever, and B. Van Ginneken, "Ridge-based vessel segmentation in color images of the retina," *IEEE transactions on medical imaging*, vol. 23, no. 4, pp. 501–509, 2004. [Online]. Available: https://www.isi.uu.nl/Research/Databases/DRIVE/

[52] A. Hoover, V. Kouznetsova, and M. Goldbaum, "Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response," *IEEE Transactions on Medical imaging*, vol. 19, no. 3, pp. 203–210, 2000. [Online]. Available: http://cecas.clemson.edu/~ahoover/stare/

[53] C. G. Owen, A. R. Rudnicka, R. Mullen, S. A. Barman, D. Monekosso, P. H. Whincup, J. Ng, and C. Paterson, "Measuring retinal vessel tortuosity in 10-year-old children: validation of the computer-assisted image analysis of the retina (caiar) program," *Investigative ophthalmology & visual science*, vol. 50, no. 5, pp. 2004–2010, 2009. [Online]. Available: https://blogs.kingston.ac.uk/retinal/chasedb1/

[54] J. Odstrcilik, R. Kolar, A. Budai, J. Hornegger, J. Jan, J. Gazarek, T. Kubena, P. Cernosek, O. Svoboda, and E. Angelopoulou, "Retinal vessel segmentation by improved matched filtering: evaluation on a new high-resolution fundus image database," *IET Image Processing*, vol. 7, no. 4, pp. 373–383, 2013. [Online]. Available: https://www5.cs.fau.de/research/data/fundus-images/

[55] B. Graham, "Kaggle diabetic retinopathy detection competition report," *University of Warwick*, 2015. [Online]. Available: https://www.kaggle.com/c/diabetic-retinopathy-detection

[56] E. Decencière, G. Cazuguel, X. Zhang, G. Thibault, J.-C. Klein, F. Meyer, B. Marcoteguí, G. Quellec, M. Lamard, R. Danno *et al.*, "Teleophta: Machine learning and image processing methods for teleophthalmology," *Irbm*, vol. 34, no. 2, pp. 196–203, 2013.

[57] E. Decencière, X. Zhang, G. Cazuguel, B. Lay, B. Cochener, C. Trone, P. Gain, R. Ordonez, P. Massin, A. Erginay *et al.*, "Feedback on a publicly distributed image database: the messidor database," *Image Analysis & Stereology*, vol. 33, no. 3, pp. 231–234, 2014. [Online]. Available: http://latim.univ-brest.fr/indexfce0.html

[58] T. Kauppi, J.-K. Kämäräinen, L. Lensu, V. Kalesnykiene, I. Sorri, H. Uusitalo, and H. Kälviäinen, "Constructing benchmark databases and protocols for medical image analysis: Diabetic retinopathy," *Computational and mathematical methods in medicine*, vol. 2013, 2013. [Online]. Available: http://www.it.lut.fi/project/imageret/diaretdb1/

[59] M. Niemeijer, M. Van Ginneken, M. J. Cree, A. Mizutani, G. Quellec, C. I. Sánchez, B. Zhang, R. Hornero, M. Lamard, C. Muramatsu *et al.*, "Retinopathy online challenge: automatic detection of microaneurysms in digital color fundus photographs," *IEEE transactions on medical imaging*, vol. 29, no. 1, pp. 185–195, 2010. [Online]. Available: http://roc.healthcare.uiowa.edu/

[60] S. Balocco, C. Gatta, F. Ciompi, A. Wahle, P. Radeva, S. Carlier, G. Unal, E. Sanidas, J. Mauri, X. Carillo *et al.*, "Standardized evaluation methodology and reference database for evaluating ivus image segmentation," *Computerized medical imaging and graphics*, vol. 38, no. 2, pp. 70–90, 2014. [Online]. Available: http://www.cvc.uab.es/IVUSchallenge2011/

[61] C. Sudlow, J. Gallacher, N. Allen, V. Beral, P. Burton, J. Danesh, P. Downey, P. Elliott, J. Green, M. Landray *et al.*, "Uk biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age," *PLoS medicine*, vol. 12, no. 3, p. e1001779, 2015. [Online]. Available: https://biobank.ctsu.ox.ac.uk/crystal/download.cgi

[62] H. Kirişli, M. Schaap, C. Metz, A. Dharampal, W. B. Meijboom, S.-L. Papadopoulou, A. Dedic, K. Nieman, M. De Graaf, M. Meijs *et al.*, "Standardized evaluation framework for evaluating coronary artery stenosis detection, stenosis quantification and lumen segmentation algorithms in computed tomography angiography," *Medical image analysis*, vol. 17, no. 8, pp. 859–876, 2013. [Online]. Available: http://www.coronary.bigr.nl/stensoses/

[63] P. H. Charlton, T. Bonnici, L. Tarassenko, D. A. Clifton, R. Beale, and P. J. Watkinson, "An assessment of algorithms to estimate respiratory rate from the electrocardiogram and photoplethysmogram," *Physiological measurement*, vol. 37, no. 4, p. 610, 2016. [Online]. Available: https://peterhcharlton.github.io/RRest/vortal_dataset.html

[64] C. Tobon-Gomez, A. J. Geers, J. Peters, J. Weese, K. Pinto, R. Karim, M. Ammar, A. Daoudi, J. Margeta, Z. Sandoval *et al.*, "Benchmark for algorithms segmenting the left atrium from 3d ct and mri datasets," *IEEE transactions on medical imaging*, vol. 34, no. 7, pp. 1460–1473, 2015. [Online]. Available: http://www.cardiacatlas.org/challenges/left-atrium-segmentation-challenge/

[65] X. Zhuang and J. Shen, "Multi-scale patch and multi-modality atlases for whole heart segmentation of mri," *Medical image analysis*, vol. 31, pp. 77–87, 2016. [Online]. Available: http://www.sdspeople.fudan.edu.cn/zhuangxiahai/0/mmwhs/

[66] S. Gopalswamy, P. J. Tighe, and P. Rashidi, "Deep recurrent neural networks for predicting intraoperative and postoperative outcomes and trends," in *Biomedical & Health Informatics (BHI), 2017 IEEE EMBS International Conference on*. IEEE, 2017, pp. 361–364.

[67] E. Choi, A. Schuetz, W. F. Stewart, and J. Sun, "Using recurrent neural network models for early detection of heart failure onset," *Journal of the American Medical Informatics Association*, vol. 24, no. 2, pp. 361–370, 2016.

[68] S. Purushotham, C. Meng, Z. Che, and Y. Liu, "Benchmarking deep learning models on large healthcare datasets," *Journal of biomedical informatics*, 2018.

[69] E. C. Polley and M. J. Van Der Laan, "Super learner in prediction," 2010.

[70] Y. J. Kim, Y.-G. Lee, J. W. Kim, J. J. Park, B. Ryu, and J.-W. Ha, "Highrisk prediction from electronic medical records via deep attention networks," *arXiv preprint arXiv:1712.00010*, 2017.

[71] H. C. Hsiao, S. H. Chen, and J. J. Tsai, "Deep learning for risk analysis of specific cardiovascular diseases using environmental data and outpatient records," in *Bioinformatics and Bioengineering (BIBE), 2016 IEEE 16th International Conference on*. IEEE, 2016, pp. 369–372.

[72] Z. Huang, W. Dong, H. Duan, and J. Liu, "A regularized deep learning approach for clinical risk prediction of acute coronary syndrome using electronic health records," *IEEE Transactions on Biomedical Engineering*, vol. 65, no. 5, pp. 956–968, 2018.

[73] J. Kim, U. Kang, and Y. Lee, "Statistics and deep belief network-based cardiovascular risk prediction," *Healthcare informatics research*, vol. 23, no. 3, pp. 169–175, 2017.

[74] S. Faziludeen and P. Sabiq, "Ecg beat classification using wavelets and svm," in *Information & Communication Technologies (ICT), 2013 IEEE Conference on*. IEEE, 2013, pp. 815–818.

[75] M. Zubair, J. Kim, and C. Yoon, "An automated ecg beat classification system using convolutional neural networks," in *IT Convergence and Security (ICITCS), 2016 6th International Conference on*. IEEE, 2016, pp. 1–5.

[76] D. Li, J. Zhang, Q. Zhang, and X. Wei, "Classification of ecg signals based on 1d convolution neural network," in *e-Health Networking, Applications and Services (Healthcom), 2017 IEEE 19th International Conference on*. IEEE, 2017, pp. 1–6.

[77] S. Kiranyaz, T. Ince, and M. Gabbouj, "Real-time patient-specific ecg classification by 1-d convolutional neural networks," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 3, pp. 664–675, 2016.

[78] A. Isin and S. Ozdalili, "Cardiac arrhythmia detection using deep learning," *Procedia Computer Science*, vol. 120, pp. 268–275, 2017.

[79] K. Luo, J. Li, Z. Wang, and A. Cuschieri, "Patient-specific deep architectural model for ecg classification," *Journal of healthcare engineering*, vol. 2017, 2017.

[80] C. Jiang, S. Song, and M. Q.-H. Meng, "Heartbeat classification system based on modified stacked denoising autoencoders and neural networks," in *Information and Automation (ICIA), 2017 IEEE International Conference on*. IEEE, 2017, pp. 511–516.

[81] J. Yang, Y. Bai, F. Lin, M. Liu, Z. Hou, and X. Liu, "A novel electrocardiogram arrhythmia classification method based on stacked sparse auto-encoders and softmax regression," *International Journal of Machine Learning and Cybernetics*, pp. 1–8, 2017.

[82] Z. Wu, X. Ding, G. Zhang, X. Xu, X. Wang, Y. Tao, and C. Ju, "A novel features learning method for ecg arrhythmias using deep belief networks," in *Digital Home (ICDH), 2016 6th International Conference on*. IEEE, 2016, pp. 192–196.

[83] M.-H. Wu, E. J. Chang, and T.-H. Chu, "Personalizing a generic ecg heartbeat classification for arrhythmia detection: A deep learning approach," in *2018 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR)*. IEEE, 2018, pp. 92–99.

[84] P. Rajpurkar, A. Y. Hannun, M. Haghpanahi, C. Bourn, and A. Y. Ng, "Cardiologist-level arrhythmia detection with convolutional neural networks," *arXiv preprint arXiv:1707.01836*, 2017.

[85] U. R. Acharya, H. Fujita, O. S. Lih, Y. Hagiwara, J. H. Tan, and M. Adam, "Automated detection of arrhythmias using different intervals of tachycardia ecg segments with convolutional neural network," *Information sciences*, vol. 405, pp. 81–90, 2017.

[86] P. Schwab, G. C. Scebba, J. Zhang, M. Delai, and W. Karlen, "Beat by beat: Classifying cardiac arrhythmias with recurrent neural networks," *Computing*, vol. 44, p. 1, 2017.

[87] Z. Yao, Z. Zhu, and Y. Chen, "Atrial fibrillation detection by multi-scale convolutional neural networks," in *Information Fusion (Fusion), 2017 20th International Conference on*. IEEE, 2017, pp. 1–6.

[88] Y. Xia, N. Wulan, K. Wang, and H. Zhang, "Detecting atrial fibrillation by deep convolutional neural networks," *Computers in biology and medicine*, vol. 93, pp. 84–92, 2018.

[89] R. S. Andersen, A. Peimankar, and S. Puthusserypady, "A deep learning approach for real-time detection of atrial fibrillation," *Expert Systems with Applications*, 2018.

[90] P. Xiong, H. Wang, M. Liu, and X. Liu, "Denoising autoencoder for eletrocardiogram signal enhancement," *Journal of Medical Imaging and Health Informatics*, vol. 5, no. 8, pp. 1804–1810, 2015.

[91] B. Taji, A. D. Chan, and S. Shirmohammadi, "False alarm reduction in atrial fibrillation detection using deep belief networks," *IEEE Transactions on Instrumentation and Measurement*, 2017.

[92] R. Xiao, Y. Xu, M. M. Pelter, R. Fidler, F. Badilini, D. W. Mortara, and X. Hu, "Monitoring significant st changes through deep learning," *Journal of electrocardiology*, 2018.

[93] M. M. Al Rahhal, Y. Bazi, H. AlHichri, N. Alajlan, F. Melgani, and R. R. Yager, "Deep learning approach for active classification of electrocardiogram signals," *Information Sciences*, vol. 345, pp. 340–354, 2016.

[94] H. Abrishami, M. Campbell, C. Han, R. Czosek, and X. Zhou, "P-qrs-t localization in ecg using deep learning," in *Biomedical & Health*

[95] J. Wu, Y. Bao, S.-C. Chan, H. Wu, L. Zhang, and X.-G. Wei, "Myocardial infarction detection and classification - a new multi-scale deep feature learning approach," in *Digital Signal Processing (DSP), 2016 IEEE International Conference on*. IEEE, 2016, pp. 309–313.

[96] T. Reasat and C. Shahnaz, "Detection of inferior myocardial infarction using shallow convolutional neural networks," in *Humanitarian Technology Conference (R10-HTC), 2017 IEEE Region 10*. IEEE, 2017, pp. 718–721.

[97] W. Zhong, L. Liao, X. Guo, and G. Wang, "A deep learning approach for fetal qrs complex detection," *Physiological measurement*, vol. 39, no. 4, p. 045004, 2018.

[98] V. J. R. Ripoll, A. Wojdel, E. Romero, P. Ramos, and J. Brugada, "Ecg assessment based on neural networks with pretraining," *Applied Soft Computing*, vol. 49, pp. 399–406, 2016.

[99] L. Jin and J. Dong, "Classification of normal and abnormal ecg records using lead convolutional neural network and rule inference," *Science China Information Sciences*, vol. 60, no. 7, p. 078103, 2017.

[100] Y. Liu, Y. Huang, J. Wang, L. Liu, and J. Luo, "Detecting premature ventricular contraction in children with deep learning," *Journal of Shanghai Jiaotong University (Science)*, vol. 23, no. 1, pp. 66–73, 2018.

[101] B. Hwang, J. You, T. Vaessen, I. Myin-Germeys, C. Park, and B.-T. Zhang, "Deep ecgnet: An optimal deep learning framework for monitoring mental stress using ultra short-term ecg signals," *TELEMEDICINE and e-HEALTH*, 2018.

[102] A. Badnjević, M. Cifrek, R. Magjarević, and Z. Džemić, *Inspection of Medical Devices: For Regulatory Purposes*. Springer, 2017.

[103] J. Pan and W. J. Tompkins, "A real-time qrs detection algorithm," *IEEE transactions on biomedical engineering*, no. 3, pp. 230–236, 1985.

[104] U. R. Acharya, H. Fujita, S. L. Oh, U. Raghavendra, J. H. Tan, M. Adam, A. Gertych, and Y. Hagiwara, "Automated identification of shockable and non-shockable life-threatening ventricular arrhythmias using convolutional neural network," *Future Generation Computer Systems*, vol. 79, pp. 952–959, 2018.

[105] U. R. Acharya, H. Fujita, O. S. Lih, M. Adam, J. H. Tan, and C. K. Chua, "Automated detection of coronary artery disease using different durations of ecg segments with convolutional neural network," *Knowledge-Based Systems*, vol. 132, pp. 62–71, 2017.

[106] U. R. Acharya, H. Fujita, S. L. Oh, Y. Hagiwara, J. H. Tan, M. Adam, and R. S. Tan, "Deep convolutional neural network for the automated diagnosis of congestive heart failure using ecg signals," *Applied Intelligence*, pp. 1–12, 2018.

[107] U. R. Acharya, H. Fujita, S. L. Oh, Y. Hagiwara, J. H. Tan, and M. Adam, "Application of deep convolutional neural network for automated detection of myocardial infarction using ecg signals," *Information Sciences*, vol. 415, pp. 190–198, 2017.

[108] J. Rubin, R. Abreu, A. Ganguli, S. Nelaturi, I. Matei, and K. Sricharan, "Recognizing abnormal heart sounds using deep learning," *arXiv preprint arXiv:1707.04642*, 2017.

[109] D. Kucharski, D. Grochala, M. Kajor, and E. Kańtoch, "A deep learning approach for valve defect recognition in heart acoustic signal," in *International Conference on Information Systems Architecture and Technology*. Springer, 2017, pp. 3–14.

[110] J. P. Dominguez-Morales, A. F. Jimenez-Fernandez, M. J. Dominguez-Morales, and G. Jimenez-Moreno, "Deep neural networks for the recognition and classification of heart murmurs using neuromorphic auditory sensors," *IEEE transactions on biomedical circuits and systems*, vol. 12, no. 1, pp. 24–34, 2018.

[111] C. Potes, S. Parvaneh, A. Rahman, and B. Conroy, "Ensemble of feature-based and deep learning-based classifiers for detection of abnormal heart sounds," in *Computing in Cardiology Conference (CinC), 2016*. IEEE, 2016, pp. 621–624.

[112] H. Ryu, J. Park, and H. Shin, "Classification of heart sound recordings using convolution neural network," in *Computing in Cardiology Conference (CinC), 2016*. IEEE, 2016, pp. 1153–1156.

[113] T.-E. Chen, S.-I. Yang, L.-T. Ho, K.-H. Tsai, Y.-H. Chen, Y.-F. Chang, Y.-H. Lai, S.-S. Wang, Y. Tsao, and C.-C. Wu, "S1 and s2 heart sound recognition using deep neural networks," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 2, pp. 372–380, 2017.

[114] S. Lee and J.-H. Chang, "Deep learning ensemble with asymptotic techniques for oscillometric blood pressure estimation," *Computer methods and programs in biomedicine*, vol. 151, pp. 1–13, 2017.

[115] F. Pan, P. He, C. Liu, T. Li, A. Murray, and D. Zheng, "Variation of the korotkoff stethoscope sounds during blood pressure measurement:

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/RBME.2018.2885714, IEEE Reviews in Biomedical Engineering

24

Analysis using a convolutional neural network," *IEEE journal of biomedical and health informatics*, vol. 21, no. 6, pp. 1593–1598, 2017.

[116] S. P. Shashikumar, A. J. Shah, Q. Li, G. D. Clifford, and S. Nemati, "A deep learning approach to monitoring and detecting atrial fibrillation using wearable technology," in *Biomedical & Health Informatics (BHI), 2017 IEEE EMBS International Conference on*. IEEE, 2017, pp. 141–144.

[117] I. Gotlibovych, S. Crawford, D. Goyal, J. Liu, Y. Kerem, D. Benaron, D. Yilmaz, G. Marcus, and Y. Li, "End-to-end deep learning from raw sensor data: Atrial fibrillation detection using wearables," *arXiv preprint arXiv:1807.10707*, 2018.

[118] M.-Z. Poh, Y. C. Poh, P.-H. Chan, C.-K. Wong, L. Pun, W. W.-C. Leung, Y.-F. Wong, M. M.-Y. Wong, D. W.-S. Chu, and C.-W. Siu, "Diagnostic assessment of a deep learning system for detecting atrial fibrillation in pulse waveforms," *Heart*, pp. heartjnl–2018, 2018.

[119] B. Ballinger, J. Hsieh, A. Singh, N. Sohoni, J. Wang, G. H. Tison, G. M. Marcus, J. M. Sanchez, C. Maguire, J. E. Olgin *et al.*, "Deepheart: Semi-supervised sequence learning for cardiovascular risk prediction," *arXiv preprint arXiv:1802.02511*, 2018.

[120] A. Jiménez-Fernández, E. Cerezuela-Escudero, L. Miró-Amarante, M. J. Domínguez-Morales, F. Gomez-Rodriguez, A. Linares-Barranco, and G. Jiménez-Moreno, "A binaural neuromorphic auditory sensor for fpga: A spike signal processing approach." *IEEE Trans. Neural Netw. Learning Syst.*, vol. 28, no. 4, pp. 804–818, 2017.

[121] G. S. Everly Jr and J. M. Lating, *A clinical guide to the treatment of the human stress response*. Springer Science & Business Media, 2012.

[122] S. Lee and J. H. Chang, "Oscillometric blood pressure estimation based on deep learning," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 2, pp. 461–472, 2017.

[123] S. Lee and J.-H. Chang, "Deep belief networks ensemble for blood pressure estimation," *IEEE Access*, vol. 5, pp. 9962–9972, 2017.

[124] S. Lee and J. H. Chang, "Deep boltzmann regression with mimic features for oscillometric blood pressure estimation," *IEEE Sensors Journal*, vol. 17, no. 18, pp. 5982–5993, 2017.

[125] G. Sebastiani and P. Barone, "Mathematical principles of basic magnetic resonance imaging in medicine," *Signal Processing*, vol. 25, no. 2, pp. 227–250, 1991.

[126] L. K. Tan, Y. M. Liew, E. Lim, and R. A. McLaughlin, "Cardiac left ventricle segmentation using convolutional neural network regression," in *Biomedical Engineering and Sciences (IECBES), 2016 IEEE EMBS Conference on*. IEEE, 2016, pp. 490–493.

[127] L. V. Romaguera, M. G. F. Costa, F. P. Romero, and C. F. F. Costa Filho, "Left ventricle segmentation in cardiac mri images using fully convolutional neural networks," in *Medical Imaging 2017: Computer-Aided Diagnosis*, vol. 10134. International Society for Optics and Photonics, 2017, p. 101342Z.

[128] R. P. Poudel, P. Lamata, and G. Montana, "Recurrent fully convolutional neural networks for multi-slice mri cardiac segmentation," in *Reconstruction, Segmentation, and Analysis of Medical Images*. Springer, 2016, pp. 83–94.

[129] C. Rupprecht, E. Huaroc, M. Baust, and N. Navab, "Deep active contours," *arXiv preprint arXiv:1607.05074*, 2016.

[130] T. Anh Ngo and G. Carneiro, "Fully automated non-rigid segmentation with distance regularized level set evolution initialized and constrained by deep-structured inference," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2014, pp. 3118–3125.

[131] M. Avendi, A. Kheradvar, and H. Jafarkhani, "A combined deep-learning and deformable-model approach to fully automatic segmentation of the left ventricle in cardiac mri," *Medical image analysis*, vol. 30, pp. 108–119, 2016.

[132] H. Yang, J. Sun, H. Li, L. Wang, and Z. Xu, "Deep fusion net for multi-atlas segmentation: Application to cardiac mr images," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2016, pp. 521–528.

[133] G. Luo, S. Dong, K. Wang, and H. Zhang, "Cardiac left ventricular volumes prediction method based on atlas location and deep learning," in *Bioinformatics and Biomedicine (BIBM), 2016 IEEE International Conference on*. IEEE, 2016, pp. 1604–1610.

[134] X. Yang, Z. Zeng, and S. Yi, "Deep convolutional neural networks for automatic segmentation of left ventricle cavity from cardiac magnetic resonance images," *IET Computer Vision*, vol. 11, no. 8, pp. 643–649, 2017.

[135] L. K. Tan, Y. M. Liew, E. Lim, and R. A. McLaughlin, "Convolutional neural network regression for short-axis left ventricle segmentation in cardiac cine mr sequences," *Medical image analysis*, vol. 39, pp. 78–86, 2017.

[136] A. H. Curiale, F. D. Colavecchia, P. Kaluza, R. A. Isoardi, and G. Mato, "Automatic myocardial segmentation by using a deep learning network in cardiac mri," in *Computer Conference (CLEI), 2017 XLIII Latin American*. IEEE, 2017, pp. 1–6.

[137] F. Liao, X. Chen, X. Hu, and S. Song, "Estimation of the volume of the left ventricle from mri images using deep neural networks," *IEEE Transactions on Cybernetics*, 2017.

[138] O. Emad, I. A. Yassine, and A. S. Fahmy, "Automatic localization of the left ventricle in cardiac mri images using deep learning," in *Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE*. IEEE, 2015, pp. 683–686.

[139] C. Zotti, Z. Luo, O. Humbert, A. Lalande, and P.-M. Jodoin, "Gridnet with automatic shape prior registration for automatic mri cardiac segmentation," in *International Workshop on Statistical Atlases and Computational Models of the Heart*. Springer, 2017, pp. 73–81.

[140] J. Patravali, S. Jain, and S. Chilamkurthy, "2d-3d fully convolutional neural networks for cardiac mr segmentation," in *International Workshop on Statistical Atlases and Computational Models of the Heart*. Springer, 2017, pp. 130–139.

[141] F. Isensee, P. F. Jaeger, P. M. Full, I. Wolf, S. Engelhardt, and K. H. Maier-Hein, "Automatic cardiac disease assessment on cine-mri via time-series segmentation and domain specific features," in *International Workshop on Statistical Atlases and Computational Models of the Heart*. Springer, 2017, pp. 120–129.

[142] P. V. Tran, "A fully convolutional neural network for cardiac segmentation in short-axis mri," *arXiv preprint arXiv:1604.00494*, 2016.

[143] W. Bai, O. Oktay, M. Sinclair, H. Suzuki, M. Rajchl, G. Tarroni, B. Glocker, A. King, P. M. Matthews, and D. Rueckert, "Semi-supervised learning for network-based cardiac mr image segmentation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2017, pp. 253–260.

[144] J. Lieman-Sifry, M. Le, F. Lau, S. Sall, and D. Golden, "Fastventricle: Cardiac segmentation with enet," in *International Conference on Functional Imaging and Modeling of the Heart*. Springer, 2017, pp. 127–138.

[145] A. Paszke, A. Chaurasia, S. Kim, and E. Culurciello, "Enet: A deep neural network architecture for real-time semantic segmentation," *arXiv preprint arXiv:1606.02147*, 2016.

[146] H. B. Winther, C. Hundt, B. Schmidt, C. Czerner, J. Bauersachs, F. Wacker, and J. Vogel-Claussen, "nu-net: Deep learning for generalized biventricular cardiac mass and function parameters," *arXiv preprint arXiv:1706.04397*, 2017.

[147] X. Du, W. Zhang, H. Zhang, J. Chen, Y. Zhang, J. C. Warrington, G. Brahm, and S. Li, "Deep regression segmentation for cardiac biventricle mr images," *IEEE Access*, 2018.

[148] A. Giannakidis, K. Kamnitsas, V. Spadotto, J. Keegan, G. Smith, B. Glocker, D. Rueckert, S. Ernst, M. A. Gatzoulis, D. J. Pennell *et al.*, "Fast fully automatic segmentation of the severely abnormal human right ventricle from cardiovascular magnetic resonance images using a multi-scale 3d convolutional neural network," in *Signal-Image Technology & Internet-Based Systems (SITIS), 2016 12th International Conference on*. IEEE, 2016, pp. 42–46.

[149] J. M. Wolterink, T. Leiner, M. A. Viergever, and I. Išgum, "Dilated convolutional neural networks for cardiovascular mr segmentation in congenital heart disease," in *Reconstruction, Segmentation, and Analysis of Medical Images*. Springer, 2016, pp. 95–102.

[150] J. Li, R. Zhang, L. Shi, and D. Wang, "Automatic whole-heart segmentation in congenital heart disease using deeply-supervised 3d fcn," in *Reconstruction, Segmentation, and Analysis of Medical Images*. Springer, 2016, pp. 111–118.

[151] L. Yu, X. Yang, J. Qin, and P.-A. Heng, "3d fractalnet: dense volumetric segmentation for cardiovascular mri volumes," in *Reconstruction, Segmentation, and Analysis of Medical Images*. Springer, 2016, pp. 103–110.

[152] C. Payer, D. Štern, H. Bischof, and M. Urschler, "Multi-label whole heart segmentation using cnns and anatomical label configurations," in *International Workshop on Statistical Atlases and Computational Models of the Heart*. Springer, 2017, pp. 190–198.

[153] A. Mortazi, J. Burt, and U. Bagci, "Multi-planar deep segmentation networks for cardiac substructures from mri and ct," in *International Workshop on Statistical Atlases and Computational Models of the Heart*. Springer, 2017, pp. 199–206.

[154] X. Yang, C. Bian, L. Yu, D. Ni, and P.-A. Heng, "Hybrid loss guided convolutional networks for whole heart parsing," in *International Workshop on Statistical Atlases and Computational Models of the Heart*. Springer, 2017, pp. 215–223.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/RBME.2018.2885714, IEEE Reviews in Biomedical Engineering

25

[155] G. Yang, X. Zhuang, H. Khan, S. Haldar, E. Nyktari, X. Ye, G. Slabaugh, T. Wong, R. Mohiaddin, J. Keegan *et al.*, "Segmenting atrial fibrosis from late gadolinium-enhanced cardiac mri by deep-learned features with stacked sparse auto-encoders," in *Annual Conference on Medical Image Understanding and Analysis.* Springer, 2017, pp. 195–206.

[156] L. Zhang, A. Gooya, B. Dong, R. Hua, S. E. Petersen, P. Medrano-Gracia, and A. F. Frangi, "Automated quality assessment of cardiac mr images using convolutional neural networks," in *International Workshop on Simulation and Synthesis in Medical Imaging.* Springer, 2016, pp. 138–145.

[157] B. Kong, Y. Zhan, M. Shin, T. Denny, and S. Zhang, "Recognizing end-diastole and end-systole frames via deep temporal regression network," in *International conference on medical image computing and computer-assisted intervention.* Springer, 2016, pp. 264–272.

[158] F. Yang, Y. He, M. Hussain, H. Xie, and P. Lei, "Convolutional neural network for the detection of end-diastole and end-systole frames in free-breathing cardiac magnetic resonance imaging," *Computational and mathematical methods in medicine*, vol. 2017, 2017.

[159] C. Xu, L. Xu, Z. Gao, S. Zhao, H. Zhang, Y. Zhang, X. Du, S. Zhao, D. Ghista, and S. Li, "Direct detection of pixel-level myocardial infarction areas via a deep-learning algorithm," in *International Conference on Medical Image Computing and Computer-Assisted Intervention.* Springer, 2017, pp. 240–249.

[160] W. Xue, G. Brahm, S. Pandey, S. Leung, and S. Li, "Full left ventricle quantification via deep multitask relationships learning," *Medical image analysis*, vol. 43, pp. 54–65, 2018.

[161] X. Zhen, Z. Wang, A. Islam, M. Bhaduri, I. Chan, and S. Li, "Multi-scale deep networks and regression forests for direct bi-ventricular volume estimation," *Medical image analysis*, vol. 30, pp. 120–129, 2016.

[162] C. Biffi, O. Oktay, G. Tarroni, W. Bai, A. De Marvao, G. Doumou, M. Rajchl, R. Bedair, S. Prasad, S. Cook *et al.*, "Learning interpretable anatomical features through deep generative models: Application to cardiac remodeling," *arXiv preprint arXiv:1807.06843*, 2018.

[163] O. Oktay, W. Bai, M. Lee, R. Guerrero, K. Kamnitsas, J. Caballero, A. de Marvao, S. Cook, D. O'Regan, and D. Rueckert, "Multi-input cardiac image super-resolution using convolutional neural networks," in *International Conference on Medical Image Computing and Computer-Assisted Intervention.* Springer, 2016, pp. 246–254.

[164] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs," *IEEE transactions on pattern analysis and machine intelligence*, vol. 40, no. 4, pp. 834–848, 2018.

[165] C.-Y. Lee, S. Xie, P. Gallagher, Z. Zhang, and Z. Tu, "Deeply-supervised nets," in *Artificial Intelligence and Statistics*, 2015, pp. 562–570.

[166] R. Girshick, "Fast r-cnn," in *Proceedings of the IEEE international conference on computer vision*, 2015, pp. 1440–1448.

[167] O. Oktay, E. Ferrante, K. Kamnitsas, M. Heinrich, W. Bai, J. Caballero, S. A. Cook, A. de Marvao, T. Dawes, D. P. O'Regan *et al.*, "Anatomically constrained neural networks (acnns): Application to cardiac image enhancement and segmentation," *IEEE transactions on medical imaging*, vol. 37, no. 2, pp. 384–395, 2018.

[168] E. Konukoglu, "An exploration of 2d and 3d deep learning techniques for cardiac mr image segmentation," in *Statistical Atlases and Computational Models of the Heart. ACDC and MMWHS Challenges: 8th International Workshop, STACOM 2017, Held in Conjunction with MICCAI 2017, Quebec City, Canada, September 10-14, 2017, Revised Selected Papers*, vol. 10663. Springer, 2018, p. 111.

[169] M. D. Abràmoff, M. K. Garvin, and M. Sonka, "Retinal imaging and image analysis," *IEEE reviews in biomedical engineering*, vol. 3, pp. 169–208, 2010.

[170] S. Wang, Y. Yin, G. Cao, B. Wei, Y. Zheng, and G. Yang, "Hierarchical retinal blood vessel segmentation based on feature and ensemble learning," *Neurocomputing*, vol. 149, pp. 708–717, 2015.

[171] L. Zhou, Q. Yu, X. Xu, Y. Gu, and J. Yang, "Improving dense conditional random field for retinal vessel segmentation by discriminative feature learning and thin-vessel enhancement," *Computer methods and programs in biomedicine*, vol. 148, pp. 13–25, 2017.

[172] Y. Chen, "A labeling-free approach to supervising deep neural networks for retinal blood vessel segmentation," *arXiv preprint arXiv:1704.07502*, 2017.

[173] D. Maji, A. Santara, P. Mitra, and D. Sheet, "Ensemble of deep convolutional neural networks for learning to detect retinal vessels in fundus images," *arXiv preprint arXiv:1603.04833*, 2016.

[174] H. Fu, Y. Xu, D. W. K. Wong, and J. Liu, "Retinal vessel segmentation via deep learning network and fully-connected conditional random fields," in *Biomedical Imaging (ISBI), 2016 IEEE 13th International Symposium on.* IEEE, 2016, pp. 698–701.

[175] A. Wu, Z. Xu, M. Gao, M. Buty, and D. J. Mollura, "Deep vessel tracking: A generalized probabilistic approach via deep learning," in *Biomedical Imaging (ISBI), 2016 IEEE 13th International Symposium on.* IEEE, 2016, pp. 1363–1367.

[176] Q. Li, B. Feng, L. Xie, P. Liang, H. Zhang, and T. Wang, "A cross-modality learning approach for vessel segmentation in retinal images," *IEEE transactions on medical imaging*, vol. 35, no. 1, pp. 109–118, 2016.

[177] A. Lahiri, A. G. Roy, D. Sheet, and P. K. Biswas, "Deep neural ensemble for retinal vessel segmentation in fundus images towards achieving label-free angiography," in *Engineering in Medicine and Biology Society (EMBC), 2016 IEEE 38th Annual International Conference of the.* IEEE, 2016, pp. 1340–1343.

[178] A. Oliveira, S. Pereira, and C. A. Silva, "Augmenting data when training a cnn for retinal vessel segmentation: How to warp?" in *Bioengineering (ENBENG), 2017 IEEE 5th Portuguese Meeting on.* IEEE, 2017, pp. 1–4.

[179] H. A. Leopold, J. Orchard, J. Zelek, and V. Lakshminarayanan, "Use of gabor filters and deep networks in the segmentation of retinal vessel morphology," in *Imaging, Manipulation, and Analysis of Biomolecules, Cells, and Tissues XV*, vol. 10068. International Society for Optics and Photonics, 2017, p. 100680R.

[180] H. A. Leopold, J. Orchard, J. S. Zelek, and V. Lakshminarayanan, "Pixelbnn: Augmenting the pixelcnn with batch normalization and the presentation of a fast architecture for retinal vessel segmentation," *arXiv preprint arXiv:1712.06742*, 2017.

[181] J. Mo and L. Zhang, "Multi-level deep supervised networks for retinal vessel segmentation," *International journal of computer assisted radiology and surgery*, vol. 12, no. 12, pp. 2181–2193, 2017.

[182] M. Melinščak, P. Prentašić, and S. Lončarić, "Retinal vessel segmentation using deep neural networks," in *VISAPP 2015 (10th International Conference on Computer Vision Theory and Applications)*, 2015.

[183] A. Sengur, Y. Guo, Ü. Budak, and L. J. Vespa, "A retinal vessel detection approach using convolution neural network," in *Artificial Intelligence and Data Processing Symposium (IDAP), 2017 International.* IEEE, 2017, pp. 1–4.

[184] M. I. Meyer, P. Costa, A. Galdran, A. M. Mendonça, and A. Campilho, "A deep neural network for vessel segmentation of scanning laser ophthalmoscopy images," in *International Conference Image Analysis and Recognition.* Springer, 2017, pp. 507–515.

[185] M. Haloi, "Improved microaneurysm detection using deep neural networks," *arXiv preprint arXiv:1505.04424*, 2015.

[186] L. Giancardo, K. Roberts, and Z. Zhao, "Representation learning for retinal vasculature embeddings," in *Fetal, Infant and Ophthalmic Medical Image Analysis.* Springer, 2017, pp. 243–250.

[187] J. I. Orlando, E. Prokofyeva, M. del Fresno, and M. B. Blaschko, "An ensemble deep learning based approach for red lesion detection in fundus images," *Computer methods and programs in biomedicine*, vol. 153, pp. 115–127, 2018.

[188] M. J. van Grinsven, B. van Ginneken, C. B. Hoyng, T. Theelen, and C. I. Sánchez, "Fast convolutional neural network training using selective data sampling: Application to hemorrhage detection in color fundus images," *IEEE transactions on medical imaging*, vol. 35, no. 5, pp. 1273–1284, 2016.

[189] F. Girard and F. Cheriet, "Artery/vein classification in fundus images using cnn and likelihood score propagation," in *Signal and Information Processing (GlobalSIP), 2017 IEEE Global Conference on.* IEEE, 2017, pp. 720–724.

[190] R. Welikala, P. Foster, P. Whincup, A. Rudnicka, C. Owen, D. Strachan, S. Barman *et al.*, "Automated arteriole and venule classification using deep learning for retinal images from the uk biobank cohort," *Computers in biology and medicine*, vol. 90, pp. 23–32, 2017.

[191] H. Pratt, B. M. Williams, J. Y. Ku, C. Vas, E. McCann, B. Al-Bander, Y. Zhao, F. Coenen, and Y. Zheng, "Automatic detection and distinction of retinal vessel bifurcations and crossings in colour fundus photography," *Journal of Imaging*, vol. 4, no. 1, p. 4, 2017.

[192] R. Poplin, A. V. Varadarajan, K. Blumer, Y. Liu, M. V. McConnell, G. S. Corrado, L. Peng, and D. R. Webster, "Predicting cardiovascular risk factors from retinal fundus photographs using deep learning," *arXiv preprint arXiv:1708.09843*, 2017.

[193] H. Leopold, J. Orchard, J. Zelek, and V. Lakshminarayanan, "Segmentation and feature extraction of retinal vascular morphology," in

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/RBME.2018.2885714, IEEE Reviews in Biomedical Engineering

26

Medical Imaging 2017: Image Processing, vol. 10133. International Society for Optics and Photonics, 2017, p. 101330V.

[194] H. Pratt, B. M. Williams, J. Ku, F. Coenen, and Y. Zheng, "Automatic detection and identification of retinal vessel junctions in colour fundus photography," in *Annual Conference on Medical Image Understanding and Analysis*. Springer, 2017, pp. 27–37.

[195] N. Lessmann, I. Išgum, A. A. Setio, B. D. de Vos, F. Ciompi, P. A. de Jong, M. Oudkerk, P. T. M. Willem, M. A. Viergever, and B. van Ginneken, "Deep convolutional neural networks for automatic coronary calcium scoring in a screening study with low-dose chest ct," in *Medical Imaging 2016: Computer-Aided Diagnosis*, vol. 9785. International Society for Optics and Photonics, 2016, p. 978511.

[196] R. Shadmi, V. Mazo, O. Bregman-Amitai, and E. Elnekave, "Fully-convolutional deep-learning based system for coronary calcium score prediction from non-contrast chest ct," in *Biomedical Imaging (ISBI 2018), 2018 IEEE 15th International Symposium on*. IEEE, 2018, pp. 24–28.

[197] C. Cano-Espinosa, G. González, G. R. Washko, M. Cazorla, and R. S. J. Estépar, "Automated agatston score computation in non-ecg gated ct scans using deep learning," in *Proceedings of SPIE–the International Society for Optical Engineering*, vol. 10574. NIH Public Access, 2018.

[198] J. M. Wolterink, T. Leiner, B. D. de Vos, R. W. van Hamersvelt, M. A. Viergever, and I. Išgum, "Automatic coronary artery calcium scoring in cardiac ct angiography using paired convolutional neural networks," *Medical image analysis*, vol. 34, pp. 123–136, 2016.

[199] G. Santini, D. Della Latta, N. Martini, G. Valvano, A. Gori, A. Ripoli, C. L. Susini, L. Landini, and D. Chiappino, "An automatic deep learning approach for coronary artery calcium segmentation," in *EMBEC & NBC 2017*. Springer, 2017, pp. 374–377.

[200] K. López-Linares, L. Kabongo, N. Lete, G. Maclair, M. Ceresa, A. García-Familiar, I. Macía, and M. Á. G. Ballester, "Dcnn-based automatic segmentation and quantification of aortic thrombus volume: Influence of the training approach," in *Intravascular Imaging and Computer Assisted Stenting, and Large-Scale Annotation of Biomedical Data and Expert Label Synthesis*. Springer, 2017, pp. 29–38.

[201] H. A. Hong and U. Sheikh, "Automatic detection, segmentation and classification of abdominal aortic aneurysm using deep learning," in *Signal Processing & Its Applications (CSPA), 2016 IEEE 12th International Colloquium on*. IEEE, 2016, pp. 242–246.

[202] H. Liu, J. Feng, Z. Feng, J. Lu, and J. Zhou, "Left atrium segmentation in ct volumes with fully convolutional networks," in *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*. Springer, 2017, pp. 39–46.

[203] B. D. de Vos, J. M. Wolterink, P. A. de Jong, M. A. Viergever, and I. Išgum, "2d image classification for 3d anatomy localization: employing deep convolutional neural networks," in *Medical Imaging 2016: Image Processing*, vol. 9784. International Society for Optics and Photonics, 2016, p. 97841Y.

[204] M. Moradi, Y. Gur, H. Wang, P. Prasanna, and T. Syeda-Mahmood, "A hybrid learning approach for semantic labeling of cardiac ct slices and recognition of body position," in *Biomedical Imaging (ISBI), 2016 IEEE 13th International Symposium on*. IEEE, 2016, pp. 1418–1421.

[205] Y. Zheng, D. Liu, B. Georgescu, H. Nguyen, and D. Comaniciu, "3d deep learning for efficient and robust landmark detection in volumetric data," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2015, pp. 565–572.

[206] J. C. Montoya, Y. Li, C. Strother, and G.-H. Chen, "Deep learning angiography (dla): three-dimensional c-arm cone beam ct angiography generated from deep learning method using a convolutional neural network," in *Medical Imaging 2018: Physics of Medical Imaging*, vol. 10573. International Society for Optics and Photonics, 2018, p. 105731N.

[207] M. Zreik, N. Lessmann, R. W. van Hamersvelt, J. M. Wolterink, M. Voskuil, M. A. Viergever, T. Leiner, and I. Išgum, "Deep learning analysis of the myocardium in coronary ct angiography for identification of patients with functionally significant coronary artery stenosis," *Medical image analysis*, vol. 44, pp. 72–85, 2018.

[208] F. Commandeur, M. Goeller, J. Betancur, S. Cadet, M. Doris, X. Chen, D. S. Berman, P. J. Slomka, B. K. Tamarappoo, and D. Dey, "Deep learning for quantification of epicardial and thoracic adipose tissue from non-contrast ct," *IEEE Transactions on Medical Imaging*, 2018.

[209] M. A. Gülsün, G. Funka-Lea, P. Sharma, S. Rapaka, and Y. Zheng, "Coronary centerline extraction via optimal flow paths and cnn path pruning," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2016, pp. 317–325.

[210] G. Carneiro, J. C. Nascimento, and A. Freitas, "The segmentation of the left ventricle of the heart from ultrasound data using deep learning architectures and derivative-based search methods," *IEEE Transactions on Image Processing*, vol. 21, no. 3, pp. 968–982, 2012.

[211] J. C. Nascimento and G. Carneiro, "Multi-atlas segmentation using manifold learning with deep belief networks," in *Biomedical Imaging (ISBI), 2016 IEEE 13th International Symposium on*. IEEE, 2016, pp. 867–871.

[212] H. Chen, Y. Zheng, J.-H. Park, P.-A. Heng, and S. K. Zhou, "Iterative multi-domain regularized deep learning for anatomical structure detection and segmentation from ultrasound images," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2016, pp. 487–495.

[213] A. Madani, R. Arnaout, M. Mofrad, and R. Arnaout, "Fast and accurate view classification of echocardiograms using deep learning," *npj Digital Medicine*, vol. 1, no. 1, p. 6, 2018.

[214] J. F. Silva, J. M. Silva, A. Guerra, S. Matos, and C. Costa, "Ejection fraction classification in transthoracic echocardiography using a deep learning approach," in *2018 IEEE 31st International Symposium on Computer-Based Medical Systems (CBMS)*. IEEE, 2018, pp. 123–128.

[215] X. Gao, W. Li, M. Loomes, and L. Wang, "A fused deep learning architecture for viewpoint classification of echocardiography," *Information Fusion*, vol. 36, pp. 103–113, 2017.

[216] A. H. Abdi, C. Luong, T. Tsang, J. Jue, K. Gin, D. Yeung, D. Hawley, R. Rohling, and P. Abolmaesumi, "Quality assessment of echocardiographic cine using recurrent neural networks: Feasibility on five standard view planes," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2017, pp. 302–310.

[217] F. C. Ghesu, E. Krubasik, B. Georgescu, V. Singh, Y. Zheng, J. Hornegger, and D. Comaniciu, "Marginal space deep learning: efficient architecture for volumetric image parsing," *IEEE transactions on medical imaging*, vol. 35, no. 5, pp. 1217–1228, 2016.

[218] D. P. Perrin, A. Bueno, A. Rodriguez, G. R. Marx, and J. Pedro, "Application of convolutional artificial neural networks to echocardiograms for differentiating congenital heart diseases in a pediatric population," in *Medical Imaging 2017: Computer-Aided Diagnosis*, vol. 10134. International Society for Optics and Photonics, 2017, p. 1013431.

[219] M. Moradi, Y. Guo, Y. Gur, M. Negahdar, and T. Syeda-Mahmood, "A cross-modality neural network transform for semi-automatic medical image annotation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2016, pp. 300–307.

[220] A. G. Roy, S. Conjeti, S. G. Carlier, K. Houissa, A. König, P. K. Dutta, A. F. Laine, N. Navab, A. Katouzian, and D. Sheet, "Multiscale distribution preserving autoencoders for plaque detection in intravascular optical coherence tomography," in *Biomedical Imaging (ISBI), 2016 IEEE 13th International Symposium on*. IEEE, 2016, pp. 1359–1362.

[221] Y. L. Yong, L. K. Tan, R. A. McLaughlin, K. H. Chee, and Y. M. Liew, "Linear-regression convolutional neural network for fully automated coronary lumen segmentation in intravascular optical coherence tomography," *Journal of biomedical optics*, vol. 22, no. 12, p. 126005, 2017.

[222] M. Xu, J. Cheng, A. Li, J. A. Lee, D. W. K. Wong, A. Taruya, A. Tanaka, N. Foin, and P. Wong, "Fibroatheroma identification in intravascular optical coherence tomography images using deep features," in *Engineering in Medicine and Biology Society (EMBC), 2017 39th Annual International Conference of the IEEE*. IEEE, 2017, pp. 1501–1504.

[223] A. Abdolmanafi, L. Duong, N. Dahdah, and F. Cheriet, "Deep feature learning for automatic tissue classification of coronary artery using optical coherence tomography," *Biomedical optics express*, vol. 8, no. 2, pp. 1203–1220, 2017.

[224] K. Lekadir, A. Galimzianova, À. Betriu, M. del Mar Vila, L. Igual, D. L. Rubin, E. Fernández, P. Radeva, and S. Napel, "A convolutional neural network for automatic characterization of plaque composition in carotid ultrasound," *IEEE journal of biomedical and health informatics*, vol. 21, no. 1, pp. 48–55, 2017.

[225] N. Tajbakhsh, J. Y. Shin, R. T. Hurst, C. B. Kendall, and J. Liang, "Automatic interpretation of carotid intima–media thickness videos using convolutional neural networks," in *Deep Learning for Medical Image Analysis*. Elsevier, 2017, pp. 105–131.

[226] F. Tom and D. Sheet, "Simulating patho-realistic ultrasound images using deep generative networks with adversarial learning," in *Biomedical Imaging (ISBI 2018), 2018 IEEE 15th International Symposium on*. IEEE, 2018, pp. 1174–1177.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/RBME.2018.2885714, IEEE Reviews in Biomedical Engineering

27

[227] J. Wang, H. Ding, F. A. Bidgoli, B. Zhou, C. Iribarren, S. Molloi, and P. Baldi, "Detecting cardiovascular disease from mammograms with deep learning," *IEEE transactions on medical imaging*, vol. 36, no. 5, pp. 1172–1181, 2017.

[228] X. Liu, S. Wang, Y. Deng, and K. Chen, "Coronary artery calcification (cac) classification with deep convolutional neural networks," in *Medical Imaging 2017: Computer-Aided Diagnosis*, vol. 10134. International Society for Optics and Photonics, 2017, p. 101340M.

[229] M. Pavoni, Y. Chang, and Ö. Smedby, "Image denoising with convolutional neural networks for percutaneous transluminal coronary angioplasty," in *European Congress on Computational Methods in Applied Sciences and Engineering*. Springer, 2017, pp. 255–265.

[230] J. J. Nirschl, A. Janowczyk, E. G. Peyster, R. Frank, K. B. Margulies, M. D. Feldman, and A. Madabhushi, "A deep-learning classifier identifies patients with clinical heart failure using whole-slide images of h&e tissue," *PloS one*, vol. 13, no. 4, p. e0192726, 2018.

[231] J. Betancur, F. Commandeur, M. Motlagh, T. Sharir, A. J. Einstein, S. Bokhari, M. B. Fish, T. D. Ruddy, P. Kaufmann, A. J. Sinusas *et al.*, "Deep learning for prediction of obstructive disease from fast myocardial perfusion spect: a multicenter study," *JACC: Cardiovascular Imaging*, 2018.

[232] N. Lessmann, B. van Ginneken, M. Zreik, P. A. de Jong, B. D. de Vos, M. A. Viergever, and I. Išgum, "Automatic calcium scoring in low-dose chest ct using deep neural networks with dilated convolutions," *IEEE Transactions on Medical Imaging*, 2017.

[233] M. Zreik, T. Leiner, B. D. de Vos, R. W. van Hamersvelt, M. A. Viergever, and I. Išgum, "Automatic segmentation of the left ventricle in cardiac ct angiography using convolutional neural networks," in *Biomedical Imaging (ISBI), 2016 IEEE 13th International Symposium on*. IEEE, 2016, pp. 40–43.

[234] Q. Le and T. Mikolov, "Distributed representations of sentences and documents," in *International Conference on Machine Learning*, 2014, pp. 1188–1196.

[235] T. Kubo, T. Akasaka, J. Shite, T. Suzuki, S. Uemura, B. Yu, K. Kozuma, H. Kitabata, T. Shinke, M. Habara *et al.*, "Oct compared with ivus in a coronary lesion assessment: the opus-class study," *JACC: Cardiovascular Imaging*, vol. 6, no. 10, pp. 1095–1104, 2013.

[236] J. E. Parrillo and R. P. Dellinger, *Critical Care Medicine E-Book: Principles of Diagnosis and Management in the Adult*. Elsevier Health Sciences, 2013.

[237] V. Mayer-Schönberger, "Big data for cardiology: novel discovery?" *European heart journal*, vol. 37, no. 12, pp. 996–1001, 2015.

[238] C. Austin and F. Kusumoto, "The application of big data in medicine: current implications and future directions," *Journal of Interventional Cardiac Electrophysiology*, vol. 47, no. 1, pp. 51–59, 2016.

[239] H. Greenspan, B. van Ginneken, and R. M. Summers, "Guest editorial deep learning in medical imaging: Overview and future promise of an exciting new technique," *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1153–1159, 2016.

[240] R. Miotto, F. Wang, S. Wang, X. Jiang, and J. T. Dudley, "Deep learning for healthcare: review, opportunities and challenges," *Briefings in bioinformatics*, 2017.

[241] C. Krittanawong, "The rise of artificial intelligence and the uncertain future for physicians," 2017.

[242] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. van der Laak, B. van Ginneken, and C. I. Sánchez, "A survey on deep learning in medical image analysis," *Medical image analysis*, vol. 42, pp. 60–88, 2017.

[243] A. Qayyum, S. M. Anwar, M. Majid, M. Awais, and M. Alnowami, "Medical image analysis using convolutional neural networks: A review," *arXiv preprint arXiv:1709.02250*, 2017.

[244] M. Henglin, G. Stein, P. V. Hushcha, J. Snoek, A. B. Wiltschko, and S. Cheng, "Machine learning approaches in cardiovascular imaging," *Circulation: Cardiovascular Imaging*, vol. 10, no. 10, p. e005614, 2017.

[245] G. W. Blair, M. V. Hernandez, M. J. Thrippleton, F. N. Doubal, and J. M. Wardlaw, "Advanced neuroimaging of cerebral small vessel disease," *Current treatment options in cardiovascular medicine*, vol. 19, no. 7, p. 56, 2017.

[246] P. J. Slomka, D. Dey, A. Sitek, M. Motwani, D. S. Berman, and G. Germano, "Cardiac imaging: working towards fully-automated machine analysis & interpretation," *Expert review of medical devices*, vol. 14, no. 3, pp. 197–212, 2017.

[247] G. Carneiro, Y. Zheng, F. Xing, and L. Yang, "Review of deep learning methods in mammography, cardiovascular, and microscopy image analysis," in *Deep Learning and Convolutional Neural Networks for Medical Image Computing*. Springer, 2017, pp. 11–32.

[248] K. W. Johnson, J. T. Soto, B. S. Glicksberg, K. Shameer, R. Miotto, M. Ali, E. Ashley, and J. T. Dudley, "Artificial intelligence in cardiology," *Journal of the American College of Cardiology*, vol. 71, no. 23, pp. 2668–2679, 2018.

[249] F. Jiang, Y. Jiang, H. Zhi, Y. Dong, H. Li, S. Ma, Y. Wang, Q. Dong, H. Shen, and Y. Wang, "Artificial intelligence in healthcare: past, present and future," *Stroke and Vascular Neurology*, pp. svn–2017, 2017.

[250] E.-J. Lee, Y.-H. Kim, N. Kim, and D.-W. Kang, "Deep into the brain: Artificial intelligence in stroke imaging," *Journal of stroke*, vol. 19, no. 3, p. 277, 2017.

[251] B. C. Loh and P. H. Then, "Deep learning for cardiac computer-aided diagnosis: benefits, issues & solutions," *mHealth*, vol. 3, 2017.

[252] C. Krittanawong, H. Zhang, Z. Wang, M. Aydar, and T. Kitai, "Artificial intelligence in precision cardiovascular medicine," *Journal of the American College of Cardiology*, vol. 69, no. 21, pp. 2657–2664, 2017.

[253] J. Gomez, R. Doukky, G. Germano, and P. Slomka, "New trends in quantitative nuclear cardiology methods," *Current Cardiovascular Imaging Reports*, vol. 11, no. 1, p. 1, 2018.

[254] K. Shameer, K. W. Johnson, B. S. Glicksberg, J. T. Dudley, and P. P. Sengupta, "Machine learning in cardiovascular medicine: are we there yet?" *Heart*, pp. heartjnl–2017, 2018.

[255] S. Shrestha and P. P. Sengupta, "Machine learning for nuclear cardiology: The way forward," 2018.

[256] A. Kikuchi and T. Kawakami, "Future of artificial intelligence and nuclear cardiology," *Annals of Nuclear Cardiology*, vol. 4, no. 1, pp. 79–82, 2018.

[257] S. E. Awan, F. Sohel, F. M. Sanfilippo, M. Bennamoun, and G. Dwivedi, "Machine learning in heart failure: ready for prime time," *Current opinion in cardiology*, vol. 33, no. 2, pp. 190–195, 2018.

[258] O. Faust, Y. Hagiwara, T. J. Hong, O. S. Lih, and U. R. Acharya, "Deep learning for healthcare applications based on physiological signals: a review," *Computer methods and programs in biomedicine*, 2018.

[259] D. S. Liebeskind, "Artificial intelligence in stroke care: Deep learning or superficial insight?" *EBioMedicine*, 2018.

[260] G. Hinton, "Deep learning—a technology with the potential to transform health care," *JAMA*, 2018.

[261] A. L. Beam and I. S. Kohane, "Big data and machine learning in health care," *Jama*, vol. 319, no. 13, pp. 1317–1318, 2018.

[262] J. A. Damen, L. Hooft, E. Schuit, T. P. Debray, G. S. Collins, I. Tzoulaki, C. M. Lassale, G. C. Siontis, V. Chiocchia, C. Roberts *et al.*, "Prediction models for cardiovascular disease risk in the general population: systematic review," *bmj*, vol. 353, p. i2416, 2016.

[263] D. Bahdanau, K. Cho, and Y. Bengio, "Neural machine translation by jointly learning to align and translate," *arXiv preprint arXiv:1409.0473*, 2014.

[264] K. Simonyan, A. Vedaldi, and A. Zisserman, "Deep inside convolutional networks: Visualising image classification models and saliency maps," *arXiv preprint arXiv:1312.6034*, 2013.

[265] S. Sabour, N. Frosst, and G. E. Hinton, "Dynamic routing between capsules," in *Advances in Neural Information Processing Systems*, 2017, pp. 3856–3866.

[266] P. Afshar, A. Mohammadi, and K. N. Plataniotis, "Brain tumor type classification via capsule networks," *arXiv preprint arXiv:1802.10200*, 2018.

[267] T. Iesmantas and R. Alzbutas, "Convolutional capsule network for classification of breast cancer histology images," in *International Conference Image Analysis and Recognition*. Springer, 2018, pp. 853–860.

[268] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Advances in neural information processing systems*, 2014, pp. 2672–2680.

[269] F. Mahmood and N. J. Durr, "Deep learning and conditional random fields-based depth estimation and topographical reconstruction from conventional endoscopy," *Medical Image Analysis*, 2018.

[270] P. F. Christ, M. E. A. Elshaer, F. Ettlinger, S. Tatavarty, M. Bickel, P. Bilic, M. Rempfler, M. Armbruster, F. Hofmann, M. D'Anastasi *et al.*, "Automatic liver and lesion segmentation in ct using cascaded fully convolutional neural networks and 3d conditional random fields," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2016, pp. 415–423.

[271] J. Ngiam, A. Khosla, M. Kim, J. Nam, H. Lee, and A. Y. Ng, "Multimodal deep learning," in *Proceedings of the 28th international conference on machine learning (ICML-11)*, 2011, pp. 689–696.