HHS Public Access

Author manuscript

Proc IEEE Inst Electr Electron Eng. Author manuscript; available in PMC 2023 October 02.

Published in final edited form as:

Proc IEEE Inst Electr Electron Eng. 2021 May ; 109(5): 820–838. doi:10.1109/JPROC.2021.3054390.

A review of deep learning in medical imaging: Imaging traits, technology trends, case studies with progress highlights, and future promises

S. Kevin Zhou[#],
School of Biomedical Engineering, University of Science and Technology of China and Institute of Computing Technology, Chinese Academy of Sciences.

Hayit Greenspan[#],
Biomedical Engineering Department, Tel-Aviv University, Israel.

Christos Davatzikos,
Radiology Department and Electrical and Systems Engineering Department, University of Pennsylvania, USA.

James S. Duncan,
Departments of Biomedical Engineering and Radiology & Biomedical Imaging, Yale University.

Bram van Ginneken,
Radboud University Medical Center, the Netherlands.

Anant Madabhushi,
Department of Biomedical Engineering, Case Western Reserve University and Louis Stokes Cleveland Veterans Administration Medical Center, USA.

Jerry L. Prince,
Electrical and Computer Engineering Department, Johns Hopkins University, USA.

Daniel Rueckert,
Klinikum rechts der Isar, TU Munich, Germany and Department of Computing, Imperial College, UK.

Ronald M. Summers
National Institute of Health Clinical Center, USA.

[#] These authors contributed equally to this work.

Abstract

Since its renaissance, deep learning has been widely used in various medical imaging tasks and has achieved remarkable success in many medical imaging applications, thereby propelling us into the so-called artificial intelligence (AI) era. It is known that the success of AI is mostly attributed to the availability of big data with annotations for a single task and the advances in high performance computing. However, medical imaging presents unique challenges that confront deep

Zhou and Greenspan are corresponding authors.

[6] https://www.acrdsi.org/DSI-Services/FDA-Cleared-AI-Algorithms

learning approaches. In this survey paper, we first present traits of medical imaging, highlight both clinical needs and technical challenges in medical imaging, and describe how emerging trends in deep learning are addressing these issues. We cover the topics of network architecture, sparse and noisy labels, federating learning, interpretability, uncertainty quantification, etc. Then, we present several case studies that are commonly found in clinical practice, including digital pathology and chest, brain, cardiovascular, and abdominal imaging. Rather than presenting an exhaustive literature survey, we instead describe some prominent research highlights related to these case study applications. We conclude with a discussion and presentation of promising future directions.

Keywords

Medical imaging; deep learning; survey

I. OVERVIEW

Medical imaging [1] exploits physical phenomena such as light, electromagnetic radiation, radioactivity, nuclear magnetic resonance, and sound to generate visual representations or images of external or internal tissues of the human body or a part of the human body in a non-invasive manner or via an invasive procedure. The most commonly used imaging modalities in clinical medicine include X-ray radiography, computed tomography (CT), magnetic resonance imaging (MRI), ultrasound, and digital pathology. Imaging data account for about 90% of all healthcare data[1] and hence is one of the most important sources of evidence for clinical analysis and medical intervention.

A. Traits of medical imaging

As described below and illustrated in Figure 1, medical imaging has several traits that influence the suitability and nature of deep learning solutions. Note that these traits are not necessarily unique to medical imaging. For example, satellite imaging shares the first trait described below with medical imaging.

Medical images have multiple modalities and are dense in pixel resolution.— There are many existing imaging modalities and new modalities such as spectral CT are being routinely invented. Even for commonly used imaging modalities, the pixel or voxel resolution has become higher and the information density has increased. For example, the spatial resolution of clinical CT and MRI has reached the sub-millimeter level and the spatial resolution of ultrasound is even better while its temporal resolution exceeds real-time.

Medical image data are isolated and acquired in non-standard settings.— Although medical imaging data exist in large numbers in the clinic, due to a lack of standardized acquisition protocols there is a large variation in terms of equipment and scanning settings, leading to the so-called “distribution drift” phenomenon. Due to patient privacy and clinical data management requirements, images are scattered among different hospitals and imaging centers, and truly centralized open source medical big data are rare.

[1]“The Digital Universe Driving Data Growth in Healthcare,” published by EMC with research and analysis from IDC (12/13).

The disease patterns in medical images are numerous and their incidence exhibits a long tail distribution.—Radiology Gamuts Ontology [2] defines 12,878 "symptoms" (conditions that lead to results) and 4,662 "diseases" (imaging findings). The incidence of disease has a typical long-tailed distribution: while a small number of common diseases have sufficient observed cases for large-scale analysis, most diseases are infrequent in the clinic. In addition, novel contagious diseases that are not represented in the current ontology, such as the outbreak of COVID-19, occur with some frequency.

The labels associated with medical images are sparse and noisy.—Labeling or annotating a medical image is time-consuming and expensive. Also, different tasks require different forms of annotation which creates the phenomenon of label sparsity. Because of variable experience and different conditions, both inter-user and intra-user labeling inconsistency is high [3] and labels must therefore be considered to be noisy. In fact, the establishment of gold standards for image labeling remains an open issue.

Samples are heterogeneous and imbalanced.—In the already labeled images, the appearance varies from sample to sample, with its probability distribution being multi-modal. The ratio between positive and negative samples is extremely uneven. For example, the number of pixels belonging to a tumor is usually one to many orders of magnitude less than that of normal tissue.

Medical image processing and analysis tasks are complex and diverse.—Medical imaging has a rich body of tasks. At the technical level, there is an array of technologies including reconstruction, enhancement, restoration, classification, detection, segmentation, and registration. When these technologies are combined with multiple image modalities and numerous disease types, a very large number of highly-complex tasks associated with numerous applications are formed and should be addressed.

B. Clinical needs and applications

Medical imaging is often a key part of the medical diagnosis and treatment process. Typically, a radiologist reviews the acquired medical images and write a summarizing report of their findings. The referring physician defines a diagnosis and treatment plan based on the images and radiologist's report. Often, medical imaging is ordered as part of a patient's follow-up to verify successful treatment. In addition, images are becoming an important component of invasive procedures, being used both for surgical planning as well as for real-time imaging during the procedure itself.

As a specific example, we can look at what we term the "radiology challenge" [4], [5]. In the past decade, with the development of technologies related to the image acquisition process, imaging devices have improved in speed and resolution. For example, in 1990 a CT scanner might acquire 50–100 slices whereas today's CT scanners might acquire 1000–2500 slices per case. A single whole slide digital pathology image corresponding to a single prostate biopsy core can easily occupy 10GB of space at 40x magnification. Overall, there are billions of medical imaging studies conducted per year, worldwide, and this number is growing.

Most interpretations of medical images are performed by physicians and, in particular, by radiologists. Image interpretation by humans, however, is limited due to human subjectivity, the large variations across interpreters, and fatigue. Radiologists that review cases have limited time to review an ever-increasing number of images, which leads to missed findings, long turn-around times, and a paucity of numerical results or quantification. This, in turn, drastically limits the medical community's ability to advance towards more evidence-based personalized healthcare.

AI tools such as deep learning technology can provide support to physicians by automating image analysis, leading to what we can term "Computational Radiology" [6], [7]. Among the automated tools that can be developed are *detection* of pathological findings, *quantification* of disease extent, *characterization* of pathologies (e.g., into benign vs malignant), and assorted software tools that can be broadly characterized as *decision support*. This technology can also extend physicians' capabilities to include the characterization of three-dimensional and time-varying events, which are often not included in today's radiological reports because of both limited time and limited visualization and quantification tools.

C. Key technologies and deep learning

Several key technologies arise from the various medical imaging applications, including: [8], [9], [10], [11]

- *Medical image reconstruction* [12], which aims to form a visual representation (aka an image) from signals acquired by a medical imaging device such as a CT or MRI scanner. Reconstruction of high quality images from low doses and/or fast acquisitions has important clinical implications.
- *Medical image enhancement,* which aims to adjust the intensities of an image so that the resultant image is more suitable for display or further analysis. Enhancement methods include denoising, super-resolution, MR bias field correction [13], and image harmonization [14]. Recently, much research has focused on modality translation and synthesis, which can be considered as image enhancement steps.
- *Medical image segmentation* [15], which aims to assign labels to pixels so that the pixels with the same label form a segmented object. Segmentation has numerous applications in clinical quantification, therapy, and surgical planning.
- *Medical image registration [16],* which aims to align the spatial coordinates of one or more images into a common coordinate system. Registration finds wide use in population analysis, longitudinal analysis, and multimodal fusion, and is also commonly used for image segmentation via label transfer.
- *Computer aided detection (CADe) and diagnosis (CADx)* [17]. CADe aims to localize or find a bounding box that contains an object (typically a lesion) of interest. CADx aims to further classify the localized lesion as benign/malignant or one of multiple lesion types.

- *Others technologies* include landmark detection [18], image or view recognition [19], automatic report generation [20], etc.

In mathematics, the above technologies can be regarded as function approximation methods, which approximate the true mapping F that takes an image (or multiple images if multimodality is accessible) as input and outputs a specific $y = F(x)$. The definition of y varies depending on the technology, which itself depends on the application or task. In CADe, y denotes a bounding box. In image registration, y is a deformation field. In image segmentation, y is a label mask. In image enhancement, y is a quality-enhanced image typically of the same size[2] of the input image x.

There are many ways to approximate F; however, deep learning (DL) [21], a branch of machine learning (ML), is one of the most powerful methods for function approximation. Since its renaissance, deep learning has been widely used in various medical imaging tasks and has achieved substantial success in many medical imaging applications. Because of its focus on learning rather than modeling, the use of DL in medical imaging represents a substantial departure from previous approaches in medical imaging. Take supervised deep learning as an example. Assume that a training dataset $\{(x_n, y_n); n = 1, \ldots, N\}$ is available and that a deep neural network is parameterized by θ, which includes the number of layers, the number of nodes of each layer, the connecting weights, the choices of activation functions, etc. The neural network that is found to approximate F can be written as $\phi_{\hat{\theta}}(x)$ where $\hat{\theta}$ are the parameters that minimize the so-called loss function

$$L(\theta) = \frac{1}{N}\sum_{n=1}^{N} l(\phi_\theta(x_n), y_n) + R_1(\phi_\theta(x_n)) + R_2(\theta)\,. \tag{1}$$

Here, $l(\phi_\theta(x), y)$ is the item-wise loss function that penalizes the prediction error, $R_1(\phi_\theta(x_n))$ reflects the prior belief about the output, and $R_2(\theta)$ is a regularization term about the network parameters. Although the neural network $\phi_{\hat{\theta}}(x)$ does represent a type of model, it is generally thought of as a “black box” since it does not represent a designed model based on well-understood physical or mathematical principles.

There are many survey papers on deep learning based key technologies for medical image analysis [22], [23], [24], [25], [26], [27], [28], [29], [30], [31]. To differentiate the present review paper from these works, we specifically omit any presentation of the technical details of DL itself, which is no longer considered new and is well-covered in numerous other works, and focus instead on the connections between the emerging DL approaches and the specific needs in medical imaging and on several case examples that illustrate the state of the art.

D. Historical perspective

Here, we briefly outline the development timeline of DL in medical imaging. Deep learning was termed one of the 10 breakthrough technologies of 2013 [32]. This followed the

[2]Image super-resolution generates an output image that has a different size from the input image.

2012 large-scale image categorization challenge that introduced the CNN superiority on the ImageNet dataset [33]. At that time, DL emerged as the leading machine-learning tool in the general imaging and computer vision domains and the medical imaging community began a debate about whether DL would be applicable in the medical imaging space. The concerns were due to the challenges we have outlined above, with the main challenge being the lack of sufficient labeled data, known as the *data challenge*.

Several steps can be pointed to as *enablers* of the DL technology within the medical imaging space: In 2015–2016, techniques were developed using "transfer learning" (TL) (or what was also called "learning from non-medical features" [34]) to apply the knowledge gained via solving a source problem to a different but related target problem. A key question was whether a network pre-trained on natural imagery would be applicable to medical images. Several groups showed this to be the case (e.g. [35], [36], [34]); using the deep network trained based on ImageNet and fine-tuning to a medical imaging task was helpful in order to speed up training convergence and improve accuracy.

In 2017–2018, *synthetic data augmentation* emerged as a second solution to process limited datasets. Classical augmentation is a key component of any network training. Still, key questions to address were whether it was possible to synthesize medical data using schemes such as generative modeling and whether the synthesized data would serve as viable medical examples and would in practice increase performance of the medical task at hand. Several works across varying domains demonstrated that this was in fact the case. In [37], for example, synthetic image augmentation based on generative adversarial network (GAN) was shown to generate lesion image samples that were not recognized as synthetic by the expert radiologists and also increased CNN performance in classifying liver lesions. GANs, variational encoders, and variations on these are still being explored and advanced in recent works, as will be described in the following Section.

For image segmentation, one of the key contributions that emerged from the medical imaging community was the U-Net architecture [38]. Originally designed for microscopic cell segmentation, the U-Net has proven to efficiently and robustly learn effective features for many medical image segmentation tasks.

E. Emerging deep learning approaches

Network architectures.—Deep neural networks have a larger model capacity and stronger generalization capability than shallow neural networks. Deep models trained on large scale annotated databases for a single task achieve outstanding performances, far beyond traditional algorithms or even human capability.

Making it deeper.: Starting from AlexNet [33], there was a research trend to make networks deeper, as represented by VGGNet [39], Inception Net [40], and ResNet [41]. The use of skip connections makes a deep network more trainable as in DenseNet [42] and U-Net [38]. U-net was first proposed to tackle segmentation while the other networks were developed for image classification. Deep supervision [43] further improves discriminative power.

Adversarial and attention mechanisms.: In the generative adversarial network (GAN) [44], Goodfellow et al. propose to accompany a generative model with a discriminator that tells whether a sample is from the model distribution or the data distribution. Both the generator and discriminator are represented by deep networks and their training is done via a minimax optimization. Adversarial learning is widely used in medical imaging [27], including medical image reconstruction [12], image quality enhancement [14], and segmentation [45].

Attention mechanism [46] allows automatic discovery of "where" and "what" to focus on when describing image contents or making a holistic decision. Squeeze and excitation [47] can be regarded as a channel attention mechanism. Attention is combined with GAN in [48] and with U-Net in [49].

Neural architecture search (NAS) and light weight design.: NAS [50] aims to automatically design the architecture of a deep network for high performance geared toward a given task. Zhu et al. [51] successfully apply NAS to volumetric medical image segmentation. Light weight design [52], [53], on the other hand, aims to design the architecture for computational efficiency on resource-constrained devices such as mobile phones while maintaining accuracy.

Annotation efficient approaches.—To address sparse and noisy labels, we need DL approaches that are efficient with respect to annotations. So, a key idea is to leverage the power and robustness of feature representation capability derived from existing models and data even though the models or data are not necessarily from the same domain or for the same task and to adapt such representation to the task at hand. To do this, there are a handful of methods proposed in the literature [28], including transfer learning, domain adaptation, self-supervised learning, semi-supervised learning, weakly/partially supervised learning, etc.

Transfer learning (TL) aims to apply the knowledge gained via solving a source problem to a different but related target problem. One commonly used TL method is to use the deep network trained based on ImageNet and fine tune it to a medical imaging task in order to speed up training convergence and improve accuracy. With the availability of a large number of annotated datasets, such TL methods [36] achieve remarkable success. However, ImageNet consists of natural images and its pretrained models are for 2D images only and are not necessarily the best for medical images, especially for small-sample settings [54]. Liu et al. [55] propose a 3D anisotropic hybrid network that effectively transfers convolutional features learned from 2D images to 3D anisotropic volumes. In [56], Chen et al. combine multiple datasets from several medical challenges with diverse modalities, target organs, and pathologies and learn one 3D network that provides an effective pretrained model for 3D medical image analysis tasks.

Domain adaptation is a form of transfer learning in which the source and target domains have the same feature space but different distributions. In [57], domain-invariant features are learned via an adversarial mechanism that attempts to classify the domain of the input data. Zhang et al. [58] propose to synthesize and segment multimodal medical volumes using generative adversarial networks with cycle- and shape-consistency. In [59], a domain

adaptation module that maps the target input to features which are aligned with source domain feature space is proposed for cross-modality biomedical image segmentation, using a domain critic module for discriminating the feature space of both domains. Huang et al. [60] propose a universal U-Net comprising domain-general and domain-specific parameters to deal with multiple organ segmentation tasks on multiple domains. This *integrated learning* mechanism offers a new possibility of dealing with multiple domains and even multiple heterogeneous tasks.

Self-supervised learning, a form of unsupervised learning, learns a representation through a proxy task, in which the data provides supervisory signals. Once the representation is learned, it is fine tuned by using annotated data. The models genesis method [61] uses a proxy task of recovering the original image using a distorted image as input. The possible distortions include non-linear gray-value transformation, local pixel shuffling, and image out-painting and in-painting. In [62], Zhu et al. propose solving a Rubik's Cube proxy task that involves three operations, namely cube ordering, cube rotating, and cube masking. This allows the network to learn features that are invariant to translation and rotation and robust to noise as well.

Semi-supervised learning often trains a model using a small set of annotated images, then generates pseudo-labels for a large set of images with annotations, and learns a final model by mixing up both sets of images. Bai et al. [63] implement such a method for cardiac MR segmentation. In [64], Nie et al. propose an attention based semi-supervised deep network for segmentation. It adversarially trains a segmentation network, from which a confidence map is computed as a region-attention based semi-supervised learning strategy to include the unlabeled data for training.

Weakly or partially supervised learning.: In [65], Wang et al. solve a weakly-supervised multi-label disease classification from a chest x-ray. To relax the stringent pixel-wise annotation for image segmentation, weakly supervised methods that use image-level annotations [66], or weak annotations like dots and scribbles [67] are proposed. For multi-organ segmentation, Shi et al. [68] learn a single multiclass network from a union of multiple datasets, each with a low sample size and partial organ label, using newly proposed marginal loss and exclusion loss. Schleg et al. [69] build a deep model from only normal images to detect abnormal regions in a test image.

Unsupervised learning and disentanglement.: Unsupervised learning does not rely on the existence of annotated images. A disentangled network structure is designed with an adversarial learning strategy that promotes the statistical matching of deep features has been widely used. In medical imaging, unsupervised learning and disentanglement have been used in image registration [70], motion tracking [71], artifact reduction [72], improving classification [73], domain adaptation [74], and general modeling [75].

Embedding knowledge into learning.—Knowledge arises from various sources such as imaging physics, statistical constraints, and task specifics and ways of embedding into a DL approach vary too. For chest x-ray disease classification, Li et al. [76], [77] encode anatomy knowledge embedded in unpaired CT into a deep network that decomposes a

chest x-ray into lung, bone and the remaining structures (see Fig. 2). With augmented bone-suppressed images, classification performance is improved in predicting 11 out of 14 common lung diseases. In [78] lung radiographs are enhanced by learning to extract lung structures from CT based simulated x-ray (DRRs) and fusing with the original x-ray image. The enhancement is shown to augment results of pathology characterization in real x-ray images. In [79], a dual-domain network is proposed to reduce metal artifacts on both the image and sinogram domains, which are seemingly integrated into one differential framework through a Radon inverse layer, rather than two separate modules.

Federated learning.—To combat issues related to data privacy, data security, and data access rights, it has become increasingly important to have the capability of learning a common, robust algorithmic model through distributed computing and model aggregation strategies so that no data are transferred outside a hospital or an imaging lab. This research direction is called federated learning (FL) [80], which is in contrast to conventional centralized learning with all the local datasets uploaded to one server. There are many ongoing research challenges related to federate learning such as reduced communication burden [81], data heterogeneity in various local sites [82], and vulnerability to attacks [83].

Despite its importance, work on FL in medical imaging has only been reported recently. Sheller et al. [84] present the first use of FL for multi-institutional DL model without sharing patient data and report similar brain lesion segmentation performances between the models trained in a federated or centralized way. In [85], Li et al. study several practical FL methods while protecting data privacy for brain tumour segmentation on the BraTS dataset and demonstrate a trade-off between model performance and privacy protection costs. Recently, FL is applied, together with domain adaptation, to train a model with boosted analysis performance and reliable discovery of disease-related biomarkers [86].

Interpretability.: Clinical decision-making relies heavily on evidence gathering and interpretation. Lacking evidence and interpretation makes it difficult for physicians to trust the ML model's prediction, especially in disease diagnosis. In addition, interpretability is also the source of new knowledge. Murdoch et al. [87] define interpretable machine learning as leveraging machine-learning models to extract relevant knowledge about domain relationships contained in data, aiming to provide insights for a user into a chosen domain problem. Most interpretation methods are categorized as model-based and post-hoc interpretability. The former is about constraining the model so that it readily provides useful information (such as sparsity, modularity, etc.) about the uncovered relationships. The latter is about extracting information about what relationships the model has learned.

Model-based interpretability.: For cardiac MRI classification [88], diagnostically meaningful concepts in the latent space are encoded. In [89], when training the model for healthy and hypertrophic cardiomyopathy classification, it leverages interpretable task-specific anatomic patterns learned from 3D segmentations.

Post-hoc interpretability.: In [90], the feature importance scores are calculated for graph neural network by comparing its interpretation ability with Random Forest. Li et al. [91] propose a brain biomarker interpretation method through a frequency-normalized sampling

strategy to corrupt an image. In [92], various interpretability methods are evaluated in the context of semantic segmentation of polyps from colonoscopy images. In [93], a hybrid RBM-Random Forest system on brain lesion segmentation is learned with the goal of enhancing interpretability of automatically extracted features.

Uncertainty quantification characterizes the model prediction with confidence measure [94], which can be regarded as a method of post-hoc interpretability, even though often the uncertainty measure is calculated along with the model prediction. Recently there are emerging works that quantify uncertainty in deep learning methods for medical image segmentation [95], [96], [97], lesion detection [98], chest x-ray disease classification [99], and diabetic retinopathy grading [100], [101]. One additional extension to uncertainty is its combination with the knowledge that the given labels are noisy. Works are now starting to emerge that take into account label uncertainty in the modeling of the network architecture and its training [102].

II. CASE STUDIES WITH PROGRESS HIGHLIGHTS

Given that DL has been used in a vast number of medical imaging applications, it is nearly infeasible to cover all possible related literature in a single paper. Therefore, we cover several selected cases that are commonly found in clinical practices, which include chest, neuro, cardiovascular, abdominal, and microscopy imaging. Further, rather than presenting an exhaustive literature survey for each studied case, we provide some prominent progress highlights in each case study.

A. Deep learning in thoracic imaging

Lung diseases have a high mortality and morbidity. In the top ten causes of death worldwide we find lung cancer, chronic obstructive pulmonary disease (COPD), pneumonia, and tuberculosis (TB). At the moment of writing this overview, COVID-19 has a death rate comparable to TB. Imaging is highly relevant to diagnose, plan treatment and learn more about the causes and mechanisms underlying these and other lung diseases. Next to that, pulmonary complications are common in hospitalized patients. As a result, chest radiography is by far the most common radiological examination, often comprising over a third of all studies in a radiology department.

Plain radiography and computed tomography are the two most common modalities to image the chest. The high contrast in density between air-filled lung parenchyma and tissue makes CT ideal for in vivo analysis of the lungs, obtaining high-quality and high-resolution images even at very low radiation dose. Nuclear imaging (PET or PET/CT) is used for diagnosing and staging of oncology patients. MRI is somewhat limited in the lungs, but can yield unique functional information. Ultrasound imaging is also difficult because sound waves reflect strongly at boundaries of air and tissue, but point-of-care ultrasound is used at the emergency department and is widely used to monitor COVID-19 patients for which the first decision support applications based on deep learning have already appeared [103].

Segmentation of anatomical structures.—For analysis and quantification from chest CT scans, automated segmentation of major anatomical structures is an important

prerequisite. Recent publications demonstrate convincingly that deep learning is now the state-of-the-art method to achieve this. This is evident from inspecting the results of LOLA11[3], a competition started in 2011 for lung and lobe segmentation in chest CT. The test dataset for this challenge include many challenging cases with lungs affected by severe abnormalities. For years, the best results were obtained by interactive methods. In 2019 and 2020, seven fully automatic methods based on U-Nets or variants thereof made the top 10 for lung segmentation, and for lobe segmentation, two recent methods obtained results outperforming the best interactive methods. Both these methods [104], [105] were trained on thousands of CT scans from the COPDGene study [106], illustrating the importance of large high quality datasets to obtain good results with deep learning. This data is publicly available on request. Both methods use a multi-resolution U-Net like architecture with several customizations. Gerard et al. integrate a previously developed method for finding the fissures [107]. Xie et al. [105] add a non-local module with self-attention and fine-tune their method on data of COVID-19 suspects to accurately segment the lobes in scans affected by ground-glass and consolidations.

Segmentation of the vasculature, separated into arteries and veins, and the airway tree, including labeling of the branches and segmentation of the bronchial wall, is another important area of research. Although methods that use convolutional networks in some of their steps have been proposed, developing an architecture entirely based on deep learning that can accurately track and segment intertwined tree structures and take advantage of the known geometry of these complex structures, is still an open challenge.

Detection and diagnosis in chest radiography.—Recently the number of publications on detecting abnormalities in the ubiquitous chest x-ray has increased enormously. This trend has been driven by the availability of large public datasets, such as ChestXRay14 [65], CheXpert [108], MIMIC [109], and PadChest [110], totaling 868K images. Labels for presence or absence of over 150 different abnormal signs were gathered by text-mining the accompanying radiology reports. This makes the labels noisy. Most publications use a standard approach of inputting the entire image in a popular convolutional network architecture. Methodological contributions include novel ways of preprocessing the images, handling the label uncertainty and the large number of classes, suppressing the bones [77], and exploiting self-supervised learning as a way of pretraining. So far, only few publications analyze multiple exams of the same patient to detect interval change or analyze the lateral views.

Decision support in lung cancer screening.—Following the positive outcome of the NLST trial, the United States has started a screening program for heavy smokers for early detection of lung cancer with annual low-dose CT scans. Many other countries worldwide are expected to follow suit. In the United States, screening centers have to use a reporting system called Lung-RADS [111]. Reading lung cancer screening CT scans is time consuming and therefore automating the various steps in Lung-RADS has received a lot of attention.

[3] https://lola11.grand-challenge.org/

The most widely studied topic is nodule detection [112]. Nodules may represent lung cancer. Many DL systems were compared in the LUNA16 challenge[4]. Lung-RADS classifies scans in categories based on the most suspicious nodule, and this is determined by the nodule type and size. DL systems to determine nodule type have been proposed [113] and measuring the size can be done by traditional methods based on thresholding and mathematical morphology but also with DL networks. Finally, Lung-RADS contains the option to directly refer scans with a nodule that is highly suspicious for cancer. Many DL systems to estimate nodule malignancy have been proposed.

The advantage of automating the LUNG-RADS guidelines step-by-step is that this leads to an explainable AI solution that can directly support radiologists in their reading workflow. Alternatively, one could ask a computer to directly predict if a CT scan contains an actionable lung cancer. This was the topic of a Kaggle challenge organized in 2017[5] in which almost 2000 teams competed for a one million dollar prize. The top 10 solutions all used deep learning and are open source. Two years later, a team from Google published an implementation [114] following the approach of the winning team in the Kaggle challenge, employing modern architectures such as a 3D inflated Inception architecture (I3D) [115]. The I3D architecture builds upon the Inception v1 model for 2D image classification but inflates the filters and pooling kernels into 3D. This enables the use of an image classification model pre-trained with 2D data for a 3D image classification task. The paper showed that the model outperformed six radiologists that followed Lung-RADS. The model was also extended to handle follow-up scans where it obtained performance slightly below human experts.

COVID-19 case study.—As an illustration how DL with pretrained elements can be used to rapidly build applications, we briefly discuss the development of two tools for COVID-19 detection, for chest radiographs and chest CT. In March 2020, many European hospitals were overwhelmed by patients presenting at the emergency care with respiratory complaints. There was a shortage of molecular testing capacity for COVID-19 and turnaround time for test results was often days. Hospitals therefore used chest x-ray or CT to obtain a working diagnosis and decide whether to hospitalize patients and how to treat them. In just six weeks, researchers from various Dutch and German hospitals, research institutes and a company managed to create a solution for COVID-19 detection from an x-ray and from a CT scan. Figure 3 shows the result of this CORADS-AI system for a COVID-19 positive case.

The x-ray solution started from a convolutional network using local and global labels, pretrained to detect tuberculosis [116], fine-tuned using public and private data of patients with and without pneumonia to detect pneumonia in general, and subsequently fine-tuned on x-ray data from patients of a Dutch hospital in a COVID-19 hotspot. The system was subsequently evaluated on 454 chest radiographs from another Dutch hospital and shown to perform comparably to six chest radiologists [117]. The system is currently being field tested in Africa.

[4] https://luna16.grand-challenge.org/
[5] https://www.kaggle.com/c/data-science-bowl-2017/

The CT solution, called CO-RADS [118], aimed to automate a clinical reporting system for CT of COVID-19 suspects. This system assesses the likelihood of COVID-19 infection on a scale from CO-RADS 1 (highly unlikely) to CO-RADS 5 (highly likely) and quantifies the severity of disease using a score per lung lobe from 0 to 5 depending on percentage affected lung parenchyma for a maximum CT severity score of 25 points. The previously mentioned lobe segmentation [105] was employed. Abnormal areas in the lung were segmented using a 3D U-net built with the nnU-Net framework [119] in a cross-validated fashion with 108 scans and corresponding reference delineations to segment ground-glass opacities and consolidation in the lungs. The CT severity score was derived from the segmentation results by computing the percentage of affected parenchymal tissue per lobe. nnU-Net was compared with several other approaches and performed best. For assessing the CO-RADS score, the previously mentioned I3D architecture [115] performed best.

B. Deep learning in neuroimaging

In recent years, deep learning has seen a dramatic rise in popularity within the neuroimaging community. Many neuroimaging tasks including segmentation, registration, and prediction now have deep learning based implementations. Additionally, through the use of deep generative models and adversarial training, deep learning has enabled new avenues of research in complex image synthesis tasks. With the increasing availability of large and diverse pooled neuroimaging studies, deep learning offers interesting prospects for improving accuracy and generalizability while reducing inference time and the need for complex preprocessing. CNNs, in particular, have allowed for efficient network parameterization and spatial invariance, both of which are critical when dealing with high-dimensional neuroimaging data. The learnable feature reduction and selection capabilities of CNNs have proven effective in high level prediction and analysis tasks and has reduced the need for highly specific domain knowledge. Specialized networks such as U-Nets [38], V-Nets [120], and GANs [44] are also popular in neuroimaging, and have been leveraged for a variety of segmentation and synthesis tasks.

Neuroimage segmentation and tissue classification.—Accurate segmentation is an important preprocessing step that informs much of the downstream analytic and predictive tasks done in neuroimaging. Commonly used tools such as FreeSurfer [122] rely on atlas based methods, whereby an atlas is deformably registered to the scan, which requires time consuming optimization problems to be solved. Proposed deep learning based approaches, however, are relatively computationally inexpensive during inference. Recent research has focused on important tasks such as deep learning based brain extraction [123], cortical and subcortical segmentation [124], [125], [126], [127], and tumor and lesion segmentation [128], [129]. Some interesting research has looked at improving the generalization performance of deep learning based segmentation methods across neuroimaging datasets imaged at different scanners. In particular, Kamnitsas et al. [57] have proposed a training schema, which leverages adversarial training to learn scanner invariant feature representations. They use an adversarial network to classify the origin of the input data based on the downstream feature representation learned by the segmentation network. By penalizing the segmentation network for improved performance of the adversarial network, they show improved segmentation generalization across datasets.

Brain tumor segmentation has been another active area of research in the neuroimaging community where deep learning has shown promise. In the past, brain tumor datasets have remained relatively small, particularly ones with subjects imaged at a single institution. The Brain Tumor Segmentation Challenge (BraTS) [130] has provided the community with an accessible dataset as well as a way to benchmark various approaches against one another. While it has been seen that deep learning has difficulty in training on datasets with relatively few scans, new architectures and training methods are becoming increasingly effective at this. Havaei et al. [128] demonstrate the performance of their glioblastoma segmentation network on the BraTS dataset, achieving high accuracy while being substantially faster than previous methods. Another task where deep networks are finding increasing success is semantic segmentation in which anatomical labels are not necessarily well-defined by image intensity changes but can be identified by relative anatomical locations. A good example is that of cerebellum parcellation where deep networks performed best in a recent comparison of methods [131]. The even newer ACAPULCO method [121] uses two cascaded deep networks to produce cerebellar lobule labels, as shown in Figure 4.

Deformable image registration.—Image registration allows for imaging analysis in a single subject across imaging modalities and time points. Deep learning based deformable registration with neuroimaging data has proven to be a difficult problem, especially considering the lack of ground truth. Still, some unique and varied approaches have achieved state-of-the-art results with relatively fast run times [132], [133], [134], [135]. Li et al. [133] propose a fully convolutional “self-supervised” approach to learn the appropriate spatial transformations at multiple resolutions. Balakrishnan et al. [132] propose a method for unsupervised image registration which attempts to directly compute the deformation field.

Neuroimaging prediction.—With many architectures being borrowed from the computer vision community, deep learning based prediction in neuroimaging has quickly gained popularity. Traditionally, machine learning based prediction on neuroimaging data has relied on careful feature selection/engineering; often taking the form of regional summary measures which may not account for all the informative variation for a particular task. Whereas, in deep learning, it is common to work with raw imaging data, where the appropriate feature representations can be learned through optimization. This can be particularly useful for high-level prediction tasks in which we do not know what imaging features will be informative. Further, by working on the raw image, reliance on complex and time consuming preprocessing can be reduced. In recent years, a large amount of work has been published on deep learning based prediction tasks such as brain age prediction [136], [137], [138], Alzheimer’s disease classification and trajectory modeling [139], [140], [141], [142], and schizophrenia classification [143], [144]. Some work has considered the use of deep Siamese networks for longitudinal image analysis. Siamese networks have gained popularity for their success in facial recognition. They work by jointly optimizing a set of weights on two images with respect to some distance metric between them. This setup makes them effective at identifying longitudinal changes on some chosen dimension. Bhagwat et al. [145] consider the use of longitudinal Siamese network for the prediction of future Alzheimer’s disease onset, using two early time points. They show substantially

improved performance in identifying future Alzheimer's cases, with the use of two time points versus only a baseline scan.

The use of GANs in neuroimaging.—GANs have enabled complex image synthesis tasks in neuroimaging, many of which have no comparable analogs in traditional machine learning. GANs and their variants have been used in neuroimaging for cross-modality synthesis [146], motion artifact reduction [147], resolution upscaling [148], [149], [150], estimating full-dosage PET images from low-dosage PET [151], [152], [153], image harmonization [14], [154], heterogeneity analysis [155], and more. To help facilitate such work, the popular MedGAN [156] proposes a series of modifications and new loss functions to traditional GANs, aimed at preserving anatomically relevant information and fine details. They use auxiliary classifiers on the translated image to ensure the resulting image feature representation is similar to the expected image representation for a given task. Additionally, they use style-transfer loss in combination with adversarial loss to ensure fine structures and textural details are matched in the translation. Some promising new work attempts to reduce the amount of radioactive tracer needed in PET imaging, potentially reducing associated costs and health risks. This problem can be framed as an image synthesis task, whereby the final image can be synthesized from the low dose image. In [157], the pixel location information is integrated into a deep network for image synthesis. Kaplan and Zhu [153] propose a deep generative based denoising method, that uses paired scans of subjects imaged with both low and full dose PET. They show that despite a ten-fold reduction in tracer material, they are able to preserve important edge, structural, and textural details. Consistent quantification in neuroimaging has been hampered for decades by the high variability in MR image intensities and resolutions between scans. Dewey et al. [14] use a U-Net style architecture and paired subjects who have been scanned with two different protocols, to learn a mapping between the two sites. Resolution differences are addressed by applying a super-resolution method to the images acquired at lower resolutions [158]. They are able to use the network to reduce site based variation, which improves consistency in segmentation between the two sites.

While deep learning in neuroimaging has certainly opened up many interesting avenues of investigation, certain areas still lack a rigorous understanding. Important lines of research such as learning from limited data, optimal hyperparameter selection, domain adaptation, semi-supervised designs, and improving robustness require further investigation.

C. Deep learning in cardiovascular imaging

The quantification and understanding of cardiac anatomy and function has been transformed by the recent progress in the field of data-driven deep learning. There has been significant recent work in a variety of sub-areas of cardiovascular imaging including image reconstruction [159], end-to-end learning of cardiac pathology from images [160] and incorporation of non-imaging information (e.g. genetics [161] and clinical information) for analysis. Here we briefly focus on three key aspects of deep learning in this field: cardiac chamber segmentation, cardiac motion/deformation analysis and analysis of cardiac vessels. Motion tracking and segmentation both play crucial roles in the detection and quantification of myocardial chamber dysfunction and can help in the diagnosis of cardiovascular disease

(CVD). Traditionally, these tasks are treated uniquely and solved as separate steps. Often times, motion tracking algorithms use segmentation results as an anatomical guide to sample points and regions of interest used to generate displacement fields [162]. In part due to this, there have also been efforts to combine motion tracking and segmentation.

Cardiac image segmentation is an important first step for many clinical applications. The aim is typically to segment the main chambers, e.g. the left ventricle (LV), right ventricle (RV), left atrium (LA), and right atrium (RA). This enables the quantification of parameters that describe cardiac morphology, e.g. LV volume or mass, or cardiac function, e.g. wall thickening and ejection fraction. There has been significant deep learning work on cardiac chamber segmentation, mostly characterized by the type of images (modalities) employed and whether the work is in 2D or 3D. One of the first efforts to apply a fully convolutional network (FCN) [163] to segment the left ventricle (LV), myocardium and right ventricle from 2D short-axis cardiac magnetic resonance (MR) images was done by Tran [164], significantly outperforming traditional methods in accuracy and speed. Since this time, a variety of other FCN-based strategies have been developed [165], especially focusing on the popular U-Net approach, often including both 2D and 3D constraints (e.g. [166]). The incorporation of spatial and temporal context has also been an important research direction, including efforts to simultaneously segment the heart in both the end-diastolic and end-systolic states [167]. Shape-based constraints had previously been found useful for LV chamber segmentation using other types of machine learning (e.g. [168]) and were nicely included in an anatomically-constrained deep learning strategy by [169]. This stacked convolutional autoencoder approach was successfully applied to LV segmentation from 3D echocardiography data as well. Other important work has been aimed at atrial segmentation from MRI [170], whole heart segmentation from CT [171] and LV segmentation from 3D ultrasound image sequences [172], the latter using a combination of atlas registration and adversarial learning. Progress in deep learning for cardiac segmentation is enabled by a number of ongoing challenges in the field [173], [174].

Cardiac motion tracking is key for deformation/strain analysis and is important for analyzing the mechanical performance of heart chambers. A variety of image registration, feature-based tracking and regularization methods using both biomechanical models and data-driven learning have been developed. One special type of dataset useful for tracking are MRI tagging acquisitions, and deep learning has recently played a role in tracking these tags and quantifying the displacement information for motion tracking and analysis [176] using a combination of recurrent neural networks (RNNs) and convolutional neural networks (CNNs) to estimate myocardial strain from short axis MRI tag image sequences. Estimating motion displacements and strain is also possible to do from both standard MR image sequences and 4D echocardiography, most often by integrating ideas of image segmentation and mapping between frames using some type of image registration. Recent efforts for cardiac motion tracking from magnetic resonance (MR) imaging have adopted approaches from the computer vision field, suggesting that the tasks of motion tracking and segmentation are closely related and information used to complete one task may complement and improve the overall performance of the other. In particular, an interesting deep learning approach proposed for joint learning of video object segmentation and optical flow (motion displacements) is termed SegFlow [177], an end-to-end unified network that

simultaneously trains both tasks and exploits the commonality of these two tasks through bi-directional feature sharing. Among the first to integrate this idea into cardiac analysis was Qin et al. [71], who successfully implemented the idea of combining motion and segmentation on 2D cardiac MR sequences by developing a dual Siamese style recurrent spatial transformer network and fully convolutional segmentation network to simultaneously estimate motion and generate segmentation masks. This work was mainly aimed at 2D MR images, which have higher SNR than echocardiographic images and, therefore, more clearly delineated LV walls. It remains challenging to directly apply this approach to echocardiography. Very recent efforts by Ta et al. [175] (see Fig. 5) propose a 4D (3D+t) semi-supervised joint network to simultaneously track LV motion while segmenting the LV wall. The network is trained in an iterative manner where results from one branch influence and regularize the other. Displacement fields are further regularized by a biomechanically-inspired incompressibility constraint that enforces realistic cardiac motion behavior. The proposed model is different from other models in that it expands the network to 4D in order to capture out of plane motion. Finally, clinical interpretability of deep learning-derived motion information will be an important topic in the years ahead (e.g. [178]).

Cardiac vessel segmentation is another important task for cardiac image analysis and includes the segmentation of the vessels including the great vessels (e.g. aorta, pulmonary arteries and veins) as well as the coronary arteries. The segmentation of large vessels such as the aorta is important for accurate mechanical and hemodynamic characterization, e.g. for assessment of aortic compliance. Several deep learning approaches have been proposed for this segmentation task, including the use of recurrent neural networks in order to track the aorta in cardiac MR image sequences in the presence of noise and artefacts [179]. A similarly important task is the segmentation of the coronary arteries as a precursor to quantitative analysis for the assessment of stenosis or the simulation of blood flow simulation for the calculation of fractional flow reserve from CT angiography (CTA). The approaches for coronary artery segmentation can be divided into those approaches that extract the vessel centerline and those that segment the vessel lumen.

One end-to-end trainable approach for the extraction of the coronary centerline has been proposed in [180]. In this approach the centerline is extracted using a multi-task fully convolutional network which simultaneously computes centerline distance maps and detects branch endpoints. The method generates single-pixel-wide centerlines with no spurious branches. An interesting aspect of this technique is that it can handle an arbitrary vessel tree with no prior assumption regarding depth of the vessel tree or its bifurcation pattern. In contrast to this, Wolterink et al. [181] propose a CNN that is trained to predict the most likely direction and radius of the coronary artery within a local 3D image patch. Starting from a seed point, the coronary artery is tracked by following the vessel centerline using the predictions of the CNN.

Alternative approaches to centerline extraction are based on techniques that instead aim to segment the vessel lumen, e.g. using CNN segmentation methods that perform segmentation by predicting vessel probability maps. One elegant approach has been proposed by Moeskops et al. [182]: Here a single CNN is trained to perform three different segmentation tasks including coronary artery segmentation in cardiac CTA. Instead of performing

voxelwise segmentation, Lee et al. [183] introduce a tubular shape prior for the vessel segments. This is implemented via a template transformer network, through which a shape template can be deformed via network-based registration to produce an accurate segmentation of the input image, as well as to guarantee topological constraints.

More recently, geometric deep learning approaches have also been applied for coronary artery segmentation. For example, Wolterink et al. [184] used graph convolutional networks for coronary artery segmentation. Here vertices on surface of the coronary artery are used as graph nodes and their locations are optimized in an end-to-end fashion.

D. Deep learning in abdominal imaging

Recently there has been an accelerating progress in automated detection, classification and segmentation of abdominal anatomy and disease using medical imaging. Large public data sets such as the MICCAI Data Decathlon and Deep Lesion data sets have facilitated progress [185], [186].

Organs and lesions.—Multi-organ approaches have been popular methods for anatomy localization and segmentation [187]. For individual organs, the liver, prostate and spine are arguably the most accurately segmented structures and the most actively investigated with deep learning. Other organs of interest to deep learning researchers include pancreas, lymph nodes and bowel.

A number of studies have used U-Net to segment the liver and liver lesions and assess for hepatic steatosis [188], [189], [190]. Dice coefficients for liver segmentation typically exceed 95%. In the prostate, gland segmentation and lesion detection have been the subject of an SPIE/AAPM challenge (competition) and numerous publications [191], [192]. Several groups have used data sets such as TCIA CT pancreas to improve pancreas segmentation with Dice coefficients reaching the mid 80 percentile [193], [194], [195], [196]. Automated detection of pancreatic cancer using deep learning has also been reported [196]. Deep learning has been used for determining pancreatic tumor growth rates in patients with neuroendocrine tumors of the pancreas [197]. The spleen has been segmented with a Dice score of 0.962 [198]. Recently marginal loss and exclusion loss [68] have been proposed to train a single multi-organ segmentation network from a union of partially labelled datasets.

Enlarged lymph nodes can indicate the presence of inflammation, infection, or metastatic cancer. Studies have assessed abdominal lymph nodes on CT in general and for specific diseases such as prostate cancer [199], [35], [200]. The TCIA CT lymph node dataset has enabled progress in this area [201].

In the bowel, CT colonography computer-aided polyp detection was a hot topic in abdominal CT image analysis over a decade ago. Recent progress with deep learning has been limited but studies have reported improved electronic bowel cleansing, and higher sensitivities and lower false-positive rates for precancerous colonic polyp detection [199], [202]. Deep learning using persistent homology has recently shown success for small bowel segmentation on CT [203]. Colonic inflammation can be detected on CT with deep learning [204]. Appendicitis can be detected on CT scans by pre-training with natural world videos

[205]. The Inception V3 convolutional neural network could detect small bowel obstruction on abdominal radiographs [206].

Kidney function can be predicted using deep learning of ultrasound images [207]. Potentially diffuse disorders such as ovarian cancer and abnormal blood collections were detectable using deep learning [208], [209]. Organs at risk for radiation therapy of the male pelvis, such as bladder and rectum, have been segmented on CT using U-Net [210].

Universal lesion detectors [186], [211] have been developed for body CT including abdominal CT (Figure 6). The universal lesion detector identifies, classifies and measures lymph nodes and a variety of tumors throughout the abdomen. This detector was trained using the publicly available Deep Lesion data set.

Opportunistic screening to quantify and detect under-reported chronic diseases has been an area of recent interest. Example deep learning methods for opportunistic screening in the abdomen include automated bone mineral densitometry, visceral fat assessment, muscle volume and quality assessment, and aortic atherosclerotic plaque quantification [212]. Studies indicate that these measurements can be done accurately and generalize well to new patient populations. These opportunistic screening assessments also enable prediction of survival and cardiovascular morbidity such as heart attack and stroke [212].

Deep learning for abdominal imaging is likely to continue to advance rapidly. For translation to the clinic, some of the most important advances sought will be in demonstrating generalizability across different patient populations and variations in image acquisition.

E. Deep learning in microscopy imaging

With the advent of whole slide scanning and the development of large digital datasets of tissue slide images, there has been a significant increase in application of deep learning approaches to digital pathology data [213]. While the initial application of these approaches in the area of digital pathology primarily focused on its utility for detection and segmentation of individual primitives like lymphocytes and cancer nuclei, they have now progressed to addressing higher level diagnostic and prognostic tasks and also the application of DL approaches to predict the underlying molecular underpinning and mutational status of the disease. Briefly below we describe the evolving applications of DL approaches to digital pathology.

Nuclei detection and segmentation.—One of the early applications of DL to whole slide pathology images was in the detection and segmentation of individual nuclei. Xu et al. [214] present an approach using stacked spare auto-encoder approach to identify the location of individual cancer nuclei on breast cancer pathology images. Subsequently work from Janowczyk et al. [215] demonstrate the utility of DL approaches for identifying and segmenting a number of different histologic primitives including lymphocytes, tubules, mitotic figures, cancer extent and also for classifying different disease categories pertaining to leukemia. The comprehensive tutorial also went into great detail with regard to best practices for annotation, network training and testing protocols. Subsequently Cruz-Roa et al. demonstrate that convolutional neural networks could be applied for accurately

identifying cancer presence and extent on whole slide breast cancer pathology images [216]. The approach is shown to have a 100% accuracy at identifying the presence or absence of cancer on a slide or patient level. Subsequently Cruz-Roa et al also demonstrate the use of a high-throughput adaptive sampling approach for improving the efficiency of the CNN presented previously in [217]. In [218] Veta and colleagues discuss the diagnostic assessment of DL algorithms for detection of lymph node metastases in women with breast cancer, as part of the CAMELYON16 challenge. The work find that at least five DL algorithms perform comparably to a pathologist interpreting the slides in the absence of time constraints and that some DL algorithms achieve better diagnostic performance than a panel of eleven pathologists participating in a simulation exercise designed to mimic a routine pathology workflow. In a related study of lung cancer pathology images, Coudray et al. [219] train a deep CNN (inception V3) on WSIs from The Cancer Genome Atlas (TCGA) to accurately and automatically classify them into lung adenocarcinoma, squamous cell carcinoma or normal lung tissue, yielding an area under the curve of 0.97.

One of the challenges with the CNN based approaches described in [216], [217], [218], [219] is the need for detailed annotations of the target of interest. This is a labor intensive task given that annotations of disease extent typically need to be provided by pathologists who have a minimal amount of time to begin with. In a comprehensive paper by Campanella et al. [220], the team employs a weakly supervised approach for training a deep learning algorithm for identifying the presence or absence of cancer on a slide level. They are able to demonstrate on a large scale study of over 44K WSIs from over 15K patients, that for prostate cancer, basal cell carcinoma and breast cancer metastases to axillary lymph nodes, the corresponding areas under the curve are all above 0.98. The authors suggest that the approach could be used by pathologists to exclude 65–75% of slides while retaining 100% cancer detection sensitivity.

Disease grading.—Pathologists can reliably identify disease type and extent on H&E slides and have made observations for decades that there are features of disease that correlate with its behavior. However they are unable to reproducibly identify or quantify these histologic hallmarks of disease behavior with enough rigor, to use these features routinely to dictate disease outcome and treatment response. One of the areas to which DL has been applied is mimicking pathologist's identification of disease hallmarks, especially in the context of cancer. For instance in prostate cancer, pathologists typically aim to place the cancer into one of five different categories, referred to as the Gleason grade groups [221]. However this grading system, as with many other cancers and diseases, is subject to inter-reader variability and disagreement. Consequently, a number of recent DL approaches for prostate cancer grading have been presented. Bulten et al. [222] and Strom et al. [223] both recently publish large cohort studies involving 1,243 and 976 patients, respectively, and show that DL approaches could be used for achieving a performance of Gleason grading that is comparable to pathologists [224].

Mutation identification and pathway association.—Disease morphology reflects the sum of all temporal genetic and epigenetic changes and alterations within the disease. Recognizing this, some groups have begun to explore the role of DL approaches for

identifying disease specific mutations and associations with biological pathways. Oncotype DX is a 21-gene expression assay that is prognostic and predictive of benefit of adjuvant chemotherapy for early stage estrogen receptor positive breast cancers. In two related studies [225], [226], Romo-Bucheli show that DL could be used to identify tubule density and mitotic index from pathology images and demonstrate a strong association between these measurements with the Oncotype DX risk categories (low, intermediate, and high) for breast cancers. Interestingly, tubule density and mitotic index are an important component of breast cancer grading. Microsatellite instability (MSI) is a condition that results from impaired DNA mismatch repair. To assess whether a tumor is MSI, genetic or immunohistochemical tests are required. A study by Kather et al. [227] shows that DL could predict MSI from histology images in gastrointestinal cancers with an AUC=0.84. Coudray et al. [219] show that a DL network can be trained to recognize many of the commonly mutated genes in non-small cell lung adenocarcinoma. They show that six of these mutated genes—TK11, EGFR, FAT1, SETBP1, KRAS, and TP53—can be predicted from pathology images, with AUCs from 0.733 to 0.856.

Survival and disease outcome prediction.—More recently, there has been an interest in applying DL algorithms to pathology images to directly predict survival and disease outcome. In a recent paper, Skrede et al. [228] perform a large study of DL involving over 12M pathology image tiles from over 2,000 patients to predict cancer specific survival in early stage colorectal cancer patients. DL yields a hazard ratio for poor versus good prognosis of 3.84 (95% CI 2.72–5.43; $p < 0.0001$) in a validation cohort of 1,122 patients and a hazard ratio of 3.04 (2.07–4.47; $p < 0.0001$) after adjusting for established prognostic markers including T and N stages. Courtiol et al. [229] present an approach employing DL for predicting patient outcome in the case of mesothelioma. Saillard et al. [230] use DL to predict survival after hepatocellular carcinoma resection.

While the studies described above clearly reflect the growing influence and impact of DL on a variety of image analysis and classification problems in digital pathology, there are still concerns with regards to its interpretability, the need for large training sets, the need for annotated data, and generalizability. Attempts have been made to use approaches like visual attention mapping [231] to provide some degree of transparency with respect to where in the image the DL network appears to be focusing its attention. Another approach to imbue interpretability is via hybrid approaches, wherein DL is used to identify specific primitives of interest (e.g. lymphocytes) in the pathology images (in other words using it as a detection and segmentation tool) and then deriving handcrafted features from these primitives (e.g. spatial patterns of arrangement of lymphocytes) to perform prognosis and classification tasks [232], [233], [234]. However, as Bera et al. [213] note in a recent review article, while DL approaches might be feasible for diagnostic indications, clinical tasks relating to outcome prediction and treatment response might still involve approaches that provide greater interpretability. While it seems highly likely that research in DL and its application to digital pathology are likely to continue to grow, it remains to be seen how these approaches fare in a prospective and clinical trial setting, which in turn might ultimately determine their translation to the clinic.

III. DISCUSSION

Technical challenges ahead.

In this overview paper, many technological challenges across several medical domains and tasks have been reviewed. In general, most challenges are met by continuous improvement of solutions to the well known *data challenge*. The community as a whole is continuously developing and improving transfer learning based solutions and data augmentation schemes. As systems are starting to be implemented across datasets, hospitals, and countries, a new spectrum of challenges is arising including *system robustness and generalization* across acquisition protocols, machines, and hospitals. Here, data pre-processing, continuous model learning, and fine-tuning across systems are a few of the new developments ahead. Detailed reviews of the topics presented herein, as well as additional topics, such as robustness to adversarial attacks [235], can be found in several recent DL review articles such as [236].

How do we get new tools into the clinic?

The question whether DL tools are used in the clinic is often raised [5]. This question is particularly relevant because results in many tasks and challenges show radiologist-level performance. In several recent works conducted to estimate the utility of AI-based technology as an aid to the radiologist, it is consistently shown that human experts with AI perform better than those without AI [237]. The excitement in the field has led to the emergence of many AI medical imaging startup companies. Although there are some earlier studies [238], [239], [240] that present evidence that early CAD tools were not necessarily helpful in the tested scenarios and, to date, the translation of technologies from research to actual clinical use does not happen at a massive scale, there are a number of technologies that obtained the FDA approval[6]. In fact, a dedicated reimbursement code was recently awarded for an AI technology[7]. There are a variety of reasons for this delayed clinical translation including: users being cautious regarding the technology, specifically the prospect of being replaced by AI; the need to prove that the technology can address real user needs and bring quantifiable benefits; regulatory pathways that are long and costly; patient safety considerations; and economic factors such as who will pay for AI tools.

The forecast going forward is that this is an emerging field, with enormous promise going forward. How would we get there? An interesting possibility is that the worldwide experience with the COVID-19 pandemic will actually serve to bridge the gap between the *need* and *AI*—with users eager to receive support, and even the regulatory steps more adaptive to facilitate a transition of general computational tools, with COVID-AI related tools in particular. Within the last several months, we witnessed several interesting advances: the ability of AI to rapidly adapt from existing pretrained models to new disease manifestations of COVID-19, using the many tools described in this article and in [241], [117]. Strong and robust DL based solutions for COVID-19 detection, localization, quantification, and characterization start to support the initial diagnosis and more so the followup of hospitalized patients. AI-based tools are developed to support the assessment of disease severity and, recently, they start to provide tools for assessing treatment and

[7]Federal Register / Vol. 85, No. 182 / Friday, September 18, 2020 / Rules and Regulations

predicting treatment success [242], [243]. Finally, numerous studies [244], [245], [246], [247] in fields like clinical neuroscience have shown that AI-based image evaluation can identify complex imaging patterns that are not perceptible with visual radiologic evaluation. For example anatomical, functional, and connectomic signatures of many neusopsychiatic and neurologic diseases not only promise to redefine many of these diseases on a neurobiological basis and in a more precise manner, but also to offer early detection and personalized risk estimates. Such capabilities can potentially significantly expand current clinical protocols and contribute to precision medicine.

Future promise.

As we envision future possibilities, one immediate step forward is to combine the image with additional clinical context, from patient record to additional clinical descriptors (such as blood tests, genomics, medications, vital signs, and non-imaging data such as ECG). This step will provide a transition from *image space* to *patient-level* information. Collecting cohorts will enable population-level statistical analysis to learn about disease manifestations, treatment responses, adverse reactions from and interactions between medications, and more. This step requires building complex infrastructure, along with the generation of new privacy and security regulations—between hospitals and academic research institutes, across hospitals, and in multi-national consortia. As more and more data become available, DL and AI will enable unsupervised explorations within the data, thus providing for new discoveries of drugs and treatments towards advancement and augmentation of healthcare as we know it.

ACKNOWLEDGEMENT

Ronald M. Summers was supported in part by the National Institutes of Health Clinical Center. Anant Madabhushi thanks research support by the National Institutes of Health under award numbers 1U24CA199374-01, R01CA202752-01A1, R01CA208236-01A1, R01CA216579-01A1, R01CA220581-01A1, 1U01CA239055-01, 1U54CA254566-01, 1U01CA248226-01, 1R43EB028736-01 and the VA Merit Review Award IBX004121A from the United States Department of Veterans Affairs Biomedical Laboratory Research and Development Service. We extend our special acknowledgement to Vishnu Bashyam for his help. The mentioning of commercial products does not imply endorsement.

REFERENCES

[1]. Beutel J, Kundel HL, and Van Metter RL, Handbook of Medical Imaging. SPEI Press, 2000, vol. 1.

[2]. Budovec JJ, Lam CA, and Kahn CE Jr, “Informatics in radiology: radiology gamuts ontology: differential diagnosis for the semantic web,” RadioGraphics, vol. 34, no. 1, pp. 254–264, 2014. [PubMed: 24428295]

[3]. Zou KH, Warfield SK, Bharatha A, Tempany CM, Kaus MR, Haker SJ, Wells III WM, Jolesz FA, and Kikinis R, “Statistical validation of image segmentation quality based on a spatial overlap index1: scientific reports,” Academic Radiology, vol. 11, no. 2, pp. 178–189, 2004. [PubMed: 14974593]

[4]. Rubin DL, “Informatics in radiology: measuring and improving quality in radiology: meeting the challenge with informatics,” Radiographics, vol. 31, no. 6, pp. 1511–1527, 2011. [PubMed: 21997979]

[5]. Recht MP, Dewey M, Dreyer K, Langlotz C, Niessen W, Prainsack B, and Smith JJ, “Integrating artificial intelligence into the clinical practice of radiology: challenges and recommendations,” European Radiology, pp. 1–9, 2020.

[6]. Börgers C. and Natterer F, Computational radiology and imaging: therapy and diagnostics. Springer Science & Business Media, 2012, vol. 110.

[7]. Tran WT, Sadeghi-Naini A, Lu F-I, Gandhi S, Meti N, Brackstone M, Rakovitch E, and Curpen B, "Computational radiology in breast cancer screening and diagnosis using artificial intelligence," Canadian Association of Radiologists Journal, p. 0846537120949974, 2020.

[8]. Beutel J, Sonka M, Kundel HL, Van Metter RL, and Fitzpatrick JM, Handbook of Medical Imaging: Medical image processing and analysis. SPIE press, 2000, vol. 2.

[9]. Prince JL and Links JM, Medical Imaging Signals and Systems. Pearson Prentice Hall Upper Saddle River, 2006.

[10]. Bankman I, Handbook of Medical Image Processing and Analysis. Elsevier, 2008.

[11]. Zhou SK, Rueckert D, and Fichtinger G, Handbook of Medical Image Computing and Computer Assisted Intervention. Academic Press, 2019.

[12]. Wang G, Zhang Y, Ye X, and Mou X, Machine learning for tomographic imaging. IOP Publishing, 2019.

[13]. Gaillochet M, Tezcan KC, and Konukoglu E, "Joint reconstruction and bias field correction for undersampled mr imaging," in International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, 2020, pp. 44–52.

[14]. Dewey BE, Zhao C, Reinhold JC, Carass A, Fitzgerald KC, Sotirchos ES, Saidha S, Oh J, Pham DL, Calabresi PA et al. , "DeepHarmony: a deep learning approach to contrast harmonization across scanner changes," Magnetic Resonance Imaging, vol. 64, pp. 160–170, 2019. [PubMed: 31301354]

[15]. Tajbakhsh N, Jeyaseelan L, Li Q, Chiang JN, Wu Z, and Ding X, "Embracing imperfect datasets: A review of deep learning solutions for medical image segmentation," Medical Image Analysis, p. 101693, 2020. [PubMed: 32289663]

[16]. Fu Y, Lei Y, Wang T, Curran WJ, Liu T, and Yang X, "Deep learning in medical image registration: a review," Physics in Medicine & Biology, 2020.

[17]. Chan H-P, Hadjiiski LM, and Samala RK, "Computer-aided diagnosis in the era of deep learning," Medical Physics, vol. 47, no. 5, pp. e218–e227, 2020. [PubMed: 32418340]

[18]. Liu D, Zhou KS, Bernhardt D, and Comaniciu D, "Search strategies for multiple landmark detection by submodular maximization," in 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE, 2010, pp. 2831–2838.

[19]. Xu Z, Huo Y, Park J, Landman B, Milkowski A, Grbic S, and Zhou S, "Less is more: Simultaneous view classification and landmark detection for abdominal ultrasound images," in International Conference on Medical Image Computing and Computer Assisted Intervention. Springer, 2018, pp. 711–719.

[20]. Jing B, Xie P, and Xing E, "On the automatic generation of medical imaging reports," in Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics *(Volume 1: Long Papers)*, 2018, pp. 2577–2586.

[21]. LeCun Y, Bengio Y, and Hinton G, "Deep learning," Nature, vol. 521, no. 7553, pp. 436–444, 2015. [PubMed: 26017442]

[22]. Greenspan H, Van Ginneken B, and Summers RM, "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.

[23]. Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, Van Der Laak JA, Van Ginneken B, and Sánchez CI, "A survey on deep learning in medical image analysis," Medical Image Analysis, vol. 42, pp. 60–88, 2017. [PubMed: 28778026]

[24]. Zhou SK, Greenspan H, and Shen D, Deep learning for medical image analysis. Academic Press, 2017.

[25]. Shen D, Wu G, and Suk H-I, "Deep learning in medical image analysis," Annual Review of Biomedical Engineering, vol. 19, pp. 221–248, 2017.

[26]. Ker J, Wang L, Rao J, and Lim T, "Deep learning applications in medical image analysis," IEEE Access, vol. 6, pp. 9375–9389, 2017.

[27]. Yi X, Walia E, and Babyn P, "Generative adversarial network in medical imaging: A review," Medical Image Analysis, vol. 58, p. 101552–2019.

[28]. Cheplygina V, de Bruijne M, and Pluim JP, “Not-so-supervised: a survey of semi-supervised, multi-instance, and transfer learning in medical image analysis,” Medical Image Analysis, vol. 54, pp. 280–296, 2019. [PubMed: 30959445]

[29]. Hesamian MH, Jia W, He X, and Kennedy P, “Deep learning techniques for medical image segmentation: Achievements and challenges,” Journal of Digital Imaging, vol. 32, no. 4, pp. 582–596, 2019. [PubMed: 31144149]

[30]. Duncan JS, Insana MF, and Ayache N, “Biomedical imaging and analysis in the age of big data and deep learning,” Proceedings of the IEEE, vol. 108, no. 1, pp. 3–10, 2019.

[31]. Haskins G, Kruger U, and Yan P, “Deep learning in medical image registration: a survey,” Machine Vision and Applications, vol. 31, no. 1, p. 8, 2020.

[32]. “MIT Technology Review, 2013,” in https://www.technologyreview.com/s/513696/deep-learning, 2013.

[33]. Krizhevsky A, Sutskever I, and Hinton GE, “ImageNet classification with deep convolutional neural networks,” in Advances in Neural Information Processing Systems, 2012, pp. 1097–1105.

[34]. Bar Y, Diamant I, Wolf L, Lieberman S, Konen E, and Greenspan H, “Chest pathology detection using deep learning with non-medical training,” in International Symposium on Biomedical Imaging (ISBI), 2015, pp. 294–297.

[35]. Shin H-C, Roth HR, Gao M, Lu L, Xu Z, Nogues I, Yao J, Mollura D, and Summers RM, “Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning,” IEEE Transactions on Medical Imaging, vol. 35, no. 5, pp. 1285–1298, 2016. [PubMed: 26886976]

[36]. Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, Venugopalan S, Widner K, Madams T, Cuadros J. et al. , “Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs,” Journal of the American Medical Association, vol. 316, no. 22, pp. 2402–2410, 2016. [PubMed: 27898976]

[37]. Frid-Adar M, Diamant I, Klang E, Amitai M, Goldberger J, and Greenspan H, “GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification,” Neurocomputing, vol. 321, pp. 321–331, 2018.

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

[39]. Simonyan K. and Zisserman A, “Very deep convolutional networks for large-scale image recognition,” arXiv:1409.1556, 2014.

[40]. Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, and Rabinovich A, “Going deeper with convolutions,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 1–9.

[41]. He K, Zhang X, Ren S, and Sun J, “Deep residual learning for image recognition,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 770–778.

[42]. Huang G, Liu Z, Van Der Maaten L, and Weinberger KQ, “Densely connected convolutional networks,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 4700–4708.

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

[44]. Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A, and Bengio Y, “Generative adversarial nets,” in Advances in Neural Information Processing Systems, 2014, pp. 2672–2680.

[45]. Yang D, Xu D, Zhou SK, Georgescu B, Chen M, Grbic S, Metaxas D, and Comaniciu D, “Automatic liver segmentation using an adversarial image-to-image network,” in International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI). Springer, 2017, pp. 507–515.

[46]. Xu K, Ba J, Kiros R, Cho K, Courville A, Salakhudinov R, Zemel R, and Bengio Y, “Show, attend and tell: Neural image caption generation with visual attention,” in International Conference on Machine Learning, 2015, pp. 2048–2057.

[47]. Hu J, Shen L, and Sun G, “Squeeze-and-excitation networks,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 7132–7141.

[48]. Zhang H, Goodfellow I, Metaxas D, and Odena A, “Self-attention generative adversarial networks,” in International Conference on Machine Learning, 2019, pp. 7354–7363.

[49]. Oktay O, Schlemper J, Folgoc LL, Lee M, Heinrich M, Misawa K, Mori K, McDonagh S, Hammerla NY, Kainz B. et al. , “Attention u-net: Learning where to look for the pancreas,” arXiv:1804.03999, 2018.

[50]. Elsken T, Metzen JH, and Hutter F, “Neural architecture search: A survey,” arXiv:1808.05377, 2018.

[51]. Zhu Z, Liu C, Yang D, Yuille A, and Xu D, “V-nas: Neural architecture search for volumetric medical image segmentation,” in International Conference on 3D Vision (3DV). IEEE, 2019, pp. 240–248.

[52]. Howard AG, Zhu M, Chen B, Kalenichenko D, Wang W, Weyand T, Andreetto M, and Adam H, “MobileNets: Efficient convolutional neural networks for mobile vision applications,” arXiv:1704.04861, 2017.

[53]. Zhang X, Zhou X, Lin M, and Sun J, “ShuffleNet: An extremely efficient convolutional neural network for mobile devices,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 6848–6856.

[54]. Raghu M, Zhang C, Kleinberg J, and Bengio S, “Transfusion: Understanding transfer learning for medical imaging,” in Advances in neural information processing systems, 2019, pp. 3347–3357.

[55]. Liu S, Xu D, Zhou SK, Pauly O, Grbic S, Mertelmeier T, Wicklein J, Jerebko A, Cai W, and Comaniciu D, “3D anisotropic hybrid network: Transferring convolutional features from 2D images to 3D anisotropic volumes,” in International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI). Springer, 2018, pp. 851–858.

[56]. Chen S, Ma K, and Zheng Y, “Med3D: Transfer learning for 3D medical image analysis,” arXiv:1904.00625, 2019.

[57]. Kamnitsas K, Baumgartner C, Ledig C, Newcombe V, Simpson J, Kane A, Menon D, Nori A, Criminisi A, Rueckert D. et al., “Unsupervised domain adaptation in brain lesion segmentation with adversarial networks,” in International conference on information processing in medical imaging. Springer, 2017, pp. 597–609.

[58]. Zhang Z, Yang L, and Zheng Y, “Translating and segmenting multimodal medical volumes with cycle-and shape-consistency generative adversarial network,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 9242–9251.

[59]. Dou Q, Ouyang C, Chen C, Chen H, and Heng P-A, “Unsupervised cross-modality domain adaptation of convnets for biomedical image segmentations with adversarial loss,” arXiv:1804.10916, 2018.

[60]. Huang C, Han H, Yao Q, Zhu S, and Zhou SK, “3D U2-net: A 3D universal u-net for multi-domain medical image segmentation,” in International Conference on Medical Image Computing and Computer Assisted Intervention (MICCIAI). Springer, 2019, pp. 291–299.

[61]. Zhou Z, Sodha V, Siddiquee MMR, Feng R, Tajbakhsh N, Gotway MB, and Liang J, “Models genesis: Generic autodidactic models for 3d medical image analysis,” in International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI). Springer, 2019, pp. 384–393.

[62]. Zhu J, Li Y, Hu Y, Ma K, Zhou SK, and Zheng Y, “Rubik’s cube+: A self-supervised feature learning framework for 3D medical image analysis,” Medical Image Analysis, p. 101746, 2020. [PubMed: 32544840]

[63]. Bai W, Oktay O, Sinclair M, Suzuki H, Rajchl M, Tarroni G, Glocker B, King A, Matthews PM, and Rueckert D, “Semi-supervised learning for network-based cardiac MR image segmentation,” in International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI). Springer, 2017, pp. 253–260.

[64]. Nie D, Gao Y, Wang L, and Shen D, “ASDNet: Attention based semi-supervised deep networks for medical image segmentation,” in International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI). Springer, 2018, pp. 370–378.

[65]. Wang X, Peng Y, Lu L, Lu Z, Bagheri M, and Summers RM, "Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 2097–2106.

[66]. Xu Y, Zhu J-Y, Eric I, Chang C, Lai M, and Tu Z, "Weakly supervised histopathology cancer image segmentation and classification," Medical Image Analysis, vol. 18, no. 3, pp. 591–604, 2014. [PubMed: 24637156]

[67]. Kervadec H, Dolz J, Tang M, Granger E, Boykov Y, and Ayed IB, "Constrained-CNN losses for weakly supervised segmentation," Medical Image Analysis, vol. 54, pp. 88–99, 2019. [PubMed: 30851541]

[68]. Shi G, Xiao L, Chen Y, and Zhou SK, "Marginal loss and exclusion loss for partially supervised multi-organ segmentation," arXiv:2007.03868, 2020.

[69]. Schlegl T, Seeböck P, Waldstein SM, Langs G, and Schmidt-Erfurth U, "f-AnoGan: Fast unsupervised anomaly detection with generative adversarial networks," Medical image analysis, vol. 54, pp. 30–44, 2019. [PubMed: 30831356]

[70]. Qin C, Shi B, Liao R, Mansi T, Rueckert D, and Kamen A, "Unsupervised deformable registration for multi-modal images via disentangled representations," in Information Processing in Medical Imaging (IPMI). Springer, 2019, pp. 249–261.

[71]. Qin C, Bai W, Schlemper J, Petersen SE, Piechnik SK, Neubauer S, and Rueckert D, "Joint learning of motion estimation and segmentation for cardiac MR image sequences," in International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI). Springer, 2018, pp. 472–480.

[72]. Liao H, Lin W-A, Zhou SK, and Luo J, "ADN: Artifact disentanglement network for unsupervised metal artifact reduction," IEEE Transactions on Medical Imaging, vol. 39, no. 3, pp. 634–643, 2019. [PubMed: 31395543]

[73]. Ben-Cohen A, Mechrez R, Yedidia N, and Greenspan H, "Improving CNN training using disentanglement for liver lesion classification in CT," in International Conference of the IEEE Engineering in Medicine and Biology Society (IEEE-EMBS 2019), 2019, pp. 886–889.

[74]. Yang J, Dvornek NC, Zhang F, Chapiro J, Lin M, and Duncan JS, "Unsupervised domain adaptation via disentangled representations: Application to cross-modality liver segmentation," in Medical Image Computing and Computer Assisted Intervention (MICCAI). Springer International Publishing, 2019, p. 255–263.

[75]. Chartsias A, Joyce T, Papanastasiou G, Semple S, Williams M, Newby DE, Dharmakumar R, and Tsaftaris SA, "Disentangled representation learning in cardiac image analysis," Medical Image Analysis, vol. 58, p. 101535, Dec 2019. [PubMed: 31351230]

[76]. Li Z, Li H, Han H, Shi G, Wang J, and Zhou SK, "Encoding CT anatomy knowledge for unpaired chest X-Ray image decomposition," in International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI). Springer, 2019, pp. 275–283.

[77]. Li H, Han H, Li Z, Wang L, Wu Z, Lu J, and Zhou SK, "High-resolution chest x-ray bone suppression using unpaired ct structural priors," IEEE Transactions on Medical Imaging, vol. 39, no. 10, pp. 3053–3063, 2020. [PubMed: 32275586]

[78]. Gozes O. and Greenspan H, "Lung structures enhancement in chest radiographs via ct based fcnn training," Lecture Notes in Computer Science, p. 147–158, 2018.

[79]. Lin W-A, Liao H, Peng C, Sun X, Zhang J, Luo J, Chellappa R, and Zhou SK, "DudoNet: Dual domain network for CT metal artifact reduction," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2019, pp. 10 512–10 521.

[80]. Yang Q, Liu Y, Chen T, and Tong Y, "Federated machine learning: Concept and applications," ACM Transactions on Intelligent Systems and Technology, vol. 10, no. 2, pp. 1–19, 2019.

[81]. Konečný J, McMahan HB, Yu FX, Richtárik P, Suresh AT, and Bacon D, "Federated learning: Strategies for improving communication efficiency," arXiv:1610.05492, 2016.

[82]. Zhao Y, Li M, Lai L, Suda N, Civin D, and Chandra V, "Federated learning with non-IID data," arXiv preprint arXiv:1806.00582, 2018.

[83]. Bagdasaryan E, Veit A, Hua Y, Estrin D, and Shmatikov V, “How to backdoor federated learning,” in International Conference on Artificial Intelligence and Statistics, 2020, pp. 2938–2948.

[84]. Sheller MJ, Reina GA, Edwards B, Martin J, and Bakas S, “Multi-institutional deep learning modeling without sharing patient data: A feasibility study on brain tumor segmentation,” in International MICCAI BrainLesion Workshop. Springer, 2018, pp. 92–104.

[85]. Li W, Milletar`ı F, Xu D, Rieke N, Hancox J, Zhu W, Baust M, Cheng Y, s S, Cardoso MJet al., “Privacy-preserving federated brain tumour segmentation,” in International Workshop on Machine Learning in Medical Imaging. Springer, 2019, pp. 133–141.

[86]. Li X, Gu Y, Dvornek N, Staib L, Ventola P, and Duncan JS, “Multi-site fMRI analysis using privacy-preserving federated learning and domain adaptation: Abide results,” arXiv:2001.05647, 2020.

[87]. Murdoch WJ, Singh C, Kumbier K, Abbasi-Asl R, and Yu B, “Interpretable machine learning: definitions, methods, and applications,” arXiv:1901.04592, 2019.

[88]. Clough JR, Oksuz I, Puyol-Antón E, Ruijsink B, King AP, and Schnabel JA, “Global and local interpretability for cardiac MRI classification,” in International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI). Springer, 2019, pp. 656–664.

[89]. Biffi C, Oktay O, Tarroni G, Bai W, De Marvao A, Doumou G, Rajchl M, Bedair R, Prasad S, Cook S. et al., “Learning interpretable anatomical features through deep generative models: Application to cardiac remodeling,” in International Conference on Medical Image Computing and Computer Assisted Intervention. Springer, 2018, pp. 464–471.

[90]. Li X, Dvornek NC, Zhou Y, Zhuang J, Ventola P, and Duncan JS, “Graph neural network for interpreting task-fMRI biomarkers,” in International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI). Springer, 2019, pp. 485–493.

[91]. Li X, Dvornek NC, Zhuang J, Ventola P, and Duncan JS, “Brain biomarker interpretation in ASD using deep learning and fMRI,” in International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI). Springer, 2018, pp. 206–214.

[92]. Wickstrøm K, Kampffmeyer M, and Jenssen R, “Uncertainty and interpretability in convolutional neural networks for semantic segmentation of colorectal polyps,” Medical Image Analysis, vol. 60, p. 101619–2020.

[93]. Pereira S, Meier R, McKinley R, Wiest R, Alves V, Silva CA, and Reyes M, “Enhancing interpretability of automatically extracted machine learning features: application to a RBM-random forest system on brain lesion segmentation,” Medical Image Analysis, vol. 44, pp. 228–244, 2018. [PubMed: 29289703]

[94]. Gal Y. and Ghahramani Z, “Dropout as a bayesian approximation: Representing model uncertainty in deep learning,” in International Conference on Machine Learning, 2016, pp. 1050–1059.

[95]. Baumgartner CF, Tezcan KC, Chaitanya K, Hötker AM, Muehlematter UJ, Schawkat K, Becker AS, Donati O, and Konukoglu E, “Phiseg: Capturing uncertainty in medical image segmentation,” in International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI). Springer, 2019, pp. 119–127.

[96]. Awate SP, Garg S, and Jena R, “Estimating uncertainty in MRF-based image segmentation: A perfect-MCMC approach,” Medical Image Analysis, vol. 55, pp. 181–196, 2019. [PubMed: 31085445]

[97]. Jungo A. and Reyes M, “Assessing reliability and challenges of uncertainty estimations for medical image segmentation,” in International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI). Springer, 2019, pp. 48–56.

[98]. Nair T, Precup D, Arnold DL, and Arbel T, “Exploring uncertainty measures in deep networks for multiple sclerosis lesion detection and segmentation,” Medical Image Analysis, vol. 59, p. 101557–2020.

[99]. Ghesu FC, Georgescu B, Gibson E, Guendel S, Kalra MK, Singh R, Digumarthy SR, Grbic S, and Comaniciu D, “Quantifying and leveraging classification uncertainty for chest radiograph assessment,” in International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI). Springer, 2019, pp. 676–684.

[100]. Ayhan MS, Kuehlewein L, Aliyeva G, Inhoffen W, Ziemssen F, and Berens P, “Expert-validated estimation of diagnostic uncertainty for deep neural networks in diabetic retinopathy detection,” Medical Image Analysis, p. 101724, 2020. [PubMed: 32497870]

[101]. Arauújo T, Aresta G, Mendonça L, Penas S, Maia C, Carneiro Â, Mendonc AM, and Campilho A, “DR|GRADUATE: uncertainty-aware deep learning-based diabetic retinopathy grading in eye fundus images,” Medical Image Analysis, p. 101715, 2020. [PubMed: 32434128]

[102]. Dgani Y, Greenspan H, and Goldberger J, “Training a neural network based on unreliable human annotation of medical images,” in International Symposium on Biomedical Imaging (ISBI). IEEE, 2018, pp. 39–42.

[103]. Roy S, Menapace W, Oei S, Luijten B, Fini E, Saltori C, Huijben I, Chennakeshava N, Mento F, Sentelli A, Peschiera E, Trevisan R, Maschietto G, Torri E, Inchingolo R, Smargiassi A, Soldati G, Rota P, Passerini A, Van Sloun RJG, Ricci E, and Demi L, “Deep learning for classification and localization of covid-19 markers in point-of-care lung ultrasound,” IEEE Transactions on Medical Imaging, 2020.

[104]. Gerard S. and Reinhardt J, “Pulmonary lobe segmentation using a sequence of convolutional neural networks for marginal learning,” in ISBI, 2019, p. 1207–1211.

[105]. Xie W, Jacobs C, Charbonnier J, and van Ginneken B, “Relational modeling for robust and efficient pulmonary lobe segmentation in CT scans,” IEEE Transactions on Medical Imaging, 2020.

[106]. Regan EA, Hokanson JE, Murphy JR, Make B, Lynch DA, Beaty TH, Curran-Everett D, Silverman EK, and Crapo JD, “Genetic epidemiology of COPD (COPDGene) study design,” COPD, vol. 7, pp. 32–43, 2010. [PubMed: 20214461]

[107]. Gerard SE, Patton TJ, Christensen GE, Bayouth JE, and Reinhardt JM, “FissureNet: A deep learning approach for pulmonary fissure detection in ct images.” IEEE Transactions on Medical Imaging, vol. 38, pp. 156–166, 2019. [PubMed: 30106711]

[108]. Irvin J, Rajpurkar P, Ko M, Yu Y, Ciurea-Ilcus S, Chute C, Marklund H, Haghgoo B, Ball R, Shpanskaya K, Seekins J, Mong DA, Halabi SS, Sandberg JK, Jones R, Larson DB, Langlotz CP, Patel BN, Lungren MP, and Ng AY, “CheXpert: A large chest radiograph dataset with uncertainty labels and expert comparison,” arXiv:1901.07031, 2019.

[109]. Johnson AEW, Pollard TJ, Greenbaum NR, Lungren MP, ying Deng C, Peng Y, Lu Z, Mark RG, Berkowitz SJ, and Horng S, “MIMIC-CXR-JPG, a large publicly available database of labeled chest radiographs,” arXiv:1901.07042, 2019.

[110]. Bustos A, Pertusa A, Salinas J-M, and de la Iglesia-Vayá M, “PadChest: A large chest x-ray image dataset with multi-label annotated reports,” arXiv:1901.07441, 2019.

[111]. American College of Radiology, “Lung CT Screening Reporting & Data System v1.1,” 2019. [Online]. Available: https://www.acr.org/Clinical-Resources/Reporting-and-Data-Systems/Lung-Rads

[112]. Liu B, Chi W, Li X, Li P, Liang W, Liu H, Wang W, and He J, “Evolving the pulmonary nodules diagnosis from classical approaches to deep learning-aided decision support: three decades’ development course and future prospect.” Journal of Cancer Research and Clinical Oncology, vol. 146, pp. 153–185, 2020. [PubMed: 31786740]

[113]. Ciompi F, Chung K, van Riel SJ, Setio AAA, Gerke PK, Jacobs C, Scholten ET, Schaefer-Prokop C, Wille MMW, Marchianò A, Pastorino U, Prokop M, and van Ginneken B, “Towards automatic pulmonary nodule management in lung cancer screening with deep learning.” Scientific Reports, vol. 7, p. 46479, Apr. 2017. [PubMed: 28422152]

[114]. Ardila D, Kiraly AP, Bharadwaj S, Choi B, Reicher JJ, Peng L, Tse D, Etemadi M, Ye W, Corrado G, Naidich DP, and Shetty S, “End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography,” Nature Medicine, vol. 25, pp. 954–961, Jun. 2019.

[115]. Carreira J. and Zisserman A, “Quo Vadis, Action Recognition? A New Model and the Kinetics Dataset,” in Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2017, pp. 4724–4733.

[116]. Murphy K, Habib SS, Zaidi SMA, Khowaja S, Khan A, Melendez J, Scholten ET, Amad F, Schalekamp S, Verhagen M, Philipsen RHHM, Meijers A, and van Ginneken B, “Computer

aided detection of tuberculosis on chest radiographs: An evaluation of the CAD4TB v6 system," Scientific Reports, vol. 10, no. 1, pp. 1–11, 2020. [Online]. Available: https://arxiv.org/abs/1903.03349 [PubMed: 31913322]

[117]. Murphy K, Smits H, Knoops AJG, Korst MBJM, Samson T, Scholten ET, Schalekamp S, Schaefer-Prokop CM, Philipsen RHHM, Meijers A, Melendez J, van Ginneken B, and Rutten M, "COVID-19 on the chest radiograph: A multi-reader evaluation of an AI system," Radiology, p. 201874, 2020.

[118]. Prokop M, van Everdingen W, van Rees Vellinga T, Quarles van Ufford J, Stöger L, Beenen L, Geurts B, Gietema H, Krdzalic J, Schaefer-Prokop C, van Ginneken B, Brink M, and the COVID-19 Standardized Reporting Working Group of the Dutch Radiological Society, "CO-RADS - a categorical CT assessment scheme for patients with suspected COVID-19: definition and evaluation," Radiology, p. 201473, 2020. [PubMed: 32339082]

[119]. Isensee F, Jäger PF, Kohl SAA, Petersen J, and Maier-Hein KH, "Automated design of deep learning methods for biomedical image segmentation," arXiv:1904.08128, 2019.

[120]. Milletari F, Navab N, and Ahmadi S-A, "V-net: Fully convolutional neural networks for volumetric medical image segmentation," in The Fourth International Conference on 3D Vision (3DV). IEEE, 2016, pp. 565–571.

[121]. Han S, Carass A, He Y, and Prince JL, "Automatic cerebellum anatomical parcellation using u-net with locally constrained optimization," NeuroImage, p. 116819, 2020. [PubMed: 32438049]

[122]. Fischl B, "FreeSurfer," NeuroImage, vol. 62, no. 2, pp. 774–781, 2012. [PubMed: 22248573]

[123]. Kleesiek J, Urban G, Hubert A, Schwarz D, Maier-Hein K, Bendszus M, and Biller A, "Deep MRI brain extraction: A 3D convolutional neural network for skull stripping," NeuroImage, vol. 129, pp. 460–469, 2016. [PubMed: 26808333]

[124]. Dolz J, Desrosiers C, and Ayed IB, "3D fully convolutional networks for subcortical segmentation in MRI: A large-scale study," NeuroImage, vol. 170, pp. 456–470, 2018. [PubMed: 28450139]

[125]. Chen H, Dou Q, Yu L, Qin J, and Heng P-A, "VoxResNet: Deep voxelwise residual networks for brain segmentation from 3D MR images," NeuroImage, vol. 170, pp. 446–455, 2018. [PubMed: 28445774]

[126]. Wachinger C, Reuter M, and Klein T, "DeepNAT: Deep convolutional neural network for segmenting neuroanatomy," NeuroImage, vol. 170, pp. 434–445, 2018. [PubMed: 28223187]

[127]. Huo Y, Xu Z, Xiong Y, Aboud K, Parvathaneni P, Bao S, Bermudez C, Resnick SM, Cutting LE, and Landman BA, "3d whole brain segmentation using spatially localized atlas network tiles," NeuroImage, vol. 194, pp. 105–119, 2019. [PubMed: 30910724]

[128]. Havaei M, Davy A, Warde-Farley D, Biard A, Courville A, Bengio Y, Pal C, Jodoin P-M, and Larochelle H, "Brain tumor segmentation with deep neural networks," Medical Image Analysis, vol. 35, pp. 18–31, 2017. [PubMed: 27310171]

[129]. Kamnitsas K, Ledig C, Newcombe VF, Simpson JP, Kane AD, Menon DK, Rueckert D, and Glocker B, "Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation," Medical Image Analysis, vol. 36, pp. 61–78, 2017. [PubMed: 27865153]

[130]. Bakas S, Reyes M, Jakab A, Bauer S, Rempfler M, Crimi A, Shinohara RT, Berger C, Ha SM, Rozycki M. et al. , "Identifying the best machine learning algorithms for brain tumor segmentation, progression assessment, and overall survival prediction in the BRATS challenge," arXiv:1811.02629, 2018.

[131]. Carass A, Cuzzocreo JL, Han S, Hernandez-Castillo CR, Rasser PE, Ganz M, Beliveau V, Dolz J, Ayed IB, Desrosiers C, Thyreau B, Romero JE, Coupe P, Manjon JV, Fonov VS, Collins DL, Ying SH, Onyike CU, Crocetti D, Landman BA, Mostofsky SH, Thompson PM, and Prince JL, "Comparing fully automated state-of-the-art cerebellum parcellation from magnetic resonance images," NeuroImage, vol. 183, pp. 150–172, 2018. [PubMed: 30099076]

[132]. Balakrishnan G, Zhao A, Sabuncu MR, Guttag J, and Dalca AV, "VoxelMorph: a learning framework for deformable medical image registration," IEEE Transactions on Medical Imaging, vol. 38, no. 8, pp. 1788–1800, 2019.

[133]. Li H. and Fan Y, "Non-rigid image registration using fully convolutional networks with deep self-supervision," arXiv:1709.00799, 2017.

[134]. Miao S, Wang ZJ, and Liao R, “A CNN regression approach for real-time 2D/3D registration,” IEEE Transactions on Medical Imaging, vol. 35, no. 5, pp. 1352–1363, 2016. [PubMed: 26829785]

[135]. Yang X, Kwitt R, Styner M, and Niethammer M, “QuickSilver: Fast predictive image registration-a deep learning approach,” NeuroImage, vol. 158, pp. 378–396, 2017. [PubMed: 28705497]

[136]. Bashyam VM, Erus G, Doshi J, Habes M, Nasralah I, Truelove-Hill M, Srinivasan D, Mamourian L, Pomponio R, Fan Y. et al. , “MRI signatures of brain age and disease over the lifespan based on a deep brain network and 14,468 individuals worldwide,” Brain, 2020.

[137]. Cole JH, Poudel RP, Tsagkrasoulis D, Caan MW, Steves C, Spector TD, and Montana G, “Predicting brain age with deep learning from raw imaging data results in a reliable and heritable biomarker,” NeuroImage, vol. 163, pp. 115–124, 2017. [PubMed: 28765056]

[138]. Jonsson BA, Bjornsdottir G, Thorgeirsson T, Ellingsen LM, Walters GB, Gudbjartsson D, Stefansson H, Stefansson K, and Ulfarsson M, “Brain age prediction using deep learning uncovers associated sequence variants,” Nature Communications, vol. 10, no. 1, pp. 1–10, 2019.

[139]. Liu S, Liu S, Cai W, Pujol S, Kikinis R, and Feng D, “Early diagnosis of Alzheimer’s disease with deep learning,” in International Symposium on Biomedical Imaging (ISBI), 2014, pp. 1015–1018.

[140]. Lu D, Popuri K, Ding GW, Balachandar R, and Beg MF, “Multimodal and multiscale deep neural networks for the early diagnosis of Alzheimer’s disease using structural MR and FDG-PET images,” Scientific Reports, vol. 8, no. 1, pp. 1–13, 2018. [PubMed: 29311619]

[141]. Islam J. and Zhang Y, “Brain MRI analysis for Alzheimer’s disease diagnosis using an ensemble system of deep convolutional neural networks,” Brain Informatics, vol. 5, no. 2, p. 2, 2018. [PubMed: 29881892]

[142]. Liu M, Zhang J, Adeli E, and Shen D, “Landmark-based deep multi-instance learning for brain disease diagnosis,” Medical Image Analysis, vol. 43, pp. 157–168, 2018. [PubMed: 29107865]

[143]. Yan W, Calhoun V, Song M, Cui Y, Yan H, Liu S, Fan L, Zuo N, Yang Z, Xu K. et al. , “Discriminating schizophrenia using recurrent neural network applied on time courses of multi-site FMRI data,” EBioMedicine, vol. 47, pp. 543–552, 2019. [PubMed: 31420302]

[144]. Zeng L-L, Wang H, Hu P, Yang B, Pu W, Shen H, Chen X, Liu Z, Yin H, Tan Q. et al. , “Multi-site diagnostic classification of schizophrenia using discriminant deep learning with functional connectivity MRI,” EBioMedicine, vol. 30, pp. 74–85, 2018. [PubMed: 29622496]

[145]. Bhagwat N, Viviano JD, Voineskos AN, Chakravarty MM, Initiative ADN et al. , “Modeling and prediction of clinical symptom trajectories in Alzheimer’s disease using longitudinal data,” PLoS Computational Biology, vol. 14, no. 9, p. e1006376, 2018. [PubMed: 30216352]

[146]. Wolterink JM, Dinkla AM, Savenije MH, Seevinck PR, van den Berg BA, and Isgum I, “Deep MR to CT synthesis using unpaired data,” in International Workshop on Simulation and Synthesis in Medical Imaging. Springer, 2017, pp. 14–23.

[147]. Tamada D, Onishi H, and Motosugi U, “Motion artifact reduction in abdominal MR imaging using the U-NET network,” in Proceedings of the ICMRM and Scientific Meeting of KSMRM, Seoul, Korea, 2018.

[148]. Chen Y, Xie Y, Zhou Z, Shi F, Christodoulou AG, and Li D, “Brain MRI super resolution using 3D deep densely connected neural networks,” in International Symposium on Biomedical Imaging (ISBI), 2018, pp. 739–742.

[149]. Pham C-H, Ducournau A, Fablet R, and Rousseau F, “Brain MRI super-resolution using deep 3D convolutional networks,” in International Symposium on Biomedical Imaging (ISBI), 2017, pp. 197–200.

[150]. Song T-A, Chowdhury SR, Yang F, and Dutta J, “Super-resolution PET imaging using convolutional neural networks,” IEEE Transactions on Computational Imaging, vol. 6, pp. 518–528, 2020. [PubMed: 32055649]

[151]. Chen KT, Gong E, de Carvalho Macruz FB, Xu J, Boumis A, Khalighi M, Poston KL, Sha SJ, Greicius MD, Mormino E. et al. , “Ultra-low-dose 18F-florbetaben amyloid PET imaging using deep learning with multi-contrast MRI inputs,” Radiology, vol. 290, no. 3, pp. 649–656, 2019. [PubMed: 30526350]

[152]. Xu J, Gong E, Pauly J, and Zaharchuk G, “200x low-dose PET reconstruction using deep learning,” arXiv:1712.04119, 2017.

[153]. Kaplan S. and Zhu Y-M, “Full-dose PET image estimation from low-dose PET image using deep learning: a pilot study,” Journal of Digital Imaging, vol. 32, no. 5, pp. 773–778, 2019. [PubMed: 30402670]

[154]. Nath V, Parvathaneni P, Hansen CB, Hainline AE, Bermudez C, Remedios S, Blaber JA, Schilling KG, Lyu I, Janve V. et al., “Inter-scanner harmonization of high angular resolution DW-MRI using null space deep learning,” in International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI). Springer, 2018, pp. 193–201.

[155]. Yang Z, Wen J, and Davatzikos C, “Smile-GANs: Semi-supervised clustering via GANs for dissecting brain disease heterogeneity from medical images,” arXiv:2006.15255, 2020.

[156]. Armanious K, Jiang C, Fischer M, Kiistner T, Hepp T, Nikolaou K, Gatidis S, and Yang B, “MedGAN: Medical image translation using GANs,” Computerized Medical Imaging and Graphics, vol. 79, p. 101684–2020.

[157]. Van Nguyen H, Zhou K, and Vemulapalli R, “Cross-domain synthesis of medical images using efficient location-sensitive deep network,” in International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, 2015, pp. 677–684.

[158]. Zhao C, Carass A, Dewey BE, Woo J, Oh J, Calabresi PA, Reich DS, Sati P, Pham DL, and Prince JL, “A deep learning based anti-aliasing self super-resolution algorithm for MRI,” in International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, 2018, pp. 100–108.

[159]. Bustin A, Fuin N, Botnar RM, and Prieto C, “From compressed-sensing to artificial intelligence-based cardiac mri reconstruction,” Frontiers in Cardiovascular Medicine, vol. 7, p. 17–2020. [Online]. Available: 10.3389/fcvm.2020.00017 [PubMed: 32158767]

[160]. Zheng Q, Delingette H, and Ayache N, “Explainable cardiac pathology classification on cine MRI with motion characterization by semi-supervised learning of apparent flow,” Medical Image Analysis, vol. 56, pp. 80–95, August 2019. [PubMed: 31200290]

[161]. de Marvao A, Dawes T, and O’Regan D, “Artificial intelligence for cardiac imaging-genetics research,” Frontiers in Cardiovascular Medicine, vol. 6, p. 195, 2020. [Online]. Available: 10.3389/fcvm.2019.00195 [PubMed: 32039240]

[162]. Parajuli N, Lu ATK, Stendahl J, Boutagy N, Alkhalil I, Eberle M, Jeng G, Zontak M, O’Donnell M, Sinusas A, and Duncan J, “Flow network tracking for spatiotemporal and periodic point matching: Applied to cardiac motion analysis,” Medical Image Analysis, vol. 55, pp. 116–135, 2019. [PubMed: 31055125]

[163]. Shelhamer E, Long J, and D. T., “Fully convolutional networks for semantic segmentation,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 39, pp. 640–651, 2017. [PubMed: 27244717]

[164]. Tran CV, “A fully convolutional neural network for cardiac segmentation in short-axis MRI,” arXiv:1604.00494, 2016.

[165]. Bai W, Sinclair M, Tarroni G, Oktay O, Rajchl M, Vaillant G, Lee AM, Aung N, Lukaschuk E, Sanghvi MM et al. , “Automated cardio-vascular magnetic resonance image analysis with fully convolutional networks,” Journal of Cardiovascular Magnetic Resonance, vol. 20, no. 1, p. 65, 2018. [PubMed: 30217194]

[166]. Isensee F, Jaeger P, Full P, Wolf I, Engelhardt S, and MeierKlein K, “Automatic cardiac disease assessment on cine-MRI via time-series segmentation and domain specific features,” in STACOM 2017: Statistical Atlases and Computational Models of the Heart, MICCAI. Springer LNCS, 2017, pp. 120–129.

[167]. Wolterink J, Leiner T, Viergever M, and Isgum I, “Automatic segmentation and disease classification using cardiac cine mr images,” in STACOM 2017: Statistical Atlases and Computational Models of the Heart, MICCAI. Springer LNCS, 2017, pp. 101–110.

[168]. Huang X, Dione D, Compas C, Papademetris X, Lin B, Bregasi A, Sinusas A, Staib L, and Duncan J, “Contour tracking in echocardiographic sequences via sparse representation and dictionary learning,” Medical Image Analysis, vol. 18, no. 2, pp. 253–271, 2014. [PubMed: 24292554]

[169]. Oktay O, Ferrante E, Kamnitsas K, Heinrich M, Bai W, Caballero J, Cook S, de Marvao A, Dawes T, O'Regan D, Kainz B, Glocker B, and Rueckert D, "Anatomically Constrained Neural Networks (AC-NNs): Application to Cardiac Image Enhancement and Segmentation," IEEE Transactions on Medical Imaging, vol. 37, no. 2, pp. 384–395, 2018. [PubMed: 28961105]

[170]. Xiong Z, Fedorov V, Fu X, Cheng E, Macleod R, and Z. J., "Fully automatic left atrium segmentation from late gadolinium enhanced magnetic resonance imaging using a dual fully convolutional neural network," IEEE Transactions on Medical Imaging, vol. 38, pp. 515–524, 2019. [PubMed: 30716023]

[171]. Ye C, Wang W, Zhang S, and Wang K, "Multi-depth fusion network for whole-heart CT image segmentation," IEEE Access, vol. 7, pp. 23 421–23 429, 2019.

[172]. Dong S, Luo G, Wang K, Cao S, Mercado A, Shmuilovich O, Zhang H, and Li S, "3D left ventricle segmentation on echocardiography with atlas guided generation and voxel-to-voxel discrimination," in International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI), MICCAI. Springer LNCS, 2018, pp. 622–629.

[173]. "Left ventricle full quantification challenge (miccai 2019)," in STACOM 2019: Statistical Atlases and Computational Models of the Heart, MICCAI. Springer LNCS, 2019.

[174]. Campello V. and et al., "Multi-centre, multi-vendor & multi-disease cardiac image segmentation challenge (m&ms))," in Medical Image Computing and Computer Assisted Intervention (MICCAI). MICCAI, 2020.

[175]. Ta K, Ahn S, Stendahl J, Sinusas A, and Duncan J, "A semi-supervised joint network for simultaneous left ventricular motion tracking and segmentation in 4d echocardiography," in International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI), MICCAI. Springer LNCS, 2020, pp. 1–10.

[176]. Ferdian E, Suinesiaputra A, Fung K, Aung N, Lukaschuk E, Barutcu A, Maclean E, Paiva J, Piechnik S, Neubauer S, Petersen S, and Young A, "Fully Automated Myocardial Strain Estimation from Cardiovascular MRI-tagged Images Using a Deep Learning Framework in the UK Biobank," Radiology: Cardiothoracic Imaging, vol. 2, no. 1, p. e190032, 2020. [Online]. Available: 10.1148/ryct.2020190032 [PubMed: 32715298]

[177]. Cheng J, Tsai Y, Wang S, and Yang M-H, "SegFlow: Joint learning for video object segmentation and optical flow," in IEEE International Conference on Computer Vision (ICCV). Ieee, 2017.

[178]. Bello G, Dawes T, Duan J, Biffi C, de Marvao A, Howard L, Gibbs S, Wilkins M, Cook S, Rueckert D, and O'Regan D, "Deep-learning cardiac motion analysis for human survival prediction," Nature Machine Intelligence, vol. 1, pp. 95–104, 2019.

[179]. Bai W, Suzuki H, Qin C, Tarroni G, Oktay O, Matthews PM, and Rueckert D, "Recurrent neural networks for aortic image sequence segmentation with sparse annotations," in International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI), Frangi AF, Schnabel JA, Davatzikos C, Alberola-López C, and Fichtinger G, Eds. Springer International Publishing, 2018, pp. 586–594. [Online]. Available: 10.1007/978-3-030-00937-3_67

[180]. Guo Z, Bai J, Lu Y, Wang X, Cao K, Song Q, Sonka M, and Yin Y, "DeepCenterline: A multi-task fully convolutional network for centerline extraction," in Information Processing in Medical Imaging (IPMI), Chung ACS, Gee JC, Yushkevich PA, and Bao S, Eds. Springer, 2019, pp. 441–453.

[181]. Wolterink JM, van Hamersvelt RW, Viergever MA, Leiner T, and Isgum I, "Coronary artery centerline extraction in cardiac CT angiography using a CNN-based orientation classifier," Medical Image Analysis, vol. 51, pp. 46–60, 2019. [PubMed: 30388501]

[182]. Moeskops P, Wolterink JM, van der Velden BH, Gilhuijs KG, Leiner T, Viergever MA, and Isgum I, "Deep learning for multi-task medical image segmentation in multiple modalities," in International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI), Ourselin S, Joskowicz L, Sabuncu MR, Unal GB, and Wells W, Eds. Springer, 2016, pp. 478–486.

[183]. Lee MCH, Petersen K, Pawlowski N, Glocker B, and Schaap M, "TETRIS: Template transformer networks for image segmentation with shape priors," IEEE Transactions on Medical Imaging, vol. 38, no. 11, pp. 2596–2606, 2019. [PubMed: 30908196]

[184]. Wolterink J, Leiner T, and Isgum I, “Graph convolutional networks for coronary artery segmentation in cardiac CT angiography,” in Workshop on Graph Learning in Medical Imaging, Zhang D, Zhou L, Jie B, and Liu M, Eds. Springer, 2019, pp. 62–69.

[185]. Simpson AL, Antonelli M, Bakas S, Bilello M, Farahani K, Van Ginneken B, Kopp-Schneider A, Landman BA, Litjens G, Menze B. et al. , “A large annotated medical image dataset for the development and evaluation of segmentation algorithms,” arXiv:1902.09063, 2019.

[186]. Yan K, Wang X, Lu L, and Summers RM, “DeepLesion: automated mining of large-scale lesion annotations and universal lesion detection with deep learning,” Journal of Medical Imaging, vol. 5, no. 3, p. 036501, 2018. [PubMed: 30035154]

[187]. Cerrolaza JJ, Picazo ML, Humbert L, Sato Y, Rueckert D, Ballester MAG, and Linguraru MG, “Computational anatomy for multi-organ analysis in medical imaging: A review,” Medical Image Analysis, vol. 56, pp. 44–67, 2019. [PubMed: 31181343]

[188]. Seo H, Huang C, Bassenne M, Xiao R, and Xing L, “Modified U-Net (mU-Net) with incorporation of object-dependent high level features for improved liver and liver-tumor segmentation in CT images,” IEEE Transactions on Medical Imaging, vol. 39, no. 5, pp. 1316–1325, 2019. [PubMed: 31634827]

[189]. Graffy PM, Sandfort V, Summers RM, and Pickhardt PJ, “Automated liver fat quantification at nonenhanced abdominal CT for population-based steatosis assessment,” Radiology, vol. 293, no. 2, pp. 334–342, 2019. [PubMed: 31526254]

[190]. Tang Y-B, Tang Y-X, Zhu Y, Xiao J, and Summers RM, “E2Net: An edge enhanced network for accurate liver and tumor segmentation on CT scans,” in International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI), 2020.

[191]. Armato SG, Huisman H, Drukker K, Hadjiiski L, Kirby JS, Petrick N, Redmond G, Giger ML, Cha K, Mamonov A. et al. , “PROSTATEx challenges for computerized classification of prostate lesions from multiparametric magnetic resonance images,” Journal of Medical Imaging, vol. 5, no. 4, p. 044501, 2018. [PubMed: 30840739]

[192]. Cheng R, Lay NS, Roth HR, Turkbey B, Jin D, Gandler W, McCreedy ES, Pohida TJ, Pinto PA, Choyke PL et al. , “Fully automated prostate whole gland and central gland segmentation on MRI using holistically nested networks with short connections,” Journal of Medical Imaging, vol. 6, no. 2, p. 024007, 2019. [PubMed: 31205977]

[193]. Roth HR, Lu L, Lay N, Harrison AP, Farag A, Sohn A, and Summers RM, “Spatial aggregation of holistically-nested convolutional neural networks for automated pancreas localization and segmentation,” Medical Image Analysis, vol. 45, pp. 94–107, 2018. [PubMed: 29427897]

[194]. Roth HR, Farag A, Turkbey EB, Lu L, Liu J, and Summers RM, “Data from Pancreas-CT,” in Cancer Imaging Archive, 2015.

[195]. Roth HR, Oda H, Zhou X, Shimizu N, Yang Y, Hayashi Y, Oda M, Fujiwara M, Misawa K, and Mori K, “An application of cascaded 3D fully convolutional networks for medical image segmentation,” Computerized Medical Imaging and Graphics, vol. 66, pp. 90–99, 2018. [PubMed: 29573583]

[196]. Chu LC, Park S, Kawamoto S, Wang Y, Zhou Y, Shen W, Zhu Z, Xia Y, Xie L, Liu F. et al. , “Application of deep learning to pancreatic cancer detection: Lessons learned from our initial experience,” Journal of the American College of Radiology, vol. 16, no. 9, pp. 1338–1342, 2019. [PubMed: 31492412]

[197]. Zhang L, Lu L, Wang X, Zhu RM, Bagheri M, Summers RM, and Yao J, “Spatio-temporal convolutional LSTMs for tumor growth prediction by learning 4d longitudinal patient data,” IEEE Transactions on Medical Imaging, vol. 39, no. 4, pp. 1114–1126, 2019. [PubMed: 31562074]

[198]. Humpire-Mamani GE, Bukala J, Scholten ET, Prokop M, van Ginneken B, and Jacobs C, “Fully automatic volume measurement of the spleen at ct using deep learning,” Radiology: Artificial Intelligence, vol. 2, no. 4, p. e190102, 2020. [PubMed: 33937830]

[199]. Roth HR, Lu L, Liu J, Yao J, Seff A, Cherry K, Kim L, and Summers RM, “Improving computer-aided detection using convolutional neural networks and random view aggregation,” IEEE Transactions on Medical Imaging, vol. 35, no. 5, pp. 1170–1181, 2016. [PubMed: 26441412]

[200]. Tang Y-B, Ke Y, Xiao J, and Summers RM, “One click lesion RECIST measurement and segmentation on CT scans,” in International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI), 2020.

[201]. Roth HR, Lu L, Seff A, Cherry KM, Wang S, Liu J, Turkbey E, and Summers RM, “CT lymph nodes,” in Cancer Imaging Archive, 2015.

[202]. Tachibana R, Nappi JJ, Ota J, Kohlhase N, Hironaka T, Kim SH, Regge D, and Yoshida H, “Deep learning electronic cleansing for single-and dual-energy CT colonography,” RadioGraphics, vol. 38, no. 7, pp. 2034–2050, 2018. [PubMed: 30422761]

[203]. Shin S, Lee S, Elton D, Gullery J, and Summers RM, “Deep small bowel segmentation with cylindrical topological constraints,” in International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI), 2020.

[204]. Liu J, Wang D, Lu L, Wei Z, Kim L, Turkbey EB, Sahiner B, Petrick NA, and Summers RM, “Detection and diagnosis of colitis on computed tomography using deep convolutional neural networks,” Medical Physics, vol. 44, no. 9, pp. 4630–4642, 2017. [PubMed: 28594460]

[205]. Rajpurkar P, Park A, Irvin J, Chute C, Bereket M, Mastrodicasa D, Langlotz CP, Lungren MP, Ng AY, and Patel BN, “AppendiXNet: Deep learning for diagnosis of appendicitis from a small dataset of CT exams using video pretraining,” Scientific Reports, vol. 10, no. 1, pp. 1–7, 2020. [PubMed: 31913322]

[206]. Cheng PM, Tejura TK, Tran KN, and Whang G, “Detection of high-grade small bowel obstruction on conventional radiography with convolutional neural networks,” Abdominal Radiology, vol. 43, no. 5, pp. 1120–1127, 2018. [PubMed: 28828625]

[207]. Kuo C-C, Chang C-M, Liu K-T, Lin W-K, Chiang H-Y, Chung C-W, Ho M-R, Sun P-R, Yang R-L, and Chen K-T, “Automation of the kidney function prediction and classification through ultrasound-based kidney imaging using deep learning,” NPJ Digital Medicine, vol. 2, no. 1, pp. 1–9, 2019. [PubMed: 31304351]

[208]. Wang S, Liu Z, Rong Y, Zhou B, Bai Y, Wei W, Wang M, Guo Y, and Tian J, “Deep learning provides a new computed tomography-based prognostic biomarker for recurrence prediction in high-grade serous ovarian cancer,” Radiotherapy and Oncology, vol. 132, pp. 171–177, 2019. [PubMed: 30392780]

[209]. Dreizin D, Zhou Y, Zhang Y, Tirada N, and Yuille AL, “Performance of a deep learning algorithm for automated segmentation and quantification of traumatic pelvic hematomas on CT,” Journal of Digital Imaging, vol. 33, no. 1, pp. 243–251, 2020. [PubMed: 31172331]

[210]. Kazemifar S, Balagopal A, Nguyen D, McGuire S, Hannan R, Jiang S, and Owrangi A, “Segmentation of the prostate and organs at risk in male pelvic CT images using deep learning,” Biomedical Physics & Engineering Express, vol. 4, no. 5, p. 055003, 2018.

[211]. Li H, Han H, and Zhou SK, “Bounding maps for universal lesion detection,” arXiv:2007.09383, 2020.

[212]. Pickhardt PJ, Graffy PM, Zea R, Lee SJ, Liu J, Sandfort V, and Summers RM, “Automated CT biomarkers for opportunistic prediction of future cardiovascular events and mortality in an asymptomatic screening population: a retrospective cohort study,” The Lancet Digital Health, 2020.

[213]. Bera K, Schalper KA, Rimm DL, Velcheti V, and Madabhushi A, “Artificial intelligence in digital pathology—new tools for diagnosis and precision oncology,” Nature Reviews Clinical Oncology, vol. 16, no. 11, pp. 703–715, 2019.

[214]. Xu J, Xiang L, Liu Q, Gilmore H, Wu J, Tang J, and Madabhushi A, “Stacked sparse autoencoder (SSAE) for nuclei detection on breast cancer histopathology images,” IEEE Transactions on Medical Imaging, vol. 35, no. 1, pp. 119–130, 2015. [PubMed: 26208307]

[215]. Janowczyk A. and Madabhushi A, “Deep learning for digital pathology image analysis: A comprehensive tutorial with selected use cases,” Journal of Pathology Informatics, vol. 7, 2016.

[216]. Cruz-Roa A, Gilmore H, Basavanhally A, Feldman M, Ganesan S, Shih NN, Tomaszewski J, Gonzalez FA, and Madabhushi A, “Accurate and reproducible invasive breast cancer detection in whole-slide images: A deep learning approach for quantifying tumor extent,” Scientific Reports, vol. 7, p. 46450–2017. [PubMed: 28418027]

[217]. Cruz-Roa A, Gilmore H, Basavanhally A, Feldman M, Ganesan S, Shih N, Tomaszewski J, Madabhushi A, and Gonzalez F, “High-throughput adaptive sampling for whole-slide histopathology image analysis (HASHI) via convolutional neural networks: Application to invasive breast cancer detection,” PloS One, vol. 13, no. 5, p. e0196828, 2018. [PubMed: 29795581]

[218]. Bejnordi BE, Veta M, Van Diest PJ, Van Ginneken B, Karssemeijer N, Litjens G, Van Der Laak JA, Hermsen M, Manson QF, Balkenhol M. et al. , “Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer,” Journal of the American Medical Association, vol. 318, no. 22, pp. 2199–2210, 2017. [PubMed: 29234806]

[219]. Coudray N, Ocampo PS, Sakellaropoulos T, Narula N, Snuderl M, Fenyö D, Moreira AL, Razavian N, and Tsirigos A, “Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning,” Nature Medicine, vol. 24, no. 10, pp. 1559–1567, 2018.

[220]. Campanella G, Hanna MG, Geneslaw L, Miraflor A, Silva VWK, Busam KJ, Brogi E, Reuter VE, Klimstra DS, and Fuchs TJ, “Clinical-grade computational pathology using weakly supervised deep learning on whole slide images,” Nature Medicine, vol. 25, no. 8, pp. 1301–1309, 2019.

[221]. Babaian RJ and Grunow WA, “Reliability of gleason grading system in comparing prostate biopsies with total prostatectomy specimens,” Urology, vol. 25, no. 6, pp. 564–567, 1985. [PubMed: 4012946]

[222]. Bulten W, Pinckaers H, van Boven H, Vink R, de Bel T, van Ginneken B, van der Laak J, Hulsbergen-van de Kaa C, and Litjens G, “Automated deep-learning system for gleason grading of prostate cancer using biopsies: a diagnostic study,” The Lancet Oncology, vol. 21, no. 2, pp. 233–241, 2020. [PubMed: 31926805]

[223]. Strom P, Kartasalo K, Olsson H, Solorzano L, Delahunt B, Berney DM, Bostwick DG, Evans AJ, Grignon DJ, Humphrey PA et al. , “Artificial intelligence for diagnosis and grading of prostate cancer in biopsies: a population-based, diagnostic study,” The Lancet Oncology, vol. 21, no. 2, pp. 222–232, 2020. [PubMed: 31926806]

[224]. Madabhushi A, Feldman MD, and Leo P, “Deep-learning approaches for gleason grading of prostate biopsies,” The Lancet Oncology, vol. 21, no. 2, pp. 187–189, 2020. [PubMed: 31926804]

[225]. Romo-Bucheli D, Janowczyk A, Gilmore H, Romero E, and Madabhushi A, “Automated tubule nuclei quantification and correlation with oncotype DX risk categories in ER+ breast cancer whole slide images,” Scientific Reports, vol. 6, p. 32706–2016.

[226]. Romo-Bucheli D, “A deep learning based strategy for identifying and associating mitotic activity with gene expression derived risk categories in estrogen receptor positive breast cancers,” Cytometry Part A, vol. 91, no. 6, pp. 566–573, 2017.

[227]. Kather JN, Pearson AT, Halama N, Jager D, Krause J, Loosen SH, Marx A, Boor P, Tacke F, Neumann UP et al. , “Deep learning can predict microsatellite instability directly from histology in gastrointestinal cancer,” Nature Medicine, vol. 25, no. 7, pp. 1054–1056, 2019.

[228]. Skrede O-J, De Raedt S, Kleppe A, Hveem TS, Liest0l K, Maddison J, Askautrud HA, Pradhan M, Nesheim JA, Albregtsen F. et al. , “Deep learning for prediction of colorectal cancer outcome: a discovery and validation study,” The Lancet, vol. 395, no. 10221, pp. 350–360, 2020.

[229]. Courtiol P, Maussion C, Moarii M, Pronier E, Pilcer S, Sefta M, Manceron P, Toldo S, Zaslavskiy M, Le Stang N. et al. , “Deep learning-based classification of mesothelioma improves prediction of patient outcome,” Nature Medicine, vol. 25, no. 10, pp. 1519–1525, 2019.

[230]. Saillard C, Schmauch B, Laifa O, Moarii M, Toldo S, Zaslavskiy M, Pronier E, Laurent A, Amaddeo G, Regnault H. et al. , “Predicting survival after hepatocellular carcinoma resection using deep-learning on histological slides,” Hepatology, 2020.

[231]. Tomita N, Abdollahi B, Wei J, Ren B, Suriawinata A, and Hassanpour S, “Attention-based deep neural networks for detection of cancerous and precancerous esophagus tissue on histopathological slides,” JAMA network open, vol. 2, no. 11, pp. e1 914645-e1 914 645, 2019.

[232]. Heindl A, Sestak I, Naidoo K, Cuzick J, Dowsett M, and Yuan Y, “Relevance of spatial heterogeneity of immune infiltration for predicting risk of recurrence after endocrine therapy of ER+ breast cancer,” Journal of the National Cancer Institute, vol. 110, no. 2, pp. 166–175, 2018.

[233]. Saltz J, Gupta R, Hou L, Kurc T, Singh P, Nguyen V, Samaras D, Shroyer KR, Zhao T, Batiste R. et al. , “Spatial organization and molecular correlation of tumor-infiltrating lymphocytes using deep learning on pathology images,” Cell Reports, vol. 23, no. 1, pp. 181–193, 2018. [PubMed: 29617659]

[234]. Corredor G, Wang X, Zhou Y, Lu C, Fu P, Syrigos K, Rimm DL, Yang M, Romero E, Schalper KA et al. , “Spatial architecture and arrangement of tumor-infiltrating lymphocytes for predicting likelihood of recurrence in early-stage non-small cell lung cancer,” Clinical Cancer Research, vol. 25, no. 5, pp. 1526–1534, 2019. [PubMed: 30201760]

[235]. Yao Q, He Z, Lin Y, Ma K, Zheng Y, and Zhou SK, “A hierarchical feature constraint to camouflage medical adversarial attacks,” arXiv preprint arXiv:2012.09501, 2020.

[236]. Rueckert D. and Schnabel J, “Model-based and data-driven strategies in medical image computing,” Proceedings of the IEEE, vol. 108, no. 1, pp. 110–124, 2020.

[237]. Brown MS, Browning P, Wahi-Anwar MW, Murphy M, Delgado J, Greenspan H, Abtin F, Ghahremani S, Yaghmai N, da Costa I, Becker M, and Goldin JG, “Integration of chest CT CAD into the clinical workflow and impact on radiologist efficiency.” Academic Radiology, vol. 26 5, pp. 626–631, 2019. [PubMed: 30097402]

[238]. de Hoop B, De Boo DW, Gietema HA, van Hoorn F, Mearadji B, Schijf L, van Ginneken B, Prokop M, and Schaefer-Prokop C, “Computer-aided detection of lung cancer on chest radiographs: effect on observer performance,” Radiology, vol. 257, no. 2, pp. 532–540, 2010. [PubMed: 20807851]

[239]. Jorritsma W, Cnossen F, and van Ooijen PM, “Improving the radiologist-cad interaction: designing for appropriate trust,” Clinical radiology, vol. 70, no. 2, pp. 115–122, 2015. [PubMed: 25459198]

[240]. Drew T, Cunningham C, and Wolfe JM, “When and why might a computer-aided detection (cad) system interfere with visual search? an eye-tracking study,” Academic radiology, vol. 19, no. 10, pp. 1260–1267, 2012. [PubMed: 22958720]

[241]. Gozes O, Frid-Adar M, Greenspan H, Browning PD, Zhang H, Ji W, Bernheim A, and Siegel E, “Rapid AI development cycle for the coronavirus (COVID-19) pandemic: Initial results for automated detection & patient monitoring using deep learning CT image analysis,” 2020.

[242]. Shi F, Wang J, Shi J, Wu Z, Wang Q, Tang Z, He K, Shi Y, and Shen D, “Review of artificial intelligence techniques in imaging data acquisition, segmentation and diagnosis for covid-19,” IEEE Reviews in Biomedical Engineering, 2020.

[243]. Greenspan H, San-Jose Estepar R, Niessen W, Siegel E, and Nielsen M, “Position paper on COVID-19 imaging and AI: From the clinical needs and technological challenges to initial ai solutions at the lab and national level towards a new era for ai in healthcare,” Medical Image Analysis - Special issue with TMI on COVID-19, 2020.

[244]. “A review on neuroimaging-based classification studies and associated feature extraction methods for alzheimer’s disease and its prodromal stages,” NeuroImage, vol. 155, pp. 530–548, 2017. [PubMed: 28414186]

[245]. Mateos-Perez JM, Dadar M, Lacalle-Aurioles M, Iturria-Medina Y, Zeighami Y, and Evans AC, “Structural neuroimaging as clinical predictor: A review of machine learning applications,” NeuroImage: Clinical, vol. 20, pp. 506–522, 2018. [PubMed: 30167371]

[246]. Nielsen AN, Barch DM, Petersen SE, Schlaggar BL, and Greene DJ, “Machine learning with neuroimaging: evaluating its applications in psychiatry,” Biological Psychiatry: Cognitive Neuroscience and Neuroimaging, 2019.

[247]. Fathi Kazerooni A, Bakas S, Saligheh Rad H, and Davatzikos C, “Imaging signatures of glioblastoma molecular characteristics: A radiogenomics review,” Journal of Magnetic Resonance Imaging, vol. 52, no. 1, pp. 54–69, 2020. [PubMed: 31456318]

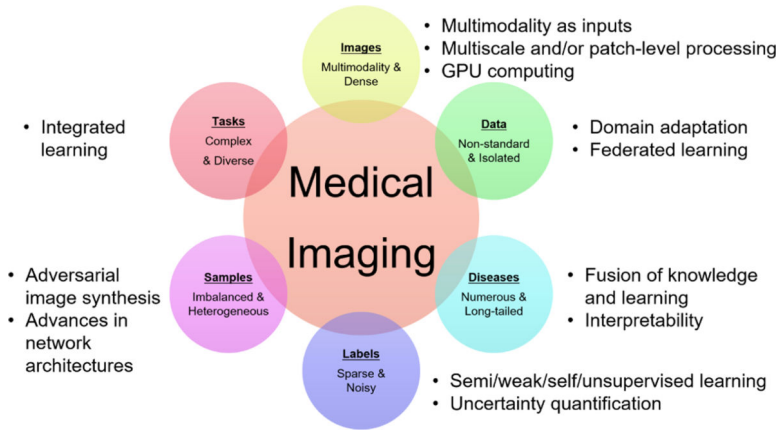

Fig. 1.

The main traits of medical imaging and the associated technological trends for addressing these traits.

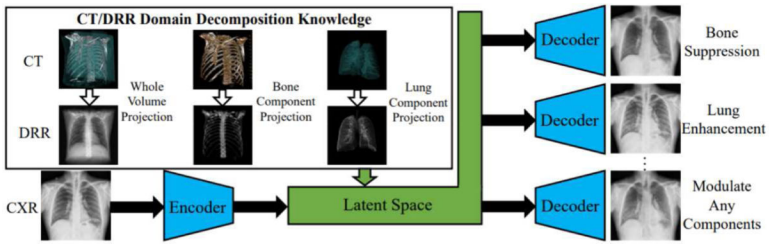

Fig. 2.

Leveraging the anatomy knowledge embedded in CT to decompose a chest x-ray [76].

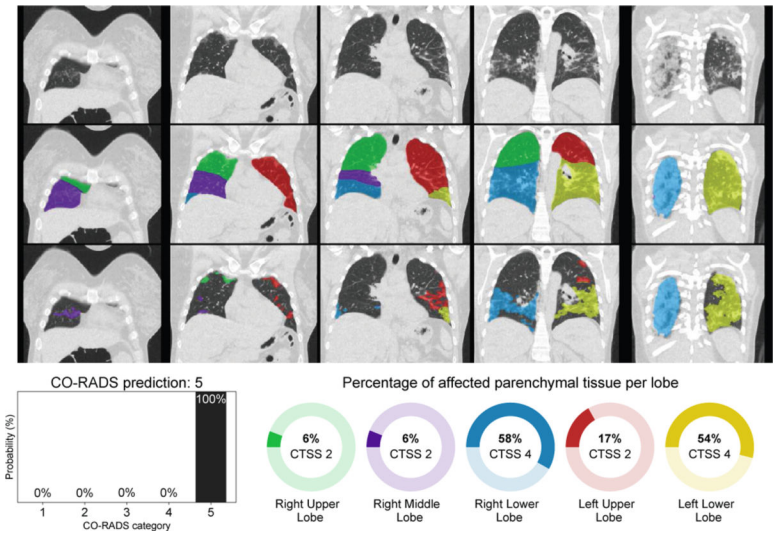

Fig. 3.

Example output of the CORADS-AI system for a COVID-19 case. Top row shows coronal slices, the second row shows lobe segmentation and bottom row shows detected abnormal areas of patchy ground-glass and consolidation typical for COVID-19 infection. The CO-RADS prediction and CT severity score per lobe are displayed below the images.

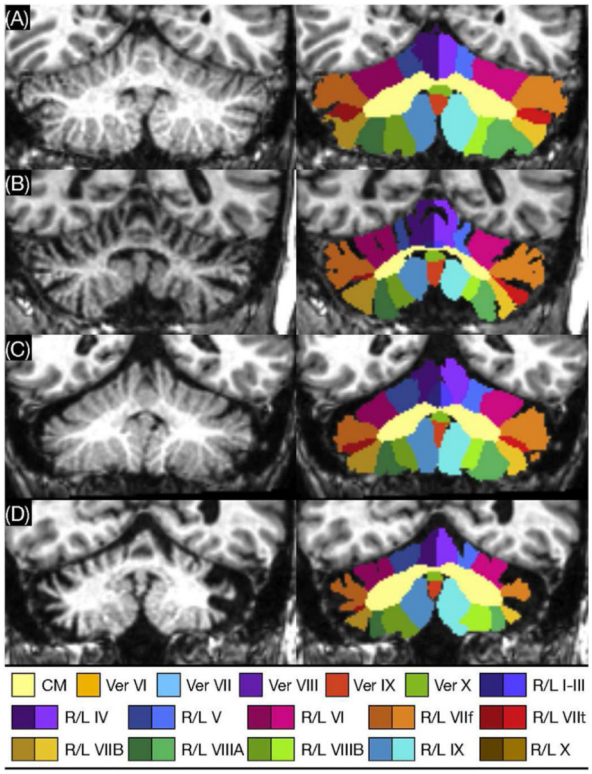

Fig. 4.

Cerebellum parcellation by the ACAPULCO cascaded deep networks method. Lobule labels are shown for (A) a healthy subject and subjects with (B) spinocerebellar ataxia (SCA) type 2, (C) SCA type 3, and (D) SCA type 6 [121].

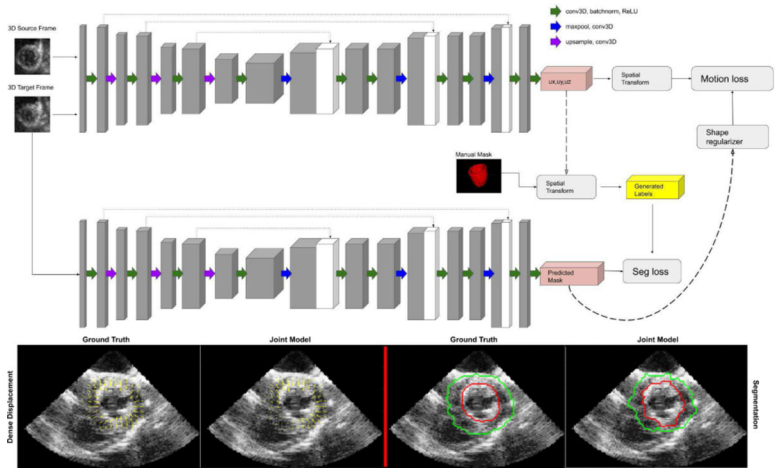

Fig. 5.

Top: 4D two-channel CNN architecture for joint LV motion tracking/segmentation network; Bottom: 2D echocardiography slices through 3D canine results (left = displacements with ground truth interpolated from sonomicrometer crystals; right = LV endocardial (red) and epicardial (green) segmented boundaries with ground truth from human expert tracing.) Reproduced from [175].

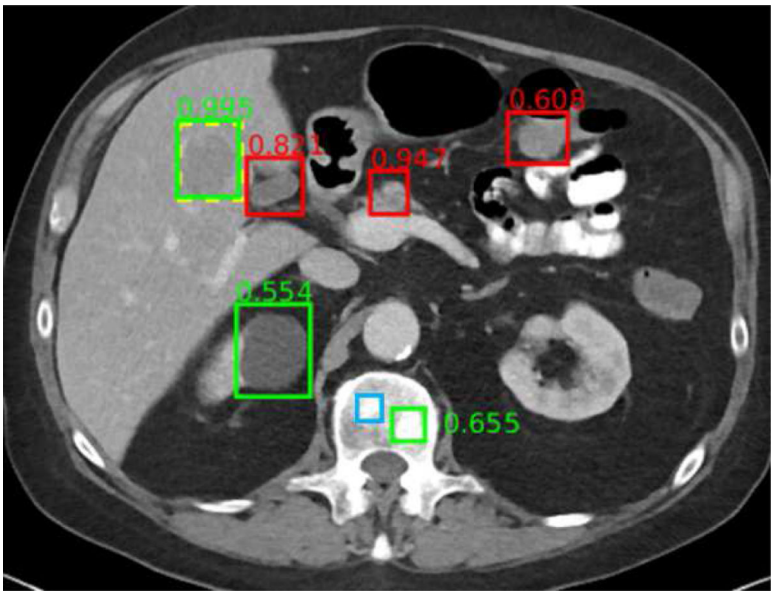

Fig. 6.

Example universal lesion detector for abdominal CT. In this axial image through the upper abdomen, a liver lesion was correctly detected with high confidence (0.995). A renal cyst (0.554) and a bone metastasis (0.655) were also detected correctly. False positives include normal pancreas (0.947), gallbladder (0.821), and bowel (0.608). A subtle bone metastasis (blue box) was missed. Reproduced from [186].

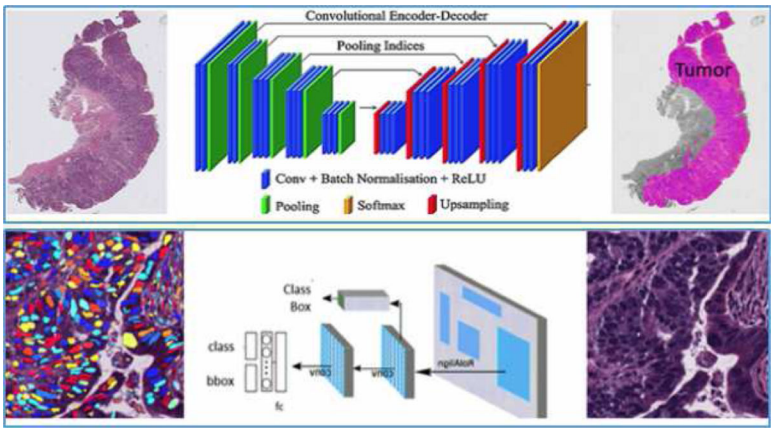

Fig. 7.

Application of deep learning for identifying cancerous regions from whole slide images as well as for identifying and segmenting different types of nuclei within the whole slide pathology images.