**TOPICAL REVIEW • OPEN ACCESS**

# Synergizing medical imaging and radiotherapy with deep learning

MACHINE LEARNING
Science and Technology

OPEN ACCESS

RECEIVED
24 September 2019

REVISED
22 March 2020

ACCEPTED FOR PUBLICATION
3 April 2020

PUBLISHED
22 June 2020



TOPICAL REVIEW

# Synergizing medical imaging and radiotherapy with deep learning

**Hongming Shan**[1,6], **Xun Jia**[2,6], **Pingkun Yan**[1], **Yunyao Li**[3], **Harald Paganetti**[4] and **Ge Wang**[1,5]

[1] Department of Biomedical Engineering, Rensselaer Polytechnic Institute, Troy, NY 12180, United States of America
[2] Department of Radiation Oncology, UT Southwestern Medical Center, Dallas, TX 75390, United States of America
[3] IBM Research, Almaden, San Jose, CA 95120, United States of America
[4] Department of Radiation Oncology, Massachusetts General Hospital and Harvard Medical School, Boston, MA 02114, United States of America
[5] Author to whom any correspondence should be addressed.
[6] H S and X J are co-first authors.

**E-mail:** wangg6@rpi.edu

## Abstract

This article reviews deep learning methods for medical imaging (focusing on image reconstruction, segmentation, registration, and radiomics) and radiotherapy (ranging from planning and verification to prediction) as well as the connections between them. Then, future topics are discussed involving semantic analysis through natural language processing and graph neural networks. It is believed that deep learning in particular, and artificial intelligence and machine learning in general, will have a revolutionary potential to advance and synergize medical imaging and radiotherapy for unprecedented smart precision healthcare.

## 1. Introduction

McCarthy *et al* [1] organized the Dartmouth workshop in 1956 to initiate artificial intelligence (AI) as a research field with a lofty goal to simulate, enhance, or even surpass human intelligence. Given the tremendous potentials and challenges, the excitements and frustrations are equally remarkable. Their interactions lead to alterations of AI springs and winters, through which the AI field has been developed step by step, and elevated to today's level, and we believe that this field will have an even brighter future.

Currently, AI is in a new spring, especially its sub-field machine learning (ML) which enjoys rapid development and constant innovations featured by deep neural networks, also known as deep learning. On August 30, 2019, the White House issued a memorandum on the Fiscal Year 2021 Administration Research and Development Budget Priorities [2], underlining that 'departments and agencies should prioritize basic and applied research investments that are consistent with the 2019 Executive Order on Maintaining American Leadership in Artificial Intelligence and the eight strategies detailed in the 2019 update of the National Artificial Intelligence Research and Development Strategic Plan.' According to the World Economic Forum [3], it is estimated in 2018 that 'by 2025, the amount of work done by machines will jump from 29% to more than 50% –but that this rapid shift will be accompanied by new labour-market demands that may result in more, rather than fewer, jobs.' 'Preparing the workforce for these changes will depend on a data-driven approach.'

In the overwhelming optimistic atmosphere of deep learning, hypes are also unavoidable that exaggerate the power of AI and over-promise the return of some AI-related investments. Furthermore, some peers wonder if the potential of AI or deep learning techniques has been largely saturated, and worry if another AI winter is coming. In 2018, a group with the MIT Technology Review analyzed 16,625 arXiv papers in the AI section published since 1993, and found that the momentum is slowing down [4].

While the future is always hard to predict such as how long the AI spring will last or how soon the AI winter will come again, as a group of medical physicists and AI engineers we believe that the AI spring will last into the near future. In addition to potential new methodological and technical progresses, we are confident that major efforts are needed to translate current (and future) deep learning methods into medical physics and clinical practice. This need is a powerful engine that will drive AI research and development productively for at least ten years.

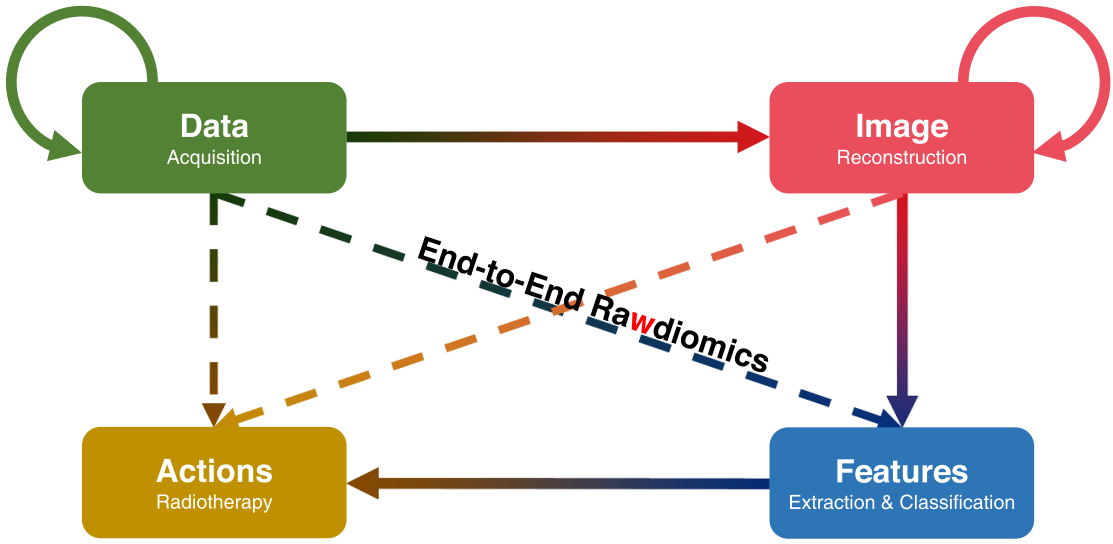


**Figure 1.** Overall big picture of deep imaging towards radiotherapy.

Two well-known AI-related areas in the field of medical physics are medical imaging and radiation therapy (radiotherapy). The penetration of AI or deep learning into these two areas has been significant. Generally speaking, more deep learning methods and results were published on medical imaging than those on radiotherapy. On the other hand, in a good sense radiotherapy is subject to more uncertainties than medical imaging, since classic medical imaging technology can generally deliver decent image quality already while imaging guided radiotherapy still suffers from insufficient margin definition, mismatches between relevant images, and difficulties in predicting outcomes and optimizing therapies. Figure 1 illustrates the big picture of deep imaging towards radiotherapy, which is adapted from our perspective on deep imaging [5]. It is widely known that machine learning especially neural networks can be used for medical image analysis from images to features. What we became first interested in is tomographic image formation. Since we use deep learning techniques, we call this area data-driven, learning-based, deep tomographic reconstruction or simply 'deep recon'. The process from data to images and the one from images to features can be integrated to form a so-called 'end-to-end' workflow in the unified machine learning framework, that is, going from raw data directly to diagnosis. Since the raw data is the starting point, we call this process 'rawdiomics', instead of 'radiomics'. This perspective of tomographic reconstruction also served as the basis for the first special journal issue dedicated to the theme of 'Machine Learning for Image Reconstruction' [6]. Those features can be used to guide radiotherapy in particular and medical interventions in general.

In the following, we first give a brief introduction to deep learning architectures in section 2. We then review medical imaging in section 3 covering image reconstruction, segmentation, registration, and radiomics, and radiotherapy in section 4 covering planning, verification, and prediction. Note that we review those papers with high numbers of citations and recent key papers to reflect the state of the art in both fields. Finally, section 5 and section 6 discuss future topics and conclude this article, respectively.

## 2. Deep learning

Deep learning (DL), as a mainstream of ML, uses trainable computational models that contain multiple processing components with adjustable parameters to learn a representation of data [7]. DL methods are mainly based on artificial neural network (ANN), which were inspired by information processing in biological neural systems, but with various abstractions of and approximations to the biological counterparts. Over the past decade, researchers proposed various deep learning architectures for different tasks. Figure 2 illustrates the DL architectures widely used in medical imaging and radiotherapy areas. For a broader and more detailed introduction of DL, please refer to these DL books [8–10].

### 2.1. Multilayer perceptron

The most vanilla neural network is a multilayer perceptron (MLP), which mimics the human brain [11]. MLP is based on a collection of connected nodes called artificial neurons, which simulates biological neurons (figure 2(a)). Each link between two nodes can transmit a signal from one to the other, which plays a role

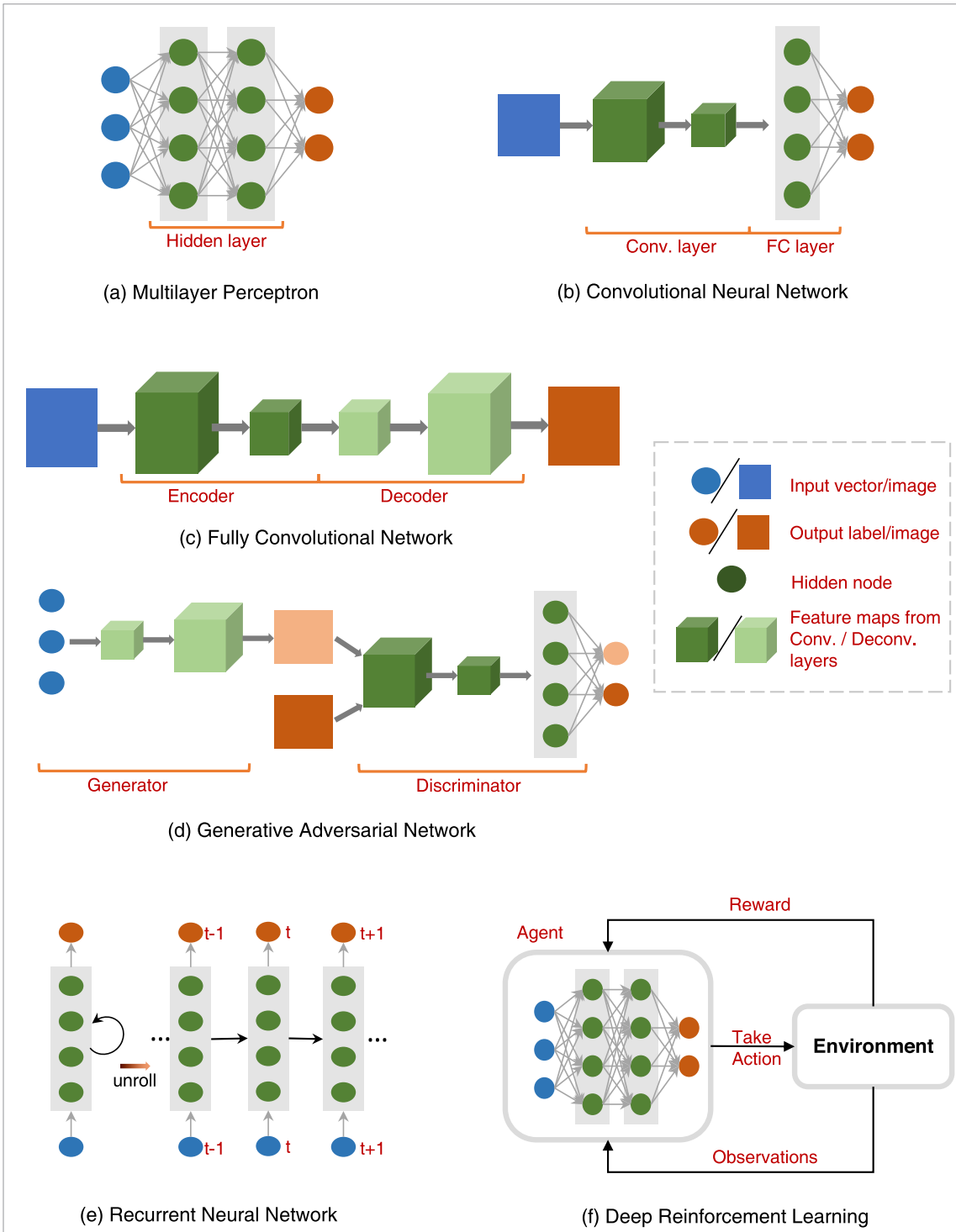


**Figure 2.** Popular deep learning architectures for medical imaging and radiotherapy.

similar to that of the synapses. A non-linear activation function, such as Sigmoid or Tanh, is used to represent the rate of action potential firing or the state of neural excitation.

The universal approximation theorem states that MLP with a single hidden layer containing a sufficiently many neurons can approximate a continuous function on compact subsets very well in a $n$-dimensional space [12]. Therefore, MLP has been used for tasks that involve a non-linear mapping, including computer vision, speech recognition, machine translation, etc. However, the fully-connected configuration would dramatically increase the number of model parameters as the problem size becomes large, bringing difficulties in training and testing.

### 2.2. Convolutional neural network

A convolutional neural network (CNN) is one of the most successful neural networks [13–15]. The CNN architecture reflects the four key ideas: local connections, shared weights, pooling, and the integration of

multiple layers [7]. A CNN consists of an input layer and an output layer, as well as multiple hidden layers [16] (figure 2(b)). The hidden layers typically consist of a series of convolutional layers convolving a local input with filters to produce the corresponding local output. The activation function is commonly a rectified linear unit (ReLU) [17], and is subsequently followed by additional operations such as pooling layers (max-pooling or average-pooling). Several fully-connected layers can be used to learn a global representation from feature maps. More precisely, those convolutional layers perform sliding dot products or cross-correlations and then non-linearly process the outcomes, while the mathematical convolutions are only linear operations.

CNN is great for capturing local features from data such as images or text, and can significantly reduce the model complexity compared to MLP. CNN is commonly used for image classification, which is to predict the label of the given input image. Note that although translational invariance introduced by pooling layers is important for object recognition, Aydore *et al* [18] pointed out that it is not applicable to brain activation images, in which meaningful structures are location specific in the brain.

### 2.3. Fully convolutional network

Just as its name implies, a fully convolutional network (FCN) is composed of convolutional layers without any fully-connected layers (figure 2(c)). Long *et al* [19] first indicates that a CNN with fully-connected layers can be implemented with all convolutional layers for semantic segmentation. A popular FCN structure is a convolutional encoder-decoder network, in which the encoder has several convolutional layers to encode an input image while the decoder has several deconvolutional layers to decode the representation. The structure of a convolutional encoder-decoder network is usually symmetric. Based on the convolutional encoder-decoder networks, variants can be obtained using different skip connections between the encoder and the decoder. For example, the residual encoder-decoder convolutional neural network (RED-CNN) with skip connection that adds the feature-maps in the encoder to later layers in the decoder [20], and U-net with skip connection that copies the feature-maps in the encoder as the input to later layers in the decoder [21]. Given the same depth and width, U-net is more flexible but it has slightly more parameters than RED-CNN.

FCN aims at learning a representation and making a decision based on its local input, which has been widely used in image-to-image translation tasks such as segmentation, registration, and image processing. Due to the use of all convolutional layers, the FCN does not depend on the input size, which means that one can train FCN on image patches and apply the trained model to full-size images.

### 2.4. Generative adversarial network

Recently, the generative adversarial network (GAN) [22] attracted a major attention, which is a powerful way to model complicated distributions of data. A GAN has a pair of two sub-networks: a generator and a discriminator (figure 2(d)). The generator takes a random noise as its input to generate a sample. The discriminator receives numerically generated and real samples, and does its best to distinguish between these two kinds of samples. This is a game between these two sub-networks: the generator learns to produce more and more realistic samples, and the discriminator needs to become smarter and smarter at distinguishing fake data from real ones. These two players are trained alternatively, and the goal is that the competition drives generated samples to be indistinguishable from real data.

In the original GAN, the input to the generator is a noise vector sampled from a predefined distribution. Now, GAN has been extended in a number of ways and applied to image processing and other tasks, playing an important role in regularizing the network outcomes. To a major extent, the adversarial loss of GAN resembles the perceptual loss that is based on a pre-trained VGG network [23], but adversarial loss is learnt from given data adaptively. GAN has rapidly evolved in recent years; see a recent survey [24]. Notably, an unpaired variant of GAN, called cycleGAN [25], attracted strong interests since it does not require paired images, which is highly desirable in many applications in which it is expensive to obtain paired images.

### 2.5. Recurrent neural network

A recurrent neural network (RNN) is a type of network in which the output from a previous step is fed into the current step (figure 2(e)). Unlike traditional neural networks whose data flows are in the feed-forward mode, RNN has a 'memory' remembering what have been seen. For example, when we predict the next word in a sentence, the previous words are required. Hence, RNN needs to remember the previous words. The most important feature of RNN is hidden states, which record some information about a sequence. It typically uses the same parameters for each input as it performs the same task to produce the output, which helps reduce the model complexity.

RNN is widely used in dealing with sequential data such as text or video. However, RNN suffers from vanishing and exploding gradients. The famous long short-term memory (LSTM) network [26] was then proposed to address the shortcomings of RNN using multiple switch gates including input, output, and

forget gates. Later on, gated recurrent units (GRU) [27] were introduced, similar to LSTM but with fewer parameters.

### 2.6. Deep reinforcement learning

Deep reinforcement learning (DRL) combines DL (function approximation) with a reinforcement learning (target optimization), which drives agents to learn the best actions to achieve their goals in a virtual environment by mapping state-action pairs to expected rewards (figure 2(f)). The combination of a deep neural network and a reinforcement learning algorithm has led to some breakthroughs such as Google Deepmind's AlphaGo [28], which beats the world champions of the Go game. In other words, DRL is a goal-oriented algorithm, which learns to achieve a complex goal by maximizing the associated objective function over many rewarding/penalizing steps. It has been widely used in video games, robotics, finance and healthcare [29].

## 3. Medical imaging

### 3.1. Image reconstruction

While many of us enthusiastically embrace the new wave of medical imaging research with deep learning, there are also doubts and concerns from some other colleagues. This conflict of options is natural and healthy. In retrospect, at the beginning of the development for analytic reconstruction, there was a major critique that given a finite number of projections, the tomographic reconstruction is not uniquely determined (ghosts) [30, 31]. Later, this was successfully addressed through regularization. When iterative reconstruction algorithms were first developed, it was observed that a reconstructed image was strongly influenced by the penalty term; in other words, it appeared that what you reconstructed is what you wanted to see; for example, in the under-determined case described in [32]. Nevertheless, by optimizing the data acquisition protocol, the reconstruction parameters, and the stopping criteria, iterative algorithms were made mature enough into commercial scanners [33]. As far as compressed sensing is concerned, it is widely known that there is a chance that a sparse solution is not the truth [34]. For example, a tumor-like structure could be introduced, or pathological vessel narrowing might be smoothed out or digitally treated if total variation is overly minimized [35]. Currently, deep learning presents issues in practice such as the black box problem, which means lack of interpretability of the successes of deep learning methods. Indeed, there is no Maxwell equations for deep learning yet, and a neural network as a black box is trained to work with big data in terms of parameter adjustment. The interpretability of neural networks remains a hot topic. Given rapid progresses in theoretical and practical aspects, we believe that machine learning algorithms will become the mainstream for medical imaging [36]. In August 2019, Tufts University organized the Conference on Modern Challenges in Imaging to celebrate the 40th anniversary of Allan Cormack's Nobel prize. Leading researchers shared their works, insights, and concerns about applying deep learning to medical image reconstruction; for more information please refer to [37].

Traditionally, tomographic image reconstruction algorithms are categorized into analytic reconstruction and iterative reconstruction [36]. Thanks to DL, now a new category of image reconstruction methods is emerging, which are ML- or DL-based [5, 6, 36, 38]. It is our point of view that in principle deep reconstruction methods ought to outperform analytic reconstruction, iterative reconstruction, and compressed sensing algorithms for tomographic image reconstruction because of the following three arguments [5, 6, 36]. First, the results from analytic and iterative algorithms can be used as the baseline, from which DL can improve image quality further in various ways; for example, the baseline and raw data can be integrated into a deep network to produce better results. Second, valuable ingredients of analytic and iterative algorithms can be used in a neural network to enhance the network instead of competing with it. Third, prior knowledge commonly used for iterative reconstruction and compressed sensing can be greatly enhanced or even replaced by neural networks equipped with much more extensive priors learnt from big data.

Table 1 summarized some recently proposed deep-learning-oriented image reconstruction algorithms, where we present those methods in terms of modality, input format, input type, network structure, skip connection, level of supervision, and if GAN is used for optimization.

**Modality:** Here we only discuss those image modalities that have been widely used in radiotherapy; *i.e.* computed tomography (CT), magnetic resonance imaging (MRI) and nuclear image techniques such as positron emission tomography (PET) and single-photon emission computed tomography (SPECT). Among those modalities, CT and MRI are two extensively studied modalities for image reconstruction due to the public availability of low dose CT and fast MRI datasets. Originally designed for a low dose CT reconstruction challenge, the Mayo low dose CT dataset contains perfectly registered low dose and normal dose CT scans, and has became a benchmark dataset in deep CT image reconstruction [39]. Deep fast MRI aims at accelerating MR imaging with AI, gaining increasingly attention [40, 41]. Nuclear imaging

**Table 1.** Recently proposed deep-learning-based image reconstruction methods.

| | | Input Format | | Input | | Skip Connection | | Supervision Level | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Ref | Modality | Img. | Raw | Type | Network Structure | Sum | Conc. | Sup. | Unsup. | GAN |
| [44] | CT | √ | | 2D | FCN | | | √ | | |
| [20] | CT | √ | | 2D | FCN(E-D) | √ | | √ | | |
| [45] | CT | √ | | 2D | FCN | | | √ | | √ |
| [42] | CT | √ | | 2.5D | FCN(E-D) | | √ | √ | | √ |
| [46] | CT | √ | | 2D | FCN | | | √ | | |
| [47] | CT | √ | | 2D | FCN(E-D) | | √ | √ | | √ |
| [48] | CT | √ | | 2D | FCN | √ | √ | √ | | |
| [49] | CT | √ | | 3D | FCN | √ | | √ | | √ |
| [50] | CT | √ | | 3D | FCN | | | √ | | √ |
| [51] | CT | | √ | 2D | FCN | √ | | √ | | |
| [52] | CT | | √ | 2D | CNN | | √ | √ | | √ |
| [53] | CT | | √ | 2D | FCN | | | √ | | |
| [54] | CT | | √ | 2D | CNN | | | √ | | |
| [55] | CT | | √ | 2D | CNN/FCN(E-D) | √ | | √ | | √ |
| [56] | CT/MRI | √ | | 2D | FCN(E-D) | | √ | | √ | |
| [57] | CT | √ | | 2D | FCN | √ | √ | | √ | √ |
| [58] | CT | √ | | 2D | FCN(E-D) | √ | √ | √ | | |
| [59] | MRI | | √ | 2D | FCN | √ | | √ | | |
| [60] | MRI | | √ | 2D | CNN | | | √ | | |
| [61] | MRI | | √ | 2D | FCN | √ | | √ | | |
| [62] | MRI | √ | | 2D | FCN(E-D) | √ | | √ | | √ |
| [63] | MRI | √ | | 2D | FCN(E-D) | | √ | √ | | |
| [64] | MRI | √ | | 2D | FCN | √ | | √ | | √ |
| [65] | MRI | √ | | 2D | FCN(E-D) | √ | √ | √ | | √ |
| [66] | PET | √ | | 2D | FCN | | | √ | | |
| [67] | PET | √ | | 2D | MLP | | | √ | | |
| [68] | PET | √ | | 3D | FCN(E-D) | | √ | | √ | |
| [69] | PET | | √ | 3D | FCN(E-D) | | √ | √ | | |
| [70] | SPECT | | √ | 2D | CNN/FCN(E-D) | | | √ | | |

techniques such as PET and SPECT are also instrumental for modern radiotherapy, particularly to detect and characterize tumors before and after treatment such as radiotherapy and immunotherapy.

**Input Format:** DL-based image reconstruction can be done in two ways: directly mapping from raw data to a tomographic image and image post-processing from a reconstructed image to an improved version. The raw data are in a sinogram data for CT/PET/SPECT and *k*-space data for MRI. Directly reconstructing image from raw data requires the deep learning model to learn the physics process involved in the reconstruction (the current models such as Radon transformation or Fourier transform are only approximations) and utilize information on the image content hidden in training data. Alternatively, the image post-processing is relatively simple yet effective, after a low-quality image is already reconstructed using a classic reconstruction algorithm such as filtered back-projection or Fourier transform. We note that the raw data are usually not available to researchers due to restrictions by vendors and anticipate that the deep learning model based on raw data should outperform the model based on reconstructed images.

**Input Type:** Currently, the input type to deep learning models is mainly 2D due to limited computing resources. However, 3D image post-processing techniques have been proved to give superior results [42, 43]. For raw data, it is also desirable to do 3D reconstruction through deep learning, given sufficient GPU memory.

**Network Structure:** The network structures are roughly divided into MLP, CNN, FCN and FCN(E-D). Here FCN(E-D) represents the convolutional encoder-decoder network, which is one of popular FCN architectures. For image post-processing, FCN or FCN(E-D) are typically used to learn an image-to-image mapping. For directly mapping from raw data to a tomographic image, most methods use FCN and FCN(E-D) in the image domain processing even though some methods use fully-connected layers to learn a domain transfer from raw data to images.

**Skip Connection:** Skip connections are widely used to enhance the network flexibility and improve the model performance. In the so-called residual dense block, it is remarkable that skip connections are wired in a sophisticated fashion to increase the model capability.

**Supervision Level:** Most deep image reconstruction methods are based on supervised learning, which differs from traditional reconstruction methods that are unsupervised. Deep image post-processing methods usually require the paired images to learn a mapping from low-quality to high-quality images. These paired images are usually unavailable in practice. In this case, the numerical noise insertion can be used for low-dose CT; for example, the Mayo low-dose CT dataset was simulated to produce quarter dose counterparts from the normal dose sinograms. Some recent methods attempt to address the reconstruction process in an unsupervised learning mode, avoiding or weakening the need for paired data, which represents a current hot topic.
**GAN:** For image reconstruction, GAN is widely used to enhance the image quality. The benefits from GAN is that it can introduce a data-driven regularizer to ensure that the learnt distribution approaches the ground-truth. That is, tomographic images are generated or reconstructed as faithfully as a numerical observer or discriminator can describe.

In order to train a deep recon model, an appropriate objective function should be chosen. The widely-used losses for image reconstruction include mean-squared error, mean absolute error, structure similarity [71], perceptual loss [72], adversarial loss [22], and so on. The combination of these losses depends on specific tasks and even datasets. For example, the mean-squared error can reduce image noise, but it tends to over-smoothen images. The adversarial loss enhances the image quality with a discriminator but it requires more training efforts. For more discussions, please refer to [73].

To validate the model, quantitative image quality metrics such as peak signal-to-noise ratio, structure similarity, root mean square error are typical employed to compute the difference between the generated and ground-truth images. However, a higher quantitative metric does not always mean a better diagnostic performance. Therefore, task-specific measures are clinically most important, such as receiver operating characteristic (ROC) and area under curve (AUC) obtained in a human reader or numerical observer study. Neural networks were recently developed to perform reader studies [74, 75].

### 3.2. Image segmentation

Medical image segmentation aims to delineate the boundary of an organ or lesion of interest in medical images. In radiotherapy, segmentation plays a crucial role, since it is needed to identify the target lesion and avoid healthy tissues during treatment. Currently, the treatment target and normal structures are commonly delineated by oncologists, often slice by slice, which is time-consuming and cumbersome. In addition, the boundaries delineated by different oncologists may vary significantly. Even for the same oncologist, he/she may not be able to reproduce his/her own delineation very well [76–78]. Therefore, automated image segmentation has been a hot research area over the past decade.

Traditional automatic segmentation methods for radiotherapy are based on analysis of image content and properties such as voxel intensities, gradients, and textures [79]. Based on the information involved, traditional automatic segmentation methods can be divided into the following categories: region-based methods [80–82], edge detection-based methods [83–85], atlas-based methods [86–90], statistical-models-based methods [91–94], and machine learning methods [95–98]. Among these methods, machine learning methods produce promising results as they learn the image prior in a data-driven manner. These methods used to be based on traditional machine learning algorithms such as support vector machine (SVM) [99], random forest [100], and Gaussian process [101], which suffer from insufficient capability and unsatisfied performance for clinical tasks.

Deep learning, as the mainstream of machine learning research, is now shaping this area rapidly. Different from the traditional machine learning algorithms, deep learning allows computational models that are composed of multiple layers to learn a representation of data and knowledge from big data [7]. As a result, deep learning methods can dramatically improve the state-of-the-art in the image segmentation field, showing a great potential to facilitate radiotherapy.

In this section, we survey the deep learning-based segmentation methods developed in recent years as summarized in table 2. For clarity, we will discuss those methods in terms of organ/tumor, modality, input, network, skip connection, supervision, and training data. We also make some notes in the last column.
**Organ/Tumor:** Segmentation of various organs and tumors has been widely reported in the literature, such as segmentation of brain tumors, rectal tumors, liver tumors, etc. Among those reports, brain image segmentation seems to be the most popular research topic, partially due to the well-organized and publicly available datasets through the reputable competitions such as Brain Tumor Segmentation (BRATS) challenges from 2012 to 2018 and Ischemic Stroke Lesion Segmentation from 2015 to 2018.
**Modality:** In radiotherapy, imaging modalities are selected for diagnostics, treatment planning, and follow-up. Such medical images are critical to identify treatment targets and to spare normal tissues/organs from significant radiation damage. CT and MRI are two popular imaging modalities used in this context. CT can be used to estimate electronic density, which enables dosimetric calculation in radiotherapy. For soft

**Table 2.** Recently proposed deep-learning-based segmentation methods.

| Ref | Tumor/Organ | Modality | Input Type | Network Structure | Skip Connection | | Supervision Level | # Training | Note |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | Sum | Conc. | | | |
| [102] | Rectal Tumor | MRI | 2D | CNN | | | Fully | 70 | |
| [103] | Renal Tumor | CT | 3D | FCN(E-D) | | √ | Fully | 89 | Dice coefficient loss |
| [104] | Portal vein | CT | 2.5D | FCN | | | Fully | 64 | Used with MRF |
| [105] | Brain Tumor | MRI | 3D | FCN(E-D) | √ | √ | Fully | 285 | Ensemble, multi-label |
| [106] | Brain Tumor | MRI | 3D | CNN | | | Fully | 48 | Used with CRF |
| [107] | Brain Tumor | MRI | 3D | FCN | | | Fully | 30 | Local and global feature, cascaded |
| [108] | Brain Tumor | MRI | 3D | FCN | √ | | Fully | 274 | |
| [109] | Brain Tumor | MRI | 2D | CNN | | | Fully | 10 | Holistically nested |
| [110] | Brain Tumor | MRI | 3D | CNN | | | Fully | 265 | Separate validation |
| [111] | Brain Tumor | MRI | 2.5D | CNN | | | Fully | 15 | Multi-source |
| [112] | Hippocampus | MRI | 3D | FCN(E-D) | | √ | Fully | 637 | Multi-task |
| [113] | Headneck | CT | 3D | CNN | | | Fully | 33 | Interleaved network |
| [114] | Brain Tumor | MRI | 2.5D | CNN | | | Fully | 15 | Atlas probability |
| [115] | Breast/fibroglandular | MRI | 2D | FCN(E-D) | | √ | Fully | 39 | Two consecutive U-nets |
| [116] | Brain structures | MRI/US | 2D | CNN | | | Fully | 40/45 | Hough voting, multi-modality |
| [117] | Prostate | MRI | 3D | FCN(E-D) | √ | √ | Fully | 50 | |
| [118] | Multi-organ | CT | 2.5D | FCN(E-D) | √ | | Fully | 230 | |
| [119] | Prostate | CT | 2D | CNN | | | Fully | 73 | Refined with Atlases |
| [120] | Thoracic organs | CT | 2D | RNN/FCN(E-D) | | √ | Fully | 25 | CRF as RNN |
| [121] | Thoracic organs | CT | 3D | FCN(E-D) | | √ | Fully | 25 | Collaborative architectures |
| [122] | Multi-organ | CT | 3D | FCN | | | Fully | 112 | |
| [123] | Spine | CT | 2D | FCN | | | Fully | 32 | Redundant class label |
| [124] | Liver tumor | CT | 2D | FCN(E-D) | | | Fully | 130 | Hierarchical |
| [125] | Head Neck | CT | 2D | CNN | | | Fully | 40 | |
| [126] | Meningiomas | MRI | 3D | FCN | √ | | Fully | 249 | |
| [127] | Lung Tumor | PET | 3D | FCN(E-D) | | √ | Fully | 50 | |
| [128] | Head Neck | CT | 3D | FCN(E-D) | √ | √ | Fully | 261 | Squeeze-and-excitation |
| [129] | Multi-organ | CT | 3D | FCN(E-D) | | √ | Fully | 48 | Dose calculation |
| [130] | Brain Tumor | MRI | 2D | RNN/FCN(E-D) | | √ | Fully | 274 | CRF as RNN |
| [131] | Multi-organ/brain | MRI | 2D | FCN | √ | √ | Fully | 10/54 | Interactive |
| [132] | Brain | MRI | 3D | FCN | √ | | Fully | 5 | multi-modality/level information |
| [133] | Liver | CT | 2D | FCN(E-D) | √ | | Fully | 131 | Squeeze-and-excitation |
| [134] | Prostate | MRI | 3D | FCN(E-D) | √ | √ | Fully | 50 | Boundary-weighted domain adaptive |
| [135] | Head Neck | PET | 3D | FCN(E-D) | | √ | Weakly | 47 | Using Bounding box |

tissue anatomies, MRI is more advantageous, depicting both anatomy and pathology of brain, prostate, and so on. A future research direction is to combine CT and MRI for radiotherapy [136, 137] or to synthesize one from another via deep learning [138, 139].

**Input Type:** The input to networks for segmentation can be classified into three types based on the dimensionality of images: 2D, 2.5D, and 3D. The use of 2D images as input allows more training examples, which is helpful in a case of limited data but these images miss 3D spatial information relevant for organ/lesion segmentation. The use of 3D input keeps the 3D spatial information and promises improved segmentation quality but the training process needs a large memory and a longer computational time. Usually, the more patients are involved for training, the more costly is the development effort. As a compromise between 2D and 3D inputs, 2.5D input can be used, which feeds multiple 2D slices from a 3D volume into a neural network. It is computationally efficient and yet utilizes 3D spatial information to a good degree. To boost a training set, a practical way is to use overlapping patches to train a deep learning model, and the testing set should not intersect with the training and validation sets.

**Network Structure:** The basic network architectures are roughly classified into the following three categories: CNN, FCN or FCN(E-D), and RNN. Previously, the segmentation problem was treated as a classification problem, where the input is an image patch, and the corresponding label is the class. Then, attempts were made to generate a segmentation map directly from the input using either fully connected, convolutional, or encoder-decoder networks. Furthermore, the recurrent neural network was also introduced to learn the spatial relationship among features extracted from CNN [140]. It should be noted that a capsule network [141], which is a new modularized architecture, was also applied to image segmentation [142].

**Skip Connection:** In addition to the typical (de)convolutional layers, skip connections play an important role in sharing information and increasing the performance of networks. Two common operations followed by a skip connection are the summation and concatenation. Skip connection with a summation operation, also called a residual skip connection, enables the network to learn the residual information between input and output, and dramatically increase the depth of network without suffering from the gradient vanishing problem [14]. Different from the residual skip connection, a concatenation skip connection copies feature-maps from earlier layers and reuse them as the input to later layers, which carries more information than the residual skip connection. The encoder-decoder network is coupled with the concatenation skip connection to form a powerful network called U-net [21].

**Supervision Level:** Most of the deep learning-based segmentation methods are trained in a fully-supervised manner, where an ideal or very accurate segmentation map of target tumors or organs are provided. However, the manual delineation of a large dataset is time-consuming and subjective with large inter- and intra-expert variablities. To address this challenge, weakly supervised or semi-supervised segmentation methods are worth of exploration. Weakly-supervised methods may not have the accurate label, but they have alternative labels that can be annotated relatively easily. For example, the precise segmented maps can be approximated by bounding boxes [135] and extreme points [143]. Semi-supervised learning only relies on labels for a few samples, but does not have labels for the other samples.

**Training Data:** A moderately sized training dataset is necessary to train a good segmentation model. As summarized in table 2, most studies take 20~300 patient scans for training a segmentation model. In a case of lacking sufficient training data, transfer learning or domain adaption can be used to transfer knowledge from related data. Alternatively, it is viable to reduce the model size to avoid over-fitting a small dataset.

The loss functions for optimizing a segmentation model can be roughly divided into two classes: cross-entropy- and dice-based measures. There are variants such as weighted cross-entropy, focal loss [144], and Tversky loss [145]. Please refer to [146] for a taxonomy in the context of medical image segmentation. To validate a trained model, image metrics such as intersection, pixel accuracy, Hausdorff distance, and dice similarity are often used. A good reference is Taha *et al* [147] which describes 20 image metrics widely used in the image segmentation community.

### 3.3. Image registration

Image registration is the process of transforming different types of images into a common coordinate system, where information gained from two or more images is usually complementary [148, 149]. Image registration plays a key role in many tasks such as multi-modality imaging, adaptive treatment planning, image-guided radiotherapy, and prognostic assessment [150]. With the rapid development of deep learning techniques, deep-learning-based methods are being constantly developed, refreshing the landscape of image registration research.

Here we highlight some recently proposed deep learning-based image registration methods in table 3, including both supervised and unsupervised methods. Many iterative methods are not included due to their inferior performance and slow speed. We will discuss the selected deep-learning-based registration methods in terms of label, transformation, input/output, modality, body region, and network architecture.

**Table 3.** Recently proposed deep-learning-based registration methods.

| Ref | Label | Transformation | Input/Output | Modality | Region | Network Structure |
|---|---|---|---|---|---|---|
| [154] | Synthetic | Rigid | 2D/3D | DRR/x-ray | Bone | CNN |
| [155] | Synthetic | Rigid | 2D/3D | CT | Thorax | SVRNet |
| [156] | Synthetic | Rigid | 2D/3D | x-ray | Bone | CNN |
| [157] | Synthetic | Rigid | 2D/2D | MRI | Brain | CNN/FCN |
| [158] | Synthetic | Rigid | 3D/3D | MRI | Brain | AIRNet |
| [159] | Synthetic | Rigid | 2D/2D | MR/TRUS | Prostate | GAN |
| [160] | Synthetic | Deformable | 3D/3D | MRI | Brain | CNN |
| [161] | Synthetic | Deformable | 3D/3D | CT | Chest | RegNet |
| [162] | Synthetic | Deformable | 3D/3D | CT/US | Liver | DVNet |
| [163] | Synthetic | Deformable | 2D/2D | MRI | Brain/Cardiac | FlowNet |
| [164] | Synthetic | Deformable | 3D/3D | CT | Chest | U-Net-Advanced |
| [152] | Real | Deformable | 2D/2D | MRI | Abdominal | CNN |
| [151] | Real | Deformable | 3D/3D | MRI | Brain | FCN |
| [165] | Real | Deformable | 3D/3D | MRI | Brain | VoxelMorph CNN |
| [166] | N/A | Deformable | 2D/3D | MRI | Brain | ICNet |
| [167] | N/A | Deformable | 3D/3D | MRI | Brain | 3D UNet |
| [168] | N/A | Deformable | 3D/3D | MRI | Brain | VoxelMorph CNN |
| [153] | N/A | Deformable | 3D/3D | CT | Liver | CycleGAN |

**Label:** The ground-truth for image registration is usually difficult to obtain clinically. Thus, in most studies synthetic ground-truth data were used to train and validate registration methods. The real ground-truth may be available for special tasks, such as registration to atlas [151] or respiratory motion correction [152]. Due to the difficulty in acquisition of ground-truth, unsupervised image registration methods attract much attention using an unpaired training strategy, such as cycleGAN [153].
**Transformation:** The transformation methods used for image registration can be either rigid or deformable. Rigid transformation is a geometric transformation in a Euclidean space that preserves the Euclidean distance between every pair of points, which includes rotation, translation, reflections or their combination. However, in practice patients may have anatomical features changed non-rigidly, due to weight loss, tumor shrinkage, and/or physiological variation, which cannot be modeled with rigid transformation. In contrast, deformable transformation has a great degree of freedom to establish the correspondence for key points before and after deformation [169].
**Input/Output:** The input and output of the network are usually of the same dimensionality; for example, in the registration between two 2D images [157] or two 3D volumes [158]. However, special cases exist, such as a 2D/3D registration problem, which means, for example, finding a best match between one or more intra-operative x-ray projections of the patient and the preoperative 3D volume [154, 170].
**Modality:** CT and MRI remain two most popular modalities being studied as they are widely used in clinical routines. The registration happens for either the same modality [153, 160] or across multiple modalities such as CT and MRI. Other imaging modalities also have registration needs; for example, ultrasound (US) to CT image registration [162] and MRI to transrectal ultrasound (TRUS) image registration [171].
**Organ:** Similar to the image segmentation methods, most image registration methods focus on brain images due to importance of the brain and public availability of large datasets and atlas. For interventional guidance applications, the targeted organs include the prostate, the liver, and the lungs.
**Network Structure:** Popular network structures are all used for image registration. Also, the type of networks depends on the transformation in use. For rigid transformation, typically CNNs are used to learn the parameters. For deformable transformation, a FCN or FCN(E-D) can be used to model an underlying deformation field. Notably, the GAN has been used to enhance the registration performance in supervised and unsupervised learning [153, 159].

When optimizing the rigid registration model, mean-squared error is commonly used. For the deformable registration model, an image similarity metric is typically preferred, including the intensity sum of squared distance (SSD), mean squared distance (MSD), correlation ratio (CR), (normalized) cross-correlation (CC/NCC), (normalized) mutual information (MI/NMI), etc [172]. Likewise, those image similarity metrics can be used to evaluate the trained registration models.

### 3.4. Radiomics

Radiomics refers to extraction and analysis of comprehensive features in medical images such as from low dose CT [173–175]. The key idea behind radiomics is that images contain more information than what can be visualized by radiologists, and sophisticated algorithms can be designed to distill hidden information.

**Table 4.** Recently proposed deep-learning-based radiomics.

| Ref | Cancer | Modality | Input | | Training | | Handcrafted | Network | | Task | | Notes |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | 2D | 3D | Transfer | Scratch | features | CNN | E-D | Single | Multi | |
| [184] | Brain | MRI | √ | | √ | | √ | √ | | √ | | multi-scale |
| [185] | Brain | MRI | √ | | √ | | √ | √ | | √ | | |
| [186] | Brain | MRI | √ | | √ | | | √ | | √ | | |
| [187] | Brain | MRI | | √ | √ | | | √ | | √ | | |
| [188] | Lung | CT | | √ | | √ | | √ | | √ | | |
| [189] | Lung | CT | | √ | | √ | | √ | √ | | √ | |
| [190] | Lung | CT | | | √ | | | √ | | √ | | |
| [191] | Lung | CT | √ | | √ | | | √ | | √ | | |
| [192] | Lung | Histopathology | √ | | √ | | | √ | | √ | | |
| [193] | Lung | CT | | √ | | √ | | √ | √ | √ | | |
| [194] | Lung | CT | | √ | √ | | | √ | | | √ | |
| [195] | Lung | CT | | √ | | | | √ | | | | multi-scale |
| [196] | Breast | Mammography | | | √ | | √ | √ | | √ | | |
| [197] | Breast | Pathology | √ | | | √ | | √ | | √ | | |
| [198] | Breast | Mammography | √ | | | √ | | √ | | √ | √ | multi-scale, multi-view |
| [199] | Breast | Mammography | | | √ | √ | | √ | | √ | | multi-scale, multi-view |
| [200] | Renal | CT | √ | | √ | | | √ | | √ | | multi-view |
| [201] | Retinal | Retinal | √ | | | √ | | √ | | √ | | |
| [182] | Retinal | Retinal | √ | | | √ | | √ | | √ | | |
| [202] | Prostate | MRI | √ | | | √ | | √ | | √ | | |
| [203] | Cervical | Histology | √ | | | √ | √ | √ | | √ | | |
| [204] | Head and neck | Hyperspectral | √ | | | √ | | √ | | √ | | |
| [205] | Bladder | CT | √ | | | √ | | √ | | √ | | |

Traditional handcrafted features can be divided into shape-based features and texture-based features. Shape-based features describe a lesion of interest heuristically; for example, the total volume, surface area, surface-to-volume ratio, and lesion compactness. Texture-based features include second-order statistics, length of run, co-occurrence matrix, histogram of oriented gradients, local binary pattern, and so on [176–180]. Different from the handcrafted features, CNNs promise comprehensive multi-scale features for deep radiomics. Currently, deep learning methods handle millions (even billions) of parameters [13, 14, 23], resulting in an ultra-high dimensional representation. In a practical sense, deep learning discovers intricate features in large datasets to deliver an unprecedented power for representation, classification and prediction.

Table 4 summarizes the recently proposed deep-learning-based radiomics results that are highly related to radiotherapy. For clarity, we cover the following aspects: cancer type, modality, input type, training strategy, involvement of handcrafted features, network structure, and single or multiple tasks.

**Cancer Type:** Radiotherapy is a cancer treatment approach that uses high dose radiation to kill cancer cells in the brain, head and neck, breast, cervix, prostate, eyes, and livers.

**Modality:** Various modalities are used to diagnose different cancers. For example, brain cancers are diagnosed by MRI, lung cancers by CT, breast cancers mainly by mamography, tomosynthesis, CT, MRI and ultrasound imaging, and retinal tumors by MRI and optical imaging. Also, algorithms can extract information from pathological images. The combination of different modalities provides complementary radiomics information.

**Input Type:** The input to deep learning algorithms includes 2D/3D patches or images/volumes of interest. However, 3D cubes usually requires a much larger memory than 2D patches.

**Training Strategy:** The deep learning model can be trained based on a pre-trained model (transfer learning) or from scratch. Considering limited data in many cases for training a large network, the transfer learning strategy is critically important to transfer the knowledge from one domain such as natural images in ImageNet [181] to another domain, in particular medical images. However, recently Raghu *et al* [182] showed that the transfer learning strategy could only offer a limited performance gain while much smaller architectures can perform comparably to the standard model. Further investigations are clearly needed.

**Involvement of Handcrafted Features:** Although the features learnt from deep neural networks are rich, the combination of deep features and handcrafted features seems advantageous in some applications; for example, with those handcrafted features that are widely used by radiologists.

**Network Structure:** CNN becomes a dominating network architecture for radiomics as it can be viewed as a classification task in computer vision. CNN networks such as ResNet [14] and DenseNet [15] can be thus adapted for radiomics. Some networks contain a decoder so that the extracted features can not only predict labels but also reconstruct the original signals. Afshar *et al* recently adapted the capsule network [141] for brain tumor radiomics [183].

**Single or Multiple Tasks:** Most algorithms solve a single task. On the other hand, multi-task algorithms were also developed to solve multiple tasks at the same time and capitalize synergies among the tasks. By doing so, the prediction accuracy and robustness can be improved.

Note that some algorithms used multi-view or multi-scale analysis of a tumor. Multi-view data offer different perspectives of the tumor in a 3D space, while multi-scale features allow a structured understanding of the tumor. They could be also viewed as ways to augment the training data and facilitate ensemble learning.

Since radiomics means classification tasks, the classification losses such as cross-entropy or focal loss can be used to train the model. Note that focal loss addresses the imbalanced classification [144]. To validate the trained model, accuracy, precision, recall, F1 score, ROC and AUC are well-justified metrics to evaluate the performance.

Now, let us revisit the concept mentioned at the beginning – 'rawdiomics' as shown in figure 1. With machine learning, radiomics has gained a new momentum. The so-called radiomics is comprehensive image analysis, or a process from an image to a list of comprehensive features that are hands-crafted and/or network-extracted. What we propose is to generalize the concept of radiomics to a new concept of rawdiomics that goes from raw tomographic data to final features, with or without explicitly reconstructed images [206–208]. For example, we could reconstruct complementary images using different algorithms from the same raw dataset, and perform radiomics of all these images systematically. In this way, each reconstructed image is a different feature channel so that the space of features is significantly enlarged. Alternatively, diagnostic features and decisions can be directly mined from raw data [209].

## 4. Radiotherapy

Application of deep learning techniques is an emerging trend in the field of radiotherapy. This section, let us review recent advancements of machine learning, particularly deep learning, along the chain of radiotherapy treatment planning and delivery steps, including treatment plan optimization, plan quality assurance, treatment delivery monitoring, and outcome prediction. As it is impossible to cover all the progresses, we will point out representative major studies in each aspect. Interested readers can find more studies in the literature.

### 4.1. Treatment planning

The typical process of treatment planning starts from the acquisition of planning images, such as CT, MRI, and PET images. These images are registered together (see section 3.3 on image registration), and then the target volumes and organs at risks are delineated manually or automatically (see section 3.2 on image segmentation.) The goal of treatment planning is to generate an optimal plan for an individual patient to meet clinical requirements, *i.e.* achieving sufficient tumor coverage of dose while keeping normal organ doses to an acceptable level or minimized.

Treatment planning techniques for modern radiotherapy, *e.g.* intensity modulated radiotherapy (IMRT) [81] or volumetric modulated arc therapy (VMAT) [210], are usually formulated as an optimization problem. The objective function for optimization has multiple terms and constrains corresponding to various clinical or practical considerations. A number of parameters exist to precisely define these terms and constraints, such as their priorities. A treatment planning system is capable of solving the optimization problem with a certain optimization algorithm. Nonetheless, deciding the exact parameter values is typically beyond the capability of the treatment planning system. Because the parameter values for the optimal solution are patient-specific, and the optimal solution for a patient cannot be perfectly known in a treatment planning practice, a planner must adjust these parameters in a trial-and-error fashion based on his/her experience to explore the solution space. Moreover, since the physician ultimately decides acceptance of a plan, the planner has to consult the physician about plan quality, as well as directions to improve the plan if the quality is not acceptable. Not only is this process tedious and time-consuming, the final plan quality is affected by many factors, such as experience of the planner, available time, and interactions between the planner and the physician. Hence, extensive efforts are being devoted to the automation of this process for high-quality plans meeting the physician's requirements in a timely fashion.

To provide the guidance about the best achievable plan for each individual patient, the regime of knowledge based planning is necessary [211–217]. Methods in this category use a certain machine learning

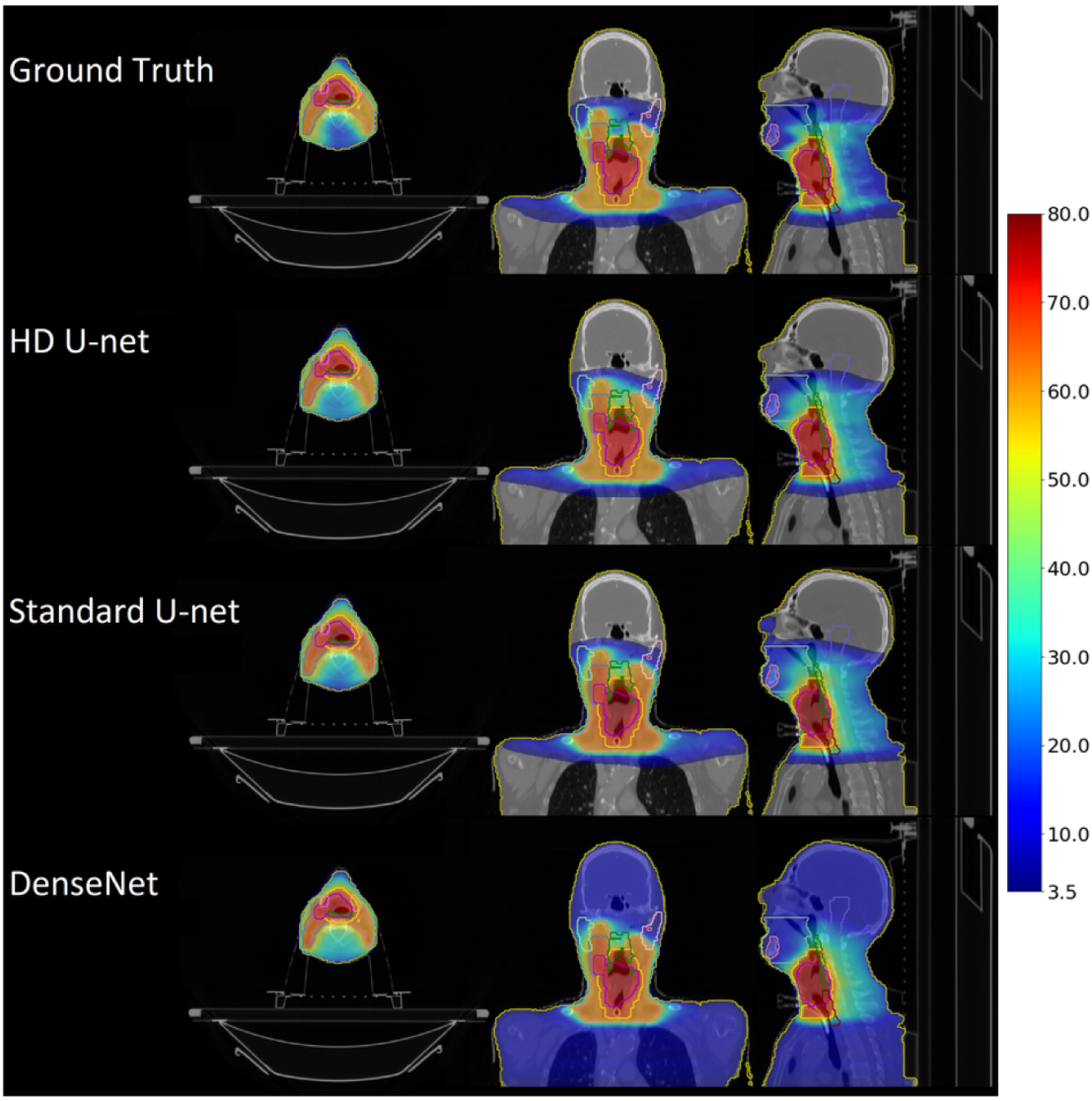


**Figure 3.** Dose prediction for treating head and neck cancer. The color bar is shown in the unit of Gy. The clinical ground truth dose is on the top row, followed by the dose predictions of the hierarchically densely connected U-net (HD U-net), Standard U-net, and DenseNet, respectively. Low dose cutoff for viewing was chosen to be 5% of the highest prescription dose (3.5 Gy). Note that all of the models predict more dose on the back of the neck than the ground-truth, which may be caused by insufficient training. Adapted with permission from [223]. Copyright 2019 IOP Publishing.

or deep learning algorithm to establish a correlation between patient anatomy and the best achievable dose volume histogram (DVH), a statistical summary of a dose distribution commonly used to evaluate a plan. Along the same direction and empowered by the capability of image-to-image mapping of deep learning methods, recent studies demonstrated the feasibility to directly predict the best achievable dose distribution, as opposed to predicting its statistical summary of DVH. Kajikawa *et al* first employed an AlexNet CNN [13] to determine if the plan for a patient with prostate cancer can meet all the dosimetric constraints in treatment planning [218]. Subsequently, a number of groups successfully used different CNNs to directly map contours of targets and organs to dose distributions in different tumor sites [219–223]. The result from an example study [223] is shown in figure 3. Lately, the feasibility is also shown for special radiotherapy modalities such as helical therapy [224]. Advanced features modeling the dose distribution, such as isodose feature-preserving voxelization, were incorporated into the CNN-based prediction model to improve model accuracy and reliability [225, 226]. These models are expected to guide the planner towards the best plan, which makes the planning process more effective and efficient.

A human planner usually has the intuition to adjust the parameters in the optimization problem. In light of recent successes in DRL, to build a computer agent for task-specific decision making in a human-like fashion, a new framework called intelligent treatment planning was recently proposed [28, 227, 228]. This type of research aimed at building a virtual treatment planner using DRL to mimic the human's behavior of parameter adjustment in treatment planning. Shen *et al* demonstrated the feasibility and potential of this

approach in a proof-of-principle study of inverse treatment planning for high dose-rate brachytherapy [229], a special form of radiotherapy. Extensions to treatment planning for external beam radiotherapy are currently in progress [230].

### 4.2. Quality assurance

To ease this quality assurance (QA) process, deep learning techniques have been utilized to perform virtual QA on plans by identifying those plans that fail the QA process. Interian *et al* [231] and Tomori *et al* [232] employed CNN models to predict a metric to quantify the success of QA, called gamma passing rate [233]. One step further, Nyflot *et al* developed a model to identify errors in the LINAC's multi-leaf collimator based on the measured beam fluence map [234]. The efforts of using deep learning techniques for the QA purpose were also made for proton radiotherapy. A model was developed to predict output factors of proton therapy treatment fields, which can be used for the sanity check of output factor measurements [235].

### 4.3. Treatment delivery

At the treatment delivery stage, it is critical to position the patient against the treatment beam following the planned geometry and to ensure accuracy of targeting the tumor with the beam under practical challenges, such as tumor motion due to respiration. Again, deep learning techniques are being developed to solve these problems.

On the positioning side, certain images of a patient has to be acquired, which should be registered to the treatment planning images to decide the correct patient position. For instance, x-ray projections of the patient anatomy using kilo-voltage (kV) or Mega-voltage (MV) beams are taken, and compared with projections computed using the treatment planning CT volume. Zhao *et al* developed a CNN method to identify the prostate target on a kV x-ray projection image, permitting accurate positioning of the target with respect to the therapy beam [236], as shown in figure 4. When comparing MV x-ray projections acquired at treatment stage with kV projections computed at the planning stage, different image characteristics, *e.g.* inconsistencies in intensity and contrast, caused by different radiation energies impede the image matching quality. To overcome this challenge, a conditional GAN approach was developed to convert MV images into kV images, which facilitates the MV-kV image matching [237].

During the treatment delivery, it is necessary to monitor the tumor and patient motion to ensure beam targeting accuracy and patient safety. For this purpose, Chen *et al* used a CNN to automatically select the patient surface area that can be monitored under a surface image camera to yield the optimal motion monitoring performance [238]. Park *et al* developed an intra- and inter-fraction fuzzy deep learning (IIFDL) method to predict lung tumor motion inter- and intra-fractionally [239]. Using an attention-aware CNN with a convolutional LSTM network, real-time motion tracking of a liver tumor was made possible via a robotic-arm-mounted ultrasound imaging system [240]. Lin *et al* also developed a LSTM-based model to predict the patient respiratory signal in real-time [241].

### 4.4. Biological effects

Deep learning can predict various biological or clinical quantities to facilitate treatment planning or treatment outcome assessment for different disease sites and treatment techniques. For example, DNNs were used to extract prognostic signatures of quantitative imaging features that can stratify patients with non-small-cell lung cancer into low and high mortality risk groups [188]. The model was developed with 7 independent datasets across 5 institutions to ensure its robustness. Peeken *et al* used deep learning based free water correction of diffusion tensor imaging scans to estimate the infiltrative gross tumor volume of patients with glioblastoma, which is the basis for the definition of the treatment target of this disease [242]. Lee *et al* proposed a survival recurrent network for patients with gastric cancer, and established the association between molecular subtype of the disease and optimal adjuvant treatment [243]. Tseng *et al* employed a DRL approach to enable the adjustment of radiation dose during the course of radiotherapy for non-small cell lung cancer, maximizing tumor local control at a reduced rate of radiation pneumonitis [244]. These studies have demonstrated the power of deep learning in terms of providing critical information to support decision-making on treatment strategies.

### 4.5. Outcome prediction

Wang *et al* developed a deep learning approach to predict the spatial and temporal evolution of lung tumor during the course of radiotherapy using longitudinal MRI scans [245]. Cui *et al* designed a network structure to take advantage of the temporal associations among longitudinal data to predict local control of lung cancer after radiotherapy [246]. In terms of treatment toxicity, CNN was applied to discover the dosimetric patterns in treatment plans associated with hepatobiliary toxicities after liver stereotactic body radiotherapy [247]. The similar technique was also used for rectal toxicity prediction for patients with cervical cancer receiving

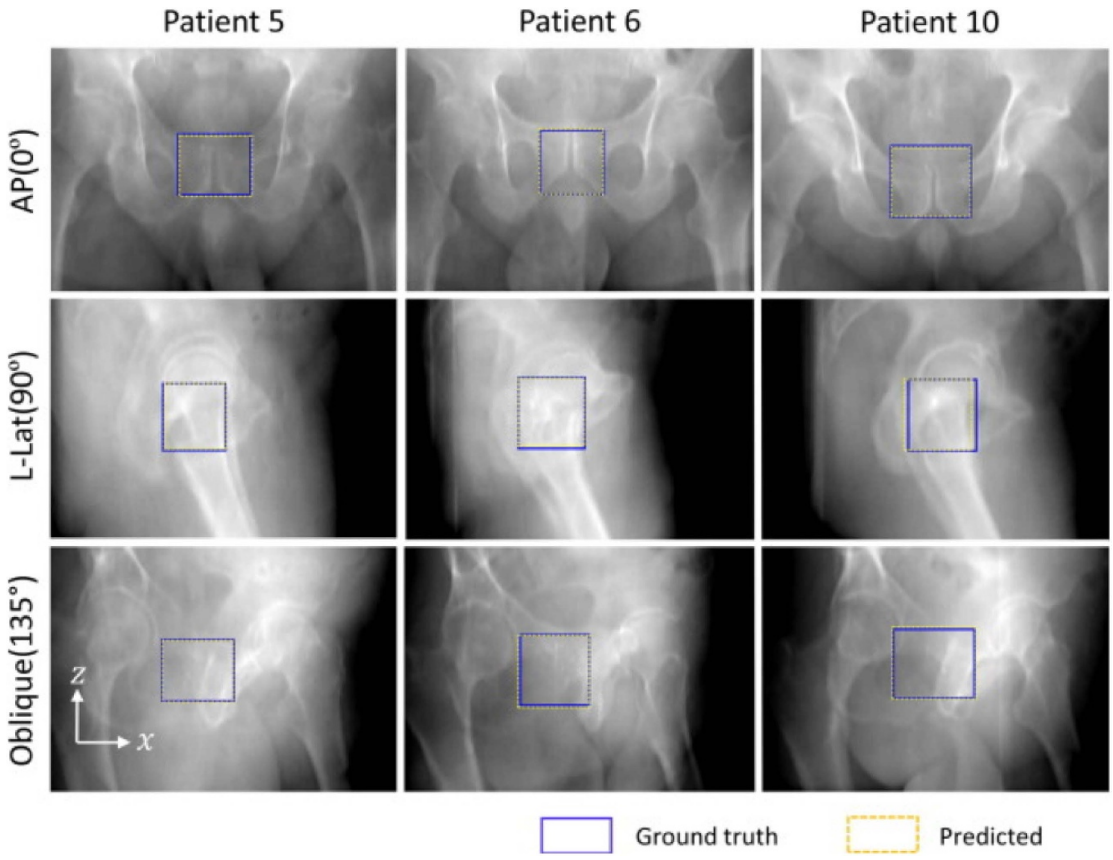


**Figure 4.** Results of identifying the prostate target from 2D x-ray projection of the patient anatomy using a CNN. Prostate boundary boxes derived from the deep learning model is shown in yellow, and their corresponding ground truth is in blue. Adapted with permission from [236]. Copyright 2019 Elsevier.

high-dose-rate brachytherapy [248] and for predicting xerostomia in patients undergoing head and neck radiotherapy [249]. Cui *et al* combined traditional machine learning methods and deep learning techniques to predict lung pneumonitis after radiotherapy [250]. In the era of personalized medicine, more and more patient-specific data other than images are incorporated for response modeling. For instance, machine learning especially deep learning techniques can be applied in radiogenomics [251].

## 5. Perspective

Machine learning techniques will soon be utilized during all stages of radiotherapy, starting with diagnostic imaging and its use for target and organ at risk delineation (potentially using image registration), followed by automated planning and outcome assessment. The latter will provide valuable input for refinements in treatment planning towards personalized medicine. It is a most promising direction to combine deep-learning-based medical imaging and AI-driven radiotherapy. It is critically important for precision radiotherapy to perform high-throughput and quantitative analysis of comprehensive features in medical images such as CT and MRI images and other relevant data. A key insight is that images and other data contain significantly more information than what can be visually extracted by radiologists and oncologists, which can be effectively harvested using sophisticated algorithms to improve treatments and quality of life. AI-based imaging guided radiotherapy will be superior than classic workflows for two main reasons. First, some hidden features the reader cannot perceive can be utilized in radiomic analysis. Second, extensive prior knowledge and domain constraints can be utilized in the imaging and radio-therapeutic processes in a data-driven and end-to-end fashion, which is powerful; for example, to transform tomographic algorithms to a new level by compensating for model mismatches of many kinds. During the radiotherapeutic process, big data are generated on anatomical, functional, metabolic, pathological, cellular and molecular features, especially in the forms of tomographic images, genetic profiles, and medical reports [252]. These data can be structured into primitives and patterns, which can be interpreted as 'biological languages' in a general sense. On the other hand, radiological, pathological and oncological reports/notes are in natural languages. A major clinical challenge is that those 'biological languages' are extremely difficult to be extracted and

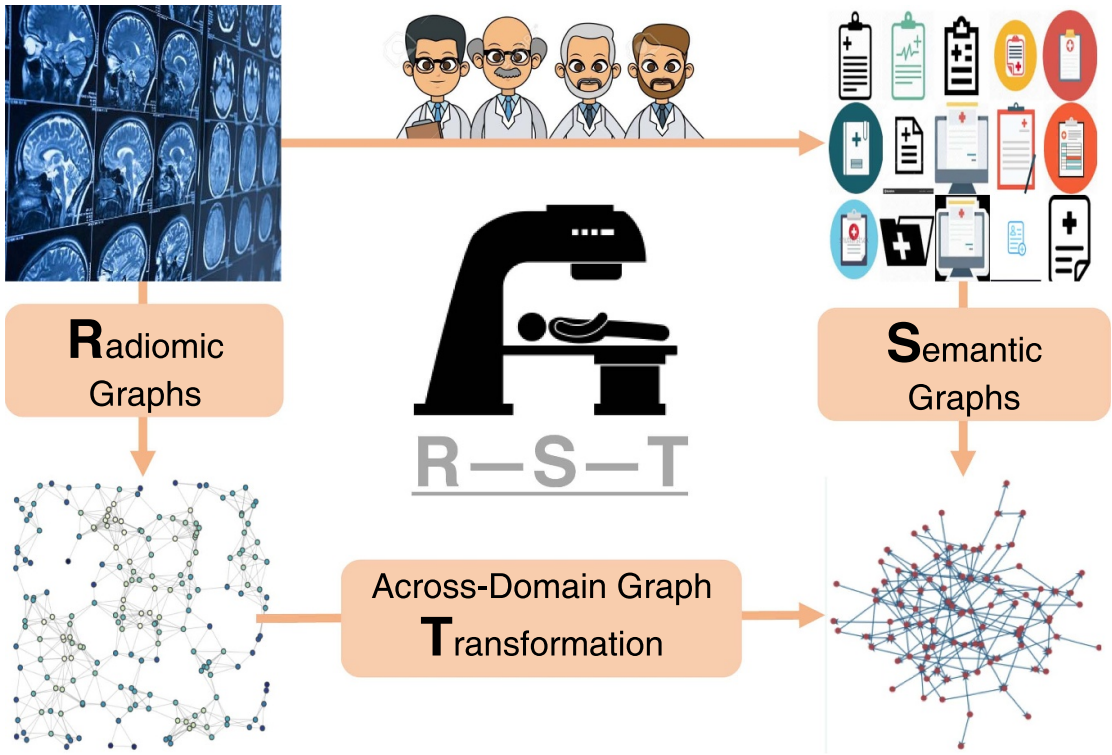


**Figure 5.** Big picture of AI-driven radiotherapy with graph neural network and natural language processing.

presented in a meaningful way to be interpreted by radiologists/oncologists, and the sensitivity and specificity of the current medical reports, associated decisions and plans need major improvements to treat cancer patients much better than the current standards. It would be fruitful to synergize expertise in tomographic imaging and image analysis as well as natural language processing (NLP), optimize the treatment planning and improve prognosis of cancer patients, which represents a generalization of language concepts, an opportunity to synergize knowledge graphs of different forms, and a convergence of inverse problems, informatics, and knowledge engineering, with a strong application background of radiotherapy.

Figure 5 presents our vision along this direction, including the key elements to (1) process medical images and extract radiomic graphs; (2) integrate clinical reports and construct semantic graphs; and (3) develop graph transformation techniques to convert radiomic graphs to semantic graphs.

There are several interesting topics demanding active investigation. First of all, we can extract treatment-related image features and organize them as graphs so that spatial and temporal dependencies among features will be captured. A graph-based learning system will be developed to quantify patient states and predict responses in terms of relevant data especially CT, nuclear and MRI images, treatment plans, and genomic profiling for personalized radiotherapy. For example, we could employ a ResNet model pre-trained on ImageNet to extract imaging biomarkers from medical images, dose distributions, and genomic data. To enhance the transferability and interpretability, we may fine-tune the pre-trained model with our real data [253] and visualize soft activation maps [193]. The visualization technique $t$-SNE [254] can also be used to visualize the learnt features in a low-dimensional manifold with respect to classification labels. Furthermore, knowledge graphs can be reconstructed for each patient using a graph neural network [255–257]. Constructed graphs will be embedded into a low-dimensional space with an adjusted weight for each edge, reflecting the strength between cancer growth/response patterns and specific data-driven features.

A medical semantic/knowledge graph from a patient's clinical reports and electronic medical records is invaluable to enable reasoning and guide planning. Given these text inputs, the domain knowledge can be extracted to construct a library of knowledge graphs using natural language processing techniques [258–262]. For example, to construct knowledge graphs from medical reports/notes, we could take advantage of online services such as the Amazon Comprehend Medical [263] or the Watson Natural Language Understanding platform [264] rather than from scratch. Based on the cloud service provided by these and other systems, we can distill and query high-quality domain-specific rules/knowledge graphs from unstructured or semi-structured contents extracted from images and data such as medical conditions,

**Table A1.** Acronyms and their full names used in this paper.

| Acronym | Full name |
|---|---|
| AI | Artificial intelligence |
| ML | Machine learning |
| DL | Deep learning |
| ANN | Artificial neural network |
| MLP | Multilayer perceptron |
| CNN | Convolutional neural network |
| ReLU | Rectified linear unit |
| FCN | Fully convolutional network |
| RED-CNN | Residual encoder-decoder convolutional neural network |
| GAN | Generative adversarial network |
| RNN | Recurrent neural network |
| LSTM | Long short-term memory |
| GRU | Gated recurrent units |
| DRL | Deep reinforcement learning |
| CT | Computed tomography |
| MRI | Magnetic resonance imaging |
| PET | Positron emission tomography |
| SPECT | Single-photon emission computed tomography |
| ROC | Receiver operating characteristic |
| AUC | Area under curve |
| US | Ultrasound |
| TRUS | Transrectal ultrasound |
| IMRT | Intensity modulated radiationtherapy |
| VMAT | Volumetric modulated arc therapy |
| DVH | Dose volume histogram |
| QA | Quality assurance |
| kV | Kilo-voltage |
| MV | Mega-voltage |
| IIFDL | Intra- and inter-fraction fuzzy deep learning |
| NLP | Natural language processing |

medication details (dosage, strength, and frequency), and other data from a variety of sources like doctors' notes, clinical trial reports, and patient health records.

With the aforementioned efforts, we will have both radiomics graphs and knowledge/semantic graphs from medical images and medical text data, respectively. These two types of graphs are in different domains: one from images and data biologically/clinically informative, and the other is in terms of professional languages directly interpretable by experts and understandable by an educated patient. Therefore, we need to bridge these two domains via an across-domain graph transformation. To this end, we would use a graph-based encoder-decoder network including the graph convolution [256], graph pooling [265], and graph unpooling [266]. The encoder will extract the information from a radiomic graph while the encoder will reconstruct a corresponding semantic graph. The bottleneck between the encoder and the decoder will serve as the bridge between the image and text domains. In the cases of dynamics changes of tumors with different reports, a graph-based RNN can be used to learn a dynamics graph mapping [267].

## 6. Conclusion

AI/ML is a scientific paradigm shift and has already generated great impacts on medical imaging and radiotherapy. Instead of claiming that deep learning is reaching its limits, we believe that deep learning is being developed into more advanced forms and enhanced with synergistic techniques especially NLP, entity resolution, and graph neural networks. In a larger perspective, deep learning demonstrates successes of connectionism but is also subject to limits of connectionism. Through NLP and its general forms enabled by knowledge graphs, it is hoped that connectionism and computionism will be unified to define the future of AI/ML, and showcase new successes in the medical physics world.

## Acknowledgment

This work was partially support by NIH/NCI under award numbers R01CA233888, R01CA237267, R01CA227289, R37CA214639, and R01CA237269, and NIH/NIBIB under award number R01EB026646.

## Data availability

Data sharing is not applicable to this article as no new data were created or analysed in this study. See table A1.

## ORCID iDs

Hongming Shan https://orcid.org/0000-0002-0604-3197
Xun Jia https://orcid.org/0000-0001-6159-2909
Pingkun Yan https://orcid.org/0000-0002-9779-2141
Harald Paganetti https://orcid.org/0000-0002-6257-2413
Ge Wang https://orcid.org/0000-0002-2656-7705

## References

[1] Moor J 2006 The Dartmouth College artificial intelligence conference: The next fifty years *AI Magazine* vol 27 pp 87–87
[2] (https://www.whitehouse.gov/wp-content/uploads/2019/08/FY-21-RD-Budget-Priorities.pdf )
[3] (https://www.weforum.org/agenda/2018/09/artificial-intelligence-shaking-up-job-market/ )
[4] (https://www.technologyreview.com/s/612768/we-analyzed-16625-papers-to-figure-out-where-ai-is-headed-next/ )
[5] Wang G 2016 A perspective on deep imaging *IEEE Access* **4** 8914–24
[6] Wang G, Ye J C, Mueller K and Fessler J A 2018 Image reconstruction is a new frontier of machine learning *IEEE Trans. Med. Imaging* **37** 1289–96
[7] LeCun Y, Bengio Y and Hinton G 2015 Deep learning *Nature* **521** 436–44
[8] Nielsen M A 2018 *Neural Networks and Deep Learning*, (CA, USA: Determination press San Francisco) 2015 http://neuralnetworksanddeeplearning.com/
[9] Goodfellow I, Bengio Y and Courville A 2016 *Deep Learning* (Cambridge, MA: MIT press) https://www.deeplearningbook.org/
[10] Aston Z, Lipton Z C, Li M and Smola A J 2020 *Dive into Deep Learning* (https://d2l.ai)
[11] Jain A K, Mao J and Mohiuddin K M 1996 Artificial neural networks: A tutorial *Computer* **29** 31–44
[12] Balázs Csanád Csáji 2001 Approximation with artificial neural networks *Masters Thesis* Faculty of Sciences, Etvs Lornd University, Hungary
[13] Krizhevsky A, Sutskever I and Hinton G E 2012 ImageNet classification with deep convolutional neural networks *Advances in Neural Information Processing Systems* pp 1097–105
[14] Kaiming H, Zhang X, Ren S and Sun J 2016 Deep residual learning for image recognition *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition* pp 770–8
[15] Huang G, Liu Z, Weinberger K Q and Laurens van der M 2017 Densely connected convolutional networks *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition* vol 4700–4708
[16] LeCun Y, Bottou Leon, Bengio Y and Haffner P 1998 Gradient-based learning applied to document recognition *Proc. of the IEEE* **86** 2278–324
[17] Nair V and Hinton G E 2010 Rectified linear units improve restricted Boltzmann machines *Proc. of the 27th International Conference on Machine Learning (ICML-10)* pp 807–14
[18] Aydore S, Thirion B and Varoquaux 2019 Feature grouping as a stochastic regularizer for high-dimensional structured data *Proc. of the 36 th Int. Conf. on Machine Learning* vol 97 pp 385–94
[19] Long J, Shelhamer E and Darrell T 2015 Fully convolutional networks for semantic segmentation *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition* pp 3431–40
[20] Chen H, Zhang Y, Kalra M K, Lin F, Chen Y, Liao P, Zhou J and Wang G 2017 Low-dose CT with a residual encoder-decoder convolutional neural network *IEEE Trans. Med. Imaging* **36** 2524–35
[21] Ronneberger O, Fischer P and Brox T 2015 U-Net: Convolutional networks for biomedical image segmentation *Med Image Comput Comput Assist Interv* Springer pp 234–41
[22] Goodfellow I, Pouget-Abadie J, Mirza M, Bing X, Warde-Farley D, Ozair S, Courville A and Bengio Y 2014 Generative adversarial nets *Advances in Neural Information Processing Systems* vol 2672–2680
[23] Simonyan K and Zisserman A Very deep convolutional networks for large-scale image recognition *Int. Conf. Learn. Representations* p 2015
[24] Pan Z, Weijie Y, Xiaokai Y, Khan A, Yuan F and Zheng Y 2019 Recent progress on generative adversarial networks (GANs): A survey *IEEE Access* **7** 36322–33
[25] Zhu J-Y, Park T, Isola P and Efros A A 2017 Unpaired image-to-image translation using cycle-consistent adversarial networks *Proc. of the IEEE International Conference on Computer Vision* pp 2223–32
[26] Hochreiter S and Schmidhuber Jurgen 1997 Long short-term memory *Neural Comput.* **9** 1735–80
[27] Cho K, Bart van Menboer, Gulcehre C, Bahdanau D, Bougares F, Schwenk H and Bengio Y 2014 Learning phrase representations using RNN encoder–decoder for statistical machine translation *Proc. of the 2014th Conf. on Empirical Methods in Natural Language Processing (EMNLP)* pp 1724–34
[28] Silver D *et al* 2016 Mastering the game of Go with deep neural networks and tree search *Nature* **529** 484
[29] Vincent Fçois-L, Henderson P, Islam R, Bellemare M G and Pineau J *et al* 2018 An introduction to deep reinforcement learning *Found. Trends® Machine Learn.* **11** 219–354
[30] Natterer F 2001 *The Mathematics of Computerized Tomography* (Philadelphia, PA: SIAM)
[31] Herman G T 2009 *Fundamentals of Computerized Tomography: Image Reconstruction From Projections* (London: Springer-Verlag)
[32] Vogel R A, Kirch D, LeFree M and Steele P 1978 A new method of multiplanar emission tomography using a seven pinhole collimator and an Anger scintillation camera *J. Nucl. Med.: Official Publ. Soc. Nucl. Med.* **19** 648–54
[33] Geyer L L *et al* 2015 State of the art: Iterative CT reconstruction techniques *Radiology* **276** 339–57
[34] Candes E J and Tao T 2005 Decoding by linear programming *IEEE Trans. Inf. Theory* **51** 4203–15
[35] Herman G T and Davidi R 2008 Image reconstruction from a small number of projections *Inverse Prob.* **24** 045011

[36] Wang G, Zhang Y, Xiaojing Y and Mou X 2019 *Machine Learning for Tomographic Imaging* (Bristol: IOP Publishing)
[37] (https://math.tufts.edu/faculty/equinto/Cormack2019/confprogram.html ).
[38] Ravishankar S, Ye J C and Fessler J A 2020 Image reconstruction: From sparsity to data-adaptive methods and machine learning *Proc. of the IEEE* vol 108 pp 86–109
[39] McCollough C H *et al* 2017 Low-dose CT for the detection and classification of metastatic liver lesions: Results of the 2016 low dose CT grand challenge *Med. Phys.* **44** e339–e352
[40] Knoll F *et al* 2020 fastMRI: A publicly available raw k-space and DICOM dataset of knee images for accelerated MR image reconstruction using machine learning *Radiol.: Art. Intell.* **2** e190007
[41] Knoll F *et al* 2020 Advancing machine learning for MR image reconstruction with an open competition: Overview of the 2019 fastMRI challenge *arXiv preprint arXiv:2001.02518*
[42] Shan H, Zhang Y, Yang Q, Kruger U, Kalra M K, Sun L, Cong W and Wang G 2018 3–D convolutional encoder-decoder network for low-dose CT via transfer learning from a 2–D trained network *IEEE Trans. Med. Imaging* **37** 1522–34
[43] Xie H, Shan H and Wang G 2019 Deep encoder-decoder adversarial reconstruction (DEAR) network for 3D CT from few-view data *Bioengineering* **6** 111
[44] Chen H, Zhang Y, Zhang W, Liao P, Ke Li, Zhou J and Wang G 2017 Low–dose CT via convolutional neural network *Biomed. Opt. Express* **8** 679–94
[45] Yang Q *et al* 2018 Low-dose CT image denoising using a generative adversarial network with Wasserstein distance and perceptual loss *IEEE Trans. Med. Imaging* **37** 1348–57
[46] Dufan W, Kim K, Fakhri G E and Quanzheng Li 2017 A cascaded convolutional nerual network for x-ray low-dose CT image denoising *arXiv preprint arXiv:1705.04267*
[47] Shan H, Padole A, Homayounieh F, Kruger U, Khera R D, Nitiwarangkul C, Kalra M K and Wang G 2019 Competitive performance of a modularized deep neural network compared to commercial algorithms for low-dose CT image reconstruction *Nat. Machine Intell.* **1** 269
[48] Kang E, Chang W, Yoo J and Ye J C 2018 Deep convolutional framelet denosing for low-dose CT via wavelet residual network *IEEE Trans. Med. Imaging* **37** 1358–69
[49] Wolterink J M, Leiner T, Viergever M A and Išgum I 2017 Generative adversarial networks for noise reduction in low-dose CT *IEEE Trans. Med. Imaging* **36** 2536–45
[50] You C *et al* 2018 Structurally-sensitive multi-scale deep neural network for low-dose CT denoising *IEEE Access* **6** 41839–55
[51] Chen H *et al* 2018 LEARN: Learned experts' assessment-based reconstruction network for sparse-data CT *IEEE Trans. Med. Imaging* **37** 1333–47
[52] Xie H, Shan H, Cong W, Zhang X, Liu S, Ning R and Wang G 2019 Dual network architecture for few-view CT– trained on ImageNet data and transferred for medical imaging *Developments in X-Ray Tomography XII* International Society for Optics and Photonics vol 11113 p 111130V
[53] Yinsheng Li, Ke Li, Zhang C, Montoya J and Chen G-H 2019 Learning to reconstruct computed tomography images directly from sinogram data under a variety of data acquisition conditions *IEEE Trans. Med. Imaging* **38** 2469–81
[54] Lin F and Bruno D M 2019 A hierarchical approach to deep learning and its application to tomographic reconstruction *15th Int. Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine* Int. Society for Optics and Photonics vol 11072 p 1107202
[55] Xie H, Shan H, Cong W, Zhang X, Liu S, Ning R and Wang G 2019 Deep efficient end-to-end reconstruction (DEER) network for low-dose few-view breast CT from projection data *arXiv preprint arXiv:1912.04278*
[56] Dufan W, Gong K, Kim K, Xiang Li and Quanzheng Li 2019 Consensus neural network for medical imaging denoising with only noisy training samples *Int. Conf. on Medical Image Computing and Computer-Assisted Intervention* Springer pp 741–9
[57] Kang E, Koo H J, Yang D H, Seo J B and Ye J C 2019 Cycle-consistent adversarial denoising network for multiphase coronary CT angiography *Med. Phys.* **46** 550–62
[58] Han Y and Ye J C 2018 Framing U-Net via deep convolutional framelets: Application to sparse-view CT *IEEE Trans. Med. Imaging* **37** 1418–29
[59] Hammernik K, Klatzer T, Kobler E, Recht M P, Sodickson D K, Pock T and Knoll F 2018 Learning a variational network for reconstruction of accelerated MRI data *Magn. Reson. Med.* **79** 3055–71
[60] Zhu B, Liu J Z, Cauley S F, Rosen B R and Rosen M S 2018 Image reconstruction by domain-transform manifold learning *Nature* **555** 487
[61] Sun J, Huibin Li and Zongben X *et al* 2016 Deep ADMM-Net for compressive sensing MRI *Advances in Neural Information Processing Systems* pp 10–18
[62] Quan T M, Nguyen-Duc T and Jeong W-K 2018 Compressed sensing MRI reconstruction using a generative adversarial network with a cyclic loss *IEEE Trans. Med. Imaging* **37** 1488–97
[63] Hyun C M, Kim H P, Lee S M, Lee S and Seo J K 2018 Deep learning for undersampled MRI reconstruction *Phys. Med. Biol.* **63** 135007
[64] Mardani M, Gong E, Cheng J Y, Vasanawala S S, Zaharchuk G, Xing L and Pauly J M 2018 Deep generative adversarial neural networks for compressive sensing MRI *IEEE Trans. Med. Imaging* **38** 167–79
[65] Yang G *et al* 2017 DAGAN: Deep de-aliasing generative adversarial networks for fast compressed sensing MRI reconstruction *IEEE Trans. Med. Imaging* **37** 1310–21
[66] Kim K, Dufan W, Gong K, Dutta J, Kim J H, Son Y D, Kim H K, Fakhri G E and Quanzheng Li 2018 Penalized PET reconstruction using deep learning prior and local linear fitting *IEEE Trans. Med. Imaging* **37** 1478–87
[67] Yang B, Ying L and Tang J 2018 Artificial neural network enhanced Bayesian PET image reconstruction *IEEE Trans. Med. Imaging* **37** 1297–309
[68] Cui J *et al* 2019 PET image denoising using unsupervised deep learning *Eur. J. Nucl. Med. Mol. Imaging* **46** 2780–9
[69] Gong K, Dufan W, Kim K, Yang J, Sun T, Fakhri G E, Seo Y and Quanzheng Li 2019 MAPEM-Net: an unrolled neural network for fully 3D PET image reconstruction *15th Int. Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine* Int. Soc. Opt. Photonics vol 11072 p110720O
[70] Shao W and Yong D 2019 SPECT image reconstruction by deep learning using a two-step training method *J. Nucl. Med.* **60** 1353–1353 supplement1
[71] Wang Z, Bovik A C, Sheikh H R and Simoncelli E P 2004 Image quality assessment: From error visibility to structural similarity *IEEE Trans. Image Process.* **13** 600–12

[72] Isola P, Zhu J-Y, Zhou T and Efros A A 2017 Image-to-image translation with conditional adversarial networks *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition* pp 1125–34
[73] Zhao H, Gallo O, Frosio I and Kautz J 2016 Loss functions for image restoration with neural networks *IEEE Trans. Comput. Imaging* **3** 47–57
[74] Gong H, Lifeng Y, Leng S, Dilger S K, Ren L, Zhou W, Fletcher J G and McCollough C H 2019 A deep learning-and partial least square regression-based model observer for a low-contrast lesion detection task in CT *Med. Phys.* **46** 2052–63
[75] De Man R, Gang G J, Xin Li and Wang G 2019 Comparison of deep learning and human observer performance for detection and characterization of simulated lesions *J. Med. Imaging* **6** 025503
[76] Zhu J, Liu Y, Zhang J, Wang Y and Chen L 2019 Preliminary clinical study of the differences between interobserver evaluation and deep convolutional neural network-based segmentation of multiple organs at risk in CT images of lung cancer *Front. Oncol.* **9** 627
[77] Lin Li *et al* 2019 Deep learning for automated contouring of primary tumor volumes by MRI for nasopharyngeal carcinoma *Radiology* **291** 677–86
[78] Kosmin M *et al* 2019 Rapid advances in auto-segmentation of organs at risk and target volumes in head and neck cancer *Radiother. Oncol.* **135** 130–40
[79] Sharp G, Fritscher K D, Pekar V, Peroni M, Shusharina N, Veeraraghavan H and Yang J 2014 Vision 20/20: perspectives on automated image segmentation for radiotherapy *Med. Phys.* **41** 050902
[80] Weszka J S 1978 A survey of threshold selection techniques *Comput. Graph. Image Process.* **7** 259–65
[81] Boykov Y Y and Jolly M-P 2001 Interactive graph cuts for optimal boundary & region segmentation of objects in ND images *Proc. Eighth IEEE International Conference on Computer Vision. ICCV 2001* vol 1 pp 105–12 IEEE
[82] Yan P and Kassim A A 2006 Segmentation of volumetric MRA images by using capillary active contour *Med. Image Anal.* **10** 317–29
[83] Mangan A P and Whitaker R T 1999 Partitioning 3D surface meshes using watershed segmentation *IEEE Trans. Vis. Comput. Graph.* **5** 308–21
[84] Lindeberg T and Meng-Xiang Li 1997 Segmentation and classification of edges using minimum description length approximation and complementary junction cues *Comput. Vis. Image Underst.* **67** 88–98
[85] Yan P and Kassim A A 2006 Medical image segmentation using minimal path deformable models with implicit shape priors *IEEE transactions on information technology in biomedicine: a publication of the IEEE Engineering in Medicine and Biology Society* **10** 677–84
[86] Rohlfing T, Brandt R, Menzel R, Russakoff D B and Maurer C R 2005 Quo vadis, atlas-based segmentation? *Handbook of Biomedical Image Analysis* (Boston, MA: Springer) pp 435–86
[87] Thirion J-P 1998 Image matching as a diffusion process: an analogy with Maxwell's demons *Med. Image Anal.* **2** 243–60
[88] Ourselin Sebastien, Roche A, Prima S and Ayache N 2000 Block matching: A general framework to improve robustness of rigid registration of medical images *Int. Conf. on Medical Image Computing And Computer-Assisted Intervention* pp 557–66 Springer
[89] Rueckert D, Sonoda L I, Hayes C, Hill D L G, Leach M O and Hawkes D J 1999 Nonrigid registration using free-form deformations: Application to breast MR images *IEEE Trans. Med. Imaging* **18** 712–21
[90] Yan P, Cao Y, Yuan Y, Turkbey B and Choyke P L 2015 Label image constrained multiatlas selection *IEEE transactions on cybernetics* **45** 1158–68
[91] Heimann T and Meinzer H-P 2009 Statistical shape models for 3D medical image segmentation: A review *Med. Image Anal.* **13** 543–63
[92] Cootes T F, Taylor C J, Cooper D H and Graham J 1995 Active shape models-their training and application *Comput. Vis. Image Underst.* **61** 38–59
[93] Cootes T F, Edwards G J and Taylor C J 2001 Active appearance models *IEEE Trans. Pattern Anal. Machine Intell.* **23** 681–5
[94] Yan P, Zhang W, Baris T, Choyke P L and Li X 2013 Global structure constrained local shape prior estimation for medical image segmentation *Comput. Vis. Image Underst.* **117** 1017–26
[95] Geremia E, Clatz O, Menze B H, Konukoglu E, Criminisi A and Ayache N 2011 Spatial decision forests for MS lesion segmentation in multi-channel magnetic resonance images *NeuroImage* **57** 378–90
[96] Wei Li, Liao S, Feng Q, Chen W and Shen D 2012 Learning image context for segmentation of the prostate in CT-guided radiotherapy *Phys. Med. Biol.* **57** 1283
[97] Wachinger C, Sharp G C and Golland P 2013 Contour-driven regression for label inference in atlas-based segmentation *Int. Conf. on Medical Image Computing and Computer-Assisted Intervention* Springer pp 211–18
[98] Yan P and Kruecker J 2010 Incremental shape statistics learning for prostate tracking in TRUS *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2010* (Berlin: Springer) pp 42–9
[99] Cortes C and Vapnik V 1995 Support-vector networks *Mach. Learn.* **20** 273–97
[100] Breiman L 2001 Random forests *Mach. Learn.* **45** 5–32
[101] Rasmussen C E 2003 Gaussian processes in machine learning *Summer School on Machine Learning* (Berlin: Springer) pp 63–71
[102] Trebeschi S, Griethuysen J J M van, Lambregts D M J, Lahaye M J, Parmar C, Bakers F C H, Peters N H G M, Beets-Tan R G H and Aerts H J W L 2017 Deep learning for fully-automated localization and segmentation of rectal cancer on multiparametric MR *Sci. Rep.* **7** 5301
[103] Jackson P, Hardcastle N, Dawe N, Kron T, Hofman M and Hicks R J 2018 Deep learning renal segmentation for fully automated radiation dose estimation in unsealed source therapy *Front. Oncol.* **8** 215
[104] Ibragimov B, Toesca D, Chang D, Koong A and Xing L 2017 Combining deep learning with anatomical analysis for segmentation of the portal vein for liver SBRT planning *Phys. Med. Biol.* **62** 8943
[105] Kamnitsas K *et al* 2017 Ensembles of multiple models and architectures for robust brain tumour segmentation *Int. MICCAI Brainlesion Workshop* pp 450–62 Springer
[106] Kamnitsas K, Ledig C, Newcombe V F J, Simpson J P, Kane A D, Menon D K, Rueckert D and Glocker B 2017 Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation *Med. Image Anal.* **36** 61–78
[107] Havaei M, Davy A, Warde-Farley D, Biard A, Courville A, Bengio Y, Pal C, Jodoin P-M and Larochelle H 2017 Brain tumor segmentation with deep neural networks *Med. Image Anal.* **35** 18–31
[108] Kamnitsas K, Ferrante E, Parisot S, Ledig C, Nori A V, Criminisi A, Rueckert D and Glocker B 2016 Deepmedic for brain tumor segmentation *Int. Workshop on Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries* Springer pp 138–49
[109] Zhuge Y, Krauze A V, Ning H, Cheng J Y, Arora B C, Camphausen K and Miller R W 2017 Brain tumor segmentation using holistically nested neural networks in MRI images *Med. Phys.* **44** 5234–43

[110] Liu Y et al 2017 A deep convolutional neural network-based automatic delineation strategy for multiple brain metastases stereotactic radiosurgery PloS one 12 e0185844
[111] de Brebisson A and Montana G 2015 Deep neural networks for anatomical brain segmentation Proc. of the Conf. on Computer Vision and Pattern Recognition Workshops pp 20–8
[112] Cao L, Long Li, Zheng J, Fan X, Yin F, Shen H and Zhang J 2018 Multi-task neural networks for joint hippocampus segmentation and clinical score regression Multimedia Tools Appl. 77 29669–86
[113] Ren X, Xiang L, Nie D, Shao Y, Zhang H, Shen D and Wang Q 2018 Interleaved 3D-CNN s for joint segmentation of small-volume structures in head and neck CT images Med. Phys. 45 2063–75
[114] Kushibar K, Valverde S, Sandra G-V, Bernal J, Cabezas M, Oliver A and Xavier L 2018 Automated sub-cortical brain structure segmentation combining spatial and deep convolutional features Med. Image Anal. 48 177–86
[115] Dalmiş M U, Litjens G, Holland K, Setio A, Mann R, Karssemeijer N and Gubern-Mérida A 2017 Using deep learning to segment breast and fibroglandular tissue in MRI volumes Med. Phys. 44 533–46
[116] Milletari F, Ahmadi S-A, Kroll C, Plate A, Rozanski V, Maiostre J, Levin J, Dietrich O and Ertl-Wagner B et al 2017 Hough-CNN: deep learning for segmentation of deep brain regions in MRI and ultrasound Comput. Vis. Image Underst. 164 92–102
[117] Milletari F, Navab N and Ahmadi S-A 2016 V-net: Fully convolutional neural networks for volumetric medical image segmentation 2016 Fourth Int. Conf. on 3D Vision (3DV) pp 565–71 IEEE
[118] Zhou X, Ito T, Takayama R, Wang S, Hara T and Fujita H 2016 Three-dimensional CT image segmentation by combining 2D fully convolutional network with 3D majority voting Deep Learning and Data Labeling for Medical Applications Springer pp 111–20
[119] Ling M, Guo R, Zhang G, Tade F, Schuster D M, Nieh P, Master V and Fei B 2017 Automatic segmentation of the prostate on CT images using deep learning and multi-atlas fusion Medical Imaging 2017 Image Processing Int. Society for Optics and Photonics vol 10133 p 101332O
[120] Trullo R, Petitjean C, Ruan S, Dubray B, Nie D and Shen D 2017 Segmentation of organs at risk in thoracic CT images using a sharpmask architecture and conditional random fields 2017 IEEE 14th Int. Symp. on Biomedical Imaging (ISBI 2017) IEEE pp 1003–6
[121] Trullo R Petitjean D Nie D Shen and Ruan S 2017 Joint segmentation of multiple thoracic organs in CT images with two collaborative deep architectures Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support Springer pp 21–9
[122] Peijun H, Fa W, Peng J, Bao Y, Chen F and Kong D 2017 Automatic abdominal multi-organ segmentation using deep convolutional neural network and time-implicit level sets Int. J. Comput. Assist. Radiol. Surg. 12 399–411
[123] Vania M, Mureja D and Lee D 2019 Automatic spine segmentation from CT images using Convolutional Neural Network via redundant generation of class labels J. Comput. Des. Eng. 6 224–32
[124] Yuan Y 2017 Hierarchical convolutional-deconvolutional neural networks for automatic liver and tumor segmentation arXiv preprint arXiv:1710.04540
[125] Ibragimov B and Xing L 2017 Segmentation of organs-at-risks in head and neck CT images using convolutional neural networks Med. Phys. 44 547–57
[126] Laukamp K R, Thiele F, Shakirin G, Zopfs D, Faymonville A, Timmer M, Maintz D, Perkuhn M and Borggrefe J 2019 Fully automated detection and segmentation of meningiomas using deep learning on routine multiparametric MRI European radiology 29 124–32
[127] Henry T et al 2018 Automated PET segmentation for lung tumors: Can deep learning reach optimized expert-based performance? J. Nucl. Med. 59 322–322
[128] Zhu W, Huang Y, Zeng L, Chen X, Liu Y, Qian Z, Nan D, Fan W and Xie X 2019 AnatomyNet: Deep learning for fast and fully automated whole-volume segmentation of head and neck anatomy Med. Phys. 46 576–89
[129] Peng Z et al 2020 A method of rapid quantification of patient-specific organ doses for CT using deep-learning-based multi-organ segmentation and GPU-accelerated Monte Carlo dose computing Med. Phys. (accepted) https://doi.org/10.1002/mp.14131
[130] Zhao X, Yihong W, Song G, Zhenye Li, Zhang Y and Fan Y 2018 A deep learning model integrating FCNNs and CRFs for brain tumor segmentation Med. Image Anal. 43 98–111
[131] Wang G et al 2018 Interactive medical image segmentation using deep learning with image-specific fine tuning IEEE Trans. Med. Imaging 37 1562–73
[132] Chen H, Dou Q, Lequan Y, Qin J and Heng P-A 2018 Voxresnet: Deep voxelwise residual networks for brain segmentation from 3D MR images NeuroImage 170 446–55
[133] Chen X, Zhang R and Yan P 2019 Feature fusion encoder decoder network for automatic liver lesion segmentation 2019 IEEE 16th Int. Symp. on Biomedical Imaging (ISBI 2019) pp 430–3 IEEE
[134] Zhu Q, Bo D and Yan P 2020 Boundary-weighted domain adaptive neural network for prostate MR image segmentation IEEE Trans. Med. Imaging 39 753–63
[135] Afshari S, BenTaieb Aıcha, Mirikharaji Z and Hamarneh G 2019 Weakly supervised fully convolutional network for PET lesion segmentation Medical Imaging 2019 Image Processing Int. Society for Optics and Photonics vol 10949 p 109491K
[136] Gjesteby L, Yan Xi, Kalra M K, Yang Q and Wang G 2016 Hybrid imaging system for simultaneous spiral MR and x-ray (MRX) scans IEEE Access 5 1050–61
[137] Gjesteby L, Cong W, Yang Q, Qian C and Wang G 2018 Simultaneous emission-transmission tomography in an MRI hardware framework IEEE Trans. Radiat. Plasma Med. Sci. 2 326–36
[138] Wolterink J M, Dinkla A M, Savenije M H F, Seevinck P R, van den Berg C A T and Ivana Išgum 2017 Deep MR to CT synthesis using unpaired data Int. Workshop on Simulation and Synthesis in Medical Imaging Springer pp 14–23
[139] Nie D, Trullo R, Lian J, Wang Li, Petitjean C, Ruan S, Wang Q and Shen D 2018 Medical image synthesis with deep convolutional adversarial networks IEEE Trans. Biomed. Eng. 65 2720–30
[140] Zhu Q, Bo D, Turkbey B, Choyke P and Yan P 2018 Exploiting interslice correlation for MRI prostate image segmentation, from recursive neural networks aspect Complexity 2018 4185279
[141] Sabour S, Frosst N and Hinton G E 2017 Dynamic routing between capsules Advances in Neural Information Processing Systems pp 3856–66
[142] LaLonde R and Bagci U 2018 Capsules for object segmentation 1st Conf. on Medical Imaging with Deep Learning (MIDL 2018) Amsterdam, The Netherlands
[143] Maninis K-K, Caelles S, Pont-Tuset J and Luc V G 2018 Deep extreme cut: From extreme points to object segmentation Proc. of the Conf. on Computer Vision and Pattern Recognition pp 616–25

[144] Lin T-Y, Goyal P, Girshick R, Kaiming H and Piotr Dar 2017 Focal loss for dense object detection *Proc. of the IEEE International Conference on Computer Vision* pp 2980–8
[145] Mohseni Salehi S S, Erdogmus D and Gholipour A 2017 Tversky loss function for image segmentation using 3D fully convolutional deep networks *Int. Workshop on Machine Learning in Medical Imaging* Springer pp 379–87
[146] (https://medium.com/@junma11/loss-functions-for-medical-image-segmentation-a-taxonomy-cefa5292eec0 )
[147] Taha A A and Hanbury A 2015 Metrics for evaluating 3D medical image segmentation: analysis, selection and tool *BMC Med. Imaging* **15** 29
[148] Antoine Maintz J B and Viergever M A 1998 A survey of medical image registration *Med. Image Anal.* **2** 1–36
[149] Haskins G, Kruger U and Yan P 2020 Deep learning in medical image registration: a survey *Mach. Vis. Appl.* **31** 8
[150] Brock K K, Mutic S, McNutt T R, Li H and Kessler M L 2017 Use of image registration and fusion algorithms and techniques in radiotherapy: Report of the AAPM radiotherapy committee task group No. 132 *Med. Phys.* **44** e43–e76
[151] Yang X, Kwitt R, Styner M and Niethammer M 2017 Quicksilver: Fast predictive image registration–a deep learning approach *NeuroImage* **158** 378–96
[152] Jun Lv, Yang M, Zhang J and Wang X 2018 Respiratory motion correction for free-breathing 3D abdominal MRI using CNN-based image registration: A feasibility study *Br. J. Radiol.* **91** 20170788
[153] Kim B, Kim J, Lee J-G, Kim D H, Park S H and Ye J C 2019 Unsupervised deformable image registration using cycle-consistent cnn *Int. Conf. on Medical Image Computing and Computer-Assisted Intervention* Springer pp 166–74
[154] Miao S, Jane Wang Z, Zheng Y and Liao R 2016 Real-time 2D/3D registration via CNN regression *2016 IEEE 13th Int. Symp. on Biomedical Imaging (ISBI)* IEEE pp 1430–4
[155] Hou B, Alansary A, McDonagh S, Davidson A, Rutherford M, Hajnal J V, Rueckert D, Glocker B and Kainz B 2017 Predicting slice-to-volume transformation in presence of arbitrary subject motion *Int. Conf. on Medical Image Computing and Computer-Assisted Intervention* pp 296–304 Springer
[156] Zheng J, Miao S, Jane Wang Z and Liao R 2018 Pairwise domain adaptation module for CNN-based 2-D/3-D registration *Journal of Medical Imaging* **5** 021204
[157] Sloan J M, Goatman K A and Paul Siebert J 2018 Learning rigid image registration-utilizing convolutional neural networks for medical image registration *Proc. of the 11th Int. Conf. on Biomedical Engineering Systems and Technologies (BIOSTEC 2018) - Volume 2: BIOIMAGING* pp 89–99 SCITEPRESS-Science and Technology Publications
[158] Chee E and Joe W 2018 AIRNet: Self-supervised affine registration for 3D medical images using neural networks *arXiv preprint arXiv:1810.02583*
[159] Yan P, Sheng X, Rastinehad A R and Wood B J 2018 Adversarial image registration with application for MR and TRUS image fusion *Int. Workshop on Machine Learning in Medical Imaging* Springer pp 197–204
[160] Cao X, Yang J, Zhang J, Nie D, Kim M, Wang Q and Shen D 2017 Deformable image registration based on similarity-steered CNN regression *Int. Conf. on Medical Image Computing and Computer-Assisted Intervention* Springer pp 300–8
[161] Sokooti H, de Vos B, Berendsen F, Lelieveldt B P F, Ivana Išgum and Staring M 2017 Nonrigid image registration using multi-scale 3D convolutional neural networks *Int. Conf. on Medical Image Computing and Computer-Assisted Intervention* Springer pp 232–9
[162] Sun Y, Moelker A, Niessen W J and Theo van W 2018 Towards robust CT-Ultrasound registration using deep learning methods *Understanding and Interpreting Machine Learning in Medical Image Computing Applications* pp 43–51 Springer
[163] Uzunova H, Wilms M, Handels H and Ehrhardt J 2017 Training CNNs for image registration from few samples with model-based data augmentation *Int. Conf. on Medical Image Computing and Computer-Assisted Intervention* pp 223–31 Springer
[164] Sokooti H, de Vos B, Berendsen F, Ghafoorian M, Yousefi S, Lelieveldt B P F, Isgum I and Staring M 2019 3D convolutional neural networks image registration based on efficient supervised learning from artificial deformations *arXiv preprint arXiv:1908.10235*
[165] Zhu Y, Zhou Z, Liao G and Yuan K 2019 Deformable registration using average geometric transformations for brain MR images *arXiv preprint arXiv:1907.09670*
[166] Zhang J 2018 Inverse-consistent deep networks for unsupervised deformable image registration *arXiv preprint arXiv:1809.03443*
[167] Dalca A V, Balakrishnan G, Guttag J and Sabuncu M R 2018 Unsupervised learning for fast probabilistic diffeomorphic registration *Int. Conf. on Medical Image Computing and Computer-Assisted Intervention* pp 729–38 Springer
[168] Balakrishnan G, Zhao A, Sabuncu M R, Guttag J and Dalca A V 2018 An unsupervised learning model for deformable medical image registration *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition* pp 9252–60
[169] Seungjong O and Kim S 2017 Deformable image registration in radiotherapy *Radiation Oncol. J.* **35** 101
[170] Van de Kraats E B, Penney G P, Tomazevic D, Walsum T V and Niessen W J 2005 Standardized evaluation methodology for 2-D-3-D registration *IEEE Trans. Med. Imaging* **24** 1177–89
[171] Haskins G, Kruecker J, Kruger U, Sheng X, Pinto P A, Wood B J and Yan P 2019 Learning deep similarity metric for 3D MR-TRUS image registration *Int. J. Comput. Assist. Radiol. Surg.* **14** 417–25
[172] Keszei Aas P, Berkels B and Deserno T M 2017 Survey of non-rigid registration tools in medicine *J. Digit. Imaging* **30** 102–16
[173] Lambin P *et al* 2012 Radiomics: extracting more information from medical images using advanced feature analysis *Eur. J. Cancer* **48** 441–6
[174] Kumar V *et al* 2012 Radiomics: the process and the challenges *Magn. Reson. Imaging* **30** 1234–48
[175] Aerts H J W L *et al* 2014 Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach *Nat. Commun.* **5** 4006
[176] Lam S W-C 1996 Texture feature extraction using gray level gradient based co-occurence matrices *1996 IEEE Int. Conf. on Systems, Man and Cybernetics. Information Intelligence and Systems (Cat. No. 96CH35929)* IEEE vol 1 pp 267–71
[177] Haralick R M, Shanmugam K and Dinstein I' H 1973 Textural features for image classification *IEEE Trans. Syst. Man Cybern. SMC-3* pp 610–21
[178] Dasarathy B V and Holder E B 1991 Image characterizations based on joint gray level—run length distributions *Pattern Recognit. Lett.* **12** 497–502
[179] Dalal N and Triggs B 2005 Histograms of oriented gradients for human detection *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)* IEEE vol 1 pp 886–93
[180] Ojala T, Pietikäinen M and Harwood D 1996 A comparative study of texture measures with classification based on featured distributions *Pattern Recognit.* **29** 51–9
[181] Deng J, Dong W, Socher R, Li-Jia Li, Kai Li and Fei-Fei Li 2009 A large-scale hierarchical image database *2009 IEEE Conference on Computer Vision and Pattern Recognition* pp 248–55
[182] Raghu M, Zhang C, Kleinberg J and Bengio S 2019 Transfusion: Understanding transfer learning for medical imaging *Advances in Neural Information Processing Systems* pp 3342–52

[183] Afshar P, Plataniotis K N and Mohammadi A 2019 Capsule networks' interpretability for brain tumor classification via radiomics analyses *2019 IEEE Int. Conf. on Image Processing (ICIP)* IEEE pp 3816–20
[184] Lao J, Chen Y, Zhi-Cheng Li, Qihua Li, Zhang J, Liu J and Zhai G 2017 A deep learning-based radiomics model for prediction of survival in glioblastoma multiforme *Sci. Rep.* **7** 1–8
[185] Han W, Qin L, Bay C, Chen X, Yu K-H, Miskin N, Li A, Xu X and Young G 2020 Deep transfer learning and radiomics feature prediction of survival of patients with high-grade gliomas *Am. J. Neuroradiol.* **41** 40–8
[186] Banerjee S, Mitra S, Masulli F and Rovetta S 2019 Deep radiomics for brain tumor detection and classification from multi-sequence MRI *arXiv preprint arXiv:1903.09240*
[187] Zhu Y *et al* 2019 A deep learning radiomics model for preoperative grading in meningioma *Eur. J. Radiol.* **116** 128–34
[188] Hosny A *et al* 2018 Deep learning for lung cancer prognostication: A retrospective multi-cohort radiomics study *PLoS Med.* **15** e1002711
[189] Lou B, Doken S, Zhuang T, Wingerter D, Gidwani M, Mistry N, Ladic L, Kamen A and Abazeed M E 2019 An image-based deep learning framework for individualising radiotherapy dose: a retrospective analysis of outcome prediction *Lancet Digit. Health* **1** e136–e147
[190] Chaunzwa T L, Christiani D C, Lanuti M, Shafer A, Diao N, Mak R H and Aerts H 2018 Using deep-learning radiomics to predict lung cancer histology *J. Clin. Oncol.* **36** 8545–45
[191] Shin H-C Roth R Gao , Le Lu , Ziyue X Nogues I Yao J Mollura D and Summers R M 2016 Deep convolutional neural networks for computer-aided detectionCNN architecturesdataset characteristics and transfer learning *IEEE Trans. Med. Imaging* **35** 1285–98
[192] Coudray N, Ocampo P S, Sakellaropoulos T, Narula N, Snuderl M, Fenyö D, Moreira A L, Razavian N and Tsirigos A 2018 Classification and mutation prediction from non–small cell lung cancer histopathology images using deep learning *Nat. Med.* **24** 1559–67
[193] Lei Y, Tian Y, Shan H, Zhang J, Wang G and Kalra M K 2020 Shape and margin-aware lung nodule classification in low-dose CT images via soft activation mapping *Med. Image Anal.* **60** 101628
[194] Hussein S, Cao K, Song Q and Bagci U 2017 Risk stratification of lung nodules using 3D CNN-based multi-task learning *Int. Conference on Information Processing in Medical Imaging* pp 249–60 Springer
[195] Kang G, Liu K, Hou B and Zhang N 2017 3D multi-view convolutional neural networks for lung nodule classification *PloS one* **12** e0188290
[196] Huynh B Q, Hui Li and Giger M L 2016 Digital mammographic tumor classification using transfer learning from deep convolutional neural networks *J. Med. Imaging* **3** 034501
[197] Wang D, Khosla A, Gargeya R, Irshad H and Beck A H 2016 Deep learning for identifying metastatic breast cancer *arXiv preprint arXiv:1606.05718*
[198] Nan W *et al* 2020 Deep neural networks improve radiologists' performance in breast cancer screening *IEEE Trans. Med. Imaging* **39** 1184–94
[199] McKinney S M *et al* 2020 Int. evaluation of an AI system for breast cancer screening *Nature* **577** 89–94
[200] Zhou L, Zhang Z, Chen Y-C, Zhao Z-Y, Yin X-D and Jiang H-B 2019 A deep learning-based radiomics model for differentiating benign and malignant renal tumors *Trans. Oncol.* **12** 292–300
[201] Welikala R A, Foster P J, Whincup P H, Rudnicka A R, Owen C G, Strachan D P and Barman S A *et al* 2017 Automated arteriole and venule classification using deep learning for retinal images from the UK biobank cohort *Comput. Biol. Med.* **90** 23–32
[202] Song Y, Zhang Y-D, Yan X, Liu H, Zhou M, Bingwen H and Yang G 2018 Computer-aided diagnosis of prostate cancer using a deep convolutional neural network from multiparametric MRI *J. Magnet. Reson. Imaging* **48** 1570–7
[203] AlMubarak H A, Stanley J, Guo P, Long R, Antani S, Thoma G, Zuna R, Frazier S and Stoecker W 2019 A hybrid deep learning and handcrafted feature approach for cervical cancer digital histology image classification *Int. J. Healthcare Inform. Syst. Inform. (IJHISI)* **14** 66–87
[204] Halicek M, Guolan L, Little J V, Wang X, Patel M, Griffith C C, El-Deiry M W, Chen A Y and Fei B 2017 Deep convolutional neural networks for classifying head and neck cancer using hyperspectral imaging *J. Biomed. Opt.* 22 060503
[205] Cha K H, Hadjiiski L, Chan H-P, Weizer A Z, Alva A, Cohan R H, Caoili E M, Paramagul C and Samala R K 2017 Bladder cancer treatment response assessment in ct using radiomics with deep-learning *Sci. Rep.* **7** 1–12
[206] Kalra M, Wang G and Orton C G 2018 Radiomics in lung cancer: Its time is here *Med. Phys.* **45** 997–1000
[207] Dufan W, Kim K, Dong B, Fakhri G E and Quanzheng Li 2018 End-to-end lung nodule detection in computed tomography *Int. Workshop on Machine Learning in Medical Imaging* p pages 37–45 Springer
[208] Lee H, Huang C, Yune S, Tajmir S H, Kim M and Synho D 2019 Machine friendly machine learning: interpretation of computed tomography without image reconstruction *Sci. Rep.* **9** 15540
[209] De Man Q *et al* 2019 A two-dimensional feasibility study of deep learning-based feature detection and characterization directly from CT sinograms *Med. Phys.* **46** e790–e800
[210] Otto K 2008 Volumetric modulated arc therapy: IMRT in a single gantry arc *Med. Phys.* **35** 310–17
[211] Binbin W, Ricchetti F, Sanguineti G, Kazhdan M, Simari P, Chuang M, Taylor R, Jacques R and Todd M 2009 Patient geometry-driven information retrieval for IMRT treatment plan quality control *Med. Phys.* **36** 5497–505
[212] Binbin W, Ricchetti F, Sanguineti G, Kazhdan M, Simari P, Jacques R, Taylor R and Todd M 2011 Data-driven approach to generating achievable dose–volume histogram objectives in intensity-modulated radiotherapy planning *Int. J. Radiat. Oncol. Biol. Phys.* **79** 1241–7
[213] Zhu X, Yaorong G, Taoran Li, Thongphiew D, Yin F-F and Jackie Q 2011 A planning quality evaluation tool for prostate adaptive IMRT based on machine learning *Med. Phys.* **38** 719–26
[214] Appenzoller L M, Michalski J M, Thorstad W L, Mutic S and Moore K L 2012 Predicting dose-volume histograms for organs-at-risk in IMRT planning *Med. Phys.* **39** 7446–61
[215] Yuan L, Yaorong G, Robert Lee W, Yin F F, John P and Q J 2012 Quantitative analysis of the factors which affect the interpatient organ-at-risk dose sparing variation in IMRT plans *Med. Phys.* **39** 6868–78
[216] Shiraishi S, Tan J, Olsen L A and Moore K L 2015 Knowledge-based prediction of plan quality metrics in intracranial stereotactic radiosurgery *Med. Phys.* **42** 908–17
[217] Shiraishi S and Moore K L 2016 Knowledge-based prediction of three-dimensional dose distributions for external beam radiotherapy *Med. Phys.* **43** 378–87
[218] Kajikawa T, Kadoya N, Ito K, Takayama Y, Chiba T, Tomori S, Takeda K and Jingu K 2018 Automated prediction of dosimetric eligibility of patients with prostate cancer undergoing intensity-modulated radiotherapy using a convolutional neural network *Radiol. Phys. Technol.* **11** 320–7

[219] Barragán-Montero A M, Nguyen D, Weiguo L, Lin M, Norouzi-Kandalan R, Geets X, Sterpin E and Jiang S 2019 Three-dimensional dose prediction for lung IMRT patients with deep neural networks: Robust learning from heterogeneous beam configurations *Med. Phys.* **46** 3679–91
[220] Chen X, Men K, Yexiong Li, Junlin Y and Dai J 2019 A feasibility study on an automated method to generate patient-specific dose distributions for radiotherapy using deep learning *Med. Phys.* **46** 56–64
[221] Fan J, Wang J, Chen Z, Chaosu H, Zhang Z and Weigang H 2019 Automatic treatment planning based on three-dimensional dose distribution predicted from deep learning technique *Med. Phys.* **46** 370–81
[222] Nguyen D, Long T, Jia X, Weiguo L, Xuejun G, Iqbal Z and Jiang S 2019 A feasibility study for predicting optimal radiotherapy dose distributions of prostate cancer patients from patient anatomy using deep learning *Sci. Rep.* **9** 1076
[223] Nguyen D, Jia X, Sher D, Lin M-H, Iqbal Z, Liu H and Jiang S 2019 3D radiotherapy dose prediction on head and neck cancer patients with a hierarchically densely connected U-net deep learning architecture *Phys. Med. Biol.* **64** 065020
[224] Liu Z, Fan J, Minghui Li, Yan H, Zhihui H, Huang P, Tian Y, Miao J and Dai J 2019 A deep learning method for prediction of three-dimensional dose distribution of helical tomotherapy *Med. Phys.* **46** 1972–83
[225] Ming M, Kovalchuk N, Buyyounouski M K, Xing L and Yang Y 2019 Incorporating dosimetric features into the prediction of 3D VMAT dose distributions using deep convolutional neural network *Phys. Med. Biol.* **64** 125017
[226] Ming M, Buyyounouski M K, Vasudevan V, Xing L and Yang Y 2019 Dose distribution prediction in isodose feature-preserving voxelization domain using deep convolutional neural network *Med. Phys.* **46** 2978–87
[227] Mnih V *et al* 2015 Human-level control through deep reinforcement learning *Nature* **518** 529
[228] Silver D *et al* 2017 Mastering the game of Go without human knowledge *Nature* **550** 354
[229] Shen C, Gonzalez Y, Klages P, Qin N, Jung H, Chen L, Nguyen D, Jiang S B and Jia X 2019 Intelligent inverse treatment planning via deep reinforcement learning, a proof-of-principle study in high dose-rate brachytherapy for cervical cancer *Phys. Med. Biol.* **64** 115013
[230] Shen C, Gonzalez Y, Chen L, Nguyen D and Jia X 2019 Automatic treatment planning in a human-like manner: Operating treatment planning systems by a deep reinforcement learning based virtual treatment planner *Ann. Meeting Am. Soc. Radiat. Oncol.* **105** S256
[231] Interian Y, Rideout V, Kearney V P, Gennatas E, Morin O, Cheung J, Solberg T and Valdes G 2018 Deep nets vs expert designed features in medical physics: An IMRT QA case study *Med. Phys.* **45** 2672–80
[232] Tomori S, Kadoya N, Takayama Y, Kajikawa T, Shima K, Narazaki K and Jingu K 2018 A deep learning-based prediction model for gamma evaluation in patient-specific quality assurance *Med. Phys.* **45** 4055–65
[233] Low D A, Harms W B, Mutic S and Purdy J A 1998 A technique for the quantitative evaluation of dose distributions *Med. Phys.* **25** 656–61
[234] Nyflot M J, Thammasorn P, Wootton L S, Ford E C and Chaovalitwongse W A 2019 Deep learning for patient-specific quality assurance: Identifying errors in radiotherapy delivery by radiomic analysis of gamma images with convolutional neural networks *Med. Phys.* **46** 456–64
[235] Sun B, Lam D, Yang D, Grantham K, Zhang T, Mutic S and Zhao T 2018 A machine learning approach to the accurate prediction of monitor units for a compact proton machine *Med. Phys.* **45** 2243–51
[236] Zhao W, Han B, Yang Y, Buyyounouski M, Hancock S L, Bagshaw H and Xing L 2019 Incorporating imaging information from deep neural network layers into image guided radiotherapy (IGRT) *Radiother. Oncol.* **140** 167–74
[237] Liu C, Zheming L, Longhua M, Wang L, Jin X and Wen S 2019 A modality conversion approach to MV-DRs and KV-DRRs registration using information bottlenecked conditional generative adversarial network *Med. Phys.* **46** 4575–87
[238] Chen H *et al* 2018 Deep-learning based surface region selection for deep inspiration breath hold (DIBH) monitoring in left breast cancer radiotherapy *Phys. Med. Biol.* **63** 245013
[239] Park S, Lee S J, Weiss E and Motai Y 2016 Intra-and inter-fractional variation prediction of lung tumors using fuzzy deep learning *IEEE J. Trans. Eng. Health Med.* **4** 1–12
[240] Huang P, Gang Y, Hua L, Liu D, Xing L, Yin Y, Kovalchuk N, Xing L and Dengwang Li 2019 Attention-aware fully convolutional neural network with convolutional long short-term memory network for ultrasound-based motion tracking *Med. Phys.* **46** 2275–85
[241] Lin H, Shi C, Wang B, Chan M F, Tang X and Wei J 2019 Towards real-time respiratory motion prediction based on long short-term memory neural networks *Phys. Med. Biol.* **64** 085010
[242] Peeken J C, Molina-Romero M, Diehl C, Menze B H, Straube C, Meyer B, Zimmer C, Wiestler B and Combs S E 2019 Deep learning derived tumor infiltration maps for personalized target definition in glioblastoma radiotherapy *Radiother. Oncol.* **138** 166–72
[243] Lee J *et al* 2018 Deep learning–based survival analysis identified associations between molecular subtype and optimal adjuvant treatment of patients with gastric cancer *JCO Clin. Cancer Inform.* **2** 1–14
[244] Tseng H-H, Luo Y, Cui S, Chien J-T, Ten Haken R K and Naqa I E 2017 Deep reinforcement learning for automated radiation adaptation in lung cancer *Med. Phys.* **44** 6690–705
[245] Wang C *et al* 2019 Towards predicting the evolution of lung tumors during radiotherapy observed on a longitudinal MR imaging study via a deep learning algorithm *Med. Phys.* **46** 4699–707
[246] Cui S, Luo Y, Tseng H-H, Ten Haken R K and Naqa I E 2018 Artificial neural network with composite architectures for prediction of local control in radiotherapy *IEEE Trans. Radiat. Plasma Med. Sci.* **3** 242–9
[247] Ibragimov B, Toesca D, Chang D, Yuan Y, Koong A and Xing L 2018 Development of deep neural network for individualized hepatobiliary toxicity prediction after liver SBRT *Med. Phys.* **45** 4763–74
[248] Zhen X, Chen J, Zhong Z, Hrycushko B, Zhou L, Jiang S, Albuquerque K and Xuejun G 2017 Deep convolutional neural network with transfer learning for rectum toxicity prediction in cervical cancer radiotherapy: A feasibility study *Phys. Med. Biol.* **62** 8246
[249] Men K, Geng H, Zhong H, Fan Y, Lin A and Xiao Y 2019 A deep learning model for predicting xerostomia due to radiotherapy for head-and-neck squamous cell carcinoma in the RTOG 0522 clinical trial *Int. J. Radiat. Oncol. Biol. Phys.* **105** 440–7
[250] Cui S, Luo Y, Tseng H-H Randall K, Haken Ten and Naqa I El 2019 Combining handcrafted features with latent variables in machine learning for prediction of radiation-induced lung damage *Med. Phys.* **46** 2497–2511
[251] Naqa I E, Kerns S L, Coates J, Luo Y, Speers C, West C M L, Rosenstein B S and Ten Haken R K 2017 Radiogenomics and radiotherapy response modeling *Phys. Med. Biol.* **62** R179
[252] Skripcak T *et al* 2014 Creating a data exchange strategy for radiotherapy research: towards federated databases and anonymised public datasets *Radiother. Oncol.* **113** 303–9

[253] Shan H, Wang G, Kalra M K, de Souza R and Zhang J 2017 Enhancing transferability of features from pretrained deep neural networks for lung nodule classification *The Proc. of the 2017 Int. Conf. on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine (Fully3D)* pp 65–8
[254] van der Maaten L and Hinton G 2008 Visualizing data using t-SNE *Journal of machine learning research* **9** 2579–605
[255] Cai H, Zheng V W and Chen-Chuan Chang K 2018 A comprehensive survey of graph embedding: Problems, techniques and applications *IEEE Trans. Knowl. Data Eng.* **30** 1616–37
[256] Kipf T N and Welling M 2017 Semi-supervised classification with graph convolutional networks *5th Int. Conf. on Learning Representations* (Toulon, France, April 24 - 26 2017)
[257] Franceschi L, Niepert M, Pontil M and Xiao H 2019 Learning discrete structures for graph neural networks *Int. Conf. on Machine Learning* pp 1972–82
[258] Zhang Z, Zhang Y and Zhou T 2018 Medical knowledge attention enhanced neural model for named entity recognition in Chinese E M R *Chinese Computational Linguistics and Natural Language Processing Based on Naturally Annotated Big Data* pp 376–85 Springer
[259] Lin Y, Liu Z, Sun M, Liu Y and Zhu X Learning entity and relation embeddings for knowledge graph completion *Twenty-Ninth AAAI Conference on Artificial Intelligence* p 2015
[260] Xie R, Liu Z, Jia J, Luan H and Sun M Representation learning of knowledge graphs with entity descriptions *Thirtieth Conf. on Artificial Intelligence* p 2016
[261] Bharadwaj S *et al* 2017 Creation and interaction with large-scale domain-specific knowledge bases *Proc. of the VLDB Endowment* vol 10 pp 1965–8
[262] Chiticariu L, Danilevsky M, Yunyao Li, Reiss F and Zhu H 2018 Systemt: Declarative text understanding for enterprise *Proc. of the 2018th Conf. of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 3 (Industry Papers)* pp 76–83
[263] Amazon Comprehend Medical (https://aws.amazon.com/comprehend/medical/ )
[264] Watson Natural Language Understanding (https://www.ibm.com/watson/services/natural-language-understanding/ )
[265] Gao H, Chen Y and Shuiwang J 2019 Learning graph pooling and hybrid convolutional operations for text representations *The World Wide Conf.* pp 2743–9
[266] Gao H and Shuiwang J 2019 Graph U-Nets *Int. Conf. on Machine Learning* pp 2083–92
[267] Liang X, Zhou H and Xing E 2018 Dynamic-structured semantic propagation network *Proc. of the Conf. on Computer Vision and Pattern Recognition* pp 752–61