# Artificial Intelligence in Radiation Therapy

Yabo Fu, Hao Zhang, Eric D. Morris, Carri K. Glide-Hurst, Suraj Pai[ID], Alberto Traverso, Leonard Wee[ID], Ibrahim Hadzic[ID], Per-Ivar Lønne[ID], Chenyang Shen[ID], Tian Liu, and Xiaofeng Yang[ID]

*Abstract*—**Artificial intelligence (AI) has great potential to transform the clinical workflow of radiotherapy. Since the introduction of deep neural networks (DNNs), many AI-based methods have been proposed to address challenges in different aspects of radiotherapy. Commercial vendors have started to release AI-based tools that can be readily integrated to the established clinical workflow. To show the recent progress in AI-aided radiotherapy, we have reviewed AI-based studies in five major aspects of radiotherapy, including image reconstruction, image registration, image segmentation, image synthesis, and automatic treatment planning. In each section, we summarized and categorized the recently published methods, followed by a discussion of the challenges, concerns, and future development. Given the rapid development of AI-aided radiotherapy, the efficiency and effectiveness of radiotherapy in the future could be substantially improved through intelligent automation of various aspects of radiotherapy.**

*Index Terms*—**Artificial intelligence (AI), image reconstruction, image registration, image segmentation, image synthesis, radiotherapy, treatment planning.**

## I. INTRODUCTION

**A**RTIFICIAL intelligence (AI) is a data-driven agent that is designed to imitate human intelligence. The concept of AI is believed to be originated from the idea of robots, which can help human perform laborious and time-consuming tasks. In recent years, the advancements of both computer hardware

This work did not involve human subjects or animals in its research.

Yabo Fu, Tian Liu, and Xiaofeng Yang are with the Department of Radiation Oncology and Winship Cancer Institute, Emory University, Atlanta, GA 30322 USA (e-mail: xiaofeng.yang@emory.edu).

Hao Zhang is with the Department of Medical Physics, Memorial Sloan Kettering Cancer Center, New York, NY 10065 USA.

Eric D. Morris is with the Department of Radiation Oncology, University of California at Los Angeles, Los Angeles, CA 90095 USA.

Carri K. Glide-Hurst is with the Department of Human Oncology, School of Medicine and Public Health, University of Wisconsin–Madison, Madison, WI 53792 USA.

Suraj Pai, Alberto Traverso, Leonard Wee, and Ibrahim Hadzic are with the Department of Radiotherapy (MAASTRO), Maastricht University Medical Centre, 6221XE Maastricht, The Netherlands.

Per-Ivar Lønne is with the Department of Medical Physics, Oslo University Hospital, Nydalen, 0424 Oslo, Norway.

Chenyang Shen is with the Department of Radiation Oncology, University of Texas Southwestern Medical Center, Dallas, TX 75002 USA.

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TRPMS.2021.3107454.

Digital Object Identifier 10.1109/TRPMS.2021.3107454

and software have enabled the development of more and more sophisticated AI agents that can excel certain complex tasks without human input. Meanwhile, the growth and sharing of data has powered the continuous evolvement of AI by machine learning (ML) and deep learning (DL).

ML, a subset of AI, enables machines to achieve AI through algorithms and statistical techniques trained with data where the training process informs decisions made by the ML framework, thus improving the end result as experience is gained [14]. Supervised ML methods for the automatic segmentation of images involve training and tuning a predictive model, often integrating prior knowledge about an image via training samples (i.e., other similarly annotated images to inform the current segmentation task). ML employs statistical tools to explore and analyze previously labeled data with image representations being built from prespecified filters tuned to a specific segmentation task. Although ML techniques are more efficient with image samples and have a less complicated structure, they are often not as accurate when compared to DL techniques [23]. DL is a subset of ML that was originally designed to mimic the learning style of the human brain using neurons. Unlike ML where the "useful" features for the segmentation process must be decided by the user, with DL, the useful features are decided by the network without human intervention.

Radiation oncology is a type of cancer treatment that requires multidisciplinary expertise, including medicine, biology, physics, and engineering. The workflow of typical radiotherapy consists of medical imaging, diagnosis, prescription, computed tomography (CT) simulation, target registration/contouring, treatment planning, treatment quality assurance, and treatment delivery. Owing to the technological advances in the past few decades, the workflow of radiotherapy has become increasingly complex, resulting in heavy reliance on human–machine interactions. Each step in the clinical workflow is highly specialized and standardized with its own technical challenges. Meanwhile, the requirement of manual input from a diverse team of healthcare professionals, including a radiation oncologist, medical physicist, medical dosimetrist, and radiation therapist, has resulted in a suboptimal treatment process that prevents patients' wider access to the scarce treatment infrastructures. The wide adoption of image-guided radiotherapy has created a massive amount of imaging data that needs to be analyzed in a short period of time. However, humans are limited in reviewing and analyzing large amounts of data due to time constraints. Machines, on the other hand, can be trained to share many repetitive workloads with humans and therefore boosting the capacity of quality healthcare. Since the introduction of deep neural networks

(DNNs), many AI-based methods have been proposed to address challenges in different aspects of radiotherapy. Given the rapid development of AI-aided radiotherapy, the efficiency and effectiveness of radiotherapy in the future could be substantially improved through intelligent automation in various aspects of radiotherapy.

Several articles have been published regarding AI in radiation oncology [30]–[32]. Huynh *et al.* [32] provided a high-level general description of AI methods, reviewed its impact on each step of the radiation therapy workflow and discussed how AI might change the roles of radiotherapy medical professionals. Siddique and Chow [31] provided a review of AI in radiotherapy, including diagnostic processes, medical imaging, treatment planning, patient simulation, and quality assurance. Vandewinckele *et al.* [30] published an overview of AI-based applications in radiotherapy, focusing on the implementation and quality assurance of AI models. In this study, to show the recent progress in AI-aided radiotherapy, we have reviewed AI-based studies in five major aspects of radiotherapy, including image reconstruction, image registration, image segmentation, image synthesis, and automatic treatment planning. In each section, we summarized and categorized the recently published methods, followed by a discussion of the challenges, concerns, and future development. Specifically, H. Zhang contributed to the image reconstruction section; Y. Fu, T. Liu, and X. Yang contributed to the image registration section; E. D. Morris and C. K. Glide-Hurst contributed to the image segmentation section; S. Pai, A. Traverso, L. Wee, and I. Hadzic contributed to the image synthesis section; and P. Lønne and C. Shen contributed to the automatic treatment planning section.

## II. IMAGE RECONSTRUCTION

Tomographic imaging plays an important role in external-beam radiation therapy for simulation and treatment planning, pretreatment and intrafractional image guidance, as well as follow-up care. Before treatment, the patient usually undergoes a CT simulation to acquire images of the area of body to be treated with radiation. The acquired CT images are used to delineate the tumors and surrounding critical structures, and then to design an optimal treatment plan for the patient. For tumors around the diaphragm, such as those in the liver and lower lung lobe, 4-D CT scans may also be performed to capture the motion of tumors in respiration. Due to the advantage of superior soft-tissue contrast, magnetic resonance imaging (MRI) scans are also prescribed for some patients with brain tumors, paraspinal tumors, head and neck cancer, prostate cancer, and extremity sarcoma. The MRI scans are fused to simulation CT images to facilitate tumor delineation and organs at risk (OAR) contouring, or in MRI-only simulation to synthesize CT images for treatment planning and dose calculation [43]. Different from anatomical imaging, such as CT and MRI, positron emission tomography (PET) provides information on tumor metabolism and is used for visualization of tumor extent and delineation of volume in need of dose-escalation, e.g., in head and neck cancers [45]. In addition, cone-beam CT (CBCT) is equipped on most C-arm linear accelerators (LINACs) and widely used in daily procedures for verifying the position of the patient and treatment target. Gated

CBCT or 4D-CBCT is sometimes utilized for positioning patient with moving tumors, such as in lung stereotactic ablative radiotherapy (SABR). Furthermore, CT, megavoltage CT (MV-CT), and MV-CBCT are also integrated in some radiotherapy machines for image guidance. MRI-LINAC systems have also been developed in the past decade [48], in which MRI helps to improve the patient setup and target localization, and enables interfraction and intrafraction radiotherapy adaptation [50]. Very recently, PET-guided radiation therapy is also ready for clinical adoption to treat advanced-stage and metastatic cancers [53]. Finally, after completing their radiation treatment course, patients may have another scan (CT, MRI, or PET) before a follow-up appointment with the radiation oncologist.

In the clinic, tomographic images are displayed on the console soon after a patient scan. Thus, making it easy to be unaware of the crucial reconstruction step which is performed in the background by dedicated reconstruction computers. In fact, reconstruction is at the heart of tomographic imaging modalities, because many clinical tasks in the radiation therapy workflow are highly dependent on the reconstruction quality, including target delineation, OARs segmentation, image registration or fusion, image synthesis, treatment planning, dose calculation, patient positioning, image guidance, and radiation therapy response assessment. Poor reconstruction quality would inevitably jeopardize the accuracy of the clinical tasks mentioned above and eventually the outcome for cancer patients. Thus, tomographic image reconstruction has always been an active area of research, with the aim of reducing radiation exposure and/or scan time, suppressing noise and artifacts, and improving image quality. Fig. 1 illustrates the image reconstruction from sensor domain to image domain for CT, MRI and PET.

After data acquisition, the detector measurements are usually preprocessed/calibrated by vendors for various degrading factors. Then, the sensor domain measurements $y \in \mathbb{R}^{I \times 1}$ and the desired image $x \in \mathbb{R}^{J \times 1}$ can be expressed as [55]

$$y = Ax \oplus \varepsilon \tag{1}$$

where $\mathbf{A} \in \mathbb{R}^{I \times J}$ is the system matrix for CT and PET and encoding matrix for MRI, $I$ is the number of sensor measurements, $J$ is the number of image voxels, $\varepsilon$ is the noise intrinsic to the data acquisition, and the operator $\oplus$ denotes the interaction between signal and noise. The operator $\oplus$ becomes $+$ when additive Gaussian noise is assumed for CT and MRI measurements and becomes a nonlinear operator when Poisson noise is assumed for PET measurements. Essentially, image reconstruction is an inverse problem where we reconstruct the unknown image $x$ from the measurements $y$.

Various image reconstruction methods have been proposed in the past few decades. Analytical reconstruction methods, such as filtered backprojection (FBP) for CT, Feldkamp–Davis–Kress (FDK) for CBCT, and inverse fast Fourier transform for MRI, are based on the mathematical inverse of the forward model. Because of their high efficiency and stability, they are still employed by most commercial scanners. However, the reconstructed images may suffer from excessive noise and streak artifacts when the sensor measurements are noisy and undersampled. Iterative reconstruction methods [56]–[61],

including algebraic reconstruction technique, statistical image reconstruction, compressed sensing, and prior-image-based reconstruction, are based on sophisticated system modeling of data acquisition and prior knowledge. They have shown advantages of reducing radiation dose or data acquisition time and improving image quality over analytical methods. One example of these iterative methods is the penalized weighted least-square (PWLS) reconstruction which is widely used for CT/CBCT, MRI, and PET

$$\hat{x} = \operatorname{argmin}\|y - Ax\|_{\mathbf{W}}^2 + \beta R(x) \tag{2}$$

where $\mathbf{W}$ is a weighting matrix accounting for the reliability of each sensor measurement, $R(x)$ is a regularization term incorporating prior knowledge or expectations of image characteristics, and $\beta > 0$ is a scalar control parameter to balance data fidelity and regularization. Commonly used regularizations are based on the Markov random field (MRF) or total variation (TV), while a comprehensive review of the regularization strategies can be found in [58].

Inspired by the successes of AI in many other fields, researchers have investigated to leverage AI, especially DL for tomographic image reconstruction [62]. Numerous papers have been published on this topic, and image reconstruction has become a new frontier of DL [63]. It is noted that many DL-based reconstruction methods can be shared by CT, MRI, and PET, thus we focus on reviewing them for CT and CBCT since they are most widely used in radiation therapy. Interested readers can refer to these review articles [64]–[68] on DL for PET and MRI reconstruction.

Patients for radiation therapy receive multiple CT and CBCT scans, and the accumulated imaging dose could be significant. Considering the harmful effects of X-ray radiation including secondary malignancies, low-dose imaging with satisfactory image quality for clinical tasks are desirable. Aside from hardware improvements, two other strategies have been investigated to achieve low-dose imaging for CT and CBCT, reducing the X-ray tube current and exposure time (low-flux acquisition) or the angular sampling per rotation (sparse-view acquisition) [58]. However, these strategies would increase noise and streak artifacts in the FBP or FDK reconstructed images. 4D-CBCT has the potential to reduce motion artifacts and improve patient setup and treatment accuracy, but the scan takes 2–4 min to acquire enough projections at each respiratory phase to achieve acceptable image quality. The long scan time leads to increased patient discomfort, intraimaging patient motion, and additional imaging radiation dose. An accelerated scan is desirable in the clinic but the FDK reconstructed images from the sparse-view acquisition are also degraded with severe streak artifacts [69]. While iterative reconstruction can tackle these challenges to some extent, the reconstruction time might be too long. Therefore, many DL-based reconstruction methods have been developed to further improve image quality and/or substantially reduce reconstruction time, which can be grouped into the following five categories.

### A. Image Domain Methods

One simple approach to improve low-dose CT image quality is post-reconstruction denoising, and researchers have
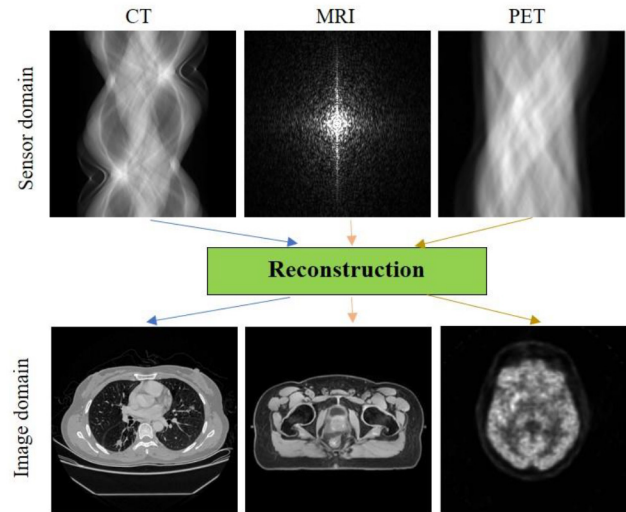


Fig. 1. Illustration of image reconstruction from sensor domain.

applied many different filters to the FBP reconstructed low-dose CT images to suppress noise and streak artifacts. Similarly, the FBP reconstructed low-dose images can be fed into a DL neural network to learn a mapping between the low-dose image and its high-quality counterpart. For example, Kang et al. [70] applied a deep convolutional neural network (CNN) to the wavelet transform coefficients of low-dose CT images, which can effectively suppress noise in the wavelet domain. Chen et al. [71] proposed using overlapped patches from low-dose and corresponding high-quality CT images to boost the number of samples, and then employed a residual encoder–decoder CNN (RED-CNN) to improve low-dose image quality. Yang et al. [72] explored a generative adversarial network (GAN)-based denoising method with Wasserstein distance and perceptual similarity to improve the GAN performance. Wang et al. [73] argued that these image-to-image mapping approaches have some limitations for ultralow-dose CT images. They proposed an iterative residual-artifact learning CNN (called IRLnet) which estimates the high-frequency details within the noise and then removes them iteratively while the residual low-frequency details can be processed through the conventional network. Fig. 2 shows the comparison of ultralow-dose CT image reconstruction by six different approaches.

The streak artifacts in FBP reconstructed images from the sparse-view acquisition are difficult to remove by conventional CNNs. Han and Ye [74] found that the existing U-Net architecture resulted in image blurring and false features for sparse-view CT reconstruction, and proposed a dual frame U-Net and tight frame U-Net to overcome limitations. Zhang et al. [75] investigated a method based on a combination of DenseNet and deconvolution for sparse-view CT, which employs the advantages of both and greatly increases the depth of the network to improve image quality. Alternatively, Jiang et al. [76] used TV-based iterative reconstruction to obtain sparse-view CBCT images (which are superior to FDK reconstructed images), and then fed them into a symmetric residual CNN to learn the mapping between TV-reconstructed images and ground truth. Then, for new

sparse-view CBCT acquisitions, one can use the network to boost TV-reconstructed images.

It is noted above that DL methods are based on a supervised learning framework, which requires both the low-dose CT images and corresponding high-quality counterparts. However, these image pairs may not be available in many clinical scenarios. Wolterink *et al.* [77] explored a GAN consists of a generator CNN and a discriminator CNN to reduce image noise, which showed a generator CNN trained with only adversarial feedback can learn the appearance of high-quality images. Li *et al.* [78] investigated a cycle-consistent GAN (CycleGAN)-based method which does not require low-dose and reference full-dose images from the same patient. In the near future, more unsupervised learning approaches which require no reference ground truth, or semisupervised learning requiring limited reference data, may be explored for low-dose CT and CBCT.

### B. Sensor Domain Methods

It is advantageous to remove noise in the CT sinogram or projection data to prevent its propagation into the reconstruction process, but the edges in the sensor domain are usually not well defined as those in the image domain, resulting in edge blurring in the final reconstructed images [79]. Thus, DL-based denoising methods are rarely applied to the sensor domain directly. Instead, efforts [80]–[82] were dedicated to sparse-view CT reconstruction, which utilize DL to interpolate or synthesize unmeasured projection views. Then, the FBP method is used to reconstruct images with substantially reduced streaking artifacts. Additionally, Beaudry *et al.* [83] proposed a DL method to reconstruct high-quality 4D-CBCT images from sparse-view acquisitions. They estimated projection data for each respiratory bin by taking projections from adjacent bins and linear interpolation and then trained a CNN model to predict full projection data which are reconstructed with the FDK method.

### C. DL for FBP

DL methods can also be combined with the FBP reconstruction. In 2016, Würfl *et al.* [84] demonstrated that FBP reconstruction can be mapped onto a DNN architecture, in which the projection filtering is reformed as a convolution layer and the backprojection is formed with a fully connected layer. They showed the advantage of learning projection-domain weights for the limited angle CT reconstruction problem. He *et al.* [85] further proposed an inverse Radon transform approximation framework which resembles the FBP reconstruction steps. They constructed a neural network with three dedicated components (a fully connected filtering layer, a sinusoidal backprojection layer, and a residual CNN) corresponding to projection filtering, backprojection, and post-processing. They demonstrated that the approach outperforms TV-based iterative reconstruction for low-flux and sparse-view CT. Li *et al.* [86] also proposed an iCT-Net which consists of four major cascaded components that are also analogous to the FBP reconstruction. This approach can achieve accurate

reconstructions under various data acquisition conditions, such as sparse-view and truncated data.

### D. DL for Iterative Reconstruction

DL is also applied to iterative reconstruction methods for different purposes, including regularization design, parameter tuning, optimization algorithms, and reconstruction results improvement. Wu *et al.* [87] proposed regularizations trained by an artificial neural network for PWLS reconstruction of low-dose CT, which can learn more complex image features and thus outperform the TV and dictionary learning regularizations. Chen *et al.* [88] learned a CNN-based regularization for PWLS reconstruction and found it can preserve both edges and regions with smooth intensity transition without staircase artifacts. Gao *et al.* [89] constructed a CNN texture prior from previous full-dose scan for PWLS reconstruction of current ultralow-dose CT images. One drawback of the conventional model-based iterative reconstruction is the manual tuning of the hyperparameter which controls the tradeoff between data fidelity and regularization. Shen *et al.* [90] used deep reinforcement learning (RL) to train a system which can automatically adjust the parameter and demonstrated that the parameter-tuning policy network is equivalent or superior to manual tuning. Chen *et al.* [91] proposed a learned experts' assessment-based reconstruction network (LEARN) for sparse-data CT that learns both regularization and parameters in the model. He *et al.* [92] also proposed a DL-based strategy for PWLS to simultaneously address regularization design and parameter selection in one optimization framework. DL can also be used to modify optimization algorithms for iterative reconstruction. Kelly *et al.* [93] incorporated DL within an iterative reconstruction framework, which utilizes a CNN as a quasi-projection operator within a least-squares minimization procedure for limited-view CT reconstruction. Gupta *et al.* [94] presented an iterative reconstruction method that replaces the projector in a projected gradient descent algorithm with a CNN, which is guaranteed to converge and under certain conditions converging to a local minimum for the nonconvex inverse problem. They also showed improved reconstruction over TV or dictionary learning-based reconstruction for sparse-view CT. Adler and Öktem [95] proposed learned primal–dual algorithm for CT iterative reconstruction which replaces the proximal operators with CNNs. They demonstrated this DL-based iterative reconstruction is superior to TV regularized reconstruction and DL-based denoising methods.

### E. Domain Transformation Methods

Researchers have also leveraged DL to map sensor domain measurements to image domain reconstruction directly. For example, Zhu *et al.* [96] proposed an automated transform by manifold approximation (AUTOMAP) approach to achieve end-to-end image reconstruction. But the approach requires high memory for storing the fully connected layer and thus, limits this approach to small size image reconstruction. Fu and De Man [97] recursively decomposed the reconstruction problem into hierarchical subproblems and each
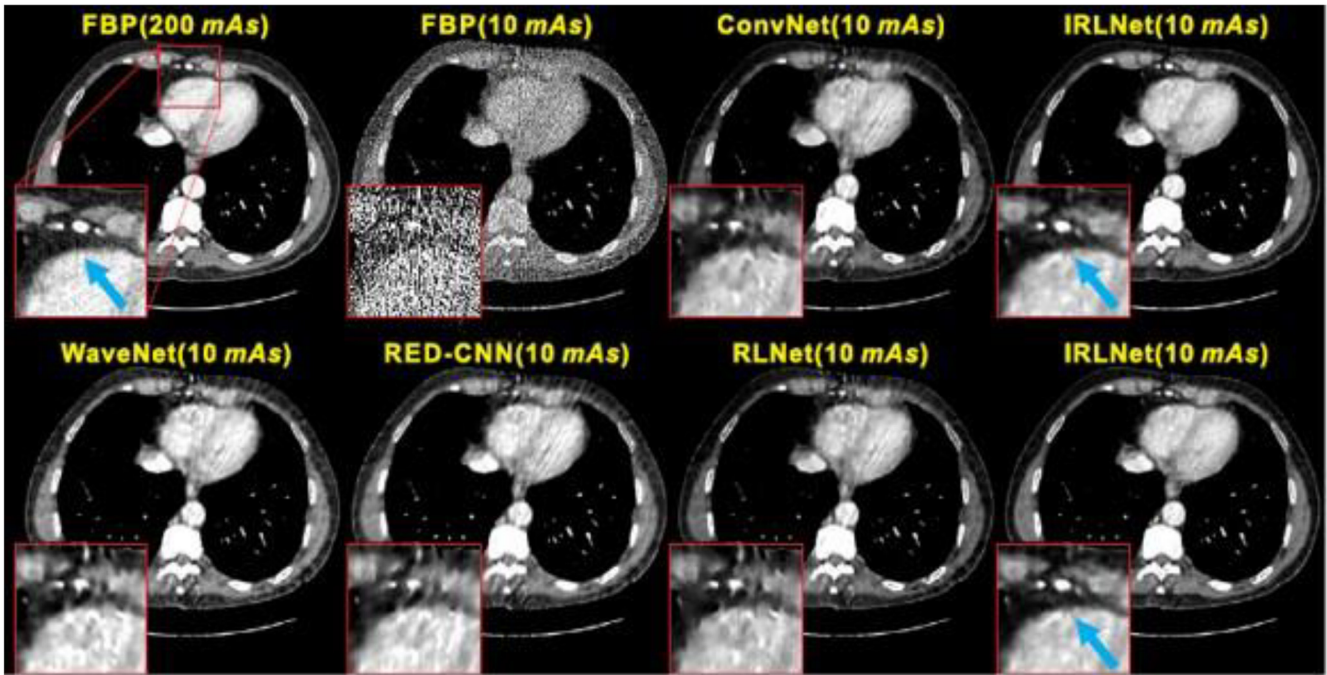
Fig. 2.　Comparison of ultralow-dose CT images (120 kVp, 10 mAs) reconstructed by six different approaches. All the images are displayed in the same window of $[-160, 240]$ HU. Figure reprinted from Wang *et al.* with permission.

can be solved by a neural network. Shen *et al.* [98] proposed a deep-learning model trained to map a single 2-D projection to 3-D CBCT images by using patient-specific prior information. They introduced a feature-space transformation between a single projection and 3-D volumetric images within a representation–generation framework. Inspired by this work, Lei *et al.* [99] investigated a GAN-based approach with perceptual supervision to generate instantaneous volumetric images from a single 2-D projection for real-time imaging in lung SABR.

It should be noted that many reconstruction methods mentioned above are generic and may be used for both diagnostic imaging and image-guided radiotherapy. We are expecting to see more evaluations and validations of these techniques on specific applications in radiotherapy. By now, there are still many concerns on the stability and reliability of the DL-based reconstruction methods because of their black-box nature. More research efforts are needed to interpret these DL models and improve the robustness and accuracy of DL-based reconstruction.

Although with great potential, we have not seen reports of using DL-based reconstruction algorithms for radiotherapy applications in the clinic. The DL-based reconstruction algorithms from GE Healthcare and Canon Medical Systems have received FDA 510(k) clearance. Solomon *et al.* [100] studied the noise and spatial resolution properties of DL-based reconstruction from GE, and found it can substantially reduce noise compared to FBP while maintaining similar noise texture and high-contrast spatial resolution but also with degraded low-contrast spatial resolution. Since the quality of the reconstructed image is so crucial in radiotherapy, more clinical assessments and comparisons

are needed before we can confidently adopt them in the clinic.

## III. Image Registration

Image registration is an important component for many medical applications, such as motion tracking [101], segmentation [101]–[103], image-guided radiotherapy [104], [105], and so on. Image registration is to seek an optimal spatial transformation which aligns the anatomical structures of two or more images based on its appearances. Traditional image registration methods include optical flow [106], demons [107], ANTs [108], HAMMER [109], ELASTIX [110], and so on. Recently, many DL-based methods have been published and achieved state-of-the-art performances in many applications [111]. CNN uses multiple learnable convolutional operations to extract features from the images. Many types of architectures exist for CNN, including the AlexNet, U-Net, ResNet, DenseNet, and so on. Due to its excellent feature extraction ability, CNN has become one of the most successful models in DL-based image processing, such as image segmentation and registration. Early works that utilized CNN for image registration attempted to train a network to predict the multimodal deep similarity metric to replace the traditional image similarity metrics such as mutual information in the iterative registration framework [112], [113]. It is important to ensure the smoothness of the first-order derivative of the learned deep similarity metrics in order to fit them into traditional iterative registration frameworks. The gradient of the deep similarity metric with respect to the transformation can be calculated using the chain rule. The major drawback of this method is that it inherits the iterative nature of the

## DL-based Medical Image Registration

### Supervised DVF Prediction

**Ground Truth DVF Generation**

- Manual/Automatic masks, landmarks
- Statistical Appearance Model generated DVF
- Traditional Registration generated DVF
- Randomly generated DVF

**Loss Function**

- Masks Overlap, Landmark distance
- DVF Error (Predicted v.s. GT)
- Deformed Image Regularization
- DVF Regularization

### Unsupervised DVF Prediction

**Image Similarity**

- Sum of Square Difference
- Normalized Correlation
- Mutual Information
- Deep Similarity

**Loss Function**

- Image Similarity Error
- Adversarial loss
- Deformed Image Regularization
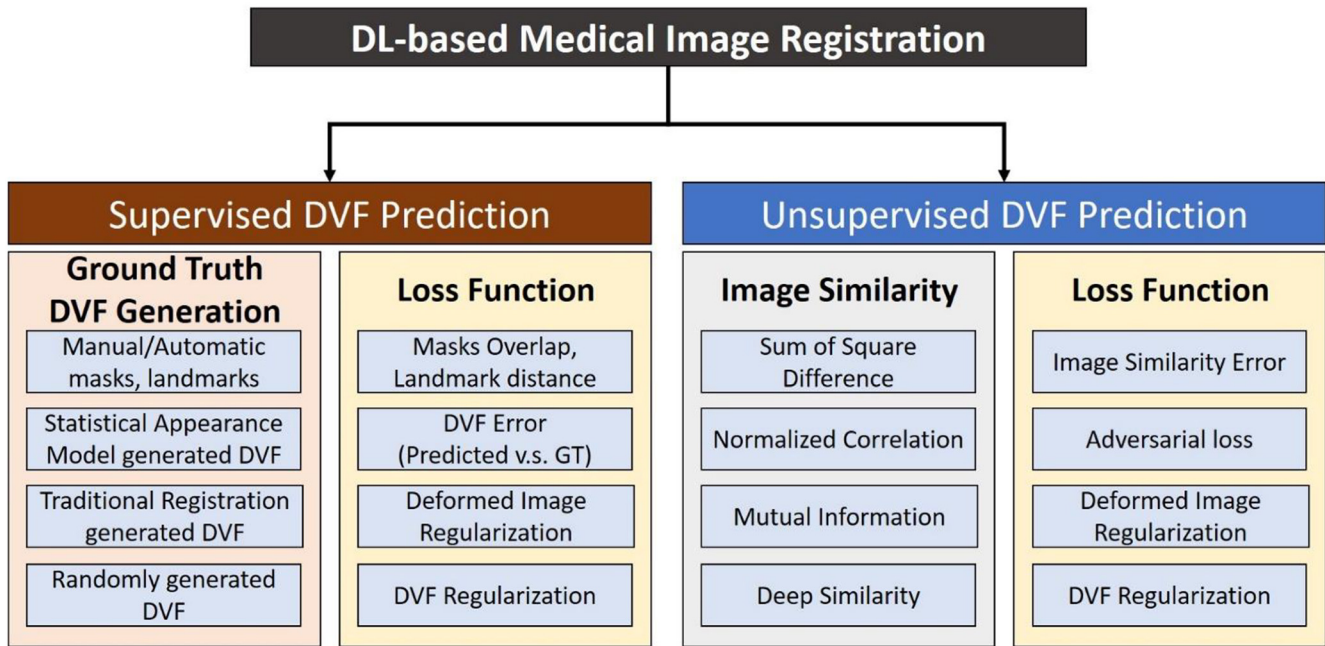- DVF Regularization

Fig. 3.  Supervised and unsupervised DVF prediction methods.

traditional registration frameworks. To enable fast registration, many CNN-based image registration methods have been proposed to directly infer the final deformation vector field (DVF) in a single/few forward predictions. In this section, we focus on this type of registration method since there is a clear trend toward direct DVF inference for DL-based fast image registration.

According to the network training strategy, this type of CNN-based image registration method can be grouped into two broad categories which are supervised DVF prediction and unsupervised DVF prediction, as shown in Fig. 3. Supervised DVF prediction refers to DL models that are trained with the known ground-truth transformation between the moving and the fixed images. On the contrary, unsupervised DVF prediction does not need the ground-truth transformation for network training. For supervised DVF prediction, the ground-truth DVF can be generated artificially using mathematical models or by traditional registration algorithms. The DVF error between the predicted and ground-truth DVFs can be minimized to train the network. For unsupervised DVF prediction, ground-truth DVF is not needed, however, robust image similarity metrics are necessary to train the network to maximize the image similarity between the deformed images and the fixed images. Over the last several years, there has been an increasing number of publications on CNN-based direct DVF inference methods.

To investigate the trend of the number of publications that used supervised and unsupervised learning methods for image registration, we have collected 100+ publications from various databases, including Google Scholar, PubMed, Web of Science, Semantic Scholar, and so on. Keywords including but not limited to ML, DL, learning-based, CNN, image registration, image fusion, and image alignment were used. Fig. 4 shows the number of publications from the year of
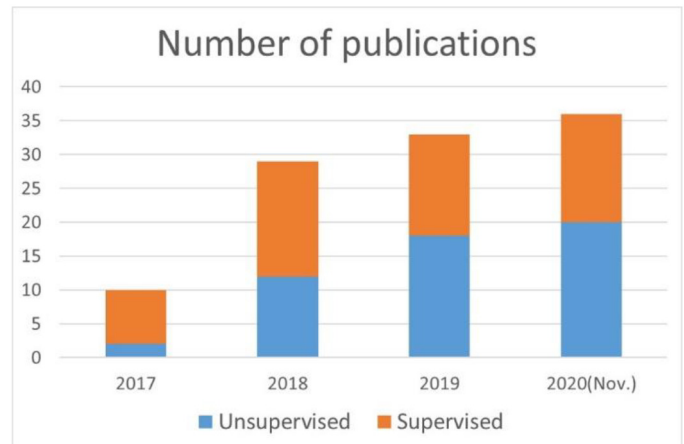


Fig. 4.  Overview of the number of publications in DL-based medical image registration.

2017 through November of 2020. In 2017 and 2018, the supervised methods are clearly more prevalent. From 2019, the unsupervised methods have become slightly more popular than the supervised methods.

### A. Supervised Transformation Prediction

Supervised transformation prediction aims to train the network with ground-truth transformations. However, the ground-truth transformation is usually unavailable in practice. Various methods have been proposed to generate/estimate the ground-truth transformation, including manual/automatic masks contouring, landmark detection/selection, artificial transformation generation, traditional registration-calculated transformation, and model-based transformation generation.

TABLE I
SELECTED SUPERVISED TRANSFORMATION PREDICTION METHODS

| References | ROI | Patch-based | Modality | Transformation |
|---|---|---|---|---|
| [8] | Cardiac | No | MR | Deformable |
| [13] | Brain | Yes | MR | Deformable |
| [15] | Brain | Yes | MR | Deformable |
| [16] | Pelvic | Yes | MR-CT | Deformable |
| [17, 18] | Prostate | No | MR-US | Deformable |
| [19] | Lung | Yes | CT | Deformable |
| [20, 21] | Lung | No | CT | Deformable |
| [12] | Brain | Yes | MR | Deformable |
| [24] | Brain | No | T1, T2, Flair | Affine |
| [25-27] | Lung | Yes | CT | Deformable |
| [29] | Prostate | No | MR-TRUS | Affine + Deformable |
| [33] | Pelvis | No | CT-CBCT | Deformable |
| [35] | Cardiac | No | MRI | Deformable |

TABLE II
SELECTED UNSUPERVISED TRANSFORMATION PREDICTION METHODS

| References | ROI | Patch-based | Modality | Transform | Similarity |
|---|---|---|---|---|---|
| [9] | Brain | No | MR | Deformable | SSD |
| [10] | Brain | No | MR | Affine | MI |
| [11, 12] | Brain | Yes | MR | Deformable | Intensity & gradient difference |
| [22] | HN | Yes | CT | Deformable | MSE |
| [28] | Cardiac | No | MR | Deformable | CC |
| [34] | Lung | No | MR | Deformable | MSE |
| [36] | Brain | No | MR-US | Deformable | MSE after intensity mapping |
| [37] | Cardiac, Lung | Yes | MR, CT | Affine and Deformable | MSE |
| [38, 39] | Brain | No | MR | Deformable | CC |
| [40] | Liver | No | CT | Deformable | CC |
| [41, 42] | Brain | No | MR | Deformable | CC |
| [44] | Liver | No | CT | Deformable | CC |
| [46] | Abdomen | Yes | CT | Deformable | CC |
| [47] | Abdomen | Yes | CT | Deformable | CC |
| [49] | Lung | Yes | CT | Deformable | CC |
| [51, 52] | Prostate | No | MR-TRUS/CBCT | Deformable | None |
| [54] | Lung, Cardiac | Yes | CT, MRI | Deformable | CC |

Table I shows a list of selected references that used supervised transformation prediction. Salehi *et al.* [114] trained a CNN-based rigid image registration for fetal brain MR scans. The network was trained to predict both rotation and translation parameters using datasets generated by randomly rotating and translating the original 3-D images. Eppenhof *et al.* [20] trained a CNN to perform 3-D lung deformable image registration (DIR) using synthetic transformations. The network was trained by minimizing the mean-square error (MSE) between the predicted DVF and the ground-truth DVF. A target registration error (TRE) of 4.02±3.08 mm was achieved on DIRLAB [115], which was worse than 1.36±0.99 mm [116] that was achieved when using the traditional DIR method. The TRE was later reduced from 4.02±3.08 mm to 2.17±1.89 mm on DIRLAB datasets using a U-Net architecture [21].

Instead of using artificially generated transformation as ground truth, Sentker *et al.* proposed generating the DVF using PlastiMatch [117], NiftyReg [118], and VarReg [119] as ground truth [19]. The authors showed that the network trained using VarReg had better performance than those trained using PlastiMatch and NiftyReg on DIRLAB [115] datasets. The best TRE values they achieved on DIRLAB was 2.50±1.16 mm, which was not better than the network trained using artificially generated transformations. Statistical appearance models (SAM) have also been used by Uzunova *et al.* [120] to generate a large and diverse set of training image pairs with known transformations from a few sample images. They showed that CNNs learned from the SAM-generated transformation outperformed CNNs learned from artificially generated and affine registration-generated transformations. Sokooti *et al.* [26] used a model of respiratory motion to simulate ground-truth DVF for 3D-CT lung image registration. They have outperformed models that were trained using artificially generated transformations. They achieved a TRE of 1.86 mm for the DIRLAB datasets. Instead of using the artificially generated dense DVF, higher-level correspondence information such as masks of anatomical organs were also used for network training [18], [121]. Networks trained using the higher level of correspondences such as organ masks or landmarks are often called weakly supervised methods since the exact dense voxel-level transformation was unknown during the training. It is called weakly supervised also because the higher level of correspondence was not required in the inference stage to facilitate fast registration.

One major limitation of supervised transformation prediction is that the generated transformation may not reflect the true physiological motion, resulting in a biased model toward the artificially generated transformation prediction. It is possible to mitigate this problem using better transformation models to generate various training image pairs which simulate realistic transformations.

### B. Unsupervised Transformation Prediction

The loss function definition in supervised transformation prediction methods was straightforward. However, for unsupervised transformation prediction, it was not so straightforward to define a proper loss function without knowing the ground-truth transformations. Fortunately, the spatial transformer network (STN) which allows spatial manipulation of data during training was proposed [122]. The STN can be readily plugged into existing CNN architectures. The STN was used to deform the moving image based on the current predicted DVF to generate the deformed images which were compared to the fixed image to calculate image similarity loss. Table II shows a list of selected references that used unsupervised transformation predictions with their respective similarity metrics. SSD stands for the sum of squared difference, MI stands for mutual information, MSE stands for mean squared error, and CC stands for cross-correlation.

An unsupervised CNN-based registration method, called VoxelMorph was proposed for MR brain atlas-based registration [38], [39]. The VoxelMorph has a U-Net like architecture. With STN, the image similarity between the deformed images and the fixed images was maximized during training. The predicted transformation was regularized to have low local spatial variations. They have achieved comparable performance to the

ANT [108] registration method in terms of the dice similarity coefficient (DSC) score of multiple anatomical structures. Zhang [123] proposed a network to predict diffeomorphic transformation using trans-convolutional layers for end-to-end MRI brain DVF prediction. An inverse-consistent regularization term was used to penalize the difference between two transformations from the respective inverse mappings. The network was trained using a combination of an image similarity loss, a transformation smoothness loss, an inverse consistent loss, and an anti-folding loss.

Lei *et al.* [46] used an unsupervised CNN to perform 3-D CT abdominal image registration. A dilated inception module was used to extract multiscale motion features for robust DVF prediction. Besides the image similarity loss and DVF regularization loss, an adversarial loss term was added by training a discriminator. De Vos *et al.* [37] proposed a fast unsupervised registration framework by stacking multiple CNNS into a larger network for cardiac cine MRI and 3-D CT lung image registration. They showed their method was comparable to conventional DIR methods while being several orders of magnitude faster. Jiang *et al.* [124] proposed a multiscale framework with unsupervised CNN for 3-D CT lung DIR. They cascaded three CNN models with each model focusing on its own scale level. The network was trained using image patches to optimize an image similarity loss and a DVF smoothness loss. They demonstrated that the network trained on SPARE datasets has good performance on the DIRLAB datasets. The same trained network could also be generalized to CT-CBCT and CBCT-CBCT registration without retraining or fine-tuning. Jiang *et al.* achieved an average TRE of 1.66±1.44 mm on DIRLAB datasets. Fu *et al.* [49] proposed an unsupervised whole-image registration for 3D-CT lung DIR. The network adopted a multiscale approach where the CoarseNet was first trained using downsampled images for global registration. Second, local image patches were registered to the image patches of the fixed image using a patch-based FineNet. A discriminator was trained to provide adversarial loss to penalize unrealistic warped images. They have outperformed some traditional registration methods with an average TRE of 1.59±1.58 mm on DIRLAB datasets.

Compared to supervised transformation prediction, unsupervised methods could alleviate the problem of lack of training datasets since the ground-truth transformation is not needed. However, without direct transformation supervision, DVF regularization terms have become more important to ensure plausible transformation prediction. So far, most of the unsupervised methods focused on unimodality registration since it is relatively easy to define image similarity metrics for unimodal registration than multimodal image registration.

Supervised transformation prediction methods are limited by the lack of known transformations in the training datasets. Artificial transformations could introduce errors due to the inherent differences between the artificial and realistic transformations. Model-based transformation generation which could simulate highly realistic transformation has been shown to alleviate the lack of realistic ground-truth transformation. On the other hand, unsupervised methods need extensive transformation regularization terms to constrain the predicted

transformation since ground-truth transformation is not available for supervision. One challenge is to efficiently determine the relative importance of each regularization term. Repeated trial and evaluation were often performed to find an optimal set of transformation regularization terms that could help generate not only physically plausible but also a physiologically realistic deformation field for a certain registration task. Another limitation for unsupervised transformation prediction is that it relies on effective and accurate image similarity metrics to calculate similarity loss and train the network. However, multimodal image similarity metrics are usually more difficult to define and calculate than unimodal image similarity metrics. Therefore, there is a lack of multimodal unsupervised image registration methods as compared to unimodal image registration methods. Deep similarity metrics could be trained for multimodal image registration tasks and used in unsupervised transformation prediction. However, the training of deep similarity metrics often requires prealigned training image pairs which are difficult to obtain.

## IV. IMAGE SEGMENTATION

In radiation therapy, image segmentation can be described as the process where each pixel in an image is assigned a label, and pixels with similar labels are linked such that a visual or logical property is realized [125]. Resultant groupings of pixels with the same label are called delineations (i.e., segmentations). Once RT images are acquired, tumors and OARs are delineated, often by a physician or dosimetrist, to be incorporated into the treatment planning process. Manual segmentation has been reported to be the most time-consuming process of radiation therapy, introducing substantial inter- and intra-observer variability [126], and dependent on image acquisition and display settings [127]. To address these limitations, auto-segmentation is often employed. Auto-segmentation may be unsupervised, where only the image itself is considered and image intensity/gradient analysis is implemented, performing best with distinct boundaries [23]. Supervised segmentation, on the other hand, integrates prior knowledge about the image often in the form of other similarly annotated images to inform the current segmentation task (i.e., training samples).

Recent DL techniques [128], [129] are well poised for the task of accurate automatic segmentation with less reliance on organ contrast [130]–[132] as the algorithm is designed to acquire higher order features from raw data [128]. DNNs learn a mapping function between an image and a corresponding feature map (i.e., segmented ground truth) by incorporating multiple hidden layers between the input and output layer. The U-Net [5] is a DNN architecture that has shown great promise for generating accurate and rapid delineations for applications in RT [133]. The U-Net "U-shaped" architecture shown in Fig. 5 was inspired by the original fully convolutional network from Long *et al.* [134] and was initially implemented by Ronneberger *et al.* [5] in 2015 to segment biomedical image data using 30 annotated image sets. The U-Net has an additional expansion pathway that replaces the maximum pooling operations with upconvolutions to increase the resolution
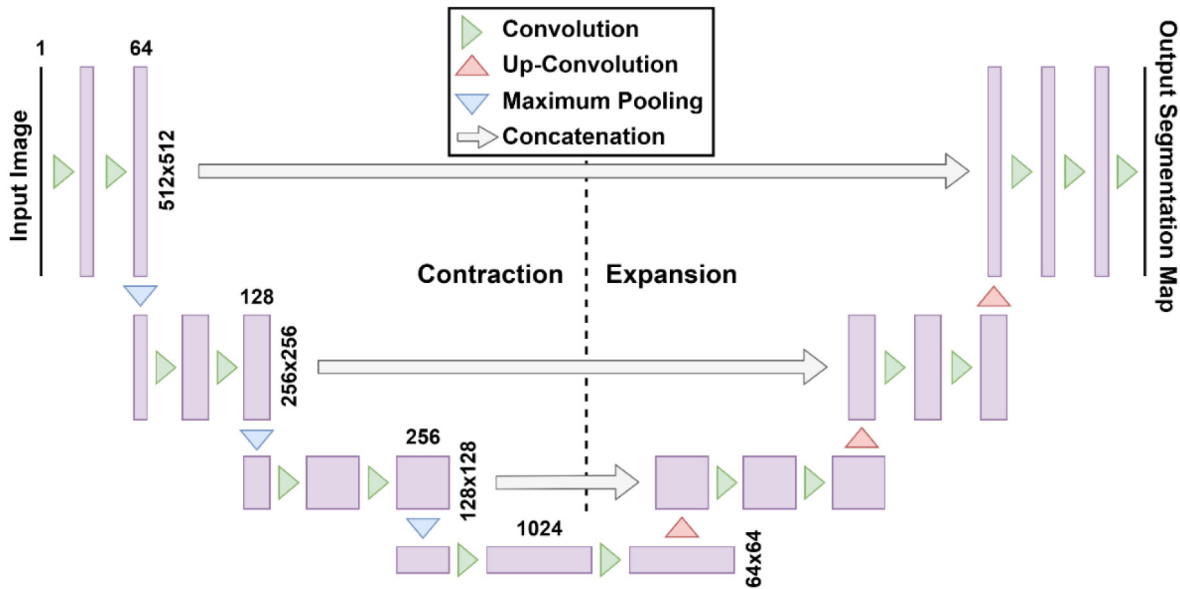
Fig. 5. Architecture for original U-Net by Ronneberger *et al.* [5] with the contraction path shown on the left and the expansion path shown on the right. The original input image has a size of 512 × 512. Feature maps are represented by purple rectangles with the number of feature maps on top of the rectangle.

TABLE III
EXAMPLE APPLICATIONS OF DL INTO MEDICAL IMAGE SEGMENTATION FOR RADIATION THERAPY PLANNING PURPOSES

| Author | Disease Site | Image Type, Number of cases | Algorithm | Outcome |
|---|---|---|---|---|
| Zhu *et al.*[1] | Head and Neck | Planning CT, 261 | 3D U-Net, whole image | -9 OAR contours generated in ~0.12 seconds<br>-Increased DSC ~3% |
| Lei *et al.*[2] | Male Pelvis | Cone-beam with synthetic MRI, 100 | CycleGAN | -DSC >0.9 and MDA <1.0 mm for bladder, prostate, rectum from ground truth delineations |
| Dong *et al.*[3] | Thorax | CT, 35 | U-Net GAN | -DSC>0.87 and MDA <1.5 mm for lungs, cord, heart<br>-DSC esophagus ~0.75<br>-Negligible dosimetric differences |
| van der Heyden *et al.*[4] | Brain | Dual-energy CT, 14 | 2-step 3D U-Net | -Quantitatively and qualitatively outperformed atlas method for all organs at risk but optic nerves |
| Chen *et al.*[6] | Abdomen | 3T MRI, 102 | 2D U-Net | -9/10 OARs had DSC 0.87-0.96<br>-Duodenum DSC ~0.8 |
| Fu *et al.*[7] | Abdomen | 0.35 T, MR-linac, 120 | CNN with correction method | -4/5 OARs had DSC >0.85 (liver, kidneys, stomach, bowel)<br>-Duodenum DSC ~0.65 |

of the feature maps, a desirable feature for medical image segmentation. The original 2-D U-NET was quickly implemented into 3-D volumetric inputs to train using the entire dataset and annotations simultaneously to improve segmentation continuity, including multichannel inputs of different image types [135] (i.e., MRI, CT, etc.). Overall, the U-Net is an end-to-end solution has shown a remarkable potential to segment medical images, even when the amount of training data is scarce [136]. Various DNNs have also been applied to medical image segmentation [133], including deep CNNs with adaptive fusion [137] or multistage [138] strategies, as well as GANs [139].

Data scarcity may be a challenge in radiation therapy. Publicly available annotated "ground-truth data" for training and validation are available through The Cancer Imaging Archive [140]. Several other strategies are employed to improve the variability and diversity of available data—without new unique samples—which is referred to as data augmentation. Data augmentation has been shown to improve auto-segmentation accuracy and prevent model overfitting [135], [137], [141]. Examples of augmentation include image flipping, rotation, scaling, and translation (pixels/axis). Other emerging areas of data augmentation include integrating interfractional data such as incorporating daily CBCTs for patients to increase segmentation accuracy in radiation therapy [142] and using transfer learning to generate new training imagesets from other modalities. Typical endpoints of AI segmentation include qualitative review or comparison with ground-truth labels using overlap metrics such as the DSC or distance metrics such as the Hausdorff distance or mean distance to the agreement.

Applications of AI for segmentation in RT planning typically fall into two main classes: 1) OAR and 2) lesions. Table III outlines a few state-of-the-art examples of each with key findings, with a comprehensive list available in other references [143].

Emerging areas of interest include segmenting substructures of OAR including the cardiac substructures [135], applications for adaptive radiation therapy [7], and longitudinal response assessment. Other areas under development include optimizing loss functions, such as integrating DSC less, unweighted DSC loss, or focal DSC loss, with tunable hyperparameters [144] to better address segmentation accuracy for small structures [145].

## V. Image Synthesis

In this section, we focus on generative modeling of information content across imaging modalities relevant to the radiotherapy workflow. Generative modeling, in this context, refers to capturing information about the data distribution associated with each modality thus enabling translation across different modalities. Within generative modeling, we limit our review to techniques that are most suitable to represent unpaired data obtained in real clinical settings where obtaining one-to-one correspondences between modalities may not be feasible. Among these techniques, GANs have proven to be very successful. GANs are a framework that optimize an objective by running a min–max two-player game between two networks. One of the networks, called the generator, tries to learn the data distribution by trying to fool the other network, called the discriminator, which simultaneously tries to differentiate between real images and fake images created by the generator.

GAN-based approaches have also shown to be superior when paired data is available where they eliminate the need to design handcrafted losses in the image space [146]. We present methods proposed to aid different stages of the workflow ranging from image acquisition to treatment delivery. Specifically, we discuss MRI to CT, CBCT to CT, and CT to PET translation. We focus our review on techniques aimed at intermodality translation due to the large variation in content representation across these modalities. For example, CT images capture electron densities through Hounsfield units whereas MRI images are generated based on the excitation and relaxation of hydrogen protons. These differences in representation allow for effective demonstration of the capacity of GAN-based image translation approaches to learn complex mappings.

### A. MRI to CT Translation

MRI-only radiotherapy can provide multiple benefits to the patient and clinic due to its immense flexibility in imaging physiological and functional characteristics of tissue, combined with much superior soft-tissue differentiation. It also avoids additional risk induced by subjecting the patient to ionizing radiation via CT. However, MRI lacks the ability to provide electron densities which is explicitly needed in radiation dosimetry transport calculations. Generative modeling of information content between MRI to CT modalities can allow for obtaining electron densities from MRI. This is done by translating it into a synthetic CT (sCT) while retaining structural information present within the original MRI scan itself. Fig. 6 shows the GAN framework to translate MRI to CT. In order to effectively translate between

the modalities, suitable input–output representations need to be determined, and this is done as follows: 1) constructing a mapping between paired MRI-CT data from the same patient (registered to ensure correspondence) or 2) unpaired MRI-CT data from the same patient or across patients. For MRI-to-CT translation in nasopharyngeal cancer treatment planning, Peng et al. [147] used conditional GANs for paired data and CycleGANs for unpaired data. CycleGANs ensure reliable translation by enforcing cycle consistency in the MRI-to-CT translation [148]. They use 2-D U-Net-based generators that operate on a slice-by-slice basis and 6-layered convolutional discriminators. Wolterink et al. [149], Lei et al. [150], and Liu et al. [151] used paired data from the same patients for MRI-to-CT translation for treatment planning of brain tumors and brain/pelvic tumors, respectively. Although they use registered MRI-CT data from the same patient, they employed CycleGANs to account for local differences between the spatial representation across these modalities. Wolterink et al. used data from sagittal slices on a slice-by-slice basis and used the default CycleGAN setup including a patch-based discriminator to preserve high-frequency features. Chen et al. [152] operated on 3-D patches of smaller sizes ($32^3$) and implemented dense blocks in the generator to capture multiscale information. They also proposed the use of novel mean P-distance ($l_p$ norm) and spatial gradient differences as cycle-consistency losses to avoid blurriness and promote sharpness, respectively.

In terms of results, Peng et al. reported mean absolute Hausdorff unit (HU) differences (MAE) within the body of 69.6±9.27 for the paired approach and 100.6±27.39 for the unpaired approach. Wolterink et al. reported 73.7±2.3 HU MAE and Lei et al. reported 57.5±4.7 HU MAE for the brain.

These methods are not directly comparable since they use different data, but they do represent an estimate of their quantitative performance. However, it is difficult to ascertain clear clinical applicability based solely on these metrics. Consider a case where the average HU values are biased by strong deviations in areas belonging to the tumor where other areas are quite accurate. If this is placed in comparison to another case where smaller but uniform deviations are present across the entire scan, which scenario would be more clinically relevant? These metrics fail to answer these questions in entirety. Peng et al. provided additional metrics such as comparing dose distributions of translated CT with the reference CT with 2%/2-mm gamma passing rates of (98.68%±0.94%) and (98.52%±1.13%) for the paired and unpaired approaches, respectively. This gives a better idea of the dosimetric accuracy of implementing their methods in a treatment planning system (TPS).

### B. CT to PET Translation

PET/CT scans can play a crucial role in combining anatomical and functional information to pinpoint metabolic activity and may provide better information in the contouring process for treatment planning. Synthesizing virtual PET from CT-only workflows can eliminate the need for the more costly PET/CT scan. Additionally, this reduces cost and complexities
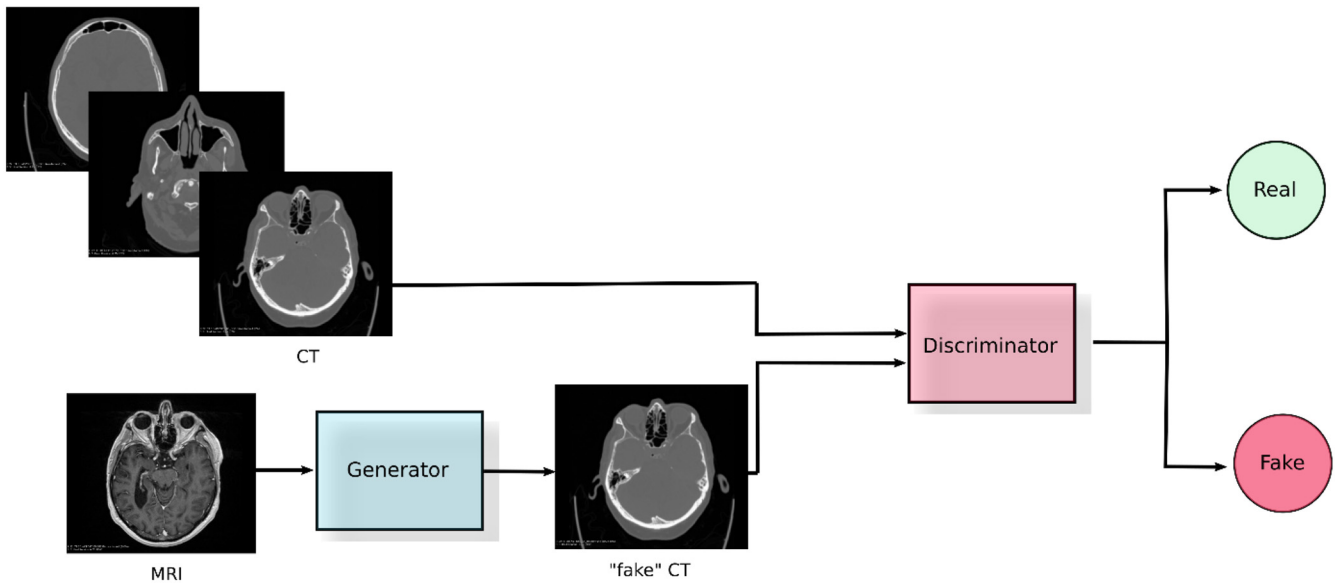
Fig. 6. GAN framework showing MRI to CT image translation. The generator tries to learn the data distribution of MRI and CT images and uses this learned representation to convert an MRI image to a fake CT image. The generator learns this representation by trying to fool the discriminator while it is comparing real and fake CTs and learning to differentiate between them.
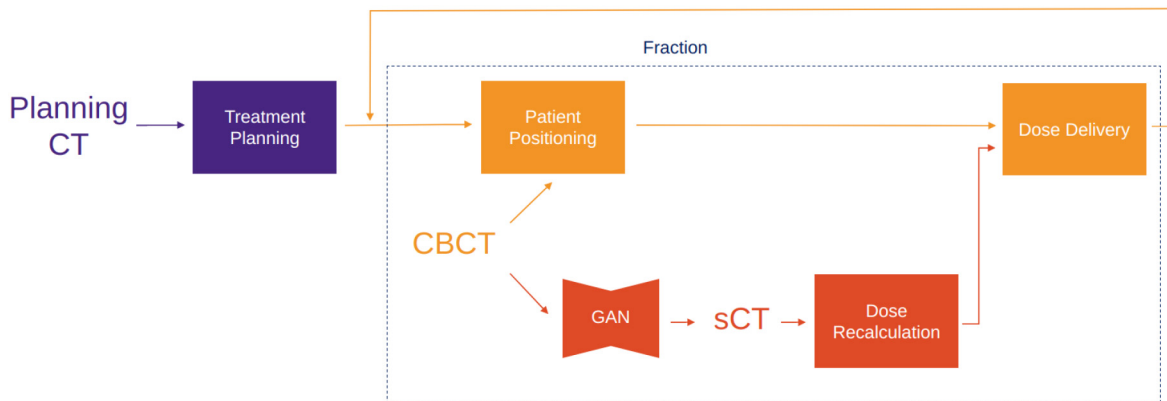


Fig. 7. Workflow for adaptive radiotherapy enabled by dose calculation from CBCT images. Prior to dose recalculation, a CBCT image is converted to an sCT, providing correct HU values and removing artifacts, while preserving the anatomy.

such as storage of radiotracers associated with PET imaging. Isola *et al.* proposed a conditional-GAN (cGAN) [146]-based method to generate PET from contrast-enhanced CT-scans for false positive reduction in lesion detection within the liver. They used paired PET/CT data and performed a transformation to align and interpolate the PET to the CT. A first synthesized estimate of PET is performed using a fully convolutional variant of VGG [153] followed by a cGAN applied on channel concatenated input comprising of CT and the previous PET estimate. In this two-stage network, while optimizing over the image losses, SUV-based weighting is applied to provide better results in PET regions with high-SUVs. The proposed method obtains a mean absolute SUV difference of $0.73\pm0.35$ across all regions and $1.33\pm0.65$ in high-SUV regions. Finally, false positives are shown to be reduced from 2.9% to 2.1% in liver lesion detection when using the generated PET information for detection. Bi *et al.* [154] explored the synthesis of PET image from paired PET/CT data for lung cancer patients and propose three different methods exploiting varied input representations.

All three methods apply a U-Net-based cGAN but vary in terms of input provided to the model: 1) binary label-map of the tumor annotations; 2) CT image; and 3) channel-combined CT with binary label-map. The last method provided the results closest to the real-PET image with an SUV MAE of 4.60. The authors suggested that this synthesized PET can be used to form training data for PET/CT-based prediction models. Further, they formulated a potential extension of their work to combine both real and synthetic PET to boost the training samples in an attempt to boost generalization.

### C. CBCT to CT Translation

Radiotherapy treatment planning starts with acquiring a CT scan of the patient, usually denoted as a planning CT (pCT). Dose calculation as well as beam energy and arrangement are derived from the pCT, comprising a treatment plan. Commonly, in order to correctly position the patient at each fraction of the dose delivery, a CBCT scan is obtained.

However, CBCT can potentially provide insight into the anatomical changes that occurred over the course of the treatment and could enable adaptive radiotherapy by leveraging that information to adjust the treatment plan. Unfortunately, treatment replanning is not possible with CBCT scans as they are noisier, contain more artifacts and have inferior soft-tissue contrast compared to fan-beam CT. One way to make use of the information in a CBCT scan is by using DIR to map pCT to the anatomy of CBCT [155], producing a scan with HU values of the pCT with the latest anatomy, in the literature referred as deformed pCT (dpCT) or virtual CT (vCT). On the other hand, generative DL methods may provide a faster and potentially superior alternative to treatment replanning based on dpCT or other techniques (lookup table [156], Monte Carlo [157], and scatter correction with pCT prior [158]) by synthesizing a CT scan from an input CBCT scan. Such sCT should have all the characteristics of a CT scan while preserving the anatomy from the CBCT. Fig. 7 shows the workflow of CBCT-based dose calculation for adaptive radiotherapy. Maspero *et al.* [159] trained four standard CycleGAN [148] models on lung, breast, and head-and-neck scans—three for each anatomical site separately and one model on all sites jointly. They showed that a single model for all three anatomical sites performs comparably against models trained per anatomical site, which would simplify its possible clinical adoption. The results are evaluated using rescanned CT (rCT) as ground truth, where rCT is a CT scan that is acquired at the same fraction as the CBCT in question. The reported HU MAE for sCT is 5312, 6618, and 8310 for head-and-neck, lung, and breast, respectively. Liang *et al.* [160] used a standard CycleGAN model, but evaluate the HU accuracy and dose calculations against dpCT instead of rCT. Furthermore, they performed an evaluation of anatomical accuracy of sCT using deformable head-and-neck phantoms. The phantom allows for a simulation of the patient at the beginning of the treatment and after a few fractions of the treatment, where the tumor shrinkage is observed. This provides a CBCT and rCT scan with identical anatomy, which can be used to assess if and how the translation of CBCT to sCT affects the representation of the anatomy. They concluded that the CycleGAN model has higher anatomical accuracy than DIR methods.

### D. Clinical Perspective and Future Applications of Synthetic Imaging

In the radiotherapy workflow, medical images are one of the most important sources of data used in decision making. Imaging is used in all the steps of patient care in oncology: diagnosis, staging, treatment planning, treatment delivery, and disease follow-up. Based on the anatomical site of a tumor and the specific properties we want to investigate, some imaging modalities might be more appropriate than others. For example, an MRI scan is suggested for malignancies located in the pelvic region, due to the large presence of soft tissues compared to bone structures. This allows an improved ability to contour lesions and surrounding organs. CT is the predominant modality for staging lung tumors, but recently MRI scans, and more specifically diffusion-weighted imaging (DWI) sequences, are used to evaluate the involvement

of mediastinal lymph nodes with higher sensitivity than CT, with implications on better staging and treatment decisions. Finally, it is well established that treatment planning always requires the acquisition of a CT scan, while PET imaging is often used prior to treatment planning to evaluate the presence of metastasis and the degree of suspicion of the identified lesions, and within/after treatment fractions to quantify treatment response [161]. Such examples highlight the need for multiple modalities of imaging to fully capture the complexity of human anatomy and tumor tissues in tandem with the specific tasks that we want to accomplish. This would eventually improve the decision-making process. In an ideal scenario, all modalities of imaging for the patient would be available, but that is far from being an achievable or practical solution. Several reasons stand behind this evidence.

*1) Cost Effectiveness:* Scanning patients costs both hospitals and the patients themselves. These costs may be reimbursed by healthcare providers or insurance companies. With an increasing number of new and diverse imaging technologies a growing demand for cost-effectiveness analysis (CEA) in imaging technology assessment is induced. As pointed out by Sailer *et al.* [162], when assessing the cost effectiveness of diagnostic imaging, the initial question is whether adding an imaging test in a medical pathway does indeed lead to improved medical decision making. One of the most significant examples was the lung cancer screening trial entitled NLST [163], which showed that participants who received low-dose helical CT scans had a 15%–20% lower risk of dying from lung cancer than participants who received standard chest X-rays. In radiotherapy, a study [164] highlighted the costs related to various radiological imaging procedures in image-guided radiotherapy of cancers, based on standard billing procedures. The median imaging cost per patients was $6197, $6183, $6358, $6428, $6535, and $6092 from 2009 to 2014, respectively. This seems to highlight an upcoming trend of reducing costs related to imaging. Unfortunately, it is not clear if this reduction in costs was associated with an optimization of the diagnostic imaging.

*2) Patient Safety and Comfort:* Recent studies indicated that repetitive imaging scans can deposit considerable radiation doses to some radiosensitive organs (e.g., heart) and could cause higher radiogenic cancer risks to the patients, with children being more impacted by this issue [165]. In a very utilitarian way, we might want to have an image with the highest possible contrast to be able to better identify suspicious structures, and to possibly obtain images of our patients within short intervals of time. This would allow using these images, for example, to perform a better evaluation of treatment response. Unfortunately, due to the physics of imaging, ionizing imaging methods give a signal to noise ratio proportional to the released dose to the body [166]. More dose leads to better contrast, but also increases the probability of radiation-induced effects to the patient. Another point is represented by images acquired with the injection of an intravenous (IV) contrast medium. IV contrast media are usually toxic substances that need to be expelled by our body. If it is true that images acquired with IV contrast media (e.g., contrast-enhanced CT) provide better resolution for specific

anatomic regions compared to images without them (e.g., conventional CT); it also stands that many patients might not be eligible for the injection of a contrast media, because of poor performance status, presence of comorbidities, or poor renal function [167]. Finally, some imaging modalities such as DWI require longer acquisition times compared to T1 and T2 sequences [168], with an impact on costs and also on patients' comfort.

*3) Differences in Imaging Acquisition Protocols and Interoperability:* Despite the presence of specific guidelines and recommendations for diagnostic imaging, each institution might adopt not only different image acquisition protocols, but might also be missing specific imaging modalities (or sequences), which might be the standard of care in another institution. This evidence has an impact when performing or designing multicenter institutional studies, especially if retrospective and with the aim of the development of image-derived biomarkers. For example, if a center has developed a prognostic model based on an image-derived biomarker obtained by processing a specific imaging modality, a large external validation of this biomarker might not be possible because this imaging modality may not be available in many institutions. Additionally, when considering quantitative imaging analysis via ML and more specifically radiomics [169], a recent review pointed how different acquisition settings (e.g., slice thickness, tube current, and reconstruction kernels) should be preferrable with respect to others, since they increase the reproducibility of the biomarkers [170]. It is, as mentioned before, not obvious that these acquisition settings are the same across all clinics. One possible solution, which is close to utopia, is to force each institution to acquire images with the same acquisition protocols. However, even if this was accomplished, the variability of scanner manufacturers, a well-known factor that impacts the stability of image-derived biomarkers, would still be difficult to tackle.

All the points presented above show that there is an open space for the application of DL-based synthetic imaging. Without going into details, applications can include the generation of multimodalities from a starting image (as explained in the case of the section MRI to CT of this article), augmentation of image quality without exposing the patient to additional dose (as explained in the case of the section CBCT to CT translation), but also the recent work that introduces fast DL reconstruction for DWI images [171]–[173]. Synthetic imaging is taking a prominent role in oncology. We refer the reader to the following publications as proof of some interesting clinical applications that DL for image generation can offer [174], [175], and to these more general reviews [176], [177].

## VI. AUTOMATIC TREATMENT PLANNING

In radiotherapy planning, a main objective is to deliver the prescribed dose accurately to the target, keeping the dose to OAR below acceptable limits and minimizing dose to surrounding, healthy tissue. The treatment planning process begins with the delineation of target volumes and OARs on a pCT. A set of dose constraints are defined for targets, OARs,
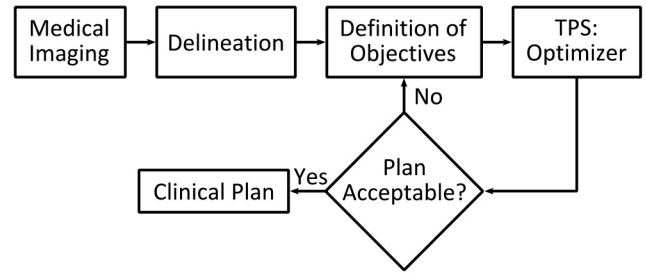


Fig. 8. Inverse treatment planning process begins with the delineation of target volumes and OARs on medical images. Optimization objectives are defined and passed to the optimizer algorithm of the TPS. If the resulting plan is not acceptable, tradeoff evaluation is performed to define new targets iteratively until a clinical plan is accepted.

and other regions of interest, typically dose–volume relations, stating the minimum or maximum dose that can be allowed to a given region. Intensity-modulated radiation therapy (IMRT) and volumetric modulated arc therapy (VMAT) are modern techniques that allow the treatment planner to create complex treatment plans by providing a set of such objectives that an optimizer algorithm will attempt to fulfill by inverse optimization, as illustrated by the workflow in Fig. 8. The optimizer often fails to achieve all the desired objectives, due to complicated patient geometry, limitations of the treatment modality or machine, etc. To improve the plan, the planner and physician discuss available options and clinical preference before adjusting the objectives for a new iteration of optimization. In recent years, automated treatment planning techniques have been developed, that aim to provide vastly improved starting points for the treatment planner, and even produce clinically acceptable treatment plans without human interaction.

The current planning workflow apparently relies heavily on humans (planner and physician). The reason for this workflow is twofold. First, there is no clear metrics to mathematically quantify plan quality. Although there are some plan quality scoring systems defined over the years [178], they might not necessarily reflect the most stringent clinical requirement for each patient. Second, for a specific patient, the best plan is unknown. The planning process has to explore the very high-dimensional solution space in a trial-and-error process to find out the optimal solution [179], [180]. This complex and cumbersome process poses substantial hurdles to plan quality and planning efficiency.

Due to extensive human involvement, the plan quality heavily depends on a number of human factors, such as the planner's experience, the planner–physician communications, the amount of effort and available time for treatment planning, the rate of human errors, etc. [181]–[183]. Suboptimal plans are often unwittingly accepted [183]–[185], deteriorating treatment outcomes [186].

Moreover, while modern computers can solve the optimization problem rapidly, the trial-and-error iterative planning process yields hours of planning time for a typical planner to generate a plan for the physician to review [187]–[189]. Multiple iterations between physician and planner are often needed, which extends the overall planning time up to one

week in some challenging tumor sites. This lengthy process strongly contributes to the delay between diagnosis and the start of RT, which has been shown to adversely affect treatment outcome. Moreover, patient's anatomy may change during the time waiting for treatment planning [190], [191], making the plan carefully designed based on the initial patient anatomy suboptimal for the changed anatomy [192], [193]. Additionally, the delayed treatment will increase the anxiety of patients who have already been overwhelmed by a cancer diagnosis and are eager to start treatment. Such delays can be particularly severe in low- and middle-income countries, with limited resources, capacity, staff, and expertise [194].

The problem of suboptimal plan quality and planning efficiency are indeed intertwined. Due to the low planning efficiency, the optimality of a plan is hard to guarantee for every patient in current practice given the strict time constraints. Heavy time pressure also increases the human error rate and may limit the availability of advanced treatment techniques. Auto-planning (AP) techniques are urgently needed to tackle these problems in the current planning process.

AP has already been around for some time and is rapidly improving. There are several research projects and in-house solutions being developed worldwide, and most commercial vendors of TPS have implemented some variety of AP tools. Studies show promising results for AP, but large-scale clinical implementation is not yet seen.

## A. Classical Auto-Planning Strategies

Classical AP tools can be divided into three categories: 1) treatment planner mimicking (TPM); 2) multicriteria optimization (MCO); and 3) knowledge-based planning (KBP).

*1) Treatment Planner Mimicking:* In TPM, the behavior and choices of the treatment planner are analyzed over time and converted into computer logic. Such logic can be a series of IF/THEN statements of decision making. Provided with a prioritized list of objectives the TPM follows its logic to create optimizer objectives that it tunes iteratively while steering the optimizer, in the same way that a human treatment planner would, pushing each objective as far as it can without degrading objectives of higher priority. The TPM is not as limited by time as a human planner, allowing it to perform more iterations, potentially leading to higher plan quality. Tol *et al.* [195] designed a system that automatically scans DVH lines in the Eclipse TPS (Varian Medical Systems, Palo Alto, USA) optimization window, and moved the mouse cursor to adjust on-screen optimization objectives. In a blinded test, automated head-and-neck cancer (HNC) plans were preferred over MP by an HNC radiation oncologist in 19/20 cases, and the method is now in clinical use. Several modern TPSs include possibilities for scripting, which have been used to develop in-house TPM by extracting DVH parameters directly from the TPS and automatically adjusting optimization objectives iteratively until the optimal solution is found [196]–[202]. A commercial TPM solution is available in Philips Pinnacle TPS (Philips Radiation Oncology Systems, Fitchburg, WI, USA) AP. Pinnacle AP works by defining a template, called
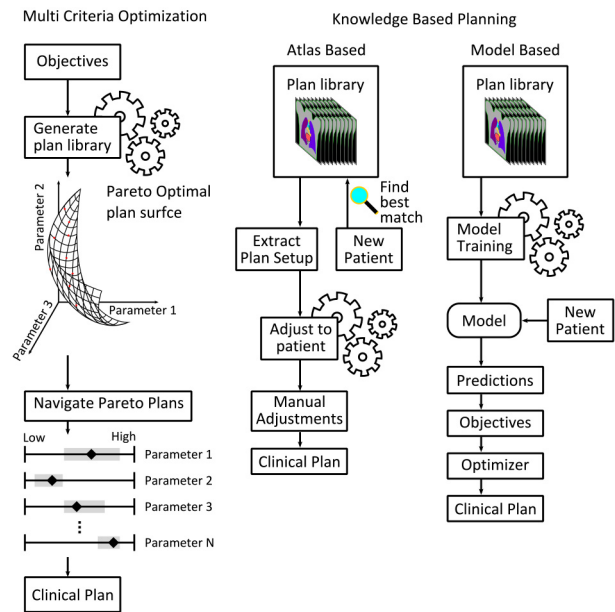


Fig. 9. Schematic view of some classical AP workflows. Left: Multicriteria optimization generates a library of Pareto-optimal plans from a set of objectives. In the schematic, the user navigates the plans *a posteriori* by use of sliders for each parameter. Right: KBP relies on a set of high-quality clinical plans. Atlas-based methods find the best matching patient in the library when introduced to a new patient, extract the plan, and adjust it to the new patient. Model-based approaches train a predictive model on the library plans. The model is used to predict parameters for new patients, which are used to generate objectives for the optimizer.

a *technique*, consisting of parameters, e.g., beam setup, dose prescriptions, and objectives for each disease site [203]. When the *technique* is applied to a new patient, the AP will iteratively optimize a treatment plan, add helper volumes with new objectives to control the dose in the same fashion as a human planner might do, lower dose to OARs and reduce hot and cold spots in the dose distribution. The use of this system has been reported in several studies [198], [203]–[214], creating APs of comparable or higher quality, reduced planning times and improved OAR sparing in comparison to MPs. In a study by Cilla *et al.* [203] Pinnacle AP produced high-quality plans for HNC and high-risk prostate and endometrial cancer while reducing planning time by 60–80 min, corresponding to 1/3 of the MP time. However, Zhang *et al.* [213] found that for nasopharyngeal carcinoma some automated plans did not fully meet dose objectives for the planning target volumes (PTV) and cautioned that AP cannot be fully trusted, and a manual selection between MP and AP should be performed for each patient.

*2) Multicriteria Optimization:* MCO is a method for AP that explores so-called Pareto-optimal (PO) plans. In a space spanned by assigning a dimension to each optimization parameter, the PO surface consists of plans that cannot be further optimized without degrading the performance of another objective. A schematic view of the MCO planning process is presented in Fig. 9. The PO surface is populated by several plans with various objective prioritizations, while the clinically ideal plan is always on the PO surface, not all plans at the PO surface will be clinically acceptable or desirable. Hussein *et al.*

described two approaches for AP with MCO, *a priori* and *a posteriori*. In the case of *a priori*, only a single PO plan is fully generated and presented to the treatment planner, while for *a posteriori* MCO, multiple plans are automatically generated, and the treatment planner can perform tradeoff navigation using, e.g., a navigation star [215] or sliders RayStation MCO (RaySearch Laboratories, Stockholm, Sweden), Varian Eclipse MCO) representing the objectives [216]. *A posteriori* MCO offers a convenient method to efficiently find an optimal treatment plan, while the treatment planner takes active choices in the process, and has been subject to several studies [217]–[229]. Kierkels *et al.* [224] found that MCO plans had similar performance to MP, but allowed inexperienced planners to make high-quality plans. Creating a large number of plans is computationally intensive, but can be done in the background while the planner attends to other tasks. Furthermore, estimations made by Bortfield and Craft suggest that as few as $N + 1$ plans are needed to populate the PO surface, where $N$ is the number of objectives [230]. The Erasmus-iCycle software [217] presents a solution for *a priori* MCO, where a disease-site-specific wish-list is defined with absolute and desired constraints, used for iterative optimization by the software. The software has been used in several studies, yielding clinically acceptable plans of similar or higher quality when compared to manually created plans [231]–[236].

*3) Knowledge-Based Planning:* KBP exploits the knowledge and expertise of treatment planners to aid the planning process for new patients, through a library of high-quality clinical treatment plans. When a new patient is presented to the system, it will be characterized based on anatomical and geometrical features. There are two branches of KBP. In *atlas-based* systems, the closest matching patient in the library is chosen, and the plan setup belonging to that patient will be duplicated to the new patient and recalculated on its planning-CT. In *model-based* systems, a predictive model is trained from the plan library, to predict parameters used for creating automated plans for new patients. The workflow for these approaches is illustrated in Fig. 9.

*a) Atlas-based KBP:* Atlas-based KBP [237]–[246] searches a library (i.e., atlas) of clinical plans to find the plan most similar in geometry to the new patient. The library plan setup is copied to the new patient, and the dose is calculated on the CT of the new patient. A general approach is to copy the plan setup and position the isocenter centered in the target volume of the new patient. The original plan can be adapted further, as demonstrated for HNC patients by Schmidt *et al.* where the atlas plan was adapted to the new patient by deforming the atlas plan beam fluences to suit the target volume in the new patient and warping the atlas primary/boost dose distribution to the new anatomy. The warped dose distribution was then used to generate dose–volume constraints as optimization constraints. In this study, it was found that AP had similar or better quality compared to MP for all objectives. The extra steps to deform and warp the plan led to improved performance compared to simply using the atlas plan directly [244]. Atlas-based KBP can either be used to provide a starting point for further optimization, or fully automate planning. Schreibmann *et al.* [245] demonstrated the use

of atlas-based KBP for whole-brain radiotherapy, generating high-quality treatment plans in 3–4 min, with reduced doses to OARs, thus reducing the clinical workload.

*b) Model-based KBP:* Model-based KBP [194], [202], [210], [214], [226], [229], [237], [238], [247]–[282] trains a model on the plan library to predict parameters for a new patient introduced to the system. Such parameters can be, e.g., beam settings, DVHs for target volumes and OARs, or full 3-D dose distributions. DVH prediction has been widely studied in recent years for most disease sites, e.g., head and neck [214], [226], [258], [272], [274], [281], prostate [210], [238], [251], [255], [264], [266]–[268], [270], [276], [281], upper GI [259], [265], [273], [280], lower GI [229], [254], [269], [277], [278], and breast [259], [262], [275]. A commercial software for DVH prediction is Varian RapidPlan. RapidPlan examines geometric and dosimetric properties of structures in each library plan and uses these to calculate a set of features. The calculated data of each plan is included in model training, where principal component analysis (PCA) is used to identify the 2–3 most important features, which are used as input for a regression model [283]. The final model is then used for DVH prediction. DVHs can be converted to objectives for the inverse planning optimizer by sampling the predicted DVH curve and creating corresponding dose–volume objectives. The majority of studies demonstrate that the dose distribution to target volumes of the APs are equally good or better than that of MPs, with no human planner interaction, while the time it takes to generate a plan is drastically reduced. Several studies also demonstrate reduced dose to OARs for APs compared to MPs, possibly due to manual planners not having sufficient time to make further improvements once a clinically acceptable plan is found. One major limitation of DVH prediction models is that they only consider dose–volume relations for delineated structures, and not spatial distributions, thus the planner must be aware of issues such as where excess dose in healthy tissue is placed.

KBP models can be trained by relatively few patients, as demonstrated by Boutillier *et al.* [284], who successfully trained a model for rectum DVH estimation using 20 library plans. In a recent review of KBP methods by Ge and Wu, it is suggested that more complex plans will require larger plan libraries. Development of larger training databases, e.g., by multi-institution collaborations, is recommended [261]. Another option to increase plan library size is to include plans from other techniques, e.g., 3D-CRT and IMRT plans for training a VMAT model [285], or plans from a different TPS [273]. However, as demonstrated by Ueda *et al.* [273], models may perform differently when used under other conditions than those of library plans, thus proper QA during commissioning is important.

### B. Modern AI in Radiation Therapy Treatment Planning

Modern AI, in particular the recent advancement in DL techniques, has achieved great success in a wide range of different disciplines, including medicine and healthcare. In the regime of RT treatment planning, a number of DL-driven AP methods have been developed recently in the literature to
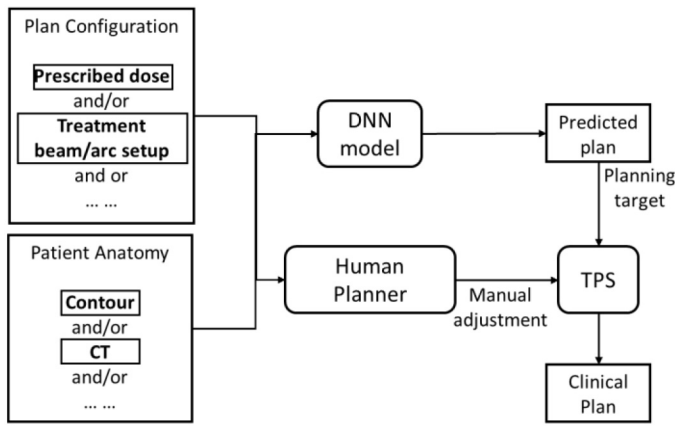
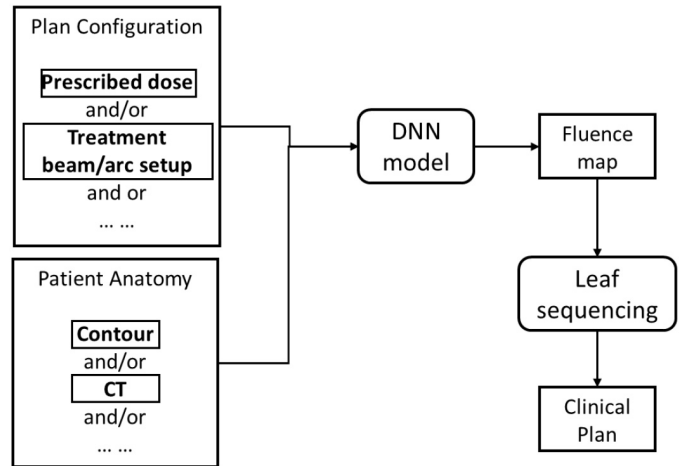Fig. 10. General workflow of DL dose prediction-based AP.



Fig. 11. General workflow of DL fluence map prediction-based AP.

address the remaining challenges in classical AP approaches. These novel methods can be roughly categorized into three groups: 1) DL-based dose prediction; 2) DL-based fluence map/aperture prediction; and 3) DL-based intelligent treatment planner.

*1) DL-Based Dose Prediction:* The DL dose prediction AP workflow is depicted in Fig. 10. A key step in this type of approaches is to build a DL-based dose prediction agent. Constructed by a carefully designed DNN, the dose prediction agent directly derives the dose distribution of a clinically acceptable plan based on the patient anatomical information (e.g., organ contours and/or CT image) and planning configurations (e.g., prescription dose and treatment beam/arc setups). Pioneer study [286] was first performed to investigate the feasibility of DL-based dose prediction using the treatment planning task of prostate cancer IMRT as a testbed. As a proof-of-principle study, a simplified 2-D dose prediction was considered. The prediction accuracy was validated on a set of dosimetric quantities of clinical interest. With feasibility well justified, DL-based dose prediction was then successfully extended to other modalities (e.g., VMAT and helical tomotherapy) [287]–[291], and more complicated treatment sites (e.g., HNC and lung cancer) [287]–[290], [292], as well as to the direct 3-D dose distribution prediction [291]–[293], which all are of more clinical and practical values. Later on, Pareto-optimal dose prediction models [228], [293], [294], which can generate multiple plans reflecting different tradeoffs among critical organs, were also developed to better address the diverse clinical needs. The clinical application of DL-based prediction methods is analog to the classical KBP methods while the major difference is that DL-based dose prediction derives dose distributions directly in contrast to the DVHs generated in KBP. Despite the great success achieved, DL-based dose prediction also suffers from the similar hurdle as KBP, i.e., the predicted dose may be dosimetrically attractive, but not be physically achievable.

*2) DL-Based Fluence Map Prediction:* As shown in Fig. 11, the second type of approach predicts the fluence map of an optimal plan from patient anatomy using DNN [196], [295]–[297], bypassing the plan optimization process in inverse treatment planning. More specifically,

Lee *et al.* [295] developed a DNN model to derive the optimal fluence map of IMRT in a beam-by-beam-fashion using the beam's eye view of PTV/organ contours, and the predicted optimal dose distribution for each beam as input. Li *et al.* [296] proposed a Dense-Res hybrid network (DRHN) to take a series of projections characterizing patient anatomy and treatment geometry as input and output the fluence intensity maps for the nine-field beam prostate cancer IMRT. Wang *et al.* [297] proposed two-stage strategy with each stage accomplished by a dedicated DNN. In the first stage, the dose of a beam is predicted from the contours and CT image of a patient, and the fluence map is then obtained based on the predicted dose of the beam in the second stage. Note that these algorithms were all developed specifically for IMRT, which typically involves only a very limited number of treatment beams ($\leq 9$). Lin *et al.* [196] proposed a DL-based fluence map prediction approach that can handle the more general VMAT planning in which fluence map of 64 treatment beams at different angles needs to be determined. By assuming the dose of an optimal plan already known, the projected dose was employed as an input of the established DNN model to predict the fluence map. All these studies have shown DL as a promising tool for fluence map prediction. However, the physical achievability and deliverability of the predicted fluence map is again not guaranteed.

*3) DL-Based Virtual Treatment Planner:* Motivated by the tremendous success of RL in achieving human-level intelligence in the decision-making process, the third group of methods focuses on developing RL-based virtual treatment planners (VPNs) to automatically adjust treatment planning setups or machine parameters [298]–[302] for high-quality treatment plans (Fig. 12). The idea of training a VPN is similar to the trial-and-error learning process of human beings. It allows the VPN to explore different ways of adjusting treatment planning setups/machine parameters and observe the impacts of these adjustments to the plan quality. VPN will gradually learn the consequences of applying different adjustments to a given plan from a wide spectrum of scenarios encountered in the training process, such that it can pick the adjustment leading to optimal plan quality.
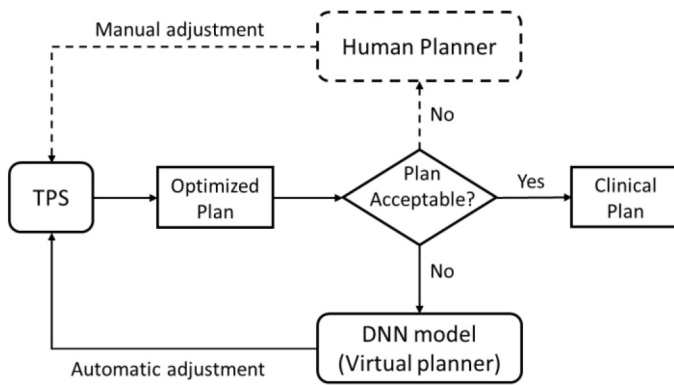
Fig. 12. General workflow of VPN-based AP, analog to conventional human planning process (dashed route).

The feasibility of RL-based VPN has been demonstrated in both a proof-of-principle context of high-dose-rate brachytherapy [300] and more complicated cases of external beam radiotherapy [298], [299], [301], [302], showing that intelligent behaviors of operating a TPS can be spontaneously generated via an end-to-end RL. Unlike the dose prediction and fluence map prediction methods, RL-based approaches naturally guarantee the achievability and deliverability since the final plans are generated by a TPS. However, the low training efficiency and poor scalability of VPN have limited its applications to only simple treatment planning problems. Although recent studies have shown that the training efficiency and model scalability can be substantially improved by incorporating human knowledge [299], and DNN of hierarchical architecture [298], respectively, the feasibility of VPN on a commercial TPS for complex clinical treatment planning tasks still needs to be further investigated. Moreover, the VPNs in existing studies are trained under the guidance of simple plan evaluation quantities, such as ProKnow score (ProKnow Systems, Sanford, FL, USA), which does not necessarily reflect the real planning objectives in the clinic. Better metrics of more clinical relevance, e.g., physicians' preference on plan quality, need to be quantified, and incorporated to guide the training process of a VPN.

### C. Challenges of Auto-Planning

Although classical/DL-based AP methods have achieved great success in many aspects, a number of technical or practical challenges remain unsolved. The first major challenge is the data size problem. Training effective AP models often require a large cohort of data of sufficient diversity to cover the variations among different patients. However, such datasets are nontrivial to collect, and their accessibility is very limited due to many concerns, such as privacy issues. This is indeed more of a concern for the DL-based algorithm since it typically requires huge amount of data to optimize DNNs with respect to the large number of learnable parameters. Lack of data often leads to severe overfitting, which may substantially deteriorate the performance of models on new patient cases never observed in model training. Second, despite their encouraging performance, most of the AP models, especially the DL-based ones, are difficult to interpret. To confidently deploy a model to automate the treatment planning process of patients, it is essential to have a good understanding about the reason behind the plan generation process. Such interpretability of a model is necessary to ensure its generality and robustness on different patients. Lack of interpretability may lead to unexpected model failure in the clinical deployment stage, posing serious risk to patients. In addition, there is also arising worries de-staffing/deskilling of human planners. However, in a study by Seo *et al.*, it was found that experienced planners outperform AP systems in complex cases [23]. In this regard, AP will be helpful to reduce the workload from simple cases and enable human planners to spend more time on complex cases, thereby further improving plan quality. To implement AP in most clinics, commonly accepted guidelines on implementation and quality assurance of AI models will play an important role, as discussed thoroughly in a recent review article by Vandewinckele *et al.* [30].

### D. Outlooks/Future Directions

AP has a great potential to automate and accelerate the tedious and time-consuming treatment planning process of RT. It is expected to improve the quality of treatment plans to a better level consistently, allowing to deliver the best of care to each patient. AP would also release human planners from simple cases, whose efforts can be focused on more challenging cases. The urgent need AP is further amplified in the regime of adaptive RT where planning efficiency critically affects the success of adaptation. Furthermore, moving toward the era of personalized care, AP permits the quick generation of treatment plans of different treatment techniques, e.g., IMRT versus VMAT, standard fractionation versus SBRT, photon versus proton treatment, etc., from which the best treatment plan can be chosen for optimal treatment outcome.

The remaining challenges of AP approaches, such as data size and interpretability, need to be addressed for the development of effective and reliable models. Collecting a large interinstitutional dataset is attractive, but less practical due to privacy and regulation concerns. One potential solution might be federated learning [303]–[305], which allows each institution to keep their own data while having the model be trained on all data. In addition, interpretable ML/DL [306], [307] has become a central topic recently due to the increasing need of model expainability and reliability in many applications including RT. More efforts are definitely needed along these directions to fulfill the urgent, but unmet need of AP.

### REFERENCES

[1] W. Zhu *et al.*, "AnatomyNet: Deep learning for fast and fully automated whole-volume segmentation of head and neck anatomy," *Med. Phys.*, vol. 46, no. 2, pp. 576–589, Feb. 2019.

[2] Y. Lei *et al.*, "Male pelvic multi-organ segmentation aided by CBCT-based synthetic MRI," *Phys. Med. Biol.*, vol. 65, no. 3, Feb. 2020, Art. no. 035013.

[3] X. Dong *et al.*, "Automatic multiorgan segmentation in thorax CT images using U-net-GAN," *Med. Phys.*, vol. 46, no. 5, pp. 2157–2168, 2019.

[4] B. van der Heyden *et al.*, "Dual-energy CT for automatic organs-at-risk segmentation in brain-tumor patients using a multi-atlas and deep-learning approach," *Sci. Rep.*, vol. 9, no. 1, p. 4126, 2019.

[5] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. MICCAI*, 2015, pp. 234–241.

[6] Y. Chen *et al.*, "Fully automated multiorgan segmentation in abdominal magnetic resonance imaging with deep neural networks," *Med. Phys.*, vol. 47, no. 10, pp. 4971–4982, 2020.

[7] Y. Fu *et al.*, "A novel MRI segmentation method using CNN-based correction network for MRI-guided adaptive radiotherapy," *Med. Phys.*, vol. 45, no. 11, pp. 5129–5137, 2018.

[8] M.-M. Rohé, M. Datar, T. Heimann, M. Sermesant, and X. Pennec, "SVF-Net: Learning deformable image registration using shape matching," in *Medical Image Computing and Computer Assisted Intervention (MICCAI)*. Cham, Switzerland: Springer, 2007, pp. 266–274.

[9] S. Ghosal and N. Rayl, "Deep deformable registration: Enhancing accuracy by fully convolutional neural net," *Pattern Recognit. Lett.*, vol. 94, pp. 81–86, Jul. 2017.

[10] E. Chee and Z. Wu, "AIRNet: Self-supervised affine registration for 3D medical images using neural networks," 2018. [Online]. Available: arXiv:1810.02583.

[11] J. Fan, X. Cao, Z. Xue, P. T. Yap, and D. Shen, "Adversarial similarity network for evaluating image alignment in deep learning based registration," *Med. Image Comput. Comput. Assist. Interv.*, vol. 11070, pp. 739–746, Sep. 2018.

[12] J. F. Fan, X. H. Cao, E. A. Yap, and D. G. Shen, "BIRNet: Brain image registration using dual-supervised fully convolutional networks," *Med. Image Anal.*, vol. 54, pp. 193–206, May 2019.

[13] X. Yang, R. Kwitt, M. Styner, and M. Niethammer, "Quicksilver: Fast predictive image registration—A deep learning approach," *NeuroImage*, vol. 158, pp. 378–396, Sep. 2017.

[14] G. Nguyen *et al.*, "Machine learning and deep learning frameworks and libraries for large-scale data mining: A survey," *Artif. Intell. Rev.*, vol. 52, no. 1, pp. 77–124, 2019.

[15] X. Cao, J. Yang, L. Wang, Z. Xue, Q. Wang, and D. Shen, "Deep learning based inter-modality image registration supervised by intra-modality similarity," in *Machine Learning in Medical Imaging (MLMI)*, vol. 11046. Cham, Switzerland: Springer, 2018, pp. 55–63.

[16] X. H. Cao, J. H. Yang, J. Zhang, Q. Wang, P.-T. Yap, and D. G. Shen, "Deformable image registration using a cue-aware deep regression network," *IEEE Trans. Biomed. Eng.*, vol. 65, no. 9, pp. 1900–1911, Sep. 2018.

[17] Y. Hu *et al.*, "Adversarial deformation regularization for training image registration neural networks," 2018. [Online]. Available: arXiv:1805.10665.

[18] Y. P. Hu *et al.*, "Weakly-supervised convolutional neural networks for multimodal image registration," *Med. Image Anal.*, vol. 49, pp. 1–13, Oct. 2018.

[19] T. Sentker, F. Madesta, and R. Werner, "GDL-FIRE$^{4D}$: Deep learning-based fast 4D CT image registration," in *Proc. MICCAI*, 2018, pp. 765–773.

[20] K. A. J. Eppenhof, M. W. Lafarge, P. Moeskops, M. Veta, and J. P. W. Pluim, "Deformable image registration using convolutional neural networks," in *Proc. Med. Imag. Image Process.*, 2018, Art. no. 105740S.

[21] K. A. J. Eppenhof, and J. P. W. Pluim, "Pulmonary CT registration through supervised learning with convolutional neural networks," *IEEE Trans. Med. Imag.*, vol. 38, no. 5, pp. 1097–1105, May 2019.

[22] V. Kearney, S. Haaf, A. Sudhyadhom, G. Valdes, and T. D. Solberg, "An unsupervised convolutional neural network-based algorithm for deformable image registration," *Phys. Med. Biol.*, vol. 63, no. 18, Sep. 2018, Art. no. 185017.

[23] H. Seo *et al.*, "Machine learning techniques for biomedical image segmentation: An overview of technical aspects and introduction to state-of-art applications," 2019. [Online]. Available: arXiv:1911.02521.

[24] A. Kori, and G. Krishnamurthi, "Zero shot learning for multi-modal real time image registration," 2019. [Online]. Available: arXiv:1908.06213.

[25] H. Sokooti, B. de Vos, F. Berendsen, B. P. F. Lelieveldt, I. Išgum, and M. Staring, "Nonrigid image registration using multi-scale 3D convolutional neural networks," in *Medical Image Computing and Computer Assisted Intervention (MICCAI)*. Cham, Switzerland: Springer, 2017, pp. 232–239.

[26] H. Sokooti *et al.*, "3D convolutional neural networks image registration based on efficient supervised learning from artificial deformations," 2019. [Online]. Available: arXiv:1908.10235.

[27] J. O. Onieva, B. Marti-Fuster, M. P. de la Puente, and R. San José Estépar, "Diffeomorphic lung registration using deep CNNs and reinforced learning," in *Image Analysis for Moving Organ, Breast, and Thoracic Images*. Cham, Switzerland: Springer, 2018, pp. 284–294.

[28] J. Krebs, T. Mansi, B. Mailhe, N. Ayache, and H. Delingette, "Unsupervised probabilistic deformation modeling for robust diffeomorphic registration," in *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, vol. 11045. Cham, Switzerland: Springer, 2018, pp. 101–109.

[29] Q. Zeng *et al.*, "Label-driven magnetic resonance imaging (MRI)-transrectal ultrasound (TRUS) registration using weakly supervised learning for MRI-guided prostate radiotherapy," *Phys. Med. Biol.*, vol. 65, no. 13, 2020, Art. no. 135002.

[30] L. Vandewinckele *et al.*, "Overview of artificial intelligence-based applications in radiotherapy: Recommendations for implementation and quality assurance," *Radiother. Oncol.*, vol. 153, pp. 55–66, Dec. 2020.

[31] S. Siddique and J. C. L. Chow, "Artificial intelligence in radiotherapy," *Rep. Pract. Oncol. Radiother.*, vol. 25, no. 4, pp. 656–666, 2020.

[32] E. Huynh *et al.*, "Artificial intelligence in radiation oncology," *Nat. Rev. Clin. Oncol.*, vol. 17, no. 12, pp. 771–781, 2020.

[33] S. Kuckertz, N. Papenberg, J. Honegger, T. Morgas, B. Haas, and S. Heldmann, "Learning deformable image registration with structure guidance constraints for adaptive radiotherapy," in *Biomedical Image Registration*. Cham, Switzerland: Springer, 2020, pp. 44–53.

[34] C. Stergios *et al.*, "Linear and deformable image registration with 3D convolutional neural networks," *Image Analysis for Moving Organ, Breast, and Thoracic Images*. Cham, Switzerland: Springer, 2018, pp. 13–22.

[35] R. R. Upendra, R. Simon, and C. A. Linte, "A supervised image registration approach for late gadolinium enhanced MRI and cine cardiac MRI using convolutional neural networks," *Med. Image Understand. Anal.*, vol. 1248, pp. 208–220, Jul. 2020.

[36] L. Sun and S. Zhang, "Deformable MRI-ultrasound registration using 3D convolutional neural network," in *Simulation, Image Processing, and Ultrasound Systems for Assisted Diagnosis and Navigation*. Cham, Switzerland: Springer, 2018, pp. 152–158.

[37] B. D. de Vos, F. F. Berendsen, M. A. Viergever, H. Sokooti, M. Staring, and I. Isgum, "A deep learning framework for unsupervised affine and deformable image registration," *Med. Image Anal.*, vol. 52, pp. 128–143, Feb. 2019.

[38] G. Balakrishnan, A. Zhao, M. R. Sabuncu, J. Guttag, and A. V. Dalca, "An unsupervised learning model for deformable medical image registration," *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2018, pp. 9252–9260.

[39] G. Balakrishnan, A. Zhao, M. R. Sabuncu, J. Guttag, and A. V. Dalca, "VoxelMorph: A learning framework for deformable medical image registration," *IEEE Trans. Med. Imag.*, vol. 38, no. 8, pp. 1788–1800, Aug. 2019.

[40] B. Kim, J. Kim, J.-G. Lee, D. H. Kim, S. H. Park, and J. C. Ye, "Unsupervised deformable image registration using cycle-consistent CNN," 2019. [Online]. Available: arXiv:1907.01319.

[41] D. Kuang, "On reducing negative Jacobian determinant of the deformation predicted by deep registration networks," 2019. [Online]. Available: arXiv:1907.00068.

[42] D. Kuang and T. Schmah, "FAIM—A ConvNet method for unsupervised 3D medical image registration," in *Machine Learning in Medical Imaging*. Cham, Switzerland: Springer, 2019, pp. 646–654.

[43] N. Tyagi *et al.*, "Clinical workflow for MR-only simulation and planning in prostate," *Radiat. Oncol.*, vol. 12, no. 1, p. 119, 2017.

[44] T. F. Lau, J. Luo, S. Zhao, E. I.-C. Chang, and Y. Xu, "Unsupervised 3D end-to-end medical image registration with volume tweening network," *IEEE J. Biomed. Health Inform.*, vol. 24, no. 5, pp. 1394–1404, May 2020.

[45] M. Unterrainer, "Recent advances of PET imaging in clinical radiation oncology," *Radiat. Oncol.*, vol. 15, p. 88, Apr. 2020.

[46] Y. Lei *et al.*, "4D-CT deformable image registration using an unsupervised deep convolutional neural network," *Artificial Intelligence in Radiation Therapy*. Cham, Switzerland: Springer, 2019, pp. 26–33.

[47] Y. Lei *et al.*, "4D-CT deformable image registration using multiscale unsupervised deep learning," *Phys. Med. Biol.*, vol. 65, no. 8, 2020, Art. no. 085003.

[48] G. P. Liney, B. Whelan, B. Oborn, M. Barton, and P. Keall, "MRI-linear accelerator radiotherapy systems," *Clin. Oncol.*, vol. 30, no. 11, pp. 686–691, 2018.

[49] Y. Fu *et al.*, "LungRegNet: An unsupervised deformable image registration method for 4D-CT lung," *Med. Phys.*, vol. 47, no. 4, pp. 1763–1774, 2020.

[50] I. J. Das, K. P. McGee, N. Tyagi, and H. Wang, "Role and future of MRi in radiation oncology," *Brit. J. Radiol.*, vol. 92, no. 1094, 2019, Art. no. 20180505.

[51] Y. Fu *et al.*, "Biomechanically constrained non-rigid MR-TRUS prostate registration using deep learning based 3D point cloud matching," *Med. Image Anal.*, vol. 67, Jan. 2021, Art. no. 101845.

[52] Y. Fu *et al.*, "Deformable MR-CBCT prostate registration using biomechanically constrained deep learning networks," *Med. Phys.*, vol. 48, no. 1, pp. 253–263, 2020.

[53] R. D. Stewart and X. A. Li, "BGRT: Biologically guided radiation therapy—The future is fast approaching!" *Med. Phys.*, vol. 34, no. 10, pp. 3739–3751, 2007.

[54] T. Fechter and D. Baltas, "One-shot learning for deformable medical image registration and periodic motion tracking," *IEEE Trans. Med. Imag.*, vol. 39, no. 7, pp. 2506–2517, Jul. 2020.

[55] H.-M. Zhang and B. Dong, "A review on deep learning in medical image reconstruction," *J. Oper. Res. Soc. China*, vol. 8, pp. 311–340, Jan. 2020.

[56] M. Beister, D. Kolditz, and W. A. Kalender, "Iterative reconstruction methods in X-ray CT," *Physica Medica*, vol. 28, no. 2, pp. 94–108, 2012.

[57] L. L. Geyer *et al.*, "State of the art: Iterative CT reconstruction techniques," *Radiology*, vol. 276, no. 2, pp. 339–357, 2015.

[58] H. Zhang, J. Wang, D. Zeng, X. Tao, and J. Ma, "Regularization strategies in statistical image reconstruction of low-dose X-ray CT: A review," *Med. Phys.*, vol. 45, no. 10, pp. e886–e907, 2018.

[59] J. A. Fessler, "Model-based image reconstruction for MRI," *IEEE Signal Process. Mag.*, vol. 27, no. 4, pp. 81–89, Jul. 2010.

[60] S. Geethanath *et al.*, "Compressed sensing MRI: A review," *Crit. Rev. Biomed. Eng.*, vol. 41, no. 3, 2013, pp. 183–204.

[61] J. Qi and R. M. Leahy, "Iterative reconstruction techniques in emission computed tomography," *Phys. Med. Biol.*, vol. 51, no. 15, pp. R541–R578, 2006.

[62] G. Wang, "A perspective on deep imaging," *IEEE Access*, vol. 4, pp. 8914–8924, 2016.

[63] G. Wang, J. C. Ye, K. Mueller, and J. A. Fessler, "Image reconstruction is a new frontier of machine learning," *IEEE Trans. Med. Imag.*, vol. 37, no. 6, pp. 1289–1296, Jun. 2018.

[64] A. Selvikvåg and A. Lundervold, "An overview of deep learning in medical imaging focusing on MRI," *Zeitschrift Medizinische Physik*, vol. 29, no. 2, pp. 102–127, 2019.

[65] A. J. Reader, G. Corda, A. Mehranian, C. da Costa-Luis, S. Ellis, and J. A. Schnabel, "Deep learning for PET image reconstruction," *IEEE Trans. Radiat. Plasma Med. Sci.*, vol. 5, no. 1, pp. 1–25, Jan. 2021.

[66] H. Ben Yedder, B. Cardoen, and G. Hamarneh, "Deep learning for biomedical image reconstruction: A survey," *Artif. Intell. Rev.*, vol. 54, pp. 215–251, Jan. 2021, doi: 10.1007/s10462-020-09861-2.

[67] G. Wang, J. C. Ye, and B. De Man, "Deep learning for tomographic image reconstruction," *Nat. Mach. Intell.*, vol. 2, pp. 737–748, Dec. 2020.

[68] S. Ravishankar, J. C. Ye, and J. A. Fessler, "Image reconstruction: From sparsity to data-adaptive methods and machine learning," *Proc. IEEE*, vol. 108, no. 1, pp. 86–109, Jan. 2020.

[69] C.-C. Shieh *et al.*, "SPARE: Sparse-view reconstruction challenge for 4D cone-beam CT from a 1-min scan," *Med. Phys.*, vol. 46, no. 9, pp. 3799–3811, 2019.

[70] E. Kang, J. Min, and J. C. Ye, "A deep convolutional neural network using directional wavelets for low-dose X-ray CT reconstruction," *Med. Phys.*, vol. 44, no. 10, pp. e360–e375, 2017.

[71] H. Chen *et al.*, "Low-Dose CT with a residual encoder-decoder convolutional neural network," *IEEE Trans. Med. Imag.*, vol. 36, no. 12, pp. 2524–2535, Dec. 2017.

[72] Q. Yang *et al.*, "Low-dose CT image denoising using a generative adversarial network with wasserstein distance and perceptual loss," *IEEE Trans. Med. Imag.*, vol. 37, no. 6, pp. 1348–1357, Jun. 2018.

[73] Y. Wang *et al.*, "Iterative quality enhancement via residual-artifact learning networks for low-dose CT," *Phys. Med. Biol.*, vol. 63, no. 21, 2018, Art. no. 215004.

[74] Y. Han and J. C. Ye, "Framing U-net via deep convolutional framelets: Application to sparse-view CT," *IEEE Trans. Med. Imag.*, vol. 37, no. 6, pp. 1418–1429, Jun. 2018.

[75] Z. Zhang, X. Liang, X. Dong, Y. Xie, and G. Cao, "A sparse-view CT reconstruction method based on combination of densenet and deconvolution," *IEEE Trans. Med. Imag.*, vol. 37, no. 6, pp. 1407–1417, Jun. 2018.

[76] Z. Jiang, Y. Chen, Y. Zhang, Y. Ge, F. F. Yin, and L. Ren, "Augmentation of CBCT reconstructed from under-sampled projections using deep learning," *IEEE Trans. Med. Imag.*, vol. 38, no. 11, pp. 2705–2715, Nov. 2019.

[77] J. M. Wolterink, T. Leiner, M. A. Viergever, and I. Išgum, "Generative adversarial networks for noise reduction in low-dose CT," *IEEE Trans. Med. Imag.*, vol. 36, no. 12, pp. 2536–2545, Dec. 2017.

[78] Z. Li, S. Zhou, J. Huang, L. Yu, and M. Jin, "Investigation of low-dose CT image denoising using unpaired deep learning methods," *IEEE Trans. Radiat. Plasma Med. Sci.*, vol. 5, no. 2, pp. 224–234, Mar. 2021.

[79] H. Zhang, D. Zeng, H. Zhang, J. Wang, Z. Liang, and J. Ma, "Applications of nonlocal means algorithm in low-dose X-ray CT image processing and reconstruction: A review," *Med. Phys.*, vol. 44, no. 3, 2017, pp. 1168–1185.

[80] H. Lee, J. Lee, H. Kim, B. Cho, and S. Cho, "Deep-neural-network based sinogram synthesis for sparse-view CT image reconstruction," 2018. [Online]. Available: https://arxiv.org/ftp/arxiv/papers/1803/1803.00694.pdf

[81] M. U. Ghani and W. C. Karl, "Deep learning-based sinogram completion for low-dose CT," in *Proc. IEEE 13th Image Video Multidimensional Signal Process. Workshop*, 2018, pp. 1–5.

[82] K. Liang, Y. Xing, H. Yang, and K. Kang, "Improve angular resolution for sparse-view CT with residual convolutional neural network," in *Proc. Med. Imag. Phys. Med. Imag.*, 2018, p. 55.

[83] J. Beaudry, P. Esquinas, and C.-C. Shieh, "Learning from our neighbours: A novel approach on sinogram completion using bin-sharing and deep learning to reconstruct high quality 4DCBCT," in *Proc. Med. Imag. Phys. Med. Imag.*, 2019, p. 153.

[84] T. Würfl, F. Ghesu, V. Christlein, and A. Maier, *Deep Learning Computed Tomography*. Cham, Switzerland: Springer, 2016.

[85] J. He, Y. Wang, and J. Ma, "Radon inversion via deep learning," *IEEE Trans. Med. Imag.*, vol. 39, no. 6, pp. 2076–2087, Jun. 2020.

[86] Y. Li, K. Li, C. Zhang, J. Montoya, and G.-H. Chen, "Learning to reconstruct computed tomography images directly from sinogram data under a variety of data acquisition conditions," *IEEE Trans. Med. Imag.*, vol. 38, no. 10, pp. 2469–2481, Oct. 2019.

[87] D. Wu, K. Kim, G. El Fakhri, and Q. Li, "Iterative low-dose CT reconstruction with priors trained by artificial neural network," *IEEE Trans. Med. Imag.*, vol. 36, no. 12, pp. 2479–2486, Dec. 2017.

[88] B. Chen, K. Xiang, Z. Gong, J. Wang, and S. Tan, "Statistical iterative CBCT reconstruction based on neural network," *IEEE Trans. Med. Imag.*, vol. 37, no. 6, pp. 1511–1521, Jun. 2018.

[89] Y. Gao *et al.*, "Constructing a tissue-specific texture prior by machine learning from previous full-dose scan for Bayesian reconstruction of current ultralow-dose CT images," *J. Med. Imag.*, vol. 7, no. 3, 2020, Art. no. 032502.

[90] C. Shen, Y. Gonzalez, L. Chen, S. B. Jiang, and X. Jia, "Intelligent parameter tuning in optimization-based iterative CT reconstruction via deep reinforcement learning," *IEEE Trans. Med. Imag.*, vol. 37, no. 6, pp. 1430–1439, Jun. 2018.

[91] H. Chen *et al.*, "LEARN: Learned experts' assessment-based reconstruction network for sparse-data CT," *IEEE Trans. Med. Imag.*, vol. 37, no. 6, pp. 1333–1347, Jun. 2018.

[92] J. He *et al.*, "Optimizing a parameterized plug-and-play ADMM for iterative low-dose CT reconstruction," *IEEE Trans. Med. Imag.*, vol. 38, no. 2, pp. 371–382, Feb. 2019.

[93] B. Kelly, T. P. Matthews, and M. A. Anastasio, "Deep learning-guided image reconstruction from incomplete data," 2017. [Online]. Available: arXiv:1709.00584.

[94] H. Gupta, K. H. Jin, H. Q. Nguyen, M. T. McCann, and M. Unser, "CNN-based projected gradient descent for consistent CT image reconstruction," *IEEE Trans. Med. Imag.*, vol. 37, no. 6, pp. 1440–1453, Jun. 2018.

[95] J. Adler and O. Öktem, "Learned primal–dual reconstruction," *IEEE Trans. Med. Imag.*, vol. 37, no. 6, pp. 1322–1332, Jun. 2018.

[96] B. Zhu, J. Z. Liu, S. F. Cauley, B. R. Rosen, and M. S. Rosen, "Image reconstruction by domain-transform manifold learning," *Nature*, vol. 555, pp. 487–492, Mar. 2018.

[97] L. Fu and B. De Man, "A hierarchical approach to deep learning and its application to tomographic reconstruction," in *Proc. 15th Int. Meeting Fully Three-Dimensional Image Reconstruct. Radiol. Nucl. Med.*, 2019, p. 41.

[98] L. Shen, W. Zhao, and L. Xing, "Patient-specific reconstruction of volumetric computed tomography images from a single projection view via deep learning," *Nat. Biomed. Eng.*, vol. 3, pp. 880–888, Oct. 2019.

[99] Y. Lei *et al.*, "Deep learning-based real-time volumetric imaging for lung stereotactic body radiation therapy: A proof of concept study," *Phys. Med. Biol.*, vol. 65, no. 23, Dec. 2020, Art. no. 235003.

[100] J. Solomon, P. Lyu, D. Marin, and E. Samei, "Noise and spatial resolution properties of a commercially available deep learning-based CT reconstruction algorithm," *Med. Phys.*, vol. 47, no. 9, pp. 3961–3971, 2020.

[101] X. Yang and B. Fei, "3D prostate segmentation of ultrasound images combining longitudinal image registration and machine learning," in *Proc. SPIE Int. Soc. Opt. Eng.*, vol. 8316, 2012, Art. no. 83162O.

[102] Y. Fu, S. Liu, H. Li, and D. Yang, "Automatic and hierarchical segmentation of the human skeleton in CT images," *Phys. Med. Biol.*, vol. 62, no. 7, pp. 2812–2833, 2017.

[103] X. Yang, P. J. Rossi, A. B. Jani, H. Mao, W. J. Curran, and T. Liu, "3D transrectal ultrasound (TRUS) prostate segmentation based on optimal feature learning framework," in *Proc. SPIE Int. Soc. Opt. Eng.*, vol. 9784, 2016, Art. no. 97842F.

[104] N. Samavati, M. Velec, and K. K. Brock, "Effect of deformable registration uncertainty on lung SBRT dose accumulation," *Med. Phys.*, vol. 43, no. 1, p. 233, 2016.

[105] I. J. Chetty and M. Rosu-Bubulac, "Deformable registration for dose accumulation," *Seminars Radiat. Oncol.*, vol. 29, no. 3, pp. 198–208, 2019.

[106] D. Yang *et al.*, "Technical note: DIRART—A software suite for deformable image registration and adaptive radiotherapy research," *Med. Phys.*, vol. 38, no. 1, pp. 67–77, 2011.

[107] T. Vercauteren, X. Pennec, A. Perchant, and N. Ayache, "Diffeomorphic demons: Efficient non-parametric image registration," *NeuroImage*, vol. 45, no. 1, pp. S61–S72, 2009.

[108] B. B. Avants, N. J. Tustison, G. Song, P. A. Cook, A. Klein, and J. C. Gee, "A reproducible evaluation of ANTs similarity metric performance in brain image registration," *NeuroImage*, vol. 54, no. 3, pp. 2033–2044, 2011.

[109] D. Shen, "Image registration by local histogram matching," *Pattern Recognit.*, vol. 40, no. 4, pp. 1161–1172, 2007.

[110] S. Klein, M. Staring, K. Murphy, M. A. Viergever, and J. P. W. Pluim, "Elastix: A toolbox for intensity-based medical image registration," *IEEE Trans. Med. Imag.*, vol. 29, no. 1, pp. 196–205, Jan. 2010.

[111] Y. Fu, Y. Lei, T. Wang, J. C. Walter, T. Liu, and X. Yang, "Deep learning in medical image registration: A review," *Phys. Med. Biol.*, vol. 65, no. 20, 2020, Art. no. 20TR01.

[112] X. Cheng, L. Zhang, and Y. Zheng, "Deep similarity learning for multimodal medical images," *Comput. Methods Biomech. Biomed. Eng. Imag. Visualization*, vol. 6, no. 3, pp. 248–252, 2018.

[113] M. Simonovsky, B. Gutiérrez-Becker, D. Mateus, N. Navab, and N. Komodakis, "A deep metric for multimodal registration," in *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*. Cham, Switzerland: Springer, 2016, pp. 10–18.

[114] S. S. M. Salehi, S. Khan, D. Erdogmus, and A. Gholipour, "Real-time deep registration with geodesic loss," 2018. [Online]. Available: arXiv:1803.05982.

[115] R. Castillo *et al.*, "A framework for evaluation of deformable image registration spatial accuracy using large landmark point sets," *Phys. Med. Biol.*, vol. 54, no. 7, pp. 1849–1870, 2009.

[116] F. F. Berendsen, A. N. T. J. Kotte, M. A. Viergever, and J. P. W. Pluim, "Registration of organs with sliding interfaces and changing topologies," in *Proc. SPIE Med. Imag.*, 2014, Art. no. 90340E.

[117] M. Modat *et al.*, "Fast free-form deformation using graphics processing units," *Comput. Methods Programs Biomed.*, vol. 98, no. 3, pp. 278–284, 2010.

[118] J. A. Shackleford, N. Kandasamy, and G. C. Sharp, "On developing B-spline registration algorithms for multi-core processors," *Phys. Med. Biol.*, vol. 55, no. 21, pp. 6329–6351, 2010.

[119] R. Werner, A. Schmidt-Richberg, H. Handels, and J. Ehrhardt, "Estimation of lung motion fields in 4D CT data by variational nonlinear intensity-based registration: A comparison and evaluation study," *Phys. Med. Biol.*, vol. 59, no. 15, pp. 4247–4260, 2014.

[120] H. Uzunova, M. Wilms, H. Handels, and J. Ehrhardt, "Training CNNs for image registration from few samples with model-based data augmentation," in *Medical Image Computing and Computer Assisted Intervention*. Cham, Switzerland: Springer, 2017, pp. 223–231.

[121] A. Hering, S. Kuckertz, S. Heldmann, and M. Heinrich, "Enhancing label-driven deep deformable image registration with local distance metrics for state-of-the-art cardiac motion tracking," 2018. [Online]. Available: arXiv:1812.01859.

[122] M. Jaderberg, K. Simonyan, A. Zisserman, and K. Kavukcuoglu, "Spatial transformer networks," 2015. [Online]. Available: arXiv:1506.02025.

[123] J. Zhang, "Inverse-consistent deep networks for unsupervised deformable image registration," 2018. [Online]. Available: arXiv:1809.03443.

[124] Z. Jiang, F. F. Yin, Y. Ge, and L. Ren, "A multi-scale framework with unsupervised joint training of convolutional neural networks for pulmonary deformable image registration," *Phys. Med. Biol.*, vol. 65, no. 1, 2020, Art. no. 015011.

[125] S. Ghosh, N. Das, I. Das, and U. Maulik, "Understanding deep learning techniques for image segmentation," *ACM Comput. Surveys*, vol. 52, no. 4, pp. 1–35, 2019.

[126] H. Zaidi and I. El Naqa, "PET-guided delineation of radiation therapy treatment volumes: A survey of image segmentation techniques," *Eur. J. Nucl. Med. Mol. Imag.*, vol. 37, no. 11, pp. 2165–2187, 2010.

[127] X. Zhuang and J. Shen, "Multi-scale patch and multi-modality atlases for whole heart segmentation of MRI," *Med. Image Anal.*, vol. 31, pp. 77–87, Jul. 2016.

[128] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[129] K. Men, J. Dai, and Y. Li, "Automatic segmentation of the clinical target volume and organs at risk in the planning CT for rectal cancer using deep dilated convolutional neural networks," *Med. Phys.*, vol. 44, no. 12, pp. 6377–6389, 2017.

[130] X. Yang *et al.*, "Automated segmentation of the parotid gland based on atlas registration and machine learning: A longitudinal MRI study in head-and-neck radiation therapy," *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 90, no. 5, pp. 1225–1233, 2014.

[131] B. Ibragimov and L. Xing, "Segmentation of organs-at-risks in head and neck CT images using convolutional neural networks," *Med. Phys.*, vol. 44, no. 2, pp. 547–557, 2017.

[132] Y. Fu, Y. Lei, T. Wang, W. J. Curran, T. Liu, and X. Yang, "A review of deep learning based methods for medical image multi-organ segmentation," *Physica Medica*, vol. 85, pp. 107–122, May 2021.

[133] G. Litjens *et al.*, "A survey on deep learning in medical image analysis," *Med. Image Anal.*, vol. 42, pp. 60–88, Dec. 2017.

[134] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proc. CVPR*, 2015, pp. 3431–3440.

[135] E. D. Morris, A. I. Ghanem, M. Dong, M. V. Pantelic, E. M. Walker, and C. K. Glide-Hurst, "Cardiac substructure segmentation with deep learning for improved cardiac sparing," *Med. Phys.*, vol. 47, no. 2, pp. 576–586, 2019.

[136] N. Ibtehaz and M. S. Rahman, "MultiResUNet: Rethinking the U-Net architecture for multimodal biomedical image segmentation," *Neural Netw.*, vol. 121, pp. 74–87, Jan. 2020.

[137] A. Mortazi, J. Burt, and U. Bagci, "Multi-planar deep segmentation networks for cardiac substructures from MRI and CT," in *Proc. STACOM MICCAI*, 2017, pp. 199–206.

[138] C. Payer, D. Štern, H. Bischof, and M. Urschler, "Multi-label whole heart segmentation using CNNs and anatomical label configurations," in *Proc. STACOM MICCAI*, 2017, pp. 190–198.

[139] Z. Zhang, L. Yang, and Y. Zheng, "Translating and segmenting multimodal medical volumes with cycle-and shapeconsistency generative adversarial network," in *Proc. CVPR*, 2018, pp. 9242–9251.

[140] J. Kirby *et al.*, "Introduction to special issue on datasets hosted in The Cancer Imaging Archive (TCIA)," *Med. Phys.*, vol. 47, no. 12, pp. 6026–6028, 2020.

[141] M. Avendi, A. Kheradvar, and H. Jafarkhani, "A combined deep-learning and deformable-model approach to fully automatic segmentation of the left ventricle in cardiac MRI," *Med. Image Anal.*, vol. 30, pp. 108–119, May 2016.

[142] E. Brion, J. Léger, U. Javaid, J. Lee, C. De Vleeschouwer, and B. Macq, "Using planning CTs to enhance CNN-based bladder segmentation on cone beam CT," in *Proc. Med. Imag. Image-Guided Procedures* 2019, Art. no. 109511M.

[143] B. Sahiner *et al.*, "Deep learning in medical imaging and radiation therapy," *Med. Phys.*, vol. 46, no. 1, pp. e1–e36, 2019.

[144] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal loss for dense object detection," in *Proc. ICCV*, 2017, pp. 2980–2988.

[145] M. Chen, L. Fang, and H. Liu, "FR-NET: Focal loss constrained deep residual networks for segmentation of cardiac MRI," in *Proc. ISBI*, 2019, pp. 764–767.

[146] P. Isola, J.-Y. Zhu, T. Zhou, and A. E. Alexei, "Image-to-image translation with conditional adversarial networks," 2016. [Online]. Available: http://arxiv.org/abs/1611.07004.

[147] Y. Peng et al., "Magnetic resonance-based synthetic computed tomography images generated using generative adversarial networks for nasopharyngeal carcinoma radiotherapy treatment planning," *Radiother. Oncol.*, vol. 150, pp. 217–224, Sep. 2020.

[148] J. Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Proc. ICCV*, vol. 2017, 2017, pp. 2242–2251.

[149] J. M. Wolterink, M. D. Anna, H. F. S. Mark, R. S. Peter, A. T. van den Berg Cornelis, and I. Igum, "Deep MR to CT synthesis using unpaired data," in *Simulation and Synthesis in Medical Imaging* (Lecture Notes in Computer Science (Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics 10557)). Cham, Switzerland: Springer, 2017, pp. 14–23.

[150] Y. Lei et al., "MRI-only based synthetic CT generation using dense cycle consistent generative adversarial networks," *Med. Phys.*, vol. 46, no. 8, pp. 3565–3581, 2019.

[151] R. Liu et al., "Synthetic dual-energy CT for MRI-only based proton therapy treatment planning using label-GAN," *Phys. Med. Biol.*, vol. 66, no. 6, 2021, Art. no. 065014.

[152] L. Chen, Y. Wu, A. M. DSouza, A. Z. Abidin, A. Wismuller, and C. Xu, "MRI tumor segmentation with densely connected 3D CNN," 2018. [Online]. Available: arXiv:1802.02427.

[153] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2015. [Online]. Available: arXiv:1409.1556.

[154] L. Bi, J. Kim, D. D. F. Feng, and M. J. Fulham, *Synthesis of Positron Emission Tomography (PET) Images via Multi-Channel Generative Adversarial Networks (GANs)* (Lecture Notes in Computer Science 10555). Cham, Switzerland: Springer, Sep. 2017, pp. 105–115.

[155] C. Veiga et al., "First clinical investigation of cone beam computed tomography and deformable registration for adaptive proton therapy for lung cancer," *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 95, no. 1, pp. 549–559, 2016.

[156] C. Kurz et al., "Comparing cone-beam CT intensity correction methods for dose recalculation in adaptive intensity-modulated photon and proton therapy for head and neck cancer," *Acta Oncologica*, vol. 54, no. 9, pp. 1651–1657, 2015.

[157] G. J. Bootsma, F. Verhaegen, and D. A. Jaffray, "Efficient scatter distribution estimation and correction in CBCT using concurrent Monte Carlo fitting," *Med. Phys.*, vol. 42, no. 1, pp. 54–68, 2014.

[158] T. Niu, M. Sun, J. Star-Lack, H. Gao, Q. Fan, and L. Zhu, "Shading correction for on-board cone-beam CT in radiation therapy using planning MDCT images," *Med. Phys.*, vol. 37, no. 10, pp. 5395–5406, 2010.

[159] M. Maspero et al., "A single neural network for cone-beam computed tomography-based radiotherapy of head-and-neck, lung and breast cancer," *Phys. Imag. Radiat. Oncol.*, vol. 14, pp. 24–31, Apr. 2020.

[160] X. Liang et al., "Generating synthesized computed tomography (CT) from cone-beam computed tomography (CBCT) using CycleGAN for adaptive radiation therapy," *Phys. Med. Biol.*, vol. 64, no. 12, 2019, Art. no. 125002.

[161] E. Y.-K. Ng and R. U. Acharya, "Imaging as a diagnostic and therapeutic tool in clinical oncology," *World J. Clin. Oncol.*, vol. 2, no. 4, pp. 169–170, 2011.

[162] A. M. Sailer, W. H. van Zwam, J. E. Wildberger, and J. P. Grutters, "Cost-effectiveness modelling in diagnostic imaging: A stepwise approach," *Eur. Radiol.*, vol. 25, no. 12, pp. 3629–3637, 2015.

[163] D. R. Aberle et al., "Reduced lung-cancer mortality with low-dose computed tomographic screening," *New Eng. J. Med.*, vol. 365, no. 5, pp. 395–409, 2011.

[164] L. Zhou, S. Bai, Y. Zhang, X. Ming, Y. Zhang, and J. Deng, "Imaging dose, cancer risk and cost analysis in image-guided radiotherapy of cancers," *Sci. Rep.*, vol. 8, no. 1, 2018, Art. no. 10076.

[165] J. Deng, Z. Chen, K. B. Roberts, and R. Nath, "Kilovoltage imaging doses in the radiotherapy of pediatric cancer patients," *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 82, no. 5, pp. 1680–1688, 2012.

[166] E. Quaia, "Comparison between 80 kV, 100 kV and 120 kV CT protocols in the assessment of the therapeutic outcome in HCC," *Liver Pan. Sci.*, vol. 1, no. 1, pp. 1–4, 2016.

[167] J. P. Heiken, "Contrast safety in the cancer patient: Preventing contrast-induced nephropathy," *Cancer Imag.*, vol. 8, Oct. 2008, pp. S124–S127.

[168] G. S. Chilla, C. H. Tan, C. Xu, and C. L. Poh, "Diffusion weighted magnetic resonance imaging and its recent trend—A survey," *Quant. Imag. Med. Surg.*, vol. 5, no. 3, pp. 407–422, 2015.

[169] V. Kumar et al., "Radiomics: The process and the challenges," *Magn. Reson. Imag.*, vol. 30, no. 9, pp. 1234–1248, 2012.

[170] A. Traverso, L. Wee, A. Dekker, and R. Gillies, "Repeatability and reproducibility of radiomic features: A systematic review," *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 102, no. 4, pp. 1143–1158, 2018.

[171] I. D. Naranjo et al., "Diagnostic value of diffusion-weighted imaging with synthetic b-values in breast tumors: Comparison with dynamic contrast-enhanced and multiparametric MRI," *Eur. Radiol.*, vol. 31, no. 1, pp. 356–367, 2021.

[172] Y. Hu et al., "RUN-UP: Accelerated multishot diffusion-weighted MRI reconstruction using an unrolled network with U-Net as priors," *Magn. Reson. Med.*, vol. 85, no. 2, pp. 709–720, 2021.

[173] M. Kawamura et al., "Accelerated acquisition of high-resolution diffusion-weighted imaging of the brain with a multi-shot echo-planar sequence: Deep-learning-based denoising," *Magn. Reson. Med. Sci.*, vol. 20, no. 1, pp. 99–105, 2021.

[174] R. G. Price, J. P. Kim, W. Zheng, I. J. Chetty, and C. Glide-Hurst, "Image guided radiation therapy using synthetic computed tomography images in brain cancer," *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 95, no. 4, pp. 1281–1289, 2016.

[175] E. D. Morris et al., "Using synthetic CT for partial brain radiation therapy: Impact on image guidance," *Pract. Radiat. Oncol.*, vol. 8, no. 5, pp. 342–350, 2018.

[176] M. V. Sokolova, M. L. Khristova, E. M. Egorenkova, S. V. Leonov, and G. M. Babkina, "Adsorption study of the virions of the influenza virus and its individual proteins on polystyrene," *Vopr. Virusol.*, vol. 33, no. 3, pp. 369–36972, 1988.

[177] E. Johnstone et al., "Systematic review of synthetic computed tomography generation methodologies for use in magnetic resonance imaging-only radiation therapy," *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 100, no. 1, pp. 199–217, 2018.

[178] Varian. *Varian Medical Affairs Case Studies and Scoring System*. Accessed Aug. 20, 2020. [Online]. Available: http://medicalaffairs.varian.com/halcyon-case-studies

[179] I. Winz, *A Decision Support System for Radiation Therapy Treatment Planning*, Dept. Eng. Sci., Univ. Auckland, Auckland, New Zealand, 2004.

[180] K. Cheung, "Intensity modulated radiotherapy: advantages, limitations and future developments," *Biomed. Imag. Intervent. J.*, vol. 2, no. 1, p. e19, 2006.

[181] V. Batumalai, M. G. Jameson, D. F. Forstner, P. Vial, and L. C. Holloway, "How important is dosimetrist experience for intensity modulated radiation therapy? A comparative analysis of a head and neck case," *Pract. Radiat. Oncol.*, vol. 3, no. 3, pp. e99–e106, 2013.

[182] S. L. Berry, A. Boczkowski, R. Ma, J. Mechalakos, and M. Hunt, "Interobserver variability in radiation therapy plan output: Results of a single-institution study," *Pract. Radiat. Oncol.*, vol. 6, no. 6, pp. 442–449, 2016.

[183] B. E. Nelms et al., "Variation in external beam treatment plan quality: An inter-institutional study of planners and planning systems," *Pract. Radiat. Oncol.*, vol. 2, no. 4, pp. 296–305, 2012.

[184] I. J. Das, C.-W. Cheng, K. L. Chopra, R. K. Mitra, S. P. Srivastava, and E. Glatstein, "Intensity-modulated radiation therapy dose prescription, recording, and delivery: Patterns of variability among institutions and treatment planning systems," *J. Nat. Cancer Inst.*, vol. 100, no. 5, pp. 300–307, 2008.

[185] K. L. Moore, R. S. Brame, D. A. Low, and S. Mutic, "Experience-based quality control of clinical intensity-modulated radiotherapy planning," *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 81, no. 2, pp. 545–551, 2011.

[186] N. Ohri, X. Shen, A. P. Dicker, L. A. Doyle, A. S. Harrison, and T. N. Showalter, "Radiotherapy protocol deviations and clinical outcomes: A meta-analysis of cooperative group clinical trials," *J. Nat. Cancer Inst.*, vol. 105, no. 6, pp. 387–393, 2013.

[187] D. L. Craft, T. S. Hong, H. A. Shih, and T. R. Bortfeld, "Improved planning time and plan quality through multicriteria optimization for intensity-modulated radiotherapy," *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 82, no. 1, pp. e83–e90, 2012.

[188] L. J. Schreiner, "On the quality assurance and verification of modern radiation therapy treatment," *J. Med. Phys.*, vol. 36, no. 4, pp. 189–191, 2011.

[189] J. Van Dye, J. Batista, and G. S. Bauman, "Accuracy and uncertainty considerations in modern radiation oncology," in *The Modern Technology of Radiation Oncology*, vol. 3. Madison, WI, USA: Med. Phys. Publ., 2013, pp. 361–412.

[190] E. Yorke, K. E. Rosenzweig, R. Wagman, and G. S. Mageras, "Interfractional anatomic variation in patients treated with respiration-gated radiotherapy," *J. Appl. Clin. Med. Phys.*, vol. 6, no. 2, pp. 19–32, 2005.

[191] F. Liu, B. Erickson, C. Peng, and X. A. Li, "Characterization and management of interfractional anatomic changes for pancreatic cancer radiotherapy," *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 83, no. 3, pp. E423–E429, 2012.

[192] R. Height *et al.*, "The dosimetric consequences of anatomic changes in head and neck radiotherapy patients," *J. Med. Imag. Radiat. Oncol.*, vol. 54, no. 5, pp. 497–504, Oct. 2010.

[193] K. R. Britton *et al.*, "Consequences of anatomic changes and respiratory motion on radiation dose distributions in conformal radiotherapy for locally advanced non-small-cell lung cancer," *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 73, no. 1, pp. 94–102, 2009.

[194] M. Tinoco *et al.*, "RapidPlan development of VMAT plans for cervical cancer patients in low- and middle-income countries," *Med. Dosimetry*, vol. 45, no. 2, pp. 172–178, 2020.

[195] J. P. Tol, M. Dahele, A. R. Delaney, P. Doornaert, B. J. Slotman, and W. Verbakel, "Detailed evaluation of an automated approach to interactive optimization for volumetric modulated arc therapy plans," *Med. Phys.*, vol. 43, no. 4, pp. 1818–1828, 2016.

[196] M. Lin, C. Mingli, G. Xuejun, and L. Weiguo, "Deep learning-based inverse mapping for fluence map prediction," *Phys. Med. Biol.*, to be published.

[197] I. B. Mihaylov, E. A. Mellon, R. Yechieli, and L. Portelance, "Automated inverse optimization facilitates lower doses to normal tissue in pancreatic stereotactic body radiotherapy," *Plos One*, vol. 13, no. 1, 2018, Art. no. e0191036.

[198] S. Speer, A. Klein, L. Kober, A. Weiss, I. Yohannes, and C. Bert, "Automation of radiation treatment planning. Evaluation of head and neck cancer patient plans created by the Pinnacle³ scripting and auto-planning functions," *Strahlentherapie Onkologie*, vol. 193, no. 8, pp. 656–665, 2017.

[199] J. R. Teruel *et al.*, "Full automation of spinal stereotactic radiosurgery and stereotactic body radiation therapy treatment planning using Varian Eclipse scripting," *J. Appl. Clin. Med. Phys.*, vol. 21, no. 10, pp. 122–131, 2020.

[200] I. Xhaferllari, E. Wong, K. Bzdusek, M. Lock, and J. Z. Chen, "Automated IMRT planning with regional optimization using planning scripts," *J. Appl. Clin. Med. Phys.*, vol. 14, no. 1, pp. 176–191, 2013.

[201] Y. W. Yang, K. N. Shao, J. Zhang, M. Chen, Y. Y. Chen, and G. P. Shan, "Automatic planning for nasopharyngeal carcinoma based on progressive optimization in RayStation treatment planning system," *Technol. Cancer Res. Treat.*, vol. 19, Jan.–Dec. 2020, Art. no. 1533033820915710.

[202] M. Zarepisheh *et al.*, "Automated intensity modulated treatment planning: The expedited constrained hierarchical optimization (ECHO) system," *Med. Phys.*, vol. 46, no. 7, pp. 2944–2954, 2019.

[203] S. Cilla *et al.*, "Template-based automation of treatment planning in advanced radiotherapy: A comprehensive dosimetric and clinical evaluation," *Sci. Rep.*, vol. 10, no. 1, p. 423, 2020.

[204] D. Gintz *et al.*, "Initial evaluation of automated treatment planning software," *J. Appl. Clin. Med. Phys.*, vol. 17, no. 3, pp. 331–346, 2016.

[205] I. Hazell *et al.*, "Automatic planning of head and neck treatment plans," *J. Appl. Clin. Med. Phys.*, vol. 17, no. 1, pp. 272–282, 2016.

[206] J. Krayenbuehl, M. D. Martino, M. Guckenberger, and N. Andratschke, "Improved plan quality with automated radiotherapy planning for whole brain with hippocampus sparing: A comparison to the RTOG 0933 trial," *Radiat. Oncol.*, vol. 12, no. 1, p. 161, Oct. 2017.

[207] J. M. A. M. Kusters *et al.*, "Automated IMRT planning in pinnacle," *Strahlentherapie Onkologie*, vol. 193, no. 12, pp. 1031–1038, 2017.

[208] K. A. McConnell *et al.*, "Dosimetric evaluation of Pinnacle's automated treatment planning software to manually planned treatments," *Technol. Cancer Res. Treat.*, vol. 17, Jun. 2018, Art. no. 1533033818780064.

[209] Z. Ouyang *et al.*, "Evaluation of auto-planning in IMRT and VMAT for head and neck cancer," *J. Appl. Clin. Med. Phys.*, vol. 20, no. 7, pp. 39–47, 2019.

[210] A. Smith *et al.*, "Can the student outperform the master? A plan comparison between pinnacle auto-planning and eclipse knowledge-based RapidPlan following a prostate-bed plan competition," *Technol. Cancer Res. Treat.*, vol. 18, Jan.–Dec. 2019, Art. no. 1533033819851763.

[211] J. Q. Wang, Z. Chen, W. W. Li, W. Qian, X. S. Wang, and W. G. Hu, "A new strategy for volumetric-modulated arc therapy planning using AutoPlanning based multicriteria optimization for nasopharyngeal carcinoma," *Radiat. Oncol.*, vol. 13, no. 1, p. 94, May 2018.

[212] S. Wang *et al.*, "Technical assessment of an automated treatment planning on dose escalation of pancreas stereotactic body radiotherapy," *Technol. Cancer Res. Treat.*, vol. 18, Jun. 2019, Art. no. 1533033819851520.

[213] Q. B. Zhang, L. Y. Ou, Y. Y. Peng, H. Yu, L. J. Wang, and S. X. Zhang, "Evaluation of automatic VMAT plans in locally advanced nasopharyngeal carcinoma," *Strahlentherapie Onkologie*, vol. 193, no. 3, pp. 177–187, 2021.

[214] A. R. Delaney *et al.*, "Automated knowledge-based intensity-modulated proton planning: An international multicenter benchmarking study," *Cancers*, vol. 10, no. 11, p. 420, Nov. 2018.

[215] M. Monz, K. H. Kufer, T. R. Bortfeld, and C. Thieke, "Pareto navigation—Algorithmic foundation of interactive multi-criteria IMRT planning," *Phys. Med. Biol.*, vol. 53, no. 4, pp. 985–998, Feb. 2008.

[216] M. Hussein, B. J. M. Heijmen, D. Verellen, and A. Nisbet, "Automation in intensity modulated radiotherapy treatment planning—A review of recent innovations," *Br J Radiol*, vol. 91, no. 1092, Dec. 2018, Art. no. 20180270.

[217] S. Breedveld, P. R. M. Storchi, P. W. J. Voet, and B. J. M. Heijmen, "iCycle: Integrated, multicriterial beam angle, and profile optimization for generation of coplanar and noncoplanar IMRT plans," *Med. Phys.*, vol. 39, no. 2, pp. 951–963, Feb. 2012.

[218] M. Buschmann *et al.*, "Automated volumetric modulated arc therapy planning for whole pelvic prostate radiotherapy," *Strahlentherapie Onkologie*, vol. 194, no. 4, pp. 333–342, 2018.

[219] H. X. Chen, D. L. Craft, and D. P. Gierga, "Multicriteria optimization informed VMAT planning," *Med. Dosimetry*, vol. 39, no. 1, pp. 64–73, 2014.

[220] J. Dias, H. Rocha, T. Ventura, B. Ferreira, and M. D. Lopes, "A heuristic based on fuzzy inference systems for multiobjective IMRT treatment planning," in *Machine Learning, Optimization, and Big Data* (Lecture Notes in Computer Science), G. Nicosia, P. Pardalos, G. Giuffrida, and R. Umeton, Eds. Cham, Switzerland: Springer, 2018, pp. 255–267.

[221] L. Engberg, A. Forsgren, K. Eriksson, and B. Hardemark, "Explicit optimization of plan quality measures in intensity-modulated radiation therapy treatment planning," *Med. Phys.*, vol. 44, no. 6, pp. 2045–2053, 2017.

[222] S. Ghandour, O. Matzinger, and M. Pachoud, "Volumetric-modulated arc therapy planning using multicriteria optimization for localized prostate cancer," *J. Appl. Clin. Med. Phys.*, vol. 16, no. 3, pp. 258–269, 2015.

[223] F. Khan and D. Craft, "Three-dimensional conformal planning with low-segment multicriteria intensity modulated radiation therapy optimization," *Pract. Radiat. Oncol.*, vol. 5, no. 2, pp. E103–E111, 2015.

[224] R. G. J. Kierkels *et al.*, "Multicriteria optimization enables less experienced planners to efficiently produce high quality treatment plans in head and neck cancer radiotherapy," *Radiat. Oncol.*, vol. 10, p. 87, Apr. 2015.

[225] Y. H. Mai, F. T. Kong, Y. W. Yang, L. H. Zhou, Y. B. Li, and T. Song, "Voxel-based automatic multi-criteria optimization for intensity modulated radiation therapy," *Radiat. Oncol.*, vol. 13, no. 1, p. 241, Dec. 2018.

[226] E. Miguel-Chumacero, G. Currie, A. Johnston, and S. Currie, "Effectiveness of multi-criteria optimization-based trade-off exploration in combination with rapidplan for head & neck radiotherapy planning," *Radiat. Oncol.*, vol. 13, no. 1, p. 229, 2018.

[227] B. S. Muller, H. A. Shih, J. A. Efstathiou, T. Bortfeld, and D. Craft, "Multicriteria plan optimization in the hands of physicians: A pilot study in prostate cancer and brain tumors," *Radiat. Oncol.*, vol. 12, p. 168, Nov. 2017.

[228] D. Nguyen *et al.*, "Incorporating human and learned domain knowledge into training deep neural networks: A differentiable dose–volume histogram and adversarial inspired framework for generating Pareto optimal dose distributions in radiation therapy," *Med. Phys.*, vol. 47, no. 3, pp. 837–849, 2020.

[229] M. J. Wang *et al.*, "An interactive plan and model evolution method for knowledge-based pelvic VMAT planning," *J. Appl. Clin. Med. Phys.*, vol. 19, no. 5, pp. 491–498, 2018.

[230] D. Craft and T. Bortfeld, "How many plans are needed in an IMRT multi-objective plan database?" *Phys. Med. Biol.*, vol. 53, no. 11, pp. 2785–2796, 2008.

[231] D. Buergy *et al.*, "Fully automated treatment planning of spinal metastases—A comparison to manual planning of Volumetric Modulated Arc Therapy for conventionally fractionated irradiation," *Radiat. Oncol.*, vol. 12, no. 1, p. 33, 2017.

[232] S. J. M. Habraken *et al.*, "The TRENDY multi-center randomized trial on hepatocellular carcinoma—Trial QA including automated treatment planning and benchmark-case results," *Radiother. Oncol.*, vol. 125, no. 3, pp. 507–513, 2017.

[233] A. W. M. Sharfo *et al.*, "Validation of fully automated VMAT plan generation for library-based plan-of-the-day cervical cancer radiotherapy," *Plos One*, vol. 11, no. 12, 2016, Art. no. e0169202.

[234] A. W. M. Sharfo *et al.*, "Automated VMAT planning for postoperative adjuvant treatment of advanced gastric cancer," *Radiat. Oncol.*, vol. 13, no. 1, p. 74, Apr. 2018.

[235] R. van Haveren, B. J. M. Heijmen, and S. Breedveld, "Automatic configuration of the reference point method for fully automated multi-objective treatment planning applied to oropharyngeal cancer," *Med. Phys.*, vol. 47, no. 4, pp. 1499–1508, 2020.

[236] P. W. J. Voet, M. L. P. Dirkx, S. Breedveld, and B. J. M. Heijmen, "Automated generation of IMRT treatment plans for prostate cancer patients with metal hip prostheses: Comparison of different planning strategies," *Med. Phys.*, vol. 40, no. 7, Jul. 2013, Art. no. 071704.

[237] P. G. Bai *et al.*, "A knowledge-based intensity-modulated radiation therapy treatment planning technique for locally advanced nasopharyngeal carcinoma radiotherapy," *Radiat. Oncol.*, vol. 15, no. 1, p. 188, 2020.

[238] V. Chanyavanich, S. K. Das, W. R. Lee, and J. Y. Lo, "Knowledge-based IMRT treatment planning for prostate cancer," *Med. Phys.*, vol. 38, no. 5, pp. 2515–2522, 2011.

[239] S. Haseai, H. Arimura, K. Asai, T. Yoshitake, and Y. Shioyama, "Similar-cases-based planning approaches with beam angle optimizations using water equivalent path length for lung stereotactic body radiation therapy," *Radiol. Phys. Technol.*, vol. 13, no. 2, pp. 119–127, 2020.

[240] E. S. F. Liu, V. W. C. Wu, B. Harris, M. Lehman, D. Pryor, and L. W. C. Chan, "Vector-model-supported approach in prostate plan optimization," *Med. Dosimetry*, vol. 42, no. 2, pp. 79–84, 2017.

[241] C. McIntosh and T. G. Purdie, "Contextual atlas regression forests: Multiple-atlas-based automated dose prediction in radiation therapy," *IEEE Trans. Med. Imag.*, vol. 35, no. 4, pp. 1000–1012, Apr. 2016.

[242] S. Petrovic, G. Khussainova, and R. Jagannathan, "Knowledge-light adaptation approaches in case-based reasoning for radiotherapy treatment planning," *Artif. Intell. Med.*, vol. 68, pp. 17–28, Mar. 2016.

[243] B. Sarkar, A. Munshi, T. Ganesh, A. Manikandan, S. K. Anbazhagan, and B. K. Mohanti, "Standardization of volumetric modulated arc therapy-based frameless stereotactic technique using a multidimensional ensemble-aided knowledge-based planning," *Med. Phys.*, vol. 46, no. 5, pp. 1953–1962, 2019.

[244] M. Schmidt, J. Y. Lo, S. Grzetic, C. Lutzky, D. M. Brizel, and S. K. Das, "Semiautomated head-and-neck IMRT planning using dose warping and scaling to robustly adapt plans in a knowledge database containing potentially suboptimal plans," *Med. Phys.*, vol. 42, no. 8, pp. 4428–4434, 2015.

[245] E. Schreibmann, T. Fox, W. Curran, H. K. Shu, and I. Crocker, "Automated population-based planning for whole brain radiation therapy," *J. Appl. Clin. Med. Phys.*, vol. 16, no. 5, pp. 76–86, 2015.

[246] Y. Sheng *et al.*, "Atlas-guided prostate intensity modulated radiation therapy (IMRT) planning," *Phys. Med. Biol.*, vol. 60, no. 18, pp. 7277–7291, 2015.

[247] C. Amaloo, L. Hayes, M. Manning, H. Liu, and D. Wiant, "Can automated treatment plans gain traction in the clinic?" *J. Appl. Clin. Med. Phys.*, vol. 20, no. 8, pp. 29–35, 2019.

[248] A. Babier, J. J. Boutilier, A. L. McNiven, and T. C. Y. Chan, "Knowledge-based automated planning for oropharyngeal cancer," *Med. Phys.*, vol. 45, no. 7, pp. 2875–2883, 2018.

[249] X. Bai, G. P. Shan, M. Chen, and B. B. Wang, "Approach and assessment of automated stereotactic radiotherapy planning for early stage non-small-cell lung cancer," *Biomed. Eng. Online*, vol. 18, p. 101, Oct. 2019.

[250] S. L. Berry, R. T. Ma, A. Boczkowski, A. Jackson, P. P. Zhang, and M. Hunt, "Evaluating inter-campus plan consistency using a knowledge based planning model," *Radiother. Oncol.*, vol. 120, no. 2, pp. 349–355, 2016.

[251] E. Bossart, M. Duffy, G. Simpson, M. Abramowitz, A. Pollack, and N. Dogan, "Assessment of specific versus combined purpose knowledge based models in prostate radiotherapy," *J. Appl. Clin. Med. Phys.*, vol. 19, no. 6, pp. 209–216, 2018.

[252] J. J. Boutilier, T. Lee, T. Craig, M. B. Sharpe, and T. C. Y. Chan, "Models for predicting objective function weights in prostate cancer IMRT," *Med. Phys.*, vol. 42, no. 4, pp. 1586–1595, 2015.

[253] E. Cagni *et al.*, "Knowledge-based treatment planning: An inter-technique and inter-system feasibility study for prostate cancer," *Physica Medica Eur. J. Med. Phys.*, vol. 36, pp. 38–45, Apr. 2017.

[254] R. Castriconi *et al.*, "Knowledge-based automatic optimization of adaptive early-regression-guided VMAT for rectal cancer," *Physica Medica Eur. J. Med. Phys.*, vol. 70, pp. 58–64, Feb. 2020.

[255] A. Chatterjee, M. Serban, S. Faria, L. Souhami, F. Cury, and J. Seuntjens, "Novel knowledge-based treatment planning model for hypofractionated radiotherapy of prostate cancer patients," *Physica Medica Eur. J. Med. Phys.*, vol. 69, pp. 36–43, Jan. 2020.

[256] A. R. Delaney, M. Dahele, J. P. Tol, B. J. Slotman, and W. Verbakel, "Knowledge-based planning for stereotactic radiotherapy of peripheral early-stage lung cancer," *Acta Oncologica*, vol. 56, no. 3, pp. 490–498, 2017.

[257] A. M. Faught *et al.*, "Functional-guided radiotherapy using knowledge-based planning," *Radiother. Oncol.*, vol. 129, no. 3, pp. 494–498, 2018.

[258] A. Fogliata *et al.*, "RapidPlan knowledge based planning: Iterative learning process and model ability to steer planning strategies," *Radiat. Oncol.*, vol. 14, no. 1, p. 187, Oct. 2019.

[259] A. Fogliata *et al.*, "Performance of a knowledge-based model for optimization of volumetric modulated arc therapy plans for single and bilateral breast irradiation," *Plos One*, vol. 10, no. 12, Dec. 2015, Art. no. e0145137.

[260] A. Fogliata *et al.*, "A broad scope knowledge based model for optimization of VMAT in esophageal cancer: Validation and assessment of plan quality among different treatment centers," *Radiat. Oncol.*, vol. 10, p. 220, Oct. 2015.

[261] Y. R. Ge and Q. J. Wu, "Knowledge-based planning for intensity-modulated radiation therapy: A review of data-driven approaches," *Med. Phys.*, vol. 46, no. 6, pp. 2760–2775, 2019.

[262] E. Inoue *et al.*, "Dose–volume histogram analysis of knowledge-based volumetric-modulated arc therapy planning in postoperative breast cancer irradiation," *In Vivo*, vol. 34, no. 3, pp. 1095–1101, 2020.

[263] N. Kishi *et al.*, "Validation of the clinical applicability of knowledge-based planning models in single-isocenter volumetric-modulated arc therapy for multiple brain metastases," *J. Appl. Clin. Med. Phys.*, Oct. 21, no. 10, pp. 141–150, 2020

[264] Y. H. Lin, L. X. Hong, M. A. Hunt, and S. L. Berry, "Use of a con-strained hierarchical optimization dataset enhances knowledge-based planning as a quality assurance tool for prostate bed irradiation," *Med. Phys.*, vol. 45, no. 10, pp. 4364–4369, 2018.

[265] C. F. Ling *et al.*, "A hybrid automated treatment planning solution for esophageal cancer," *Radiat. Oncol.*, vol. 14, no. 1, p. 232, 2019.

[266] O. Nwankwo, H. Mekdash, D. S. K. Sihono, F. Wenz, and G. Glatting, "Knowledge-based radiation therapy (KBRT) treatment planning versus planning by experts: Validation of a KBRT algorithm for prostate cancer treatment planning," *Radiat. Oncol.*, vol. 10, p. 111, May 2015.

[267] C. Schubert *et al.*, "Intercenter validation of a knowledge based model for automated planning of volumetric modulated arc therapy for prostate cancer. The experience of the German RapidPlan Consortium," *Plos One*, vol. 12, no. 5, May 2017, Art. no. e0178034.

[268] Y. Sheng, Y. R. Ge, L. L. Yuan, T. R. Li, F. F. Yin, and Q. J. Wu, "Outlier identification in radiation therapy knowledge-based planning: A study of pelvic cases," *Med. Phys.*, vol. 44, no. 11, pp. 5617–5626, 2017.

[269] M. Shepherd *et al.*, "Developing knowledge-based planning for gynaecological and rectal cancers: A clinical validation of RapidPlan ((TM))," *J. Med. Radiat. Sci.*, vol. 67, no. 3, pp. 217–224, 2020.

[270] S. Shiraishi and K. L. Moore, "Knowledge-based prediction of three-dimensional dose distributions for external beam radiotherapy," *Med. Phys.*, vol. 43, no. 1, pp. 378–387, 2016.

[271] S. Shiraishi, J. Tan, L. A. Olsen, and K. L. Moore, "Knowledge-based prediction of plan quality metrics in intracranial stereotactic radiosurgery," *Med. Phys.*, vol. 42, no. 2, pp. 908–917, 2015.

[272] J. P. Tol, M. Dahele, A. R. Delaney, B. J. Slotman, and W. Verbakel, "Can knowledge-based DVH predictions be used for automated, individualized quality assurance of radiotherapy treatment plans?" *Radiat. Oncol.*, vol. 10, p. 234, Nov. 2015.

[273] Y. Ueda *et al.*, "Knowledge-based planning for oesophageal cancers using a model trained with plans from a different treatment planning system," *Acta Oncologica*, vol. 59, no. 3, pp. 274–283, 2020.

[274] T. Uehara, H. Monzen, M. Tamura, K. Ishikawa, H. Doi, and Y. Nishimura, "Dose–volume histogram analysis and clinical evaluation of knowledge-based plans with manual objective constraints for pharyngeal cancer," *J. Radiat. Res.*, vol. 61, no. 3, pp. 499–505, 2020.

[275] J. Q. Wang *et al.*, "Is it possible for knowledge-based planning to improve intensity modulated radiation therapy plan quality for planners with different planning experiences in left-sided breast cancer patients?" *Radiat. Oncol.*, vol. 12, p. 85, May 2017.

[276] A. Q. Wu *et al.*, "Robustness comparative study of dose–volume-histogram prediction models for knowledge-based radiotherapy treatment planning," *J. Radiat. Res. Appl. Sci.*, vol. 13, no. 1, pp. 390–397, 2020.

[277] H. Wu, F. Jiang, H. Z. Yue, S. Li, and Y. B. Zhang, "A dosimetric evaluation of knowledge-based VMAT planning with simultaneous integrated boosting for rectal cancer patients," *J. Appl. Clin. Med. Phys.*, vol. 17, no. 6, pp. 78–85, 2016.

[278] H. Wu, F. Jiang, H. Z. Yue, H. Zhang, K. Wang, and Y. B. Zhang, "Applying a RapidPlan model trained on a technique and orientation to another: A feasibility and dosimetric evaluation," *Radiat. Oncol.*, vol. 11, p. 108, Aug. 2016.

[279] K. C. Younge *et al.*, "Improving quality and consistency in NRG oncology radiation therapy oncology group 0631 for spine radiosurgery via knowledge-based planning," *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 100, no. 4, pp. 1067–1074, 2018.

[280] G. Yu *et al.*, "Knowledge-based IMRT planning for individual liver cancer patients using a novel specific model," *Radiat. Oncol.*, vol. 13, p. 52, Mar. 2018.

[281] Y. J. Zhang *et al.*, "A knowledge-based approach to automated planning for hepatocellular carcinoma," *J. Appl. Clin. Med. Phys.*, vol. 19, no. 1, pp. 50–59, 2018.

[282] B. P. Ziemer, P. Sanghvi, J. Hattangadi-Gluth, and K. L. Moore, "Heuristic knowledge-based planning for single-isocenter stereotactic radiosurgery to multiple brain metastases," *Med. Phys.*, vol. 44, no. 10, pp. 5001–5009, 2017.

[283] A. Fogliata *et al.*, "On the pre-clinical validation of a commercial model-based optimisation engine: Application to volumetric modulated arc therapy for patients with lung or prostate cancer," *Radiother. Oncol.*, vol. 113, no. 3, pp. 385–391, 2014.

[284] J. J. Boutilier, T. Craig, M. B. Sharpe, and T. C. Y. Chan, "Sample size requirements for knowledge-based treatment planning," *Med. Phys.*, vol. 43, no. 3, pp. 1212–1221, 2016.

[285] K. C. Snyder *et al.*, "Development and evaluation of a clinical model for lung cancer patients using stereotactic body radiotherapy (SBRT) within a knowledge-based algorithm for treatment planning," *J. Appl. Clin. Med. Phys.*, vol. 17, no. 6, pp. 263–275, 2016.

[286] D. Nguyen *et al.*, "A feasibility study for predicting optimal radiation therapy dose distributions of prostate cancer patients from patient anatomy using deep learning," *Sci. Rep.*, vol. 9, no. 1, p. 1076, 2019.

[287] X. Chen, K. Men, Y. Li, J. Yi, and J. Dai, "A feasibility study on an automated method to generate patient-specific dose distributions for radiotherapy using deep learning," *Med. Phys.*, vol. 46, no. 1, pp. 56–64, 2019.

[288] J. Fan, J. Wang, Z. Chen, C. Hu, Z. Zhang, and W. Hu, "Automatic treatment planning based on three-dimensional dose distribution predicted from deep learning technique," *Med. Phys.*, vol. 46, no. 1, pp. 370–381, 2019.

[289] D. Nguyen *et al.*, "3D radiotherapy dose prediction on head and neck cancer patients with a hierarchically densely connected U-net deep learning architecture," *Phys. Med. Biol.*, vol. 64, no. 6, 2019, Art. no. 065020.

[290] Z. Liu *et al.*, "A deep learning method for prediction of three-dimensional dose distribution of helical tomotherapy," *Med. Phys.*, vol. 46, no. 5, pp. 1972–1983, 2019.

[291] V. Kearney, J. W. Chan, S. Haaf, M. Descovich, and T. D. Solberg, "DoseNet: a volumetric dose prediction algorithm using 3D fully-convolutional neural networks," *Phys. Med. Biol.*, vol. 63, no. 23, 2018, Art. no. 235022.

[292] A. M. Barragán-Montero *et al.*, "Three-dimensional dose prediction for lung IMRT patients with deep neural networks: Robust learning from heterogeneous beam configurations," *Med. Phys.*, vol. 46, no. 8, pp. 3679–3691, 2019.

[293] D. Nguyen, A. S. Barkousaraie, C. Shen, X. Jia, and S. Jiang, "Generating Pareto optimal dose distributions for radiation therapy treatment planning," 2019. [Online]. Available: arXiv:1906.04778.

[294] G. Bohara, A. S. Barkousaraie, S. Jiang, and D. Nguyen, "Using deep learning to predict beam-tunable Pareto optimal dose distribution for intensity-modulated radiation therapy," *Med. Phys.*, vol. 47, no. 9, pp. 3898–3912, 2020.

[295] H. Lee *et al.*, "Fluence-map generation for prostate intensity-modulated radiotherapy planning using a deep-neural-network," *Sci. Rep.*, vol. 9, no. 1, 2019, Art. no. 15671.

[296] X. Li *et al.*, "Automatic IMRT planning via static field fluence prediction (AIP-SFFP): A deep learning algorithm for real-time prostate treatment planning," *Phys. Med. Biol.*, vol. 65, no. 17, 2020, Art. no. 175014.

[297] W. Wang *et al.*, "Fluence map prediction using deep learning models–direct plan generation for pancreas stereotactic body radiation therapy," *Front. Artif. Intell.*, vol. 3, p. 68, Sep. 2020.

[298] C. Shen, L. Chen, Y. Gonzalez, D. Nguyen, S. Jiang, and X. Jia, "Hierarchical deep reinforcement learning for intelligent automatic radiotherapy treatment planning," in *Proc. Annu. Meeting Amer. Assoc. Physicists Med. (AAPM)*, vol. 66, 2020.

[299] C. Shen, C. Liyuan, Y. Gonzalez, and X. Jia, "Improving efficiency of training a virtual treatment planner network via knowledge-guided deep reinforcement learning for intelligent automatic treatment planning of radiotherapy," 2020. [Online]. Available: arXiv:2007.12591.

[300] C. Shen *et al.*, "Intelligent inverse treatment planning via deep reinforcement learning, a proof-of-principle study in high dose-rate brachytherapy for cervical cancer," *Phys. Med. Biol.*, vol. 64, no. 11, 2019, Art. no. 115013.

[301] W. T. Hrinivich and J. Lee, "Artificial intelligence-based radiotherapy machine parameter optimization using reinforcement learning," *Med. Phys.*, vol. 47, no. 12, pp. 6140–6150, 2020.

[302] J. Zhang *et al.*, "An interpretable planning bot for pancreas stereotactic body radiation therapy," 2020. [Online]. Available: arXiv:2009.07997.

[303] Q. Yang, Y. Liu, T. Chen, and Y. Tong, "Federated machine learning: Concept and applications," *ACM Trans. Intell. Syst. Technol.*, vol. 10, no. 2, pp. 1–19, 2019.

[304] K. Bonawitz *et al.*, "Towards federated learning at scale: System design," 2019. [Online]. Available: arXiv:1902.01046.

[305] T. M. Deist *et al.*, "Distributed learning on 20000+ lung cancer patients—The personal health train," *Radiother. Oncol.*, vol. 144, pp. 189–200, Mar. 2020.

[306] Q. Zhang, Y. Nian Wu, and S.-C. Zhu, "Interpretable convolutional neural networks," in *Proc. CVPR*, 2018, pp. 8827–8836.

[307] F. Doshi-Velez and B. Kim, "Towards a rigorous science of interpretable machine learning," 2017. [Online]. Available: arXiv:1702.08608.