Recent Advancements of Artificial Intelligence in Particle Therapy

Hao Peng, Chao Wu, Dan Nguyen, Jan Schuemann, Andrea Mairani, Yuehu Pu, and Steve Jiang

Abstract—We are in a golden age of progress in the field of artificial intelligence (AI). Radiotherapy is perfectly suited to benefit from AI to enhance accuracy and efficiency due to its technology-intensive nature and direct human–machine interactions. While large amount of AI research have recently been published in the field of photon therapy, applications of AI specifically targeted for particle therapy remain scarcely investigated. There are two distinct differences between the photon therapy and particle therapy: 1) beam interaction physics (photons versus charged particles) and 2) beam delivery mode (e.g., IMRT/VMAT versus pencil beam scanning). Consequently, different strategies of AI deployment are required for these two radiotherapy modalities. In this article, we aim to present a comprehensive survey of recent literature exclusively focusing on AI-powered particle therapy. Six major aspects are included: 1) treatment planning; 2) dose calculation; 3) range and dose verification; 4) image guidance; 5) quality assurance; and 6) adaptive replanning. A number of perspectives, as well as potential challenges and common pitfalls, are also discussed.

Index Terms—Adaptive radiotherapy, artificial intelligence (AI), cone-beam CT (CBCT), dose calculation, dose verification, image guidance, positron emission tomography (PET), proton therapy, quality assurance (QA), treatment planning.

I. INTRODUCTION

In the last several years, a rising tide of artificial intelligence (AI) has been changing our world in many fields, especially healthcare. This new dawn of AI is largely attributable to the availability of big data, high-performance computing, and learning algorithms [1], [2]. AI has already entered the stage of applied technology and offers exciting opportunities for creating advanced tools in many healthcare research areas, particularly radiotherapy. Moving toward the goal of precise and personalized cancer treatment, the community has witnessed many technology advancements, including intensity-modulated radiation therapy (IMRT), volumetric modulated arc therapy (VMAT), image-guided radiation therapy (IGRT), stereotactic body radiation therapy (SBRT), as well as particle therapy (proton/carbon/helium). In radiotherapy, a typical workflow comprises numerous complex tasks: 1) imaging; 2) segmentation; 3) treatment planning; 4) patient positioning/immobilization; 5) treatment delivery; 6) quality assurance (QA); and 7) post-treatment follow-up [1], [2], [3], [4]. These tasks are quite labor and resource intensive, lending themselves well to AI integration in order to boost efficiency.

Compared to photon therapy, particle therapy has a unique feature called the Bragg peak. Its depth-dose profile starts with a relatively flat region on the proximal end, followed by a sharp Bragg peak over the distal end [5], [6], [7]. This crucial feature allows particle therapy to achieve superior dose conformity and healthy tissue protection compared to photon therapy. This physical advantage also imposes a much more stringent requirement on delivery accuracy, presenting itself as a double-edged sword. For instance, if unexpected range and dose errors occur, proton beams will yield a dose deposition deviating from planned results, causing either overdose or undertose patterns. Driven by this challenge, a number of AI-related research topics have attracted extensive interest in the community, including converting dual-energy CT (DECT) images or cone-beam CT (CBCT) images to stopping power (SP) ratio (SPR) maps, converting MR images to CT images for accurate dose calculation, converting CBCT images to CT images for adaptive replanning, performing accurate and fast dose calculations, and online range and dose verification. The selection of works covered herein aims at presenting readers with the most current research and exemplifying the deployment of AI as a promising tool, with much emphasis placed on the rational, motivation, and upcoming opportunities.

It is noteworthy that although many reviews on AI applications in radiotherapy have already been published [1], [2], [3], [4], [8], [9], [10], there is none exclusively focusing on particle therapy to the best of our knowledge. To fill that gap, the focus of this review will be narrowed down to the conjoining of AI with particle therapy. Some general topics applicable to both photon and particle therapy will thus be excluded to maintain the focus. Meanwhile, we would like to point out: 1) this review is designed to be a state-of-the-art type review summarizing the most recent developments; 2) AI in particle therapy itself is an emerging line of research and the comparison with non-AI techniques is often not easily found in most literature; and 3) we intend to engage readers’ broad interest.
in exploring AI as a potential tool, rather than delving into the technical details of different models. Out of these three considerations, we strive to make an appropriate balance between the breadth and depth of this review.

II. AI FOR TREATMENT PLANNING IN PARTICLE THERAPY

A. Stopping Power Ratio Mapping With Dual-Energy CT

The accurate calculation of SP is an indispensable component in particle therapy. Dose calculation based on computed tomography (CT) images requires the conversion from Hounsfield Units (HU) to SPR. Over the past decade, a widely used approach is a piecewise linear fit based on the HU numbers in a single-energy CT (SECT) measured with tissue surrogates [11], [12], [13], [14], [15], [16], [17], [18]. However, using such an empirical fitting for HU–SPR calibration causes inaccuracy due to three challenges: 1) many-to-one mapping; 2) intrinsic fitting error; and 3) patient-specific tissue heterogeneities. To tackle these challenges, DECT has been recently utilized to improve the estimation of SPR by several research groups. Among these studies, the analytical method based on interaction cross-sections (e.g., photoelectric, Compton scattering) derived from DECT images in a voxelwise manner has already been used in clinical routine.

Recently, several groups have started to leverage AI tools for this task [19], [20], [21], [22]. For example, Charyyev et al. proposed to generate synthetic DECT (sDECT) from SECT and then to derive SPR from sDECT [19] using a residual attention generative adversarial network (GAN). The corresponding SPR maps generated from sDECT are found to be comparable to the counterparts from original DECTs. In parallel, Wang et al. [20] successfully integrated a cycle-consistent GAN to predict SPR maps directly from DECTs. The datasets used in the above-mentioned two studies [19], [20] included 20 head-and-neck patients and used a leave-one-out cross validation. Despite the promising findings, a more thorough evaluation using, including a larger number of patients and different DECT acquisition, protocols is still necessary. Liu et al. [21] proposed the use of two convolutional neural networks (CNNs), UNet, and ResNet, for deriving elemental concentration with DECT. The main purpose of this study was to determine carbon and oxygen concentrations, but the method could easily be applied for hydrogen. When no noise was present, two AI models were able to obtain <2% mean absolute errors (MAEs) and <4% root mean-square errors in the derivation of carbon/oxygen concentration for brain images. Furthermore, different accuracy and noise immunity were indeed observed in the comparison between the UNet and ResNet.

The rationale of these AI-based approaches is elaborated below. Essentially, the role of AI is to map from one domain (CT images) to another domain (e.g., SPR and elemental concentration). Though conventional DECT-based approaches have been deployed for material decomposition and virtual enhancement, they face several well-known challenges, such as ill-posed inverse problems, nonuniqueness, and stability. AI-based models are able to avoid, or at least mitigate, these challenges. Furthermore, the AI models may be advantageous on the following two aspects: 1) geometric prior and 2) noise/artifact immunity [21]. Geometric prior is with regard to the difference between pixelwise and imagewise operations. For example, either UNet or ResNet extracts image features through a series of CNNs and downsampling steps. As a result, not only local features but also global features can be extracted. Said differently, spatial information can be utilized as prior information to enforce geometrical correlation among voxels, such as organ type and/or boundary. On the other hand, noise/artifact immunity can also benefit from AI and machine learning, a process closely tied to geometric priors (e.g., serving as a regularization tool). Considering the physical process of CT imaging, X-rays pass through all voxels staying on a pathway to obtain projected data, and an algorithm is used to reconstruct the image. As a result, voxels are not completely independent. Besides noise propagation in the backward projection, image artifacts, such as beam hardening and scattering make such a process more complicated. Once the machine-learning model was trained with high-SNR images, low-SNR images were tested and accurate derivation was still achieved, clearly suggesting good noise immunity and robustness of AI models [21].

B. MR-Based Treatment Planning

Radiation therapy heavily relies on CT as the imaging modality for treatment planning. However, the delineation of clinical target volumes (CTVs) based on CT images suffers from poor soft tissue contrast. Magnetic resonance imaging (MRI) has been proposed as a complementary modality to address this limitation. To help avoid the complexity and inaccuracy of image registration between separately acquired CT and MR images, as well as to help reduce the workload, MR-only treatment planning has emerged as a new line of research. Because MRI signals are not directly linked to CT HU numbers, one central step is to generate synthetic CT (sCT) images that can subsequently be used for planning.

The existing methods of sCT generation can be broadly categorized into atlas-based [23], [24], [25], [26], [27], [28], [29], segmentation-based [30], [31], [32], [33], [34], and AI-based methods [35], [36], [37], [38], [39], [40], [41], [42], [43]. One challenge faced by atlas-based methods is the handling of irregular anatomical structures in combination with registration errors. Segmentation-based methods are not only time consuming but also prone to errors in manual contouring [24]. To address these limitations, several groups have explored various AI approaches, including dictionary learning [35], random forest [36], and deep learning methods [37], [38], [39], [40], [41], [42], [43]. Unlike the other two approaches, deep learning methods offer more power in feature extraction and image mapping, while requiring minimum human interaction. For instance, Spada et al. [44] adopted a U-Net using 15 pairs of MRI/CT head scans to predict HU values for each voxel, and Neppl et al. [45] compared the performances of 2-D and 3-D U-Nets for head MRI images. Very recently, the feasibility of a U-Net-based method for pediatric patients with abdominal tumors was evaluated based on T1- and T2-weighted MR images [46]. The dosimetric difference between standard CT
and sCT is reported to be within 2%, possibly caused by interscan differences (e.g., bowel filling).

For the time being, a majority of MR-to-CT conversion depends on paired MR and CT images, like the U-Net CNN models mentioned above. However, possible misalignment between paired images can lead to errors in synthesizing CT images. To overcome this drawback, GAN has been introduced recently. For instance, Liu et al. [37], [38] used a 3-D dense cycle-GAN to generate abdominal and pelvic sCTs and Kazemifar et al. [43] also demonstrated the feasibility of using GAN with brain images. To pinpoint the impact of heterogeneous tissues, Maspero et al. assessed the feasibility of using conditional GAN with a dataset comprised of 60 pediatric brain MRI images [47]. In parallel, Shafai-Erfani et al. [48] adopted a 3-D cycle-GAN with brain MRI images.

From a machine-learning perspective, the advantage of choosing GAN over CNN for this specific task is apparent. CNNs are trained by minimizing voxelwise differences with respect to reference CT images that are rigidly aligned with MR images. As a result, voxelwise misalignment between these two image sets will lead to the blurring of synthesized images. One distinct benefit of GAN is its capability to work with unpaired images, due to the bidirectional generator/discriminator workflow and the enforcement of cycle consistency [49]. Such a strength has been proven in several computer vision tasks, such as image generation, style transfer, and realistic rendering. In radiotherapy, a training set including MR–CT pairs may sometimes be difficult to obtain and a GAN-based model may play its role in two scenarios: 1) MR or CT images are scanned under different protocols and 2) MR images are available, but without paired CT images.

MRI-only treatment planning in particle therapy emerges as an active field offering several potential benefits, such as eliminating MR–CT co-registration errors, reducing CT radiation exposure, and simplifying clinical workflow. The rational for AI deployment is not that different from Section II-A, and the same limitation related to the small dataset sizes applies. Future work should consider using larger and more heterogeneous datasets for more rigorous validation. Special attention should also be paid to examine to what extent an AI-based approach outperforms conventional approaches, in terms of quantitative dosimetric parameters.

III. AI FOR DOSE CALCULATION IN PARTICLE THERAPY

Dose calculation is a critical component of radiotherapy. With the current pace of technological advances, AI-powered dose calculation has already matured to the point of clinical translation, and it will surely continue to receive wider attention. Differing from photon therapy, two factors make the task of dose calculation more demanding in particle therapy: 1) the shape of Bragg peaks and 2) the physics of charged particle interaction with tissues. Two approaches for dose calculation in proton therapy have been widely studied: 1) pencil beam algorithm (PBA) and 2) Monte-Carlo simulation (MC) [50], [51], [52]. An important tradeoff is between the accuracy and efficiency. In a PBA approach, a number of pencil beams are ray traced individually as they deposit energies inside a patient. Its advantages include ease of use, commercial availability, and high speed. On the contrary, an MC approach models the step-by-step interaction of individual particles down to the very fundamental level of physics. It is believed that this approach delivers the gold standard with high accuracy, albeit at the cost of a heavier computational workload and longer calculation times. The limited accuracy of PBA stems from two problems. The first is that it does not fully embody tissue heterogeneity. Patients are modeled as a stack of semi-infinite layers, and the materials encountered by each pencil beam are assumed to be laterally homogenous. Such an assumption is not valid in highly heterogeneous sites, such as the head, neck, and lung. The second aspect is the impacts resulting from Coulomb scattering with orbital electrons and nuclear interactions, both of which are dependent on beam energy and depth. Numerous studies have already investigated the dosimetric discrepancy between the PBA and MC models [53], [54], [55], [56]. Over the years, the utilization of graphical process units (GPUs) has also been extensively investigated in order to speed up MC-based dose calculation [57], [58], [59], [60].

In a series of pioneering studies, our UT Southwestern group proposed the use of AI (a densely connected U-Net) for dose calculation in radiotherapy [61], [62], [63], starting with photon therapy and then expanding to proton therapy. As mentioned above, one dilemma in dose calculation is that fast algorithms are generally less accurate, while accurate dose engines are often time consuming. The group proposed to resolve this dilemma by exploring deep learning for the first time. Essentially, the AI model is expected to convert the less accurate results from a fast algorithm to the MC results with high accuracy. In one study [63], the model uses the PB doses and CT images as inputs to generate the MC doses as outputs. For a dataset of 290 patients (90 head and neck, 93 liver, 75 prostate, and 32 lung), the average gamma passing rate (1 mm/1%) between the AI-predicated dose maps and the counterparts directly from MC is 92.8% (head and neck), 92.7% (liver), 89.7% (lung), and 99.6% (prostate). The calculation for a single field takes less than 4 s.

Along the same path, another group used a 3-D CNN for head and neck cases [64]. Beam parameters, such as spot weight, spot position, and beam energies were used as inputs, and the dose maps obtained through MC were used as outputs. Moreover, a transfer learning technique was used in order to speed up training and improve the generalization capability of the CNN model, though the rational of taking such a step is not fully convincing in our view. Furthermore, as the authors pointed out, there exist two obstacles to be further addressed [64]: 1) multiple intermediate stages are required for pre- and post-processing and 2) the voxel resolution was set to 4 mm due to limited GPU memory capacity, larger than a typical value in clinical settings (2 to 3 mm).

Recently, Neishabouri et al. [65] proposed a long short-term memory (LSTM) recurrent neural network (RNN) model for pencil beams. Relative SP values of tissues were used as inputs, and dose maps obtained through MC simulation were used as outputs. The authors tested the model with digital phantoms and lung images, both of which exhibit a high degree
of tissue inhomogeneity (e.g., interfaces between lung tissues and high-density rib cages). When compared to MC results (the benchmark), agreement is found and the gamma-index pass rates stay between 94\% and 98\% for three beam energies (67.85, 104.25, and 134.68 MeV). The calculation time was about 1.5 ms for a single beamlet with a consumer GPU. The remaining task in this situation is understanding how to extend the current work to a complete treatment plan comprising of multiple pencil beamlets. Another study developed a hierarchically densely connected U-Net model [66], with dose maps calculated with the PBA and patient CT images as inputs. The outputs were the dose maps from MC simulations at different organ sites: head and neck, liver, prostate, and lung, similar to [63]. The model clearly showcases its strengths over PBA with regard to multiple evaluation criteria, including Gamma index, mean-square error, dose volume histogram (DVH), and dose difference. Moreover, the authors claimed that with additional efforts to improve the model efficiency (i.e., model compression), the calculation time can be further shortened.

Besides CNN and U-Net, one promising study investigated dose calculation using a discovery cross-domain GAN (DiscoGAN) [67]. The training data was generated using the MC simulations. In essence, the DiscoGAN was designed to perform the mapping between two domains: 1) beam parameters and 2) dose, while HU values from CT images and an analytically derived SP kernel were incorporated as auxiliary features. Information such as beam energy and cross-section can be implicitly embodied through the SP kernel. At the heart of this process is the capability of the AI model to capture the complicated relationship between dose and HU/SP. For a single beamlet, the calculation time is about 0.5 s. Accuracy was comparatively evaluated in terms of mean relative error (MRE) and MAE. The mean MRE is consistently below 3\% for all three sites (head and neck, thoracic, and abdomen). Furthermore, no systematic deviation, either overdose or underdose, is found between the AI and MC approaches.

When considering the design of AI models, one study [67] provided a good example in regards to two aspects: 1) choosing a suitable network model tailored for a specific task in particle therapy (e.g., the nature of generative capability) and 2) combining AI with physics to interpret the findings and potential benefits (e.g., analytically derived SP as prior information). The second aspect is often overlooked in AI-related papers in our field. As the authors stated, a heuristic view may help illuminate how the AI model functions. First, the SP modeling behaves quite like a PBA to yield a preliminary dose profile close to the true dose profile, as would be obtained by the MC simulation. Second, the AI model fine-tunes the preliminary dose profile to address minor discrepancies resulting from tissue heterogeneity and physical processes (Coulomb scattering and nuclear interactions), bringing the result even closer to the true dose profile.

IV. AI FOR RANGE AND DOSE VERIFICATION IN PARTICLE THERAPY

Range and dose verification in particle therapy is a task distinct from other sections in several aspects. First, there are secondary signals induced by particle beams, resulting from unique physical processes not seen with high-energy photon beams. Second, the task is closely tied to the instrumentation, such as hardware design and signal processing, making it more challenging compared to a task involving software only. Third, while the adoption of AI in this research area is of great promise, it is still in the very early phase and has limited literature available for discussion.

A major challenge in particle therapy is how to accurately monitor the location of the Bragg peak and real dose distribution. Possible uncertainty may result from a number of factors, such as beam profile, SP conversion (e.g., from CT images), patient positioning, and anatomical changes. The underlying principle is that the spatial distribution of secondary signals correlates with dose distribution. Several types of secondary signals have been examined, including positron emitter, prompt gamma, secondary electron bremsstrahlung (SEB) X-ray, acoustic wave, and water luminesces. Each category has its own strengths and limitations. Among them, the use of positron emission tomography (PET) has been extensively studied for detecting positron emitters (e.g., $^{11}$C, $^{13}$O), either with or without AI [68], [69], [70], [71], [72], [73], [74], [75], [76], [77], [78]. In our opinion, PET-based verification is the most promising tool to find clinical translation. From the perspective of AI and machine learning, the rational of AI deployment is highly similar among different verification approaches. To maintain the focus of this review, PET-related studies are given more weight and we encourage readers to delve into other applications that might be of interest to them.

A. Positron Emitter

In a series of studies, Peng et al. proposed the use of multiple AI models for range/dose verification in proton therapy for the first time [68], [69], [70], [71]. The logic behind these studies is clear: as the models become more complicated with increased amounts of physics-related information incorporated, AI unleashes its power. For instance, Li et al. [68] started with a simple forward neural network and a LSTM RNN to predict 1-D dose distribution for mono-energetic beams. Later, Liu et al. compared the performance of five RNN models: 1) LSTM; 2) bidirectional LSTM; 3) GRU; 4) bidirectional GRU; and 5) Seq2seq. The impact of including anatomical information (HU numbers) was also thoroughly examined, with the results suggesting that the bidirectional GRU structure achieves the most accurate prediction and the best generalization capability, especially with the presence of HU features. Built upon these two studies, the team gradually enhanced the AI framework on the following four aspects [70], [71]: 1) adding SP along with HU as prior information to enhance noise immunity and generalization capability; 2) realizing 3-D verification for both center and off-center voxels; 3) assessing spread-out Bragg peak (SOBP) cases in addition to mono-energetic cases; and 4) testing the performance of two input scenarios (reconstructed PET signals and raw positron emitter signals). Nevertheless, one limitation of these studies is that the dataset for model training/testing was based on a single CT image. The performance for different organ sites...
needs to be rigorously checked, even though the framework is devised for patient-specific verification.

When conducting AI-related studies in the field of particle therapy, or radiotherapy in general, two general questions should be answered. First, what is the rational or motivation for introducing AI to a given task? Second, are there proven benefits in comparison to non-AI approaches? To a certain degree, the answers to these two questions carry more weight than the development of an AI model itself. The studies in [68], [69], [70], [71] exemplify how to effectively do so. For instance, the authors fully explained why the mapping from a dose profile to an activity profile (e.g., positron emitter signals) is not a trivial task. The reason is because dose profiles depend on SP and medium, while the yield of proton-induced positron emitters depend on an additional parameter, the nuclear reaction cross-section. Consequently, a typical activity versus depth profile exhibits complicated fluctuation patterns, distinct from a smooth dose versus depth profile. The selection of RNN-based models is due to its strength in extracting the underlying correlation between an input sequence (activity profile) and an output sequence (dose profile), while requiring a reduced number of parameters. Furthermore, the authors thoroughly explained the necessity of including HU and/or SP as extra features. When no anatomical information is incorporated, an RNN model focuses largely on global features. The inclusion of HU and SP helps the RNN model to capture local correlation. With regard to the individual contribution of HU and SP, the former is associated with the cross-section of the X-ray (i.e., information of carbon and oxygen concentration in tissues) and correlates with the activity profile in an indirect way, while SP is directly linked to the dose profile.

With regard to potential benefits relative to non-AI models, the authors provided plausible explanations as well [69], [70], [71]. AI is easy to use and has the potential to provide an end-to-end solution. It extracts features automatically, different from preselected functions and fitting routine in kernel-based models (e.g., the convolution of a Gaussian function and a power-law function) [72], [73], [74], [75]. It demonstrates good robustness, generalization capability, and noise immunity. Some quantitative comparisons are also presented in dose [76] and range verification [72], [73], [76], [77], [78], between the AI and non-AI studies. Though a direct comparison is sometimes difficult to make, indirect comparisons like these are still conducive to other colleagues working on the same topic.

### B. Prompt Gamma

During the interaction of charged particles with tissues, characteristic photons, also known as prompt gamma (PG), are emitted. Similar to PET-based verification, PG signals are highly correlated with dose distribution [75], [79], [80], [81]. Gueth et al. proposed an approach to detect possible discrepancies between planned and delivered dose [80] based upon a combined classifier using distal falloff and registered correlation as features. Liu and Huang applied a U-Net model to tackle the same task with brain phantoms [81], selecting MC-simulated PG signals as inputs and dose maps as outputs. Recently, Schumann et al. [75] proposed to combine a filtering procedure based on Gaussian-power law convolution with an evolutionary algorithm. For all these studies, however, a number of physical factors that will degrade the quality of PG signals (i.e., counting statistics, limited spatial resolution, image artifacts), have yet to be examined.

#### C. Secondary Electron Bremsstrahlung X-Rays

SEB X-rays can also be utilized for range and dose verification [82], [83], [84], [85], [86], with a prototype system already developed and tested in both proton therapy [83] and carbon therapy [84], [85]. The goal is to reveal how to convert SEB X-ray to dose images, while facing the presence of two challenges: 1) limited spatial resolution and 2) poor counting statistics. To address these challenges, Yamaguchi et al. [86] proposed the use of two U-Net models, one for X-ray to dose conversion, and the other for resolution enhancement. Instead of using MC simulations, a predefined analytical function with parameters extracted from experimental measurements was adopted to ease the workload of data generation. Despite the promising results, it can be foreseen that a formidable obstacle exists when the AI approach is applied to heterogeneous tissues.

#### D. Acoustic Signals

Numerous studies have exploited thermoacoustic signals for range/dose verification [87], [88], [89], [90], [91], [92], [93], [94]. Due to the characteristics of signal production, two types of pressure waves are emitted by the prepeak and the peak dose deposition. The time-of-flight (TOF) method has been investigated for a uniform water medium [87], [88]. In parallel, the time-reversal (TR) reconstruction method in both 2-D and 3-D heterogeneous tissues has been proposed to tackle two challenges associated with the TOF method [91], [92]: 1) complicated extraction of arrival time and 2) incapability to provide complete dose information. However, compared with the TOF method, the TR method is much more time consuming, taking up to several minutes for 3-D calculation even with GPU acceleration. To address the time constraint, Yao et al. developed an AI model (Bi-LSTM RNN) for both 2-D and 3-D scenarios [93], [94], which can identify the correlation between the acoustic waveforms and dose profiles, in combination with advanced signal processing techniques (e.g., Hilbert transform, wavelet decomposition, etc.). Compared to the TR approach, the AI model requires much shorter computational time (several seconds) and a reduced number of sensors for detecting acoustic waves.

The advantage of an RNN model for the task lies in its unique strength of extracting sequential correlation, as well as requiring a small number of hyperparameters. As illustrated in [93] and [94], every single dose profile bears a close correlation with a unique 2-D map comprising of multichannel acoustic waveforms. The RNN model then analyzes the characteristics of time-series signals for feature extraction. For example, one feature of particular importance is the relative time offset of amplitude peaks among individual detector channels. The authors also mentioned that treating multichannel signals all at once using a CNN is also a viable solution,
but two potential concerns make such a choice less appealing. First, a CNN requires a much larger number of hyperparameters to be trained (i.e., demanding more data samples for training). Second, a CNN may not provide the same strength in identifying temporal correlation as an RNN model. Despite these potential advantages, data generation for model training remains a formidable challenge. Specifically, the propagation of acoustic waves and their shapes strongly rely on the set of physical parameters (e.g., sound speed in heterogeneous tissue). Both input and output for machine learning have originated from simulation, for both non-AI and AI approaches [91], [92], [93], [94]. How closely the simulated dataset represents the signals in practice needs to be carefully examined once a prototype system is built.

E. Luminescence

The luminescence signal from water during particle therapy is another line of research [95], [96], [97]. The luminescence of water is produced through a process similar to Cerenkov light, which can be detected with a cooled charge-coupled device (CCD) camera. Yabe et al. [97] used a U-Net to predict 2-D dose distributions from the measured luminescence images of a water medium for both proton and carbon beams. Similar to the SEB X-rays described above, this approach will encounter obstacles when applied to heterogeneous and/or deep-seated tissues. Furthermore, we would like to reiterate our belief that for this type of research, as well as many others mentioned in this section, the dominating challenge is not tied to AI itself, as even a very basic AI model may fulfill the requirements of mapping detected signals to dose distribution. Instead, the major challenge stays within other aspects, such as instrumentation and signal detection. For this reason, a justification of choosing AI over a non-AI approach, as well as the plan for performance comparison between two avenues, should be offered.

V. AI for CBCT Image Guidance in Particle Therapy

Image guidance is another important component for precision radiotherapy, particularly toward the realization of adaptive radiotherapy. The need for adaptation in particle therapy is more critical than in photon therapy, due to its unique Bragg peak and steep dose gradient [98], [99], [100]. Any patient-related deviation from the initial treatment plan (positioning, anatomy, etc.), will cause dose inconsistencies and consequently compromise treatment efficacy. CBCT is a widely used tool for image guidance, and one already integrated with a number of proton therapy systems. CBCT has manifested its role in the verification of patient positioning, monitoring of anatomical changes (intrafractional changes), and adjusting treatment plans.

However, the image quality of CBCT is not sufficient for yielding accurate HU numbers [101], [102], [103], [104], [105]. Over the past decade, numerous non-AI approaches have been proposed to tackle this issue, including look-up tables [101], histogram matching [102], deformable image registration (DIR) [103], [104], and improved scatter correction [105]. Recently, leveraging the strength of AI has attracted increasing attention. The discussion in this section focuses exclusively on two aspects: 1) CBCT to CT conversion [106], [107], [108], [109], [110], [111], [112] and 2) CBCT to SPR mapping [113], [114], [115]. As a matter of fact, there is a very fine line between these two aspects since SPR derivation is the ultimate goal, and several studies consist of both for completeness. Furthermore, we would like to clarify two points. First, the tasks of image conversion/synthesis and SPR derivation share many common features, as shown in the examples presented in Section II; therefore, the emphasis below is placed on CBCT (including scatter correction). Second, besides image guidance, the creation of an adaptive treatment needs to take into account beam-related uncertainties (spot position, monitor unit (MU), beam energy, etc.) and replanning algorithms, a topic to be discussed in Section VI.

A. CBCT to CT Image Conversion

Acquiring paired CT and CBCT images is a difficult task in practice. As discussed in Section II-B, GAN is able to alleviate the need for paired CT and CBCT images. For instance, Liang et al. [106] adopted a cycle-GAN to generate sCTs from CBCT images for head-and-neck patients. Similarly, Kurz et al. [110] evaluated the feasibility for both VMAT and proton therapy. The authors conclude that the accuracy of dose calculation based upon sCTs is sufficient for VMAT, but not for proton therapy.

AI approaches have also been reported when paired CT and CBCT images are available. Hansen et al. [107] used a 2-D U-Net to correct CBCT images for both IMPT and VMAT plans. For pelvic patients, the results exhibit satisfactory accuracy only for VMAT plans. In another study [108], Landry et al. trained a U-Net with three types of datasets for prostate images: 1) raw CBCT and CT images; 2) raw CBCTs and DIR-sCTs; and 3) raw CBCT and scatter-corrected CBCT. Such design intends to decouple the impact of scatter correction and deformation, with the third dataset achieving the best performance. Thummerer et al. [109] used a dataset of 27 head and neck patients containing planning CT, repeated CTs, CBCTs, and MRs to train a U-Net for sCT generation. The attained results suggest that CBCT-based sCTs have a higher degree of similarity relative to planning CTs than MR-based sCTs, and that both are equally suited for daily adaptive proton therapy.

B. CBCT for SPR Mapping

The goal of this task is to improve SPR estimation and dose calculation, similar to Section II-A. Harms et al. [113] recently proposed a cycle-GAN to extend its role from CBCT to CT conversion to CBCT to SPR mapping. The authors conclude that the AI approach achieves comparable, if not superior, performance to that of a DIR method for head-and-neck patients, in terms of MAE, mean error (ME), peak signal-to-noise ratio (PSNR), and structural similarity (SSIM). The MAE between CT-based and CBCT-based SPRs is 0.06 ± 0.01 and the ME is −0.01 ± 0.01. There are two limitations in the current study [113]: 1) the empirically selected HU–SPR curve was used
for data generation, which had an intrinsic inaccuracy up to a root-mean-squared error of 5.5% and 2) the rational of the direct inclusion of SPR for learning demands a clearer justification.

One crucial task to improve the CBCT image quality is scatter correction. Conventional scatter correction strategies require either complicated analytical models with ad hoc assumptions, or heavy computational burdens such as the well-known scatter kernel superposition algorithm. The role of AI models for more effective scatter correction and accurate SPR derivation, is exemplified in the following two examples. Lalonde et al. [114] evaluated the performance of a U-Net to evaluate scatter correction for 48 head and neck CBCT images. Dosimetric performance and proton range were compared among three scenarios: 1) scatter-free (ground truth obtained through MC simulation); 2) uncorrected; and 3) scatter-corrected CBCT images. For AI-powered scatter correction, the mean HU difference with regard to the ground-truth decreases from −28.6 HU (uncorrected images) down to −0.8 HU (corrected images). The root-mean-square error of proton range between the ground truth and the scatter-corrected scenario is 0.73 mm. The correction for the complete image volume can be completed in less than 5 s. In parallel, Nomura et al. [115] evaluated a U-Net for two sites, head/neck CBCT images. Dosimetric performance and proton range were compared among three scenarios: 1) scatter-free (ground truth obtained through MC simulation); 2) uncorrected; and 3) scatter-corrected CBCT images. For AI-powered scatter correction, the mean HU difference with regard to the ground-truth decreases from −28.6 HU (uncorrected images) down to −0.8 HU (corrected images). The root-mean-square error of proton range between the ground truth and the scatter-corrected scenario is 0.73 mm. The correction for the complete image volume can be completed in less than 5 s.

Three recent studies considered the AI-powered scatter correction for 64 head and neck CBCT images. For AI-powered scatter correction, the mean HU difference with regard to the ground-truth decreases from −28.6 HU (uncorrected images) down to −0.8 HU (corrected images). The root-mean-square error of proton range between the ground truth and the scatter-corrected scenario is 0.73 mm. The correction for the complete image volume can be completed in less than 5 s.

A. Patient-Specific Quality Assurance

QA is a critical component in radiotherapy and its main goal is to guarantee the accuracy of dose delivery [116]. Since QA is not only a labor-intensive process but also prone to measurement uncertainties, AI tools are of great potential in this area. Back in 2011, Zhu et al. [117] developed a support vector regression (SVR) model to establish the correlation between DVHs of OARs (bladder and rectum) and anatomical information (e.g., organ volumes, distance-to-target histogram). Recently, AI techniques have also been adopted as a secondary check tool for the prediction of MUs and dose output.

For example, before the wide spread of the pencil beam scanning mode in particle therapy, treatment planning systems did not have built-in modules to calculate MUs in passive scattering mode. Phantom measurements are thus required to determine the field-specific dose output, a routine subjected to measurement errors and limited machine time. To address this problem, Sun et al. [118] proposed an AI-based approach to predict dose output (cGy/MU) using gantry angle and field size as features. The authors conclude that all three models (Random-forest, XGBoost, and Cubist) outperform an empirical model, for a dataset comprising 1754 treatment fields and phantom measurements. A similar work was also reported for the pencil beam scanning mode [119] in which Grewal et al. used Gaussian process regression (GPR) and shallow neural networks to predict multiple QA parameters (range, modulation, field size, and output factor). Again, for a training dataset of 4231 patient-specific QA measurements, both models outperform an empirical model.

Very recently, the use of AI for the prediction of treatment delivery errors has built momentum. Maes et al. [120] developed three models (linear regression, random-forest, and neural network) where the planned spot parameters (e.g., spot position, MUs and energy) were extracted from TPS as inputs, and the delivered spot parameters were extracted from log files as outputs. The dataset contained treatment plans of 20 prostate patients. An intriguing comparison presented is that in terms of standard deviation, the uncertainty in X/Y positions (difference between plan and delivery) reduces from 0.39/0.44 mm, down to 0.22/0.11 mm (difference between prediction and delivery). The random-forest model offered the best predictive power in this case. Two comments deserve mention here: first, although the preliminary feasibility of AI-based identification of delivery errors was proven, a much larger dataset is needed. Second, a logical question arises immediately—whether and how such predicted error be taken into account before beam delivery.

B. Decision Support System

The idea of leveraging AI in the development of decision support tools has been an exciting part of healthcare for decades. These tools have the potential to provide valuable insights on diagnosis, treatment options, and prognosis. Unfortunately, many efforts failed in the migration process from research to clinical practice, largely due to the deficiency in the design of human–computer interaction and the consideration of collaborative nature of clinical workflow [121], [122], [123].

When narrowed down to particle therapy, one potential application of a clinical decision system (CDS) is connecting...
past treatment decisions with current assessments, to help clinicians efficiently identify the optimum course of treatment with regard to the selection of treatment modality and dose prescription. Valdes et al. [124] proposed an AI approach for early-stage lung and postoperative oropharyngeal cancer patients treated with photon or proton. A library of historical treatment plans and patient-specific feature sets were used to construct the classifiers. The authors claim that based upon the learning curves, 45, 60, and 30 patients are needed for developing a sufficiently accurate classification model for early-stage lung, postoperative oropharyngeal (photon), and postoperative oropharyngeal (proton), respectively. Given both the complexity of the proposed task and the large degree of variations among patients, the above-mentioned small data sizes seem to be overly optimistic.

C. Replanning for Adaptive Therapy

This is a task interleaved with multiple tasks covered in this review, centering on how to make a treatment adaptive. The generation of an adapted plan involves two subtasks: 1) fast dose calculation and 2) fast plan optimization. The computational speed previously presented as a bottleneck, but GPUs have recently been deployed for increasing speed [125], [126], [127], [128]. The total time for generating a plan can be reduced to five minutes (MC dose calculation) [125] and 10 s (analytical dose calculation) [127]. Optimization of a plan can be devised along two avenues. One avenue is to moderately alter the position of Bragg peaks based on the latest geometry. For instance, first adapting the energy of each pencil beam to the new water equivalent path length, and then reoptimizing beam weights using a standard optimization solver [126]. The other avenue is to produce a plan with newly added beamlets/spots (i.e., less constraints in replanning), which can potentially improve dose conformity and spare OARs. This would be beneficial in cases where the relative distance between the PTV and OAR significantly differs from the original plan [129], [130].

Range and dose verification in Section IV is able to help form a closed loop in the adaptive workflow, a critical step to evince the dosimetric advantage of particle therapy over other modalities. Besides proton-induced secondary signals, this can also be achieved utilizing log files [127], [131], [132], similar to the studies in Section VI.

In our view, being adaptive comprises both before- and after-delivery adaptation. Specifically, if the anatomy has not changed above a threshold, the delivery can proceed as planned and no adaptation is needed. If the anatomy does change noticeably, the steps of dose influence recalculation and plan reoptimization steps are required before delivery. Finally, after-delivery dose verification can be performed to check whether fine-tuning spot weights between fractions is necessary. When synthesizing the above-mentioned issues altogether, two intriguing questions arise. First, can an AI framework be used for online adaptation immediately following after-delivery verification while the patient is still on the treatment bed? Second, can an AI framework be used for the fully automatic generation of adapted plans? More information about this topic can be found in [133]. As the field of particle therapy becomes more prevalent and sees an increasing number of AI-powered applications (e.g., dose calculation in Section III, dose verification in Section IV, CBCT image guidance in Section V), AI-powered replanning will be the next step.

VII. COMMON CHALLENGES AND PITFALLS

In spite of the promising progress summarized in this review, we must point out that AI is no silver bullet and faces its own challenges. As models become increasingly complex, it becomes more challenging to inspect how inputs and outputs have been manipulated, as well as to interpret and check results. A number of common pitfalls to be avoided can be found in [134]. Below, we briefly discuss three challenges and pitfalls highly prevalent in the field of radiotherapy (both particle and photon therapy), largely based on our own experiences.

1) Generating and Splitting Data Appropriately: Compared to other fields, such as computer vision and natural language processing, data scarcity is a huge bottleneck in particle therapy as encountered in a majority of examples we have discussed so far. The limited size of a dataset may cause several well-known problems, such as overfitting. To address such a limitation, data augmentation through MC simulation has been frequently used. How to effectively generate sufficient datasets, being both task and model specific, is critical. Three commitment issues related to data generation should also be considered. First, whether MC simulation produces exact or comparable results as would be expected in practical settings (e.g., system setup, beam profile, and image acquisition protocols), should be carefully examined. Second, as its name suggests, a patient-specific task (e.g., dose calculation, dose verification, and QA) demands data generation per patient basis. How to balance a tradeoff between being patient-specific and obtaining good generalization capability is important, since the latter will have larger sample complexity and higher efficiency of data utilization. Transfer learning can also be considered here. Third, AI practitioners typically split data into training, validation, and test sets. The splitting may or may not be done in a completely random manner. Convergence behaviors and bias/variance tradeoff should always be reported. Careful consideration and exploration of different approaches is highly encouraged to ensure research findings are consistent.

2) Interpretation With Respect to Features and Hidden Variables: This is an extremely important aspect in AI-related studies, but is often overlooked. Based on our experiences, this aspect greatly helps avoid the overfitting problem. A well-known story in the machine-learning field is the “tank problem” [134], [135]. Researchers developed a model to spot tanks in pictures provided by the military. The model found the tanks successfully in the test dataset, but failed later with real pictures. Simply put, other image features, such as morning light or clouds drove the model, instead of the
presence of tanks. In practice, a good way to check this is to use the same model to make other predictions. If it succeeds, the results and study design may be skeptical. Furthermore, by inspecting hidden variables, a more in-depth comparison can be made among different models as exemplified in [69]. With regard to input features and their respective contributions, several rounds of cross-validation should be considered. For example, the AI model was deployed for dose verification with three inputs: 1) activity profile; 2) HU from CT images; and 3) SP [70]. The authors intentionally altered each input and evaluated its impact on the overall performance, which was a good way to better understand how the model works inside a black box. In another study, the order of acoustic waveforms was randomly permuted to pinpoint whether the time difference and sequential information were actually utilized for identifying the location of Bragg peaks [93], [94]. Many AI papers in the field of particle therapy fail to make the final step and perform such experiments.

3) **Starting With the Simplest Model:** It seems natural for researchers to go with complex models, but sometimes a basic model without the use of neural networks, or with just a shallow neural network, performs equally well as a deep neural network with many layers. It is important to keep the overall aim in mind. Otherwise, one may be setting up an unnecessarily complicated model to solve a simple problem, or even the wrong problem. If the purpose is only to predict the Bragg peak, a basic forward neural network [68] or a regression classifier [80] will suffice. Unless the purpose is to perform 3-D dose verification, there is no need to choose advanced models such as RNN or GAN [68], [69], [70], [71], [80]. Trying a simple baseline model is also useful in several regards: better understanding the dataset, establishing a performance baseline, choosing correct figures-of-merit, and guiding ablation studies if necessary.

Researchers conducting AI studies in the field of particle therapy should familiarize themselves with these common challenges and pitfalls and hold themselves to higher standards. The soundness check should always be rigorously conducted. A clear standard on how to perform and report AI results should be established.

VIII. **Conclusion**

Computational power, big data, and advanced algorithms are coming together to unleash the power of AI in radiotherapy. Due to two distinct differences between the photon therapy and particle therapy (beam interaction physics and beam delivery mode), different strategies should be devised accordingly. Based on the examples considered in this review, AI deployment in particle therapy will not be “old wine in new bottles,” but has great potential to address a number of unique unmet needs for boosting both accuracy and efficiency. Data generation, result interpretation, incorporation of fundamental physical processes, and rigorous validation are four aspects as critical as the AI model development itself. Clear standards about how to conduct and publish AI results need to be clearly established in the community. A clear pathway lies ahead of us to push the limits of AI tools toward more effective particle therapy.

**References**

[1] R. F. Thompson et al., “Artificial intelligence in radiation oncology: A specialty-wide disruptive transformation?” *Radiother. Oncol.*, vol. 129, no. 3, pp. 421–426, 2018.

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

[3] P. Meyer, V. Noblet, C. Mazzara, and A. Lallement, “Survey on deep learning for radiotherapy,” *Comput. Biol. Med.*, vol. 98, pp. 126–146, Jul. 2018.

[4] K. J. Kiser, C. D. Fuller, and V. K. Reed, “Artificial intelligence in radiation oncology treatment planning: A brief overview,” *J. Med. Artif. Intell.*, vol. 2, p. 9, May 2019.

[5] H. Paganietti, “Range uncertainties in proton therapy and the role of Monte Carlo simulations,” *Phys. Med. Biol.*, vol. 57, no. 11, pp. R99–R117, 2012.

[6] A. Lomax, “Intensity modulation methods for proton radiotherapy,” *Phys. Med. Biol.*, vol. 44, no. 1, pp. 185–205, 1999.

[7] W. D. Newhauser and R. Zhang, “The physics of proton therapy,” *Phys. Med. Biol.*, vol. 60, no. 8, pp. R155–R209, 2015.

[8] A. Hoffmann et al., “MR-guided proton therapy: A review and a preview,” *Radiat. Oncol.*, vol. 15, no. 1, p. 129, 2020.

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

[10] D. Jarrett, E. Stride, K. Vallis, and M. J. Gooding, “Applications and limitations of machine learning in radiation oncology,” *Brit. J. Radiol.*, vol. 92, no. 1100, 2019, Art. no. 20190001.

[11] R. L. Grant et al., “Relative stopping power measurements to aid in the design of anthropomorphic phantoms for proton radiotherapy,” *J. Appl. Clin. Med. Phys.*, vol. 15, no. 2, pp. 121–126, 2014.

[12] V. T. Tausti et al., “Comparison of single and dual energy CT for stopping power determination in proton therapy of head and neck cancer,” *Phys. Imag. Radiat. Oncol.*, vol. 6, pp. 14–19, Apr. 2018.

[13] U. Schneider, E. Pedroni, and A. Lomax, “The calibration of CT Hansfield units for radiotherapy treatment planning,” *Phys. Med. Biol.*, vol. 41, no. 1, pp. 111–124, 1996.

[14] D. C. Hansen et al., “A simulation study on proton computed tomography (CT) stopping power accuracy using dual energy CT scans as benchmark,” *Acta Oncologica*, vol. 54, no. 9, pp. 1638–1642, 2015.

[15] N. Hunemohr, B. Krauss, C. Tremmel, B. Ackermann, O. Jäkel, and S. Greilich, “Experimental verification of ion stopping power prediction from dual energy CT data in tissue surrogates,” *Phys. Med. Biol.*, vol. 59, no. 1, pp. 83–96, 2014.

[16] M. Yang, G. Virshup, J. Clayton, X. R. Zhu, R. Mohan, and L. Dong, “Theoretical variance analysis of single- and dual- energy computed tomography methods for calculating proton stopping power ratios of biological tissues,” *Phys. Med. Biol.*, vol. 55, no. 5, pp. 1343–1362, 2010.

[17] A. E. Bourque et al., “A stoichiometric calibration method for dual energy computed tomography,” *Phys. Med. Biol.*, vol. 59, no. 8, pp. 2059–2088, 2014.

[18] J. Zhu and S. N. Penfold, “Dosimetric comparison of stopping power calibration with dual-energy CT and single-energy CT in proton therapy treatment planning,” *Med. Phys.*, vol. 43, no. 6, pp. 2845–2854, 2016.

[19] S. Charyyev et al., “Learning-based synthetic dual energy CT imaging from single energy CT for stopping power ratio calculation in proton radiation therapy,” 2020, arXiv:2005.12908.

[20] T. Wang et al., “Learning-based stopping power mapping on dual-energy CT for proton radiotherapy,” 2020, arXiv:2003.09058.

[21] Y. Liu, L. Zhou, and H. Peng, “Machine learning based oxygen and carbon concentration derivation using dual-energy CT for PET-based dose verification in proton therapy,” *Med. Phys.*, vol. 49, no. 5, pp. 3347–3360, 2022, doi: 10.1002/mp.15581.

[22] C.-W. Chang et al., “Dual-energy CT based mass density and relative stopping power estimation for proton therapy using physics-informed deep learning,” *Phys. Med. Biol.*, vol. 67, no. 11, 2022, Art. no. 115010, doi: 10.1088/1361-6560/ac6bec,
[68] Z. Li, Y. Wang, Y. Yu, K. Fan, L. Xing, and H. Peng, “Machine learning approaches for range and dose verification in proton therapy using proton-induced positron emitters,” Med. Phys., vol. 46, no. 12, pp. 5748–5757, 2019.

[69] C. Liu, Z. Li, W. Hu, L. Xing, and H. Peng, “Range and dose verification in proton therapy using proton-induced positron emitters and recurrent neural networks (RNNs),” Phys. Med. Biol., vol. 64, no. 17, 2019, Art. no. 175009.

[70] Z. Hu, G. Li, X. Zhang, K. Ye, J. Lu, and H. Peng, “A machine learning framework with anatomical prior for Online dose verification using positron emitters and PET in proton therapy,” Phys. Med. Biol., vol. 65, no. 18, 2020, Art. no. 185003.

[71] S. Ma, Z. Hu, K. Ye, X. Zhang, Y. Wang, and H. Peng, “Feasibility study of patient-specific dose verification in proton therapy utilizing positron emission tomography (PET) and generative adversarial network (GAN),” Med. Phys., vol. 47, no. 10, pp. 5194–5208, 2020.

[72] K. Parodi and T. Bortfeld, “A filtering approach based on Gaussian-powell law convolutions for local PET verification of proton radiotherapy,” Phys. Med. Biol., vol. 51, no. 8, pp. 1991–2009, 2006.

[73] F. Attanasii et al., “Extension and validation of an analytical model for in vivo PET verification of proton therapy—A phantom and clinical study,” Phys. Med. Biol., vol. 56, pp. 5079–5098, Aug. 2011.

[74] S. Remmele, J. Hesser, H. Paganetti, and T. Bortfeld, “A deconvolution approach for PET-based dose reconstruction in proton radiotherapy,” Phys. Med. Biol., vol. 56, no. 23, pp. 7601–7619, 2011.

[75] A. Schumann, M. Prieznitz, S. Schoene, W. Enghardt, H. Rohling, and F. Fiedler, “From prompt gamma distribution to dose: A novel approach combining an evolutionary algorithm and filtering based on Gaussian-powell convolutions,” Phys. Med. Biol., vol. 61, pp. 6919–6934, Oct. 2016.

[76] K. Parodi et al., “Patient study of in vivo verification of beam delivery and range, using positron emission tomography and computed tomography imaging after proton therapy,” Int. J. Radiat. Oncol. Biol. Phys., vol. 68, pp. 920–934, Jul. 2007.

[77] K. Frey et al., “TPS(PET)—A TPS-based approach for in vivo dose verification with PET in proton therapy,” Phys. Med. Biol., vol. 59, no. 1, pp. 1–21, 2014.

[78] T. Hofmann et al., “Dose reconstruction from PET images in carbon ion therapy: A deconvolution approach,” Phys. Med. Biol., vol. 64, no. 2, 2019, Art. no. 250111.

[79] C. Richter et al., “First clinical application of a prompt gamma based in vivo proton range verification system,” Radiother. Oncol., vol. 118, no. 2, pp. 232–237, 2016.

[80] P. Gueth et al., “Machine learning-based patient specific prompt-gamma dose monitoring in proton therapy,” Phys. Med. Biol., vol. 58, no. 13, pp. 4563–4577, 2013.

[81] C.-C. Liu and H.-M. Huang, “A deep learning approach for converting prompt gamma images to proton dose distributions: A Monte Carlo simulation study,” Phys. Med., vol. 69, pp. 110–119, Jan. 2020.

[82] M. Yamaguchi et al., “Secondary-electron-bremsstrahlung imaging for proton therapy,” Nucl. Instrum. Methods Phys. Res. A Accelerat. Spectrometers Detectors Assoc. Equip., vol. 833, pp. 199–207, Oct. 2016.

[83] K. Ando, M. Yamaguchi, S. Yamamoto, T. Toshito, and M. Komori, “Luminescence imaging of water during carbon-ion irradiation during range estimation,” Med. Phys., vol. 42, pp. 6498–6506, Nov. 2015.

[84] S. Yamamoto et al., “Luminescence imaging of water during carbon-ion irradiation for range estimation,” Med. Phys., vol. 43, pp. 2455–2463, May 2016.

[85] T. Yabe, S. Yamamoto, M. Oda, K. Mori, T. Toshito, and T. Akagi, “Prediction of dose distribution from luminescence image of water using a deep convolutional neural network for particle therapy,” Med. Phys., vol. 47, pp. 3882–3891, 2020.

[86] A. Thummerer et al., “Comparison of CBCT based synthetic CT methods suitable for proton dose calculations in adaptive proton therapy,” Phys. Med. Biol., vol. 65, Apr. 2020, Art. no. 95002.

[87] J. J. Sonke, M. Aznar, and C. Rasch, “Adaptive radiotherapy for anatomical changes,” Seminars Radiat. Oncol., vol. 29, no. 3, pp. 245–257, 2019.

[88] F. Albertini, M. Matter, L. Nenoﬀ, Y. Zhang, and A. Lomax, “Online daily adaptive proton therapy,” Brit. J. Radiol., vol. 93, Mar. 2020, Art. no. 20190594.

[89] 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.

[90] K. Araï et al., “Feasibility of CBCT-based proton dose calculation using a histogram-matching algorithm in proton beam therapy,” Physica Medica, vol. 33, pp. 68–76, Jan. 2017.

[91] 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, pp. 549–559, May 2016.

[92] G. Landry et al., “Investigating CT to CBCT image registration for head and neck proton therapy as a tool for daily dose recalculation,” Med. Phys., vol. 42, pp. 1334–1366, Mar. 2015.

[93] Y. K. Park, G. C. Sharp, J. Phillips, and B. A. Wney, “Proton dose calculation on scatter-corrected CBCT image: Feasibility study for adaptive proton therapy,” Med. Phys., vol. 42, pp. 4449–4449, Aug. 2015.

[94] 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.

[95] D. C. Hansen et al., “ScatterNet: A convolutional neural network for cone-beam CT intensity correction,” Med. Phys., vol. 45, pp. 4916–4926, Nov. 2018.

[96] G. Landry et al., “Comparing U-net training with three different datasets to correct CBCT images for prostate radiotherapy dose calculations,” Phys. Med. Biol., vol. 64, Jan. 2019, Art. no. 35011.

[97] A. Thummerer et al., “Comparison of the suitability of CBCT- and MR-based synthetic CTs for daily adaptive proton therapy in head and neck patients,” Phys. Med. Biol., vol. 65, no. 23, 2020, Art. no. 235036.

[98] C. Kurz et al., “CBCT correction using a cycle-consistent generative adversarial network and unpaired training to enable photon and proton dose calculation,” Phys. Med. Biol., vol. 64, Nov. 2019, Art. no. 225004.
S. Devaraj, S. K. Sharma, D. J. Fausto, S. Viernes, and H. Kharrazi

H. S. Grewal, M. S. Chacko, S. Ahmad, and H. Jin, “Prediction of

B. Sun et al., “A machine learning approach to the accurate prediction

G. Elwyn et al., “Many miles to go…,” A systematic review of the

D. Maes et al., “A machine learning-based framework for delivery error

A. Lalonde, B. Winey, J. Verburg, H. Paganetti, and G. C. Sharp,

Y. Nomura, Q. Xu, H. Shirato, S. Shimizu, and L. Xing, “Projection-

L. Chen, X. Liang, X. Liang, and S. Jiang, “Synthetic CT generation

G. Valdes et al., “Clinical decision support of radiotherapy treatment

M. Matter, L. Nenoff, G. Meier, D. C. Weber, A. J. Lomax,

P. Botas, J. Kim, B. Winey, and H. Paganetti “Online adaption

A. M. Kalet, S. M. H. Luk, and M. H. Phillips, “Radiation therapy

X. Zhu, Y. Ge, T. Li, D. Thongphiew, F.-F. Yin, and Q. J. Wu, “A

A. M. Kalet, S. M. H. Luk, and M. H. Phillips, “Radiation therapy

S. Kida et al., “Cone beam computed tomography image quality

S. Kida et al., “Cone beam computed tomography image quality

S. Toscano et al., “Impact of machine log-files uncertainties on the

S. Toscano et al., “Impact of machine log-files uncertainties on the

G. L. Zhang, L. Zhou, Z. Han, W. Zhao, and H. Peng, “SWFT-Net: A

D. Scandurra et al., “Assessing the quality of proton PBS treatment

D. Maa et al., “A machine learning-based framework for delivery error

S. Devraj, S. K. Sharma, D. J. Fausto, S. Viernes, and H. Kharrazi

G. Elwyn et al., “Many miles to go…: A systematic review of the

M. A. Musen, B. Middleton, and R. A. Greenes, “Clinical decision-

M. A. Musen, B. Middleton, and R. A. Greenes, “Clinical decision-

S. Kida et al., “Cone beam computed tomography image quality

X. Zhu, Y. Ge, T. Li, D. Thongphiew, F.-F. Yin, and Q. J. Wu, “A

L. Nenoff et al., “Daily adaptive proton therapy – The key to innova-

J. Thyrza, S. Breedveld, R. van Haveren, B. Heijmen, and

P. Riley, “Three pitfalls to avoid in machine learning,” Nature, vol. 572,

L. Nenoff et al., “Daily adaptive proton therapy – The key to innova-

L. Nenoff et al., “Daily adaptive proton therapy – The key to innova-

D. Scandurra et al., “Assessing the quality of proton PBS treatment

J. Harms et al., “Cone-beam CT-derived relative stopping power map
generation via deep learning using machine and deep learning approach in uni-

L. Chen, X. Liang, X. Liang, and S. Jiang, “Synthetic CT generation

J. Bus. Administr. Res., vol. 3, no. 2, p. 36, 2014.

J. Harms et al., “Cone-beam CT-derived relative stopping power map
generation via deep learning for proton radiotherapy,” Med. Phys.,
vol. 47, pp. 4416–4427, Sep. 2020.

A. Lalande, B. Winey, J. Verburg, H. Paganetti, and G. C. Sharp,
"Evaluation of CBCT scatter correction using deep convolutional neural
networks for head and neck adaptive proton therapy," Phys. Med. Biol.,
vol. 65, no. 24, 2020, Art. no. 245022, doi: 10.1088/1361-6560/ab9fcb.

Y. Nomura, Q. Xu, H. Shirato, S. Shimizu, and L. Xing, “Projection-
domain scatter correction for cone beam computed tomography using
a residual convolutional neural network,” Med. Phys. vol. 46, no. 7,
pp. 3142–3155, 2019.

A. M. Kalet, S. M. H. Luk, and M. H. Phillips, “Radiation therapy
quality assurance tasks and tools: The many roles of machine learning," Med. Phys., vol. 47, pp. e168–e177, Jun. 2020.

X. Zhu, Y. Ge, T. Li, D. Thongphiew, F.-F. Yin, and Q. J. Wu, “A
planning quality evaluation tool for prostate adaptive IMRT based on
machine learning,” Med. Phys., vol. 38, no. 2, pp. 719–726, 2011.

B. Sun et al., “A machine learning approach to the accurate prediction
of monitor units for a compact proton machine,” Med Phys, vol. 45,
pp. 2243–2251, May 2018.

H. S. Grewal, M. S. Chacko, S. Ahmad, and H. Jin, “Prediction of
the output factor using machine and deep learning approach in uni-
form scanning proton therapy,” J. Appl. Clin. Med. Phys., vol. 21,
pp. 128–134, Jul. 2020.

D. Maa et al., “A machine learning-based framework for delivery error
prediction in proton pencil beam scanning using irradiation log-files,” Physica Medica, vol. 78, pp. 179–186, Oct. 2020.

S. Devraj, S. K. Sharma, D. J. Fausto, S. Viernes, and H. Kharrazi
“Barriers and facilitators to clinical decision support systems adoption: A systematic review,” J. Bus. Administr. Res., vol. 3, no. 2, p. 36, 2014.

G. Elwyn et al., “Many miles to go…: A systematic review of the
implementation of patient decision support interventions into routine
clinical practice,” BMC Med. Informat. Decis. Making, vol. 13, p. S14,
Nov. 2013.

M. A. Musen, B. Middleton, and R. A. Greenes, “Clinical decision-
support systems,” in Biomedical Informatics. Cham, Switzerland: Springer, 2014, pp. 643–674.