

Deep learning in MRI-guided radiation therapy: A systematic review

Zach Eidex^{1,2} | Yifu Ding¹ | Jing Wang¹ | Elham Abouei¹ | Richard L. J. Qiu¹  | Tian Liu³ | Tonghe Wang⁴ | Xiaofeng Yang^{1,2}

¹Department of Radiation Oncology and Winship Cancer Institute, Emory University, Atlanta, Georgia, USA

²School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, Georgia, USA

³Department of Radiation Oncology, Icahn School of Medicine at Mount Sinai, New York, New York, USA

⁴Department of Medical Physics, Memorial Sloan Kettering Cancer Center, New York, New York, USA

Correspondence

Xiaofeng Yang, Department of Radiation Oncology, Emory University School of Medicine, 1365 Clifton Road NE, Atlanta, GA 30322, USA.
Email: xiaofeng.yang@emory.edu

Funding information

National Institutes of Health under Award Numbers, Grant/Award Numbers: R01CA215718, R56EB033332, R01EB032680, P30 CA008748

Abstract

Recent advances in MRI-guided radiation therapy (MRgRT) and deep learning techniques encourage fully adaptive radiation therapy (ART), real-time MRI monitoring, and the MRI-only treatment planning workflow. Given the rapid growth and emergence of new state-of-the-art methods in these fields, we systematically review 197 studies written on or before December 31, 2022, and categorize the studies into the areas of image segmentation, image synthesis, radiomics, and real time MRI. Building from the underlying deep learning methods, we discuss their clinical importance and current challenges in facilitating small tumor segmentation, accurate x-ray attenuation information from MRI, tumor characterization and prognosis, and tumor motion tracking. In particular, we highlight the recent trends in deep learning such as the emergence of multi-modal, visual transformer, and diffusion models.

KEYWORDS

deep learning, MRI-guided, radiation therapy, radiotherapy, review

1 | INTRODUCTION

Recent innovations in magnetic resonance imaging (MRI) and deep learning are complementary and hold great promise for improving patient outcomes. With the advent of the Magnetic Resonance Imaging Guided Linear Accelerator (MRI-LINAC) and MR-guided radiation therapy (MRgRT), MRI allows for accurate and real-time delineation of tumors and organs at risk (OARs) that may not be visible with traditional CT-based plans.¹ Deep learning methods augment the capabilities of MRI by reducing acquisition times, generating electron density information crucial to treatment planning, and increasing spatial resolution, contrast, and image quality. In addition, MRI auto-segmentation and dose calculation

methods greatly reduce the required human effort on tedious treatment planning tasks, enabling physicians to further optimize treatment outcomes. Finally, deep learning methods offer a powerful tool in predicting the risk of tumor recurrence and adverse effects. These advancements in MRI and deep learning usher in the era of fully adaptive radiation therapy (ART) and the MRI-only workflow.²

Deep learning methods represent a broad class of neural networks which derive abstract context through millions of sequential connections. While applicable to any imaging modality, these algorithms are especially well suited to MRI due to its high information density.³ Deep learning demonstrates state of the art performance over traditional hand-crafted and machine

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2023 The Authors. *Journal of Applied Clinical Medical Physics* published by Wiley Periodicals, LLC on behalf of The American Association of Physicists in Medicine.

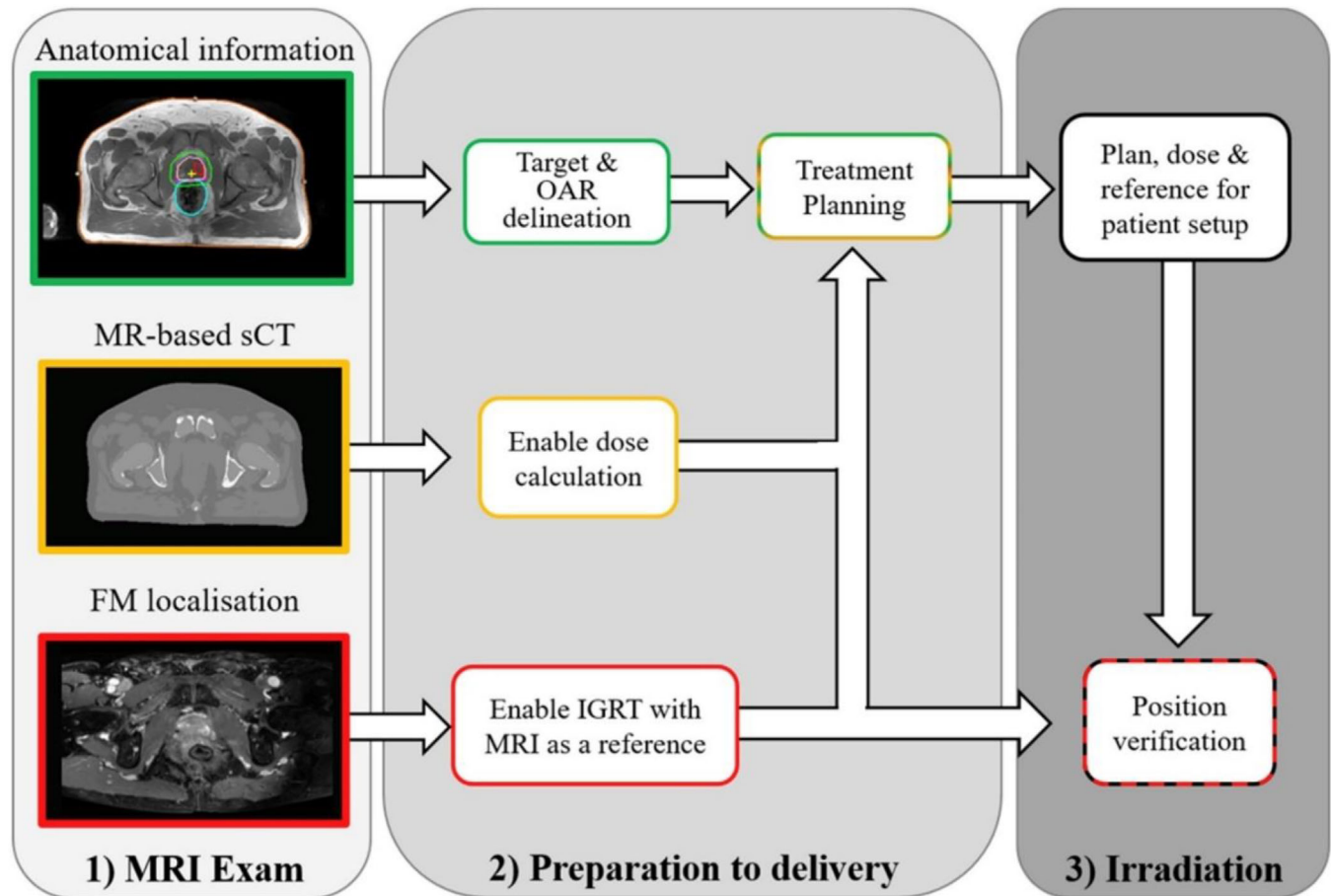


FIGURE 1 An MRI-only workflow for prostate cancer. From top to bottom, a diagnostic MRI is taken to determine the sites of the target volume and organs at risk (OARs) which can be aided by segmentation models and prognostic radiomics models. Simultaneously, sCT enables x-ray attenuation information. Finally, fiducial markers (FMs) are identified to define the prostate position which can be monitored during treatment with real time MRI. Reprinted by permission from Elsevier: Clinical Oncology, Magnetic Resonance Imaging only Workflow for Radiotherapy Simulation and Planning in Prostate Cancer by Kerkmeijer et al. 2018.¹⁶⁰

learning methods but are computationally intensive and require large datasets. For MRI and other imaging tasks, convolutional neural networks (CNNs), built on local context, have traditionally dominated the field. However, advancements in network architecture, availability of more powerful computers, large high-quality datasets, and increased academic interest have led to rapid innovation. Especially exciting are the rapid adaptation of cutting-edge transformer and generative methods, which utilize data from multiple input modalities.

Deep learning techniques can be organized according to their applications in MRgRT in the following groups: segmentation, synthesis, radiomics (classification), and real-time/4D MRI. Shown in Figure 1 is an example of all of these groups working together for an MRI-only workflow for prostate cancer. Segmentation methods automatically delineate tumors, organs at risk (OARs), and other structures. However, deep learning approaches face challenges when adapting to small tumors, multiple organs, low contrast, and differing ground truth contour quality and style. These

challenges differ greatly depending on the region of the body, so segmentation methods are primarily organized by anatomical region.⁴

Synthesis methods are best understood by their input and output modalities. Going from MRI to CT, synthetic CT (sCT) provides accurate attenuation information not apparent in MRI, augmenting the information of co-registered CT images. In an MRI-only workflow, sCT avoids registration errors and the radiation exposure associated with traditional CT.⁵ In addition, synthetic relative proton stopping power (sRPSP) maps can be generated to directly obtain dosimetric information for proton radiation therapy.⁶ The dosimetric uncertainty can be further enhanced with deep learning dose calculation methods, which greatly reduce inference time and could yield lower dosimetric uncertainties compared to traditional Monte Carlo (MC) methods. Synthetic MRI (sMRI), generated from CT, is appealing by combining the speed and dosimetric information of CT with MRI's high soft tissue contrast. However, CT's lower soft tissue contrast makes this application much more challenging,

but sMRI has still found success in improving CT-based segmentation accuracy.^{7–9} Alternatively, there are rich intra-modal applications by generating one MRI sequence from another. For example, the spatial resolution of clinical MRI can be increased by predicting a higher resolution image^{10,11} and applying contrast can be avoided with synthetic contrast MRI.¹²

Radiomics represents an eclectic body of works but can be divided into studies which classify structures in an MRI image¹³ or prognostic models which use MR images to predict treatment outcomes such as tumor recurrence or adverse effects.^{14,15} Deep learning methods in real-time and 4D MRI overcome MRI's long acquisition time and the low field strengths of the MRI-LINAC by reconstructing images from undersampled k-space,¹⁶ synthesizing additional MRI slices,¹⁷ and exploiting periodic motion to improve image quality.¹⁸

In this review, we systematically examine studies that apply deep learning to MRgRT, categorizing them based on their application and highlighting interesting or important contributions. We identify four distinct areas of deep learning methods which enhance the clinical workflow: segmentation, synthesis, radiomics (classification), and real-time/4D MRI. For each category, the sections are ordered as follows: foundational deep learning methods, challenges specific to MRI, commonly used evaluation metrics, and finally subcategories including a brief overview followed by interesting or influential studies in that subcategory. Since deep learning methods build on each other, the unfamiliar reader is encouraged to read the explanations on deep learning methods sequentially. In Section 7, we discuss current trends in deep learning architectures and how they may benefit new clinical techniques and MRI technologies like the MRI-LINAC and higher strength MRI scanners.

2 | LITERATURE SEARCH

This systematic review surveys literature which implements deep learning methods and MRI for radiation therapy research. “Deep learning” is defined to be any method which includes a neural network directly or indirectly. These include machine learning models and other hybrid architectures which take deep learning derived features as input. Studies including MRI as at least part of the dataset are included. Studies must list their purpose as being for radiation therapy and include patients with tumors. Studies on immunotherapy and chemotherapy without radiation therapy are excluded. Conference abstracts and proceedings are excluded due to an absence of strict peer review.

The literature search was performed on PubMed on December 31, 2022, with the following search criteria in the title or abstract: “deep learning and (MRI or MR) and radiation therapy” and is displayed in Table 1. This search yielded 335 results. Of these results,

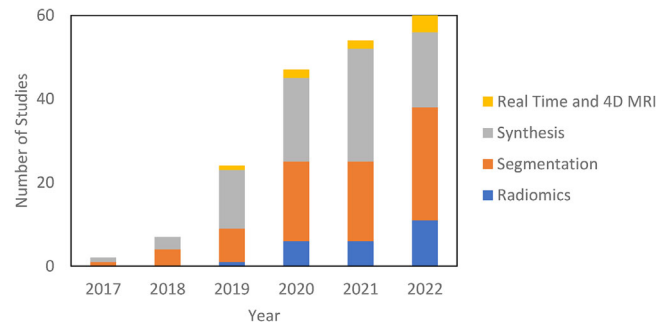


FIGURE 2 Number of deep learning studies with applications towards MRgRT per year by category including references 161–277 in [Supporting Information](#).

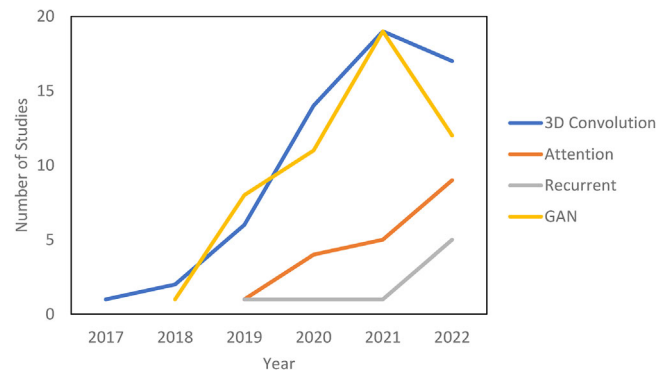


FIGURE 3 Technical trends in deep learning.

197 were included based on manual screening using the aforementioned criteria. Seventy-eight were classified as segmentation, 81 as synthesis, 24 as radiomics (classification), and 14 as real-time or 4D MRI. There is inevitably some overlap in these categories. In particular, studies which use sMRI for the purposes of segmentation are classified as synthesis and papers which deal with real-time or 4D MRI are placed in Section 6: Real-Time and 4D MRI. Figure 2 shows the papers sorted by category and year. Compared to other review papers, this review paper is more comprehensive in its literature search and is the first specifically on the topic of deep learning in MRgRT. In addition, this work uniquely focuses on the underlying deep learning methods as opposed to their results. Figure 3 shows technical trends in deep learning methods implementing 3D convolution, attention, recurrent, and GAN techniques.

3 | IMAGE SEGMENTATION

Contouring (segmentation) in MRgRT is the task of delineating targets of interest on MR images, which can be broadly divided into distinct categories: contouring of organs at risk and other anatomical structures expected to receive radiation dose and contouring of individual tumors. Both tumor and multi-organ

segmentation suffer from intra- and inter- observer variability.¹⁹ MRI does not capture the true extent of the tumor volume, as well as poorly defined boundaries and similar structures like calcifications lead to institutional and intra-observer variability. Physician contouring conventions and styles further complicate the segmentation task and lead to inter-observer variability.^{20,21} Multi-organ segmentation is mostly challenged by the large number of axial slices and OARs which make the task tedious and prone to error. Automated solutions to MRI segmentation have been proposed to reduce physician-workload and provide expert-like performance. In this section, we first review foundational deep learning methods starting with CNN and moving into recurrent and transformer architectures.

3.1 | Deep learning methods

Since the application of CNNs to MRI-based segmentation in 2017²², fully convolutional networks (FCNs) have outperformed competing atlas-based and hand-crafted auto-segmentation methods, often matching the intra-observer variability among physicians.²³ FCNs employ convolutional layers which are trained to detect patterns in either nearby voxels or feature maps output from previous convolutional layers. In contrast with traditional CNNs, FCNs forgo densely connected layers. This design choice enables voxel-wise segmentation, allows for variable sized images, and reduces model complexity and training time. Different types of convolutions include atrous and separable convolutions. Atrous convolutions sample more sparsely to gain a wider field of view and can be mix-and-matched to capture large and small features in the same layer. Separable convolutions divide a 2D convolution into two 1D convolutions to use fewer parameters for similar results. By connecting multiple convolutional layers together with non-linear activation functions, larger and more abstract regions of the input image are analyzed to form the encoder. For pixelwise segmentation, the final feature map is expanded to the original image resolution through a corresponding series of transposed convolutional layers forming the decoder. All FCNs include pooling layers to conserve computational resources whereby the resolution of feature maps is reduced by choosing the largest (max-pooling) or average local pixel.²⁴

Advances from the field of natural language processing (NLP) have had a tremendous impact on segmentation tasks. Recurrent neural networks (RNNs) are defined by the output of their node being connected to the input of their node. To avoid an infinite loop, the output is only allowed to connect to its input a set number of times. This property allows for increased context and the ability to handle sequential data which is especially important in language translation. Applied to CNNs, each recurrent convolutional layer (convolution + activa-

tion function) is performed multiple times which creates a wider field of view and more context with each subsequent convolution. However, recurrent layers can suffer from a vanishing gradient problem. Long short-term memory blocks (LSTM) solve this by adding a “forget” gate, which forgets irrelevant information. In addition, LSTMs are more capable of making long range connections. Similar to the LSTM gate, the gated recurrent unit (GRU) has an update and reset gate which decide which information to pass on and which to forget. Both LSTM and GRU also have bidirectional versions which pass information forward and backwards.^{25,26} Relative performance between the LSTM and GRU gates are situational with the GRU gate being less computationally expensive.²⁷

Related developments from NLP are the concepts of attention and the transformer. In terms of MRI, attention is the idea that certain regions of the MRI volume are more important to the segmentation task and should have more resources allocated to them. ROI schemes can then be defined as a form of hard attention by only considering the region around a tumor. A version of soft attention would weight the region around the tumor heavily and process the information in high resolution but also give a smaller weighting to nearby organs and process it in lower resolution.²⁸ In practice, attention modules include a fully connected feedforward neural network to generate weights between a feature map of the encoder and a shallower feature map in the decoder. These weights are improved upon through backpropagation of the entire network to give higher representational power to contextually significant areas of the image. This fully connected network can also be replaced with other models such as the RNN, GRU, or LSTM.²⁹ If the same feature map is compared with itself, this is called self-attention and is the basis for the transformer architecture.³⁰ The transformer can be thought of as a global generalization of the convolution and can even replace convolutional layers. The advantages of the transformer are explicit long-range context and the transformer’s multi-head attention block allows for attention to be focused on different structures in parallel. However, transformers require more data to train and can be very computationally expensive. Such computational complexity can be remedied by including convolutional layers in hybrid CNN-transformer architectures,³¹ by making long range connections between voxels sparse,³² or by implementing more efficient self-attention models like FlashAttention.³³

From the field of neuroscience, deep spiking neural networks (DSNNs) attempt to more closely model biological neurons by connecting neurons with asynchronous time dependent spikes instead of the continuous connections between neurons of traditional neural networks. Potential advantages include lower power use, real-time unsupervised learning, and

new learning methods. However, these advantages are only fully realized with special neuromorphic hardware, are difficult to train, and currently lag conventional approaches. For these reasons, they are currently only represented by one paper in this review.³⁴

3.2 | Challenges in MRI

The properties of MRI datasets have driven innovation. Multiple MRI sequences, with and without contrast, are often available. To capture all data, the different sequences are co-registered and input as multiple channels yielding multiple segmentations. These segmentations are combined to produce a final segmentation using an average, weighted average, or more advanced method. To account for MRI's high through-plane resolution relative to its in-plane resolution, 3D convolutional layers are often utilized to capture features not apparent in 2D convolution. However, 3D convolutions are computationally expensive, so numerous 2.5D architectures have been proposed.^{35–37} In a 2.5D architecture, adjacent MRI slices are input as channels, and 2D convolutions are performed. It is also common to see new papers forgo the 3D convolution to save resources for new computationally intense methods. An unfortunate fact is that high-quality MRI datasets are often small. To remedy this, data augmentation methods such as rotating and flipping the MR images are ubiquitous. In addition, the generation of synthetic images to increase dataset size and generalizability is an exciting field of research.³⁸ Public datasets and competitions have also helped in this regard. For example, the Brain Tumor Segmentation Challenge (BraTS) dataset,³⁹ updated since 2012, has been a primary contributor to brain segmentation progress, spawning the popular DeepMedic framework.⁴⁰ Another approach for small datasets is transfer learning. In transfer learning, a model is trained on a large dataset, and then retrained on a smaller dataset with the idea that many of the previously found features are transferable.⁴¹

A major issue faced in MRI-segmentation can be characterized as “the small tumor problem”. Small structures like tumors or brachytherapy fiducial markers represent a small fraction of the total MRI volume, where CNNs can struggle to find them or be confused by noise. Further exacerbating the problem is that applying a deep CNN to whole MR images consumes extensive computational resources, so the MRI must be downsampled. In this case, the down sampling is very likely to cause small tumors to be missed entirely. One of the simplest ways to improve performance is to alter the loss function. Standard loss functions are cross-entropy and dice loss which seek to maximize voxel wise classification accuracy and overlap between the predicted and ground truth contours, respectively. These can be modified to achieve higher sensitivity to small structures at the

expense of accuracy. Focal loss is the cross-entropy loss modified for increased sensitivity⁴² and Tversky loss does the same for the dice loss.⁴³ In addition, borders of the contours are the most important part of the segmentation, so boundary loss functions seek to improve model performance by placing increased emphasis on regions near the contour edge.^{44,45} Another approach to solve the problem, albeit at the expense of long-range context, is with two stage networks. In the first stage, regions of interest (ROIs) are identified, and target structures are then contoured in the ROIs in the second stage. Notable efforts include Mask R-CNN⁴⁶ and Retina U-Net⁴⁷ which implement convolution-based ROI sub-networks with advanced correction algorithms. Seqseg instead replaces the correction algorithms with a reinforcement learning based model.⁴⁵ An agent is guided by a reward function to iteratively improve the conformity of the bounding box. Seqseg reported comparable performance with higher bounding box recall and intersection over union (IoU) compared to Mask R-CNN.

Many new models for MRI segmentation have been created by modifying U-Net. U-Net derives its name from its shape which features convolutional layers in the encoder and transposed convolutional layers in the decoder. Its main innovation, however, is its long-range skip connections between the encoder and decoder. Dense U-Net densely connects convolutional layers in blocks,⁴⁸ ResU-Net includes residual connections,⁴⁹ Retina U-Net is a two-stage network, RU-Net includes recurrent connections, R2U-Net adds residual recurrent connections.⁵⁰ Attention modules have also been added at the skip connections.^{51,52} Both V-Net⁵³ and nnUNet⁵⁴ were designed with 3D convolutional layers with nnUNet additionally automating preprocessing and learning parameter optimization. Pix2pix uses U-Net as the generator with a convolutional discriminator (PatchGAN).⁵⁵ Other state-of-the-art architectures include Mask R-CNN, DeepMedic, and DeepLabV3+.⁵⁶ Mask R-CNN is a two-stage network with a ResNet backbone. Mask Scoring RCNN (MS-RCNN) improves upon Mask R-CNN by adding a module which penalizes ROIs with high classification accuracy but low segmentation performance.⁵⁷ DeepMedic, designed for brain tumor segmentation, is an encoder-only CNN which inputs a ROI and features two independent row-resolution and normal resolution channels. These channels are joined in a fully connected convolutional layer to predict the final segmentation. The convolutions in the encoder-only style reduce the final segmentation map dimensions compared to the original ROI ($25 \times 25 \times 25$ vs. $9 \times 9 \times 9$ voxels). DeepLabV3+ leverages residual connections and multiple separable atrous convolutions. Xception improves upon the separable convolution by reversing the order of the convolutions and including ReLU blocks after each operation for non-linearity.⁵⁸

3.3 | Evaluation metrics

To evaluate performance, various evaluation metrics are employed with the Dice similarity coefficient (DSC) being the most prevalent. The DSC is defined in Equation 1 as the overlap between the ground truth physician contours and the predicted algorithmic volumes with a value of 0 corresponding to no overlap and 1 corresponding to complete overlap. Mathematically, it is defined as follows where VOLGT is the ground truth volume and VOLPT is the predicted volume⁵⁹:

$$DSC = \frac{2|VOLGT \cap VOLPT|}{|VOLGT| + |VOLPT|} \quad (1)$$

Additional metrics include the Hausdorff distance⁵⁹ which measures the farthest distance between two points of the ground truth and algorithmic volumes, volume difference,⁶⁰ which is simply the difference in volumes, and the Jaccard Index,⁶¹ which is similar to the DSC and measures the overlap between VOLPT and VOLGT relative to their combined volumes. A discussion of these metrics is found in Müller et al.⁶² However, performance between datasets must be evaluated with caution due to high inter-observer variation between physicians and dataset quality.

3.4 | Brain

Largely unaffected by patient motion and comprised of detailed soft tissue structures, the brain is an ideal site to benchmark segmentation performance for MRI and represents the dominant category in MRI segmentation research. Unique to brain MRI preprocessing is skull stripping, where the skull and other non-brain tissue are removed from the image. This can significantly improve results, especially for networks with limited training data.⁶³ Shown in Table 2, the majority of the studies focus on segmenting different brain tumors such as glioma, Glioblastoma Multiforme (GBM), and metastases. A small minority of studies focuses on OARs like the hippocampus. Advancements in brain segmentation have come, in large part, from the yearly Multimodal Brain Tumor Image Segmentation Benchmark (BraTS) challenge, which includes high quality T1-weighted (T1W), T2-weighted (T2W), T1-contrast (T1C), and T2 -Fluid-Attenuated Inversion Recovery (FLAIR) sequences with the purpose of segmenting the whole tumor (WT), tumor core (TC), and enhancing tumor (ET) volumes. The WT is defined as the entire spread of the tumor visible on MRI; The ET is the inner core which shows significant contrast compared to healthy brain tissue, and the TC is the entire core including low contrast tissue. The most popular architectures are DeepMedic, created for the BraTS challenge, and U-Net.

Notable efforts in the BraTS challenge include Momin *et al* achieving an exceptional WT dice score of 0.97 ± 0.03 with a Retina U-Net based model and mutual enhancement strategy. In their model, Retina U-Net finds a ROI and segments the tumor. This feature map is fed into the classification localization map (CLM module) which further classifies the tumor into subregions. The CLM shares the encoding path with a segmentation module, so classification and segmentation share information and are improved iteratively.⁶⁴ Huang et al. focuses on correctly segmenting small tumors. Based on DeepMedic, the method incorporates a prior scan and custom loss function, the volume-level sensitivity–specificity (VSS), which rates and significantly improves the metastasis sensitivity and specificity to segment small brain metastases.⁶⁵ Another paper improves small tumor detection by 2.5 times compared to the standard dice loss by assigning a higher weight to small tumors.⁶⁶ Both Tian et al.⁶⁷ and Ghaffari et al.⁶⁸ utilize transfer learning datasets to cope with limited data. Ahmadi et al. achieves competitive results in the BraTS challenge with a DSNN.³⁴

3.5 | Head and neck

The head and neck (H&N) region contains many small structures, making high-resolution and high-contrast imaging of great importance. MRI is especially preferred over CT imaging for patients with amalgam dental fillings due to the metallic content that can cause intense streaking artifacts on CT.⁶⁹ In addition, MRI is the standard of care for nasopharyngeal carcinoma (NPC), leading to significant research attention on auto-segmentation algorithms for H&N MR images (Table 3). Other research efforts include segmentation of oropharyngeal cancer, glands, and lymph nodes in the American Association of Physicists in Medicine (AAPM)'s RT-MAC challenge,⁷⁰ as well as multi-organ segmentation.

Notable efforts include the two-stage multi-channel Seqseg architecture for NPC segmentation.⁷¹ Seqseg uses reinforcement learning to refine the position of the bounding box, implements residual blocks, recurrent channel and region-wise attention, and a custom loss function that emphasizes segmentation of the edges of the tumor. Outier et al.⁷² improves the dice score by 0.10 with a two-stage approach compared to single-state 3D U-Net for oropharyngeal cancer segmentation. For multiparametric MRI (mp-MRI), Deng et al.⁷³ concludes that the union output from T1W and T2W sequences has similar performance to T1C MRI, suggesting that contrast may not be necessary for NPC segmentation. Similarly, Wahid et al.⁷⁴ finds that T1W and T2W sequences significantly improve performance, but dynamic contrast-enhanced MRI (DCE) and diffusion-weighted imaging (DWI) have little effect. The

first stage segments the OARs in low resolution to create a bounding box, followed by U-Net segmenting the ROI in high resolution. Jiang et al. segments the parotid glands using T2W MRI and unpaired CT images with ground truth contours. First, sMRI is generated from the CT volumes using a GAN. In the second step, U-Net generates probabilistic segmentation maps for both the sMRI and MRI based on the CT ground truth contours. These maps, along with sMRI and MRI data, are then input into the organ attention discriminator, which is designed to learn finer details during training, ultimately producing the final segmentations.⁷⁵

3.6 | Abdomen, heart, and lung

In contrast to the brain, the abdomen is susceptible to respiratory and digestive motion of the patient often leading to poorly defined boundaries. While motion management techniques like patient breath-hold and not eating or drinking before treatment can mitigate these effects, the long acquisition time of MRI will inevitably lead to errors. Often physicians must rely on anatomical knowledge to deduce the boundaries of OARs. This makes segmentation challenging for CNN-based architectures, which build from local context. In addition, registration errors make including multiple sequences impractical. OARs segmented in the abdomen include the liver, kidneys, stomach, bowel, and duodenum. The liver and kidneys are not associated with digestion and are relatively stable while the stomach, bowel, and duodenum are considered unstable. The duodenum is the most difficult for segmentation algorithms due to its small size, low contrast, and variability in shape. In addition, radiation induced duodenal toxicity is often dose-limiting in dose escalation studies making accurate segmentation of high importance.⁷⁶ Similar problems occur in the heart and lung because of their periodic motion with the lung being particularly challenging since it is filled with low-signal air. However, MR segmentation of cardiac subregions have shown growing interest as these are not visible on CT and have different tolerances to radiation.⁷⁷

The results are summarized in Table 4. Due to the large number of organs segmented in several of these studies, only the stomach and duodenum dice scores are reported to establish how the algorithms handle unstable organs. Zhang et al.⁷⁸ generates a composite image from the current slice, prior slice, and contour map to predict the current segmentation with U-Net. Luximon et al.⁷⁸ takes a similar approach by having a physician contour every 8th slice. These contours are then linearly interpolated and improved upon with a 2D Dense U-Net. The remaining studies do not require previous information and struggle to segment the duodenum. Ding et al.⁷⁹ improves upon a physician-defined acceptable contour rate by up to 39% with an active contour model. Mor-

ris et al. segments heart substructures with a 2 channel 3D U-Net.⁸⁰ Wang et al. segments lung tumors with high accuracy relying on segmentation maps from previous weeks with the aim of adaptive radiation therapy (ART).⁸¹ An addition study by the same group feeds the features from the CNN into a GRU based RNN to predict tumor position over the next 3 weeks. Attention is included to weigh the importance of the prior weeks' segmentation maps.⁸²

3.7 | Pelvis

The anatomy of the pelvis allows both external beam radiation therapy (EBRT) and brachytherapy approaches for radiation therapy. Therefore, MRI segmentation studies have proposed methods to contour fiducial markers and catheters for cervical and prostate therapy, as well as tumors and OARs. However, a current challenge is that fiducials and catheters are designed for CT and are not optimal for MRI segmentation. For example, in prostate EBRT, gold fiducial markers localize the prostate with high contrast and correct for motion. However, metal does not emit a strong signal on MRI, so fiducials on MRI are characterized by an absence of signal, which can be confused with calcifications. Despite this, MRI is enabling treatments with higher tumor conformality. For instance, the gross tumor volume (GTV) of prostate cancer is not well delineated on CT but is often visible on MRI. In addition, the prostate apex is significantly clearer on MRI.⁸³ MRI-based focal boost radiation therapy, in addition to a single dose level to the whole prostate, escalates additional dose to the GTV to reduce tumor recurrence.^{84,85}

Table 5 shows relevant auto-segmentation techniques applied to the pelvic region. Shaaer et al.⁸⁶ segments catheters with a T1W and T2W MRI-based U-Net model and takes advantage of catheter continuity to refine the contours in post processing. Zabihollahy et al.⁸⁷ creates an uncertainty map of cervical tumors by retraining the U-Net model with a randomly set dropout layer. This technique is called Monte Carlo Dropout (MCDO). Cao et al.²³ takes pre-implant MRI and post-implant CT as input channels to their network. After performing intra-observer variability analysis, they achieve performance more similar to a specialist radiation oncologist for cervical tumors in brachytherapy than a non-specialist. Eidex et al.⁶¹ segments dominant intraprostatic lesions (DILs) and the prostate for focal boost radiation therapy with a Mask R-CNN based architecture. Sensitivity is found to be an important factor in evaluating model performance because weak models can appear strong by missing difficult lesions entirely. Figure 4 shows an example of automatic contours of the prostate and DIL on T2w MRI which would not be visible on CT. STRAINet⁸⁸ realizes exceptional performance by utilizing a GAN with stochastic residual

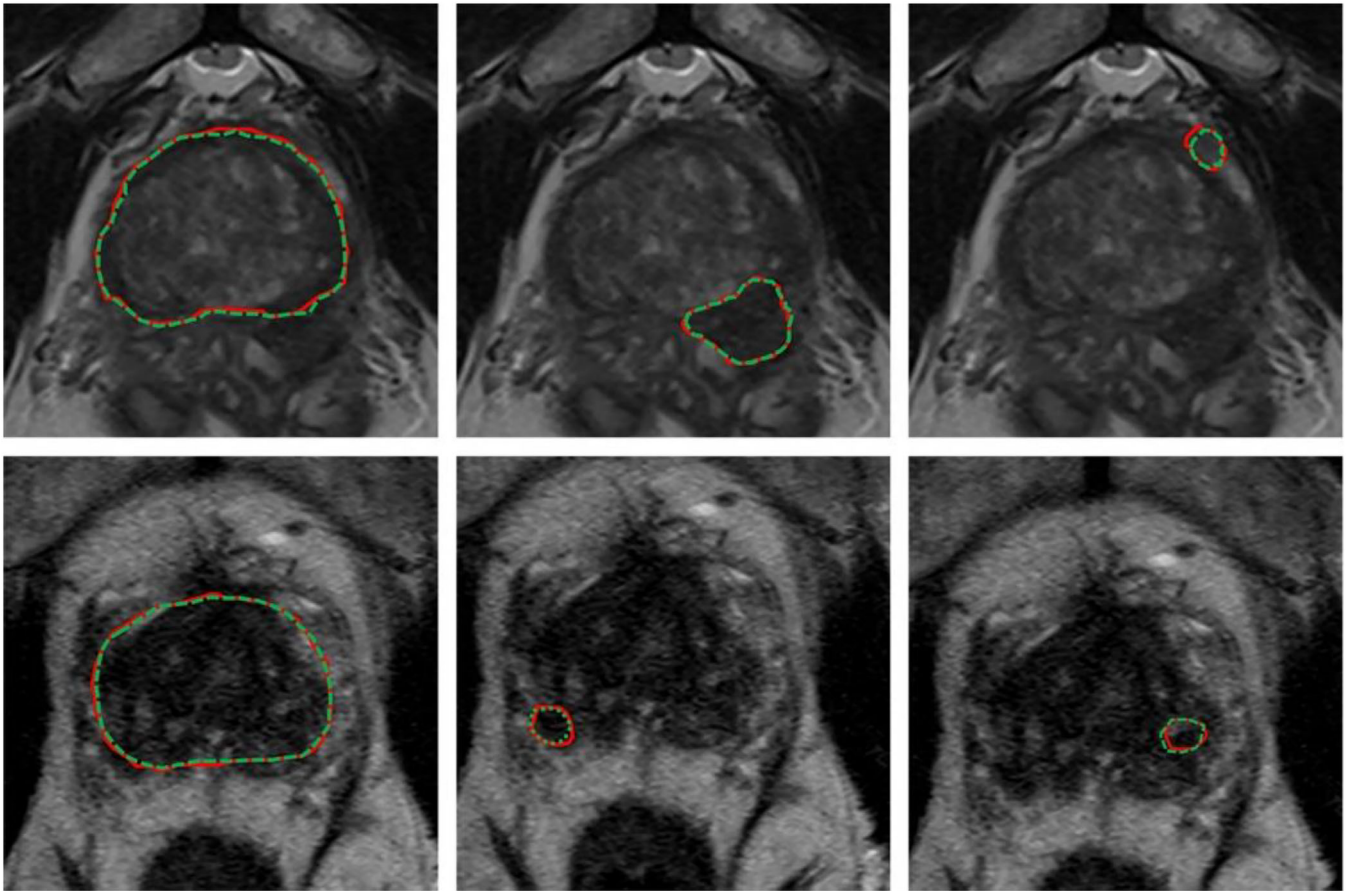


FIGURE 4 Expert (red) versus proposed auto-segmented (green dashed) prostate and DIL contours on axial MRI. From left to right: prostate manual and auto-segmented contours overlaid on MRI, and two DIL manual and auto-segmented contours overlaid on MRI. The upper and lower rows are representative of two patients. Reprinted by permission from John Wiley and Sons: Medical Physics, MRI-based prostate and dominant lesion segmentation using cascaded scoring convolutional neural network by Eidex et al.⁶¹ © 2022.

and atrous convolutions. In contrast with standard residual connections, each element of the input feature map which does not undergo convolution has a 1% chance of being set to zero. Singhrao et al.⁸⁹ implements a pix2pix architecture for fiducial detection achieving 96% detection with the misses caused by calcifications.

4 | IMAGE SYNTHESIS

Image synthesis is an exciting field of research, defined as translating one imaging modality into another. Benefits of synthesis include avoiding potential artifacts, reducing patient cost and discomfort, and avoiding radiation exposure.⁹⁰ In addition, utilizing multiple modalities introduces registration errors which can be avoided with synthetic images. Current methods in MRgRT include synthesis of sCT from MRI, sMRI from CT, and relative proton stopping power images from MRI. Other areas of synthesis research include creating higher resolution MRI (super-resolution) and predicting organ displacement based on periodic motion in 4D MRI. Segmentation can also be thought of as a special case of synthe-

sis because the input MRI is translated into voxel-wise masks which assume discrete values according to their class. The distinction between synthesis and segmentation is particularly muddled when the segmentation ground truth is from a different imaging modality.⁹¹

4.1 | Generative models

Synthesis architectures are fundamentally interchangeable with segmentation architectures but have diverged in practice. For example, U-Net, described in detail in Section 3, is the predominant backbone in both areas. However, synthesis models require that the entire image be translated, so that they do not include two-stage architectures and are dominated by generational adversarial network (GAN)-based architectures. The GAN is comprised of a CNN or self-attention-based generator which generates synthetic images. The generator competes with a discriminator which attempts to correctly classify synthetic and real images. As the GAN trains, a loss function is applied to the discriminator when it mislabels the image, whereas a loss function is applied to the

generator when the discriminator is correct. The model is ideally considered trained once the discriminator can no longer correctly identify the synthetic images. Conditional GANs (cGANs) expand on the standard GAN by also inputting a vector with random values or additional information into both the generator and discriminator.⁹² In the case of MRI, the values of the vector can correspond to the MRI sequence type and clinical data to account for differences in patient population and setup. The CycleGAN adds an additional discriminator and generator loop.⁹³ For example, an MRI would be translated into a sCT. The sCT would then be translated into a sMRI. Since the input is ultimately tested against itself, this allows for training with unpaired data. The need for co-registration is eliminated but requires significantly more data to achieve comparable results with paired training.

Despite their success, GANs can at times be unstable during training and may encounter difficulties with complex synthesis problems. One approach to enhancing training performance and stability is the implementation of the Wasserstein GAN (WGAN). Contrary to the discriminator in traditional GANs, which classifies images as either real or fake, the WGAN evaluates the probability distributions of the real and synthetic images and calculates the Wasserstein distance, or the distance between these distributions. The discriminator strives to maximize this distance while the generator endeavors to minimize it. Although not exclusive to WGANs, spectral normalization is frequently incorporated to constrain the training weights of the discriminator, thereby preventing gradient explosion.⁹⁴ The Wasserstein GAN with Gradient Penalty (WGAN-GP) further amends the WGAN by adding a gradient penalty to the loss function, which helps to stabilize training and improve the model's performance.⁹⁵ Another innovative approach that claims superior performance to the WGAN is the Relativistic GAN (RGAN). The RGAN postulates that the generator should not only increase the likelihood of synthetic images appearing realistic but also enhance the probability that real images appear fake to the discriminator. Absent this condition, in the late stages of training with a well-trained generator, the discriminator may conclude that every image it encounters is real, contradicting the a priori knowledge that half of the images are synthetic. A standard GAN can be converted to a RGAN by modifying its loss function.⁹⁶

Diffusion models, another approach to synthesis, are more recent entrants to the field. These models gradually add and remove noise from the image to better learn the latent space. The key advantage of diffusion models is increased stability during training. Despite these advantages, diffusion models are computationally intensive since the noise is added in small steps and the model must learn to reconstruct the images at varying noise levels.⁹⁷ Given their promising performance and stability, diffusion models represent an

exciting avenue for future exploration in MRgRT applications and achieve state-of-the-art performance in many computer vision tasks.

4.2 | Evaluation metrics

To evaluate image synthesis performance, various metrics are used to compare voxel values between the ground truth and synthesized volume. The most common metric is the mean absolute error (MAE),^{98,99} which is reported in Tables 6 and 7 if available. The MAE is defined below in Equation 2, where x_i and y_i are the corresponding voxel values of the ground truth and synthesized volume, respectively, and n is the number of voxels.

$$MAE = \sum_{i=1}^n \frac{|y_i - x_i|}{n} \quad (2)$$

For sCT studies, the MAE is typically reported in Hounsfield units (HU) but can also be dimensionless if reported with normalized units. In contrast, MRI intensity is only relative and not in definitive units like CT, so MAE is less clinically meaningful than other metrics for MRI synthesis studies. Therefore, peak signal to noise ratio (PSNR) is preferentially reported.¹⁰⁰ Other common metrics in literature are the mean error,¹⁰¹ which forgoes the absolute value in MAE, the mean squared error (MSE),¹⁰² which substitutes absolute value for the square, and the Structural Similarity Index (SSIM), which varies from -1 to 1 where -1 represents extremely dissimilar images and 1 represents identical images.¹⁰³ A full discussion of these metrics can be found in Necasova et al.¹⁰⁴ Since sCT is primarily intended for treatment planning, dosimetric quantities which measure the deviation between CT- and sCT-derived plans are often reported. One of the most common metrics is gamma analysis. Repurposed as a metric to compare treatment plan dose to actual dose on LINACs, gamma analysis looks at each point on the dose distribution and evaluates if the acceptance criteria are met. The American Association of Physicists in Medicine (AAPM) Task Group 119 recommends a low dose threshold of 10%, meaning that points, which receive less than 10% of the maximum dose are excluded from the calculation. Other metrics include the mean dose difference and the minimum dose delivered to 95% of the clinical treatment volume (D95) difference.

4.3 | MRI-based synthetic CT

MRI-based sCT is the most extensively researched and influential application of synthesis models in radiation therapy. While MR images provide excellent soft tissue

contrast, they do not contain the necessary attenuation information for dose calculation that is embedded in CT images. Owing to this limitation, CT has traditionally been the workhorse for treatment planning while MRI has been relegated to diagnostic applications. However, CT suffers from lower soft tissue contrast and imparts a non-negligible radiation dose, especially for patients receiving standard fractionated image guided radiation therapy (IGRT). In addition, metallic materials found in dental work and implants can lead to severe artifacts in CT, reducing the quality of the treatment plan. By augmenting CT with sCT, these problems can be avoided. Furthermore, according to the “As Low As Reasonably Achievable” (ALARA) principle, the replacement of CT with sCT for an MRI only workflow could be justified with its high accuracy, especially in radiosensitive populations like pediatric patients.^{105,106}

The primary challenge to sCT methods is the accurate reconstruction of bone and air, due to their low proton density and weak signal. This can make it difficult for sCT to distinguish between the two, leading to large errors. In addition, further complicating the issue is that bone makes up a small fraction of the patient volume in radiation therapy tasks or applications which is similar to the “small tumor problem” seen in segmentation. Other issues that can arise include small training sets, misalignment between CT and MRI, and causes of high imaging variability such as intestinal gas.

Calculation of dose distribution using MRI-based sCT can be enhanced by replacing traditional Monte Carlo simulation (MC) techniques with deep learning. MC accurately predicts the dose distribution based on physical principles, including the electron return effect (ERE), which adds additional dose to boundaries with different proton densities in the presence of a magnetic field. However, the technique can be extremely slow, as it relies on randomly generating paths of tens of thousands of particles. The higher number of particles reduces dosimetric uncertainty. This problem is particularly noticeable in proton therapy, where MC or pencil beam algorithm (PBA) calculations can take several minutes on a CPU, and it can take hours to optimize a single treatment plan.¹⁰⁷ As a result, compromises must be made in clinical practice between dosimetric uncertainty, MC run time, and treatment plan optimization. Deep learning methods show exceptional potential to improve upon MC dose calculation models. Once trained, deep learning algorithms take only a few seconds to synthesize a dose distribution. In addition, they can be trained on extremely high accuracy MC generated dose distributions that would be impractical in everyday clinical practice.

Sampling notable MRI-based sCT works for photon radiation therapy, several take advantage of cGANs to include additional information. Liu et al. improve upon the CycleGAN by including a dense block, which captures structural and textural information and better

handles local mismatches between MRI and ground truth CT images. In addition, a compound loss function with adversarial and distance losses improves boundary sharpness. An example patient is shown in Figure 5.¹⁰⁸ A conditional CycleGAN in Boni et al. passes in MR manufacturer information and achieves good results despite using unpaired data and different centers for their training and test sets.¹⁰⁹ Many studies also experiment with multiple sequences. Massa et al. train a U-Net with Inception-V3 blocks on 1.5T T1W, T2W, T1C, and FLAIR sequences separately and finds no statistical difference.¹¹⁰ However, Koike et al. use multiple MR sequences for sCT generation employing a cGAN to provide better image quality and dose distribution results compared with those from only a single T1W sequence.¹¹¹ Dinkla et al. find that sCT removes dental artifacts.¹¹² Reaungamornrat et al. decompose features into modality specific and modality invariant spaces between high- and low-resolution Dixon MRI with the Huber distance. In addition, separable convolutions are used to reduce parameters, and a relativistic loss function is applied to improve training stability.¹¹³ Finally, Zhao et al. represent the first MRI-based sCT paper to implement a hybrid transformer-CNN architecture outperforming other state-of-the-art methods. Their method implements a conditional GAN. The generator consists of CNN blocks in the shallow layers to capture local context and save computational resources, while transformers are used in deeper layers to provide better global context.¹¹⁴

4.4 | Synthetic CT for proton radiation therapy

Generating sCTs from MRI for the purposes of proton therapy is not fundamentally different from the process for photon therapy. However, proton therapy takes advantage of the Bragg peak, which concentrates the radiation in a small region to spare healthy tissue. While this is beneficial, this puts a tighter constraint on sCT errors. Another difference is that sCT images must first be converted to relative proton stopping power maps before they can be used in treatment planning. Therefore, directly generating synthetic proton relative stopping power (sRPSP) maps instead of sCT would be ideal. Boron therapy is a form of targeted radiation therapy in which boronated compounds are delivered to the site of the tumor and irradiated with neutrons. The boron undergoes a fission reaction, releasing alpha particles that kill the tumor cells. However, the targeting mechanism typically relies on targeting cancer cells' high metabolic rate. Epidermal tissue that also has a high metabolic rate uptakes boron, making skin dose an important concern in boron therapy. Therefore, methods for generating sCT images for boron therapy should emphasize accurate reconstruction around the skin.

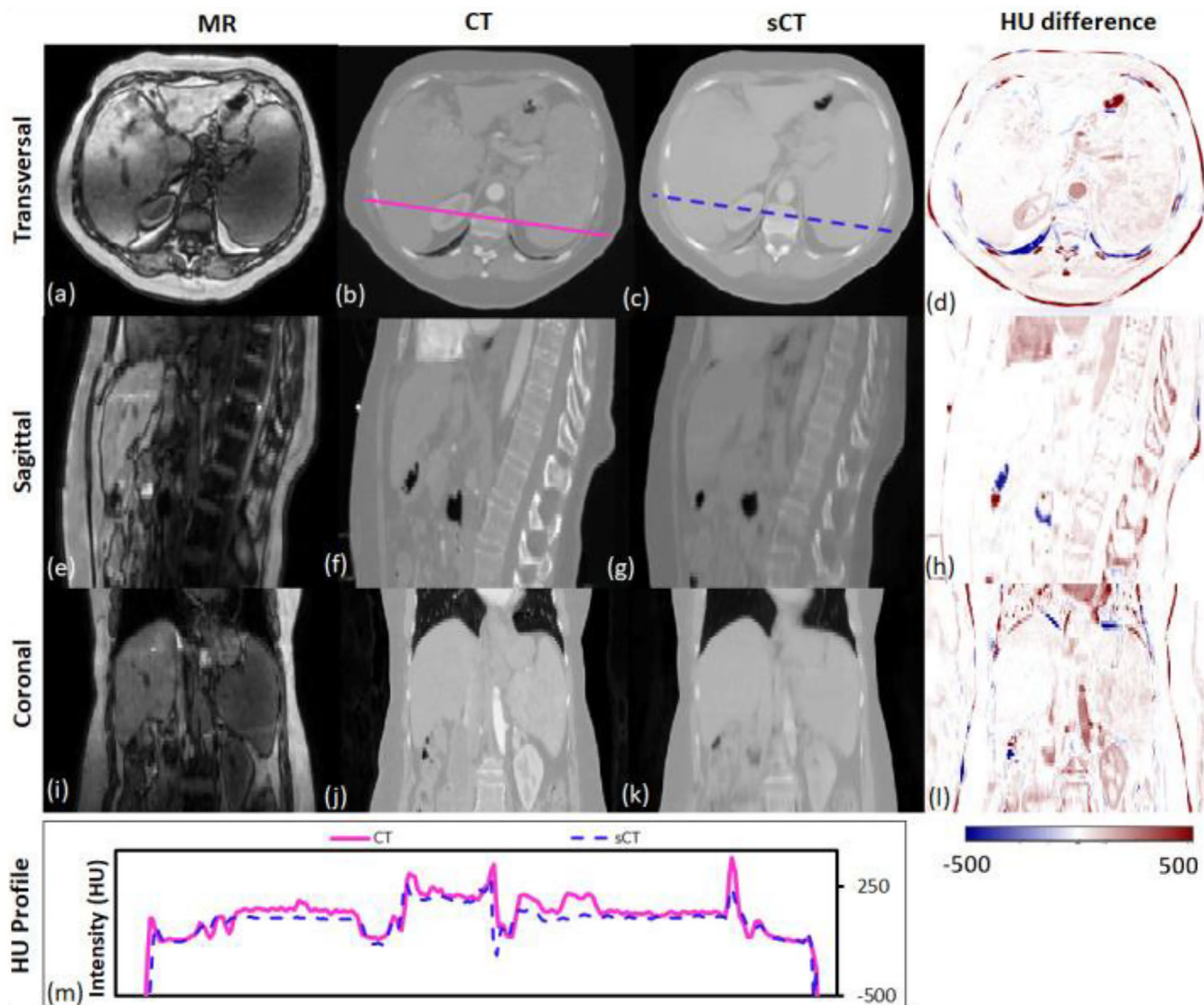


FIGURE 5 Traverse, sagittal, and coronal images of a representative patient. MRI, CT, and sCT images and the HU difference map between CT and sCT are presented. The CT (solid line) and sCT (dashed line) voxel-based HU profiles of the traverse images are compared in the lowermost panel. Reprinted by permission from British Journal of Radiology, MRI-based treatment planning for liver stereotactic body radiotherapy: validation of a deep learning-based synthetic CT generation method by Liu et al.¹⁰⁸© 2019.

Shown in Table 7, many methods show high dosimetric accuracy for proton therapy. Liu *et al* develops a conditional cycleGAN to synthesize both high and lower energy CT.¹¹⁵ Wang et al. create the first synthetic relative proton stopping power maps from MRI with a cycleGAN and loss function to take advantage of paired data. Their method achieves an excellent MAE of 42 ± 13 HU, but struggles with dosimetric accuracy.⁶ Maspero et al. achieve a 2%/2 mm gamma pass rate above 99% for proton therapy by averaging predictions from three separate GANs trained on axial, sagittal, and coronal views, respectively.¹¹⁶ Replacing traditional MC dose calculation methods, Tsekas et al. generate VMAT (volumetric modulated arc therapy) dose distributions in static positions with sCT.¹¹⁷ Finally, SARU, a self-attention Res-UNet, lowers skin dose for

boron therapy, achieving better results than the pix2pix method.¹¹⁸

4.5 | CT and CBCT-based synthetic MRI

Generating sMRI from CT leverages MRI's high soft tissue contrast for improved segmentation accuracy and pathology detection for CT-only treatment planning. In addition, the ground truth x-ray attenuation information is maintained compared to an MRI-only workflow. Cone beam CT (CBCT) is primarily used for patient positioning before each fraction of radiation therapy. Kilovoltage (kV) and megavoltage (MV) energies are standard in CBCT with kV images providing superior contrast and MV images providing superior tissue penetration.

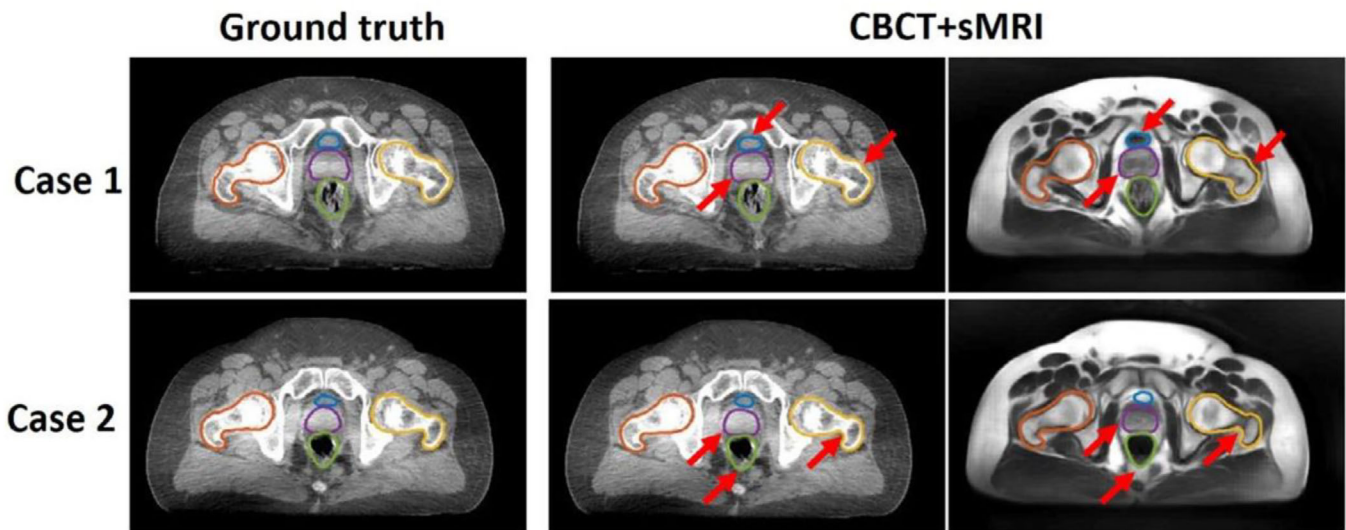


FIGURE 6 Contours of segmented pelvic organs for two representative patients. Ground truth contours are overlaid onto CBCT. The predicted contours of the proposed method are overlaid on CBCT and sMRI. Red arrows highlight regions in which CBCT and sMRI provide complementary information for bony structure and soft tissue segmentation. Reprinted by permission from John Wiley and Sons: Medical Physics, Pelvic multi-organ segmentation on cone-beam CT for prostate adaptive radiotherapy by Fu et al.¹²³ © 2020.

However, noise and artifacts can often reduce CBCT image quality.¹¹⁹ Generating CBCT-based sMRI can yield higher image quality and soft-tissue contrast while also retaining CBCT's fast acquisition speed. CT and CBCTs' rapid acquisition time can make it preferable over MRI for patients with claustrophobia during the MR simulation or for pediatric patients who would require additional sedation. In addition, MRI is not suitable for patients with metal implants such as pacemakers. However, sMRI is significantly more challenging to generate compared to sCT. This is primarily due to the recovery of soft tissue structures visible only in MRI. For this reason, sMRI is often used to improve segmentation results in CT and CBCT. However, some studies report direct use of sMRI for segmentation.

The studies of CT and CBCT-based synthetic MRI are summarized in Table 8. For CT-based sMRI, Dae *et al* implements a cycleGAN for sMRI synthesis with dense blocks in the generator. The sMRIs are input into MS-RCNN improving segmentation performance.¹²⁰ Lei *et al* incorporates dual pyramid networks to extract features from both sMRI and CT and includes attention to achieve exceptional results.¹²¹ BPGAN synthesizes both sMRI and sCT bidirectionally with a cycleGAN. Pathological prior information, an edge retention loss, and spectral normalization improve accuracy and training stability.¹²² Both CBCT-based sMRI studies, from Emory's Deep Biomedical Imaging Lab, significantly improve CBCT segmentation results. In their first paper, Lei *et al* generates sMRI with a CycleGAN, then inputs this into an attention U-Net.⁸ Fu et al. makes additional improvements by generating the segmentations with inputs from both CBCT and sMRI and also including

additional pelvic structures. Example contours overlaid onto CBCT and sMRI are shown in Figure 6.¹²³

4.6 | Intramodal MRI synthesis and super resolution

It can be beneficial to synthesize MRI sequences from other MRI sequences. Intra-modal applications include generating synthetic contrast MRI to prevent the need for injected contrast, super-resolution MRI to improve image quality and reduce acquisition time, and synthetic 7T MRI due to its lack of widespread availability and improve spatial resolution and contrast.¹²⁴ To reduce complexity and cost, a potential approach to radiation therapy is to rotate the patient instead of using a gantry. However, the patient's organs deform under gravity, requiring multiple MRIs at different angles for MRgRT. MR images of patients rotated at different angles can better enable gantry free radiation therapy. In this section, synthesis studies which synthesize other MRI sequences are discussed and summarized in Table 9.

Preetha et al. synthesize T1C images with a multi-channel T1W, T2W, and FLAIR MRI sequences using the pix2pix architecture.¹² Another study included a A cycleGAN with a ResUNet generator to generate lateral and supine MR images for gantry-free radiation therapy.¹²⁵ ResUNet is also implemented to generate ADC uncertainty maps from ADC maps for prostate cancer and mesothelioma.¹²⁶ Studies designed explicitly for super-resolution include Chun et al. and Zhao et al. In the former study, a U-Net based denoising

autoencoder is trained to remove noise from clinical MRI.¹¹ The same architecture is employed in Kim *et al.*¹²⁷ for real-time 3D MRI to increase spatial resolution. In addition, dynamic keyhole imaging is formulated to reduce acquisition time by only sampling central k-space data associated with contrast. The peripheral k-space data associated with edges is added from previously generated super-resolution images in the same position.¹²⁷ Zhao *et al.* make use of super-resolution for brain tumor segmentation, increasing the dice score from 0.724 to 0.786 with 4x super resolution images generated from a GAN architecture. The generator has low- and high-resolution paths and dense blocks.¹⁰ Often in clinical practice, the through plane resolution is increased to reduce the MRI scan time. Xie *et al.* achieves near perfect accuracy in recovering 1 from 3 mm through plane resolution by training parallel CycleGANs, which predict the higher resolution coronal and sagittal slices, respectively. These predictions are then fused to create the final 3D prediction.¹²⁸

5 | RADIOMICS (CLASSIFICATION)

Unlike synthesis which maps one imaging modality to another, radiomics extracts imaging data to classify structures or to predict a value. Deep learning applications to MRI-based radiomics often achieve state-of-the-art performance over hand-crafted methods in detection and treatment outcome prediction tasks. Traditional radiomics algorithms apply various hand-crafted matrices based on shape, intensity, texture, and imaging filters to generate features. The majority of these features have no predictive power, and would confuse the model if all were directly implemented. Therefore, an important step is feature reduction which screens out features without statistical significance. Typically, this is done with a regression such as analysis of variance (ANOVA), Least Absolute Shrinkage and Selection Operator (LASSO), or ridge regression. Alternatively, a CNN or other neural network can learn significant features. The advantage of the deep learning approach is that the network can learn any relevant features including handcrafted ones. However, this assumes a large enough dataset which can be problematic for small medical datasets. Hand-crafted features have no such constraint and are easily interpretable. It is often the case that a hybrid approach including both hand-crafted and deep learning features yields the highest performance. Biometric data like tumor grade, patient age, and biomarkers can also be included as features. Once the significant features are found, supervised machine learning algorithms like support vector machines, artificial neural networks, and random forests are employed to make a prediction from these features. Recently, CNNs like Xception and InceptionResNet,¹²⁹ recurrent neural networks with GRU and LSTM blocks, and trans-

formers have also found favor in this task, as introduced in Section 3. Radiomics can also be done purely with deep learning as it is done with segmentation and synthesis. In this section, we divide the studies into those detecting or classifying objects in the image and studies predicting a value such as the likelihood of distant metastases, treatment response, and adverse effects. While detection is traditionally under the purview of segmentation, the architectures of detection methods and the classification task are in common with other radiomics methods, and so are discussed here.

While radiomics algorithms can excel on local datasets, the main concern for MRI applications is the generalizability of the methods. Variability in MR imaging characteristics such as field strength, scanner manufacturer, pulse sequence, ROI or contour quality, and the feature extraction method can result in different features being significant. This variability can largely be mitigated by normalizing the data to a reference MRI and including data from multiple sources.¹³⁰

5.1 | Evaluation metrics

Classification accuracy is an appealing evaluation metric due to its simplicity, but accuracy can be misleading with unbalanced data. For example, if 90% of tumors in the dataset are malignant, a model can achieve 90% accuracy by labeling every tumor as malignant. Precision,¹³¹ the ratio of true positives to all examples labeled as positive by the classifier, and recall,¹⁵ the ratio of true positives to all actual positives, will also both differ if given imbalanced data. The F1 score¹³² is defined in Equation 3, ranging from 0 to 1 and combining precision and recall to provide a single metric. A high F1 value indicates both high precision and recall and is resilient towards unbalanced data.

$$F1 = 2 \frac{\text{Recall} * \text{Precision}}{\text{Recall} + \text{Precision}} \quad (3)$$

The most common evaluation metric resistant to unbalanced data is the area under the curve (AUC) of a receiver operating characteristics (ROC) curve.^{133–135} In a ROC curve, the x-axis represents the false positive (FP) rate, while the y-axis relates the true positive (TP) rate. In addition, the ROC curve can be viewed as a visual representation to help find the best trade-off between sensitivity and specificity for the clinical application by comparing one minus the specificity versus the sensitivity of the model. The AUC value provides a measurement for the overall performance of the model with a value of 0.5 representing random chance and a value of 1 being perfect classification. If the AUC value is below 0.5, the classifier would simply need to invert its predictions to achieve higher accuracy. It is important to note that all these metrics are for binary classification but are

commonly used in multi-class classification by comparing a particular class with an amalgamation of every other category. Finally, the concordance index (C-index) measures how well a classifier predicts a sequence of events and is most appropriate for prognostic models which predict the timing of adverse effects, tumor recurrence, or patient survival times. The C-index ranges from 0 to 1 with a value of 1 being perfect prediction.^{136,137} A full discussion of evaluation metrics for classification tasks is found in Hossin and Suliaman.¹³⁸

5.2 | Cancer detection and staging

Effectively detecting and classifying tumors is vital for treatment planning. Deep learning detection methods supersede segmentation algorithms when the tumors are difficult to accurately segment or cannot easily be distinguished from other structures. In addition, detection models can further improve segmentation results by eliminating false positives. When applied to MRI, detection studies also have the potential to differentiate between cancer types and tumor stage to potentially avoid unnecessary invasive procedures like biopsy.

As shown in Table 10, The majority of works in detection are for brain lesion classification. Chakrabarty et al. attain exceptional results in differentiating between common types of brain tumors with a 3D CNN and outperforms traditional hand-crafted methods.¹³³ Radiation-induced cerebral microbleeds appear as small dark spots in 7T time of flight magnetic resonance angiography (TOF MRA) and can be difficult to distinguish from look-a-like structures. Chen et al. utilize a 3D ResNet model to differentiate between true cerebral microbleeds and mimicking structures with high accuracy.¹³⁹ Finally, Gao et al. distinguish between radiation necrosis and tumor recurrence for gliomas, significantly outperforming experienced neurosurgeons with a CNN.¹⁴⁰

5.3 | Treatment response

The decision to treat with radiation therapy is often definitive. Since radiation dose will unavoidably be delivered to healthy tissue, treatment response and the risk of adverse effects are heavily considered. Further compounding the decision, dose to healthy tissue is cumulative that is complicating any subsequent treatments. In addition, unknown distant metastasis can derail radiation therapy's curative potential. Therefore, predicting treatment response and adverse effects are of high importance, and significant work has gone into applying deep learning algorithms to prognostic models.

Diffusion-weighted imaging (DWI) has attracted strong interest in studies which predict the outcome of radiation therapy. DWI measures the diffusion of water through tissue often yielding high contrast for tumors.

Cancers can be differentiated by altering DWI's sensitivity to diffusion with the b value, in which higher b values correspond to an increased sensitivity to diffusion. By sampling at multiple b-values, the attenuation of the MR signal can be measured locally in the form of apparent diffusion coefficient (ADC) values. A drawback of DWI is that the spatial resolution is often significantly worse than T1W and T2W imaging.¹⁴¹ Unlike segmentation and synthesis which require highly accurate structural information, high spatial resolution is not necessary for treatment outcome prediction, so the functional information from DWI is most easily exploited in predictive algorithms.

The majority of studies summarized in Table 11 seek to predict treatment outcomes and tumor recurrence. Zhu et al. take the interesting approach of concatenating DWI histograms across twelve b values to create a "signature image." A CNN is then applied to the signature image to achieve exceptional performance in predicting pathological complete response.¹⁴ Jing et al., in addition to MRI data includes clinical data like age, gender, and tumor stage to improve predictive performance.¹⁴² Keek et al. achieves better results in predicting adverse effects by combining hand-crafted radiomics and deep learning features.¹⁵ Other notable papers include Huisman et al., which uses an FCN suggesting that radiation therapy accelerates brain aging by 2.78 times,¹⁴³ Hua et al., which predicts distant metastases with an AUC of 0.88,¹⁴⁴ and Jalalifar et al., which achieves excellent results by feeding in clinical and deep learning features into an LSTM model.¹⁴⁵ An additional study by Jalalifar et al. finds the best performance for local treatment response prediction using a hybrid CNN-transformer architecture when compared to other methods. Residual connections and algorithmic hyperparameter selection further improve results.¹⁴⁶

6 | REAL-TIME AND 4D MRI

Real-time MRI during treatment has recently been made possible in the clinical setting with the creation of the MRI-LINAC. Popular models include the Viewray MRIdian (ViewRay Inc, Oakwood, Ohio, USA) and the Elekta Unity (Elekta AB, Stockholm). Electron return effect (ERE), which increasing dose at boundaries with differing proton densities such as the skin at an external magnetic field, guides the architecture of these models.¹⁴⁷ At higher field strengths, the ERE becomes more significant, but MR image quality increases. In addition, a higher field strength can reduce the acquisition time for real-time MRI. Therefore, a balance must be struck. Both the Elekta Unity and Viewray MRidian with 1.5T and 0.35T magnetic fields, respectively, compromise by choosing lower field strengths. The Elekta Unity prioritizes image quality and real-time tracking capabilities at the expense of a more

severe ERE.¹⁴⁸ The MRI-LINAC has enabled an exciting new era of ART wherein anatomical changes and changes to the tumor volume can be accurately discerned and optimized between treatment fractions. In addition, unique to MRgRT, the position of the tumor can be directly monitored during treatment, potentially leading to improved tumor conformality and improved patient outcomes.¹⁴⁹

Periodic respiratory and cardiac motion are common sources of organ deformation and should be accounted for optimal dose delivery to the PTV. Tracking these motions is problematic with conventional MRI since scans regularly take approximately 2 min per slice leading to a total typical scan time of 20–60 min.¹⁵⁰ In addition to motion restriction techniques like patient-breath hold, cine MRI accounts for motion in real-time by reducing acquisition times to 15 seconds or less. This is achieved by only sampling one (2D) or more (3D) slices with short repetition times, increasing slice thickness, and undersampling. In addition, the MR signal is sampled radially in k-space to reduce motion artifacts. Capturing a 3D volume across multiple timesteps of periodic motion is known as 4D MRI.¹⁵¹

Deep learning methods can further reduce acquisition time by reconstructing intensely undersampled cine MRI slices. In addition to reconstructing from undersampled k-space MRI sequences, several approaches further reduce acquisition time. In the first approach, cine MRI and/or k-space trajectories are used to predict the timestep of a previously taken 4D MRI. However, this method requires a lengthy 4D MRI and does not adapt to changes in the tumor volume over the course of the treatment. Additional approaches include synthesizing a larger volume than cine MRI slice captures to reduce acquisition time, predicting the deformation vector field (DVF) which relays real-time organ deformation information, or determining the 3D iso-probability surfaces of the organ to stochastically determine tumor position if real-time motion adaptation is not possible.

Shown in Table 12, this category is experiencing rapid growth with majority of papers being published within the current year. Notable works include Gulamhussene et al., which predict a 3D volume from 2D cine MRI or a 4D volume from a sequence of 2D cine MR slices. A simple U-Net, introduced in Section 3, is implemented to reduce inference time. The performance degrades for synthesized slices far away from the input slices but achieves an exceptional target registration error.¹⁷ Nie et al. instead uses autoregression and the LSTM time series modeling to predict the diaphragm position and to find the matching 4D MRI volume. Autoregression outperforms an LSTM model which could be attributed to a low number of patients.¹⁵² Patient motion is alternatively predicted in Terpestra et al. by using undersampled 3D cine MRI to generate the DVF with a CNN with low target registration error.¹⁵³ Similarly, Romaguera et al.

predict liver deformation using a residual CNN and prior 2D cine MRI. This prediction is then input into a transformer network to predict the next slice.¹⁵⁴ Driever et al. simply segments the stomach with U-Net and constructs iso-probability surfaces centered about the center of mass to isolate respiratory motion. These probability distributions can then be implemented in treatment planning.¹⁵⁵

7 | OVERVIEW AND FUTURE DIRECTIONS

Over the last 6 years, we have seen growing adoption of MRgRT and the rapid development of powerful deep learning techniques which encourage an efficient, adaptive MRI only workflow. Shown in Figure 2, powerful models which better exploit MRI's 3D and long-range context and generative learning continue to gain research interest and be improved upon. In addition, the MRI-LINAC has spawned the exciting new field of real time MRI. In this section, we discuss the progress of deep learning applications to MRgRT as well as promising future trends of clinical interest.

In our literature search, we identify three overarching trends for deep learning models in MRgRT:

(1) multimodal approaches—Methods which leverage many different types of information such as different MR sequences, clinical data, and synthetically generated information have demonstrated state-of-the-art results and often outperform models using only one source of data. Following this trend, we anticipate that data sources commonly applied to radiomics methods like genomics data,¹⁵⁶ biomarkers,¹⁴² and additional imaging sources to have an increased role in other MRgRT applications. Including synthetic data can also enhance performance leading to the blurring of the subfields of MRgRT (image segmentation, image synthesis, radiomics, and real time MRI). For example, studies achieved higher performance on CT segmentation tasks by performing the contouring on sMRI images.⁸ Therefore, it is foreseeable that future models may consider information from all subfields for optimal adaptive treatment planning.

(2) transformer models—Transformer models have been proved to be powerful by directly learning global relations in MRI but can become computationally expensive. Currently a balance is often struck with hybrid CNN-Transformer models in which convolutional layers capture fine detail in early layers while transformers capture global context in deeper layers. These may transition to purely transformer models as computational resources and more efficient approaches are developed. Attention mechanisms for multi-modal segmentation and synthesis have improved upon multi-modal image synthesis by preferentially weighting input channels with stronger context. We predict that this success will be

improved upon with transformers. An additional exciting property of transformer models is that they first divide images into a 1D input sequence of patches, so it trivial to add additional patches to represent diverse data sources. Along with recent computationally efficient multimodal approaches,¹⁵⁷ this feature makes transformers a strong contender to effectively employ data from many sources.

Finally, transformer models show promise in better handling small tumor volumes. While state-of-the-art CNN models first identify a ROI so that the network can better focus on relevant features, transformers can directly identify important regions of the image while still considering long-range context. This emphasis on global context may also help when significant motion blurring or other artifacts prevent CNNs from learning meaningful local features and a holistic understanding is required to achieve accurate results.

(3) generative models—Generative models create data and have proved especially powerful in image segmentation and synthesis tasks. GANs and its variants such as the WGAN-GP and CycleGAN have been the dominant model in MRgRT applications but are difficult to train and often suffer from instability. Although not represented in any publications at the time of this review, diffusion models solve these issues and have achieved state state-of-the art results in computer vision tasks. We expect diffusion models to quickly gain interest for MRgRT applications as these algorithms improve and become less computationally intensive.

Alongside the development of novel algorithms, breakthroughs in MRI technology and the clinical workflow will undoubtedly benefit from advanced deep learning architectures. Image segmentation models mitigate tedious contouring and intra-observer variability. This need will rise with the advent of the MRI-LINAC and real-time MRI, as tracking tumor motion will necessitate real-time contouring or prediction of the deformable vector field. In addition, the increased soft tissue contrast of MRI has allowed for the differentiation of substructures with differing radiation tolerances, requiring additional contouring. Although already finding success in a variety of tasks like generating sCT, sMRI, and super-resolution, synthesis models are expected to continuously adapt to emerging technologies. For instance, the image quality of MR-LINACs is comparably poor due to lower field strengths necessary to mitigate the ERE and could be enhanced with image synthesis models based on high quality diagnostic scans. Furthermore, 7T MRI is gaining clinical interest due to its higher resolution and image contrast, but scanners are not widespread and may cause side effects like nausea.¹⁵⁸ Image synthesis models could increase the quality of existing 1.5T and 3T scanners in a similar fashion. Finally, radiomics models might also find new applications in providing insights into treatment progress and outcomes.¹⁵⁹

8 | CONCLUSION

New deep learning approaches to MRgRT are rapidly improving state-of-the-art performance in segmentation, synthesis, radiomics, and real-time MRI. Trends such as multimodal approaches, transformer models, and generative models demonstrate great potential in tackling current areas of research such as generating accurate x-ray attenuation information, the “small tumor problem” in image segmentation, and generating high quality radiomics predictions. In addition, these approaches pave the way to better integrate real time MRI into the clinical workflow and improve image quality at shorter acquisition times and lower field strengths.

AUTHOR CONTRIBUTIONS

Conception and design: Zach Eidex, Tian Liu, Xiaofeng Yang; Data collection: Zach Eidex, Yifu Ding, Jing Wang; Analysis and interpretation: Zach Eidex, Elham Abouei, Richard L.J. Qiu; Draft manuscript preparation: Zach Eidex, Richard L.J. Qiu, Tonghe Wang, Xiaofeng Yang.

ACKNOWLEDGMENTS

This research is supported in part by the National Cancer Institute of the National Institutes of Health under Award Numbers R01CA215718, R56EB033332, R01EB032680 and P30 CA008748.

CONFLICT OF INTEREST STATEMENT

The author declares no conflicts of interest.

ORCID

Richard L. J. Qiu 

<https://orcid.org/0000-0002-7877-1900>

REFERENCES

- Chin S, Eccles CL, McWilliam A, et al. Magnetic resonance-guided radiation therapy: a review [published online ahead of print 20191023]. *J Med Imaging Radiat Oncol*. 2020;64(1):163-177.
- Murgic J, Chung P, Berlin A, et al. Lessons learned using an MRI-only workflow during high-dose-rate brachytherapy for prostate cancer [published online ahead of print 20160129]. *Brachytherapy*. 2016;15(2):147-155.
- Johnstone E, Wyatt JJ, Henry AM, et al. Systematic review of synthetic computed tomography generation methodologies for use in magnetic resonance imaging-only radiation therapy [published online ahead of print 20170908]. *Int J Radiat Oncol Biol Phys*. 2018;100(1):199-217.
- Fu Y, Lei Y, Wang T, Curran WJ, Liu T, Yang X. A review of deep learning based methods for medical image multi-organ segmentation [published online ahead of print 20210513]. *Phys Med*. 2021;85:107-122.
- Fu Y, Lei Y, Wang T, Curran WJ, Liu T, Yang X. Deep learning in medical image registration: a review. *Phys Med Biol*. 2020;65(20).
- Wang C, Uh J, Patni T, et al. Toward MR-only proton therapy planning for pediatric brain tumors: synthesis of relative proton stopping power images with multiple sequence MRI and development of an online quality assurance tool [published

- online ahead of print 20220211]. *Med Phys.* 2022;49(3):1559-1570.
7. Dong X, Lei Y, Tian S, et al. Synthetic MRI-aided multi-organ segmentation on male pelvic CT using cycle consistent deep attention network [published online ahead of print 20191017]. *Radiother Oncol.* 2019;141:192-199.
 8. Lei Y, Wang T, Tian S, et al. Male pelvic multi-organ segmentation aided by CBCT-based synthetic MRI [published online ahead of print 20200204]. *Phys Med Biol.* 2020;65(3):035013.
 9. Kieselmann JP, Fuller CD, Gurney-Champion OJ, Oelfke U. Cross-modality deep learning: contouring of MRI data from annotated CT data only [published online ahead of print 20201213]. *Med Phys.* 2021;48(4):1673-1684.
 10. Zhou Z, Ma A, Feng Q, et al. Super-resolution of brain tumor MRI images based on deep learning [published online ahead of print 20220915]. *J Appl Clin Med Phys.* 2022;23(11):e13758.
 11. Chun J, Zhang H, Gach HM, et al. MRI super-resolution reconstruction for MRI-guided adaptive radiotherapy using cascaded deep learning: in the presence of limited training data and unknown translation model [published online ahead of print 20190807]. *Med Phys.* 2019;46(9):4148-4164.
 12. Jayachandran Preetha C, Meredig H, Brugnara G, et al. Deep-learning-based synthesis of post-contrast T1-weighted MRI for tumour response assessment in neuro-oncology: a multicentre, retrospective cohort study [published online ahead of print 20211020]. *Lancet Digit Health.* 2021;3(12):e784-e794.
 13. Yang Z, Chen M, Kazemimoghadam M, et al. Deep-learning and radiomics ensemble classifier for false positive reduction in brain metastases segmentation [published online ahead of print 20220119]. *Phys Med Biol.* 2022;67(2).
 14. Zhu HT, Zhang XY, Shi YJ, Li XT, Sun YS. The conversion of MRI data with multiple b-Values into signature-like pictures to predict treatment response for rectal cancer [published online ahead of print 20211216]. *J Magn Reson Imaging.* 2022;56(2):562-569.
 15. Keek SA, Beuque M, Primakov S, et al. Predicting adverse radiation effects in brain tumors after stereotactic radiotherapy with deep learning and handcrafted radiomics [published online ahead of print 20220713]. *Front Oncol.* 2022;12:920393.
 16. Shao HC, Li T, Dohopolski MJ, et al. Real-time MRI motion estimation through an unsupervised k-space-driven deformable registration network (KS-RegNet) [published online ahead of print 20220629]. *Phys Med Biol.* 2022;67(13).
 17. Gulamhussene G, Meyer A, Rak M, et al. Predicting 4D liver MRI for MR-guided interventions [published online ahead of print 20220909]. *Comput Med Imaging Graph.* 2022;101:102122.
 18. Frueh M, Kuestner T, Nachbar M, Thorwarth D, Schilling A, Gatidis S. Self-supervised learning for automated anatomical tracking in medical image data with minimal human labeling effort [published online ahead of print 20220827]. *Comput Methods Programs Biomed.* 2022;225:107085.
 19. Fiorino C, Reni M, Bolognesi A, Cattaneo GM, Calandrino R. Intra- and inter-observer variability in contouring prostate and seminal vesicles: implications for conformal treatment planning. *Radiother Oncol.* 1998;47(3):285-292.
 20. Behjatnia B, Sim J, Bassett LW, Moatamed NA, Apple SK. Does size matter? Comparison study between MRI, gross, and microscopic tumor sizes in breast cancer in lumpectomy specimens [published online ahead of print 20100222]. *Int J Clin Exp Pathol.* 2010;3(3):303-309.
 21. Balagopal A, Morgan H, Dohopolski M, et al. PSA-Net: deep learning-based physician style-aware segmentation network for postoperative prostate cancer clinical target volumes. *Artificial Intelligence in Medicine.* 2021;121:102195.
 22. Liu Y, Stojadinovic S, Hrycushko B, et al. A deep convolutional neural network-based automatic delineation strategy for multiple brain metastases stereotactic radiosurgery [published online ahead of print 20171006]. *PLoS One.* 2017;12(10):e0185844.
 23. Cao Y, Vassantachart A, Ragab O, et al. Automatic segmentation of high-risk clinical target volume for tandem-and-ovoids brachytherapy patients using an asymmetric dual-path convolutional neural network [published online ahead of print 20220204]. *Med Phys.* 2022;49(3):1712-1722.
 24. Alzubaidi L, Zhang J, Humaidi AJ, et al. Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. *Journal of Big Data.* 2021;8(1):53.
 25. Cui Z, Ke R, Pu Z, Wang Y. Deep Bidirectional and Unidirectional LSTM Recurrent Neural Network for Network-wide Traffic Speed Prediction. 2018. doi:10.48550/ARXIV.1801.02143
 26. Liu X, Wang Y, Wang X, Xu H, Li C, Xin X. Bi-directional gated recurrent unit neural network based nonlinear equalizer for coherent optical communication system. *Opt Express.* 2021;29(4):5923-5933.
 27. Yang S, Yu X, Zhou Y. LSTM and GRU Neural Network Performance Comparison Study: taking Yelp Review Dataset as an Example. *Paper presented at: 2020 International Workshop on Electronic Communication and Artificial Intelligence (IWECAI).* 2020:2020.
 28. Simon AB, Hurt B, Karunamuni R, et al. Automated segmentation of multiparametric magnetic resonance images for cerebral AVM radiosurgery planning: a deep learning approach [published online ahead of print 20220117]. *Sci Rep.* 2022;12(1):786.
 29. Zhong G, Yue G, Ling X. Recurrent Attention Unit. 2018. doi:10.48550/ARXIV.1810.12754
 30. Vaswani A, Shazeer N, Parmar N, et al. Attention Is All You Need. 2017. doi:10.48550/ARXIV.1706.03762
 31. Li C, Tang T, Wang G, et al. BossNAS: Exploring Hybrid CNN-transformers with Block-wisely Self-supervised Neural Architecture Search. 2021. doi:10.48550/ARXIV.2103.12424
 32. Child R, Gray S, Radford A, Sutskever I. Generating Long Sequences with Sparse Transformers. 2019. doi:10.48550/ARXIV.1904.10509
 33. Dao T, Fu DY, Ermon S, Rudra A, Ré C. FlashAttention: fast and memory-efficient exact attention with IO-Awareness. *arXiv [cs.LG].* 2022.
 34. Ahmadi M, Sharifi A, Hassantabar S, Enayati S. QAIS-DSNN: tumor area segmentation of MRI image with optimized quantum matched-filter technique and deep spiking neural network [published online ahead of print 20210118]. *Biomed Res Int.* 2021;2021:6653879.
 35. Yoo SK, Kim TH, Chun J, et al. Deep-learning-based Automatic detection and segmentation of brain metastases with small volume for stereotactic ablative radiotherapy [published online ahead of print 20220523]. *Cancers (Basel).* 2022;14(10).
 36. Hänsch A, Hendrik Moltz J, Geisler B, et al. Hippocampus segmentation in CT using deep learning: impact of MR versus CT-based training contours [published online ahead of print 20201111]. *J Med Imaging (Bellingham).* 2020;7(6):064001.
 37. Wang J, Lu J, Qin G, et al. Technical Note: a deep learning-based autosegmentation of rectal tumors in MR images [published online ahead of print 20180503]. *Med Phys.* 2018;45(6):2560-2564.
 38. Billot B, Greve DN, Puonti O, et al. SynthSeg: segmentation of brain MRI scans of any contrast and resolution without retraining. *Medical Image Analysis.* 2023;86:102789.
 39. Menze BH, Jakab A, Bauer S, et al. The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS) [published online ahead of print 20141204]. *IEEE Trans Med Imaging.* 2015;34(10):1993-2024.
 40. Kamnitsas K, Ledig C, Newcombe VFJ, et al. Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. *Med Image Anal.* 2017;36:61-78.
 41. Zhuang F, Qi Z, Duan K, et al. A Comprehensive Survey on Transfer Learning. 2019. doi:10.48550/ARXIV.1911.02685

42. Lin TY, Goyal P, Girshick R, He K, Dollár P. Focal loss for dense object detection. *IEEE Trans Pattern Anal Machine Intell.* 2020;42(2):318-327.
43. Salehi SSM, Erdogmus D, Gholipour A. Tversky loss function for image segmentation using 3D fully convolutional deep networks. 2017. doi:10.48550/ARXIV.1706.05721
44. Hsu DG, Ballangrud Å, Shamseddine A, et al. Automatic segmentation of brain metastases using T1 magnetic resonance and computed tomography images [published online ahead of print 20210826]. *Phys Med Biol.* 2021;66(17).
45. Tao G, Li H, Huang J, et al. SeqSeg: a sequential method to achieve nasopharyngeal carcinoma segmentation free from background dominance. *Med Image Anal.* 2022;78:102381.
46. He K, Gkioxari G, Dollár P, Girshick R. Mask R-CNN. 2017. doi:10.48550/ARXIV.1703.06870
47. Jaeger PF, Kohl SAA, Bickelhaupt S, et al. Retina U-Net: Embarrassingly Simple Exploitation of Segmentation Supervision for Medical Object Detection. 2018. doi:10.48550/ARXIV.1811.08661
48. Cai S, Tian Y, Lui H, Zeng H, Wu Y, Chen G. Dense-UNet: a novel multiphoton in vivo cellular image segmentation model based on a convolutional neural network. *Quant Imaging Med Surg.* 2020;10(6):1275-1285.
49. Diakogiannis FI, Waldner F, Caccetta P, Wu C. ResUNet-a: a deep learning framework for semantic segmentation of remotely sensed data. *ISPRS J Photogram Remote Sensing.* 2020;162:94-114.
50. Alom MZ, Hasan M, Yakopcic C, Taha TM, Asari VK. Recurrent Residual Convolutional Neural Network based on U-Net (R2U-Net) for Medical Image Segmentation. 2018. doi:10.48550/ARXIV.1802.06955
51. Zhang J, Gu L, Han G, Liu X. AttR2U-Net: a fully automated model for MRI nasopharyngeal carcinoma segmentation based on spatial attention and residual recurrent convolution [published online ahead of print 20220128]. *Front Oncol.* 2021;11:816672.
52. Oktay O, Schlemper J, Folgoc LL, et al. Attention U-Net: Learning Where to Look for the Pancreas. 2018. doi:10.48550/ARXIV.1804.03999
53. Milletari F, Navab N, Ahmadi SA. V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation. 2016. doi:10.48550/ARXIV.1606.04797
54. Isensee F, Jaeger PF, Kohl SAA, Petersen J, Maier-Hein KH. nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. *Nat Methods.* 2021;18(2):203-211.
55. Isola P, Zhu J-Y, Zhou T, Efros AA. Image-to-Image Translation with Conditional Adversarial Networks. 2016. doi:10.48550/ARXIV.1611.07004
56. Chen L-C, Zhu Y, Papandreou G, Schroff F, Adam H. Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation. 2018. doi:10.48550/ARXIV.1802.02611
57. Huang Z, Huang L, Gong Y, Huang C, Wang X. Mask Scoring R-CNN. 2019. doi:10.48550/ARXIV.1903.00241
58. Chollet F. Xception: Deep Learning with Depthwise Separable Convolutions. 2016. doi:10.48550/ARXIV.1610.02357
59. Liang Y, Schott D, Zhang Y, et al. Auto-segmentation of pancreatic tumor in multi-parametric MRI using deep convolutional neural networks [published online ahead of print 20200208]. *Radiother Oncol.* 2020;145:193-200.
60. Bousabarah K, Ruge M, Brand JS, et al. Deep convolutional neural networks for automated segmentation of brain metastases trained on clinical data [published online ahead of print 20200420]. *Radiat Oncol.* 2020;15(1):87.
61. Eidex ZA, Wang T, Lei Y, et al. MRI-based prostate and dominant lesion segmentation using cascaded scoring convolutional neural network [published online ahead of print 20220517]. *Med Phys.* 2022;49(8):5216-5224.
62. Müller D, Soto-Rey I, Kramer F. Towards a guideline for evaluation metrics in medical image segmentation [published online ahead of print 20220620]. *BMC Res Notes.* 2022;15(1):210.
63. Kalavathi P, Prasath VB. Methods on skull stripping of MRI head scan images-a review. *J Digit Imaging.* 2016;29(3):365-379.
64. Momin S, Lei Y, Tian Z, et al. Cascaded mutual enhancing networks for brain tumor subregion segmentation in multiparametric MRI [published online ahead of print 20220411]. *Phys Med Biol.* 2022;67(8).
65. Huang Y, Bert C, Sommer P, et al. Deep learning for brain metastasis detection and segmentation in longitudinal MRI data [published online ahead of print 20220722]. *Med Phys.* 2022;49(9):5773-5786.
66. Chartrand G, Emiliani RD, Pawlowski SA, et al. Automated detection of brain metastases on T1-weighted MRI using a convolutional neural network: impact of volume aware loss and sampling strategy [published online ahead of print 20220527]. *J Magn Reson Imaging.* 2022;56(6):1885-1898.
67. Tian S, Wang C, Zhang R, et al. Transfer learning-based autosegmentation of primary tumor volumes of glioblastomas using preoperative MRI for radiotherapy treatment [published online ahead of print 20220414]. *Front Oncol.* 2022;12:856346.
68. Ghaffari M, Samarasinghe G, Jameson M, et al. Automated post-operative brain tumour segmentation: a deep learning model based on transfer learning from pre-operative images [published online ahead of print 20211027]. *Magn Reson Imaging.* 2022;86:28-36.
69. Diehn FE, Michalak GJ, DeLone DR, et al. CT dental artifact: comparison of an iterative metal artifact reduction technique with weighted filtered back-projection [published online ahead of print 20171126]. *Acta Radiol Open.* 2017;6(11):2058460117743279.
70. Cardenas CE, Mohamed ASR, Yang J, et al. Head and neck cancer patient images for determining auto-segmentation accuracy in T2-weighted magnetic resonance imaging through expert manual segmentations. *Med Phys.* 2020;47(5):2317-2322.
71. Tao G, Li H, Huang J, et al. SeqSeg: a sequential method to achieve nasopharyngeal carcinoma segmentation free from background dominance [published online ahead of print 20220211]. *Med Image Anal.* 2022;78:102381.
72. Rodríguez Outeiral R, Bos P, van der Hulst HJ, et al. Strategies for tackling the class imbalance problem of oropharyngeal primary tumor segmentation on magnetic resonance imaging [published online ahead of print 20220813]. *Phys Imaging Radiat Oncol.* 2022;23:144-149.
73. Deng Y, Li C, Lv X, et al. The contrast-enhanced MRI can be substituted by unenhanced MRI in identifying and automatically segmenting primary nasopharyngeal carcinoma with the aid of deep learning models: an exploratory study in large-scale population of endemic area [published online ahead of print 20220216]. *Comput Methods Programs Biomed.* 2022;217:106702.
74. Wahid KA, Ahmed S, He R, et al. Evaluation of deep learning-based multiparametric MRI oropharyngeal primary tumor auto-segmentation and investigation of input channel effects: results from a prospective imaging registry [published online ahead of print 20211016]. *Clin Transl Radiat Oncol.* 2022;32:6-14.
75. Jiang J, Hu YC, Tyagi N, et al. Self-derived organ attention for unpaired CT-MRI deep domain adaptation based MRI segmentation [published online ahead of print 20201007]. *Phys Med Biol.* 2020;65(20):205001.
76. Kelly P, Das P, Pinnix CC, et al. Duodenal toxicity after fractionated chemoradiation for unresectable pancreatic cancer [published online ahead of print 20121129]. *Int J Radiat Oncol Biol Phys.* 2013;85(3):e143-149.
77. Steponavičienė R, Jonušas J, Griškevičius R, Venius J, Cicėnas S. A pilot study of safer radiation dosage to the heart and it

- subregions [published online ahead of print 20210331]. *Medicina (Kaunas)*. 2021;57(4).
78. Luximon DC, Abdulkadir Y, Chow PE, Morris ED, Lamb JM. Machine-assisted interpolation algorithm for semi-automated segmentation of highly deformable organs [published online ahead of print 20211127]. *Med Phys*. 2022;49(1):41-51.
 79. Ding J, Zhang Y, Amjad A, Xu J, Thill D, Li XA. Automatic contour refinement for deep learning auto-segmentation of complex organs in MRI-guided adaptive radiation therapy [published online ahead of print 20220420]. *Adv Radiat Oncol*. 2022;7(5):100968.
 80. Morris ED, Ghanem AI, Dong M, Pantelic MV, Walker EM, Glide-Hurst CK. Cardiac substructure segmentation with deep learning for improved cardiac sparing [published online ahead of print 20191229]. *Med Phys*. 2020;47(2):576-586.
 81. Wang C, Tyagi N, Rimner A, et al. Segmenting lung tumors on longitudinal imaging studies via a patient-specific adaptive convolutional neural network [published online ahead of print 20181231]. *Radiother Oncol*. 2019;131:101-107.
 82. Wang C, Rimner A, Hu YC, et al. Toward predicting the evolution of lung tumors during radiotherapy observed on a longitudinal MR imaging study via a deep learning algorithm [published online ahead of print 20190906]. *Med Phys*. 2019;46(10):4699-4707.
 83. Wang B, Lei Y, Tian S, et al. Deeply supervised 3D fully convolutional networks with group dilated convolution for automatic MRI prostate segmentation. *Medical Physics*. 2019;46(4):1707-1718.
 84. Kerkmeijer LGW, Groen VH, Pos FJ, et al. Focal boost to the intraprostatic tumor in external beam radiotherapy for patients with localized prostate cancer: results from the FLAME Randomized Phase III Trial [published online ahead of print 20210120]. *J Clin Oncol*. 2021;39(7):787-796.
 85. Simeth J, Jiang J, Nosov A, et al. Deep learning-based dominant index lesion segmentation for MR-guided radiation therapy of prostate cancer. *Med Phys*. 2023. n/a(n/a).
 86. Shaaer A, Paudel M, Smith M, Tonolet F, Ravi A. Deep-learning-assisted algorithm for catheter reconstruction during MR-only gynecological interstitial brachytherapy [published online ahead of print 20211210]. *J Appl Clin Med Phys*. 2022;23(2):e13494.
 87. Zabihollahy F, Viswanathan AN, Schmidt EJ, Lee J. Fully automated segmentation of clinical target volume in cervical cancer from magnetic resonance imaging with convolutional neural network [published online ahead of print 20220727]. *J Appl Clin Med Phys*. 2022;23(9):e13725.
 88. Nie D, Wang L, Gao Y, Lian J, Shen D. STRAINet: spatially varying sTochastic Residual Adversarial Networks for MRI pelvic organ segmentation [published online ahead of print 20181009]. *IEEE Trans Neural Netw Learn Syst*. 2019;30(5):1552-1564.
 89. Singhrao K, Fu J, Parikh NR, et al. A generative adversarial network-based (GAN-based) architecture for automatic fiducial marker detection in prostate MRI-only radiotherapy simulation images [published online ahead of print 20201019]. *Med Phys*. 2020;47(12):6405-6413.
 90. Wang T, Lei Y, Fu Y, et al. A review on medical imaging synthesis using deep learning and its clinical applications. *J Appl Clin Med Phys*. 2021;22(1):11-36.
 91. Seetharaman A, Bhattacharya I, Chen LC, et al. Automated detection of aggressive and indolent prostate cancer on magnetic resonance imaging. *Med Phys*. 2021;48(6):2960-2972.
 92. Mirza M, Osindero S. Conditional Generative Adversarial Nets. 2014. doi:10.48550/ARXIV.1411.1784
 93. Zhu J-Y, Park T, Isola P, Efros AA. Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks. 2017. doi:10.48550/ARXIV.1703.10593
 94. Arjovsky M, Chintala S, Bottou L. WassersteinGAN.
 95. Gulrajani I, Ahmed F, Arjovsky M, Dumoulin V, Courville A. Improved Training of Wasserstein GANs.
 96. Jolicoeur-Martineau A. The relativistic discriminator: a key element missing from standard GAN.
 97. Yang L, Zhang Z, Song Y, et al. Diffusion Models: A Comprehensive Survey of Methods and Applications.
 98. Zimmermann L, Knäusel B, Stock M, Lütgendorf-Caucig C, Georg D, Kuess P. An MRI sequence independent convolutional neural network for synthetic head CT generation in proton therapy [published online ahead of print 20211215]. *Z Med Phys*. 2022;32(2):218-227.
 99. Li W, Li Y, Qin W, et al. Magnetic resonance image (MRI) synthesis from brain computed tomography (CT) images based on deep learning methods for magnetic resonance (MR)-guided radiotherapy. *Quant Imaging Med Surg*. 2020;10(6):1223-1236.
 100. Wang T, Lei Y, Fu Y, et al. A review on medical imaging synthesis using deep learning and its clinical applications [published online ahead of print 20201211]. *J Appl Clin Med Phys*. 2021;22(1):11-36.
 101. Largent A, Marage L, Gicquiau I, et al. Head-and-Neck MRI-only radiotherapy treatment planning: from acquisition in treatment position to pseudo-CT generation [published online ahead of print 20200314]. *Cancer Radiother*. 2020;24(4):288-297.
 102. Yuan J, Fredman E, Jin JY, et al. Monte Carlo dose calculation using MRI based synthetic CT generated by fully convolutional neural network for Gamma Knife radiosurgery. *Technol Cancer Res Treat*. 2021;20:15330338211046433.
 103. Sara U, Akter M, Uddin M. Image Quality Assessment through FSIM, SSIM, MSE and PSNR—A comparative study. *J Comput Commun*. 2019;07:8-18.
 104. Nečasová T, Burgos N, Svoboda D. Chapter 25 - Validation and evaluation metrics for medical and biomedical image synthesis. In: Burgos N, Svoboda D, eds. *Biomedical Image Synthesis and Simulation*. Academic Press; 2022:573-600. doi:10.1016/B978-0-12-824349-7.00032-3
 105. Ogbole GI. Radiation dose in paediatric computed tomography: risks and benefits. *Ann Ib Postgrad Med*. 2010;8(2):118-126.
 106. Wang T, Manohar N, Lei Y, et al. MRI-based treatment planning for brain stereotactic radiosurgery: dosimetric validation of a learning-based pseudo-CT generation method [published online ahead of print 20180814]. *Med Dosim*. 2019;44(3):199-204.
 107. Padilla-Cabal F, Georg D, Fuchs H. A pencil beam algorithm for magnetic resonance image-guided proton therapy. *Med Phys*. 2018;45(5):2195-2204.
 108. Liu Y, Lei Y, Wang T, et al. MRI-based treatment planning for liver stereotactic body radiotherapy: validation of a deep learning-based synthetic CT generation method [published online ahead of print 20190620]. *Br J Radiol*. 2019;92(1100):20190067.
 109. Brou Boni KND, Klein J, Gulyban A, Reynaert N, Pasquier D. Improving generalization in MR-to-CT synthesis in radiotherapy by using an augmented cycle generative adversarial network with unpaired data [published online ahead of print 20210424]. *Med Phys*. 2021;48(6):3003-3010.
 110. Massa HA, Johnson JM, McMillan AB. Comparison of deep learning synthesis of synthetic CTs using clinical MRI inputs [published online ahead of print 20201223]. *Phys Med Biol*. 2020;65(23).
 111. Koike Y, Akino Y, Sumida I, et al. Feasibility of synthetic computed tomography generated with an adversarial network for multi-sequence magnetic resonance-based brain radiotherapy. *J Radiat Res*. 2020;61(1):92-103.
 112. Dinkla AM, Florkow MC, Maspero M, et al. Dosimetric evaluation of synthetic CT for head and neck radiotherapy generated by a patch-based three-dimensional convolutional neural network [published online ahead of print 20190709]. *Med Phys*. 2019;46(9):4095-4104.
 113. Reaungamornrat S, Sari H, Catana C, Kamen A. Multimodal image synthesis based on disentanglement representations of anatomical and modality specific features, learned using

- uncooperative relativistic GAN [published online ahead of print 20220611]. *Med Image Anal.* 2022;80:102514.
114. Zhao B, Cheng T, Zhang X, et al. CT synthesis from MR in the pelvic area using Residual Transformer Conditional GAN [published online ahead of print 20221129]. *Comput Med Imaging Graph.* 2022;103:102150.
 115. Liu R, Lei Y, Wang T, et al. Synthetic dual-energy CT for MRI-only based proton therapy treatment planning using label-GAN [published online ahead of print 20210309]. *Phys Med Biol.* 2021;66(6):065014.
 116. Maspero M, Bentvelzen LG, Savenije MHF, et al. Deep learning-based synthetic CT generation for paediatric brain MR-only photon and proton radiotherapy [published online ahead of print 20200923]. *Radiother Oncol.* 2020;153:197-204.
 117. Tsekas G, Bol GH, Raaymakers BW. Robust deep learning-based forward dose calculations for VMAT on the 1.5T MR-linac [published online ahead of print 20221118]. *Phys Med Biol.* 2022;67(22).
 118. Zhao S, Geng C, Guo C, Tian F, Tang X. SARU: a self-attention ResUNet to generate synthetic CT images for MR-only BNCT treatment planning [published online ahead of print 20220921]. *Med Phys.* 2022. doi:10.1002/mp.15986
 119. Srinivasan K, Mohammadi M, Shepherd J. Applications of linac-mounted kilovoltage Cone-beam Computed Tomography in modern radiation therapy: a review [published online ahead of print 20140703]. *Pol J Radiol.* 2014;79:181-193.
 120. Dai X, Lei Y, Wang T, et al. Automated delineation of head and neck organs at risk using synthetic MRI-aided mask scoring regional convolutional neural network [published online ahead of print 20210818]. *Med Phys.* 2021;48(10):5862-5873.
 121. Lei Y, Wang T, Tian S, et al. Male pelvic CT multi-organ segmentation using synthetic MRI-aided dual pyramid networks [published online ahead of print 20210416]. *Phys Med Biol.* 2021;66(8).
 122. Xu L, Zeng X, Zhang H, Li W, Lei J, Huang Z. BPGAN: bidirectional CT-to-MRI prediction using multi-generative multi-adversarial nets with spectral normalization and localization [published online ahead of print 20200508]. *Neural Netw.* 2020;128:82-96.
 123. Fu Y, Lei Y, Wang T, et al. Pelvic multi-organ segmentation on cone-beam CT for prostate adaptive radiotherapy [published online ahead of print 20200511]. *Med Phys.* 2020;47(8):3415-3422.
 124. Shaffer A, Kwok SS, Naik A, et al. Ultra-high-field MRI in the diagnosis and management of gliomas: a systematic review [published online ahead of print 20220405]. *Front Neurol.* 2022;13:857825.
 125. Chen X, Cao Y, Zhang K, et al. Technical note: a method to synthesize magnetic resonance images in different patient rotation angles with deep learning for gantry-free radiotherapy [published online ahead of print 20220922]. *Med Phys.* 2022. doi:10.1002/mp.15981
 126. Zormpas-Petridis K, Tunariu N, Collins DJ, Messiou C, Koh DM, Blackledge MD. Deep-learned estimation of uncertainty in measurements of apparent diffusion coefficient from whole-body diffusion-weighted MRI [published online ahead of print 20220913]. *Comput Biol Med.* 2022;149:106091.
 127. Kim T, Park JC, Gach HM, Chun J, Mutic S. Technical Note: real-time 3D MRI in the presence of motion for MRI-guided radiotherapy: 3D Dynamic keyhole imaging with super-resolution [published online ahead of print 20190827]. *Med Phys.* 2019;46(10):4631-4638.
 128. Xie H, Lei Y, Wang T, et al. Synthesizing high-resolution magnetic resonance imaging using parallel cycle-consistent generative adversarial networks for fast magnetic resonance imaging. *Med Phys.* 2022;49(1):357-369.
 129. Szegedy C, Ioffe S, Vanhoucke V, Alemi A. Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning. 2016. doi:10.48550/ARXIV.1602.07261
 130. Cui Y, Yin FF. Impact of image quality on radiomics applications [published online ahead of print 20220722]. *Phys Med Biol.* 2022;67(15).
 131. Bouget D, Pedersen A, Jakola AS, et al. Preoperative brain tumor imaging: models and software for segmentation and standardized reporting [published online ahead of print 20220727]. *Front Neurol.* 2022;13:932219.
 132. Gustafsson CJ, Swärd J, Adalbjörnsson SI, Jakobsson A, Olsson LE. Development and evaluation of a deep learning based artificial intelligence for automatic identification of gold fiducial markers in an MRI-only prostate radiotherapy workflow [published online ahead of print 20201112]. *Phys Med Biol.* 2020;65(22):225011.
 133. Chakrabarty S, Sotiras A, Milchenko M, LaMontagne P, Hileman M, Marcus D. MRI-based identification and classification of major intracranial tumor types by using a 3D convolutional neural network: a retrospective multi-institutional analysis [published online ahead of print 20210811]. *Radiol Artif Intell.* 2021;3(5):e200301.
 134. Jalalifar SA, Soliman H, Sahgal A, Sadeghi-Naini A. Predicting the outcome of radiotherapy in brain metastasis by integrating the clinical and MRI-based deep learning features [published online ahead of print 20220706]. *Med Phys.* 2022;49(11):7167-7178.
 135. Nunez-Gonzalez L, van Garderen KA, Smits M, et al. Pre-contrast MAGic in treated gliomas: a pilot study of quantitative MRI [published online ahead of print 20221217]. *Sci Rep.* 2022;12(1):21820.
 136. Harrell FE Jr, Lee KL, Califf RM, Pryor DB, Rosati RA. Regression modelling strategies for improved prognostic prediction. *Stat Med.* 1984;3(2):143-152.
 137. Longato E, Vettoretti M, Di Camillo B. A practical perspective on the concordance index for the evaluation and selection of prognostic time-to-event models. *J Biomed Inform.* 2020;108:103496.
 138. Hossin MMNS. A review on evaluation metrics for data classification evaluations. *Int J Data Mining Knowledge Manage Process.* 2015;5:01-11.
 139. Chen Y, Villanueva-Meyer JE, Morrison MA, Lupo JM. Toward automatic detection of radiation-induced cerebral microbleeds using a 3D Deep Residual Network. *J Digit Imaging.* 2019;32(5):766-772.
 140. Gao Y, Xiao X, Han B, et al. Deep learning methodology for differentiating glioma recurrence from radiation necrosis using multimodal magnetic resonance imaging: algorithm development and validation [published online ahead of print 2021117]. *JMIR Med Inform.* 2020;8(11):e19805.
 141. Messina C, Bignone R, Bruno A, et al. Diffusion-weighted imaging in oncology: an update [published online ahead of print 20200608]. *Cancers (Basel).* 2020;12(6).
 142. Jing B, Deng Y, Zhang T, et al. Deep learning for risk prediction in patients with nasopharyngeal carcinoma using multi-parametric MRIs [published online ahead of print 20200802]. *Comput Methods Programs Biomed.* 2020;197:105684.
 143. Huisman SI, van der Boog ATJ, Cialdella F, Verhoeff JJC, David S. Quantifying the post-radiation accelerated brain aging rate in glioma patients with deep learning [published online ahead of print 20220810]. *Radiother Oncol.* 2022;175:18-25.
 144. Hua HL, Deng YQ, Li S, et al. Deep learning for predicting distant metastasis in patients with nasopharyngeal carcinoma based on pre-radiotherapy magnetic resonance imaging [published online ahead of print 20220919]. *Comb Chem High Throughput Screen.* 2022. doi:10.2174/1386207325666220919091210
 145. Jalalifar A, Soliman H, Sahgal A, Sadeghi-Naini A. A cascaded deep-learning framework for segmentation of metastatic brain

- tumors before and after stereotactic radiation therapy(). *Annu Int Conf IEEE Eng Med Biol Soc.* 2020;2020:1063-1066.
146. Jalalifar SA, Soliman H, Sahgal A, Sadeghi-Naini A. A self-attention-guided 3D deep residual network with big transfer to predict local failure in brain metastasis after radiotherapy using multi-channel MRI [published online ahead of print 20221104]. *IEEE J Transl Eng Health Med.* 2023;11:13-22.
 147. Xia W, Zhang K, Li M, et al. Impact of magnetic field on dose distribution in MR-Guided radiotherapy of head and neck cancer. *Front Oncol.* 2020;10.
 148. Rammohan N, Randall JW, Yadav P. History of technological advancements towards MR-Linac: the future of image-guided radiotherapy [published online ahead of print 20220812]. *J Clin Med.* 2022;11(16).
 149. Otazo R, Lambin P, Pignol J-P, et al. MRI-guided radiation therapy: an emerging paradigm in adaptive radiation oncology. *Radiology.* 2021;298(2):248-260.
 150. Edelstein WA, Mahesh M, Carrino JA. MRI: time is dose—and money and versatility. *J Am Coll Radiol.* 2010;7(8):650-652.
 151. Stankovic Z, Allen BD, Garcia J, Jarvis KB, Markl M. 4D flow imaging with MRI. *Cardiovasc Diagn Ther.* 2014;4(2):173-192.
 152. Nie X, Li G. Real-Time 2D MR Cine from Beam Eye's View with tumor-volume projection to ensure beam-to-tumor conformality for MR-guided radiotherapy of lung cancer [published online ahead of print 20220629]. *Front Oncol.* 2022;12:898771.
 153. Terpstra ML, Maspero M, Bruijnen T, Verhoeff JJC, Lagendijk JJW, van den Berg CAT. Real-time 3D motion estimation from undersampled MRI using multi-resolution neural networks [published online ahead of print 20211026]. *Med Phys.* 2021;48(11):6597-6613.
 154. Romaguera LV, Plantefève R, Romero FP, Hébert F, Carrier JF, Kadoury S. Prediction of in-plane organ deformation during free-breathing radiotherapy via discriminative spatial transformer networks [published online ahead of print 20200613]. *Med Image Anal.* 2020;64:101754.
 155. Driever T, Hulshof M, Bel A, Sonke JJ, van der Horst A. Quantifying intrafractional gastric motion using auto-segmentation on MRI: deformation and respiratory-induced displacement compared [published online ahead of print 20221224]. *J Appl Clin Med Phys.* 2022. doi:10.1002/acm2.13864:e13864
 156. Li S, Zhou B. A review of radiomics and genomics applications in cancers: the way towards precision medicine. *Radiat Oncol.* 2022;17(1):217.
 157. Nagrani A, Yang S, Arnab A, Jansen A, Schmid C, Sun C. *Attention Bottlenecks for Multimodal Fusion.*
 158. Hoff MN, McKinney A, Shellock FG, et al. Safety considerations of 7-T MRI in clinical practice. *Radiology.* 2019;292(3):509-518.
 159. Du R, Lee VH, Yuan H, et al. Radiomics model to predict early progression of nonmetastatic nasopharyngeal carcinoma after intensity modulation radiation therapy: a multicenter study. *Radiol Artif Intell.* 2019;1(4):e180075.
 160. Kerkmeijer LGW, Maspero M, Meijer GJ, van der Voort van Zyp JRN, de Boer HCJ, van den Berg CAT. Magnetic resonance imaging only workflow for radiotherapy simulation and planning in prostate cancer. *Clin Oncol.* 2018;30(11):692-701.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Eidex Z, Ding Y, Wang J, et al. Deep learning in MRI-guided radiation therapy: A systematic review. *J Appl Clin Med Phys.* 2024;25:e14155.
<https://doi.org/10.1002/acm2.14155>