



Applications of generative adversarial networks (GANs) in radiotherapy: narrative review

Zhixiang Wang^{1#}, Glauco Lorenzut^{1#}, Zhen Zhang^{1,2}, Andre Dekker¹, Alberto Traverso¹

¹Department of Radiation Oncology (Maastr), GROW-School for Oncology, Maastricht University Medical Centre+, Maastricht, The Netherlands;

²Department of Radiation Oncology, Tianjin Medical University Cancer Institute and Hospital, National Clinical Research Center for Cancer, Key Laboratory of Cancer Prevention and Therapy, Tianjin's Clinical Research Center for Cancer, Tianjin, China

Contributions: (I) Conception and design: Z Wang, G Lorenzut; (II) Administrative support: A Dekker, A Traverso; (III) Collection and assembly of data: Z Zhang; (IV) Data analysis and interpretation: Z Wang, Z Zhang; (V) Manuscript writing: All authors; (VI) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work and should be considered as co-first authors.

Correspondence to: Alberto Traverso. Department of Radiation Oncology (Maastr), GROW-School for Oncology, Maastricht University Medical Centre+, Maastricht, The Netherlands. Email: alberto.traverso@maastro.nl.

Background and Objective: Radiation therapy (RT) is the dominant method for clinical cancer treatment, which aims to ensure that planning target volume (PTV) receives a sufficient dose while organs-at-risk (OARs) are exposed to little or no radiation. However, obtaining a clinically acceptable radiotherapy plan often requires a long time, tedious work, and a high level of physician experience. The general steps to perform RT include planning [computed tomography (CT)/magnetic resonance imaging (MRI)/positron emission tomography (PET)] image acquisition, contouring the treatment area (gross tumor volume, OARs, etc.), and developing a treatment plan and treatment implementation. But there are still some challenges that need to be overcome. Fortunately, with the development of the computer science, Generative Adversarial Network (GAN) which is composed of a generator and discriminator with opposing optimized goals has been widely used by an increasing number of applications in various fields, especially in CT, MRI, and other images and plays a great role in RT. This review aims to provide an up-to-date snapshot of GAN applications in radiotherapy.

Methods: We searched for studies published from January 2018 to March 2022, with English language restrictions on PubMed and IEEE Xplore databases.

Key Content and Findings: (I) GAN is an active field of research to support the automation of the radiotherapy workflow; (II) clinical validation is still needed to guarantee the usability of GANs in radiotherapy.

Conclusions: GAN model has already been widely used in RT. Thanks to their ability to automatically learn the anatomical features from different modalities images, improve quality images, generate synthetic images and make less time consumption automatic dose and plan calculation. Even though the GAN model cannot replace the radiotherapy doctors' work, it still has great potential to enhance the radiologists' workflow. There are lots of opportunities to improve the diagnostic ability and decrease potential risks during radiotherapy and time cost for plan calculation.

Keywords: Generative Adversarial Network (GAN); radiotherapy; applications

Received: 13 June 2022; Accepted: 17 October 2022; Published: 30 December 2022.

doi: [10.21037/pcm-22-28](https://doi.org/10.21037/pcm-22-28)

View this article at: <https://dx.doi.org/10.21037/pcm-22-28>

Introduction

Radiation therapy (RT) is the most used method for cancer treatment, which aims to deliver the prescribed dose to the planning target volumes (PTVs), while simultaneously reducing at minimum the dose to organs-at-risk (OARs) (1). Obtaining a clinically acceptable RT treatment plan often requires a long time, tedious work, and a high level of physician/technician experience (2).

The general steps to perform RT planning include [computed tomography (CT)/magnetic resonance imaging (MRI)/positron emission tomography (PET)] image acquisition, contouring of treatment area (gross tumor volume, OARs, etc.), treatment plan optimization, and treatment delivery.

However, there are still some challenges: first, sometimes a diagnosis is performed on magnetic resonance (MR) scans because of better soft-tissue contrast, but a CT is always required to make a plan. Second, contouring and treatment planning is time-consuming and dependent on expertise which is prone to inter-observer variability. Finally, in-room imaging is of a lower quality than diagnostic CT.

In the last 10 years, the RT research community has focused on optimizing and automizing the above-presented steps by using artificial intelligence (AI). With the development of computer science, deep learning (DL) algorithms, a branch of AI, are widely applied by researchers to solve the above-mentioned issues. Generative Adversarial Networks (GANs), a subset of DL, became popular in the medical imaging domain, mainly for synthetic data generation (2). Since GAN was proposed in 2014 by Ian Goodfellow (3), it has been widely used by an increasing number of applications in standard of care medical imaging, especially in CT, MRI, and plays a great role in RT (4). A deep understanding of GANs requires specific knowledge of computer science, often not available at RT clinics.

Therefore, in this review, we will introduce the development of the GAN models, their structures, their improvements, and their applications in RT which can help the researchers have a preliminary knowledge about these DL models.

For readers interested in a specific application of GANs, we have grouped the GAN applications into three clusters: CT translation and synthesis (see later GANs for synthetic imaging), dose and plan calculation (see later GANs for dose and plan calculation), and image quality improvement (see later GANs for quality improvement). Finally, we have discussed the limitations and future directions to

give some hints for the following researchers who want to develop GAN applications in RT. We present the following article in accordance with the Narrative Review reporting checklist (available at <https://pcm.amegroups.com/article/view/10.21037/pcm-22-28/rc>).

Methods

The search was performed from PubMed and IEEE Xplore datasets according to multiple keyword combinations and the related MeSH terms including: “Radiotherapy”, “generative adversarial network (GAN)”, and “application”. January 1, 2018, was set as the cut-off date because we only considered the research within 5 years. The inclusion criteria were: original research articles (proceedings included), English language, and the development of a GAN model using a RT dataset. One hundred publications were extracted according to the search string above. Two researchers with expertise in DL, quickly scanned the abstracts to exclude irrelevant articles. Subsequently, we scanned the reference list from the selected articles to include the related ones which were not found by the initial search. Finally, 23 articles refer to the applications of GANs in radiotherapy, most of which were appropriately referenced in this review (*Table 1*).

Development of generative adversarial models

GANs are inspired by game theory. The basic structure of the GANs model is shown in *Figure 1A*. It is composed of a generator G and discriminator D and it aims to optimize these components alternately according to the Minimax game logic until they can't beat each other (Nash equilibrium).

The generator G takes as input a vector z obeying the standard normal distribution $N(0,1)$ and creates the target data distribution $G(z)$. The goal of generator G is to synthesize new data in such a way that the discriminator D cannot distinguish it from a real one. The discriminator D can be seen as a classification network that distinguishes whether this new input data is real or not.

The Nash equilibrium is reached when the generator G synthesizes data $G(z)$ hard to distinguish from the real ones and the discriminator D can classify real and fake data with high precision.

In the process of training, indeed, the update of the generator G tries to make the synthetic data classified as the real ones, so that the synthetic data is closer to the decision

Table 1 The search strategy summary

Items	Specification
Date of search	April 14, 2022
Databases and other sources searched	PubMed & IEEE Xplore
Search terms used	“Radiotherapy”, “generative adversarial network (GAN)”, and “application”
Timeframe	January 1, 2018–April 14, 2022
Inclusion and exclusion criteria	Original research articles (proceedings included), English language, and the development of a GAN model using a RT dataset are included
Selection process (who conducted the selection, whether it was conducted independently, how consensus was obtained, etc.)	Two researchers with expertise in deep learning, quickly scanned the abstracts to exclude irrelevant articles
Any additional considerations, if applicable	The reference list from the selected articles to include the related ones which were not found by the initial search

RT, radiation therapy.

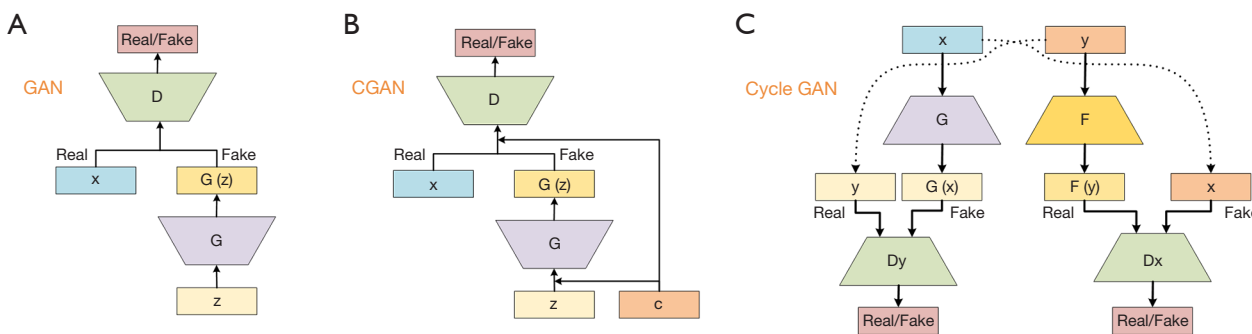


Figure 1 The structure of GANs model. (A) GAN; (B) CGAN; (C) Cycle GAN. G and D are respectively the generator and discriminator. x and y present the real data come from different dataset. z present random noise. G and F are the Generator. D, Dx, and Dy are discriminators. c is the class information of x. GAN, Generative Adversarial Network; CGAN, Conditional GAN; Cycle GAN, Cycle-consistent Generative Adversarial Networks.

boundary and the real image. While, the discriminator D plays the role of a binary classifier, and each update of the discriminator enhances its ability to distinguish between real data and synthetic data, which means dividing the correct decision boundary between the two kinds of data. With the continuous training of alternate iterations, the synthetic data will close to the real image, which will eventually make them indistinguishable to the discriminator D, so that the generator G can fit the real data with a high degree of realism.

However, the basic GAN generates data from noise so there are still some shortcomings: first, the class of the generated data cannot be controlled and, secondly, the transferring between two different clinical imaging

modalities cannot be done by a basic GAN. To overcome these problems, some improvements to GAN based model were made.

Conditional GAN (CGAN)

In the training processing of GAN, the random noise Z is used as a priori information in the comparison training process, which greatly improves the calculation efficiency when the amount of data is too large. However, too much random noise will lead to the uncontrollability of the training process and experimental results, which greatly reduces the accuracy of the network. To solve this problem, supervised learning or semi-supervised learning is added

based on GAN to effectively restrict the generation process and increase the stability of the network during training. CGAN is such an improved model (1). The CGAN structure shown in *Figure 1B*.

Cycle-consistent Generative Adversarial Networks (Cycle GAN)

In the real world, it is difficult to obtain a large amount of paired image data that arises from the same individual at different modalities or machines. Therefore, *Zhu et al.* proposed the Cycle GAN in 2017 to solve the problem of converting the images between different modalities with unpaired data (2).

The Cycle GAN consists of two identical GAN models with a generator and discriminator respectively. The generator is trained for getting a mapping between data source distribution x and y .

The discriminators are the same as the traditional GAN model to determine whether the data is real or synthetic. The structure of the Cycle GAN structure is shown in *Figure 1C*.

x , z , and $G(z)$ represent respectively the real data, the input data, and the synthetic data generated by generator G .

For *Figure 1C*, F and G present two generators that generate fake data from y and x . The D_x and D_y present the discriminators for distinguishing between real and fake data created by the generators.

Evaluation metrics

The evaluation metrics needed to evaluate the quality of the synthetic data in GAN-based radiotherapy applications are divided into three different groups: image similarity, dose performance, and plan evaluation (shown in *Table 2*). They can be selected according to the specific RT tasks.

Evaluation metrics for calculation similarity between synthetic data and target

Normally, it is difficult for human eyes to evaluate the similarity between synthetic images and the target ones. These metrics can quantify the similarity between them are shown in *Table 2* image-related metrics.

The mean squared error (MSE) and mean average error (MAE) are the metrics that refer to the expected value of the difference between the synthetic and target data. The higher score means the bigger difference between them. The peak signal to noise ratio (PSNR) represents the ratio

of the energy of the peak signal to the average energy of the noise. The higher the score, the smaller the distinction between target and synthetic data.

The above metrics only calculate the gap between one-to-one correspondence pixels without considering the other positions. This treats the image as isolated pixels, while ignoring visual features, especially the local structural information. The structural information has great influence on the subjective evaluation of medical images.

Conversely, to address above shortage, the structural similarity index measure (SSIM) considers a region of pixels when calculating the difference between two images. When the two images are identical, the value of SSIM is equal to 1.

To align more the quantitative evaluation of image similarity to the visual inspection, the learned perceptual image patch similarity (LPIPS) (5) was proposed also known as perceptual loss. It is used to measure the difference between two images in subject feeling contains rich image information such as texture color and texture primitives. The lower the value of LPIPS, the more similar the two images are.

Evaluation metrics for synthetic dose performance

This type of evaluation is different from the evaluation method which compares the similarity between the real target and the synthetic data. The similarity metrics are not the best method considering the evaluation method between synthetic and target doses. Based on the aforementioned motivation, to evaluate the performance of the synthetic dose, the commonly used methods are shown in *Table 2* (dose-related metrics).

Dose-volume histogram (DVH) difference compares the difference of the DVH in RT planning between the generated and the real one. The lower the score, the more realistic the generated doses are. Moreover, the Hausdorff distance also can be used to calculate the difference between the synthetic DVH and the real one.

Evaluation metrics for plan evaluation

For evaluation of the feasibility of the generated plan, a comparison between the real and generated plan is required (shown in *Table 2*, plan-related metrics). Furthermore, the conformity index (CI) and homogeneity index (HI) scores that can evaluate the conformity and uniformity of dose distribution also should be considered. Averaged prediction error (APE) calculates the averaged ratio of the prescription dose and the difference between the ground truth and the prediction (4).

Table 2 Evaluation metrics for GAN applications

Category	Metrics	Full name	Definition	Function
Image	ME	Mean error	The average difference between the estimated values and the actual value	$ME = \frac{1}{n} \sum_{i=1}^n f_i - y_i$
Image	MSE	Mean square error	The average squared difference between the estimated values and the actual value	$MSE = \frac{1}{n} \sum_{i=1}^n (f_i - y_i)^2$
Image	MAE	Mean absolute error	The average absolute difference between the estimated values and the actual value	$MAE = \frac{1}{n} \sum_{i=1}^n f_i - y_i $
Image	MRE	Mean relative error	The ratio of the mean absolute error to the mean value of the quantity being measured	$MRE = \frac{1}{n} \sum_{i=1}^n \frac{ f_i - y_i }{y_i}$
Image	SNU	Spatial Nonuniformity	The maximum and minimum percentage differences from the mean irradiance	$SNU = \frac{\overline{HU}_{max} - \overline{HU}_{min}}{1000} \times 100\%$
Image	PSNR	Peak signal to noise ratio	The ratio between the maximum value of an image and the value of corrupting noise affects the fidelity of its quality	$PSNR = 10 \log_{10} \left(\frac{MAX_i}{MSE} \right)$
Image	SSIM	Structural similarity metric	The method to evaluate the quality of images	$SSIM = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}$
Image	NCC	Normalized cross correlation	The normalized inverse Fourier transform of the convolution of the Fourier transform of two images	$NCC = \frac{cov(I_0, I)}{\sigma_{I_0} \sigma_I}$
Image	LPIPS	Learned Perceptual Image Patch Similarity	The method to measure the perceptual difference between two images	$d(x, x_0) = \sum_1 \frac{1}{H_1 W_1} \sum_{h,w} \left\ w_1 \square (y_{hw}^1 - y_{0hw}^1) \right\ _2^2$
Dose	DVH difference	Dose-volume histogram difference	Difference between the DVH is a histogram relating radiation dose to tissue volume in radiation therapy planning	$DVH \text{ difference} = \sum_{i=1}^n DVH_{xi} - DVH_{yi}$
Dose	Hausdorff distance	-	The method measures the distance between two-point sets	$h(A,B) = \max(h(A,B), h(B,A))$ $h(A,B) = \max_{a \in A} \left\{ \min_{b \in B} \ a - b\ \right\}$ $h(B,A) = \max_{b \in B} \left\{ \min_{a \in A} \ b - a\ \right\}$
Plan	CI	Conformity index	The method to quantitatively assess the quality of radiotherapy treatment plans, and represents the relationship between isodose distributions and target volume	$CI = \frac{V_{RI}}{TV}$
Plan	HI	Homogeneity Index	The method to calculate the uniformity of dose distribution in the target volume	$HI = D_{2\%} - D_{98\%}$

f_i is the pixel value of the synthetic image. y_i is the pixel value of the target image. MAX_i is the maximum possible pixel value of the image. x is the target images, y is the synthetic images. μ_x, μ_y are the mean value of x and y . σ_x^2 is the variance of x . σ_y^2 is the variance of y . σ_{xy} is the covariance of x and y . c_1 and c_2 are two variables to stabilize the division with weak denominator $c_1=(k_1L)^2$, $c_2=(k_2L)^2$, L is the dynamic range of the pixel-values, $k_1=0.01$ and $k_2=0.03$ by default. $cov(.,.)$ presents covariance, σ_{I_0} and σ_I are the standard deviation of target and generated images, respectively. $h(A,B)$ and $h(B,A)$ calculate the maximum distance between point groups. V_{RI} is the reference dose volume, and TV is the target volume. $D_{2\%}$ and $D_{98\%}$ are the percentage dose to 2% and 98% target volume. DVH_{xi} and DVH_{yi} present the dose-volume histogram difference distribution of two sets. \overline{HU}_{max} and \overline{HU}_{min} are the averaged maximum and the minimum intensity of region of interests' data. n , the amount of the pixel in the images. ME, mean error; MSE, mean square error; MAE, mean absolute error; MRE, mean relative error; SNU, spatial nonuniformity; PSNR, peak signal to noise ratio; SSIM, structural similarity metric; NCC, normalized cross correlation; LPIPS, learned perceptual image patch similarity; DVH, dose-volume histogram; CI, conformity index; HI, homogeneity index.

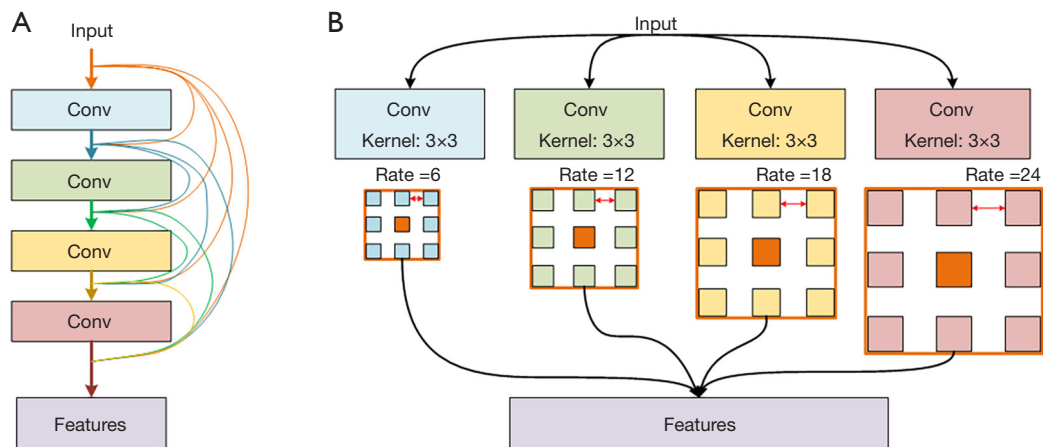


Figure 2 The structure of blocks. (A) Dense block; (B) ASPP block. Conv presents the convolutional layer. ASPP, atrous spatial pyramid pooling.

Generative adversarial models for CT translation and synthesis

Different modalities of medical images can provide multimodal information, that can be used for a better diagnosis and RT planning.

However, in a realistic situation, limitations due to unnecessary costs and radiation protection of the patient, make it hard to collect all the desired imaging modalities from a single patient. Fortunately, while there is a different focus and range distribution between modalities, there is still some hidden information in one type of image that may prevent the need to take another one. This is why the cross-mode image synthesis method is feasible (6).

For treatment planning, CT is always required, while delineations are often performed on MR, for example for pelvic or head and neck tumors. However, the transformation from MR to CT will lead to an undesirable 2–5 mm systematic error (7).

To address the systematic error, MR-only treatment planning was proposed, in which only MRI is required as the sole input modality. It can protect the patient from CT radiation doses and benefit a pediatric patient who has less dose upper (8).

The GAN-based method has the feasibility of mapping the information and generating the image from different modalities. GAN can generate synthetic CT (sCT) from MRI, thanks to allow performing the calculation of the dose accuracy with a single MRI-only workflow.

For sCT methods, Liu *et al.* [2019] tried to integrate dense block into three-dimensional (3D) Cycle GAN to

effectively generate the CT from T2-weighted MRI. Dense block connects all blocks that make up the model directly into each other, leading to each block gets additional inputs from all previous blocks and passes on its own output to all subsequent blocks. This ensures the maximum transmission of information between blocks in the model (see in *Figure 2A*). The proposed method achieved 51.32 ± 16.91 Hounsfield unit (HU) and less than 1% DVH difference compare to the one generated by real CT. This demonstrates the feasibility of GAN-based applications for the development of the MRI-only workflow for prostate proton radiotherapy (9).

To evaluate whether the sCT is accurate enough for MRI-only treatment planning, Kazemifar *et al.* [2020] used mutual information (MI) which is used to evaluate non-linear relations between two variables as the loss function in GAN to overcome the misalignment between two modalities in training model processing. In fact, one of the largest issues when translating MR to CT is that the two images belong to different spaces: the frequency one and the tissue density respectively. The above methods achieve a mean absolute difference below 0.5% (0.3 Gy) for the prescription dose for CTV and below 2% (1.2 Gy) for OARs. The excellent result illustrates that GAN is a potentially promising method to generate sCT for MRI-only treatment planning of patients with brain tumors in intensity-modulated proton therapy (10,11).

Zhang *et al.* [2018] proposed a Cycle- and Shape-Consistency GAN to synthetic realistic looking 3D images using unpaired data and improve the volume segmentation by generated data. With the extensive experiment on a 4,496

CT and MRI dataset, it proves that both tasks are beneficial to each other, and coupling them has better performance than exclusively (12).

In another study, Olberg *et al.* [2019] designed an atrous spatial pyramid pooling (ASPP) structure in GAN model. ASPP is a structure that captures objects and image features on multiple scales, thanks to the introduction of multiple filters that have complementary effective fields of view (as shown in *Figure 2B*).

The proposed method achieved a root mean square error (RMSE) (17.7 ± 4.3), a SSIM (0.9995 ± 0.0003), a PSNR (71.7 ± 2.3), and great dose performance based on sCT with more than 98% passing rates in the 1,042 images test set. The excellent result illustrates the designed structures can improve the performance of traditional GAN (13).

Meanwhile, different from traditional sCT which is generated from a single MR sequence, Koike *et al.* [2020] tried to generate and assess the feasibility of sCT from multi-sequence MRI using GAN for brain radiotherapy treatment planning. With the small, clinically negligible difference (less than 0.1% DVH difference and 0.6 ± 1.9 mm overall equivalent path length difference), CGAN is feasible to generate sCT from multi-sequence include T1-weighted, T2-weighted and fluid-attenuated inversion recovery (Flare) MRI (1).

Bourbonne *et al.* [2021] was the first study which demonstrate the GAN-generated CT from diagnostic brain MRIs have comparable performance to initial CT for the planning of brain stereotactic RT. In their study, the two-dimensional (2D)-UNet was selected as the backbone of generator. Through experiments on a dataset of 184 patients, there were no significant statistical differences regarding International Commission on Radiation Units and Measurements (ICRU) 91 s endpoints, which means the sCT and initial CT has high similarity for both the organs at risk and the target volumes (14).

On the other hand, the time cost is also worth considering the feature of radiotherapy applications.

Maspero *et al.* [2018] tried to assess whether the GAN method can rapidly generate sCT to be used for accurate MR-based dose calculations in the entire pelvis. As the result, the CGAN required 5.6 s and 21 s for a single patient on graphics processing unit (GPU) and central processing unit (CPU), respectively. It achieves less than $\pm 2.5\%$ calculated DVH differences on sCT and CT. Results suggest that the sCT generation was sufficiently fast and accurate to be integrated into an MR-guided radiotherapy workflow (15).

The 8 key publications focusing on GAN application for CT translation and synthesis are presented in *Table 3*.

Generative adversarial models for dose and plan calculation

The RT planning is highly dependent on the clinical experience and skills of the radiotherapy physicist or dosimetrist, as well as their knowledge of radiotherapy physics and understanding of the treatment planning system (TPS). With advances in DL, especially GANs, automatically generating 3D RT dose distributions from medical images like CTs and MRIs became possible. In the past few years, several methods can generate dose distribution or intensity-modulated radiation therapy (IMRT) plans from different kinds of inputs.

To overcome a limited dataset situation for a DL task, Liao *et al.* [2021] proposed an Auxiliary Classifier GAN (ACGAN) to synthesize dose distribution according to tumor types and beam types. The proposed with excellent PSNR (75.6032) and multi-scale SSIM (MS-SSIM) (0.95120) results, that demonstrate the synthetic dose distribution is close to the real one which can be used for increasing the training set for dose prediction tasks (16).

In order to different organs to jointly constrain the dose distribution of each organ in model training to achieve better PTV dose coverage and OARs sparing. Cao *et al.* [2021] (4) designed an adaptive multi-organ loss (AML)-based Generative Adversarial Network (AML-GAN). The AML loss can measure the gap between synthetic dose and real one on whole dose, OAR and PTV distribution. The experiment demonstrates the proposed method achieves state-of-art APE (0.021 ± 0.014) in terms of OARs and PTV (16).

Dose calculation is a time-consuming task, which sharply decreases the RT workflow efficiency. Therefore, some GAN-based dose simulation methods to decrease the time cost and generate accurate dose distribution were proposed.

In another work, Li *et al.* [2021] designed a CGAN-based model which can real-time generate fluence map from CT. This model containing a novel PyraNet that implements 28 classic ResNet blocks in pyramid-line concatenations as the generator. The proposed method was evaluated on 15 plans, the AI model only cost 3 s to predict a fluence map without statistical significance from the real one. This approach holds great potential for clinical applications in real-time planning (17).

While Zhang *et al.* [2021] proposed a discovery cross-domain GAN (DiscoGAN) to generate comparable accuracy

Table 3 Key publications focusing on GAN application for CT translation and synthesis

Authors	Year	Country	Dimension	Model information	Target	Region	No. of patient	Evaluation method	Conclusion
Yingzi Liu	2019	USA	3D	Cycle GAN	T2-weighted MRI to CT	Pelvic	17	MAE, DVH difference	We applied a novel learning-based approach to integrating dense-block into Cycle GAN to synthesize pelvic sCT images from routine MR images (Lei <i>et al.</i> 2019) for potential MRI-only prostate proton therapy. The proposed method demonstrated a comparable level of precision in reliably generating sCT images for dose calculation, which supports further development of MRI-only treatment planning. Unlike photon therapy, the accuracy of proton dose calculation is highly dependent on stopping power rather than HU values. Therefore, the future directions of MRI-only proton treatment planning include prediction of the stopping power map based on the MR images and generating elemental concentration maps that can be used for Monte Carlo simulations
Samaneh Kazemifar	2019	USA	2D	GAN	MRI to CT	Brain	77	MAE	In conclusion, MRI-only treatment planning will reduce radiation dose, patient time, and imaging costs associated with CT imaging, streamlining clinical efficiency and allowing high-precision radiation treatment planning. Despite these advantages, several challenges prevent clinical implementation of MRI-only radiation therapy. Through the method we have proposed here, synthetic CT images can be generated from only one pulse sequence of MRI images of a range of brain tumors. This method is a step toward using artificial intelligence to establish MRI-only radiation therapy in the clinic
Samaneh Kazemifar	2020	USA	2D	GAN with MI as the loss function	MRI to CT	Brain	77	DVH difference	This work explained the feasibility of using sCT images generated with a deep learning method based on GANs for intensity-modulated proton therapy. We tested the method in brain tumors—some of them located close to complex bone, air, and soft-tissue interfaces—and obtained excellent dosimetric accuracy even in those challenging cases. The proposed method can generate sCT images in around 1 s without any manual pre- or post-processing operations. This opens the door for online MRI-guided adaptation strategies for IMPT, which would eliminate the dose burden issue of current adaptive CT-based workflows, while providing the superior soft-tissue contrast characteristic of MRI images

Table 3 (Continued)

Table 3 (Continued)

Authors	Year	Country	Dimension	Model information	Target	Region	No. of patient	Evaluation method	Conclusion
Zizhao Zhang	2018	USA	3D	Cycle and Shape-Consistency GAN	MRI to CT and segmentation task	Heart	4,496 images	Segmentation score	We present a method that can simultaneously learn to translate and segment medical 3D images, which are two significant tasks in medical imaging. Training generators for cross-domain volume-to-volume translation is more difficult than that on 2D images. We address three key problems that are important in synthesizing realistic 3D medical images: (I) learn from unpaired data, (II) keep anatomy (i.e., shape) consistency, and (III) use synthetic data to improve volume segmentation effectively. We demonstrate that our unified method that couples the two tasks is more effective than solving them exclusively. Extensive experiments on a 3D cardiovascular dataset validate the effectiveness and superiority of our method
Sven Olberg	2019	USA	2D	ASPP, GAN	MRI to CT	Breast	2,400	RMSE, SSIM, PSNR	In this study, we have evaluated the robustness of the conventional pix2pix GAN framework that is ubiquitous in the image-to-image translation task as well as the novel deep spatial pyramid framework we propose here. The proposed framework demonstrates improved performance in metrics of training time and image quality, even in cases when training data are limited. The success of the framework in sCT generation is a promising step toward an MR-only RT workflow that eliminates the need for CT simulation and setup scans while enabling online adaptive therapy applications that are becoming ever more prevalent in MR-IGRT
Yuhei Koike	2020	Japan	3D	CGAN	Multi-sequence Brain MRI to CT		580	DVH difference, clinically negligible difference	Images from multi-sequence brain MR images using an adversarial network. The performance of the model was evaluated by comparing the image quality and the treatment planning with those of the original CT images. The use of multiple MR sequences for sCT generation using cGAN provided better image quality and dose distribution results compared with those from only a single T1w sequence. The CT number of the generated sCT images showed good agreement with the original images, but not in the bone regions. Impacts on the dose calculations were within 1%. These findings demonstrate the feasibility and utility of sCT-based treatment planning and support the use of deep learning for MR-only radiotherapy

Table 3 (Continued)

Table 3 (Continued)

Authors	Year	Country	Dimension	Model information	Target	Region	No. of patient	Evaluation method	Conclusion
Matteo Maspero	2018	The Netherlands	2D	CGAN	MRI to CT	Abdominal, pelvic	91	DVH	To conclude, this study shows, for the first time, that sCT images generated with a deep learning approach employing a CGAN and multi-contrast MR images acquired with a single acquisition facilitated accurate dose calculations in prostate cancer patients. It was further shown that without retraining the network, the CGAN could generate sCT images in the pelvic region for accurate dose calculations for rectal and cervical cancer patients. A particularly attractive feature of our method is its speed as it allows sCT generation within 6 s on a GPU and within 21 s on a CPU. This could be of particular benefit for MRgRT applications
Yingzi Liu	2019	USA	3D	Cycle GAN	MRI to CT	Abdominal	21	MAE, DVH difference	We applied a novel learning-based approach to integrate dense-block into cycle GAN to synthesize abdominal sCT images from routine MR images for potential MRI-only liver proton therapy. The proposed method demonstrated a comparable level of precision in reliably generating sCT images for dose calculation, which supports further development of MRI-only treatment planning. Unlike photon therapy, the accuracy of proton dose calculation is highly dependent on stopping power rather than HU values. Therefore, the future directions of MR-only proton treatment planning include prediction of the stopping power map based on the MR images or generating elemental concentration maps that can be used for Monte Carlo simulations

ASPP, atrous spatial pyramid pooling; CGAN, conditional GAN; RMSE, root mean square error; SSIM, structural similarity metric; PSNR, peak signal to noise ratio; MAE, mean absolute error; MI, mutual information; DVH, dose-volume histogram; sCT, synthetic CT; MRgRT, MRI-guided radiotherapy; GPU, graphics processing unit; CPU, central processing unit; GAN, Generative Adversarial Network; IGRT, image-guided radiotherapy; Cycle GAN, Cycle-consistent Generative Adversarial Networks; RT, radiation therapy; CT, computed tomography; MRI, magnetic resonance imaging; 2D, two-dimensional; 3D, three-dimensional; HU, Hounsfield unit.

to Monte Carlo simulation without much time cost. The proposed method was evaluated on abdominal cases, thoracic cases, and head cases by mean relative error (MRE) and achieved no systematic deviation. It demonstrates the proposed method has great potential for accurate dose calculation compared to the Monte Carlo simulation method (18).

For relative stopping power (RSP) prediction application, Harms *et al.* [2020] used a Cycle GAN, relying on a

compound loss function designed for structural and contrast preservation, to predict RSP maps from cone beam CT (CBCT). With the result of a MAE (0.06 ± 0.01) and a ME (0.01 ± 0.01) between RSP generation from CT and CBCT, this method provides sufficiently accurate prediction which makes CBCT-guided adaptive treatment planning for IMPT become feasible (19).

The 5 key publications focusing on GAN application for dose and plan calculation and synthesis are presented in Table 4.

Table 4 Key publications focusing on GAN application for dose and plan calculation and synthesis

Authors	Year	Country	Dimension	Model information	Target	Region	No. of patient	Evaluation method	Conclusion
Wentao Liao	2021	China	2D	ACGAN	Synthesis of Radiotherapy dose	head and neck	110	MS-SSIM and PSNR	We proposed the Dose-ACGAN for data enhancement work of radiotherapy DL. The dose distribution of specified tumor category or beam number category can be customized successfully, and the desired dose distribution map can be customized by controlling two variables together. One purpose of Dose-ACGAN is to generate multi-classification dose distribution enhancement data, train and generate dose distribution data of specified tumor type or beam number type. Used for training dose data required by radiotherapy plan using AI model. The next step in future work is to introduce CT data, contour information and beam angle information to customize the dose distribution corresponding to the predicted CT contour. Provide other better ideas for automatic planning, by using the dose distribution data of normal and effective radiotherapy plans, a large number of high-quality tagged data can be generated, such as tumor types, beam types, etc. The reliability and accuracy of automatic dose prediction model for radiotherapy will be improved effectively. A further plan is to enhance the data in this paper for a comparative study of different AI of predicting dose tasks
Xinyi Li	2021	USA	4D	CGAN contains a novel PyraNet that implements 28 classic ResNet blocks in pyramid-line concatenations as generator	CT to IMRT planning	Oropharyngeal	231	CI, HI	In this work, an AI agent was successfully developed as a DL approach for oropharyngeal IMRT planning. Without time-consuming inverse planning, this AI agent could automatically generate an oropharyngeal IMRT plan for the primary target with acceptable plan quality. With its high implementation efficiency, the developed AI agent holds great potentials for clinical application after future development validation studies
Chongyang Cao	2021	China	3D	An AML-GAN	Automatic dose prediction of cervical cancer,	Cervical cancer	75	CI, OAR, APE	We propose an adaptive multi-organ loss based Generative Adversarial Network, namely AML-GAN, to predict the dose distribution map from CT images automatically. Innovatively, besides the global dose prediction loss, we have also considered the dose prediction losses of PTV and individual OAR separately, making sure that the predicted dose distributions of local areas are as accurate as possible. Extensive experiments demonstrate that our proposed AML-GAN outperforms all state-of-the-art approaches

Table 4 (Continued)

Table 4 (Continued)

Authors	Year	Country	Dimension	Model information	Target	Region	No. of patient	Evaluation method	Conclusion
Xiaoke Zhang	2021	China	3D	A DiscoGAN	Synthesis of Radiotherapy dose	Head, abdomen, thorax	36	MRE, MAE	We developed a novel machine learning model based on DiscoGAN for dose calculation in proton therapy, which offers comparable accuracy (below 5%) to MC simulation but of reduced computational workload. The relationship between MRE and other factors such as dose, beam energy and location within the beam cross-section was examined. The proposed DiscoGAN has proven effective in identifying the relationship among dose, SP and HU in three dimensions. If successful, our approach is expected to find its potential use in more advanced applications such as inverse planning and adaptive proton therapy
Joseph Harms	2020	Atlanta	2D	Cycle GAN	RSP prediction	Head-and-neck	23	MAE, ME, PSNR, SSIM	This work presents the use of a deep-learning algorithm for generation of RSP maps directly from cone-beam CTs. The proposed method closely matches the quantitative values of the planning CT and the geometric qualities of the daily CBCT. When used for dose calculation, the method shows strong agreement to a DIR-based method that is in clinical use for dose evaluation while patients are under treatment. The proposed method was validated on head-and-neck patient CT images, a particularly difficult image set to work with due to the presence of several soft tissue structures, changing body shapes, and the frequent presence of metal artifacts. However, the proposed method still produced and median MAE of around 0.06 when compared to the planning CT and a median structural similarity of 0.88. Gamma analysis between the proposed method and the DPCT method using 3% dose difference and 3 mm distance-to-agreement had an average passing rate around 96% showing that the method can be used for dose evaluation

AML, adaptive multi-organ loss; ACGAN, auxiliary classifier GAN; CGAN, conditional GAN; GAN, Generative Adversarial Network; MS-SSIM, multi-scale structural similarity index measure; PSNR, peak signal to noise ratio; CGAN, conditional GAN; CI, conformity index; HI, homogeneity index; AI, artificial intelligence; DL, deep learning; IMRT, intensity-modulated radiotherapy; OAR, organ at risk; APE, averaged prediction error; PTV, planning target volume; MRE, mean relative error; MAE, mean absolute error; RSP, relative stopping power; ME, mean error; CBCT, cone beam computed tomography; DiscoGAN, discovery cross-domain GAN; MC, MonteCarlo; DIR, deformable image registration; DPCT, dose plan on computed tomography.

Generative adversarial models for quality improvement

Traditional medical image enhancement methods are mainly used to improve medical images with low contrast, narrow dynamic range, uneven intensity distribution, and blurred edges. This is given by studying effective image enhancement

algorithms to improve the image quality of existing medical images, improve their resolution or emphasize the important texture information and suppress noise. After those passages, the images become more standard and suited for computer-aided diagnosis (CAD) systems.

Medical images can be acquired with different methods

according to the need that doctors have to treat a particular case. Moreover, medical image data is much more complicated than natural image data, and it is difficult to get detailed information directly on the original data. These characteristics mean that medical images have a relatively greater demand for image enhancement algorithms that are beyond the capabilities of traditional algorithms.

In these cases, the GAN-based models can improve the quality of medical images (e.g., denoising).

Several GAN-based methods are trained using paired data. Charyyev *et al.* [2022] designed a residual attention GAN to synthetic dual-energy CT (DECT) from single energy CT (SECT). The MAE, PSNR, and normalized cross correlation (NCC) were applied to evaluate the performance of the synthetic high- and low-energy CT was 36.9 HU, 29.3 dB, 0.96 and 35.8 HU, 29.2 dB, and 0.96, respectively. The proposed method has potential feasibility for proton radiotherapy by generating DECT from SECT (20).

In another study, Zhao *et al.* [2020] designed a supervised GAN with the Cycle-consistency loss, Wasserstein distance loss, and an additional supervised learning loss, named S-Cycle GAN, to synthesize full-dose PET (FDPET) from low-dose PET (LDPET). The model was evaluated in 10 testing datasets and 45 simulated datasets by NRMSE, SSIM, PSNR, LPIPS, SUVmax and SUVmean, and the results show this method achieves accurate, efficient, and robust performance (21).

A study by Lee *et al.* [2021] used Cycle GAN to synthetic kilovoltage CT (KVCT) from megavoltage CT (MVCT). With the excellent average MAE, RMSE, PSNR, and SSIM values were 18.91 HU, 69.35 HU, 32.73 dB, and 95.48, respectively. This Cycle GAN can improve the MVCT to KVCT while maintaining the anatomical structure in radiation therapy treatment planning (22).

For CBCT improvement, some methods can synthesize target images from unpaired data.

Uh *et al.* [2021] applied Cycle GAN to correct abdominal and pelvic CBCT between children and young adults in the presence of diverse patient sizes, anatomic extent, and scan parameters. The performance of the model has significantly outperformed performance in the 14 patients' test set (47 ± 7 versus 51 ± 8 HU; paired Wilcoxon signed-rank test, $P < 0.01$). This proposed method can decrease the impact of anatomic variations in CBCT images for proton dose calculation (23).

Park *et al.* [2020] designed a spectral blending technique to combine trained sagittal and coronal directions Cycle GAN to sCT from CBCT. The proposed method achieves better performance than the existing Cycle GAN on PSNR

(30.6027 versus 29.4991), NMSE (1.3442 versus 1.5874), and SSIM (0.8977 versus 0.8674) (24).

Liu *et al.* [2020] designed a self-attention Cycle GAN to sCT from CBCT. There is no significant different performance between the CT-based contours and treatment plans from sCT on MAE and DVH differences. The result indicates that the sCT from CBCT can be used for accurate dose calculation (25).

Except self-attention, Gao *et al.* [2021] proposed an attention-guided Cycle GAN which contains two equipped with attention module generators to generate attention mask. It can make generator pays attention to the important part of images to eliminate numerous artifacts. By training and testing on a dataset of 170 patients, the proposed method has similar quality with real CT in MAE (43.5 ± 6.69), SSIM (93.7 ± 3.88), PSNR (29.5 ± 2.36), mean and standard deviation (SD) HU values ($P < 0.05$). Besides that, sCT provided the highest gamma passing rates (91.4 ± 3.26) in dose calculation compared with GAN methods. These demonstrate that the proposed method can trained by unpaired data to generate high-quality CT from CBCT (26).

The 6 key publications focusing on GAN application for quality improvement are presented in *Table 5*.

Discussion

This review clusters the recent GANs application in RT articles into three groups: CT translation and synthesis, dose and plan calculation, and image quality improvement.

Among the included studies, in terms of treatment sites, the majority focused on the head and neck (10/23). This was followed by the abdomen (8/23) and the thorax (5/23).

In RT applications, compared with the traditional handcraft method, GAN brings a potential significant improvement. The GAN model was trained to attain target high-level information distribution rather than simple geometrical and texture deformations. That makes GAN capable of establishing a nonlinear mapping between two different modalities for image translation tasks such as MR and CT or CBCT and CT.

Besides the accuracy prediction, the GAN model has less time consumption compared with traditional methods such as Monte Carlo simulation. The GAN applications can only cost a few seconds per patient which is an unimaginable performance using traditional methods. This will enormously increase the efficiency of RT.

All the publications mentioned in this review prove that GAN applications have great performance in modality

Table 5 key publications focusing on GAN application for quality improvement

Authors	Year	Country	Dimension	Model information	Target	Region	No. of patient	Evaluation method	Conclusion
Serdar Charyyev	2022	USA	–	A residual attention GAN	SECT to synthetic DECT	Head-and-neck	70	PSNR, NCC, ME	We applied a novel deep learning- based approach, namely residual attention GAN, to synthesize sLECT and sHECT images from SECT images for potential applications in the clinic where a DECT scanner is not available. The proposed method demonstrated a comparable level of precision in reliably generating synthetic images when compared to ground truth, and noise robustness in derived SPR maps
Kui Zhao	2020	China	3D	A supervised Cycle GAN	LDPET to FDPET	Brain	109	NRMSE, SSIM, PSNR, LPIPS, SUVmax and SUVmean	In conclusion, we have introduced a novel deep learning based generative adversarial model with the cycle consistent to estimate the high-quality image from the LDPET image. The proposed S-Cycle GAN approach has produced comparable image quality as corresponding FDPET images by suppressing image noise and preserving structure details in a supervised learning fashion. Systemic evaluations further confirm that the S-Cycle GAN approach can better preserve mean and maximum SUV values than other two deep learning methods, and suggests the amount of dose reduction should be carefully decided according to the acquisition protocols and clinical usages
Dongyeon Lee	2021	Korea	2D	Cycle GAN	MVCT to KVCT	Prostate	11	Hausdorff distances, DVH difference, OAR	In this study, we developed a synthetic approach based on Cycle GAN to produce sKVCT images from MVCT images for applying MVCT to adaptive helical tomotherapy treatment. The proposed method generates clear CT images by including the anatomical features of MVCT images through a deep learning algorithm without an additional calibration process. The Cycle GAN employed in this study was optimized using augmented training data derived from a small number of CT images. The proposed method successfully enhanced the quality of the MVCT images, preserving the anatomical structures of MVCT and restoring the HU to values similar to those of KVCT, along with providing reduced noise and improved contrast. The MVCT can be better utilized for aligning both the patient setup for daily treatment and the dose re-calculations for the ART process by considering the distributions of the HU values of sKVCT images approach and those of the planning KVCT images
Jinsoo Uh	2021	USA	2D	Cycle GAN	Correct CBCT between children and young adults	Abdominal, pelvic	64	MAE and ME	Using both abdominal and pelvic images for training a single deep learning model and normalizing age-dependent body sizes helped mitigate the impact of anatomic variations in CBCT images. Delivered proton dose can be accurately estimated from the corrected CBCT for children and young adults with abdominal or pelvic tumors

Table 5 (Continued)

Table 5 (Continued)

Authors	Year	Country	Dimension	Model information	Target	Region	No. of patient	Evaluation method	Conclusion
Sangjoon Park	2020	Korea	2D	Cycle GAN	CBCT to CT	Lung	10	PSNR, NMSE, SSIM	In this paper, we proposed a novel unsupervised synthetic approach based on Cycle GAN to produce CT images from CBCT images, which requires only unpaired CBCT and CT images for training. The proposed method properly combined Cycle GAN and spectral blending technique, generating CT images by Cycle GAN and further reducing the artifacts from missing frequency problem by spectral blending. Our method outperforms the existing Cycle GAN-based method both qualitatively and quantitatively
Yingzi Liu	2020	USA	3D	Cycle GAN	CBCT to CT	Pancreas	30	MAE, DVH, SNU, NCC	The image similarity and dosimetric agreement between the CT and sCT-based plans validated the dose calculation accuracy carried by sCT. The CBCT-based sCT approach can potentially increase treatment precision and thus minimize gastrointestinal toxicity

ART, adaptive radiotherapy; SECT, single energy CT; DECT, dual energy CT; LDPET, low-dose PET; FDPET, full-dose PET; CT, computed tomography; PET, positron emission tomography; GAN, Generative Adversarial Network; PSNR, peak signal to noise ratio; NCC, normalized cross correlation; ME, mean error; sLECT, synthetic low-energy CT; sHECT, synthetic high-energy CT; SPR, stopping power ratio; NRMSE, normalized root mean squared error; SSIM, structural similarity metric; LPIPS, learned perceptual image patch similarity; SUV, standardized uptake value; MVCT, megavoltage CT; KVCT, kilovoltage CT; DVH, dose-volume histogram; OAR, organ at risk; HU, Hounsfield unit; sKVCT, synthetic kilovoltage CT; CBCT, cone beam CT; MAE, mean absolute error; NMSE, normalized mean square error; SNU, spatial nonuniformity; sCT, synthetic CT.

translation, dose calculation, and image quality improvement tasks by maintaining anatomical and functional information, which has great benefits in RT workflow.

However, the training of GAN models is a challenging task, as it contains two models (generator and discriminator) with their own opposite targets. It differs from traditional DL model which contain a clear target, such as classification model that can stop training when it can achieve high accuracy in validation sets.

Therefore, in order to train the model, the discriminator should train first to make it has a preliminary classification ability to recognize real data from just noise images. Thanks to this preliminary training it is then possible to train the generator and so the whole GAN model. Then, the final training (adding generator) can stop when the accuracy of discriminator retains 0.5 or the loss function of discriminator is unable to continue descent, that means the discriminator can no longer distinguish between real and synthetic data given by generator.

The setting of hyper parameters also significantly affects whether the model can be successfully trained. Important hyper parameters to take in consideration are batch size (BS), learning rate (LR), epochs and optimizer.

BS means the number of data feed to the model in per step of training, which should be the first determined hyper parameter. Too small and too large BS could cause problems of too long training time or difficult convergence of the model, respectively. Though the ratio of image size and GPU's memory will significantly limit the large of BS, the number around 10 is recommended as BS initial setting (27).

The epochs represent the number of times that model is trained on the whole data set. The larger epochs, the more time consumption is required for training. Therefore, 100 epochs were recommended to make the model get sufficient training without too much time consumption.

The optimizer is the algorithm that modifies the weights of the DL model during the training phase and the LR determines how much every iteration influences the weights of the model.

For the optimizer and LR, the adaptive moment estimation (Adam) optimizer and LR at $1e-4$ are recommend and widely used as initial settings.

And, it is worth noting that, addition to the hyper parameters of the GANs model, the physical difference (contrasts, scanners, and patients, etc.) which often not included during training also have a significant impact

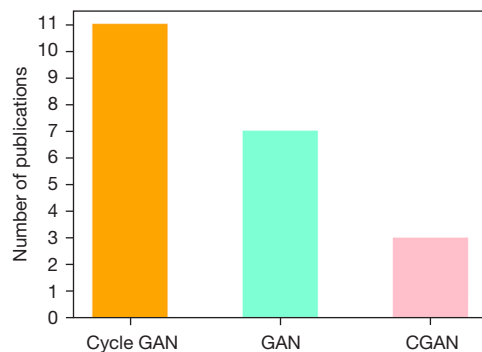


Figure 3 The distribution of the GAN applications in RT. GAN, Generative Adversarial Network; CGAN, Conditional GAN; Cycle GAN, Cycle-consistent Generative Adversarial Networks; RT, radiation therapy.

on the performance of the generated images. Especially for MR images, image contrast, or absolute image pixel values highly rely on scanning parameters and scanners, which makes the algorithm more difficult to train, less robustness, harder to migrate the model to data generated by other scanners, making it difficult to be widely used. Fortunately, there are many traditional and DL based image harmonization methods (adjusting the distribution of images from different sources so that they are close to each other) can overcome the above problem (28-30).

In order to develop and train a GAN model, there are many open-source DL framework can be selected such as Pytorch, Keras, Tensorflow and Caffe. Among them, Pytorch developed by Facebook is the most widely recommended framework by DL researchers.

Thank to convolutional neural networks (CNN), the GAN model can learn the specific tissue and organ textures from the training sets such as brain, breast, and pelvic which is impossible in the traditional handcraft method. It makes GAN automatic maintain useful anatomical and functional details to achieve excellent performance.

But there are still some potential risks for GAN applications in medical imaging tasks. sCT still remain some dimly visible artifacts in top and bottom areas in some applications (24). And GAN's tasks are highly dependent on data quality and quantity which make them have difficulty with nonstandard patient anatomies (7).

Although many studies (9,13,15,25,31,32) have conducted distribution comparison between synthetic and real one or validated their models in simulated clinical settings and have shown great potential to apply it in the real world. It still needs more evidence to apply it in the clinic, such as conducting clinical

trials or embedding it into TPSs to validate its application in daily clinical practice. Setting these exciting results aside, there are still some technical barriers that need to be overcome. For example, for MRI-based generation of sCT, organ effects such as organ motion (13,32) and organs containing cavities (15) will have an impact on the accuracy of the results or require manual intervention (11,13), which needs to be supported by more relevant studies, such as advances in deformable alignment techniques. Therefore, we consider that the application of GANs techniques requires a series of synergistic developments from a technical point of view to be accomplished.

Another important consideration is the choice of the GAN model, that is influenced by the typology of task it has to face. As the most common selected model (shown in *Figure 3*), the Cycle GAN has great advantages over the other two GAN models for unpaired data sets. It is composed of two independent GANs which makes obtaining two different modalities of information distribution become feasible without requiring a one-to-one correspondence. This will significantly decrease the difficulty for the researchers to build a large enough dataset for training. The basic GAN instead, is the second most selected model. It has the simplest structure which makes it the researchers easier to build their own model to address specific tasks. Finally, the CGAN with additional information as input can be used in the multi-class transformation tasks.

As mentioned above, when the researchers have paired data but do not require to generate the specific image according to the input class, the basic GAN was recommended as the start. Otherwise, the CGAN is your recommended selection. Finally, when only unpaired data, the Cycle GAN is the only choice for this task (27).

Limitations and future work

Though the GAN application has a strong power over image generation. There are still some limitations that need to be discussed. First, 3D GANs have a higher requirement in hardware which makes their deployment difficult in most hospitals, so, model compression or small models need to be considered in the future model design and deployment. Second, for plan and dose calculation, GAN-based applications have less time consumption compared with the traditional methods. However, most of the methods still cannot achieve real-time performance. So, the acceleration methods for GANs are still worth studying. Third, the GAN-based application has not been applied in the real-world clinical trial, which means it is still unclear how much the GAN can help radiotherapy

doctors. In this way, the GAN deployment in the real world needs to be done in future works.

Conclusions

In conclusion, the GAN model has already been widely used in RT. Thanks to their ability to automatically learn the anatomical features from different modalities images, improve quality images, generate synthetic images and make less time consumption automatic dose and plan calculation. Even though the GAN model still cannot replace the radiotherapy doctors' work, it still has great potential to enhance the radiologists' workflow. There are lots of opportunities to improve the diagnostic ability and decrease potential risks during radiotherapy and time cost for plan calculation.

Acknowledgments

We thank the Chinese Scholarship Council (CSC) for their financial support of studying abroad.

Funding: None.

Footnote

Reporting Checklist: The authors have completed the Narrative Review reporting checklist. Available at <https://pcm.amegroups.com/article/view/10.21037/pcm-22-28/rc>

Peer Review File: Available at <https://pcm.amegroups.com/article/view/10.21037/pcm-22-28/prf>

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at <https://pcm.amegroups.com/article/view/10.21037/pcm-22-28/coif>). AD reports grants from Varian, Janssen, Philips and BMS; royalties from Mirada Medical; personal fees from Medtronic, Janssen, Roche; and is board member of MD Anderson Advisory Board, Hanarth Fund Advisory Board and Peter Munk Advisory Board. The other authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Open Access Statement: This is an Open Access article

distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

1. 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;61:92-103.
2. Zhu JY, Park T, Isola P, et al. Unpaired image-to-image translation using cycle-consistent adversarial networks. 2017 IEEE International Conference on Computer Vision (ICCV); 22-29 October 2017; Venice, Italy. IEEE, 2017.
3. Goodfellow IJ, Pouget-Abadie J, Mirza M, et al. Generative Adversarial Networks. *arXiv:1406.2661v1 [stat. ML]*, 2014. doi: 10.48550/arXiv.1406.2661.
4. Cao C, Xiao J, Zhan B, et al. Adaptive Multi-Organ Loss Based Generative Adversarial Network for Automatic Dose Prediction in Radiotherapy. 2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI); 13-16 April 2021; Nice, France. IEEE, 2021.
5. Zhang R, Isola P, Efros AA, et al. The unreasonable effectiveness of deep features as a perceptual metric. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition; 18-23 June 2018; Salt Lake City, UT, USA. IEEE, 2018.
6. Sorin V, Barash Y, Konen E, et al. Creating Artificial Images for Radiology Applications Using Generative Adversarial Networks (GANs) - A Systematic Review. *Acad Radiol* 2020;27:1175-85.
7. Edmund JM, Nyholm T. A review of substitute CT generation for MRI-only radiation therapy. *Radiat Oncol* 2017;12:28.
8. Dougeni E, Faulkner K, Panayiotakis G. A review of patient dose and optimisation methods in adult and paediatric CT scanning. *Eur J Radiol* 2012;81:e665-83.
9. Liu Y, Lei Y, Wang Y, et al. Evaluation of a deep learning-based pelvic synthetic CT generation technique for MRI-based prostate proton treatment planning. *Phys Med Biol* 2019;64:205022.
10. Kazemifar S, McGuire S, Timmerman R, et al. MRI-only brain radiotherapy: Assessing the dosimetric accuracy

- of synthetic CT images generated using a deep learning approach. *Radiother Oncol* 2019;136:56-63.
11. Kazemifar S, Barragán Montero AM, Souris K, et al. Dosimetric evaluation of synthetic CT generated with GANs for MRI-only proton therapy treatment planning of brain tumors. *J Appl Clin Med Phys* 2020;21:76-86.
 12. Zhang Z, Yang L, Zheng Y. Translating and Segmenting Multimodal Medical Volumes with Cycle- and Shape-Consistency Generative Adversarial Network. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition; 18-23 June 2018; Salt Lake City, UT, USA. IEEE, 2018.
 13. Olberg S, Zhang H, Kennedy WR, et al. Synthetic CT reconstruction using a deep spatial pyramid convolutional framework for MR-only breast radiotherapy. *Med Phys* 2019;46:4135-47.
 14. Bourbonne V, Jaouen V, Hognon C, et al. Dosimetric Validation of a GAN-Based Pseudo-CT Generation for MRI-Only Stereotactic Brain Radiotherapy. *Cancers (Basel)* 2021;13:1082.
 15. Maspero M, Savenije MHF, Dinkla AM, et al. Dose evaluation of fast synthetic-CT generation using a generative adversarial network for general pelvis MR-only radiotherapy. *Phys Med Biol* 2018;63:185001.
 16. Liao W, Pu Y. Dose-Conditioned Synthesis of Radiotherapy Dose With Auxiliary Classifier Generative Adversarial Network. *IEEE Access* 2021;9:87972-81.
 17. Li X, Wang C, Sheng Y, et al. An artificial intelligence-driven agent for real-time head-and-neck IMRT plan generation using conditional generative adversarial network (cGAN). *Med Phys* 2021;48:2714-23.
 18. Zhang X, Hu Z, Zhang G, et al. Dose calculation in proton therapy using a discovery cross-domain generative adversarial network (DiscoGAN). *Med Phys* 2021;48:2646-60.
 19. Harms J, Lei Y, Wang T, et al. Cone-beam CT-derived relative stopping power map generation via deep learning for proton radiotherapy. *Med Phys* 2020;47:4416-27.
 20. Charyyev S, Wang T, Lei Y, et al. Learning-based synthetic dual energy CT imaging from single energy CT for stopping power ratio calculation in proton radiation therapy. *Br J Radiol* 2022;95:20210644.
 21. Zhao K, Zhou L, Gao S, et al. Study of low-dose PET image recovery using supervised learning with CycleGAN. *PLoS One* 2020;15:e0238455.
 22. Lee D, Jeong SW, Kim SJ, et al. Improvement of megavoltage computed tomography image quality for adaptive helical tomotherapy using cycleGAN-based image synthesis with small datasets. *Med Phys* 2021;48:5593-610.
 23. Uh J, Wang C, Acharya S, et al. Training a deep neural network coping with diversities in abdominal and pelvic images of children and young adults for CBCT-based adaptive proton therapy. *Radiother Oncol* 2021;160:250-8.
 24. Park S, Ye JC. Unsupervised Cone-Beam Artifact Removal Using CycleGAN and Spectral Blending for Adaptive Radiotherapy. 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI); 03-07 April 2020; Iowa City, IA, USA. IEEE, 2020.
 25. Liu Y, Lei Y, Wang T, et al. CBCT-based synthetic CT generation using deep-attention cycleGAN for pancreatic adaptive radiotherapy. *Med Phys* 2020;47:2472-83.
 26. Gao L, Xie K, Wu X, et al. Generating synthetic CT from low-dose cone-beam CT by using generative adversarial networks for adaptive radiotherapy. *Radiat Oncol* 2021;16:202.
 27. He F, Liu T, Tao D. Control batch size and learning rate to generalize well: Theoretical and empirical evidence. In: Wallach H, Larochelle H, Beygelzimer A, et al. *Advances in Neural Information Processing Systems 32 (NeurIPS 2019)*, 2019;32.
 28. Song S, Zhong F, Qin X, et al. Illumination harmonization with gray mean scale. *Computer Graphics International Conference*. Springer, 2020.
 29. Pinto MS, Paoletta R, Billiet T, et al. Harmonization of Brain Diffusion MRI: Concepts and Methods. *Front Neurosci* 2020;14:396.
 30. Jiang Y, Zhang H, Zhang J, et al. SSH: A self-supervised framework for image harmonization. 2021 IEEE/CVF International Conference on Computer Vision (ICCV); 10-17 October 2021; Montreal, QC, Canada. IEEE, 2021.
 31. Babier A, Mahmood R, McNiven AL, et al. Knowledge-based automated planning with three-dimensional generative adversarial networks. *Med Phys* 2020;47:297-306.
 32. Liu Y, Lei Y, Wang Y, et al. MRI-based treatment planning for proton radiotherapy: dosimetric validation of a deep learning-based liver synthetic CT generation method. *Phys Med Biol* 2019;64:145015.

doi: 10.21037/pcm-22-28

Cite this article as: Wang Z, Lorenz G, Zhang Z, Dekker A, Traverso A. Applications of generative adversarial networks (GANs) in radiotherapy: narrative review. *Precis Cancer Med* 2022;5:37.