Radiation Medicine and Protection 1 (2020) 171–178

Contents lists available at ScienceDirect

Radiation Medicine and Protection

journal homepage: www.radmp.org

Review

A review on 3D deformable image registration and its application in dose warping

Haonan Xiao, Ge Ren, Jing Cai *

Department of Health Technology and Informatics, The Hong Kong Polytechnic University, Hong Kong, China

ARTICLE INFO

Keywords:
Deformable image registration (DIR)
Deep learning
Dose summation
Dose accumulation

ABSTRACT

Deformable image registration (DIR) has been well explored in recent decades, and it is widely utilized in clinical tasks, especially dose warping. Nowadays, as deep learning (DL) develops rapidly, many DL-based methods were also applied in DIR. This paper reviews DL-based DIR methods in recent years and the application of DIR in dose warping. We collected and categorized the latest DL-based DIR studies. A thorough review of each category was presented, in which studies were discussed based on their supervision, advantage, and challenges. Then, we reviewed DIR-based dose warping and discussed its rationale, feasibility, successes, and difficulties. Lastly, we summarized the review on both parts and discussed their future development trend.

1. Introduction

Deformable image registration (DIR) is a common technique in the clinic, mainly applied in motion tracking and modeling,[1] image segmentation,[2] image-guided treatment,[3] and adaptive radiotherapy (ART).[4–10] In terms of transformation methods, algorithms can be grouped into rigid, affine, and deformable registration methods. Rigid and affine registration include limited degrees of freedom (6 for rigid and 12 for affine in 3D) and therefore, cannot account for complex and subtle transformation in medical images. DIR, also known as non-parametric registration, provides each voxel a corresponding vector, and the degrees of freedom rise to 3 times of the number of voxels. Therefore, DIR can deal with complex image registration tasks. Conventional DIR methods have already obtained huge success on registration accuracy, however, they usually take a long computation time and this limits its clinical application.[11] Recently, as artificial intelligence rapidly grows, some deep learning (DL)-based DIR methods were proposed, and they provided slightly worse or even comparable registration accuracy with much shorter implementation time.

DIR warps not only images but also other information attached to images, such as radiation dose. DIR-based dose warping includes dose summation and dose accumulation, and they are mainly applied to radiation treatments during which anatomical changes need managing. DIR-based dose summation usually accounts for larger anatomical changes, including re-treatment and external beam radiation therapy (EBRT) boosts after brachytherapy,[12,13] while DIR-based dose accumulation is primarily used in adaptive radiotherapy (ART) to account for daily bodyweight loss.[4–6,9,10,14–16] Such clinical application requires DIR methods to be not only accurate but also plausible, which raised bigger challenges. Currently, to overcome the problem of too long computation time for obtaining satisfactory deformation using iterative DIR methods, many ART protocols use rigid and affine registration only and suffer from inaccurate alignment,[17] or conduct image registration with graphic processing unit (GPU)-acceleration,[18] which can complete the computation within a minute but requires professional GPU cards that are usually not accessible in the clinic. Therefore, DL-based DIR methods have great potential in ART for their high registration accuracy and efficient implementation even if GPUs are not available.

In this review, the basic idea of conventional DIR methods was introduced and some latest progress, challenges, and possible future developments on DL-based DIR were introduced and discussed. Then the classical conventional DIR method was summarized, and the basic idea of convolutional neural network (CNN) and several main categories of CNN-based DIR methods were introduced. After the application of DIR on dose summation and accumulation was reviewed, the challenges and future development trend of CNN-based DIR methods and its application in dose wrapping were discussed.

2. Conventional DIR methods

DIR has been a hot research topic for some time, and various methods and toolboxes were proposed, including elastix,[19] advanced

* Corresponding author.
E-mail address: jing.cai@polyu.edu.hk (J. Cai).

https://doi.org/10.1016/j.radmp.2020.11.002
Received 3 October 2020; Received in revised form 20 November 2020; Accepted 20 November 2020
Available online 25 November 2020

normalization tools (ANTS),[20] demons,[21] and so on. When registering, these conventional methods define a cost function and achieve optimal transforms in each iteration by minimizing it. A general format of the cost function $\mathscr{C}$ is

$$\mathscr{C} = \frac{1}{N}\sum_{x} Sim(I_n, J) + Reg(\mathscr{T}_n), \quad (1)$$

where N is the total number of voxels (or pixels in 2 dimensional images). The cost function comprises a similarity metrics and a transform regularization term.

Commonly used similarity terms are the mean square difference (MSD) or mean square error (MSE), normalized cross-correlation (NCC), and mutual information (MI). MSE is sensitive to absolute values and is mainly used when registering images from the same modalities, i.e., CT-CT registration or MRI-MRI registration, while NCC and MI are mainly used for inter-modality registration, e.g., CT-MRI registration. Optimizing similarity term only often leads to an ill-posed problem and makes the transforms unreal,[21] since numerous transforms can give the same results. To obtain a smooth and reasonable transform, l_1 and l_2 norm of the transform and its gradient are usually included in the cost function as regularization terms to penalize overall transform strength and first-order or second-order gradient.

3. DL-based methods

3.1. The history of convolutional neural network

Machine learning has gained great success in computer vision. As one of the machine learning models, convolutional neural networks (CNN) becomes a hot topic and draws attention in various fields. Instead of taking the whole image as input, which is what artificial neural network (ANN) does, CNNs try to minimize the computational cost by using shift-invariant perceptrons in multiple layers. The idea of perceptrons could be traced back to brain studies in the 1950s and 1960s. Hubel and Wiesel[22] identified two types of visual cells in brains: simple cells (S-cells) and complex cells (C-cells). S-cells extract local features, while C-cells tolerate features' deformation, such as shifting and rotation. Every neuron takes the visual stimuli within a certain range, which is known as the receptive field. The receptive field of neighboring neurons are similar and overlapping, and the systematical combination of the size and location of receptive fields forms a complete visual space.

The first convolutional neural network was the time-delay neural network (TDNN) that was used to recognize phonemes in speech signals.[23] It utilized shared weights along the temporal dimension, making the speech processing time-invariant. This was the early prototype of the shift-invariant filters in image recognition. However, TDNN was a feed-forward network and all the weights had to be updated in a separate back-propagation process. LeCun et al[24] later proposed a CNN to recognize hand-written ZIP code numbers and made all the weights automatically trainable via back-propagation. Later, they further improved the network to a 7-layer CNN named as LeNet-5.[25] LeNet-5 had shown great performance on the recognition of hand-written numbers and was adopted by several banks to deposit checks. However, the input layer of LeNet-5 could only receive 32 by 32 black-and-white images. The processing of higher resolution and colorful images requires larger and more convolutional layers, so this network was constrained by computational cost.

The development of processors, especially GPUs, boosted the development of CNNs. In 2004, Oh and Jung[26] first implemented neural network training with GPUs instead of central processing units (CPUs) and got a 20 times faster performance. In 2012, Krizhevsky et al[27] proposed a deep CNN model, AlexNet. It was considered a classical CNN model and many recent CNN models are fine-tuning or modification of AlexNet. After the success of AlexNet in computer vision, researchers introduced CNNs into medicine, including computer-aided diagnosis, image segmentation, and image registration.

3.2. CNN in medical image segmentation and registration

Clinical tasks are usually repetitive and labor-intensive. Implementing artificial intelligence (AI) in medical tasks has been a popular topic for years. The computer system or AI was first introduced to diagnosis, which could be traced back to about 50 years ago.[28] In more recent years, as the CNN models have been mature, CNNs were implemented in various disease diagnosis and showed comparable or even higher accuracy than human experts.[29,30] Later, with more powerful CNNs, the model prediction could be in a higher dimension, e.g., an image or a displacement vector field (DVF), rather than Boolean variables. Therefore, CNNs were also implemented in image-wise tasks, i.e., image segmentation and registration. The applications of CNN in the medical image field was even more after the "encoder-and-decoder" architecture was proposed. U-net was the pioneer and a successful example of such architecture.[31] Another commonly used category of CNN architecture is generative adversarial networks (GANs) and its derivatives,[32] which comprise generators and discriminators and the two parts improve adversarially in training. GANs have obtained huge success in image synthesis and translation, super resolution, and registration.[33,34] Recently, as the development of hardware keeps going, CNNs can deal with 3D convolutional kernels and the input/output can be a 3D patch or whole volume. Therefore, some researchers also implemented CNNs in 3D medical image segmentation and registration.[11,33–37]

CNN-based registration methods could be categorized based on various properties, including CNN architecture (plain-CNN, ResNet, U-Net, GAN, etc.), supervision methods (unsupervised, weakly supervised, supervised, dual-supervised, etc.), input size (point sets, patches or whole images), and pre-processing procedures. In this review, DIR methods were classified into 3 categories, as shown in Fig. 1: models supervised with reference DVFs; models supervised with artificial DVFs; unsupervised models. In following sections, CNN-based DIR methods of the 3 categories were discussed based on the abovementioned properties.

3.2.1. Models supervised by reference DVFs

Supervised training is a straightforward idea and DL-based DIR models were started from it. The typical training process of DL-based DIR models with the supervision of reference DVFs is shown in Fig. 1: a moving volume and a fixed volume are fed to the CNN model and a DVF is predicted; a loss function is defined between the predicted DVF and reference DVF, and the parameters of the CNN model are optimized toward minimizing the loss function via gradient backpropagation.

Some DL-based reference DVF-supervised models were listed in Table 1. They mostly focused on brain MRI-MRI intra-modality registration. The reference DVFs were obtained from various conventional DIR methods, including ANTs, Syn, Demons, surface matching, and large deformation diffeomorphic metric mapping (LDDMM).[20,21,38,39] Despite the variance of the abovementioned conventional DIR methods, the registration accuracy of these methods made them eligible to be ground truth for deep learning model training.

The CNN architecture used in these works were mostly U-Net or U-Net with some modifications so that the output was of the similar or same size of the input. For example, Fan et al[45] implemented a U-Net with gap filling, which replaced the straight concatenations of U-Net with several convolutional layers. This made the network more complex and powerful. Cao et al[42,43] built an AlexNet-like model with three neurons representing dx, dy, and dz in the output layer. This model only predicted the motion at the center of a $15 \times 15 \times 15$ voxel cube and the final DVF was obtained by interpolation.

As for experiment implementation, although Rohé et al.[41] used whole images as input, they were segmented heart images instead of the entire chest region. Consequently, almost all these works registered small patches or a small portion of an image. The input size might be limited by the available GPU memory and the complexity of CNN models, varying

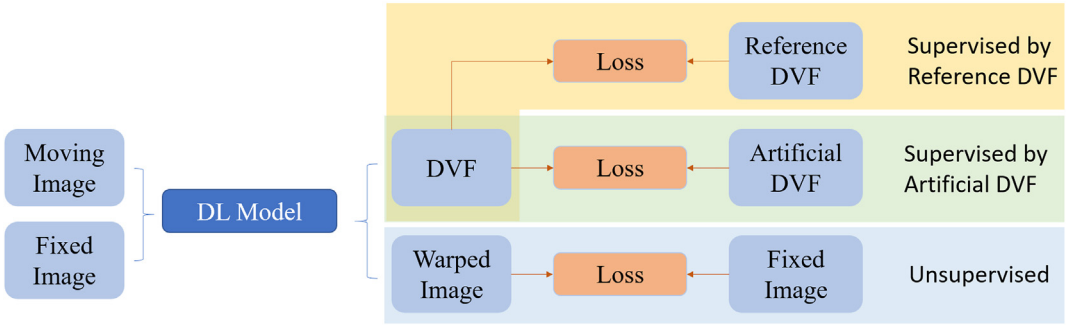

Fig. 1. Three categories of DL-based DIR methods.

Table 1
DL-based DIR models supervised by reference DVFs.

Authors	Publication Year	Region of interest (ROI)	Modality	Patch-based	Reference DVF Obtained by
Yang et al.[40]	2017	Brain	MR-MR	Yes	LDDMM
Rohé et al.[41]	2017	Heart	MR-MR	No	Surface matching
Cao et al.[42]	2017	Brain	MR-MR	Yes	Syn and Demons
Cao et al.[43]	2018	Brain	MR-MR	Yes	Syn and Demons
Onieva et al.[44]	2018	Lung	CT-CT	No	ANTs
Fan et al.[45]	2019	Brain	MR-MR	Yes	Syn and Demons

from 15 × 15 × 15 to 64 × 64 × 64 voxels. To avoid inter-patch discontinuity, the patched were overlapped with a certain pad. Loss functions used for supervision in these works were primarily MSE between the predicted and reference DVF, except Yang et al[40] chose the mean absolute error (MAE) as the loss function to better tolerant outliers. Fan et al.[45] added the MSE between the warped image and fixed image in the loss function to compensate for the imperfect reference DVF. To expand and fully utilize the dataset, they applied a data augmentation strategy, in which the moving images were warped by 20%, 40%, 60%, 80%, and 100% of the reference DVF to train the CNN model with different motion amplitude. Onieva et al.[44] utilized reinforcement learning strategy and reuse the best and worst prediction cases after each epoch.

Reference DVFs were only required in the training process and only image pairs were needed in testing or actual implementation, which makes the implementation efficient and convenient. The computation time could be reduced from tens of minutes in conventional DIR methods to seconds with CNNs and GPU available. In terms of registration performance, all these works were evaluated on public datasets and showed comparable or slightly higher image similarity and structure overlapping compared to classical methods, including Syn and Demons.

3.2.2. *Models supervised with artificial DVF*

Since obtaining reference DVFs could be time-consuming, and the reference DVFs did not always represent the actual motion, some researchers turned to supervise CNN models with artificial DVFs, as shown in Fig. 1. The primary difference is that DL models are supervised with artificial and maybe not plausible DVFs instead of reference DVFs. DL-based artificial DVF-supervised models were listed in Table 2.

The key idea was to generate known DVFs with different spatial frequencies as ground truth to deform the images and train the DL model to learn the artificial DVFs from the input images. The advantage of using artificial DVFs was that these DVFs perfectly described the deformation between the image pairs, eliminating the uncertainties introduced by imperfect reference DVFs. The disadvantage of training with artificial

Table 2
DL-based DIR models supervised by artificial DVFs.

Authors	Publication Year	ROI	Modality	Patch-based	Reference DVF Obtained by
Sokooti et al.[46]	2017	Lung	CT-CT	No	Mixed spatial frequency
Krebs et al.[47]	2017	Pelvic	MR-MR	No	Statistical deformation models
Eppenhof et al.[48,49]	2018	Lung	CT-CT	No	Random numbers
Sokooti et al.[50]	2019	Lung	CT-CT	No	Mixed spatial frequency

DVFs, however, was that artificial DVFs were different from real motions, and the primary challenge of the models in this category was to generate plausible DVFs in training data. Krebs et al[47] obtained a small number of reference DVFs by matching ROI contours and then augmented them by random deformation samples from the statistical deformation models (SDM). Eppenhof et al[48,49] generated artificial DVFs by randomly sampling from certain ranges in a coarse-to-fine grid. Sokooti et al[46,50] proposed two similar approaches in 2017 and 2019, in which random values were assigned to several control points and the final DVF was smoothed by Gaussian filters with different standard deviation, representing different spatial frequencies. The modification in the latter work was an additional binary mask and Gaussian smoothing on the DVF, making the spatial frequency of DVF even higher.

Regarding the performance of these works, artificial DVFs provided good supervision. Studies demonstrated that their registration accuracy was comparable with state-of-the-art methods and achieved average registration error around 2 mm in lungs.[48,50] Sokooti et al[50] also demonstrated the artificial DVF generation strategy could heavily affect the model performance. Instead of generating DVFs from random numbers, some other studies investigated the feasibility of modeling deformation and generating more realistic training samples based on 2D MR images.[51] However, such a strategy was not seen on 3D images. Research on this would be interesting and encouraged.

3.2.3. *Unsupervised and weakly supervised models*

Though reference data shortage is not a problem for artificial DVF-supervised models, preparing training samples is still a time-consuming process. It would be more convenient if only the moving and fixed image pairs are provided as training data and the model learns the deformation without the supervision of reference DVFs (unsupervised) or with supervisions other than reference DVFs (weakly supervised). Some unsupervised and weakly supervised DL-based DIR studies were listed in Table 3.

As shown, the training process is supervised by loss functions like those in conventional methods, usually comprising an image dissimilarity item and a DVF regularization item. A more straightforward demonstration can be found in Fig. 1. Despite the abovementioned MSE,

Table 3
Unsupervised and weakly supervised DL-based DIR models.

Authors	Publication Year	ROI	Modality	Patch-based	Supervised by
Sentker et al.[52]	2018	Lung	CT-CT	No	DIS
Cao et al.[43]	2018	Prostate	CT-MR	Yes	DIS, DVF
Hu et al.[53]	2018	Prostate	MR-MR	No	CON, ADV
Fan et al.[54]	2018	Brain	MR-MR	Yes	DIS, DVF, ADV
Kearney et al.[55]	2018	Head and Neck	CBCT-CT	Yes	DIS
Li et al.[56]	2018	Brain	MR-MR	No	DIS, DVF
Krebs et al.[57]	2018	Heart	MR-MR	No	DIS, DVF
Stergios et al.[58]	2018	Lung	MR-MR	No	DIS, DVF
Sun et al.[59]	2018	Brain	MR-US	No	DIS
Zhang et al.[60]	2018	Brain	MR-MR	No	DIS, DVF, INV
Fan et al.[61]	2019	Brain Pelvic	MR-MR	Yes	DIS, DVF, ADV
de Vos et al.[36]	2019	Heart Lung	CT-CT MR-MR	No	DIS, DVF
Balakrishnan et al.[11]	2019	Brain	MR-MR	No	DIS, DVF, CON
Kim et al.[33]	2019	Liver	CT-CT	No	DIS, CYC, IDE
Elmahdy et al.[62]	2019	Prostate	CT-CT	Yes	DIS, CON
Kuang et al.[63]	2019	Brain	MR-MR	No	DIS, DVF, CYC
Yu et al.[64]	2019	Abdomino-pelvic	PET-CT	Yes	DIS
Jiang et al.[35]	2020	Lung	CT-CT	No	DIS, DVF
Fu et al.[34]	2020	Lung	CT-CT	Yes	DIS, DVF, ADV
Fechter et al.[65]	2020	Lung Heart	CT-CT MR-MR	Yes	DIS, DVF, CYC
Lei et al.[66]	2020	Abdomen	CT-CT	Yes	DIS, DVF, ADV

Notes: DIS, Dissimilarity; DVF. DVF regularization; ADV. Adversarial loss; CON. Contour overlapping; INV. Inverse consistency; CYC. Cycle consistency; IDE. Identity loss.

MAE, NCC, etc., another popular loss function for dissimilarity is local NCC (LNCC), which utilizes the convolutional nature of CNNs and computes the cross-correlation among small patches.[11,35] As for DVF regularization, the most used loss function is DVF spatial gradient,[36] which can be mathematically expressed as

$$L_{gradient}(\varphi) = \sum_{x \in V} || \nabla \varphi(x) ||^2 \quad (2)$$

where φ represented the DVF and x was a voxel in the volume V. In some works, the DVF regularization term was further extended to the second order or the combination of first and second order of DVF spatial gradient.[11,34,35] For computation efficiency, some studies directly utilized the amplitude of DVFs as regularization.[43] Krebs et al[57] regularized the predicted DVF to a prior possibility distribution with Kullback-Leibler (KL) divergence. To achieve symmetric and more realistic registration, Kuang et al[63] regularized the DVF with cycle-consistency loss, which deformed the warped image back to the moving one and penalized the dissimilarity between them; Zhang et al[60] adopted the similar idea and included an inverse-consistent constrain to penalize the differences between the negative flow and the DVF from switched input images. In addition to these regular items, Balakrishnan et al[11] provided a weakly supervision option and structure contours could also be included in training, and Hu et al chose contour overlapping as the primary loss function.[53]

Unsupervised and weakly supervised learning is more challenging than supervised learning and therefore takes more effort, including more complex pre-processing and more complicated yet powerful models. Pre-processing of training samples usually involves ROI segmentation and windowing, which masks out all tissues outside ROI to emphasizes this region.[34,35] Furthermore, Fu et al.[34] implemented an automatic vessel segmentation and enhanced vessel region with a factor of 1000 to enrich the fine structures in CTs. Jiang et al.[67] also included the optical flow estimated by the first iteration of the Lucas-Kanada method as part of inputs to improve registration accuracy. As for deep learning models, most of them are based on typical deep learning convolutional neural network (DCNN) or U-Net with more trainable parameters and multiple scales in which the moving and fixed images are registered following a coarse-to-fine approach.[35–37] Some also combined affine and deformable image registration for higher accuracy.[36,58] Kim et al.[33] utilized a cycle-consistent CNN and regularized the predicted DVFs with cycle consistency and identity loss. Some models utilized GANs and integrated adversarial (ADV) loss in loss functions to better supervise model training and achieved high registration accuracy.[34,53,66]

Unsupervised DIR models nowadays draws great attention for its convenient training and broad application potentials. The number of papers in this field grows very rapid.[68,69] With more powerful deep learning models, especially GANs, the unsupervised models obtained comparable or even better registration accuracy with state-of-art conventional methods.[34–36] For lung CT registration specifically, so far the most accurate registration was achieved by Fu et al.[34] in lung CT registration with a target registration error (TRE) of 1.59 ± 1.58 mm. However, the high accuracy of unsupervised models is at the cost of complicated image pre-processing in training and sometimes even in implementation.

4. Applications of DIR in dose warping

With efficient and accurate DIR, multiple clinic tasks can be completed more easily. Particularly, as ART becoming more and more popular, DIR serves as the backbone to register the daily images to the planning image and provides accumulated dose for plan evaluation and re-planning.[14] Re-treatment is also often needed and the patient's anatomy may change largely due to the years or decades time separation.[70] Various DIR methods have been applied to multiple ROIs, including pulmonary, abdominal, and cervical regions.

As reviewed above, DL-based DIR methods achieved great registration accuracy as well as efficiency. Their application in ART, especially dose warping, is very promising. However, DL-based DIR has not yet been applied in dose warping; even for conventional DIR methods, debates are still going on about their feasibility on dose warping. Unlike matching images, in which pixel intensity is primarily focused, dose warping also requires the deformations to be physically plausible. Therefore, previously reviewed DL-based DIR models may encounter failures if they are directly utilized in dose warping. It is important to review related conventional DIR-based dose warping studies and identify the challenges they met and the solutions they proposed. Those challenges and solutions could provide experience for DL-based DIR model's improvement before their implementation.

4.1. Dose warping: fundamentals

The utility of imaging improves the treatment accuracy and precision in radiation oncology. Various imaging modalities, including CT, MRI, and ultrasound, provided detailed information on the patient's anatomy, aiding both planning and dose delivery.[71–73] However, the patient's anatomy could be inconsistent from fraction to fraction, resulting from breathing pattern variation, weight loss, and tumor shrinkage/growth during treatment.[8,14–16] The idea of adaptive radiotherapy (ART) was proposed to monitor the daily anatomical change and tune the treatment plan periodically.[14] The dose delivered to structures in each fraction varies, and the fraction doses need accumulating via DVFs obtained between daily CBCT and planning CT.[8] Also, some tumors like locally advanced cervical carcinoma require intracavitary brachytherapy (ICBT)

as a boost after external beam radiotherapy (EBRT),[74] and a composite dose is needed to meet the prescription. Besides, patient retreatments are sometimes needed, and lack of accurate registration and dose summation can cause dose error up to14 Gy.[70] It is necessary to include prior doses to avoid overdose to organs at risk.[75]

Georg et al[15,76] have shown a consistent correlation between dose toxicity and dose-volume histogram (DVH) parameters D_{2cm^3}, D_{1cm^3}, and $D_{0.1cm^3}$ (minimal dose to the most irradiated $2cm^3$, $1cm^3$, and $0.1cm^3$, respectively). Based on the "worst-case assumption", which assumes a static distribution of hotspots in all fractions, a composite dose and toxicity could be obtained by the summation of the abovementioned DVH parameters.[77] The assumption holds because intuitively the total hotspot dose cannot exceed the summation of hotspots of all fractions. However, doses could be even higher if organs move closer to applicators in ICBT or other high dose regions in EBRT within a fraction, resulting in cases worse than the assumed "worst" one. Consequently, the terminology was changed to "DVH parameter addition".[12] This effect varies from different organs and patients, introducing uncertainties in dose assessment.

A more recent approach for dose summation is to register all the moving images in different treatment fractions and warp doses to the fixed image using the DVF from registration,[12] which can be expressed mathematically as

$$D_{accu}(i) = D_F(i) + \sum_{M=1,\ M\neq F}^{N} D_M(i) = D_F(i) + \sum_{M=1,\ M\neq F}^{N} D_M(DVF_{MF}(i)) \qquad (3)$$

where $D_{accu}(i)$ is the accumulated dose at position i; D_F and D_M are the dose from moving and fixed dataset; $DVF_{MF}(i)$ is the DVF that maps the corresponding location from the moving dataset to the position i in the fixed one.

Various studies have been done to compare the difference, especially the hotspot difference in the accumulated dose from DIR-based dose summation and other methods including DVH parameter addition. Andersen et al. reported that in the composite of EBRT and ICBT, DVH parameter addition resulted in (0.4 ± 0.3) Gy (or $1.5\% \pm 1.8\%$) higher dose for D_{2cm^3} and (1.9 ± 1.6)Gy(or $5.2\% \pm 4.2\%$) for $D_{0.1\ cm^3}$ in bladders, respectively.[12] Similar results were also shown by Teo et al. that when combining fractions of ICBT, the $D_{2\ cm^3}$ of dose accumulated via DIR was $10.1\% \pm 9.5\%$ lower for rectums and $7.2\% \pm 6.3\%$ lower for bladders compared to DVH parameter addition; when combining EBRT and ICBT, the difference reduced to $-2.9\% \pm 4.0\%$ for rectums and $-3.2\% \pm 3.3\%$ for bladders.[74] However, limited dose difference for bladder was also reported both in brachytherapy[78] and EBRT,[13] indicating DVH parameter addition was an efficient and effective method for dose accumulation for most patients and patients with large organ variation between fractions benefit more from DIR-based dose accumulation.[13] In liver radiotherapy, which is mainly affected by breathing motion, Velec et al[79] reported that relative to static plans, DIR-based dose accumulation could account for the dose difference up to 8% to minimum tumors and 7% to maximum normal tissues. As for head and neck, brain, and mediastinum, the difference of $D_{0.1cm^3}$ between DIR-based and rigid dose accumulation varied from −8 Gy to 2 Gy.[70]

4.2. Challenges in DIR-based dose accumulation and summation

Despite the abovementioned promising results, challenging problems also exist in the implementation of DIR-based dose accumulation, including DIR accuracy and consistency, the feasibility of DIR-based dose accumulation, and the evaluation or measurement of it. These problems imply the immaturity of this method and studies address them would be meaningful and encouraged.

DIR-based dose accumulation is based on the DVFs generated in registration, therefore it is heavily affected by the accuracy of DVFs. However, for real clinical image pairs, the ground truth deformation is often unavailable. With the intrinsic ill-posed property of registration problems, the point-wise true deformation may not even exist.[80–82] As recommended by AAPM Task Group Report No. 132 and some other publications, landmark registration errors between 1 and 2 voxels or structure overlap greater than 80% would be acceptable in radiotherapy since such error could be accounted for by treatment margins.[80,81,83] For intra-modality registration, most commercially available software and in-house algorithms could meet these criteria, including Velocity, MIM, Syn, ANTs, and so on.[7,34–37,84] However, as Mogadas et al[85] reported, even if different DIR algorithms gave comparable registration errors, the hotspot dose difference could be as high as 29 Gy, especially for targets with low image contrast like livers. Inter-modality DIR could be more difficult, and some commercial software like MIM was reported to give errors up to 41.6 mm when doing kV-MV image registration.[7] In the clinic, though various registration tools were provided, human visual evaluation and adjustment of registration results were still considered necessary, especially for patients with significant tumor volume change.[5,86] For example, Andersen et al[12] reported 5 DIR failure cases out of 47 patients with obviously implausible DVFs; Hardcastle et al[87] found 27% GTVs propagated by DIR needed major edit. Besides the absolute registration error, inverse inconsistency is also a concern in registration accuracy, which means the deformation between the same image pair could be different via different pathways.[88] Bender et al[89] verified this effect also exists in DIR-based dose accumulation and can be reduced by combining DVFs from different registration pathways.

Precise registration is already difficult, but accurate dose accumulation is not guaranteed even with perfect registration provided. Tissue masses vary from fraction to fraction and sometimes even disappear, and DIR-based dose accumulation may violate the energy conservation principle.[90] This is because radiation dose reflects the energy density instead of the absolute amount of energy. When the constant energy density is deformed to a greater or smaller mass, energy is created or eliminated. To address this problem, Heath et al[91] proposed a voxel tracking method, in which the deformation of each rectangular voxel was tracked and the energy deposition of particles was calculated by Monte Carlo simulation inside the voxel of arbitrary shape. The density of each voxel was also adjusted so that the conservation of energy and mass was held. Certainly, such a process is very computationally expensive. Efforts have been made to simplify and speed up the calculation, and with modern algorithms and hardware, it can even be real-time with pre-calculated DVFs.[92–95]

Another challenge in DIR-based dose accumulation is the hotspot inverse inconsistency. Despite the DVF inverse inconsistency discussed above, dose indexes like $D_{0.1cm^3}$ and D_{2cm^3} in the accumulated dose can still be different, depending on the choice of the reference frame.[12] This phenomenon is because of the volume change between frames. For example, if the bladder in fraction No.1 is larger than that in fraction No.2, the hotspot D_{2cm^3} of fraction No.1 will be warped to a volume smaller than 2 cm^3 in fraction No.2; consequently, the hotspot D_{2cm^3} of fraction No.2 includes not only the corresponding volume but an additional margin, resulting in a larger value. An approximate approach is to accumulate dose in both pathways and report the mean value.[12] More accurate computation requires voxel-wise energy tracking as discussed above.

Though computer simulation looks accurate and convincing, dose measurement is still the most solid evidence for the feasibility of DIR-based dose accumulation. Graves et al[4] designed a 2D deformable head and neck phantom in the shape of a single axial CT slice with 4 diode holders and reported a DIR error of (2.1 ± 2.2) mm and a maximum dose error of 3.1% between the accumulated and prescription dose. Zhong et al[10] built a deformable phantom with thermoluminescent dosimeters (TLDs) placed in different locations in each irradiation. They found a difference within 5.1% between the measured doses and accumulated doses using Monte Carlo simulation; however, the difference went up to 11.8% when the accumulated doses were obtained by Pinnacle. The abovementioned 2 measurements were both point-wise, while Niu et al[96]

implemented a 3D gel measurement to obtain volumetric dose distribution and showed an average difference of $(1.5 \pm 13.4)\%$ between the accumulated and measured doses. Though the mean dose difference is acceptable, the dose difference in the control group, in which no deformation was applied, reached 4.7% compared to the treatment planning system (TPS).[96]

5. Discussion

This review introduced the basic idea of conventional DIR methods and some latest papers on DL-based DIR models. One clear difference between these two is that specific goals are achieved by modifying objective function in conventional DIR methods, while in DL-based models these goals can be achieved by modifying both loss function and data. For example, to make the methods inverse-consistent, the objective function needs to be modified so that dissimilarity from both directions is included; while in DL-based models, inverse consistency could be achieved by switching moving and fixed volumes in the training data.[35] Therefore, deep learning models depend on the data that they were trained on, and their performance can be different even if their architectures are the same. It may be helpful to include as much as possible data in training, especially data from multiple centers to improve data diversity and generalize the model. On the other hand, standardizing all the test data may also be a useful strategy, which makes the images in the testing data set similar to those in the training dataset.

In terms of the registration accuracy, although DL-based models are generally worse than conventional methods, they are getting better as new models are proposed. It is optimistic that these DL-based models can consistently provide comparable accuracy with conventional methods or even outperformed them soon. Deep learning is an active field and integrating the latest CNN architectures may lead to even better results. Another noteworthy point is the pre-processing of training and testing data in DL-based models. In the deep learning models mentioned in this review, especially unsupervised ones, the pre-processing becomes more and more complex. It does not only include regular isotropic resampling, intensity normalization, but also some advanced processes like lung segmentation and pulmonary vessel enhancement.[34,35] Although the registration is more accurate in the new deep learning models, it is difficult to conclude whether such improvement is from the model or the pre-processing.

Most of the presented studies focused on intra-modality registration, like CT-CT and MR-MR registration, and few studies did inter-modality registration. Cao et al[43] registered CT to MR images, however, they utilized the pre-aligned CT-MR image pairs and supervised the training with MR-MR intra-modality image similarity. This may because of the difficulty of measuring cross-modality image similarity, for example, Sun et al[59] designed a very complex similarity metrics for MR-US registration. In scenarios where no pre-aligned image pairs are available, a possible approach is to translate the input images to the same image modality and conduct intra-modality registration.[97,98] DL model's application in this field is worth further exploration.

Conventional DIR methods have been applied in many clinical tasks, and all of them could be potential applications for DL-based DIR methods. For example, Kurugol et al[99] used DIR in abdominal diffusion-weighted MRI reconstruction by registering images with different b values; Wang et al[100] did segmentation with DIR by deforming a segmented atlas to other subjects; Ebner et al[101] used DIR in abdominal image super-resolution by correcting organ motions. In principle, DL-based DIR models could also be used in abovementioned applications. However, model generalizability may be a potential obstacle for them since different deep learning models have to be trained individually for every specific application. The application of DL-based DIR models will be facilitated if general-purpose models or training strategies are proposed.

Dose warping is a promising application for DL-based DIR models. Current solutions either take a long computation time and are not eligible for online ART or require pre-calculated DVFs that may not accurately reflect patient's anatomical change at the time of treatment. If DL-based DIR models can provide accurate enough deformation in real-time, ART will be more convenient and the application of it may be more popular. However, all the difficulties conventional DIR methods encountered also apply to DL-based DIR models. A major one of them is the lack of dose-level evaluation. Now commonly used evaluation metrics for DL-based DIR models are intensity- or landmark-based, which focus on overall image similarity or anatomical structure matching. However, as reviewed previously, even several DIR methods can provide comparable registration accuracy, the dose accumulation difference of them can be up to 29 Gy in hotspots. Therefore, including dose information in model evaluation or training may be helpful when developing DIR models for dose warping. Another difficulty is the mass and energy conservation in deformation. In reviewed studies, voxel tracking and Monte Carlo simulation were used to compensate the body mass change. In DL-based DIR models, integrating Monte Carlo simulation may be difficult, but body density map can be an additional input to DL models and supervise the output warped dose to be energy conserved.

Despite the advantages and promising results, DIR-based dose accumulation and summation is still controversial, and it requires a thorough evaluation with accurate measurement before widely applied in the clinic. Though efforts have been made to develop advanced simulation models and deformable phantoms, expensive computational cost or point-wise measurement is still the major barrier.[81,102] 3D gel dosimeter seems promising in volumetric measurement, yet the intrinsic uncertainty makes it not convincing enough.[96] Research on more accurate and representative measurements on deformable phantoms will be interesting.

In addition to the feasibility of DIR-based dose accumulation and summation, new technical terminologies on the dose report are also needed. Conventional terminologies for dose hot spots are $D_{0.1\ cm^3}$ and $D_{2\ cm^3}$, describing the amount of energy deposit in a certain volume. However, inter-frame volume change can lead to a consistently higher or lower hot spot dose, depending on the choice of the reference frame.[12,81] This kind of variation can be misleading. New dose report terminologies or a standardized protocol for reference frame choice is needed to address this issue.

In summary, deep learning models are promising in medical image registration tasks. Efforts can be done on exploring new deep learning models, new pre-processing methods to enhance contrast, and more data from various centers. Though various conventional DIR methods have been applied and evaluated on dose accumulation and summation, the application of DL-based DIR methods on it is limited. This is partly because deep learning is a new technique and not fully explored. Future research on this topic would be interesting and encouraged.

Declaration of competing interest

None.

Acknowledgement

This research was partly supported by Hong Kong research grants (General Research Fund (GRF) from University Grants Committee: GRF 151021/18M and GRF 151022/19M; and Health and Medical Research Fund (HMRF) from Food and Health Bureau: HMRF 06173276 and HMRF 07183266).

References

1. Christensen GE, Song JH, Lu W, et al. Tracking lung tissue motion and expansion/compression with inverse consistent image registration and spirometry. *Med Phys*. 2007;34(6Part1):2155–2163. https://doi.org/10.1118/1.2731029.
2. Reed VK, Woodward WA, Zhang L, et al. Automatic segmentation of whole breast using atlas approach and deformable image registration. *Int J Radiat Oncol Biol Phys*. 2009;73(5):1493–1500. https://doi.org/10.1016/j.ijrobp.2008.07.001.

3. Schreibmann E, Thorndyke B, Li T, et al. Four-dimensional image registration for image-guided radiotherapy. *Int J Radiat Oncol Biol Phys*. 2008;71(2):578–586. https://doi.org/10.1016/j.ijrobp.2008.01.042.
4. Graves YJ, Smith AA, McIlvena D, et al. A deformable head and neck phantom with in-vivo dosimetry for adaptive radiotherapy quality assurance. *J Med Phys*. 2015; 42(4):1490–1497. https://doi.org/10.1118/1.4908205.
5. Olteanu LA, Madani I, De Neve W, et al. Evaluation of deformable image coregistration in adaptive dose painting by numbers for head-and-neck cancer. *Int J Radiat Oncol Biol Phys*. 2012;83(2):696–703. https://doi.org/10.1016/j.ijrobp.2011.07.037.
6. Schwartz DL, Garden AS, Thomas J, et al. Adaptive radiotherapy for head-and-neck cancer: initial clinical outcomes from a prospective trial. *Int J Radiat Oncol Biol Phys*. 2012;83(3):986–993. https://doi.org/10.1016/j.ijrobp.2011.08.017.
7. Singhrao K, Kirby N, Pouliot J. A three-dimensional head-and-neck phantom for validation of multimodality deformable image registration for adaptive radiotherapy. *J Med Phys*. 2014;41(12):121709. https://doi.org/10.1118/1.4901523.
8. Veiga C, Lourenço AM, Mouinuddin S, et al. Toward adaptive radiotherapy for head and neck patients: uncertainties in dose warping due to the choice of deformable registration algorithm. *J Med Phys*. 2015;42(2):760–769. https://doi.org/10.1118/1.4905050.
9. Yang D, Brame S, El Naqa I, et al. DIRART–A software suite for deformable image registration and adaptive radiotherapy research. *Med Phys*. 2011;38(1):67–77. https://doi.org/10.1118/1.3521468.
10. Zhong H, Adams J, Glide-Hurst C, et al. Development of a deformable dosimetric phantom to verify dose accumulation algorithms for adaptive radiotherapy. *J Med Phys*. 2016;41(2):106. https://doi.org/10.4103/0971-6203.181641.
11. Balakrishnan G, Zhao A, Sabuncu MR, et al. VoxelMorph: a learning framework for deformable medical image registration. *IEEE Trans Med Imag*. 2019. https://doi.org/10.1109/TMI.2019.2897538.
12. Andersen ES, Noe KØ, Sørensen TS, et al. Simple DVH parameter addition as compared to deformable registration for bladder dose accumulation in cervix cancer brachytherapy. *Radiother Oncol*. 2013;107(1):52–57. https://doi.org/10.1016/j.radonc.2013.01.013.
13. Andersen E, Muren L, Sørensen TS, et al. Bladder dose accumulation based on a biomechanical deformable image registration algorithm in volumetric modulated arc therapy for prostate cancer. *Phys Med Biol*. 2012;57(21):7089. https://doi.org/10.1088/0031-9155/57/21/7089.
14. Yan D, Vicini F, Wong J, et al. Adaptive radiation therapy. *Phys Med Biol*. 1997; 42(1):123. https://doi.org/10.1088/0031-9155/42/1/008.
15. Georg P, Lang S, Dimopoulos JC, et al. Dose–volume histogram parameters and late side effects in magnetic resonance image–guided adaptive cervical cancer brachytherapy. *Int J Radiat Oncol Biol Phys*. 2011;79(2):356–362. https://doi.org/10.1016/j.ijrobp.2009.11.002.
16. Simone II CB, Ly D, Dan TD, et al. Comparison of intensity-modulated radiotherapy, adaptive radiotherapy, proton radiotherapy, and adaptive proton radiotherapy for treatment of locally advanced head and neck cancer. *Radiother Oncol*. 2011;101(3): 376–382.
17. de Jong R, Crama K, Visser J, et al. Online adaptive radiotherapy compared to plan selection for rectal cancer: quantifying the benefit. *Radiat Oncol*. 2020;15(1):1–9. https://doi.org/10.1186/s13014-020-01597-1.
18. Samant SS, Xia J, Muyan-Özçelik P, et al. High performance computing for deformable image registration: towards a new paradigm in adaptive radiotherapy. *Med Phys*. 2008;35(8):3546–3553. https://doi.org/10.1118/1.2948318.
19. Klein S, Staring M, Murphy K, et al. Elastix: a toolbox for intensity-based medical image registration. *IEEE Trans Med Imag*. 2009;29(1):196–205. https://doi.org/10.1109/TMI.2009.2035616.
20. Avants BB, Tustison NJ, Song G, et al. A reproducible evaluation of ANTs similarity metric performance in brain image registration. *Neuroimage*. 2011;54(3): 2033–2044. https://doi.org/10.1016/j.neuroimage.2010.09.025.
21. Vercauteren T, Pennec X, Perchant A, et al. Diffeomorphic demons: efficient non-parametric image registration. *Neuroimage*. 2009;45(1):S61–S72. https://doi.org/10.1016/j.neuroimage.2008.10.040.
22. Hubel DH, Wiesel TN. Receptive fields and functional architecture of monkey striate cortex. *J Physiol*. 1968;195(1):215–243. https://doi.org/10.1113/jphysiol.1968.sp008455.
23. Waibel A. Modular construction of time-delay neural networks for speech recognition. *Neural Comput*. 1989;1(1):39–46. https://doi.org/10.1162/neco.1989.1.1.39.
24. LeCun Y, Boser B, Denker JS, et al. Backpropagation applied to handwritten zip code recognition. *Neural Comput*. 1989;1(4):541–551. https://doi.org/10.1162/neco.1989.1.4.541.
25. Lecun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition. *Proc IEEE*. 1998;86(11):2278–2324. https://doi.org/10.1109/5.726791.
26. Oh K-S, Jung K. GPU implementation of neural networks. *Pattern Recogn*. 2004; 37(6):1311–1314. https://doi.org/10.1016/j.patcog.2004.01.013.
27. Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. *Commun ACM*. 2017;60(6):84–90. https://doi.org/10.1145/3065386.
28. De Dombal F, Leaper D, Staniland JR, et al. Computer-aided diagnosis of acute abdominal pain. *Br Med J*. 1972;2(5804):9–13. https://doi.org/10.1136/bmj.2.5804.9.
29. Nakao T, Hanaoka S, Nomura Y, et al. Deep neural network-based computer-assisted detection of cerebral aneurysms in MR angiography. *J Magn Reson Imaging*. 2018;47(4):948–953. https://doi.org/10.1002/jmri.25842.
30. Chang K, Bai HX, Zhou H, et al. Residual convolutional neural network for the determination of IDH status in low- and high-grade gliomas from MR imaging. *Clin Canc Res*. 2018;24(5):1073–1081. https://doi.org/10.1158/1078-0432.Ccr-17-2236.
31. Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation. Medical image computing and computer-assisted intervention – MICCAI, 2015:234-241 https://doi.org/10.1007/978-3-319-24574-4_28; 2015.
32. Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets. *Adv Neural Inf Process Syst*. 2014:2672–2680.
33. Kim B, Kim J, Lee J-G, et al. Unsupervised deformable image registration using cycle-consistent cnn. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. 2019:166–174.
34. Fu Y, Lei Y, Wang T, et al. LungRegNet: an unsupervised deformable image registration method for 4D-CT lung. *Med Phys*. 2020;47(4):1763–1774. https://doi.org/10.1002/mp.14065.
35. Jiang Z, Yin F-F, Ge Y, et al. A multi-scale framework with unsupervised joint training of convolutional neural networks for pulmonary deformable image registration. *Phys Med Biol*. 2020;65(1), 015011. https://doi.org/10.1088/1361-6560/ab5da0.
36. de Vos BD, Berendsen FF, Viergever MA, et al. A deep learning framework for unsupervised affine and deformable image registration. *Med Image Anal*. 2019;52: 128–143. https://doi.org/10.1016/j.media.2018.11.010.
37. de Vos BD, Berendsen FF, Viergever MA, et al. End-to-end unsupervised deformable image registration with a convolutional neural network. Deep learning in medical image analysis and multimodal learning for clinical decision support, 204-212 https://doi.org/10.1007/978-3-319-67558-9_24; 2017.
38. Durrleman S, Prastawa M, Charon N, et al. Morphometry of anatomical shape complexes with dense deformations and sparse parameters. *Neuroimage*. 2014;101: 35–49. https://doi.org/10.1016/j.neuroimage.2014.06.043.
39. Beg MF, Khan A. Computing an average anatomical atlas using LDDMM and geodesic shooting. In: *3rd IEEE International Symposium on Biomedical Imaging*. vol. 2006. Nano to Macro; 2006:1116–1119. https://doi.org/10.1109/ISBI.2006.1625118.
40. Yang X, Kwitt R, Styner M, et al. Quicksilver: fast predictive image registration – a deep learning approach. *Neuroimage*. 2017;158:378–396. https://doi.org/10.1016/j.neuroimage.2017.07.008.
41. Rohé M-M, Datar M, Heimann T, et al. SVF-Net: learning deformable image registration using shape matching. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. 2017:266–274. https://doi.org/10.1007/978-3-319-66182-7_31.
42. Cao X, Yang J, Zhang J, et al. Deformable image registration based on similarity-steered CNN regression. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. 2017:300–308. https://doi.org/10.1007/978-3-319-66182-7_35.
43. Cao X, Yang J, Wang L, et al. Deep learning based inter-modality image registration supervised by intra-modality similarity. In: *International Workshop on Machine Learning in Medical Imaging*. 2018:55–63. https://doi.org/10.1007/978-3-030-00919-9_7.
44. Onieva JO, Marti-Fuster B, de la Puente MP, et al. *Diffeomorphic Lung Registration Using Deep Cnns and Reinforced Learning. Image Analysis for Moving Organ, Breast, and Thoracic Images*. 2018:284–294. https://doi.org/10.1007/978-3-030-00946-5_28.
45. Fan J, Cao X, Yap P-T, et al. BIRNet: brain image registration using dual-supervised fully convolutional networks. *Med Image Anal*. 2019;54:193–206. https://doi.org/10.1016/j.media.2019.03.006.
46. Sokooti H, De Vos B, Berendsen F, et al. Nonrigid image registration using multi-scale 3D convolutional neural networks. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. 2017:232–239. https://doi.org/10.1007/978-3-319-66182-7_27.
47. Krebs J, Mansi T, Delingette H, et al. Robust non-rigid registration through agent-based action learning. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. 2017:344–352. https://doi.org/10.1007/978-3-319-66182-7_40.
48. Eppenhof KA, Pluim JP. Pulmonary CT registration through supervised learning with convolutional neural networks. *IEEE Trans Med Imag*. 2018;38(5):1097–1105. https://doi.org/10.1109/TMI.2018.2878316.
49. Eppenhof KA, Lafarge MW, Moeskops P, et al. Deformable image registration using convolutional neural networks. *Med Imaging*. 2018;10574:105740S. https://doi.org/10.1117/12.2292443. Image Processing. 2018.
50. Sokooti H, de Vos B, Berendsen F, et al. *3D Convolutional Neural Networks Image Registration Based on Efficient Supervised Learning from Artificial Deformations*. 2019. *arXiv preprint arXiv:10235*.
51. Uzunova H, Wilms M, Handels H, et al. Training CNNs for image registration from few samples with model-based data augmentation. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. 2017:223–231. https://doi.org/10.1007/978-3-319-66182-7_26.
52. Sentker T, Madesta F, Werner R. GDL-FIRE (4D): deep learning-based fast 4D CT image registration. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. 2018:765–773. https://doi.org/10.1007/978-3-030-00928-1_86.
53. Hu Y, Gibson E, Ghavami N, et al. Adversarial deformation regularization for training image registration neural networks. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. 2018:774–782. https://doi.org/10.1007/978-3-030-00928-1_87.
54. Fan J, Cao X, Xue Z, et al. Adversarial similarity network for evaluating image alignment in deep learning based registration. In: *International Conference on*

Medical Image Computing and Computer-Assisted Intervention. 2018:739–746. https://doi.org/10.1007/978-3-030-00928-1_83.
55. Kearney V, Haaf S, Sudhyadhom A, et al. An unsupervised convolutional neural network-based algorithm for deformable image registration. *Phys Med Biol*. 2018; 63(18):185017. https://doi.org/10.1088/1361-6560/aada66.
56. Li H, Fan Y. Non-rigid image registration using self-supervised fully convolutional networks without training data. In: *IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*. 2018:1075–1078. https://doi.org/10.1109/ISBI.2018.8363757, 2018.
57. Krebs J, Mansi T, Mailhé B, et al. Unsupervised probabilistic deformation modeling for robust diffeomorphic registration. In: *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*. 2018:101–109. https://doi.org/10.1007/978-3-030-00889-5_12.
58. Stergios C, Mihir S, Maria V, et al. *Linear and Deformable Image Registration with 3d Convolutional Neural Networks. Image Analysis For Moving Organ, Breast, and Thoracic Images*. 2018:13–22. https://doi.org/10.1007/978-3-030-00946-5_2.
59. Sun L, Zhang S. *Deformable Mri-Ultrasound Registration Using 3d Convolutional Neural Network. Simulation, Image Processing, and Ultrasound Systems For Assisted Diagnosis and Navigation*. 2018:152–158. https://doi.org/10.1007/978-3-030-01045-4_18.
60. Zhang J. *Inverse-consistent Deep Networks for Unsupervised Deformable Image Registration*. 2018. *arXiv preprint arXiv:180903443*.
61. Fan J, Cao X, Wang Q, et al. Adversarial learning for mono-or multi-modal registration. *Med Image Anal*. 2019;58:101545. https://doi.org/10.1016/j.media.2019.101545.
62. Elmahdy MS, Jagt T, Zinkstok RT, et al. Robust contour propagation using deep learning and image registration for online adaptive proton therapy of prostate cancer, 46(8):3329-3343 https://doi.org/10.1002/mp.13620; 2019.
63. Kuang D. Cycle-Consistent training for reducing negative Jacobian determinant in deep registration networks. In: *International Workshop on Simulation and Synthesis in Medical Imaging*. 2019:120–129. https://doi.org/10.1007/978-3-030-32778-1_13.
64. Yu H, Zhou X, Jiang H, et al. *Learning 3D Non-rigid Deformation Based on an Unsupervised Deep Learning for PET/CT Image Registration. Medical Imaging 2019: Biomedical Applications in Molecular, Structural, and Functional Imaging*. vol. 10953. 2019:109531X. https://doi.org/10.1117/12.2512698.
65. Fechter T, Baltas D. One shot learning for deformable medical image registration and periodic motion tracking. *IEEE Trans Med Imag*. 2020;39(7):2506–2517. https://doi.org/10.1109/TMI.2020.2972616.
66. Lei Y, Fu Y, Wang T, et al. 4D-CT deformable image registration using multiscale unsupervised deep learning. *Phys Med Biol*. 2020;65(8), 085003. https://doi.org/10.1088/1361-6560/ab79c4.
67. Jiang P, Shackleford JA. Cnn driven sparse multi-level b-spline image registration. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2018:9281–9289.
68. Fu Y, Lei Y, Wang T, et al. Deep learning in medical image registration: a review. *Phys Med Biol*. 2020;65(20):20TR01. https://doi.org/10.1088/1361-6560/ab843e.
69. Haskins G, Kruger U, Yan P. Deep learning in medical image registration: a survey. *Mach Vis Appl*. 2020;31(1):8. https://doi.org/10.1007/s00138-020-01060-x.
70. Boman E, Kapanen M, Pickup L, et al. Importance of deformable image registration and biological dose summation in planning of radiotherapy retreatments. *Med Dosim*. 2017;42(4):296–303. https://doi.org/10.1016/j.meddos.2017.06.006.
71. Verellen D, De Ridder M, Storme G. A (short) history of image-guided radiotherapy. *Radiother Oncol*. 2008;86(1):4–13. https://doi.org/10.1016/j.radonc.2007.11.023.
72. Dawson LA, Sharpe MB. Image-guided radiotherapy: rationale, benefits, and limitations. *Lancet Oncol*. 2006;7(10):848–858. https://doi.org/10.1016/S1470-2045(06)70904-4.
73. Mackie TR, Kapatoes J, Ruchala K, et al. Image guidance for precise conformal radiotherapy. *Int J Radiat Oncol Biol Phys*. 2003;56(1):89–105. https://doi.org/10.1016/s0360-3016(03)00090-7.
74. Teo B-K, Millar LPB, Ding X, et al. Assessment of cumulative external beam and intracavitary brachytherapy organ doses in gynecologic cancers using deformable dose summation. *Radiother Oncol*. 2015;115(2):195–202. https://doi.org/10.1016/j.radonc.2015.04.002.
75. Zhang G, Huang T-C, Feygelman V, et al. Generation of composite dose and biological effective dose (BED) over multiple treatment modalities and multistage planning using deformable image registration. *Med Dosim*. 2010;35(2):143–150. https://doi.org/10.1016/j.meddos.2009.05.001.
76. Georg P, Pötter R, Georg D, et al. Dose effect relationship for late side effects of the rectum and urinary bladder in magnetic resonance image-guided adaptive cervix cancer brachytherapy. *Int J Radiat Oncol Biol Phys*. 2012;82(2):653–657. https://doi.org/10.1016/j.ijrobp.2010.12.029.
77. Potter R, Haie-Meder C, Van Limbergen E, et al. Recommendations from gynaecological (GYN) GEC ESTRO working group (II): concepts and terms in 3D image-based treatment planning in cervix cancer brachytherapy - 3D dose volume parameters and aspects of 3D image-based anatomy, radiation physics, radiobiology. *Radiother Oncol*. 2006;78(1):67–77. https://doi.org/10.1016/j.radonc.2005.11.014.
78. Sabater S, Andres I, Sevillano M, et al. Dose accumulation during vaginal cuff brachytherapy based on rigid/deformable registration vs. single plan addition. *Brachytherapy*. 2014;13(4):343–351. https://doi.org/10.1016/j.brachy.2013.11.006.
79. Velec M, Moseley JL, Eccles CL, et al. Effect of breathing motion on radiotherapy dose accumulation in the abdomen using deformable registration, 80(1):265-272 https://doi.org/10.1016/j.ijrobp.2010.05.023; 2011.
80. Brock KK, Mutic S, McNutt TR, et al. Use of image registration and fusion algorithms and techniques in radiotherapy: report of the AAPM radiation therapy committee task group No. 132. *Med Phys*. 2017;44(7):e43–e76. https://doi.org/10.1002/mp.12256.
81. Chetty IJ, Rosu-Bubulac M. Deformable registration for dose accumulation. *Semin Radiat Oncol*. 2019;29(3):198–208. https://doi.org/10.1016/j.semradonc.2019.02.002.
82. Crum WR, Hartkens T, Hill D. Non-rigid image registration: theory and practice. *Br J Radiol*. 2004;77(suppl_2):S140–S153. https://doi.org/10.1259/bjr/25329214.
83. Oh S, Kim SJRoj. Deformable image registration in radiation therapy. *Radiat Oncol J*. 2017;35(2):101. https://doi.org/10.3857/roj.2017.00325.
84. Lowther NJ, Marsh SH, Louwe RJJR, et al. Quantifying the dose accumulation uncertainty after deformable image registration in head-and-neck radiotherapy. *Radiother Oncol*. 2020;143:117–125. https://doi.org/10.1016/j.radonc.2019.12.009.
85. Mogadas N, Sothmann T, Knopp T, et al. Influence of deformable image registration on 4D dose simulation for extracranial SBRT: a multi-registration framework study. *Radiother Oncol*. 2018;127(2):225–232. https://doi.org/10.1016/j.radonc.2018.03.015.
86. Boswell S, Tomé W, Jeraj R, et al. Automatic registration of megavoltage to kilovoltage CT images in helical tomotherapy: an evaluation of the setup verification process for the special case of a rigid head phantom. *Med Phys*. 2006; 33(11):4395–4404. https://doi.org/10.1118/1.2349698.
87. Hardcastle N, Tomé WA, Cannon DM, et al. A multi-institution evaluation of deformable image registration algorithms for automatic organ delineation in adaptive head and neck radiotherapy. *Radiat Oncol*. 2012;7(1):90. https://doi.org/10.1186/1748-717X-7-90.
88. Christensen GE, Johnson HJ. Consistent image registration. *IEEE Trans Med Imag*. 2001;20(7):568–582. https://doi.org/10.1109/42.932742.
89. Bender ET Hardcastle N, WAJMp Tome. On the dosimetric effect and reduction of inverse consistency and transitivity errors in deformable image registration for dose accumulation. *Med Phys*. 2012;39(1):272–280. https://doi.org/10.1118/1.3666948.
90. Zhong H, Chetty IJ. Caution must be exercised when performing deformable dose accumulation for tumors undergoing mass changes during fractionated radiation therapy. *Int J Radiat Oncol Biol Phys*. 2017;97(1):182–183. https://doi.org/10.1016/j.ijrobp.2016.09.012.
91. Heath E, Seuntjens J. A direct voxel tracking method for four-dimensional Monte Carlo dose calculations in deforming anatomy. *Med Phys*. 2006;33(2):434–445. https://doi.org/10.1118/1.2163252.
92. Siebers JV, Zhong H. An energy transfer method for 4D Monte Carlo dose calculation. *Med Phys*. 2008;35(9):4096–4105. https://doi.org/10.1118/1.2968215.
93. Li HS, Zhong H, Kim J, et al. Direct dose mapping versus energy/mass transfer mapping for 4D dose accumulation: fundamental differences and dosimetric consequences. *Phys Med Biol*. 2013;59(1):173. https://doi.org/10.1088/0031-9155/59/1/173.
94. Zhong H, Siebers JV. Monte Carlo dose mapping on deforming anatomy. *Phys Med Biol*. 2009;54(19):5815. https://doi.org/10.1088/0031-9155/54/19/010.
95. Ziegenhein P, Kamerling CP, Fast MF, et al. Real-time energy/mass transfer mapping for online 4D dose reconstruction. *Sci Rep*. 2018;8(1):1–10. https://doi.org/10.1038/s41598-018-21966-x.
96. Niu CJ, Foltz WD, Velec M, et al. A novel technique to enable experimental validation of deformable dose accumulation. *Med Phys*. 2012;39(2):765–776. https://doi.org/10.1118/1.3676185.
97. Roy S, Carass A, Jog A, et al. MR to CT registration of brains using image synthesis. *Med Imaging*. 2014:903419. https://doi.org/10.1117/12.2043954. Image Processing. 2014;9034.
98. Chen M, Carass A, Jog A, et al. Cross contrast multi-channel image registration using image synthesis for MR brain images. *J Med Imag Anal*. 2017;36:2–14. https://doi.org/10.1016/j.media.2016.10.005.
99. Kurugol S, Freiman M, Afacan O, et al. Motion-robust parameter estimation in abdominal diffusion-weighted MRI by simultaneous image registration and model estimation. *Med Image Anal*. 2017;39:124–132. https://doi.org/10.1016/j.media.2017.04.006.
100. Wang Y, Seguro F, Kao E, et al. Segmentation of lumen and outer wall of abdominal aortic aneurysms from 3D black-blood MRI with a registration based geodesic active contour model. *Med Image Anal*. 2017;40:1–10. https://doi.org/10.1016/j.media.2017.05.005.
101. Ebner M, Chouhan M, Patel PA, et al. Point-spread-function-aware slice-to-volume registration: application to upper abdominal MRI super-resolution. Reconstruction, Segmentation, and Analysis of Medical Images, 3-13 https://doi.org/10.1007/978-3-319-52280-7_1; 2016.
102. Schultheiss TE, Tomé WA, Orton CG. It is not appropriate to "deform" dose along with deformable image registration in adaptive radiotherapy. *J Med Phys*. 2012; 39(11):6531–6533. https://doi.org/10.1118/1.4722968.