

RESEARCH

Open Access



Deformable registration for nasopharyngeal carcinoma using adaptive mask and weight allocation strategy based CycleFCNs model

Yi Guo^{1,2,3}, Jun Chen^{1,2}, Lin Lu^{1,2}, Lingna Qiu^{1,2}, Linzhen Lan^{1,2}, Feibao Guo^{1,2,3*} and Jinsheng Hong^{1,2,3*}

Abstract

Background Deformable registration plays an important role in the accurate delineation of tumors. Most of the existing deep learning methods ignored two issues that can lead to inaccurate registration, including the limited field of view in MR scans and the different scanning angles that can exist between multimodal images. The purpose of this study is to improve the registration accuracy between CT and MR for nasopharyngeal carcinoma cases.

Methods 269 cases were enrolled in the study, and 188 cases were designated for training, while a separate set of 81 cases was reserved for testing. Each case had a CT volume and a T1-MR volume. The treatment table was removed from their CT images. The CycleFCNs model was used for deformable registration, and two strategies including adaptive mask registration strategy and weight allocation strategy were adopted for training. Dice similarity coefficient, Hausdorff distance, precision, and recall were calculated for normal tissues of CT-MR image pairs, before and after the registration. Three deformable registration methods including RayStation, Elastix, and VoxelMorph were compared with the proposed method.

Results The registration results of RayStation and Elastix are essentially consistent. Upon employing the VoxelMorph model and the proposed method for registration, a clear trend of increased dice similarity coefficient and decreased hausdorff distance can be observed. It is noteworthy that for the temporomandibular joint, pituitary, optic nerve, and optic chiasma, the proposed method has improved the average dice similarity coefficient from 0.86 to 0.91, 0.87 to 0.93, 0.85 to 0.89, and 0.77 to 0.83, respectively, as compared to RayStation. Additionally, within the same anatomical structures, the average hausdorff distance has been decreased from 2.98 mm to 2.28 mm, 1.83 mm to 1.53 mm, 3.74 mm to 3.56 mm, and 5.94 mm to 5.87 mm. Compared to the original CycleFCNs model, the improved model has significantly enhanced the dice similarity coefficient of the brainstem, pituitary gland, and optic nerve ($P < 0.001$).

Conclusions The proposed method significantly improved the registration accuracy for multi-modal medical images in NPC cases. These findings have important clinical implications, as increased registration accuracy can lead to more precise tumor segmentation, optimized treatment planning, and ultimately, improved patient outcomes.

Keywords Deformable registration, Multimodal images, Deep learning, Nasopharyngeal carcinoma

*Correspondence:

Feibao Guo
gfb803@fjmu.edu.cn
Jinsheng Hong
hjs703@126.com

¹Department of Radiation Therapy, Cancer Center, The First Affiliated Hospital of Fujian Medical University, Fuzhou 350005, China

²Department of Radiotherapy, National Regional Medical Center, Binhai Campus of The First Affiliated Hospital, Fujian Medical University, Fuzhou 350212, China

³Key Laboratory of Radiation Biology of Fujian Higher Education Institutions, Fuzhou 350005, China



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

Background

The International Agency for Research on Cancer projects a 63.4% escalation in the global incidence of cancer by the year 2040, relative to the figures reported in 2018 [1]. Radiotherapy, characterized by its precision in targeting tumors while concurrently minimizing exposure to adjacent healthy tissues, stands as a pivotal modality in the treatment of cancer. The accuracy of tumor segmentation is paramount to the efficacy of precise radiotherapy, necessitating the utilization of multimodal medical imaging, including Computed Tomography (CT) and Magnetic Resonance (MR) scans, to furnish physicians with a comprehensive dataset for delineating the tumor margins. However, the fusion of multimodal images poses a challenge due to variations in image dimensions, pixel spacing, the number of slices, slice thickness, and patient positioning. These discrepancies can impede the direct application of multimodal imaging data by medical professionals. Image registration, a technique that aligns two images with distinct anatomical structures, is instrumental in addressing this issue, facilitating the fusion of multimodal images and the precise segmentation of regions of interest (ROI) [2–3]. While rigid registration is prevalent in clinical workflow and is capable of effecting translations and rotations, it is limited in its ability to address the complexities of organ deformation. In contrast, deformable registration offers a more sophisticated solution by employing spatially deformable grids to achieve a more nuanced alignment of the images [4–5].

Conventional deformable registration methods obtain the transformation parameters by optimizing an objective function, which is time-consuming and inaccurate for complex three-dimensional image cases. In recent years, the advent of deep learning models has introduced a paradigm shift in the field of medical image deformable registration. These models encompass a spectrum of supervised, weakly supervised, and unsupervised neural networks, each with distinct advantages and methodologies [6–10]. (1) Supervised and weakly supervised neural networks are predominantly informed by prior knowledge during the training phase. This includes leveraging pre-aligned images, synthetic datasets, and ROIs delineated by clinicians [11–17]. For instance, Van Eijnatten et al. solved the problem of insufficient CT data based on synthetic deformable fields, and the results showed that although their method was not superior to other common tools in terms of the registration accuracy of vertebrae, it greatly improved the registration speed [18]. Similarly, Eppenhof et al. utilized a stochastic synthetic deformable field to augment lung CT datasets for the training of supervised neural networks [19]. Despite the supervised or weakly supervised methods to a certain extent, alleviate or avoid the optimization problem of metrics. They require synthetic deformable fields or

manually defined markers as the registration basis. (2) Unsupervised methods, on the other hand, directly calculate the similarity loss function between the fixed and moving images. While this approach sidesteps the need for prior knowledge, it may encounter issues related to local extrema during optimization [20–24].

Nasopharyngeal Carcinoma (NPC) is one of the head and neck malignancies with high incidence in East Asia and Southeast Asia, and radiotherapy is recommended because it is sensitive to ionizing radiation [25]. The delineation of the NPC tumor target, characterized by its complexity and irregularity, necessitates the utilization of multimodal images such as planning CT and MR [26]. Rigid registration can align images through translation and rotation, yet it fails to address deformations of soft tissues, which are particularly common between multimodal images [27]. Deformable registration offers a distinct advantage over its rigid counterpart by providing superior alignment of soft tissue structures and yielding more precise information regarding the tumor's margins, which is crucial for effective treatment planning. Despite a few neural network models were proposed for NPC deformable registration [28–29], several challenges remain to be addressed: (1) Supervised neural networks are used in some studies, which require prior knowledge, such as ROIs delineated in advance by physicians, as labels for training. These labels are difficult to obtain and difficult to ensure accuracy. (2) Most studies proposed a deformable registration network between CT images but ignored the registration between multimodal images. (3) The limited field of view (FOV) in MR scans may result in an incomplete representation of anatomical structures, such as the shoulder, as depicted in Fig. 1. Existing methods struggle to reconcile inconsistencies in anatomical structure information between multimodal images. (4) Discrepancies in the scanning inclination angles between CT and MR acquisitions may introduce registration errors at the edge slices of the two image volumes, as shown in Fig. 2. None of the above studies has considered this issue. To date, this potential source of error has not been adequately considered within the existing literature.

In this study, an unsupervised model named cycle-consistent fully convolutional network (CycleFCNs) was used as the deep learning framework [30] for deformable registration between MR and CT in NPC cases. To address the challenge of limited FOV, we introduce an adaptive mask registration strategy by extracting the outer contour mask of MR image as the registration domain. During the training process, CycleFCNs only focuses on the registration domain but ignores the other regions in the CT image. Furthermore, to mitigate the registration discrepancies precipitated by edge information mismatches, we proposed a weight allocation strategy for the assessment of different image slices. Specifically, higher weights

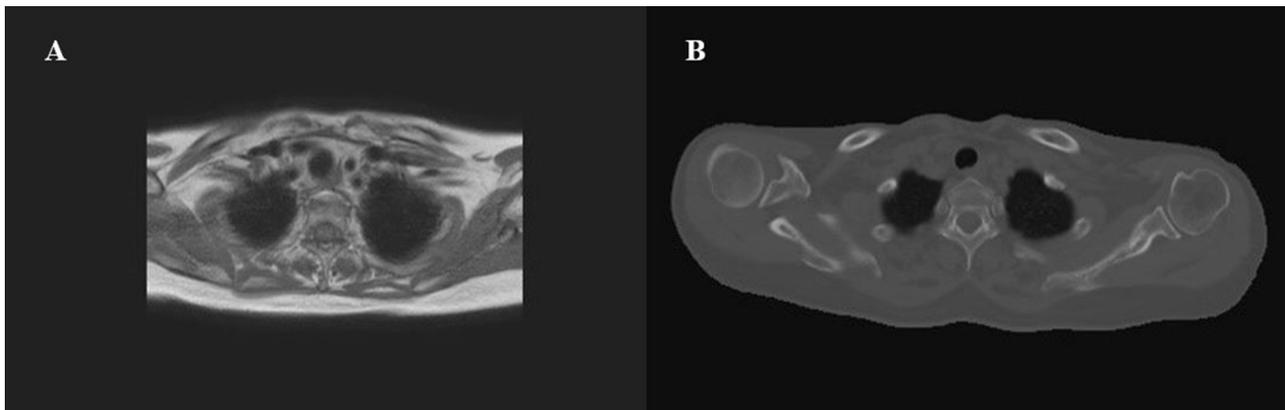


Fig. 1 The mismatch of anatomical structure information between CT (B) and MR (A)

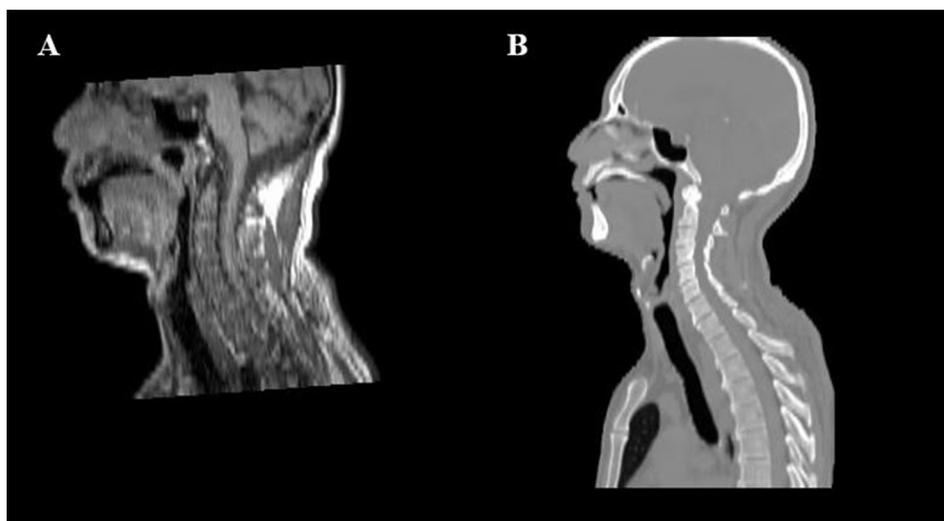


Fig. 2 The mismatch of scanning inclination angles between MR (A) and CT (B)

were allocated to the central slice of image volume, and lower weights were allocated to the edge slice to reduce the registration errors caused by the mismatch of image edge information. By providing more precise alignment of multimodal imaging data, the improved deformable registration method can lead to better tumor segmentation, optimized dose delivery, and more effective adaptive radiotherapy, ultimately improving patient outcomes and quality of life.

Methods

Data acquisition

NPC patients who underwent intensity-modulated radiotherapy in our hospital from April 2020 to December 2021 were randomly selected as the research objects. Each case had a CT volume and a contrast enhanced T1-MR volume. CT scans were performed on an 80 cm Siemens large aperture CT simulator, with parameters set to an acquisition voltage of 120 kV, a current of 35 mA, and a slice thickness of 3 mm. Within

three days after the CT scan, corresponding MR images were obtained utilizing a Prodiva 1.5 T magnetic resonance imaging system, with a FOV of 240×240 mm. Subsequently, all imaging data were transferred to the RayStation treatment planning system (TPS), and ROIs on both CT and MR were delineated by a senior physician, and the delineation results were reviewed and confirmed by another senior physician. A total of 269 cases were enrolled in the study, including 216 males and 83 females, with a median age of 50. The clinical staging distribution of all cases is summarized as follows: Stage I, 7 cases; Stage II, 32 cases; Stage III, 132 cases; Stage IVA, 84 cases; and Stage IVB, 14 cases [31].

Data preprocessing

Removal of CT scanning table

A known discrepancy between CT and MR images pertains to the presence of a scanning table in the former, which is absent in the latter. This inconsistency can lead to a misestimation by neural networks of the alignment

degree within the table region during the deformable registration process, thereby potentially diminishing the overall accuracy of registration. Therefore, a semi-automatic method was proposed for the removal of the scanning table from the CT images. Firstly, the outer contour could automatically be created in the TPS, and then the scanning table region was manually removed to obtain the patient's body mask. Subsequently, the mask region of the CT image was retained, and the region outside the mask was set as the background pixel value, as shown in Fig. 3.

Unification of image parameters

The variability in image resolution and slice thickness across different modalities can significantly impact the robustness of registration results. To mitigate this, it is essential to standardize image parameters. In this study, a dataset comprising 188 cases was designated for training, while a separate set of 81 cases was reserved for testing purposes. The initial step involved the resampling of all images to a uniform resolution of $1\text{ mm} \times 1\text{ mm} \times 3\text{ mm}$. Given the inconsistencies in image sizes post-resampling, the resampled images were uniformly cropped to a size of $448 \times 448 \times 64$ voxels. Subsequently, due to the limitation of the computer video memory size, each image slice was further resized to 224×224 voxels before the training process. To simplify the neural network training phase, 3Dslicer software [32] was employed to perform rigid registration on all cases, thereby reducing the complexity of the registration task. Finally, the image data were normalized to be distributed in the range of $(-1, 1)$.

Neural network model

The structure of improved CycleFCNs

The original CycleFCNs model consists of two fully convolutional networks (FCN) [33]. The outputs of the two FCNs were used as new inputs to each other to obtain

the reconstructed image for loss calculation, which can prevent some excessive deformation and make the network easier to converge. In addition, the model also used ResNet blocks to increase the depth of the network. The use of modality-independent neighborhood descriptor (MIND) metric can extract distinctive image structure by comparing each image patch with all its neighbors in a non-local region, and help to align the anatomical structures between multimodal images [34]. Regularization loss was also applied to deformable grids to prevent unreasonable deformation [35]. Based on the original CycleFCNs model, 3 ResNet blocks were added after each down-sampling layer, and skip connections were introduced to solve the problem of vanishing gradients, as shown in Fig. 4. The input of each FCN was a pair of CT and MR, and the image size was $224 \times 224 \times 24$. Three down-sampling layers were used to extract the features of the input images and the activation function was linear rectification function (ReLU). Then, three up-sampling layers were adopted to restore the high-dimensional features into a transform grid with the same size as the input images. Training and testing were performed on a computer with an Intel i9-10700KF CPU, 16GB Memory, NVIDIA GeForce RTX 3090 GPU, and 24 GB Video Memory.

Training strategy

(1) Adaptive mask registration strategy. Given the expansive and irregularly shaped target volumes characteristic of NPC cases, MR and CT imaging often necessitates a scanning range that extends from the chest to the crown of the head. The limited FOV may result in incomplete shoulder anatomical structure in MR scan, as shown in Fig. 1. Such inconsistencies in structural representation between multimodal images can adversely impact the accuracy of the overall registration process. To solve this problem, we have devised an adaptive mask registration

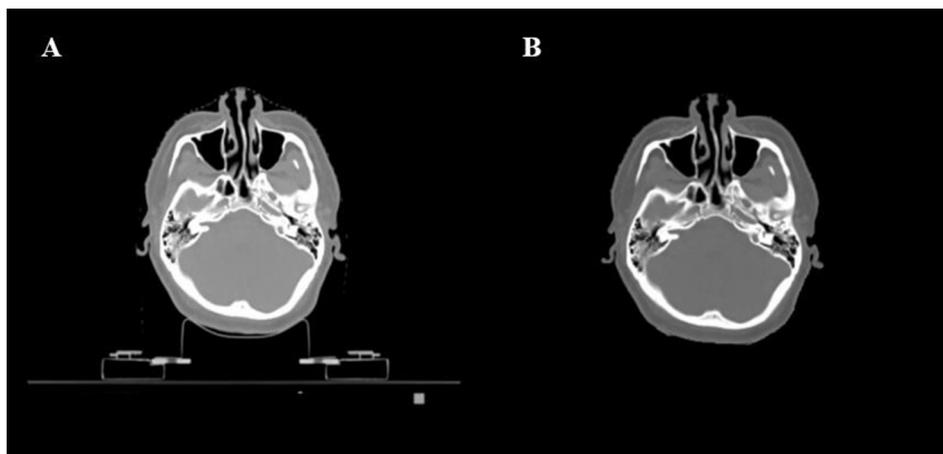


Fig. 3 CT image before and after scanning table removal. **A:** CT with table, **B:** CT without table

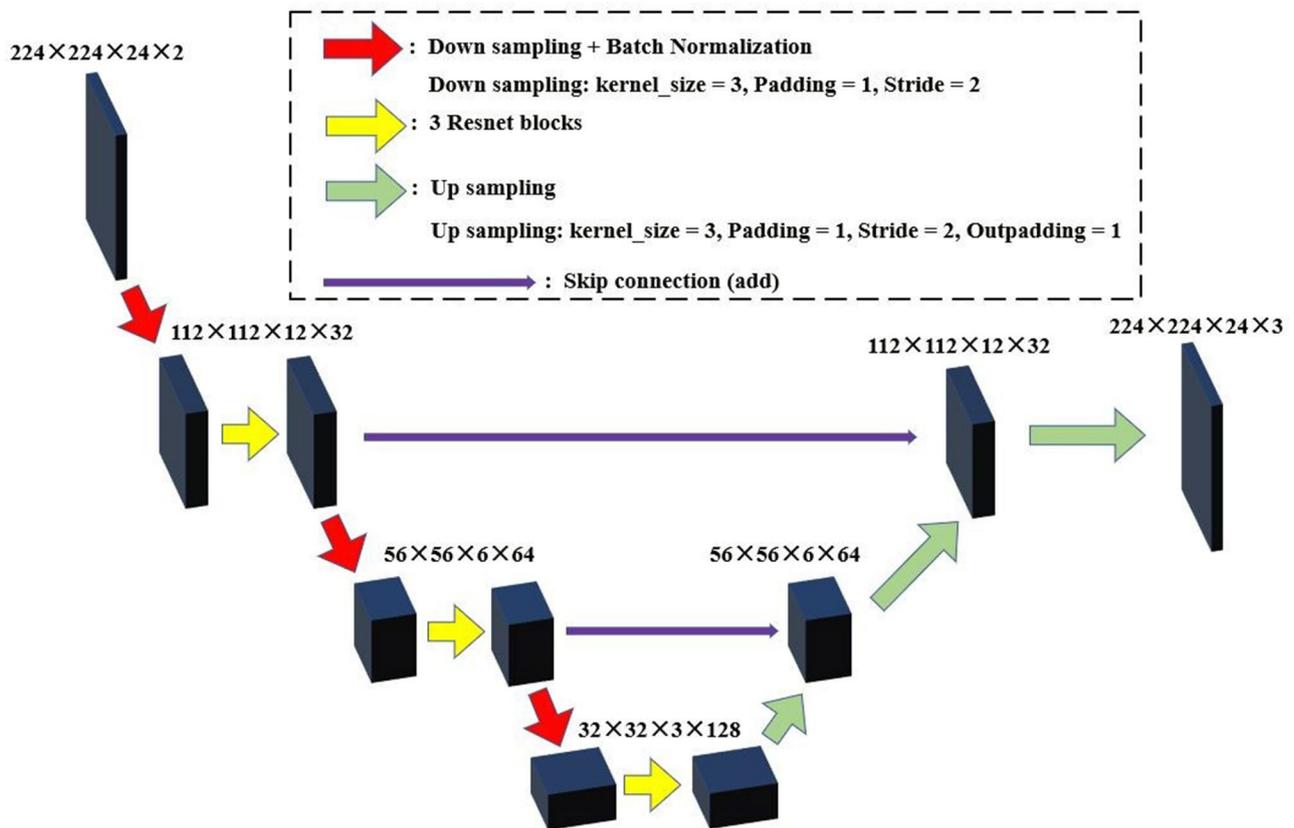


Fig. 4 The structure of the FCN in improved CycleFCNs

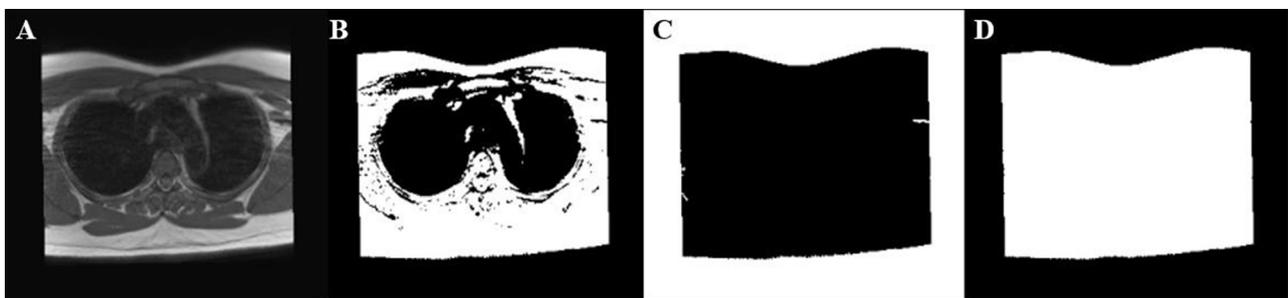


Fig. 5 The outer contour mask extraction process. **A:** origin MR, **B:** foreground and background, **C:** background mask, **D:** outer contour mask

strategy that involves the extraction of an outer contour mask from the MR, which is then defined as the ROI for registration. This approach ensures that the neural network concentrates on the deformable registration results within the mask during the training phase. The methodology for extracting the outer contour mask was as follows: Initially, the foreground and background of MR image were segmented based on the OTSU algorithm [36]. Subsequently, a background mask was obtained by flood-fill algorithm, and the edge cavity of the inverted background mask was repaired by morphological closure operation to get the outer contour mask as shown in Fig. 5. While some studies have also utilized mask-based training strategies in lung deformable registration to limit

changes in lung volume, this study differs by employing an adaptive mask to prevent registration errors caused by incomplete scanning of MR images. During the training process, both MR image and its mask was mapped based on the deformable grid. This adaptive mask prevents unreasonable deformation for MR image and also prevents the issue of mismatch between the mapped MR and the mask. Furthermore, the mask loss term and the mask constraints on other loss terms were introduced.

(2) Weight allocation strategy. Since most image registration algorithms only consider global registration, the mismatch of image edges as shown in Fig. 2 may lead to unreasonable registration results. The adaptive mask registration strategy was incapable of correcting registration

errors caused by the discrepancies in the scanning inclination angles between CT and MR. To ameliorate this issue, a weighted allocation strategy was introduced to improve the stability and precision of the registration results by increasing the registration error weight of the central slice region and decreasing the registration error weight of the edge slice region. Specifically, the weights of different regions were set as follows: The middle 18 slices of MR were set as the central slice region, while the three slices of the upper or lower boundary of the MR were set as the edge slice region. The weight of the central region was set to 2, and the weight of the edge region was set to 0.5.

Loss functions

The original loss functions of CycleFCNs were employed, including content loss, regularization loss, and cycle loss [30]. The MIND values of different registration regions were calculated based on the weight allocation strategy in the content loss to compare the image gradient difference between registered image pairs, as shown in formula (1). The N represents the number of image voxels, R is the spatial search region and we set the region size to 7×7 during model training. The regularization loss made the deformable grid smoother and prevented unreasonable registration results. The cycle loss could transform the deformed image back to the original image, making the registration results more reliable. In addition to the above loss terms, a mask loss has been added to constrain the occurrence of unreasonable deformations at the outer contours, as shown in formula (2), where V represents the volume and ϕ represents the deformable field. All loss terms were calculated only within the image mask, rather than across the entire image.

$$\begin{aligned} loss_{content} = & \frac{2}{N_1|R|} \sum_{n \in N_1} \sum_{r \in R} \\ & |MIND(CT, n, r) - MIND(MR, n, r)| \\ & + \frac{1}{2N_2|R|} \sum_{n \in N_2} \sum_{r \in R} \\ & |MIND(CT, n, r) - MIND(MR, n, r)| \end{aligned} \quad (1)$$

$$loss_{mask} = |V(\phi(mask_{MR})) - V(mask_{MR})| \quad (2)$$

Evaluation methods

Other models

The proposed model was compared with the following three deformable registration methods: (1) the hybrid intensity and structure-based deformable registration method in RayStation TPS (9.0 version), with the deformation strategy set to default, similarity measure set to mutual information, and resolution grid set to 0.25 cm, (2) the open-source image registration tool Elastix [37], (3) the classical image registration network VoxelMorph

[38]. The registration parameters in Elastix; interpolator is “BSpline Interpolator”, Optimizer is “Adaptive Stochastic Gradient Descent”, Transform is “BSpline Transform”, Metric is “Advanced Mattes Mutual Information” and “Maximum Number of Iterations” is 5000.

To evaluate the importance of the two proposed training strategies, we conducted a series of ablation experiments. The experimental design is as follows: (1) Baseline model, which includes all two strategies: adaptive mask registration strategy, and weight allocation strategy. (2) WAS model, which removes the adaptive mask registration strategy, while retaining the weight allocation strategy. (3) AMS model, which removes the weight allocation strategy, while retaining the adaptive mask registration strategy. (4) NAS model, which removes the two training strategies.

Evaluation metrics

For evaluation, the ROIs including spinal cord, thyroid, musculus constrictor pharynges, oral cavity, brainstem, temporomandibular joint, temporal lobe, pituitary, eyes, optic nerve, and optic chiasma were delineated on both CT and MR by a senior physician and reviewed by another senior physician before registration. The assessment of registration accuracy was based on several established indices: dice similarity coefficient (DSC), Hausdorff distance (HD), precision, and recall. The formulas for precision and recall are presented as Eqs. (2) and (3), respectively, where V_{MR} and V_{CT} denote the volumes of the respective ROIs within the deformed MR and CT. The greater the accuracy of registration results, the greater the DSC, precision, and recall, and the smaller the HD. Wilcoxon signed-rank test in SPSS 19 software was used to compare the registration results between CycleFCNs and RayStation.

$$Precision = \frac{V_{MR} \cap V_{CT}}{V_{MR}} \quad (2)$$

$$Recall = \frac{V_{MR} \cap V_{CT}}{V_{CT}} \quad (3)$$

Results

Registration results using rigid registration and deformable registration

Figure 6 shows the axial, sagittal, and coronal planes of the registration results using different methods. Although the rigid registration method provided a coarse alignment between CT and MR, it is evident that the local regional details of the two image sets were not completely matched, as depicted in Fig. 6A. The intensity-based deformable registration method demonstrated a certain degree of success in rectifying the soft tissue displacement between the multimodal images, with particularly

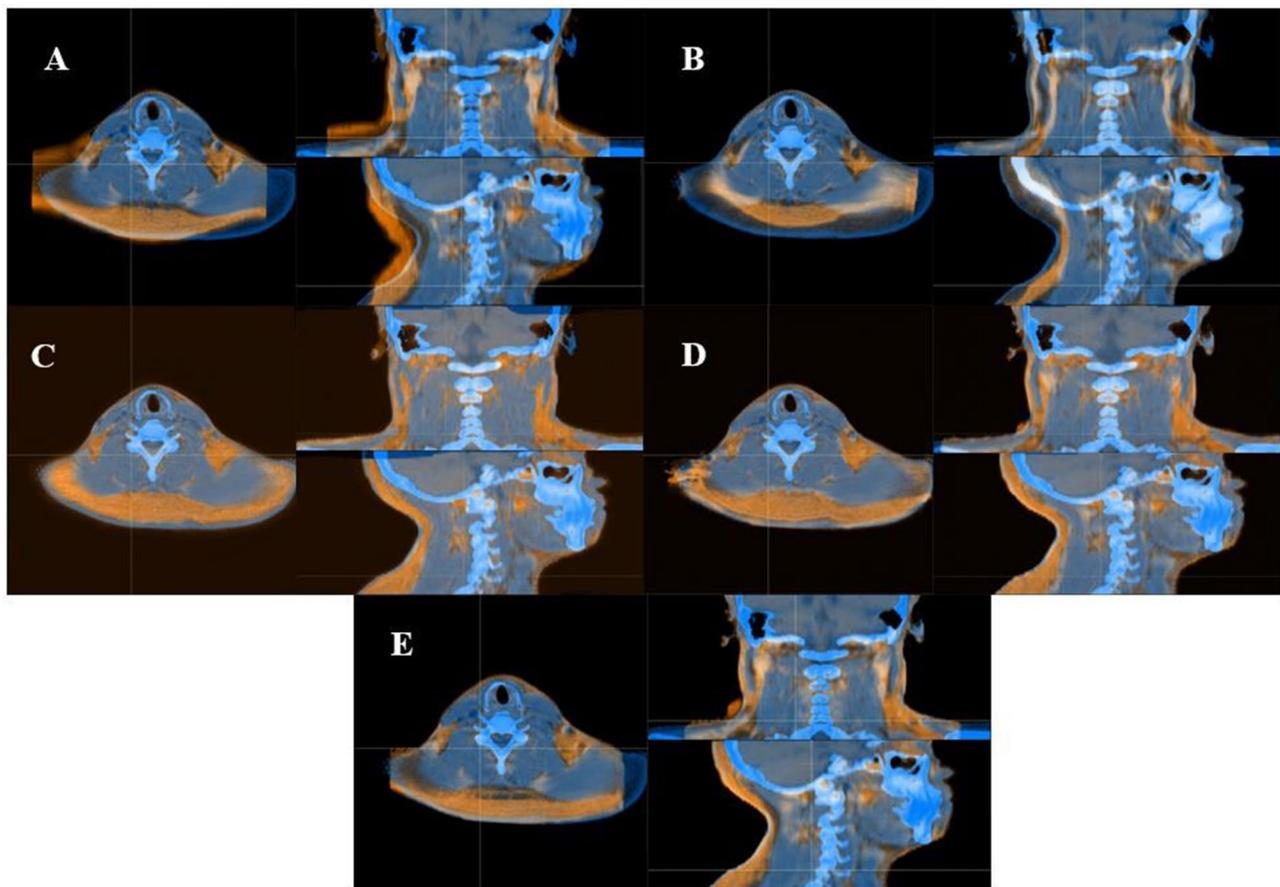


Fig. 6 Overlay fusion images using different registration methods. **A:** rigid registration in RayStation, **B:** hybrid intensity and structure-based deformable registration in RayStation, **C:** Elastix, **D:** VoxelMorph, **E:** improved CycleFCNs

notable enhancements in the alignment observed in the coronal and sagittal planes. Nevertheless, a shortcoming is observed in the axial plane, as illustrated in Fig. 6B, where the intensity-based approach encounters difficulties in accurately aligning the body's periphery. In contrast to the preceding methods, Elastix, VoxelMorph, and the improved CycleFCNs model exhibited marked improvements in the overall alignment between CT and MR.

To further illustrate the local registration effect of the three methods, the checkerboard fusion images are shown in Fig. 7. Images (1) and (2) represent two different slices, with three regions outlined in red in images (1) and three regions outlined in green in images (2) being highlighted to emphasize areas with significant registration differences. Although the intensity-based deformable registration method offered an improvement in the registration accuracy of soft tissues, it occasionally introduced unreasonable deformations in localized areas, leading to image distortion, as observed in Fig. 7A (1) and Fig. 7A (2). Conversely, Elastix, VoxelMorph, and the improved CycleFCNs model all achieved a more substantial enhancement in overall image registration accuracy.

Among them, the CycleFCNs model avoided excessive stretching of MR images in edge regions, which is reasonable given the limited FOV of MR images. However, in Elastix and VoxelMorph, areas outside the FOV of MR images were filled in, and their accuracies were questionable, as shown in Fig. 7B (2) and Fig. 7B (3). The CycleFCNs model also had some issues. In order to achieve higher registration accuracy in the central slices, there are some unreasonable deformation results in the peripheral slices, as shown in Fig. 7B (4).

Segmentation results using different deformable registration

Tables 1 and 2 respectively presented the DSC and HD values obtained through different deformable registration methods. The registration results of RayStation and Elastix were essentially consistent, showing no significant differences. A discernible trend was observed where the DSC has increased and the HD has decreased following the application of deep learning-based methods, as compared to the metrics derived from RayStation. Notably, for the temporomandibular joint, pituitary, optic nerves, and optic chiasma, there was a significant improvement

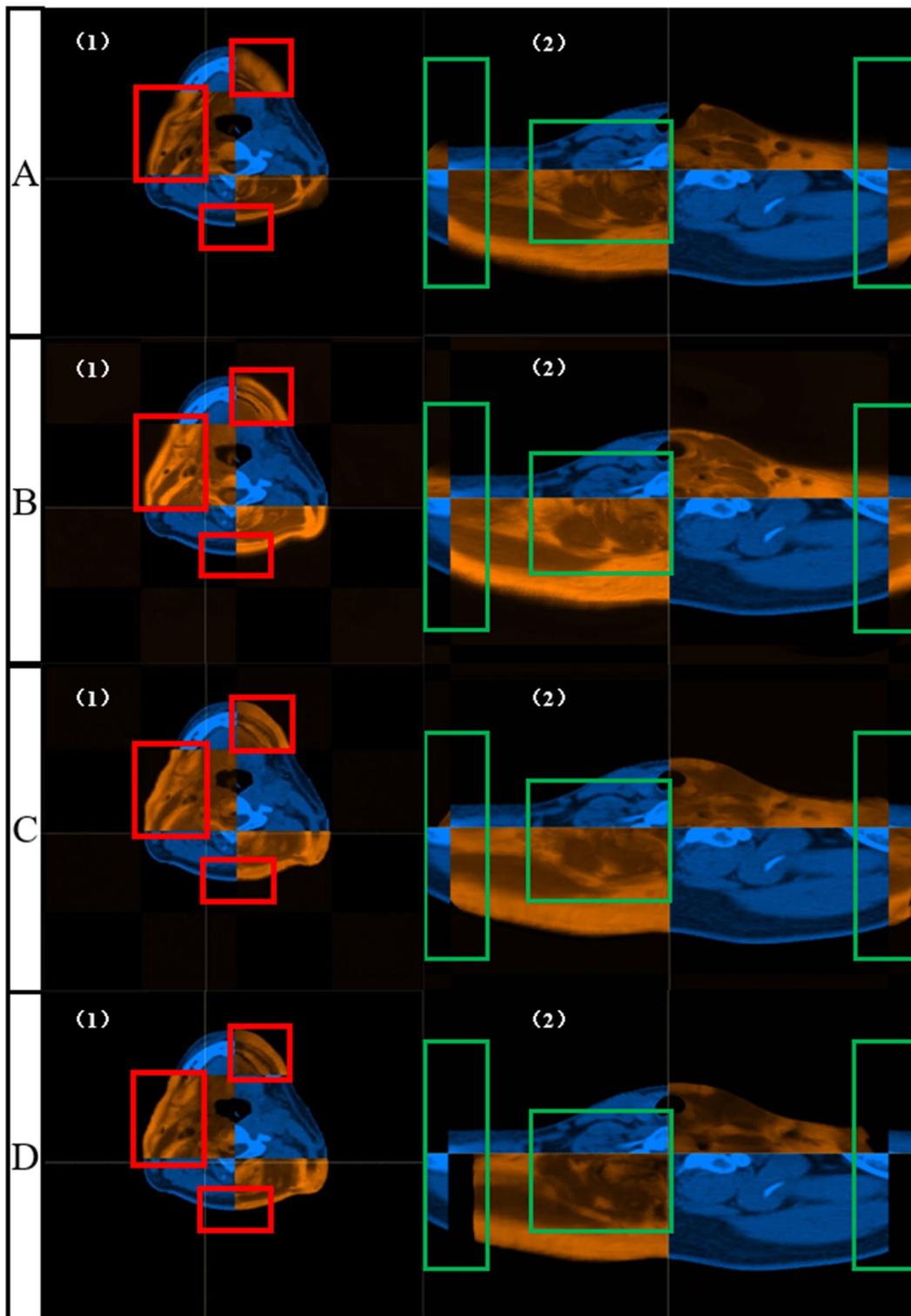


Fig. 7 Checkboard fusion images using different registration methods. **A:** hybrid intensity and structure-based deformable registration in RayStation, **B:** Elastix, **C:** VoxelMorph, **D:** improved CycleFCNs

Table 1 DSC of different organs in test cases based on different deformable registration

organ	Deformable registration (Mean ± SD)			
	RayStation	Elastix	VoxelMorph	CycleFCNs
Spinal cord	0.69±0.07	0.69±0.06	0.72±0.06	0.71±0.08
Thyroid	0.82±0.05	0.84±0.05	0.84±0.04	0.84±0.04
Musculus constrictor pharynges	0.82±0.04	0.81±0.06	0.83±0.05	0.85±0.05
Oral cavity	0.92±0.03	0.94±0.02	0.92±0.02	0.93±0.02
Brainstem	0.92±0.01	0.93±0.02	0.94±0.02	0.95±0.01
Temporomandibular joint	0.86±0.06	0.87±0.05	0.88±0.04	0.91±0.03
Temporal lobe	0.88±0.05	0.87±0.04	0.90±0.03	0.91±0.05
Pituitary	0.87±0.05	0.86±0.04	0.89±0.03	0.93±0.02
Eyes	0.90±0.04	0.88±0.04	0.91±0.02	0.91±0.04
Optic nerves	0.85±0.06	0.86±0.05	0.86±0.04	0.89±0.05
Optic chiasma	0.77±0.07	0.79±0.06	0.80±0.07	0.83±0.06

Table 2 HD of different organs in test cases based on different deformable registration

organ	Deformable registration (Mean ± SD) / mm			
	RayStation	Elastix	VoxelMorph	CycleFCNs
Spinal cord	4.11±0.91	4.08±0.92	3.89±0.99	3.64±0.93
Thyroid	3.71±0.87	3.75±0.88	3.24±0.80	3.10±0.84
Musculus constrictor pharynges	2.25±0.63	2.02±0.57	2.03±0.48	1.92±0.49
Oral cavity	7.21±1.38	8.01±1.44	6.55±1.46	6.17±1.78
Brainstem	2.86±0.94	2.57±1.01	2.33±0.87	2.35±0.83
Temporomandibular joint	2.98±1.28	2.81±1.14	2.56±0.97	2.28±0.86
Temporal lobe	13.74±3.52	12.56±3.45	12.44±3.61	12.43±3.57
Pituitary	1.83±0.49	1.86±0.56	1.69±0.51	1.53±0.54
Eyes	2.95±1.19	2.82±1.04	2.89±1.09	3.01±1.15
Optic nerves	3.74±0.88	3.91±0.99	3.78±0.84	3.56±0.83
Optic chiasma	5.94±1.45	5.89±1.50	5.88±1.60	5.87±1.72

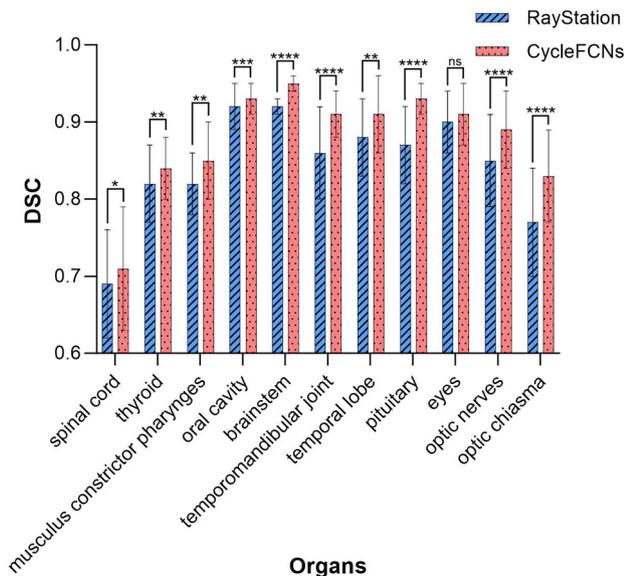


Fig. 8 DSC histogram of different organs using RayStation deformable registration and CycleFCNs model

in the average DSC, which rose from 0.86 to 0.91, 0.87 to 0.93, 0.85 to 0.89, and 0.77 to 0.83, respectively, after the deformable registration based on the improved CycleFCNs model. Concurrently, the average HD exhibited a reduction, from 2.98 mm to 2.28 mm, 1.83 mm to 1.53 mm, 3.74 mm to 3.56 mm, and 5.94 mm to 5.87 mm, for the same anatomical structures. When compared with the deformable registration results produced by RayStation, the CycleFCNs model achieved a higher average DSC for most organs and a correspondingly lower average HD.

RayStation is one of the commonly used treatment planning systems in clinical practice. The Wilcoxon signed-rank test was employed to compare the registration accuracy between RayStation and the proposed model, to assess whether the proposed model can enhance the registration accuracy in the clinical workflow. Statistical analysis revealed that, for most normal tissues, the differences in DSC and HD between the two deformable registration methods were statistically significant ($P < 0.05$). However, for eyes, optic nerves, and optic chiasma, the registration results of CycleFCNs were comparable to those of RayStation, with the P values for HD being 0.827, 0.385, and 0.808, respectively. Figures 8

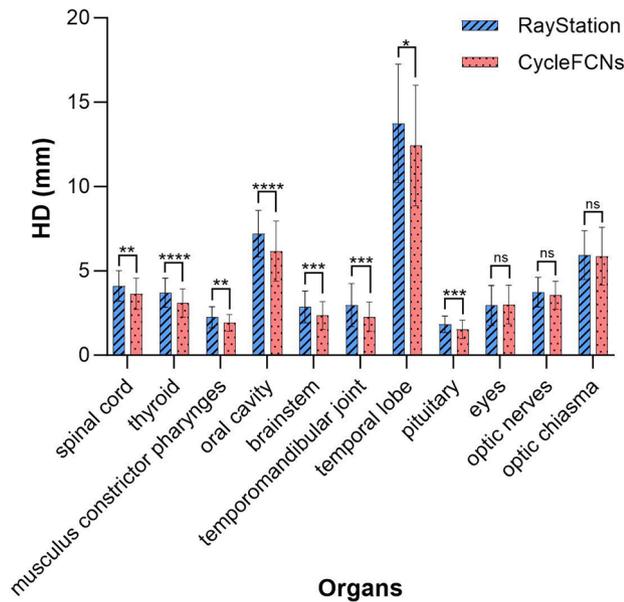


Fig. 9 HD histogram of different organs using RayStation deformable registration and CycleFCNs model

and 9 graphically represented the DSC and HD histograms for various organs, as determined by RayStation's

deformable registration and the CycleFCNs model. The CycleFCNs model consistently demonstrated higher DSC values and lower HD values for the majority of the organs analyzed.

The precision and recall values of different normal tissues in test cases based on different deformable registration methods are shown in Tables 3 and 4. Similar to the results of DSC and HD, the registration accuracy of the two deep learning-based methods was better than that of RayStation and Elastix. Furthermore, the precision and recall values derived from the CycleFCNs model were observed to be superior to those obtained through VoxelMorph for most organs. However, it is important to note that for specific anatomical structures, including the spinal cord, temporal lobe, and eyes, the differences in precision between RayStation and CycleFCNs did not reach statistical significance ($P > 0.05$). Similarly, for the musculus constrictor pharynges, the differences in recall were not statistically significant ($P > 0.05$). The precision and recall histograms depicted in Figs. 10 and 11 further substantiated the superior registration accuracy of the CycleFCNs model when compared to the conventional deformable registration method.

Table 3 Precision of different organs in test cases based on different deformable registration

organ	deformable registration (Mean ± SD)			
	RayStation	Elastix	VoxelMorph	CycleFCNs
Spinal cord	0.71 ± 0.06	0.70 ± 0.05	0.73 ± 0.05	0.72 ± 0.06
Thyroid	0.81 ± 0.07	0.84 ± 0.06	0.84 ± 0.04	0.83 ± 0.07
Musculus constrictor pharynges	0.80 ± 0.04	0.83 ± 0.07	0.84 ± 0.05	0.88 ± 0.07
Oral cavity	0.91 ± 0.03	0.93 ± 0.03	0.94 ± 0.01	0.94 ± 0.02
Brainstem	0.90 ± 0.02	0.94 ± 0.02	0.94 ± 0.02	0.94 ± 0.02
Temporomandibular joint	0.84 ± 0.07	0.88 ± 0.04	0.90 ± 0.04	0.94 ± 0.03
Temporal lobe	0.91 ± 0.07	0.86 ± 0.05	0.90 ± 0.03	0.91 ± 0.07
Pituitary	0.86 ± 0.04	0.87 ± 0.04	0.91 ± 0.02	0.95 ± 0.01
Eyes	0.90 ± 0.04	0.89 ± 0.04	0.90 ± 0.02	0.91 ± 0.04
Optic nerves	0.75 ± 0.10	0.83 ± 0.06	0.85 ± 0.04	0.82 ± 0.05
Optic chiasma	0.75 ± 0.08	0.82 ± 0.05	0.79 ± 0.06	0.82 ± 0.06

Table 4 Recall of different organs in test cases based on different deformable registration

organ	deformable registration (Mean ± SD)			
	RayStation	Elastix	VoxelMorph	CycleFCNs
Spinal cord	0.66 ± 0.08	0.67 ± 0.06	0.71 ± 0.06	0.71 ± 0.07
Thyroid	0.82 ± 0.04	0.84 ± 0.04	0.84 ± 0.04	0.85 ± 0.04
Musculus constrictor pharynges	0.84 ± 0.06	0.80 ± 0.06	0.82 ± 0.05	0.84 ± 0.05
Oral cavity	0.92 ± 0.04	0.94 ± 0.01	0.90 ± 0.03	0.93 ± 0.02
Brainstem	0.93 ± 0.01	0.92 ± 0.02	0.94 ± 0.02	0.96 ± 0.01
Temporomandibular joint	0.87 ± 0.05	0.86 ± 0.07	0.86 ± 0.05	0.89 ± 0.06
Temporal lobe	0.87 ± 0.05	0.88 ± 0.03	0.90 ± 0.03	0.92 ± 0.05
Pituitary	0.87 ± 0.05	0.86 ± 0.04	0.87 ± 0.03	0.91 ± 0.03
Eyes	0.89 ± 0.04	0.88 ± 0.03	0.91 ± 0.02	0.93 ± 0.02
Optic nerves	0.82 ± 0.08	0.88 ± 0.05	0.86 ± 0.04	0.90 ± 0.04
Optic chiasma	0.78 ± 0.05	0.76 ± 0.07	0.80 ± 0.07	0.83 ± 0.05

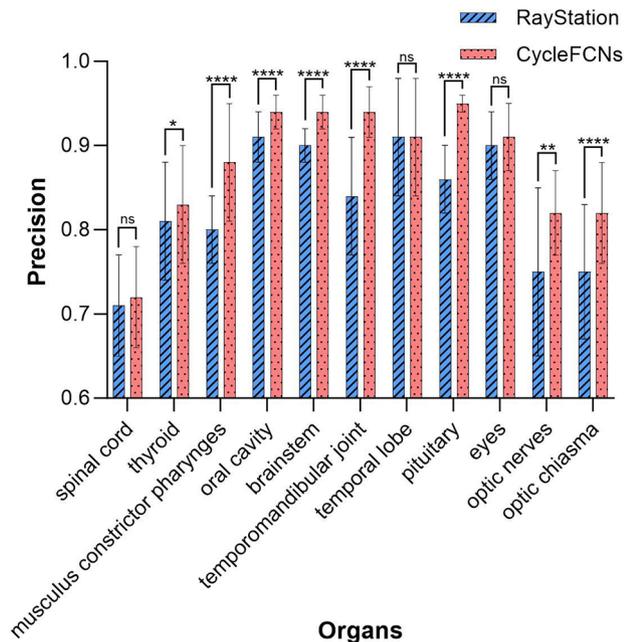


Fig. 10 Precision histogram of different organs using RayStation deformable registration and CycleFCNs model

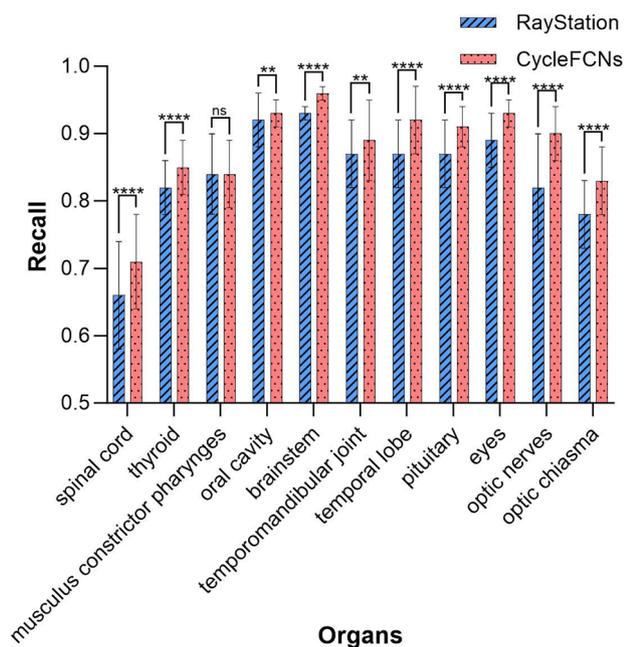


Fig. 11 Recall histogram of different organs using RayStation deformable registration and CycleFCNs model

Ablation experiments

A series of ablation experiments was conducted to verify the effectiveness of the adaptive mask registration strategy and weight allocation strategy. It was observed that the NAS model while attempting to align the MR and CT, resulted in an undesirable stretching effect as highlighted within the white box in Fig. 12A. Although the adaptive

mask registration strategy has improved the alignment of the spine to some extent, the accuracy remained insufficient. The weight allocation strategy significantly enhanced the registration accuracy of bony structure. However, this approach, by focusing more computational weight of the loss function on the central slices, resulted in poorer alignment of the edge regions, as illustrated in the green boxes of Fig. 12C and D.

In terms of DSC, the proposed method exhibited superior performance compared to the other three models. Owing to the weight allocation strategy, the WAS model demonstrated performance comparable to the baseline model for OARs primarily located in the central slices. However, the registration accuracy for the spinal cord, optic nerves, and optic chiasma was slightly inferior. The AMS model exhibited slightly lower registration accuracy for most organs compared to the WAS model, except for OARs in more peripheral slices, such as the spinal cord. The NAS model performed the worst. In terms of HD, precision, and recall, we also obtained similar results, as shown in Tables 6, 7 and 8.

Discussion

Image registration between planning CT and MR is the key to accurate target segmentation of NPC cases. Deformable registration is suitable for multi-modal image registration due to its ability to correct the deformation and displacement of internal soft tissues. Although many deep learning models have been proposed for the deformable registration to increase the registration accuracy, several difficulties prevented the widespread application of these models in clinical practice. Specifically, differences in FOV and scanning inclination angle between CT and MR may lead to registration errors [39–40]. In this study, the CycleFCNs model, which adopts two synchronously trained fully convolutional networks to ensure the registration results maintaining inverse consistency, was used to perform deformable registration between planning CT and MR for NPC cases. Adaptive mask registration strategy and weight allocation strategy were proposed to improve the registration accuracy.

In this study, both the intensity-based iterative algorithm and the deep learning model were capable of rectifying the soft tissue deformation and displacement between multimodal images for NPC cases. Compared with the registration results in RayStation, the proposed method improved the contour similarity of each ROI between CT and MR, as shown in Tables 1, 2 and 3, and 4. The American Association of Physicists in Medicine TG-132 report summarized the quantitative metrics and their tolerance for image registration, with a recommended DSC value of 0.8 to 0.9 [41]. Although the four deformable registration methods can achieve adequate accuracy for most organs, deep learning methods

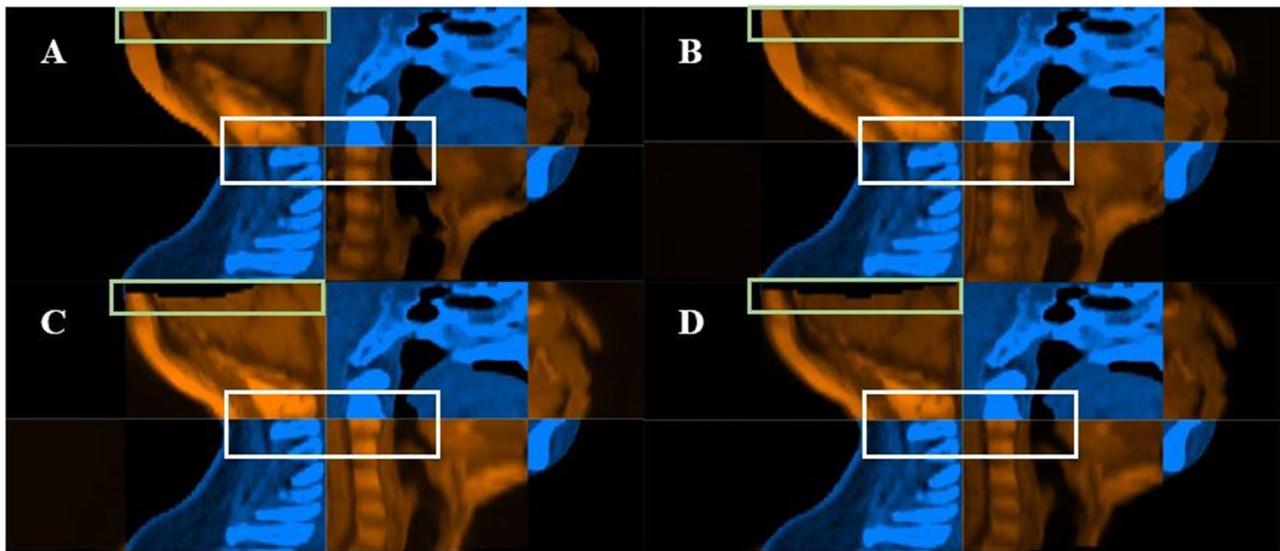


Fig. 12 Checkboard fusion images. **A:** NAS model (origin CycleFCNs), **B:** AMS model (CycleFCNs with adaptive mask registration strategy), **C:** AMS model (CycleFCNs with weight allocation strategy), **D:** Baseline model (CycleFCNs with two strategies)

exhibited higher DSC values compared to the two conventional iterative algorithms, with CycleFCNs demonstrating the best performance.

Figures 6 and 7 showed the results of rigid registration and the four deformable registration methods. Although the intensity-based deformable registration methods implemented in RayStation improved the registration accuracy compared to the rigid algorithm, several challenges remained. First of all, the conventional algorithm did not fully rectify the displacement and deformation of soft tissues, as depicted in Fig. 6B. Additionally, the emergence of some unreasonable deformations in localized regions raised concerns regarding the reliability of the registration results. Elastix and VoxelMorph stretched the edges of the MR images to fill in the missing shoulder parts, but such unreasonable deformations may affect the reliability of registration in other areas. In contrast, CycleFCNs ensured the alignment of the central region while avoiding excessive stretching of the peripheral areas of the MR images.

The Wilcoxon signed-rank test results presented from Figs. 8, 9, 10 and 11 indicated that the four quantitative metrics of RayStation and the method described in this study exhibited statistical significance in the majority of normal tissues. Notably, for the oral cavity, brainstem, temporomandibular joint, and pituitary, the improved CycleFCNs model has demonstrated significant advantages. Zhao et al. developed a ViT-Morph model for MR-MR deformable registration and evaluated 7 manually contoured organs (brainstem, spinal cord, mandible, left/right parotids, left/right submandibular glands) by comparing with the traditional registration methods in Monaco treatment planning system [42]. Their findings

also corroborated the notion that deformable registration based on deep learning can achieve higher accuracy in comparison to conventional methods.

To further verify the effectiveness of the adaptive mask registration strategy and weight allocation strategy in improving registration accuracy, a series of ablation experiments was conducted. As illustrated in Fig. 12, it is evident that both training strategies have their respective advantages. The adaptive mask strategy can prevent excessive stretching of MR images, but it is insufficient in terms of registration accuracy for the central slices. The weight allocation strategy improves the registration accuracy of the central slices and is more effective for cases where ROIs are concentrated in the central slices of image. However, this strategy can also lead to undesirable stretching effects at the edges of the MR images. The reliability of such registration results remain further validation, primarily since the FOV in MR is limited compared to CT, resulting in the absence of information outside the FOV. As shown in Tables 5, 6, 7 and 8, the AMS model demonstrated slightly lower registration accuracy for most organs compared to the WAS model. This may be attributed to the fact that the adaptive mask strategy was primarily designed to address the issue of limited FOV in MR images. Since most OARs for NPC are located in the head and neck region, where MR images are complete, the adaptive mask strategy could not exert its effect. For areas below the neck, where MR images suffer from insufficient FOV, the AMS model could enhance the registration accuracy of OARs in these regions. For the temporal lobe and eyes, the average DSC and HD values among the four models were not significantly different, indicating that the additional training strategies exerted

Table 5 DSC of different organs in test cases based on different models

organ	DSC (Mean ± SD)			
	Baseline	WAS	AMS	NAS
Spinal cord	0.71 ± 0.08	0.69 ± 0.07	0.71 ± 0.07	0.68 ± 0.06
Thyroid	0.84 ± 0.04	0.84 ± 0.04	0.83 ± 0.03	0.82 ± 0.07
Musculus constrictor pharynges	0.85 ± 0.05	0.85 ± 0.06	0.84 ± 0.05	0.82 ± 0.04
Oral cavity	0.93 ± 0.02	0.93 ± 0.02	0.93 ± 0.03	0.93 ± 0.03
Brainstem	0.95 ± 0.01	0.95 ± 0.02	0.94 ± 0.02	0.94 ± 0.01
Temporomandibular joint	0.91 ± 0.03	0.92 ± 0.05	0.89 ± 0.04	0.88 ± 0.05
Temporal lobe	0.91 ± 0.05	0.90 ± 0.04	0.90 ± 0.03	0.89 ± 0.06
Pituitary	0.93 ± 0.02	0.92 ± 0.03	0.91 ± 0.02	0.90 ± 0.03
Eyes	0.91 ± 0.04	0.90 ± 0.03	0.91 ± 0.04	0.91 ± 0.03
Optic nerves	0.89 ± 0.05	0.87 ± 0.05	0.88 ± 0.06	0.84 ± 0.06
Optic chiasma	0.83 ± 0.06	0.81 ± 0.07	0.83 ± 0.08	0.79 ± 0.08

Note: baseline model includes all two strategies; WAS model removes the adaptive mask registration strategy; AMS model removes the weight allocation strategy; NAS model removes all the two training strategies

Table 6 HD of different organs in test cases based on different models

organ	HD (Mean ± SD) / mm			
	Baseline	WAS	AMS	NAS
Spinal cord	3.64 ± 0.93	3.84 ± 0.98	3.71 ± 0.89	3.92 ± 1.07
Thyroid	3.10 ± 0.84	3.27 ± 0.81	3.33 ± 0.85	3.71 ± 0.86
Musculus constrictor pharynges	1.92 ± 0.49	1.93 ± 0.50	1.95 ± 0.46	2.15 ± 0.44
Oral cavity	6.17 ± 1.78	6.06 ± 1.61	6.62 ± 1.70	7.02 ± 1.49
Brainstem	2.35 ± 0.83	2.42 ± 0.95	2.67 ± 0.82	2.80 ± 0.81
Temporomandibular joint	2.28 ± 0.86	2.12 ± 0.78	2.43 ± 0.82	2.55 ± 1.02
Temporal lobe	12.43 ± 3.57	12.45 ± 3.68	12.87 ± 3.88	13.29 ± 4.11
Pituitary	1.53 ± 0.54	1.78 ± 0.71	1.61 ± 0.62	1.70 ± 0.67
Eyes	3.01 ± 1.15	3.13 ± 1.25	2.98 ± 1.01	3.02 ± 1.19
Optic nerves	3.56 ± 0.83	3.76 ± 0.90	3.60 ± 0.77	3.47 ± 0.94
Optic chiasma	5.87 ± 1.72	6.02 ± 1.80	5.90 ± 1.71	6.16 ± 1.70

Note: baseline model includes all two strategies; WAS model removes the adaptive mask registration strategy; AMS model removes the weight allocation strategy; NAS model removes all the two training strategies

a negligible impact on the registration accuracy for these particular organs.

The proposed method minimizes the necessity for manual adjustments and offers more dependable anatomical updates, ensuring that the treatment plan stays consistent with the patient’s current anatomical state. This is crucial for maintaining the accuracy and effectiveness of full-auto planning and adaptive radiotherapy [43]. Given the general nature of the proposed strategies, the proposed method has the potential to be scalable to other anatomical regions or cancer types. The adaptive mask registration strategy, which focuses on the ROIs by extracting an outer contour mask, can be adapted to other anatomical regions. For example, in prostate

Table 7 Precision of different organs in test cases based on different models

organ	Precision (Mean ± SD)			
	Baseline	WAS	AMS	NAS
Spinal cord	0.72 ± 0.06	0.70 ± 0.07	0.72 ± 0.06	0.71 ± 0.05
Thyroid	0.83 ± 0.07	0.83 ± 0.05	0.82 ± 0.05	0.83 ± 0.06
Musculus constrictor pharynges	0.88 ± 0.07	0.86 ± 0.06	0.84 ± 0.05	0.82 ± 0.04
Oral cavity	0.94 ± 0.02	0.94 ± 0.02	0.92 ± 0.02	0.91 ± 0.03
Brainstem	0.94 ± 0.02	0.95 ± 0.02	0.94 ± 0.02	0.95 ± 0.02
Temporomandibular joint	0.94 ± 0.03	0.93 ± 0.04	0.90 ± 0.03	0.89 ± 0.06
Temporal lobe	0.91 ± 0.07	0.90 ± 0.03	0.92 ± 0.04	0.90 ± 0.07
Pituitary	0.95 ± 0.01	0.94 ± 0.03	0.92 ± 0.02	0.92 ± 0.05
Eyes	0.91 ± 0.04	0.89 ± 0.03	0.90 ± 0.04	0.90 ± 0.04
Optic nerves	0.82 ± 0.05	0.84 ± 0.05	0.84 ± 0.07	0.86 ± 0.05
Optic chiasma	0.82 ± 0.06	0.80 ± 0.06	0.82 ± 0.08	0.81 ± 0.08

Note: baseline model includes all two strategies; WAS model removes the adaptive mask registration strategy; AMS model removes the weight allocation strategy; NAS model removes all the two training strategies

Table 8 Recall of different organs in test cases based on different models

organ	Recall (Mean ± SD)			
	Baseline	WAS	AMS	NAS
Spinal cord	0.71 ± 0.07	0.68 ± 0.07	0.70 ± 0.07	0.67 ± 0.07
Thyroid	0.85 ± 0.04	0.84 ± 0.04	0.84 ± 0.04	0.82 ± 0.05
Musculus constrictor pharynges	0.84 ± 0.05	0.84 ± 0.06	0.84 ± 0.05	0.82 ± 0.06
Oral cavity	0.93 ± 0.02	0.92 ± 0.02	0.94 ± 0.03	0.94 ± 0.04
Brainstem	0.96 ± 0.01	0.95 ± 0.02	0.95 ± 0.02	0.93 ± 0.01
Temporomandibular joint	0.89 ± 0.06	0.91 ± 0.06	0.88 ± 0.05	0.86 ± 0.05
Temporal lobe	0.92 ± 0.05	0.90 ± 0.05	0.89 ± 0.03	0.89 ± 0.06
Pituitary	0.91 ± 0.03	0.91 ± 0.03	0.90 ± 0.03	0.89 ± 0.05
Eyes	0.93 ± 0.02	0.91 ± 0.03	0.92 ± 0.03	0.94 ± 0.04
Optic nerves	0.90 ± 0.04	0.89 ± 0.05	0.90 ± 0.05	0.84 ± 0.08
Optic chiasma	0.83 ± 0.05	0.81 ± 0.07	0.83 ± 0.07	0.77 ± 0.06

Note: baseline model includes all two strategies; WAS model removes the adaptive mask registration strategy; AMS model removes the weight allocation strategy; NAS model removes all the two training strategies

cancer, where the prostate gland is the primary target, a similar mask can be created to focus the registration on the prostate region. This strategy can help address issues related to limited FOV in MR scans, which is a common problem in many anatomical regions. The weight allocation strategy, which assigns higher weights to central slices and lower weights to edge slices, can be applied to other regions where edge mismatches are a concern. For instance, in brain tumor registration, where the brain’s surface can introduce registration errors, this strategy can help improve the registration accuracy by focusing on the central regions of the brain.

Deformable registration process still has some other potential risks, particularly concerning tumor boundaries, which can significantly impact clinical

decision-making. For instance, deformable registration may struggle with partial volume effects, where the size of the imaging voxels exceeds that of the structures being imaged. This can result in inaccuracies in the delineation of tumor boundaries, especially in regions where the tumor is small or has an irregular shape. Moreover, variations in image intensity within the tumor or surrounding tissues can introduce artifacts during the registration process, leading to misalignment of tumor boundaries by the deformable registration algorithm. In the future, we plan to conduct relevant research to address these issues. A limitation in this study is that the semi-automatic table removal process, while effective, may require additional effort from medical physicists and may be challenging to automate due to variability in CT scanners. Our future measures will focus on the development of automated tools to improve the efficiency of the table removal process.

Conclusions

In this work, adaptive mask registration strategy and weight allocation strategy were proposed to perform CT-MR deformable registration for NPC cases based on the CycleFCNs model. Compared with RayStation, Elastix, and VoxelMorph, the proposed method improved the registration accuracy for multi-modal medical images. Furthermore, we demonstrated that the adaptive mask registration strategy and weight allocation strategy allow for better deformable image registration performance.

Abbreviations

CT	Computed Tomography
MR	Magnetic Resonance
ROI	Regions of interest
NPC	Nasopharyngeal Carcinoma
FOV	Field of view
CycleFCNs	Cycle-consistent fully convolutional network
TPS	Treatment planning system
FCN	Fully convolutional network
MIND	Modality-independent neighborhood descriptor
ReLU	Linear rectification function
DSC	Dice similarity coefficient
HD	Hausdorff distance

Author contributions

Yi Guo designed the study. Jun Chen, Lin Lu, Lingna Qiu, and Linzhen Lan were responsible for data collection and preprocessing. Yi Guo performed the model training. Yi Guo wrote the manuscript. Feibao Guo and Jinsheng Hong revised the manuscript.

Funding

This research was supported by Natural Science Foundation of Fujian Province of China (2021J05146), Medical Science and Technology Project of Fujian Province of China (2022GGA025), and Leading Project Foundation of Science and Technology, Fujian Province of China (2023Y0017).

Data availability

No datasets were generated or analysed during the current study.

Declarations

Ethical approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 27 August 2024 / Accepted: 14 February 2025

Published online: 25 February 2025

References

1. IARC. Estimated number of incident cases from 2018 to 2040. 2018. <https://gco.iarc.fr/tomorrow/graphic-isotype>
2. Mahapatra D, Antony B, Sedai S et al. Deformable medical image registration using generative adversarial networks. In: 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018). Washington, DC, USA: IEEE; 2018:1449–1453.
3. Sloan JM, Goatman KA, Siebert JP. Learning rigid image registration-utilizing convolutional neural networks for medical image registration. In: 11th International Joint Conference on Biomedical Engineering Systems and Technologies. Funchal, Madeira, Portugal: SciTePress; 2018:89–99.
4. Yang F, Ding MY, Zhang XM, et al. Non-rigid multi-modal 3D medical image registration based on foveated modality independent neighborhood descriptor. *Sensors*. 2019;19(21):4675.
5. Zhu ZY, Cao YQ, Qin CC, et al. Joint affine and deformable three-dimensional networks for brain MRI registration. *Med Phys*. 2021;48(3):1182–96.
6. Sengupta D, Gupta P, Biswas A. A survey on mutual information based medical image registration algorithms. *Neurocomputing*. 2022;486:174–88.
7. Fu YB, Lei Y, Wang TH, et al. Deep learning in medical image registration: a review. *Phys Med Biol*. 2020;65(20):20TR01.
8. Boveiri HR, Khayami R, Javidan R, et al. Medical image registration using deep neural networks: a comprehensive review. *Comput Electr Eng*. 2020;87:106767.
9. Abbasi S, Tavakoli M, Boveiri HR, et al. Medical image registration using unsupervised deep neural network: a scoping literature review. *Biomed Signal Proces*. 2022;73:103444.
10. Villena-Martinez V, Oprea S, Saval-Calvo M, et al. When deep learning meets data alignment: a review on deep registration networks (DRNS). *Appl Sci-Basel*. 2020;10(21):7524.
11. Gu DD, Liu GC, Cao XH, et al. A consistent deep registration network with group data modeling. *Comput Med Imag Grap*. 2021;90:101904.
12. Yang BN, Chen XY, Li JW, et al. A feasible method to evaluate deformable image registration with deep learning-based segmentation. *Phys Med*. 2022;95:50–6.
13. Han X, Hong J, Reyngold M, et al. Deep-learning-based image registration and automatic segmentation of organs-at-risk in cone-beam CT scans from high-dose radiation treatment of pancreatic cancer. *Med Phys*. 2021;48(6):3084–95.
14. He ZQ, He YP, Cao WM. Deformable image registration with attention-guided fusion of multi-scale deformation fields. *Appl Intell*. 2023;53(3):2936–50.
15. Han R, Jones CK, Lee J, et al. Joint synthesis and registration network for deformable MR-CBCT image registration for neurosurgical guidance. *Phys Med Biol*. 2022;67(12):125008.
16. Mansilla L, Milone DH, Ferrante E. Learning deformable registration of medical images with anatomical constraints. *Neural Netw*. 2020;124:269–79.
17. Sun KC, Simon S. FDRN: A fast deformable registration network for medical images. *Med Phys*. 2021;48(10):6453–63.
18. van Eijnatten M, Rundo L, Batenburg KJ, et al. 3D deformable registration of longitudinal abdominopelvic CT images using unsupervised deep learning. *Comput Meth Prog Bio*. 2021;208:106261.
19. Eppenhof KAJ, Lafarge MW, Veta M, et al. Progressively trained convolutional neural networks for deformable image registration. *IEEE T Med Imaging*. 2020;39(5):1594–604.
20. Yang J, Yang JH, Zhao F, et al. An unsupervised multi-scale framework with attention-based network (MANet) for lung 4D-CT registration. *Phys Med Biol*. 2021;66(13):135008.

21. Cui KP, Lin YS, Liu Y, et al. Deep residual-SVD network for brain image registration. *Phys Med Biol*. 2022;67(14):144002.
22. Xu Z, Luo J, Yan JP, et al. F3RNet: full-resolution residual registration network for deformable image registration. *Int J Comput Ass Rad*. 2021;16(6):923–32.
23. Zheng YN, Jiang S, Yang ZY. Deformable registration of chest CT images using a 3D convolutional neural network based on unsupervised learning. *J Appl Clin Med Phys*. 2021;22(10):22–35.
24. Ma YJ, Niu DM, Zhang JS, et al. Unsupervised deformable image registration network for 3D medical images. *Appl Intell*. 2022;52(1):766–79.
25. Huang WJ, Yang H, Liu XF, et al. A coarse-to-fine deformable transformation framework for unsupervised multi-contrast MR image registration with dual consistency constraint. *IEEE T Med Imaging*. 2021;40(10):2589–99.
26. Johnstone E, Wyatt JJ, Henry AM, et al. Systematic review of synthetic computed tomography generation methodologies for use in magnetic resonance imaging-only radiation therapy. *Int J Radiat Oncol*. 2018;100(1):199–217.
27. Hirotaki K, Tomizawa K, Moriya S, et al. Impact of anatomical position errors on dose distribution in head and neck radiotherapy and robust image registration against anatomical changes. *Anticancer Res*. 2023;43(4):1827–34.
28. Lei Y, Fu YB, Tian Z, et al. Deformable CT image registration via a dual feasible neural network. *Med Phys*. 2022;49(12):7545–54.
29. McKenzie EM, Santhanam A, Ruan D, et al. Multimodality image registration in the head-and-neck using a deep learning derived synthetic CT as a bridge. *Med Phys*. 2020;47(3):1094–104.
30. Guo Y, Wu X, Wang Z, et al. End-to-end unsupervised cycle-consistent fully convolutional network for 3D pelvic CT-MR deformable registration. *J Appl Clin Med Phys*. 2020;21(9):193–200.
31. Amin MB, Edge S, Greene F, et al. *AJCC Cancer staging Manual*. 8th ed. New York, NY: Springer; 2016.
32. Pieper S, Michael H, Ron K. 3D Slicer. In: 2004 2nd IEEE international symposium on biomedical imaging: nano to macro (IEEE Cat No. 04EX821). Arlington, VA, USA: IEEE; 2004:632–635.
33. Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation. *IEEE T Pattern Anal*. 2015;39(4):640–51.
34. Yang H, Sun J, Carass A et al. Unpaired brain MR-to-CT synthesis using a structure-constrained CycleGAN. In: Jorge C, Tal A, Gustavo C Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. 2018 ed. Cham, Switzerland: Springer; 2018:174–182.
35. He K, Zhang X, Ren S et al. Deep residual learning for image recognition. In: proceedings of the IEEE conference on computer vision and pattern recognition 2016 (CVPR 2016). Las Vegas, Nevada, USA: IEEE; 2016:770–778.
36. Otsu N. A threshold selection method from gray-level histograms. *IEEE T Syst Man Cy B*. 1979;9(1):62–6.
37. Klein S, Staring M, Murphy K, et al. Elastix: a toolbox for intensity-based Medical Image Registration. *IEEE T Med Imaging*. 2010;29(1):196–205.
38. Balakrishnan G, Zhao A, Sabuncu MR, et al. VoxelMorph: a learning framework for deformable medical image registration. *IEEE T Med Imaging*. 2019;38(8):1788–800.
39. Ou YM, Akbari H, Bilello M, et al. Comparative evaluation of registration algorithms in different brain databases with varying difficulty: results and insights. *IEEE T Med Imaging*. 2014;33(10):2039–65.
40. Hamilton A, Singh A, Friedland B, et al. The impact of cone beam computer tomography field of view on the precision of digital intra-oral scan registration for static computer-assisted implant surgery: a CBCT analysis. *Clin Oral Implan Res*. 2022;33(12):1273–81.
41. 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 132. *Med Phys*. 2017;44(7):E43–76.
42. Zhao Y, Chen XR, Mcdonald B, et al. A transformer-based hierarchical registration framework for multimodality deformable image registration. *Comput Med Imag Grap*. 2023;108:102286.
43. Hirotaki K, Tomizawa K, Moriya S, et al. Fully automated volumetric modulated arc therapy planning for locally advanced rectal cancer: feasibility and efficiency. *Radiat Oncol*. 2023;18(1):147.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.