

Article

2D/3D Non-Rigid Image Registration via Two Orthogonal X-ray Projection Images for Lung Tumor Tracking

Guoya Dong^{1,2,3,†}, Jingjing Dai^{1,2,3,4,†}, Na Li⁵, Chulong Zhang⁴, Wenfeng He⁴, Lin Liu⁴, Yiping Chan⁴, Yunhui Li⁴, Yaoqin Xie⁴  and Xiaokun Liang^{4,*} 

¹ School of Health Sciences and Biomedical Engineering, Hebei University of Technology, Tianjin 300130, China

² Hebei Key Laboratory of Bioelectromagnetics and Neural Engineering, Tianjin 300130, China

³ Tianjin Key Laboratory of Bioelectromagnetic Technology and Intelligent Health, Tianjin 300130, China

⁴ Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China

⁵ Department of Biomedical Engineering, Guangdong Medical University, Dongguan 523808, China

* Correspondence: xk.liang@siat.ac.cn

† These authors contributed equally to this work.

Abstract: Two-dimensional (2D)/three-dimensional (3D) registration is critical in clinical applications. However, existing methods suffer from long alignment times and high doses. In this paper, a non-rigid 2D/3D registration method based on deep learning with orthogonal angle projections is proposed. The application can quickly achieve alignment using only two orthogonal angle projections. We tested the method with lungs (with and without tumors) and phantom data. The results show that the Dice and normalized cross-correlations are greater than 0.97 and 0.92, respectively, and the registration time is less than 1.2 seconds. In addition, the proposed model showed the ability to track lung tumors, highlighting the clinical potential of the proposed method.

Keywords: 2D/3D registration; orthogonal X-ray; deep learning



Citation: Dong, G.; Dai, J.; Li, N.; Zhang, C.; He, W.; Liu, L.; Chan, Y.; Li, Y.; Xie, Y.; Liang, X. 2D/3D Non-Rigid Image Registration via Two Orthogonal X-ray Projection Images for Lung Tumor Tracking. *Bioengineering* **2023**, *10*, 144. <https://doi.org/10.3390/bioengineering10020144>

Academic Editors: Paolo Zaffino and Maria Francesca Spadea

Received: 26 December 2022

Revised: 10 January 2023

Accepted: 16 January 2023

Published: 21 January 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Medical imaging has helped a lot with diagnosing and treating diseases as modern medical technology has grown quickly. Image registration is crucial in medical image processing because it helps predict, diagnose, and treat diseases. For the images to be registered, three-dimensional (3D) medical images with rich anatomical and structural information are an inevitable choice for clinical problems. Unfortunately, 3D images have a higher radiation dose and a slower imaging speed, which inconveniences real-time clinical problems, such as image-guided radiotherapy and interventional surgery. On the other hand, two-dimensional (2D) images lack some spatial structure information, while the imaging speed is very fast. Therefore, in recent years, 2D/3D image registration with faster speed and simple imaging equipment has attracted much attention. The types of 2D images are usually X-ray [1–4], fluoroscopic [5], digital subtraction angiography (DSA) [6,7], or ultrasound [8], whereas 3D images are chosen from computed tomography (CT) [1–4] or magnetic resonance imaging (MRI) [8].

2D/3D registration methods can be divided into traditional and deep learning-based image registration. In traditional image registration, 2D/3D alignment usually translates into the problem of solving for the maximum similarity between digitally reconstructed radiographs (DRR) and X-ray images. Similarity metrics are usually based on intensity-based mutual information [9–11], normalized cross-correlation (NCC) [12] and Pearson correlation coefficients [13], or gradient-based similarity metrics [14]. To minimize the dimensionality of the transformation parameters, regression models that rely on a priori information are usually built using B-spline [15] or principal component analysis (PCA) [16–19]. However, organ motion and deformation can cause errors in regression models, which rely too much on prior information. By incorporating finite element information into the

regression model, Zhang et al. [18,19] obtained more realistic and effective deformation parameters. However, adding finite element information makes the model-driven method of finding the optimal solution iteratively more inefficient. Therefore, this process is a constraint for developing real-time 2D/3D registration and tumor tracking algorithms.

With the development of artificial intelligence and deep learning, learning-based methods replace the tedious iterative optimization process with predicted values in the testing process, greatly improving computing efficiency. Zhang [20] proposed an unsupervised 2D-3D deformable registration network that addresses 2D/3D registration based on finite angles. Li et al. [4] proposed an unsupervised multiscale encode decode framework to achieve non-rigid 2D/3D registration based on a single 2D lateral brain image and 3D CBCT image. Ketcha et al. [21] used multi-stage rigid registration based on convolutional neural networks (CNN) to obtain a deformable spine model. Finally, Zhang et al. [22] achieved a deformable registration of the skull surface. Unfortunately, the above learning-based approach evaluates the similarity between DRR and X-ray, a 2D/3D registration reduced dimension to 2D/2D registration. Therefore, it is inevitable that spatial information will be lost to some extent. In addition, even with Graphic Processing Unit (GPU) support, forward projection, backward projection, and DRR generation involved in the above methods are computationally expensive. Then, the researchers completed end-to-end 2D/3D registration by integrating the forward/inverse projection spatial transformation layer into a neural network [3,23]. Frysch et al. [2] used Grangeat's relation instead of expensive forward/inverse projection to complete the 2D/3D registration method based on a single projection of arbitrary angle, which greatly accelerated the computational speed. However, this is a rigid transformation which is difficult to apply to elastic organs. Likewise, deep learning researchers have attempted to use statistical deformation models to build deep learning-based regression models. Using a priori information to build patient-specific deformation spaces, convolutional neural networks are used to accomplish regression on PCA coefficients [1,24,25] or B-spline parameter coefficients [26,27] to achieve patient-specific registration networks. Tian et al. [28] obtained the predicted deformation field based on the regression coefficients. However, this deformation space, which is completely based on a priori information, may lead to mistakes in the clinical application stage. In addition, some researchers [29,30] also accomplished 2D/3D image registration by extracting feature points. With the maturity of point cloud technology, many researchers have also built point-to-plane alignment models by extracting global point clouds to complete 2D/3D alignment models, but the anomaly removal for 2D/3D alignment models presents a challenge [31–33]. Graphical neural networks are also used for 2D/3D registration in low-contrast conditions [34]. Shao et al. [35] tracked liver tumors by adding finite element modeling. Still, the introduction of finite elements also brought some trouble to the registration time.

Therefore, we developed a deep learning-based method for non-rigid 2D/3D image registration of the same subject. Compared with traditional algorithms based on iterative optimization, this approach significantly improves the registration speed. Compared with the downscaled optimization of DRR and X-ray similarity, we optimized the similarity of 3D/3D images, which can effectively moderate the loss of spatial information. Additionally, only two projections based on orthogonal angles were chosen for 2D images to reduce the irradiation dose further. The proposed method is used to study the process of changes in the elastic organ as respiratory motion proceeds. More significantly, we also investigated the change in tumor position with respiratory motion, which can be used to achieve tracking of tumors based on orthogonal angular projections during radiotherapy.

The contributions of our work are summarized as follows:

1. We propose a 2D/3D elastic alignment framework based on deep learning, which can be applied to achieve organ shape tracking at lower doses using only two orthogonal angles of X-rays.
2. Our framework is expected to be used for tumor tracking with tumor localization accuracy up to 0.97 and registration time within 1.2 s, which may be a potential solution for image-guided surgery and radiotherapy.

The organizational structure of this article is as follows. Section 2 describes the experimental method. Section 3 describes the experiment setups. Section 4 shows the Result. Section 5 is the discussion and Section 6 concludes the paper and the references.

2. Methods

2.1. Overview of the Proposed Method

The framework of this method is shown in Figure 1. We design a non-rigid 2D/3D registration framework based on deep learning of orthogonal angle projection. Since it is a deep learning-based model, a large amount of data is needed to participate in training. The real paired 2D/3D medical images at the same time are very scarce, so the first task that needs to be done is data augmentation. We chose 4D CT of the lungs as the experimental subject. The expiratory end was used as a moving image M_{CT} and hybrid data augmentation [36,37] was used to obtain a large number of CT F_{CT} representing each respiratory phase of the lung (this procedure will be described in Section 3.1). Then, the ray casting method obtains a pair of 2D DRRs of F_{CT} with orthogonal angles. After that, the orthogonal DRR and the moving image M_{CT} are input into the 2D/3D registration network. The network outputs a 3D deformation field ϕ_p . Then, the moving image M_{CT} is transformed by the spatial transformation layer [38] to obtain the corresponding predicted CT image. The maximum similarity between the predicted CT image and the ground truth F_{CT} is calculated. Through continuous iterative optimization, we can complete the model training. In the inference phase, only the X-ray projections or DRRs and the moving image must be input to the trained network to get corresponding 3D images.

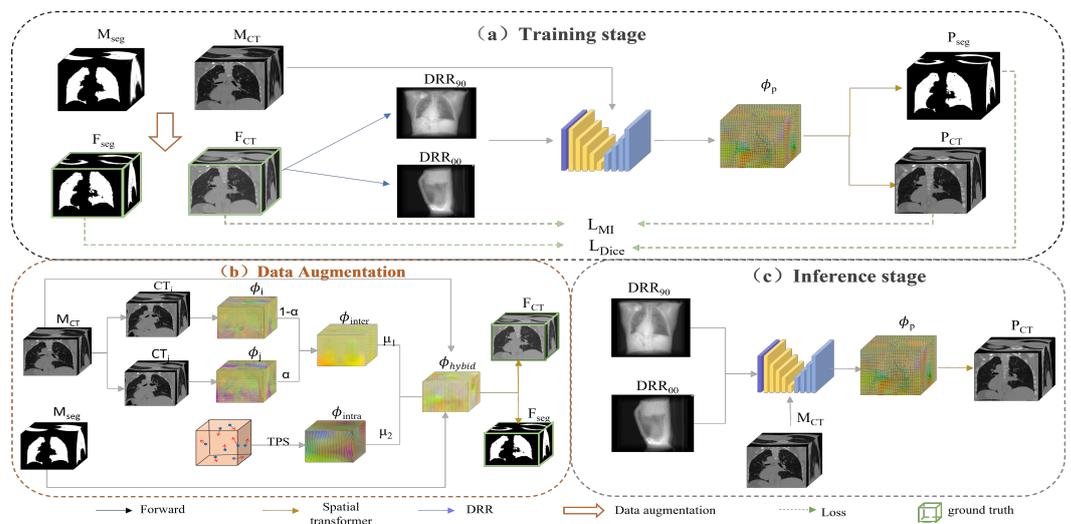


Figure 1. Overview of the proposed method. (a) Flowchart of the training phase of the method. First, a large number of CT F_{CT} and segmentation F_{seg} representing each phase are obtained by performing hybrid data augmentation of the moving image M_{CT} and the corresponding segmentation image M_{seg} . Then, the F_{CT} images are projected to obtain the 2D DRR_{90} and DRR_{00} . After that, they are fed into the registration network with the moving image M_{CT} to obtain the predicted deformation field ϕ_p . Finally, the moving image M_{CT} and the moving segmentation map M_{seg} are transformed to obtain the corresponding predicted images, P_{CT} and P_{seg} . (b) The process of hybrid data augmentation. The deformation field ϕ_{inter} is first obtained by inter-phase registration using traditional image registration. The small deformation ϕ_{intra} is simulated by TPS interpolation. The hybrid deformation field ϕ_{hybrid} is obtained by summing with random weights for data augmentation. (c) Inference stage. The 2D projection and moving images are directly input to the trained network to get the prediction ϕ_p , and then the registration can be completed by transformation.

2.2. 2D/3D Registration Network

Figure 2 shows the registration network. For 2D/3D image registration, the first thing to consider is the consistency of spatial dimension. As a result, we use the extracted feature up-dimensional approach to transform the 2D/3D registration problem into the 3D/3D registration problem. We used the residual network to get the 2D features. The most important step is identity mapping, stopping the gradient from going away, and helping train the network. Thus, when two DRRs with orthogonal angles are input to the network, they are first concatenated in the channel layer as the input of the residual network and then passed through the convolution layer, the max pooling layer, and two output channels with 64 and 128 residual blocks in turn. The channel layer is the third dimension to form a 3D feature map, which is input to the feature extraction network together with the moving image.

We selected the 3D Attention-U-net [37,39] (3D Attu) as the feature extraction network in this study. It can be called the 3D/3D matching network. The network 3D Attu adds an attention gate mechanism to the original U-net, which can automatically distinguish the target shape and scale, and learn more useful information. It also employs encoding and decoding mechanisms and skips connection mechanisms. It effectively blends high- and low-level semantic information while widening the perceptual domain. It has been used in many medical image processing tasks with excellent results. As a result, in this model, we feed the moving image M_{CT} and the 3D feature map into the 3D Attu. The output is the predicted deformation field.

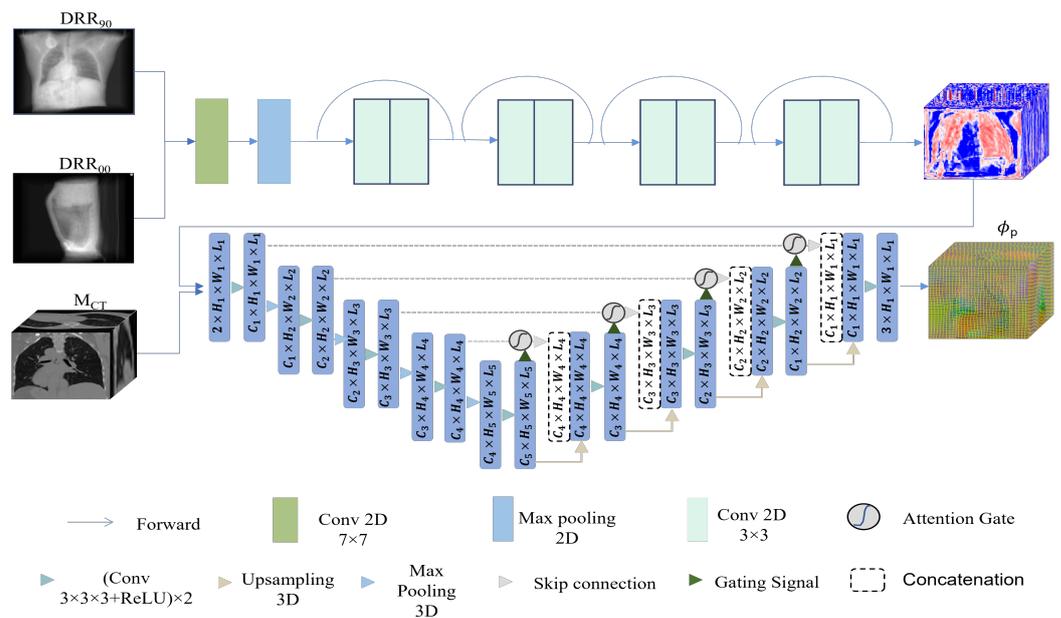


Figure 2. 2D/3D registration network. First, 2D DRRs at orthogonal angles are processed by residual blocks to obtain 3D feature maps. Then, the feature maps and moving images are fed into a 3D Attu-based encode–decode network. The final output of this network is the predicted 3D deformation field.

2.3. Loss Function

The mutual information (MI) between the ground truth F_{CT} and the predicted 3D CT P_{CT} obtained by the registration network constitute the loss function $L_{MI}(F_{CT}, P_{CT})$. The other part of the loss function is $L_{Dice}(F_{seg}, P_{seg})$, obtained by computing the Dice between the

corresponding segmented images, which allows the model to focus more on the lung region. Lastly is the regularized smoothing constraint $L_{Reg}(\phi_p)$ for the deformation field.

$$L_{Dice}(F_{seg}, P_{seg}) = \sum_{i=0}^n \frac{1}{n} \frac{2 |F_{seg}^i \cap P_{seg}^i|}{|F_{seg}^i + P_{seg}^i|} \quad (1)$$

$$L = \lambda_1 L_{Dice}(F_{seg}, P_{seg}) + \lambda_2 L_{MI}(F_{CT}, P_{CT}) + \lambda_3 L_{Reg}(\phi_p) \quad (2)$$

where n denotes the number of categories in the image, i denotes the i -th category of the image. ϕ denotes all elements in the entire deformation field. $\lambda_1, \lambda_2, \lambda_3$ denote the weights of $L_{Dice}, L_{MI}, L_{Reg}$ respectively, which were chosen as 0.5, 0.5, and 0.1 in this experiment.

3. Experiment Setups

3.1. Data and Augmentation

We conducted experiments on three different types of lung data, TCIA [40–43] patient with a tumor, Dirlab [44] lung CT without tumor, and CIRS phantom. ITK-SNAP is used for automatic segmentation to obtain labels. In the TCIA patient data, we selected one of the patients for the experiment. In Dirlab, we selected the first five sets of data for the experiment. In the CIRS phantom, we simulated the lung tumor with a water sphere. In the experiment, we resampled the 3D CT image to $128 * 128 * 128$ with a voxel spacing of $1 \text{ mm} * 1 \text{ mm} * 1 \text{ mm}$. Since our experiment is a 2D/3D registration, paired 2D projections and 3D medical images of the same moment are rare. It is unethical to expose the human body to additional radiation doses, so the first task is data augmentation. However, for 2D/3D registration of the treatment phase (e.g., radiotherapy, surgical navigation), it is obvious that the focus is more on the specific person. Therefore, we chose a hybrid data augmentation approach to train a deep learning-based 2D/3D registration model for a specific human body.

In the hybrid data augmentation shown in Figure 1b, we first selected the end-expiratory phase of 4D CT as the moving image M_{CT} and the remaining phases as the fixed image $CT_{1,\dots,i,j,9}$. Then, we used the conventional intensity-based image registration method to obtain nine deformation fields in the order of $\phi_{1,\dots,i,j,9}$. The deformation fields used for data augmentation were arbitrarily selected from two of the nine deformation fields and superimposed with random weights to obtain many inter-phase deformations. The lung may also change during respiratory motion. Therefore, we use thin plate spline (TPS) interpolation to simulate small changes in specific phases. The number of control points N was randomly chosen between 20 and 60. The movement distance of control points was chosen between 0 mm and 20 mm to obtain many phase-specific random deformations. To obtain more morphologically diverse images, we combined inter-phase and intra-phase specific deformation with random weights to obtain many hybrid deformation fields. Spatial warping of the moving and segmented images was performed to obtain CT and segmented images representing each respiratory phase of the lung.

3.2. DRR Image Generation

The orthogonal angle X-ray projection system in this experiment is shown in Figure 3. Two-point light sources at orthogonal angles emit rays through the object and project them on two detectors perpendicular to the central axis. We assume that the initial intensity of I_0 at the light source, μ is the internal attenuation coefficient of the object to the rays, l is the thickness of the ray through the object, and I_n is the intensity of the ray after passing through the object. The formula $I_n = I_0 e^{-\int \mu(l) dl}$ arises. After the projection of one ray is finished, the attenuation coefficient obtained by accumulating the whole path and then converting it to CT value is the X-ray image. In this experiment, like most researchers, DRR images with the same imaging principle are used instead of X-ray. Virtual X-rays were used to pass through the CT images, and after attenuation, they were projected onto

the imaging plane to reconstruct the DRR images. The 3D CT images representing each respiratory phase after data augmentation are projected using this method to obtain the DRR images at the corresponding moment. This technique has been widely used for 2D/3D registration methods.

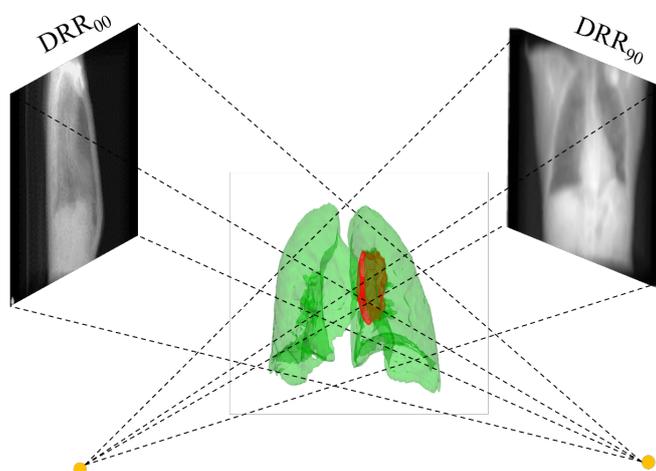


Figure 3. Schematic diagram of DRR image generation.

3.3. Experiment Detail

We used hybrid data augmentation to obtain 6000 samples from the three types of experimental data. Of these, 5400 were used as the training set, 300 as the validation set, and 300 as the test set. Our experiment was implemented using the deep learning framework Pytorch 1.10 on a NVIDIA A6000 GPU with 48 G of memory, and an AMD Ryzen 7 3700X 8-core processor with 128 GB of internal memory. The learning rate is set to 10^{-4} . For all datasets, the batch size was set to 8 and the optimization algorithm is Adam.

3.4. Experiment Evaluation

In order to verify that our model can achieve 2D/3D registration by two orthogonal angular projections, we selected the end of expiration as the moving image and aligned it toward the remaining phases. We evaluated the three-lung data using NCC, MI, 95% Hausdorff surface distance, and Dice. In addition, to explore the tracking of lung tumors that can be achieved by our model, we compared between predicted and ground truth values for the dataset with tumors and quantitatively evaluated using Dice and the tumor center of mass.

4. Result

4.1. Registration from the Expiratory End to Each Phase

Here, we demonstrate the registration results of each phase from the end of expiration to the end of inspiration for the TCIA, Dirlab, and phantom. For the qualitative assessment, Figure 4a shows the results of our selected experiments on patients with tumors on TCIA, Figure 4b shows a randomly selected set of experiments from Dirlab, and Figure 4c shows the effect of registration of the phantom data. The odd rows are the unaligned ones, and the even rows are the aligned results. Based on the results, both TCIA patients with tumors and without tumors in Dirlab, as well as the phantom model with water balloons that simulate tumors, can achieve registration from the end of expiration to the rest of the stages.

For the quantitative analysis, we used Dice of the segmentation map, 95% Hausdorff surface distance, NCC, and MI of grayscale images to evaluate our scheme separately. The results are shown in Table 1. It can be seen that good registration results are obtained for all three types of data, not only on the grayscale images, but also on the lung of interest. The Dice values of all three data types are above 0.97, the Hausdorff surface distances are below 2 mm, NCC are above 0.92 and MI are above 0.90. Compared with the real

human lung, the NCC and MI of the phantom data are relatively small because the lung of the phantom itself does not change. Only the internal water sphere changes, which is more rigidly transformed relative to the real patient, so the NCC and MI are relatively small at higher Dice. However, the total accuracies are still above 0.92 and 0.90. Therefore, quantitative and qualitative results show that the proposed method can achieve non-rigid 2D/3D registration for a specific subject by two orthogonal angular projections.

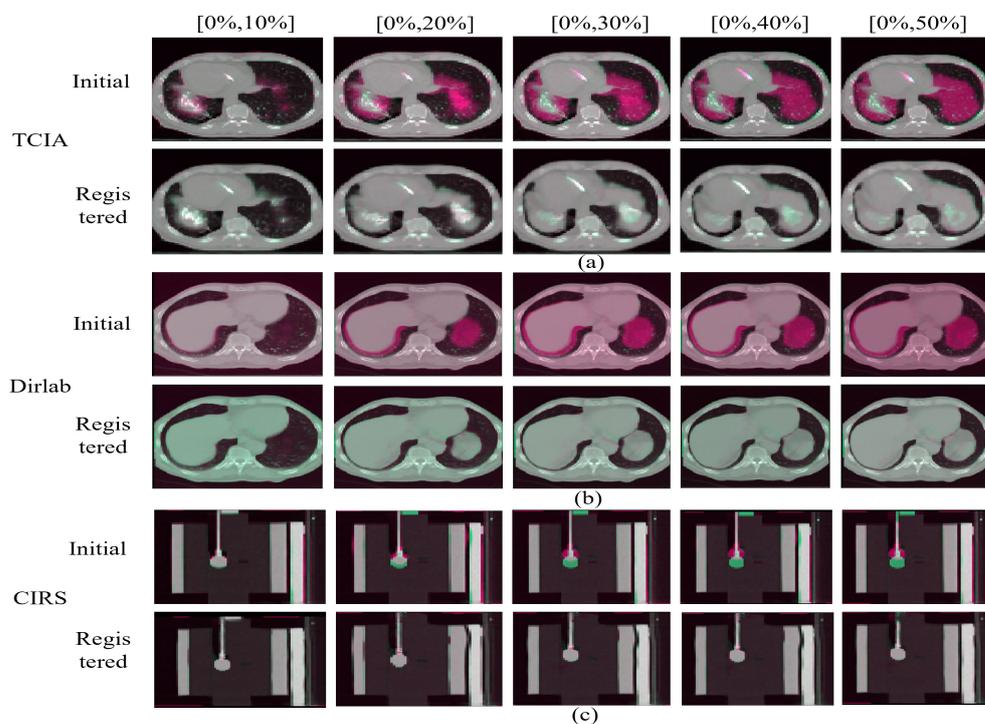


Figure 4. Registration from the exhalation end to the other stages. (a) shows the results of our registration on TCIA, (b) a randomly selected set of experiments from Dirlab, and (c) the registration results of the phantom data. The odd-numbered rows are the unregistered contrast images, and the even-numbered rows are the registered contrast images.

Table 1. The accuracy of subjects’ registration from the expiratory end to each phase.

		Dice	Hauf (95%)	NCC	MI
TCIA	[0%,10%]	0.9814	1.1210	0.9846	0.9796
	[0%,20%]	0.9791	1.3811	0.9846	0.9760
	[0%,30%]	0.9795	1.3811	0.9847	0.9540
	[0%,40%]	0.9785	1.3811	0.9846	0.9578
	[0%,50%]	0.9806	1.4196	0.9848	0.9650
Dirlab	[0%,10%]	0.9857	1.8839	0.9762	0.9590
	[0%,20%]	0.9857	1.8620	0.9723	0.9535
	[0%,30%]	0.9853	1.8280	0.9691	0.9511
	[0%,40%]	0.9853	1.8290	0.9680	0.9424
	[0%,50%]	0.9854	1.8290	0.9753	0.9608
CIRS	[0%,10%]	0.9862	0.9043	0.9338	0.9135
	[0%,20%]	0.9907	1.6713	0.9291	0.9023
	[0%,30%]	0.9888	2.0000	0.9360	0.9064
	[0%,40%]	0.9885	1.9087	0.9348	0.9065
	[0%,50%]	0.9894	1.9087	0.9349	0.9178

4.2. Tumor Location

Both TCIA patient and the phantom contained tumors. The accuracy of tumor localization was evaluated qualitatively and quantitatively.

Figure 5 shows the qualitative evaluation of the 3D tumor with two types of data, where (a) is a 3D visualization image of the patient’s overall lung and tumor and (b) is of the phantom data. Table 2 presents the quantitative results, where we evaluated the tumor center mass and Dice. The tumor center of mass deviation is within 0.15 mm for the real patient. The phantom tumor center of mass is less than 0.05 mm. The Dice of both are above 0.88. It can be seen that the proposed method can achieve registration both for the whole lung and for the tumor. In addition, the fact that local tumors are well aligned suggests that our model could be useful for clinical applications such as tracking tumors.

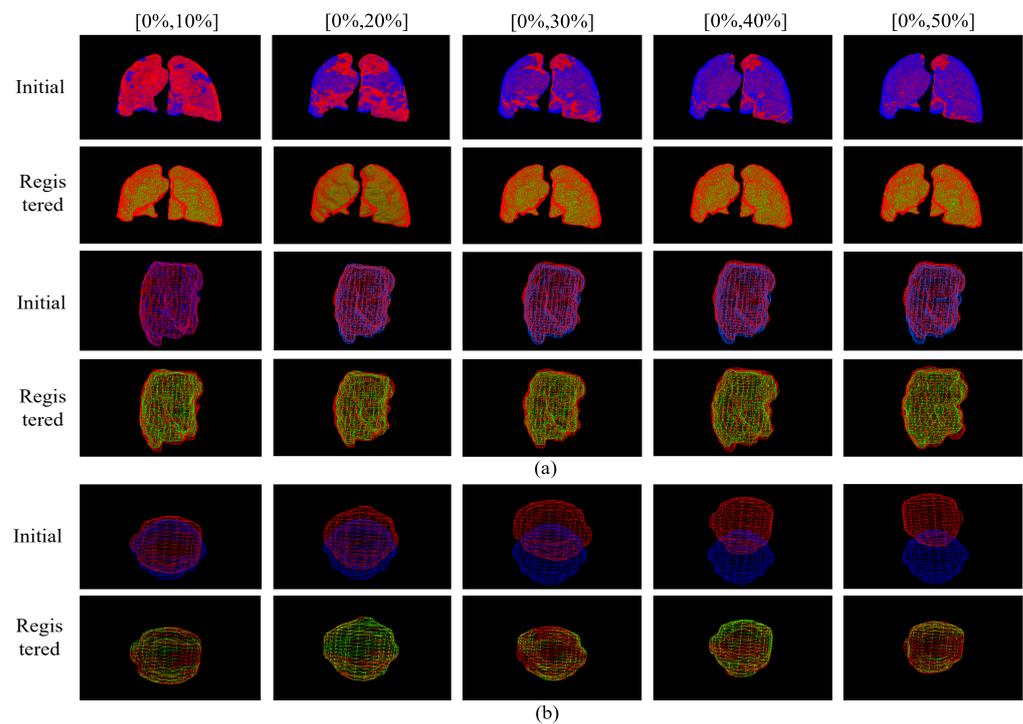


Figure 5. Tumor registration results from the exhalation end to other stages. Where (a) is the 3D presentation of the results before and after a real patient’s lung and tumor registration, and (b) is of the phantom data of the 3D results of the tumor display. The odd rows are the unregistered images, and the even rows are the post-registered images. The red image indicates the ground truth. Blue is the moving image, and green is the predicted result obtained by the model.

Table 2. The accuracy of tumor location from the expiratory end to each phase.

		X(LR)	Center Mass (mm)		Center	Dice Tomor
			Y(AP)	Z(LR)		
TCIA	[0%,10%]	0.0003	0.0473	0.0844	0.0968	0.9440
	[0%,20%]	0.0117	0.0133	0.0260	0.0315	0.9434
	[0%,30%]	0.0031	0.0023	0.0473	0.0022	0.9023
	[0%,40%]	0.0032	0.0078	0.0339	0.0350	0.9080
	[0%,50%]	0.0227	0.0510	0.1251	0.1370	0.8984
CIRS	[0%,10%]	0.0224	0.0011	0.0270	0.0351	0.9717
	[0%,20%]	0.0061	0.0025	0.0081	0.0104	0.9764
	[0%,30%]	0.0086	0.0018	0.0158	0.0181	0.9609
	[0%,40%]	0.0174	0.0177	0.0084	0.0262	0.9702
	[0%,50%]	0.0022	0.0043	0.0477	0.0479	0.8826

5. Discussion

5.1. Traditional Registration in Data Augmentation

We used traditional intensity-based image registration for data augmentation to complete the registration between phases. Here we present the two most distorted parts of the three data, the end of expiration and the end of inspiration, for evaluation. The experimental results are shown in Figure 6.

Figure 6 shows the results from three directions before and after the registration. The odd columns are the unregistered images. The even columns are the results after the traditional registration method. Conventional image registration can be accomplished for real patients and models from exhalation to the end of inspiration, ensuring that our augmentation data encompasses all respiratory phases of the lung. In addition, the registration covers the larger deformations at both ends.

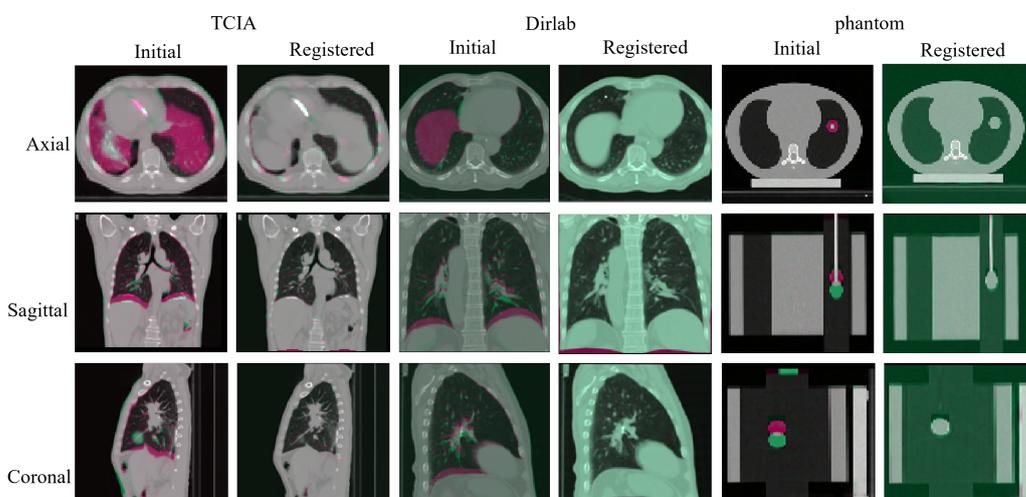


Figure 6. The traditional registration method results from the end of exhalation to the end of inhalation. The first two columns are the registration results on the patient, the middle two are the registration results on the normal human lung, and the last two are the registration results on the phantom. The odd columns are the unregistered images, and the even columns are the results of the registered images.

5.2. Landmark Error

For the Dirlab data, the landmark points and the deformation field for performing data augmentation are known. Thus, the landmark points of the image generated after data augmentation are also known and used as our ground truth. The mean target registration error (mTRE) is evaluated with our model-predicted images.

The data obtained from the evaluation are shown in Table 3. The corresponding box plot is shown in Figure 7, in which green indicates the data before registration, yellow is the data after registration of the proposed method, and purple represents the data after 3D/3D registration using Demons [45]. The results show that the proposed model can achieve effective 2D/3D registration, but the accuracy is lower than the existing advanced 3D/3D registration models because the experimental data are only two 2D X-rays with orthogonal angles. Although the proposed method transforms 2D/3D registration into a 3D/3D registration problem, some image details are indeed lost compared with the 3D images, resulting in the loss of information on tiny details, such as capillaries, leading to a lower accuracy of landmark error based on detailed information. However, in the 2D/3D registration mission, more attention is paid to the global overall changes in the lung and the tumor location. The proposed method greatly reduces the irradiation dose and improves the registration speed.

Table 3. Mean target registration error of landmarks in Dirlabs.

(mm)	Initial	Proposed (2D/3D)	Demons (3D/3D)
Dirlab1	3.9776 (1.8616)	2.0065 (0.6748)	1.6297 (0.3196)
Dirlab2	6.3989 (2.1719)	4.0079 (1.0077)	3.3807 (0.5089)
Dirlab3	6.2138 (1.7843)	2.9219 (0.7237)	2.1556 (0.2991)
Dirlab4	7.6437 (2.3978)	4.0682 (0.9898)	3.2525 (0.3918)
Dirlab5	6.6075 (2.1448)	2.6253 (0.7719)	1.6408 (0.4824)

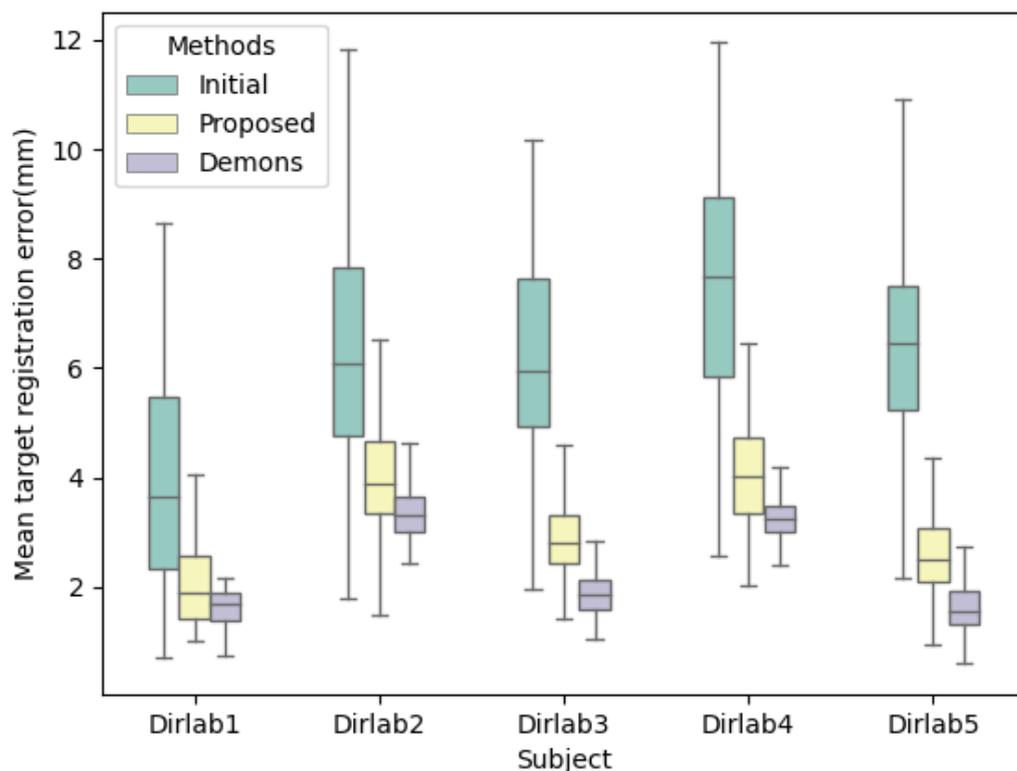


Figure 7. Box plot of landmark points in Dirlab; green indicates the data before registration, yellow is the data after registration of the proposed method, and purple represents the data after 3D/3D registration using Demons.

In addition, our method can complete 2D/3D registration in 1.2 s. In contrast, other data-driven 2D/3D registration models, such as [20], may take a few seconds. On the other hand, traditional image registration methods may take tens of minutes or even hours. Our method also only needs two different angles of X-rays, which greatly reduces the amount of radiation and makes the hardware in the clinic easier to use. Of course, our method also has some limitations. First of all, since real medical images do not exist at the same moment of paired orthogonal angles of X-ray and corresponding 3D CT, we use 2D DRR. Although DRR and real X-ray use the same imaging way, it is undeniable that there are some grayscale and noise differences between the two. However, it can be corrected by using existing methods, such as histogram matching [24], network of GAN [25], etc., which is not the main focus of our study. We will also make some improvements to the program to speed up the processing speed for radiotherapy or interventional procedures that require more real-time, etc. In addition, since there are few non-rigid 2D/3D registration articles, the code is not open source. We have yet to choose a suitable comparison experiment, and we will continue to look for it in the future.

6. Conclusions

This study proposes a deep learning-based 2D/3D registration method using two orthogonal angular X-ray projection images. The proposed algorithm has been verified on lung data with and without tumor and phantom data, and obtained high registration accuracy, where Dice and NCC are greater than 0.97 and 0.92. In addition, we evaluated the accuracy on the data containing tumor, and the tumor center-of-mass error was within 0.15 mm, which indicates the promising use of our model for tumor tracking. The registration time is within 1.2 s, and this is promising for clinical applications, such as radiotherapy or surgical navigation, to track the shape of organs in real time. Moreover, we only need to use two orthogonal angles of X-rays to achieve 2D/3D deformable image registration, which can greatly reduce the extra dose during treatment and simplify the hardware system required.

Author Contributions: Conceptualization, G.D., J.D., Y.X. and X.L.; methodology, J.D. and X.L.; software, J.D.; validation, N.L.; investigation, N.L., H.W. and L.X.; data curation, G.D. and J.D.; writing—original draft preparation, G.D. and J.D.; writing—review and editing, C.Z., W.H., L.L., Y.C., L.L. Y.L. and X.L.; visualization, L.L. and X.L.; supervision, X.L. All authors have read and agreed to the published version of the manuscript.

Funding: This work is partly supported by grants from the National Natural Science Foundation of China (U20A201795, U21A20480, 82202954, 61871374, 62001464, 11905286), Young S&T Talent Training Program of Guangdong Provincial Association for S&T, China (SKXRC202224), and the Chinese Academy of Sciences Special Research Assistant Grant Program.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The dataset images used for this study is publicly available on TCIA at <https://wiki.cancerimagingarchive.net/pages/viewpage.action?pageId=21267414> (accessed on 27 October 2022) and Dirlab at <https://med.emory.edu/departments/radiation-oncology/research-laboratories/deformable-image-registration/downloads-and-reference-data/4dct.html> (accessed on 27 October 2022).

Acknowledgments: We thank The Cancer Imaging Archive (TCIA) and Emory University School of Medicine (4D CT) for public access and for sharing their datasets.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

2D	Two-dimensional
3D	Three-dimensional
DSA	Digital Subtraction Angiography
CT	Computed Tomography
MRI	Magnetic Resonance Imaging
DRR	Digitally Reconstructed Radiographs
NCC	Normalized cross-correlation
PCA	Principal Component Analysis
CNN	Convolutional Neural Network
GPU	Graphic Processing Unit(GPU)
3D Attu	3D Attention-U-net
MI	Mutual Information
TPS	Thin plate spline
mTRE	Mean Target Registration Error

References

1. Foote, M.D.; Zimmerman, B.E.; Sawant, A.; Joshi, S.C. Real-time 2D-3D deformable registration with deep learning and application to lung radiotherapy targeting. In Proceedings of the International Conference on Information Processing in Medical Imaging, Hong Kong, 2–7 June 2019; Springer: Berlin/Heidelberg, Germany, 2019; pp. 265–276.
2. Frysck, R.; Pfeiffer, T.; Rose, G. A novel approach to 2D/3D registration of X-ray images using Grangeat's relation. *Med. Image Anal.* **2021**, *67*, 101815. [[CrossRef](#)]
3. Van Houtte, J.; Audenaert, E.; Zheng, G.; Sijbers, J. Deep learning-based 2D/3D registration of an atlas to biplanar X-ray images. *Int. J. Comput. Assist. Radiol. Surg.* **2022**, *17*, 1333–1342. [[CrossRef](#)]
4. Li, P.; Pei, Y.; Guo, Y.; Ma, G.; Xu, T.; Zha, H. Non-rigid 2D-3D registration using convolutional autoencoders. In Proceedings of the 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI), Iowa City, IA, USA, 3–7 April 2020; IEEE: New York, NY, USA, 2020; pp. 700–704.
5. Wang, C.; Xie, S.; Li, K.; Wang, C.; Liu, X.; Zhao, L.; Tsai, T.Y. Multi-View Point-Based Registration for Native Knee Kinematics Measurement with Feature Transfer Learning. *Engineering* **2021**, *7*, 881–888. [[CrossRef](#)]
6. Guan, S.; Wang, T.; Sun, K.; Meng, C. Transfer learning for nonrigid 2d/3d cardiovascular images registration. *IEEE J. Biomed. Health Inform.* **2020**, *25*, 3300–3309. [[CrossRef](#)]
7. Miao, S.; Wang, Z.J.; Liao, R. A CNN regression approach for real-time 2D/3D registration. *IEEE Trans. Med. Imaging* **2016**, *35*, 1352–1363. [[CrossRef](#)]
8. Markova, V.; Ronchetti, M.; Wein, W.; Zettinig, O.; Prevost, R. Global Multi-modal 2D/3D Registration via Local Descriptors Learning. *arXiv* **2022**, arXiv:2205.03439.
9. Zheng, G. Effective incorporating spatial information in a mutual information based 3D–2D registration of a CT volume to X-ray images. *Comput. Med. Imaging Graph.* **2010**, *34*, 553–562. [[CrossRef](#)]
10. Zollei, L.; Grimson, E.; Norbash, A.; Wells, W. 2D-3D rigid registration of X-ray fluoroscopy and CT images using mutual information and sparsely sampled histogram estimators. In Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2001, Kauai, HI, USA, 8–14 December 2001; IEEE: New York, NY, USA, 2001; Volume 2, p. II.
11. Gendrin, C.; Furtado, H.; Weber, C.; Bloch, C.; Figl, M.; Pawiro, S.A.; Bergmann, H.; Stock, M.; Fichtinger, G.; Georg, D.; et al. Monitoring tumor motion by real time 2D/3D registration during radiotherapy. *Radiother. Oncol.* **2012**, *102*, 274–280. [[CrossRef](#)]
12. Gao, C.; Grupp, R.B.; Unberath, M.; Taylor, R.H.; Armand, M. Fiducial-free 2D/3D registration of the proximal femur for robot-assisted femoroplasty. *IEEE Trans. Med. Robot. Bionics* **2020**, *11315*, 350–355.
13. Munbodh, R.; Knisely, J.P.; Jaffray, D.A.; Moseley, D.J. 2D–3D registration for cranial radiation therapy using a 3D kV CBCT and a single limited field-of-view 2D kV radiograph. *Med. Phys.* **2018**, *45*, 1794–1810. [[CrossRef](#)]
14. De Silva, T.; Uneri, A.; Ketcha, M.; Reaungamornrat, S.; Kleinszig, G.; Vogt, S.; Aygun, N.; Lo, S.; Wolinsky, J.; Siewerdsen, J. 3D–2D image registration for target localization in spine surgery: Investigation of similarity metrics providing robustness to content mismatch. *Phys. Med. Biol.* **2016**, *61*, 3009. [[CrossRef](#)]
15. Yu, W.; Tannast, M.; Zheng, G. Non-rigid free-form 2D–3D registration using a B-spline-based statistical deformation model. *Pattern Recognit.* **2017**, *63*, 689–699. [[CrossRef](#)]
16. Li, R.; Jia, X.; Lewis, J.H.; Gu, X.; Folkerts, M.; Men, C.; Jiang, S.B. Real-time volumetric image reconstruction and 3D tumor localization based on a single x-ray projection image for lung cancer radiotherapy. *Med. Phys.* **2010**, *37*, 2822–2826. [[CrossRef](#)]
17. Li, R.; Lewis, J.H.; Jia, X.; Gu, X.; Folkerts, M.; Men, C.; Song, W.Y.; Jiang, S.B. 3D tumor localization through real-time volumetric x-ray imaging for lung cancer radiotherapy. *Med. Phys.* **2011**, *38*, 2783–2794. [[CrossRef](#)]
18. Zhang, Y.; Tehrani, J.N.; Wang, J. A biomechanical modeling guided CBCT estimation technique. *IEEE Trans. Med. Imaging* **2016**, *36*, 641–652. [[CrossRef](#)]
19. Zhang, Y.; Folkert, M.R.; Li, B.; Huang, X.; Meyer, J.J.; Chiu, T.; Lee, P.; Tehrani, J.N.; Cai, J.; Parsons, D. 4D liver tumor localization using cone-beam projections and a biomechanical model. *Radiother. Oncol.* **2019**, *133*, 183–192. [[CrossRef](#)]
20. Zhang, Y. An unsupervised 2D–3D deformable registration network (2D3D-RegNet) for cone-beam CT estimation. *Phys. Med. Biol.* **2021**, *66*, 074001. [[CrossRef](#)]
21. Ketcha, M.; De Silva, T.; Uneri, A.; Jacobson, M.; Goerres, J.; Kleinszig, G.; Vogt, S.; Wolinsky, J.; Siewerdsen, J. Multi-stage 3D–2D registration for correction of anatomical deformation in image-guided spine surgery. *Phys. Med. Biol.* **2017**, *62*, 4604. [[CrossRef](#)]
22. Zhang, Y.; Qin, H.; Li, P.; Pei, Y.; Guo, Y.; Xu, T.; Zha, H. Deformable registration of lateral cephalogram and cone-beam computed tomography image. *Med. Phys.* **2021**, *48*, 6901–6915. [[CrossRef](#)]
23. Gao, C.; Liu, X.; Gu, W.; Killeen, B.; Armand, M.; Taylor, R.; Unberath, M. Generalizing spatial transformers to projective geometry with applications to 2D/3D registration. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Lima, Peru, 4–8 October 2020; Springer: Berlin/Heidelberg, Germany, 2020; pp. 329–339.
24. Wei, R.; Liu, B.; Zhou, F.; Bai, X.; Fu, D.; Liang, B.; Wu, Q. A patient-independent CT intensity matching method using conditional generative adversarial networks (cGAN) for single x-ray projection-based tumor localization. *Phys. Med. Biol.* **2020**, *65*, 145009. [[CrossRef](#)]
25. Wei, R.; Zhou, F.; Liu, B.; Bai, X.; Fu, D.; Liang, B.; Wu, Q. Real-time tumor localization with single X-ray projection at arbitrary gantry angles using a convolutional neural network (CNN). *Phys. Med. Biol.* **2020**, *65*, 065012. [[CrossRef](#)]

26. Van Houtte, J.; Gao, X.; Sijbers, J.; Zheng, G. 2D/3D registration with a statistical deformation model prior using deep learning. In Proceedings of the 2021 IEEE EMBS international conference on biomedical and health informatics (BHI), Athens, Greece, 27–30 July 2021; IEEE: New York, NY, USA, 2021, pp. 1–4.
27. Pei, Y.; Zhang, Y.; Qin, H.; Ma, G.; Guo, Y.; Xu, T.; Zha, H. Non-rigid craniofacial 2D-3D registration using CNN-based regression. In *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*; Springer: Berlin/Heidelberg, Germany, 2017; pp. 117–125.
28. Tian, L.; Lee, Y.Z.; San José Estépar, R.; Niethammer, M. LiftReg: Limited Angle 2D/3D Deformable Registration. In Proceedings of the Medical Image Computing and Computer Assisted Intervention—MICCAI 2022, Singapore, 18–22 September 2022; Wang, L., Dou, Q., Fletcher, P.T., Speidel, S., Li, S., Eds.; Springer Nature: Cham, Switzerland, 2022; pp. 207–216.
29. Liao, H.; Lin, W.A.; Zhang, J.; Zhang, J.; Luo, J.; Zhou, S.K. Multiview 2D/3D rigid registration via a point-of-interest network for tracking and triangulation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 12638–12647.
30. Markova, V.; Ronchetti, M.; Wein, W.; Zettinig, O.; Prevost, R. Global Multi-modal 2D/3D Registration via Local Descriptors Learning. In Proceedings of the Medical Image Computing and Computer Assisted Intervention—MICCAI 2022, Singapore, 18–22 September 2022; Wang, L., Dou, Q., Fletcher, P.T., Speidel, S., Li, S., Eds.; Springer Nature: Cham, Switzerland, 2022; pp. 269–279.
31. Schaffert, R.; Wang, J.; Fischer, P.; Borsdorf, A.; Maier, A. Learning an attention model for robust 2-D/3-D registration using point-to-plane correspondences. *IEEE Trans. Med. Imaging* **2020**, *39*, 3159–3174. [[CrossRef](#)]
32. Schaffert, R.; Wang, J.; Fischer, P.; Borsdorf, A.; Maier, A. Metric-driven learning of correspondence weighting for 2-D/3-D image registration. In Proceedings of the German Conference on Pattern Recognition, Stuttgart, Germany, 9–12 October 2019; Springer: Berlin/Heidelberg, Germany, 2019, pp. 140–152.
33. Schaffert, R.; Wang, J.; Fischer, P.; Maier, A.; Borsdorf, A. Robust multi-view 2-d/3-d registration using point-to-plane correspondence model. *IEEE Trans. Med. Imaging* **2019**, *39*, 161–174. [[CrossRef](#)]
34. Nakao, M.; Nakamura, M.; Matsuda, T. Image-to-Graph Convolutional Network for 2D/3D Deformable Model Registration of Low-Contrast Organs. *IEEE Trans. Med. Imaging* **2022**, *41*, 3747–3761. [[CrossRef](#)]
35. Shao, H.C.; Wang, J.; Bai, T.; Chun, J.; Park, J.C.; Jiang, S.; Zhang, Y. Real-time liver tumor localization via a single x-ray projection using deep graph neural network-assisted biomechanical modeling. *Phys. Med. Biol.* **2022**, *67*, 115009. [[CrossRef](#)]
36. Liang, X.; Zhao, W.; Hristov, D.H.; Buyyounouski, M.K.; Hancock, S.L.; Bagshaw, H.; Zhang, Q.; Xie, Y.; Xing, L. A deep learning framework for prostate localization in cone beam CT-guided radiotherapy. *Med. Phys.* **2020**, *47*, 4233–4240. [[CrossRef](#)]
37. Liang, X.; Li, N.; Zhang, Z.; Xiong, J.; Zhou, S.; Xie, Y. Incorporating the hybrid deformable model for improving the performance of abdominal CT segmentation via multi-scale feature fusion network. *Med. Image Anal.* **2021**, *73*, 102156. [[CrossRef](#)]
38. Jaderberg, M.; Simonyan, K.; Zisserman, A.; Kavukcoglu, K. Spatial transformer networks. *Adv. Neural Inf. Process. Syst.* **2015**, *28*, 2017–2025.
39. Schlemper, J.; Oktay, O.; Schaap, M.; Heinrich, M.; Kainz, B.; Glocker, B.; Rueckert, D. Attention gated networks: Learning to leverage salient regions in medical images. *Med. Image Anal.* **2019**, *53*, 197–207. [[CrossRef](#)]
40. Hugo, G.D.; Weiss, E.; Sleeman, W.C.; Balik, S.; Keall, P.J.; Lu, J.; Williamson, J.F. Data from 4d lung imaging of nscl patients. *Med. Phys.* **2017**, *44*, 762–771. [[CrossRef](#)]
41. Balik, S.; Weiss, E.; Jan, N.; Roman, N.; Sleeman, W.C.; Fatyga, M.; Christensen, G.E.; Zhang, C.; Murphy, M.J.; Lu, J.; et al. Evaluation of 4-dimensional computed tomography to 4-dimensional cone-beam computed tomography deformable image registration for lung cancer adaptive radiation therapy. *Int. J. Radiat. Oncol. Biol. Phys.* **2013**, *86*, 372–379. [[CrossRef](#)] [[PubMed](#)]
42. Clark, K.; Vendt, B.; Smith, K.; Freymann, J.; Kirby, J.; Koppel, P.; Moore, S.; Phillips, S.; Maffitt, D.; Pringle, M.; et al. The Cancer Imaging Archive (TCIA): Maintaining and operating a public information repository. *J. Digit. Imaging* **2013**, *26*, 1045–1057. [[CrossRef](#)]
43. Roman, N.O.; Shepherd, W.; Mukhopadhyay, N.; Hugo, G.D.; Weiss, E. Interfractional positional variability of fiducial markers and primary tumors in locally advanced non-small-cell lung cancer during audiovisual biofeedback radiotherapy. *Int. J. Radiat. Oncol. Biol. Phys.* **2012**, *83*, 1566–1572. [[CrossRef](#)]
44. Castillo, R.; Castillo, E.; Guerra, R.; Johnson, V.E.; McPhail, T.; Garg, A.K.; Guerrero, T. A framework for evaluation of deformable image registration spatial accuracy using large landmark point sets. *Phys. Med. Biol.* **2009**, *54*, 1849. [[CrossRef](#)] [[PubMed](#)]
45. Vercauteren, T.; Pennec, X.; Perchant, A.; Ayache, N. Diffeomorphic demons: Efficient non-parametric image registration. *NeuroImage* **2009**, *45*, S61–S72. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.