RESEARCH ARTICLE

An Unsupervised Learning Approach to 3D Rectal MRI Volume Registration

CHI-JUI HO[1], **SOAN T. M. DUONG**[2,3], **YIQIAN WANG**[1], **CHANH D. TR. NGUYEN**[2,4], **BIEU Q. BUI**[5], **STEVEN Q. H. TRUONG**[2], **TRUONG Q. NGUYEN**[1], **(Fellow, IEEE)**, **AND CHEOLHONG AN**[1]

[1]Department of Electrical and Computer Engineering, University of California at San Diego, La Jolla, CA 92093, USA
[2]Vinbrain JSC, Hanoi 11619, Vietnam
[3]Department of Computer Science, Le Quy Don Technical University, Hanoi 11917, Vietnam
[4]College of Engineering and Computer Science, VinUniversity, Hanoi 12406, Vietnam
[5]Department of Radiotherapy and Radiosurgery, 108 Military Central Hospital, Hanoi 11610, Vietnam

Corresponding author: Chi-Jui Ho (chh009@ucsd.edu)

ABSTRACT Accurate alignment of multi-session medical imaging is essential to the analysis of disease progression. By comparing the magnetic resonance imaging (MRI) data captured before and after a course of neoadjuvant chemoradiation (nCRT) treatment, physicians are able to evaluate the tumor response for further treatment of the disease. However, rectal MRI data captured in multi-session are often misaligned and not guaranteed to have one-to-one correspondence, making it challenging for physicians to observe the treatment response of tumor. To address this issue, we propose an unsupervised learning based volume registration framework, which enables accurate alignment even under a high degree of deformation between multi-session rectal data. Moreover, it works without the assumption of one-to-one correspondence between multi-session data, and hence is a general solution to rectal MRI volume registration. The experimental results show that the proposed registration framework accurately aligns rectal cancer images and outperforms other state-of-the-art methods in medical image registration. By providing accurate registration, it can potentially increase the efficiency and reduce the workload for physicians to evaluate the rectal tumor response to nCRT.

INDEX TERMS Magnetic resonance imaging, image registration, rectal cancer, deep learning, convolutional neural network.

I. INTRODUCTION

Rectal cancer is one of the most common and life-threatening cancers, especially for older people and people in developing countries. According to the statistics by World Health Organization (WHO), about 732,210 new cases and 339,022 deaths were caused by rectal cancer around the world in 2020 [47]. The prognosis of rectal cancer is directly related to tumor staging, and the treatment strategy is based on the location and extent of the disease. For instance, transanal endoscopic microsurgery and the extra levator abdominoperineal resection are performed for patients whose rectal tumors are in early and late stages, respectively [34]. According

to the National Comprehensive Cancer Network (NCCN) guidelines, neoadjuvant chemoradiotherapy (nCRT) is the standard treatment for patients with locally advanced rectal cancer before surgery [46]. Moreoer, it has been demonstrated that nCRT and total mesorectal excision had resulted in substantial improvement in local disease control for rectal cancer [35], [36], [37].

It is necessary to restage the rectal tumor post-nCRT before making surgical decisions [33], [46]. Specifically, a ''watch-and-wait'' process is widely adopted to evaluate tumor treatment response and decide whether the rectal is preservable [34], [39], [40], [42]. The mainstay of staging and restaging rectal tumors is magnetic resonance scanning of the pelvis, which helps physicians accurately access tumor location, tumor extent, invasion to nearby organs, and regional

---

The associate editor coordinating the review of this manuscript and approving it for publication was Orazio Gambino.

(a) 3D view

(b) SI   (c) LR   (d) AP
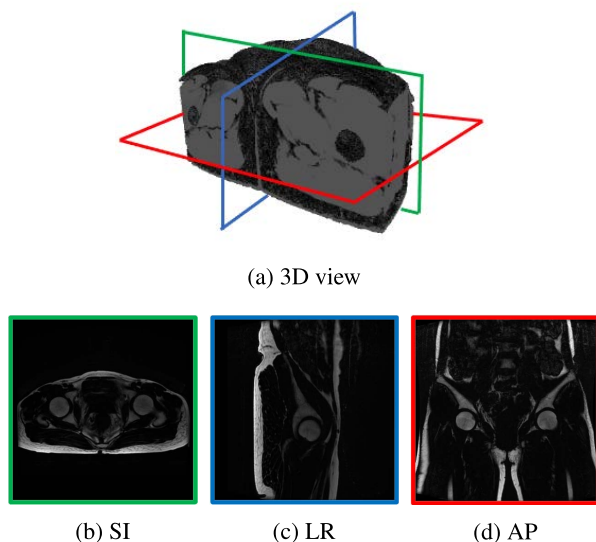
FIGURE 1. An illustration of three possible scanning directions of MRI data in our dataset. The green, blue, and red rectangular in 3D view represents the scanning in SI, LR, and AP directions, respectively. Example images in these three directions are shown in the bottom row in order. In our rectal data, only one of the three directions are captured in full-resolution. For illustration purpose, we showed full-resolution images for all three directions. Besides, all the images shown in this figure are from session 1, which is pre-treatment.

lymph node metastasis. Nowadays, several methods for evaluating tumor response to treatment have been developed. The most popular one is the Response Evaluation Criteria In Solid Tumors (RECIST). Comparing with methods based on functional or metabolic changes [43], [45], RECIST is easier to perform because it is based on anatomical change. To support RECIST in rectal cancer, an accurate alignment for volumes pre and post-nCRT, called multi-session volumes, is needed. Therefore, volume registration approaches, which estimates a transformation function between multiple volumes, is a good fit to support RECIST.

However, our rectal MRI data presents a challenging situation to volume registration. First of all, the ground truth transformation functions among multi-session volumes are unavailable, making supervised learning an infeasible way to train a registration network. In addition, it is difficult to set a fixed scanning position in the pelvic region in multi-session, so we usually face a high degree of transformation. It is also notable that the organs usually undergo uneven changes over time, e.g., the bladder filling or weight change after a period of nCRT, which takes about three months. Therefore, it is hard to adopt one-to-one correspondence between multi-session rectal data as an assumption, which is made by many existing methods. Due to these limitations, an unsupervised method without the assumption of one-to-one correspondence is needed. An example of 3D rectal data is shown in Fig. 1. We can see that a total of three different scanning directions: superior-inferior (SI), left-right (LR), and anterior-posterior (AP) are used for data collection.

To address these issues, we propose a rectal registration model trained by unsupervised learning. Specifically,

we perform data synthesis to provide pseudo paired data for model training. The proposed model consists of four components: feature extraction, feature description, affine matrix estimation, and reweighed deformable registration. The first component applies SuperPoint Network (SPN) to detect key points from MRI volumes. Then, each key point and its surrounding features are extracted by the feature descriptor, a ten-layer 3D convolutional neural network (CNN). The affine matrix estimate is completed by two-way-nearest-neighborhood (TWNN) and random sample consensus (RANSAC), which matches the key points from multi-session volumes according to corresponding descriptions. The last component further aligns the boundaries but preserves the structure of interior features of rectal. By leveraging these components, the proposed method provides accurate alignment between fixed and moving data regardless of a high degree of deformation between them.

Our contributions are highlighted as follows:
- We propose a 3D rectal registration approach that outperforms state-of-the-art methods for this application.
- The proposed approach accurately aligns important features such as Femur head and Internal obturator muscle between multi-session rectal volumes.
- The proposed approach does not rely on ground truth transformation function or one-to-one correspondence between multi-session rectal volumes.

The remainder of the paper is organized as follows: Sec. II reviews the existing approaches to medical image registration. The proposed method is described in detail in Sec. III. The experimental results and possible extensions are presented in Sec. IV and Sec. V, respectively. Finally, we give a conclusion in Sec. VI.

II. RELATED WORKS

There has been a substantial amount of research in the 3D medical image registration problem, including both conventional and deep learning approaches. Conventional rigid or deformable registration methods usually iteratively optimize parameters of transformation models according to a predefined energy function that evaluates the similarity of source and target imaging volumes [15], [16], [17], [22]. Specifically, rigid transformation is often used to register bones-like features [22]. Reuter *et al.* proposed an algorithm iteratively computes the affine transformation matrix that preserves inverse consistency [2]. It can be implemented by FreeSurfer [1], a common software for MRI image processing. Moreover, considering the deformable nature of most tissues, many approaches adopt a deformable model to pursue further improvements, including Maxwell's or Diffeomorphic demons [15], [16], free form deformation (FFD) b-splines [17], [18], and fluid flow model [19], [20]. Also, mutual information [48] and kernel regression [49] based methods have been applied to deformable medical image registration.

Recently, deep learning methods have achieved significant improvement in runtime without sacrificing registration

performance [9], [24], [25], [26], [28], [29], [31]. Vos *et al.* proposed a coarse-to-fine pipeline that cascades an affine estimation network with B-spline estimation network based on deep learning [24]. Although making registration 350 times faster than conventional methods, the applicable resolution is limited due to fully connected layers. Balakrishnan *et al.* proposed VoxelMorph (VXM), a fully convolutional network for deformable MRI image registration that is able to process a variety of resolution [31]. It demonstrated comparable accuracy to conventional methods while drastically decreased the runtime to be less than a second [31]. However, the method only works well on images pre-aligned by the algorithm proposed by Reuter *et al.* [2]. Yang *et al.* proposed a neural network called LapIRN that estimates the diffeomorphic maps instead of displacement field to preserve topology and smoothness of deformation [29]. ProsRegNet (PRN) is a deep learning based approach to coarse-to-fine registration between MRI and 2D histopathology prostate cancer images [32]. However, PRN requires manual pre-processing to crop and mask the region of interest for every input pair, which significantly limits the applicability of the approach in clinical practice. Among these methods, both PRN and VXM claimed themselves the state-of-the-art approach for coarse-to-fine medical image registration, while LapIRN claimed that it is the state-of-the-art approach for deformable medical image registration.

Unlike previous approaches, the proposed 3D image registration framework does not require any pre-processing before estimating the correspondence between the paired data. In addition, it does not require the presence of one-to-one correspondence between paired data. With these properties, our proposed framework is a more general choice than existing ones for 3D MRI volume registration.

III. PROPOSED APPROACH

In this section, we introduce every process of the proposed pipeline in detail. An overview of our proposed pipeline is illustrated in Fig. 2. It takes a pair of multi-session volumes, also called fixed and moving volumes (V_f and V_m), as input, and warps V_m toward V_f. It contains four main procedures: key point detection, feature description, affine matrix estimation, and reweighed deformable registration. They are achieved by SPN, feature descriptor (FD), the combination of TWNN and RANSAC 3D [3], and reweighed deformable registration (R-VXM), in order. Specifically, as shown in the figure, their outputs are key points $\{p_1, \ldots, p_n\}$ and $\{q_1, \ldots, p_m\}$, feature vectors $\{f_1, \ldots, f_n\}$ and $\{g_1, \ldots, g_m\}$, corresponding pairs $\{(p_{i1}, q_{j1}, \ldots, p_{ik}, q_{jk})\}$, the affine matrix estimate $\tilde{A}$, and the final registration result V_o, which is the warping result of V_m. We also describe how we synthesize training data for unsupervised learning and how triplet learning is deployed for model training.

A. KEY POINT DETECTION
The first component of the pipeline is the key point detection, which detects the position with rich information and extracts

TABLE 1. Network architecture of feature descriptor.

Layer #	Input shape	Output shape	(k, s, p)
1	(1, 64, 64, 16)	(64, 64, 64, 16)	(5,1,2)
2	(64, 64, 64, 16)	(64, 32, 32, 8)	(5,2,2)
3	(64, 32, 32, 8)	(128, 32, 32, 8)	(3,1,1)
4	(128, 32, 32, 8)	(128, 16, 16, 4)	(3,2,1)
5	(128, 16, 16, 4)	(256, 16, 16, 4)	(3,1,1)
6	(256, 16, 16, 4)	(256, 8, 8, 2)	(3,2,1)
7	(256, 8, 8, 2)	(256, 8, 8, 2)	(3,1,1)
8	(256, 8, 8, 2)	(256, 4, 4, 1)	(3,2,1)
9	(256, 4, 4, 1)	(256, 4, 4, 1)	(3,1,1)
10	(256, 4, 4, 1)	(256, 2, 2, 1)	(3,2,1)*

k: kernel size, s: strides, p: padding. Input and output shapes show channel, width, height, and depth in order.
*: the stride along width and height is 2 while is 1 along depth.

its corresponding features. We apply the SuperPoint Network (SPN) [14] as a key point detector. SPN consists of three components: VGG-style encoder [12], interest point decoder, and descriptor decoder. Specifically, the VGG-style encoder consists of 8 convolutional layers, followed by two parallel decoders, both consist of 2 convolutional layers. The two decoders generate a set of key point $p_i \in \mathbb{R}^3$ and corresponding feature vectors $f_{pi} \in \mathbb{R}^{256}$, respectively. In particular, for each slice $V_{[k]} \in \mathbb{R}^{w \times h}$ obtained from volume $V \in \mathbb{R}^{w \times h \times d}$, we got key points $P_k : \{p_{k0}, \ldots, p_{ka}\} = SPN(V_{[k]})$ at that slice, where $p_{ki} \in \mathbb{R}^3$. Then, we merge all the P_k, $1 \leq k \leq d$ to form the set of key point $\{p_1, \ldots, p_n\}$ distributed over the volume V. Mathematically speaking,

$$P = \{p_i | p_i \in P_k, 1 \leq k \leq d\} \tag{1}$$

B. FEATURE DESCRIPTION
Since most of existing feature description approaches accompanied by the key point detection are developed for 2D data, we develop our own 3D feature description approaches (FD). For every key point p_i detected by the SPN, we crop a surrounding volume $v(p_i) \in \mathbb{R}^{48 \times 48 \times 16}$ to be the input of FD. Specifically, the feature descriptor network is a 10-layer 3D convolutional network where every convolutional layer is followed by 3D batch normalization [13] and a scaled exponential linear unit (SELU) [8]. In the last layer, we apply normalization and flatten the output to be a feature vector $f_{id} = FD(v(p_i)) \in \mathbb{R}^{1024}$ as a descriptor of the corresponding key point. The detailed architecture of the feature descriptor is shown in Table 1. In addition, to incorporate complete 2D information, we concatenate f_{id} with feature vector $f_{pi} \in \mathbb{R}^{256}$ extracted along with p_i by SPN, mentioned in Sec. III. A. Therefore, we obtain a 1280 dimensional feature vector for each key point.

C. AFFINE MATRIX ESTIMATION
Affine matrix estimation is completed in two steps: TWNN and RANSAC 3D. Denote the feature vectors of key points from fixed and moving images, obtained by FD, as fixed and moving vectors, respectively. The two groups of vectors are matched by TWNN as follows. For each moving feature
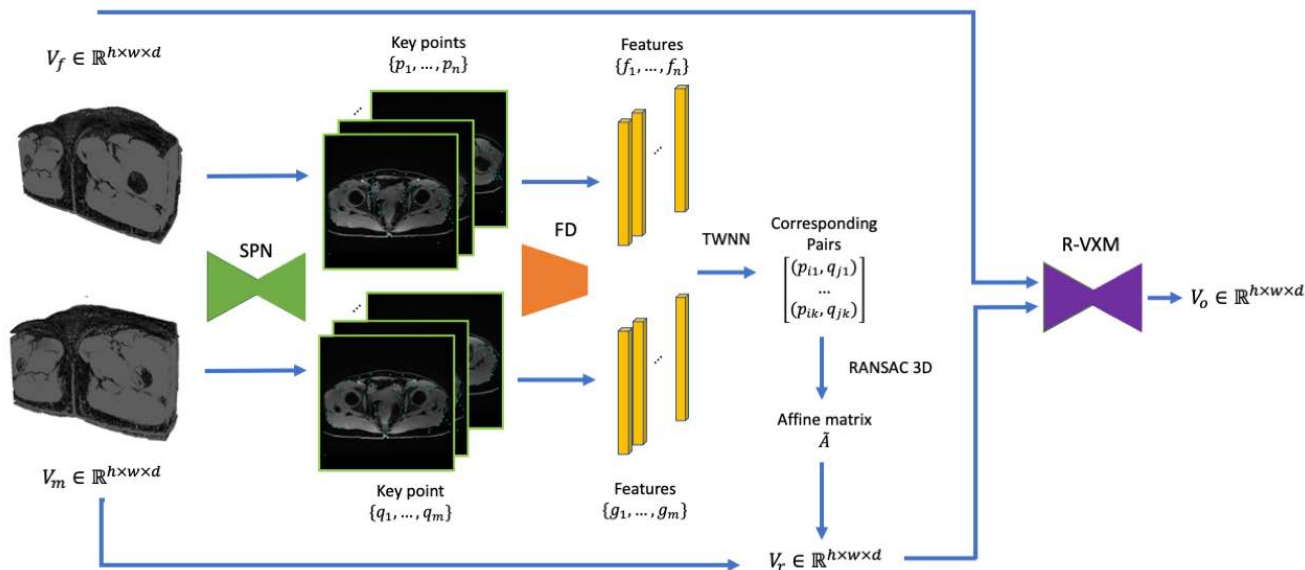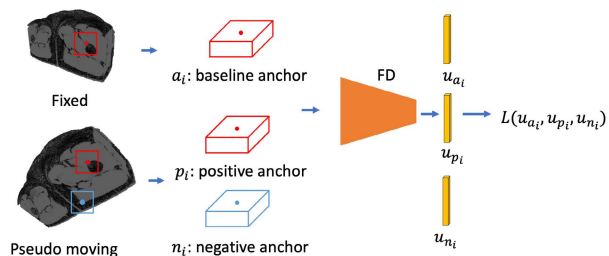
FIGURE 2. An overview of our proposed pipeline. V_f and V_m are the input multi-session volumes, called fixed and moving volumes, respectively. SPN, FD, and TWNN denote Superpoint network, feature descriptor, and two-way-nearest-neighborhood, respectively. The outputs of these three components are key points $\{p_i\}$ and $\{q_i\}$, feature vectors $\{f_i\}$ and $\{g_i\}$, and corresponding pairs $\{p_{ik}, q_{jk}\}$. $\tilde{A}$ is the affine matrix generated by RANSAC 3D, and R-VXM denotes reweighed deformable registration completed by voxelmorph. V_r nad V_o are preliminary and final registration results of the pipeline, respectively.
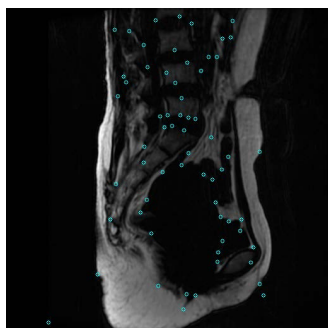
FIGURE 3. An example of key points detected from a slice of rectal volume scanned in LR direction. Key points are marked in cyan circles.


FIGURE 4. A configuration of triplet learning performed in this paper. The red point denotes the baseline anchor from the fixed volume, while red and blue points denote positive and negative anchors from the pseudo moving volume. FD denotes feature descriptor, and u_{a_i}, u_{p_i}, and u_{n_i} denote the feature extracted by baseline, positive, and negative anchor, respectively. $L(u_{a_i}, u_{p_i}, u_{n_i})$ is the loss function that guides the learning process.

vector, we find a fixed vector that is the closest to it, called its nearest neighbor. Specifically, we use cosine distance $d(u, v) = 1 - u^T v$ to measure the distance between feature vectors u and v. If a moving and fixed feature are the nearest neighbor of each other, we call them a matching pair. Accordingly, a set of k matching pairs $\{(p_{i\triangle}, q_{j\triangle}), \ldots, (p_{ik}, q_{jk})\}$ is generated, where p_i and q_j are 3-dimensional vector consisting of coordinate $p_i = (x_i, y_i, z_i)$ and $q_j = (v_j, u_j, w_j)$ from fixed and moving volumes, respectively.

RANSAC 3D is an iterative process that computes the affine matrix that represents the transformation between matching pairs. In each iteration t, it randomly picks six matching pairs and applies the six-point-algorithm [5] to compute a 4×4 affine matrix A_t that best describes the transformation of these eight pairs. Then, we measure the distance between p'_i and $A_t^T q'_i$ for all $1 \leq i \leq n$, where $p'_i = [p_i, 1]$ and $q'_i = [q_i, 1]$. If p'_i satisfies $||p'_i - A_t^T q'_i||_2 \leq d$, where $d = 5$ is the tolerant distance, it is counted as an

inlier of matrix A_t. RANSAC 3D terminates when the number of inliers is beyond a threshold, or a maximum iteration is reached. When it terminates, the matrix A_t that allows the most inliers becomes the output of RANSAC 3D, which is denoted as $\tilde{A}$ in Fig 2. In other words, the affine transformation matrix $\tilde{A}$ represents the rigid transformation estimate between fixed and moving volumes.
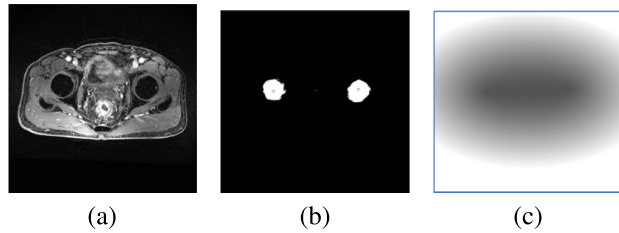
D. REWEIGHED DEFORMABLE REGISTRATION

It is notable that the transformations of rectal boundaries are too complex to be characterized by an affine matrix. Hence, deformable registration is needed to fine-tune the results from rigid registration. However, we note that if they are direct combined, undesirable artifacts are easily resulted because of the lack of one-to-one correspondence between fixed and moving data. Furthermore, when unexpected
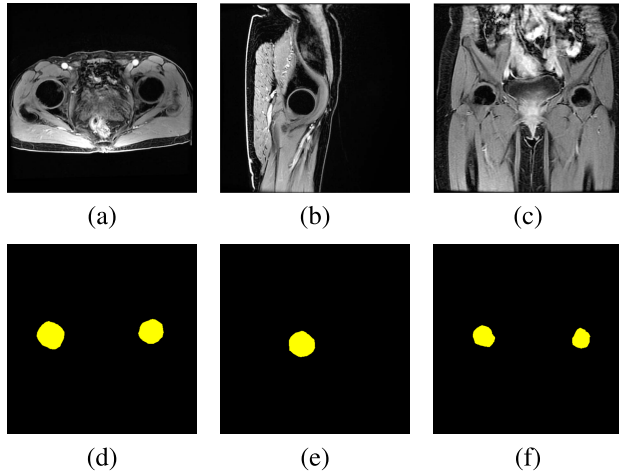
FIGURE 5. An example of manual feature labeling of data in three different scanning directions. Subfigure (d), (e), and (f) depict features labeled from (a), (b), and (c), respectively.

FIGURE 6. Examples of (a) raw image, (b) segmentation map, and (c) reweighing mask for deformable registration.

distortions occur on the interior features, a performance degradation will be caused. Therefore, when deformable registration is applied to align the boundaries, the preservation of interior features should be carefully considered.

To address this issue, we apply a reweighed deformable registration to rewarp V_r, the preliminary registration results from $\tilde{A}$. The mask to reweigh the flow field generated by VXM is computed as follows: We first obtain a segmentation map that identifying Femur head and Internal obturator muscle, an interior feature of rectal data. Then, a mask is generated according to the distance from each pixel to the center of labeled features. The closer a pixel is to the center of features, the lower value is assigned on the mask. An example of segmentation map and its corresponding mask is shown in Fig. 6. Specifically, the mathematical form of mask $M_{V,k}$ of slice $V_{[k]}$ is

$$M_{V,k} = max(1, (d_0 + d_1)/s), \quad (2)$$

where d_0 and d_1 are the distance to the two bones, and s is an hyper-parameter for scaling. Note that we only consider d_0 in LR scanning because at least one of the circle bones appear in slices scanned in that direction.

As observed, by restricting the impact of flow field on interior features, we can fine-tune the structure without causing serious distortion. It is notable that the flow field at boundaries is preserved, so it has enough intensity to drive the alignment at boundaries. In addition, when running reweighed

VXM, we concatenate V_r and the fixed volume V_f as input. Then, the output flow field is multiplied by the mask and guides the deformable registration to finalize the whole registration process.

E. DATA SYNTHESIS
Recall that a challenge to conduct learning-based registration approaches is the lack of one-to-one correspondence and ground truth transformation function. To address the limitation, we simulate an affine matrix T to generate pseudo paired data for 10 MRI volumes, covering all three scanning directions. A total of three different types of transformations are considered in the simulation: translation, rotation, and resizing, whose degrees are within $[-50, 50]$, $[-30°, 30°]$, and $[0.8, 1.2]$, respectively, in all three dimensions. By considering these three types of deformations, we are able to generate pseudo paired data with diverse transformations randomly. When pseudo paired data with labeled transformation T become available, model training is achievable.

F. MODEL TRAINING
We use synthesized paired data described in Sec. III-E to train the feature descriptor network. In particular, triplet learning [11] is employed to train the feature descriptor network. Fig. 4 depicts the scheme of triplet learning discussed in this paper. For synthesized paired data, the transformation matrix T between fixed and pseudo moving data is known as prior. Hence, for a key points a_i from fixed volume (called baseline anchor), we pick a positive anchor p_i satisfying $||a_i' - Tp_i'||_2 \leq \tau$, where $\tau = 5$ is a hyperparameter denoting tolerant distance. We also pick a negative anchor n_i that does not satisfy this condition.

In the training phase, we first randomly sample baseline anchors and their associated positive and negative anchors. For each anchor at the center, we crop a volume with size $(48 \times 48 \times 16)$ around it. After feeding the volumes to the feature descriptor, corresponding feature vectors $u_{a_i}, u_{p_i}, u_{n_i}$ are obtained. The loss function $L(u_{a_i}, u_{p_i}, u_{n_i})$ is defined as follows:

$$L(u_{a_i}, u_{p_i}, u_{n_i}) = \max(d(u_{a_i}, u_{p_i}) - d(u_{a_i}, u_{n_i}) + m, 0), \quad (3)$$

where $d(u, v) = 1 - u^T v$ measures the cosine distance between u and v and $m = 1$ is the margin. Through the loss function, it is encouraged to have u_{a_i}, u_{p_i} close to each other, while u_{a_i}, u_{n_i} far away from each other. In other words, it encourages features extracted from an anchor to be close to its positive anchor, but be far away from the negative one. The Adam optimizer [10] is applied to guide the optimization process and the batch size is set as 128.

IV. EXPERIMENTS
In this section, we describe how the experiments are conducted to compare the registration performance of the proposed pipeline with four existing frameworks: FreeSurfer (FS) [1], VoxelMorph (VXM) [31], ProsRegNet (PRN) [32], and LapIRN [29]. Specifically, we took pre-trained model for

all learning-based baseline methods. Note that before performing VXM, we pre-align images using FS as suggested by the literature. We also perform ablation test by only using the rigid registration part of the proposed pipeline, denoted as Ablated. The registration performance is evaluated in both objective and subjective metrics.

A. EVALUATION METRICS

In subjective tests, we use chessboard images, which alternatively show the patches from fixed and moving images, to evaluate the continuity between fixed and registered image [9]. We also manually label binary feature maps of Femur head and Internal obturator muscle, whose shape and size are invariant over treatment, from MRI data. The shape of these features are two balls in 3D views. In AP and SI scanning, two circles are shown in slices of similar or nearby locations. In LR scanning, one of the circle is shown in the first few slices and the other one is in the last few slices. Examples of labeled features of MRI data in different scanning directions are shown in Fig. 5.

Then, we compute the Dice score, a metric widely used in image registration [4], [29], [31], to perform objective evaluation. It computes the overlapped region between binary feature maps binary maps I_1 and I_2 by

$$Dice(I_1, I_2) = \frac{2\Sigma(I_1 \circ I_2)}{\Sigma I_1 + \Sigma I_2}. \tag{4}$$

B. MRI DATA COLLECTION AND PREPROCESSING

Using an institutional review board approved protocol, we retrospectively reviewed data on patients with rectal cancer who were treated with nCRT from February 2017 to October 2020 at the 108 Military Central Hospital, Vietnam. Data from 16 patients were identified for this study. All patients underwent MRI scans for the rectum and pelvic regions using a Discovery MR750w 3.0 Tesla GEM scanner. With the setup, tumor status in different sessions can be easily accessed to all the three scanning directions.

The size of a 3D rectal MRI data is $(512, 512, d)$, where d is the depth that varies over samples. Specifically, d ranges from 70 to 160 in our rectal data. In each volume, only one of the three scanning directions shown in Fig. 1 is captured in full resolution, 512×512, while other two directions are captured in lower resolution, $512 \times d$. For each patient, we scanned rectal data before and after nCRT, obtaining fixed and moving volumes, respectively. It is notable that the depth ranges of fixed and moving volumes may be inconsistent. For data partition, we use ten paired MRI data to train our feature descriptor, and 24 paired data for testing, 8 for each scanning direction.

C. QUANTITATIVE RESULTS

Table 2 shows the resulted Dice score of feature maps under different registration approaches. We note that the proposed method yields the highest Dice score among all competitors in all directions. The superiority of our proposed method is further demonstrated in Fig. 7. As observed, even the features of paired data are at a distance from each other before registration, the proposed method is still able to provide accurate alignment, which is not the case for other methods. Compared with other methods, the proposed method yields a higher overlapped feature maps after registration. The analysis about the inferior performance caused by the competitors is discussed in Sec. IV-E.

D. QUALITATIVE RESULTS

The chessboard-like images for different registration approaches are shown in Fig. 8. As observed, the proposed algorithm yields the most accurate alignment over the entire image. The boundaries of most patches are continuous, including the rectal surface and its interior circular bone. Furthermore, when comparing the results from the full and ablated version of proposed pipeline, it is observed that the continuities at boundaries are enhanced after reweighed deformable registration is applied. Therefore, although only slight improvement on Dice score is resulted by applying reweighed deformable registration, the contribution to the alignment over boundaries is unignorable. One may note that the results from VXM also show high continuities over images. Nonetheless, VXM is still inferior to our proposed method due to its fragility to paired data with inconsistent depth. More analysis is described in Sec. IV-E.

We also present chessboard images along other axis to verify the 3D integrity of registration results. Fig. 9 demonstrates that the proposed method is able to register MRI data scanned in all three directions accurately. In other words, an accurate alignment is achieved no matter which dimension that the volume is projected in. The results verify the robustness of our registration method to deformations in all three dimensions.

E. DISCUSSION

Among all the competitors to our proposed method, we first observe that FS yields moderate Dice score. However, we also note that it usually fails to provide accurate registration when the depths of source and moving data differ significantly, while our proposed algorithm is robust to such inconsistency. A typical example is shown in Fig. 10, where the depths of fixed and moving volumes are 92 and 124, respectively. In this case, our method is robust to the inconsistent depth and provides accurate alignment after rigid registration. Nonetheless, FS fails to recover the transformation between paired data and results in a serious mismatching. Therefore, Femur head and Internal obturator muscle of fixed and warped volumes are not aligned under FS, causing a low Dice score. It is notable that inconsistent depth is unavoidable in rectal data collection, and our proposed method is a better choice than FS due to its robustness to this issue.

We also notice that VXM achieves higher Dice score than FS because it fine-tunes the output generated by FS. However, when encountering a serious mismatching due to inconsistent depth, such as the case in Fig. 10, VXM is not able to recover the mismatching. Recall that VXM only works well when
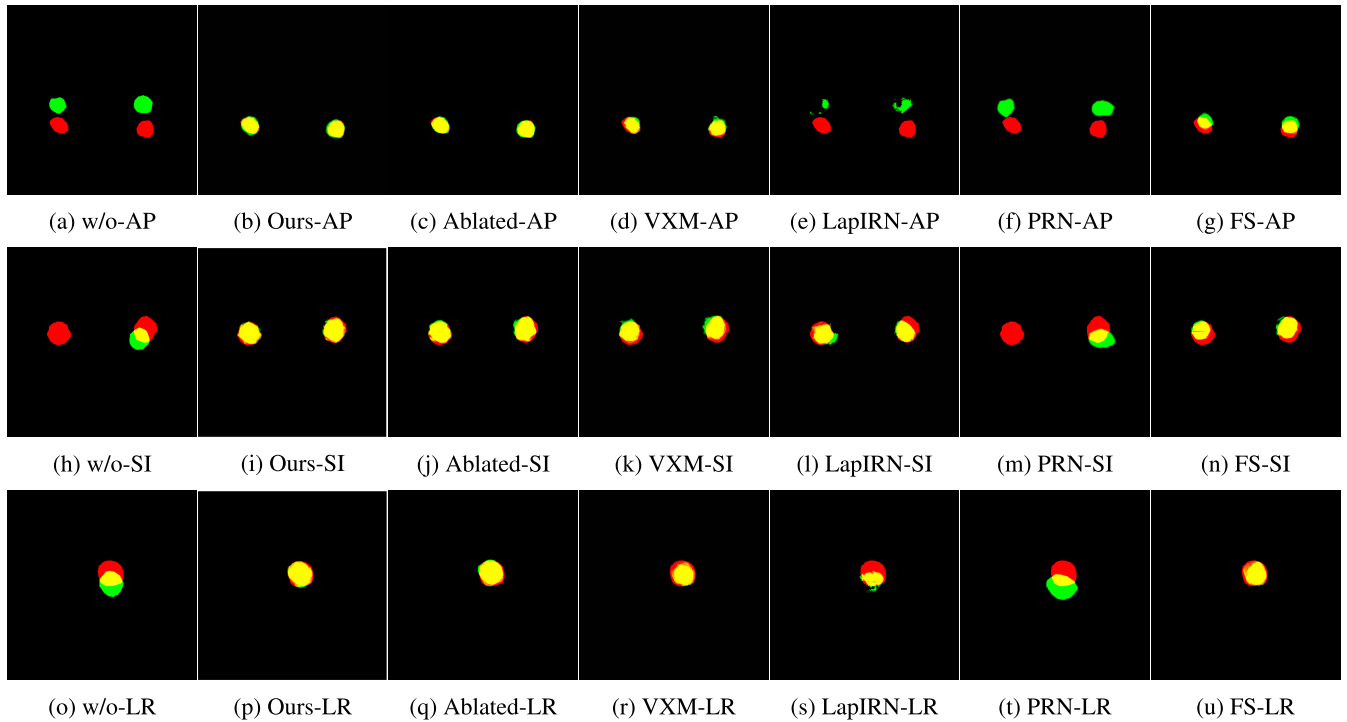
FIGURE 7. Examples of overlapped segmentation maps that demonstrate the registration results done by different registration methods. The red and green feature maps are from slices of fixed and moving volume, respectively, and the yellow maps represent the overlapped region between the maps from two slices. A higher overlapping region is better. The combination of method and direction is written as method-direction" (i.e., VXM-SI denotes VoxelMorph in SI direction).

TABLE 2. Dice score of feature-based binary map with different registration approaches.

Scanning Direction	w/o registration	Ours	Ablated	FS	VXM	LapIRN	PRN
SI	0.149	**0.806**	0.799	0.598	0.642	0.294	0.118
AP	0.202	**0.737**	0.726	0.557	0.580	0.471	0.218
LR	0.306	**0.816**	0.799	0.779	0.807	0.449	0.358

The highest score is shown in boldface.

the rigid registration is accurate. Moreover, a serious performance degradation will be encountered if FS is not applied to pre-align the paired volume before performing VXM. Three examples of registration results by directly applying VXM without FS are shown in Fig. 11. As observed, undesirable artifacts are easily generated over the rectal volume if FS is not applied before VXM.

Among all the competitors, PRN has the worst performance. Although it took MRI data from prostate, which is closer than brain images to our rectal data, for model training, it provides inaccurate alignment in almost all paired data. The inferior performance was mainly due to three reasons. First, PRN is designed for 2D registration, assuming no offset in the vertical direction, which is not the case in our dataset. Second, PRN assumes multi-modality registration, taking histology and MRI data as input to estimate the correspondence. However, both fixed and moving images in our dataset are in the same modality. Third, the input data for PRN are manually pre-processed, while we work on raw images. Consequently,

PRN is not suitable to register rectal volumes in our dataset and the poor Dice scores are resulted.

LapIRN achieves the best performance among deformable approaches. It is able to recover overall transformation between fixed and moving volumes, when the degree is moderate. However, it usually fails to preserve the integrity of interior features, such as examples shown in Fig. 7, when warping the moving volume toward the fixed one. As LapIRN is a deformable approach, whose transformation function is with few restrictions, it easily overly transforms the moving volume and causes artifacts like examples shown in Fig. 11. Moreover, it usually incorrectly deforms the tumor of moving images, which is misleading when RECIST is conducted according to the registration results. As presented in Sec. II, all the competitor approaches are state-of-the-art approaches for deformable or coarse-to-fine registration. Based on the qualitative and quantitative comparison, our proposed unsupervised approach outperforms these approaches and is a better choice for rectal data registration.
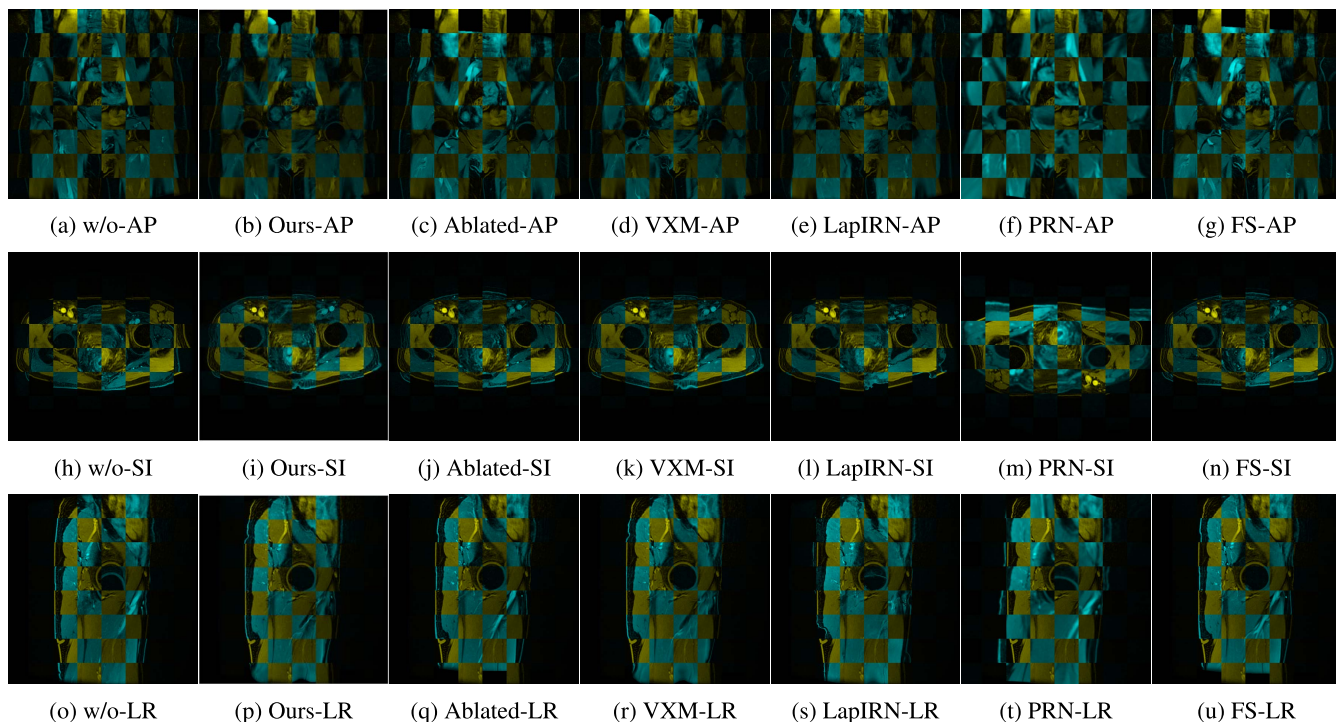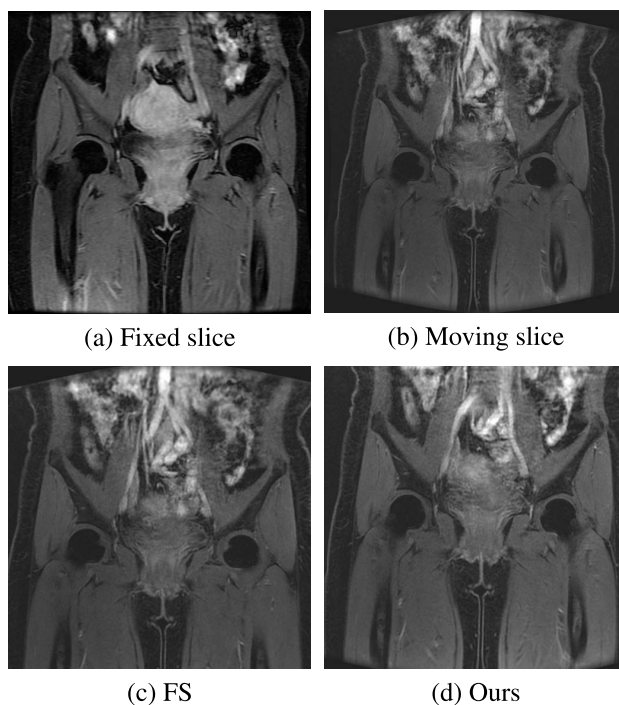
FIGURE 8. Examples of chessboard images that demonstrate the registration results for different methods. Yellow and cyan patches are from slice of fixed and moving volume, respectively. Similar to Fig. 7, the combination of method and direction is written in "method-direction".
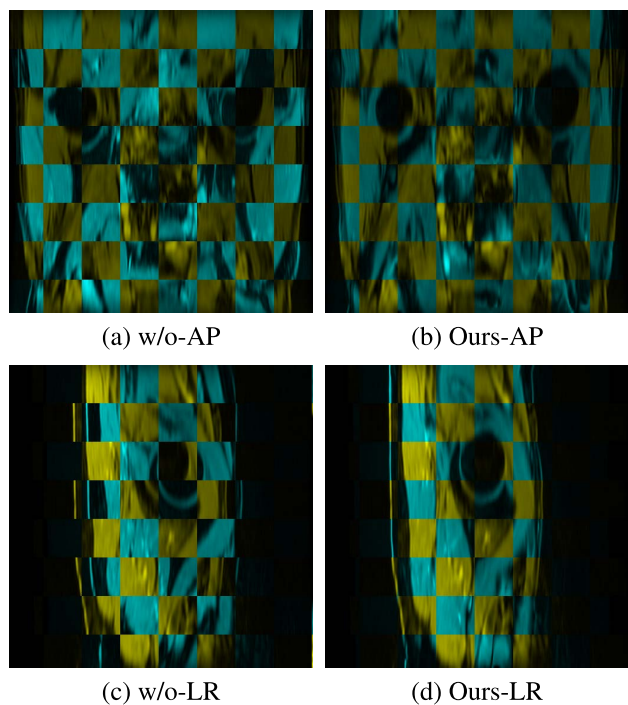
FIGURE 9. Chessboard images in (a) (b) AP and (c) (d) LR projection of a rectal data whose full-resolution slices are scanned in SI direction. Yellow and cyan patches are from slice of fixed and moving volume, respectively.

FIGURE 10. An example of registration results from FS (c) and the proposed algorithm (d) when the depths of fixed (a) and moving (b) data are inconsistent. Depths of fixed and moving data are 92 and 124, respectively.

When comparing chessboard images generated by full and ablated version of the proposed method, we observe that introducing reweighed deformable registration as fine-tuning greatly improves alignment at rectal surfaces and brings the
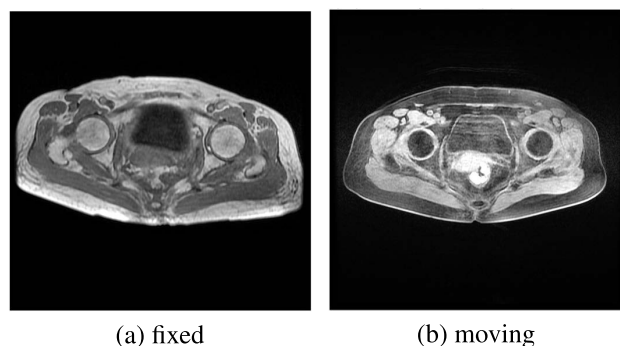
paired data even close to each other without causing artifacts. This is because rigid transformation focuses on translation, scaling, and rotation, while deformable transformation
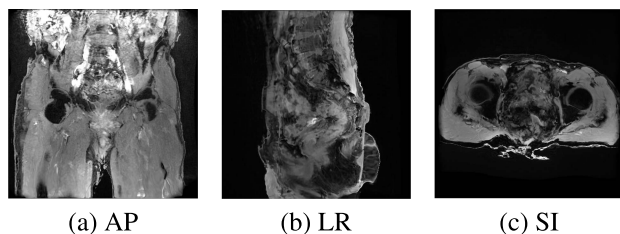
FIGURE 11. Examples of artifacts in all three scanning directions generated by directly applying VXM after rigid registration.

(a) fixed   (b) moving

FIGURE 12. An example of paired data with color mismatching, where (a) and (b) are slice from fixed and moving volume, respectively.

has more flexibility to characterize the change of thickness and shape of outermost layer of rectal. However, due to the lack of one-to-one correspondence, changing the structure of rectal surfaces is sometime risky. For example, the surface may become thicker after nCRT due to normal growth, but a deformable registration approach may yield a thickness reduction to reach alignment, which may wipe out information in moving data. Thus, the integration of rigid and deformable registration should be carefully conducted in the applications where surfaces over time is important for clinical analysis.

Furthermore, it is notable that the superior performance is achieved by using only ten pairs of rectal volumes as training data. In other words, it is an economical choice in model development. As opposed to the amount of training data used in previous approaches [29], [31], [32], the proposed method takes much less data for model training, and hence has less computational burden in model training. Moreover, as accurate registration results is provided with acceptable computational cost, physicians are able to efficiently evaluate the stage propagation of rectal cancer and make decisions accordingly, making it a practical tool for clinical analysis. Therefore, we believe the proposed pipeline has high potential in supporting RECIST to evaluate the effect of treatment.

V. POSSIBLE EXTENSIONS
In this section, we describe possible extensions to our registration approach, with a focus on paired data with color mismatching or different modalities.

It is still an open problem to register rectal data across multi-modality. Because the settings for image collection

usually change overtime, color of data in multi-session may mismatch, as the example shown in Fig. 12. As observed, except for the background, the range of voxel intensities of slices from fixed and moving volume differs a lot. When the colors of the two images are mismatched, feature description usually becomes erroneous and affects the accuracy of the affine matrix estimate. A possible solution is applying volume transformation before feeding mismatched paired data to the proposed pipeline. Specifically, an option is to adopt style transfer network, which has been successfully applied to retina image registration [4], to bring the paired data into the same modality. Once the transformation works well, the problem can be reduced to that discussed in this paper. Therefore, we believe the collaboration between style transfer network and our method can broaden the range of applications of image registration.

Multi-modality also happens when T1w and T2w are both deployed in data collection. The former with contrast enhanced, and the latter has various resolutions to easily access the tumor stage. In this paper, we focus on registering T1w volumes in multi-session, while making the registration of T2w scanning an open problem. Specifically, it is challenging to our proposed method due to two reasons. First, because the field of view in T2w scanning is much smaller than that of T1w, a higher degree of distortion is observed in data scanned in T2w. Moreover, T2w scanning captures smaller areas than T1w does, so fewer key points are detected, and the clues for image registration may be insufficient. Therefore, we leave it as a future work to capture sufficient features even in T2w rectal data and enhance the robustness to different field of views. Hopefully, multi-modality registration for rectal data could reveal more information of rectal cancer for clinical analysis.

VI. CONCLUSION
Rectal cancer has long raised attention in medical science due to its high mortality, and its staging is essential to diagnosis. When images in multiple sessions are well-aligned, the propagation of the tumor and the efficacy of nCRT treatment can be easily observed. In this paper, we present a registration model that leverages data synthesis, feature detection, description, and matching, and reweighed deformable registration. It achieves accurate image registration without the need for ground-truth labeling or one-to-one correspondence, making itself a general approach to rectal MRI volume registration. Moreover, in terms of Dice score of Femur head and Internal obturator muscle, the proposed model outperforms existing registration methods by achieving 0.806, 0.737, and 0.816 in SI, AP, and LR scanning direction, respectively. Its importance to clinical analysis is demonstrated through the high continuity shown in chessboard images and its potential value to increase the efficiency of RECIST evaluation.

REFERENCES

[1] B. Fischl, "Freesurfer," *NeuroImage*, vol. 62, no. 2, pp. 774–781, Aug. 2012.

[2] M. Reuter, H. D. Rosas, and B. Fischl, "Highly accurate inverse consistent registration: A robust approach," *NeuroImage*, vol. 53, no. 4, pp. 1181–1196, Dec. 2010.

[3] M. Fischler and R. Bolles, "Randomphsample consensus: A paradigmphfor model fitting with applications to image analysis and automated cartography," *Commun. ACM*, vol. 24, pp. 381–395, Jan. 1981.

[4] J. Zhang, C. An, J. Dai, M. Amador, D.-U. Bartsch, S. Borooah, W. R. Freeman, and T. Q. Nguyen, "Joint vessel segmentation and deformable registration on multi-modal retinal images based on style transfer," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2019, pp. 839–843.

[5] A. Torii, Z. Kukelova, M. Bujnak, and T. Pajdla, "The six point algorithmphrevisited," in *Proc. Asian Conf. Comput. Vis.*, 2010, pp. 184–193.

[6] Society, A. Key Statistics for Rrostate Cancer. Accessed: Jan. 12. 2022. [Online]. Available: https://www.cancer.org/cancer/prostate-cancer/about/key-statistics.html

[7] W. Bi, A. Hosny, M. Schabath, M. Giger, N. Birkbak, A. Mehrtash, T. Allison, O. Arnaout, C. Abbosh, and I. Dunn, "Artificial intelligence in cancer imaging: Clinical challenges and applications," *CA, A Cancer J. Clinicians*, vol. 69, pp. 127–157, Jan. 2019.

[8] G. Klambauer, T. Unterthiner, A. Mayr, and S. Hochreiter, "Self-normalizing neural networks," in *Proc. Int. Conf. Neural Inf. Process. Syst.*, 2017, pp. 972–981.

[9] Y. Wang, J. Zhang, M. Cavichini, D.-U.-G. Bartsch, W. R. Freeman, T. Q. Nguyen, and C. An, "Study on correlation between subjective and objective metrics for multimodal retinal image registration," *IEEE Access*, vol. 8, pp. 190897–190905, 2020.

[10] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*.

[11] E. E. Hoffer and N. Ailon, "Deep metric learning using triplet network," in *Proc. Int. Workshop Similarity Pattern Recognit.*, 2015, pp. 84–92.

[12] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*.

[13] S. S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *Proc. Int. Conf. Mach. Learn.*, 2015, pp. 448–456.

[14] D. DeTone, T. Malisiewicz, and A. Rabinovich, "SuperPoint: Self-supervised interest point detection and description," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2018, pp. 224–236.

[15] J.-P. Thirion, "Image matching as a diffusion process: An analogy with Maxwell's demons," *Med. Image Anal.*, vol. 2, no. 3, pp. 243–260, Sep. 1998.

[16] T. Vercauteren, X. Pennec, A. Perchant, and N. Ayache, "Diffeomorphic demons: Efficient non-parametric image registration," *NeuroImage*, vol. 45, no. 1, pp. S61–S72, Mar. 2009.

[17] D. Rueckert, L. I. Sonoda, C. Hayes, D. L. G. Hill, M. O. Leach, and D. J. Hawkes, "Nonrigid registration using free-form deformations: Application to breast MR images," *IEEE Trans. Med. Imag.*, vol. 18, no. 8, pp. 712–721, Sep. 1999.

[18] D. Rueckert, P. Aljabar, R. Heckemann, J. Hajnal, and A. Hammers, "Diffeomorphic registration using B-splines," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.*, 2006, pp. 702–709.

[19] G. E. Christensen, R. D. Rabbitt, and M. I. Miller, "3D brain mapping using a deformable neuroanatomy," *Phys. Med. Biol.*, vol. 39, no. 3, pp. 609–618, Mar. 1994.

[20] E. D'Agostino, F. Maes, D. Vandermeulen, and P. Suetens, "A viscous fluid model for multimodal non-rigid image registration using mutual information," *Med. Image Anal.*, vol. 7, no. 4, pp. 565–575, Dec. 2003.

[21] Y. Ou, A. Sotiras, N. Paragios, and C. Davatzikos, "DRAMMS: Deformable registration via attribute matching and mutual-saliency weighting," *Med. Image Anal.*, vol. 15, no. 4, pp. 622–639, Aug. 2011.

[22] H. Livyatan, Z. Yaniv, and L. Joskowicz, "Gradient-based 2-D/3-D rigid registration of fluoroscopic X-ray to CT," *IEEE Trans. Med. Imag.*, vol. 22, no. 11, pp. 1395–1406, Nov. 2003.

[23] B. Vos, F. Berendsen, M. Viergever, M. Staring, and I. Isgum, "End-to-end unsupervised deformable image registration with a convolutional neural network," in *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*. Cham, Switzerland: Springer, 2017, pp. 204–212.

[24] B. D. de Vos, F. F. Berendsen, M. A. Viergever, H. Sokooti, M. Staring, and I. Išgum, "A deep learning framework for unsupervised affine and deformable image registration," *Med. Image Anal.*, vol. 52, pp. 128–143, Feb. 2019.

[25] G Wu, M Kim, Q Wang, Y Gao, S Liao, and D. Shen, "Unsupervised deep feature learning for deformable registration of MR brain images," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.*, 2013, pp. 649–656.

[26] M. Simonovsky, B. Gutiérrez-Becker, D. Mateus, N. Navab, and N. Komodakis, "A deep metric for multimodal registration," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.*, 2016, pp. 10–18.

[27] W. P. Risk, G. S. Kino, and H. J. Shaw, "Fiber-optic frequency shifter using a surface acoustic wave incident at an oblique angle," *Opt. Lett.*, vol. 11, no. 2, pp. 115–117, Feb. 1986.

[28] X. Yang, R. Kwitt, M. Styner, and M. Niethammer, "Quicksilver: Fast predictive image registration—A deep learning approach," *NeuroImage*, vol. 158, pp. 378–396, 2017.

[29] T. Mok and A. Chung, "Large deformation diffeomorphic image registration with Laplacian pyramid networks," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.*, 2020, pp. 211–221.

[30] X. Cheng, L. Zhang, and Y. Zheng, "Deep similarity learning for multi-modal medical images," *Comput. Methods Biomech. Biomed. Eng., Imag. Vis.*, vol. 6, no. 3, pp. 248–252, May 2018.

[31] G. Balakrishnan, A. Zhao, M. R. Sabuncu, J. Guttag, and A. V. Dalca, "VoxelMorph: A learning framework for deformable medical image registration," *IEEE Trans. Med. Imag.*, vol. 38, no. 8, pp. 1788–1800, Aug. 2019.

[32] W. Shao, L. Banh, C. A. Kunder, R. E. Fan, S. J. C. Soerensen, J. B. Wang, N. C. Teslovich, N. Madhuripan, A. Jawahar, P. Ghanouni, J. D. Brooks, G. A. Sonn, and M. Rusu, "ProsRegNet: A deep learning framework for registration of MRI and histopathology images of the prostate," *Med. Image Anal.*, vol. 68, Feb. 2021, Art. no. 101919.

[33] American Cancer Society. *About Colorectal Cancer*. Accessed: Sep. 22, 2021. [Online]. Available: https://www.cancer.org/content/dam/CRC/PDF/Public/8604.00.pdf

[34] N. Horvat, C. C. T. Rocha, B. C. Oliveira, I. Petkovska, and M. J. Gollub, "MRI of rectal cancer: Tumor staging, imaging techniques, and management," *RadioGraphics*, vol. 39, no. 2, pp. 367–387, Mar. 2019.

[35] R. Sauer, H. Becker, W. Hohenberger, C. Rödel, C. Wittekind, R. Fietkau, P. Martus, J. Tschmelitsch, E. Hager, C. F. Hess, and J. H. Karstens, "Preoperative versus postoperative chemoradiotherapy for rectal cancer," *New England J. Med.*, vol. 351, pp. 1731–1740, Oct. 2004.

[36] J. E. Krook, C. G. Moertel, L. L. Gunderson, H. S. Wieand, R. T. Collins, R. W. Beart, T. P. Kubista, M. A. Poon, W. C. Meyers, J. A. Mailliard, D. I. Twito, R. F. Morton, M. H. Veeder, T. E. Witzig, S. Cha, and S. C. Vidyarthi, "Effective surgical adjuvant therapy for high-risk rectal carcinoma," *New England J. Med.*, vol. 324, no. 11, pp. 709–715, Mar. 1991.

[37] R. J. Heald and R. D. H. Ryall, "Recurrence and survival after total mesorectal excision for rectal cancer," *Lancet*, vol. 327, no. 8496, pp. 1479–1482, Jun. 1986.

[38] S. Rosati, C. M. Gianfreda, G. Balestra, V. Giannini, S. Mazzetti, and D. Regge, "Radiomics to predict response to neoadjuvant chemotherapy in rectal cancer: Influence of simultaneous feature selection and classifier optimization," in *Proc. IEEE Life Sci. Conf. (LSC)*, Oct. 2018, pp. 65–68.

[39] G. Simpson, P. Hopley, J. Wilson, N. Day, A. Haworth, A. Montazeri, D. Smith, L. Titu, J. Anderson, D. Agbamu, and C. Walsh, "Long-term outcomes of real world 'watch and wait' data for rectal cancer after neoadjuvant chemoradiotherapy," *Colorectal Disease*, vol. 22, no. 11, pp. 1568–1576, Nov. 2020.

[40] R. Cianci, G. Cristel, A. Agostini, R. Ambrosini, L. Calistri, G. Petralia, and S. Colagrande, "MRI for rectal cancer primary staging and restaging after neoadjuvant chemoradiation therapy: How to do it during daily clinical practice," *Eur. J. Radiol.*, vol. 131, Oct. 2020, Art. no. 109238.

[41] H. Chen, L. Shi, K. Nguyen, A. Monjazeb, K. Matsukuma, T. Loehfelm, H. Huang, J. Qiu, and Y. Rong, "MRI radiomics for prediction of tumor response and downstaging in rectal cancer patients after preoperative chemoradiation," *Adv. Radiat. Oncol.*, vol. 5, no. 6, pp. 1286–1295, 2020.

[42] N. Seo, H. Kim, M. S. Cho, and J. S. Lim, "Response assessment with MRI after chemoradiotherapy in rectal cancer: Current evidences," *Korean J. Radiol.*, vol. 20, pp. 1003–1018, Jul. 2019.

[43] H. Öztürk, "PET/MRI: The future of cancer restaging," *Cancer Treatment Res. Commun.*, vol. 25, Jan. 2020, Art. no. 100250.

[44] J. S. Davids, K. Alavi, J. Andres Cervera-Servin, C. S. Choi, P. R. Sturrock, W. B. Sweeney, and J. A. Maykel, "Routine preoperative restaging CTs after neoadjuvant chemoradiation for locally advanced rectal cancer are low yield: A retrospective case study," *Int. J. Surgery*, vol. 12, no. 12, pp. 1295–1299, Dec. 2014.

[45] W. Cai and G. Hong, "Quantitative image analysis for evaluation of tumor response in clinical oncology," *Chronic Diseases Transl. Med.*, vol. 4, no. 1, pp. 18–28, Mar. 2018.

[46] National Comprehensive Cancer Network. *NCCN Guidelines for Patients Rectal Cancer.* Accessed: Sep. 22, 2021. [Online]. Available: https://www.nccn.org/patients/guidelines/content/PDF/rectal-patient.pdf

[47] J. Ferlay, I. Soerjomataram, F. Lyon, and F. Lyon, "Global cancer statistic 2020: GLOBOCAL estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA, A Cancer J. Clinicians.* vol. 71, pp. 209–249, Jun. 2021.

[48] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens, "Multimodality image registration by maximization of mutual information," *IEEE Trans. Med. Imag.*, vol. 16, no. 2, pp. 187–198, Apr. 1997.

[49] E. Ardizzone, R. Gallea, O. Gambino, and R. Pirrone, "Multi-modal image registration using fuzzy kernel regression," in *Proc. 16th IEEE Int. Conf. Image Process. (ICIP)*, Nov. 2009, pp. 193–196.

CHI-JUI HO was born in Taipei, Taiwan, in 1996. He received the bachelor's degree in electrical engineering from the National Taiwan University, Taipei, in 2019. He is currently pursuing the Ph.D. degree in electrical and computer engineering with the University of California at San Diego, San Diego, CA, USA. He was a Summer Intern with MediaTek, Hsinchu, Taiwan, in 2018. His research interests include computational imaging and medical image registration.

SOAN T. M. DUONG received the B.Eng. degree in information technology from Le Quy Don Technical University, Vietnam, in 2010, the M.Eng. degree in computer science from Dongguk University, South Korea, in 2014, and the Ph.D. degree in computer engineering from the University of Wollongong, Australia, in 2020. Her research interests include image processing, medical image registration, medical image processing, machine learning, and neural architecture search. She was awarded the Examiners' Commendation for Outstanding Thesis in 2020.

YIQIAN WANG received the B.S. degree in electrical engineering from the Beijing Institute of Technology, Beijing, China, in 2018. She is currently pursuing the Ph.D. degree with the Electrical and Computer Engineering Department, University of California at San Diego, San Diego, CA, USA. Her research interests include medical image processing, signal processing, and deep learning.

CHANH D. TR. NGUYEN received the Ph.D. degree in mechanical engineering from the Korea Advanced Institute of Science and Technology (KAIST), in 2018. He is currently the Head of the Computer Vision Department, VinBrain JSC, Vietnam. His current research interest includes medical image analysis.

BIEU Q. BUI received the M.D. and Ph.D. degrees. He is currently the Vice Director of the Cancer Center and the Head of the Department of Radiation Oncology and Radiosurgery, 108 Military Central Hospital. He is an expert in radiation oncology field. He is specialized in advanced gastrointestinal, liver and lung cancers, stereotactic radiosurgery, stereotactic body radiation therapy, and PET/CT-radiotherapy planning. He is a Vietnam NPC of RAS/6085 and RAS/6096 projects involved radiation therapy with the International Atomic Energy Agency. He has published several articles in *Asia Oceania Journal of Nuclear Medicine and Biology*, *European Journal of Nuclear Medicine and Molecular Imaging*, *Hepatoma Research*, *International Journal of Hepatology*, and *Nuclear Medicine and Molecular Imaging*.

STEVEN Q. H. TRUONG received the B.Sc. degree in engineering science from the University of Toronto, Canada, in 1989, and the M.B.A. degree in information technology from the University of Phoenix, USA, in 1997. He has been with the AI and software industry for more than 26 years and has held senior leadership positions in top technology companies in USA and Canada, including Microsoft, Honeywell, and IntelliCommunities. He is currently the Founder and the CEO of VinBrain JSC, which is an AI healthcare company and a subsidiary of Vingroup, the largest conglomerate in Vietnam. His research interests include developing AI healthcare, smarthome, and smartcity systems.

TRUONG Q. NGUYEN (Fellow, IEEE) is currently a Distinguished Professor at the ECE Department, UCSD. He is the coauthor (with Prof. Gilbert Strang) of a popular textbook *Wavelets and Filter Banks* (Wellesley-Cambridge Press, 1997). He has over 450 publications. His current research interests include 3D video processing, machine learning with applications in health monitoring/analysis, and 3D modeling. He received the IEEE TRANSACTIONS ON SIGNAL PROCESSING Paper Award in 1992 and the NSF Career Award in 1995. He served as an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING, IEEE SIGNAL PROCESSING LETTERS, IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS—I: REGULAR PAPERS, and IEEE TRANSACTIONS ON IMAGE PROCESSING.

CHEOLHONG AN received the B.S. and M.S. degrees in electrical engineering from Pusan National University, Busan, South Korea, in 1996 and 1998, respectively, and the Ph.D. degree in electrical and computer engineering, in 2008. He is currently an Assistant Adjunct Professor at the Department of Electrical and Computer Engineering, University of California at San Diego, San Diego, CA, USA. Earlier, he worked at Samsung Electronics, South Korea; and Qualcomm, USA. His current research is focused on the medical image processing and the real-time bio image processing. His research interests include 2D and 3D image processing with machine learning and sensor technology.

• • •