Progress in
Medical Devices

Zentime

REVIEW ARTICLE

Deep learning for prostate intervention: Recent advances in non-rigid magnetic resonance imaging–transrectal ultrasound image registration

Peiyu Chen, Xudong Guo

School of Health Science and Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China.

Corresponding author: Xudong Guo.

Address correspondence to: Xudong Guo, School of Health Science and Engineering, University of Shanghai for Science and Technology, No. 516 Jungong Road, Yangpu District, Shanghai 200093, China.
E-mail: guoxd@usst.edu.cn.

Received December 29, 2025; Accepted March 6, 2026; Published June 26, 2026

DOI: 10.61189/692164snwggk

Abstract

The treatment of prostate cancer (PCa) is shifting towards the use of highly accurate image-guided procedures in order to achieve better oncologic results. A common strategy consists of using both pre-operative multiparametric magnetic resonance imaging and intra-operative transrectal ultrasound during a prostate biopsy or focal ablation procedure, thus offering enhanced localisation information through high spatial resolution and dynamic response, respectively. However, reliable non-rigid registration is still technically challenging owing to differences in cross-modal imaging physics as well as large deformations between the two modalities caused by rectal probe compression; this paper reviews how deep learning has evolved, focusing on convolutional neural networks, generative models, including generative adversarial networks and diffusion models, and transformer-based architectures. We discuss the extent to which they utilise biomechanical priors to inform the solution of registration problems, against standardized challenges such as μ-RegPro. State-of-the-art approaches achieve sub-millimetre target registration errors with real-time inference times for intra-operative deployment. Addressing outstanding challenges related to interpretation and generalization, this review provides an outlook of the road map to develop "physics-aware" smart interventional systems. All these developments represent important steps toward a fully automated, precise, and minimally invasive PCa management pipeline.

Keywords: Deep learning, Image registration, Non-rigid, Transrectal ultrasound, Magnetic resonance imaging

Highlights

- This review systematically reviews the evolution of deep learning-based non-rigid prostate magnetic resonance imaging–transrectal ultrasound registration.
- This review analyzes dominant paradigms: hybrid convolutional neural networks, generative adversarial networks/diffusion models, and transformers.
- This review explores integrating anatomical priors and physical constraints to address label scarcity.
- This review critically evaluates the generalization gap between state-of-the-art benchmarks and clinical workflows.
- This review proposes future directions in physics-aware artificial intelligence and intelligent robotic interventions.

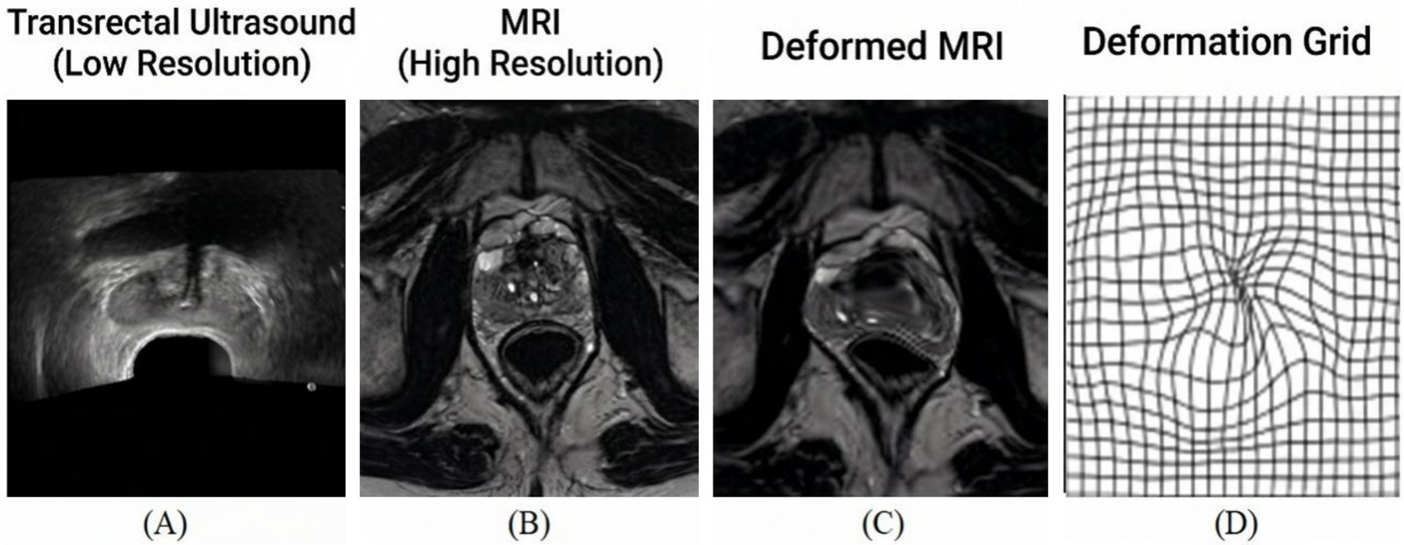

Figure 1. Visualization of core challenges in cross-modal registration. (A) Intraoperative TRUS: Target ultrasound image characterized by low contrast, speckle noise, and a limited FOV, highlighting the significant modality gap; (B) Preoperative MRI: High-resolution T2-weighted image showing clear anatomical details; (C) Warped MRI: MRI image after non-rigid spatial transformation, aligned to match the prostate anatomy under probe compression; (D) Deformation Grid: Visualization of the dense displacement field; grid distortion reflects the non-linear spatial mapping. MRI, magnetic resonance imaging; TRUS, transrectal ultrasound; FOV, field of view; T2, transverse relaxation time 2.

1 INTRODUCTION

Prostate cancer (PCa) remains one of the most prevalent malignancies among men worldwide, ranking as a leading cause of both cancer incidence and mortality [1]. Based on Global Cancer Observatory statistics for the year 2022, more than one and a half million new cases of cancer were diagnosed worldwide, leading to about 400,000 deaths [2]. Early diagnosis and accurate treatment are crucial to improve the five-year survival rate and quality of life in PCa patients. Currently, clinical diagnosis relies heavily on systematic biopsy guided by transrectal ultrasound (TRUS). However, conventional TRUS-guided biopsy is hindered by two intrinsic limitations [3, 4]. First, as illustrated in **Figure 1A**, TRUS is inherently plagued by inferior image quality characteristics such as a low signal-to-noise ratio, speckle noise, and poor soft-tissue contrast, which obscure tumour borders and hamper the differentiation of malignant lesions from benign hyperplasia. Second, the process is highly dependent on the operator, and the accurate placement of a needle depends on the clinician's experience and hand steadiness. All these factors lead to a high false-negative rate, with a risk of missing clinically significant prostate cancer (csPCa), resulting in a delay in care. In contrast, multiparametric magnetic resonance imaging (MRI) offers superior soft-tissue resolution and multi-sequence imaging capabilities (**Figure 1B**), which enable accurate detection, localization, and risk stratification of csPCa. However, MRI is expensive and time-consuming, and it cannot guide the procedure in real time. Consequently, preoperative MRI-TRUS image fusion-targeted biopsy has been proposed. This technology aligns high-precision preoperative three-dimensional (3D) MRI with real-time intraoperative two-dimensional (2D)/3D TRUS, "projecting" MRI-identified suspicious lesions onto the TRUS space to provide clinicians with precise navigation, thereby significantly enhancing the diagnostic yield and detection rate of biopsy [5, 6].

In such a clinical workflow, image registration is the key technology to realize accurate navigation. Image registration refers to the task of searching for a spatial transformation that can map images of either the same subject or different subjects into each other. MRI-TRUS registration is a classic cross-modal problem fraught with severe technical hurdles. First, the disparate imaging physics of MRI and TRUS result in vast differences in intensity distributions, texture features, and tissue contrast. This makes conventional intensity-based similarity measures—such as normalized mutual information, mean squared error, and modality-independent neighborhood descriptor—struggle to establish effective spatial correspondences [7-10]. Second, the prostate is a compliant, deformable organ; intraoperative probe pressure, changes in patient positioning, and respiratory motion cause significant non-rigid deformation [11]. This physical shift renders preoperative anatomy inconsistent with intraoperative observations. Registration algorithms must therefore apply complex spatial transformations to the MRI to fit the TRUS geometry (**Figure 1C**). This spatial mapping is highly non-linear and non-uniform, as visualized by the dense deformation grid in **Figure 1D**. Moreover, the low signal-to-noise ratio of TRUS, acoustic artifacts, as well as the difference in spatial resolution and field of view among modalities add extra hurdles for obtaining an exact voxel-wise correspondence. Conventional non-rigid registration approach-

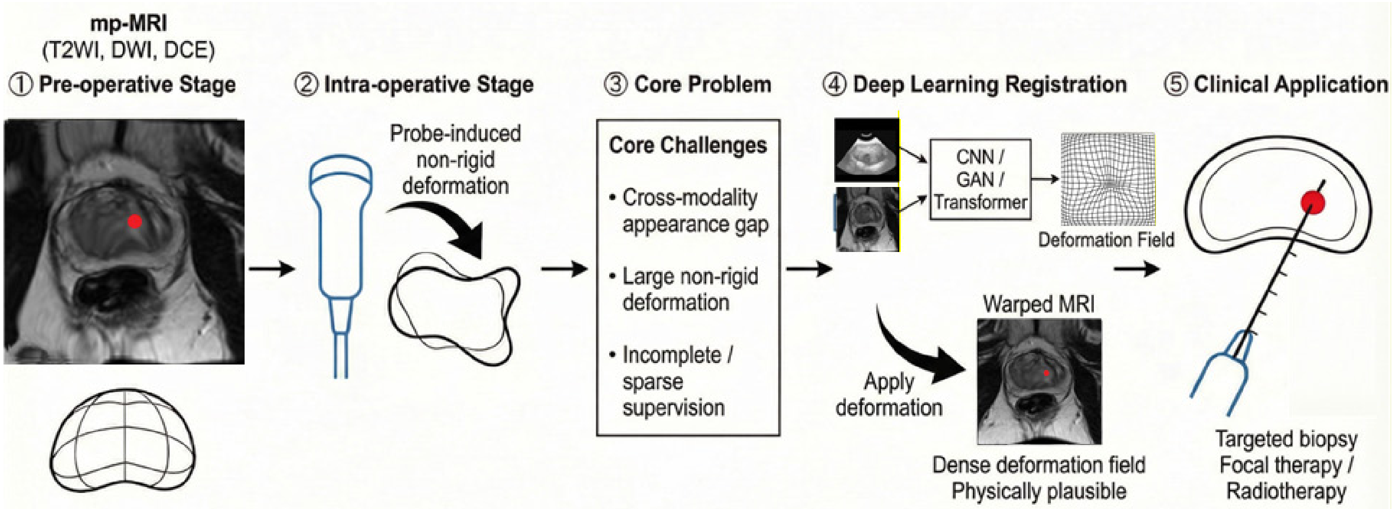

Figure 2. Overview of the DL-based prostate MRI-TRUS registration clinical workflow and technical framework. The process comprises five stages: ① Preoperative acquisition of mp-MRI to localize tumors; ② Intraoperative TRUS imaging, during which non-rigid deformation occurs due to probe pressure; ③ Identification of core registration challenges; ④ DL registration module (using CNN, GAN, or Transformer backbones) to predict dense deformation fields; ⑤ Final clinical application, providing precision navigation for targeted biopsy or focal therapy. DL, deep learning; mp-MRI, multiparametric magnetic resonance imaging; MRI, magnetic resonance imaging; TRUS, transrectal ultrasound; CNN, convolutional neural network; GAN, generative adversarial network; T2WI, T2-weighted imaging; DWI, diffusion-weighted imaging; DCE, dynamic contrast-enhanced imaging.

es are typically classified into spline-based methods (e.g., B-spline Free-Form Deformation and Thin-Plate Splines) and physics-based models [12, 13]. However, both paradigms have significant disadvantages. Spline-based approaches are often unable to model very local and non-uniform deformation as they assume a global smoothness constraint. Physics-based models are biomechanically plausible, but often require tedious iterative optimisation of complicated energy functions, making them prone to local minima and unsuitable for the real-time requirements of clinical interventions [14, 15].

To address such issues, deep learning (DL), in particular convolutional neural network (CNN)-based architectures, has enabled a new era of learning-based medical image (MI) registration, which does not need to optimize for each specific pair: DL methods consider registration as an end-to-end (E2E) regression problem. By training deep neural networks with a large dataset, this network learns a complicated non-linear mapping from an input image pair to the output deformation field [16]. After learning, the network is able to predict displacement fields in less than a millisecond and therefore enables online registration for clinicians with better reliability. In recent years, researchers have also widely used DL in prostate MRI-TRUS registration with customised designs according to the particularities of this task. Among these, we can mention the use of CNNs and weakly-supervised approaches to extract features at different levels, adopting generative models, including generative adversarial networks (GANs) and diffusion models, to address the cross-modal discrepancy issue and large deformation generation, and incorporating transformers for modeling long-range context relations. These developments are moving prostate-directed biopsy away from "experience-dependent" cognitive fusion and closer to fully automatic, precise, and intelligent "software fusion".

The aim of this review is to provide a systematic summary and critical discussion on the status quo of DL-based non-rigid prostate MRI-TRUS registration. First, we describe the key technical paradigms and representative work in the domain. Second, we summarize the datasets and evaluation metrics for performance assessment and quantitative comparison. Third, we examine how to integrate DL technologies into a clinical surgical workflow and introduce the open problems for their clinical deployment. Last but not least, we present our perspectives regarding the directions of future work and clinical translation so as to serve as a useful guide to researchers and clinicians alike. We summarize the overall DL-based prostate MRI-TRUS registration clinical workflow and technical pipeline presented throughout this manuscript in **Figure 2**.

2 DL PARADIGMS FOR PROSTATE MRI-TRUS REGISTRATION

Over the past few years, DL has revolutionized prostate MRI-TRUS image registration and generated multiple types of approaches that tackle the problem from different angles. To better understand how each technology has evolved over time, in this subsection we divide the existing work into three categories according to what they are innovating: first, we present CNN-based approaches, which are the most established para-

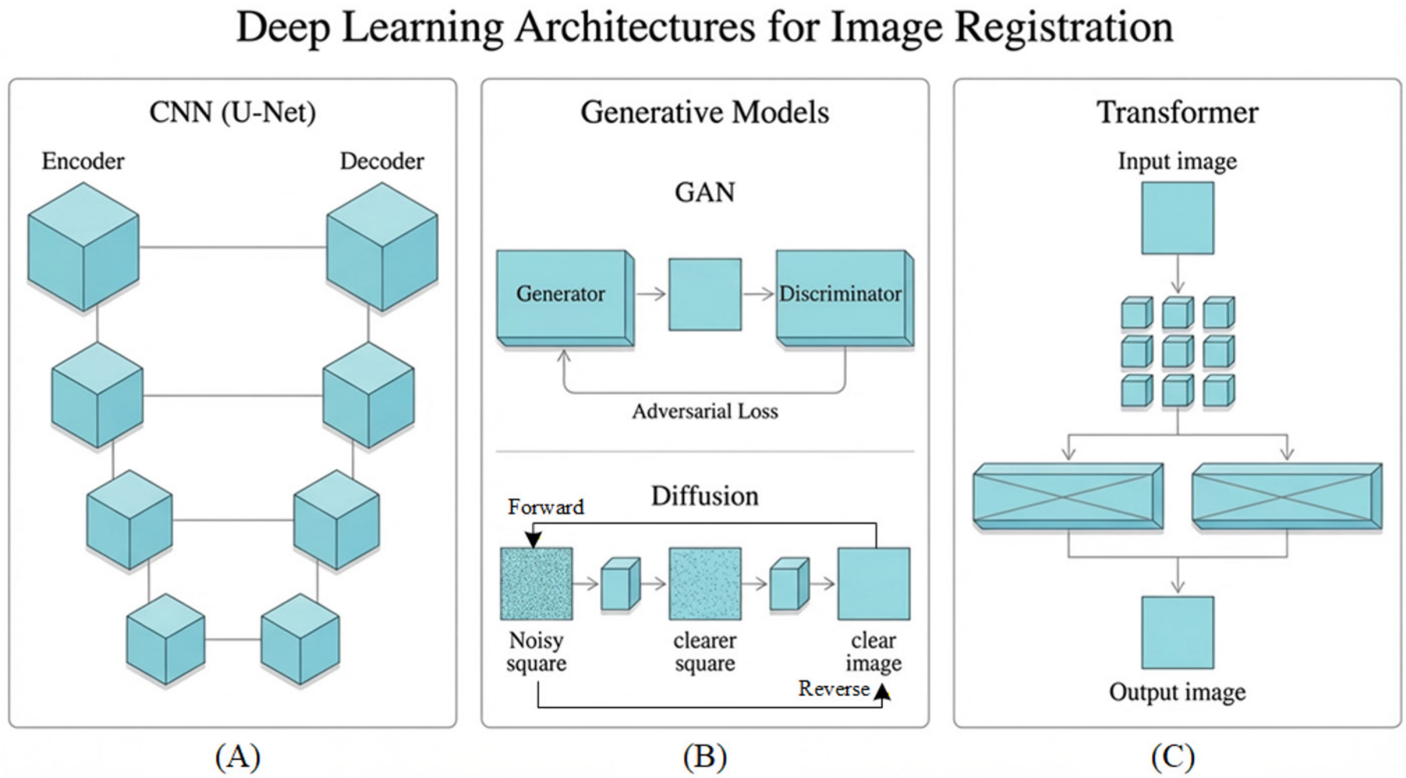

Figure 3. Schematic overview of deep learning architectures for prostate MRI-TRUS registration. (A) CNN Architecture: Represented by the U-Net backbone, utilizing encoder-decoder paths and skip connections for end-to-end regression; (B) Generative Models: Illustrating two distinct paradigms: GANs (top), which use adversarial training between a generator and discriminator to bridge modality gaps; and diffusion models (bottom), which employ an iterative denoising process (from noisy to clear) to generate high-fidelity deformation fields or images; (C) Transformer Architecture: Demonstrating the tokenization of input images into patches and the use of self-attention mechanisms to capture global anatomical dependencies. MRI, magnetic resonance imaging; TRUS, transrectal ultrasound; CNN, convolutional neural network; GAN, generative adversarial network.

digm, including advances in real-time architecture optimization and improved feature representation by incorporating geometry and physics priors (e.g., segmentation contours, biomechanical models). The second category is the generative model-based methods, including GANs that use adversarial learning to close the modality gap, and the more recent diffusion models, which are known for having strong distribution modeling capabilities. Lastly, we include attention mechanism and transformer-based models, where we explain how such an approach solves the local receptive field problem that comes with the convolution operation by adding a self-attention module or as a hybrid model. This architecture can therefore capture global, long-range dependencies as well as prominent anatomical features to perform this classification task. An illustration of the classification architecture used is shown in **Figure 3**.

2.1 CNN-based methods

As one of the first technically mature DL-driven image registration methods, CNN-based methods are designed based on a simple assumption that strong local feature representation capability and nonlinear transformation capacity can be utilized by modeling the image registration problem as an E2E regressor. Pioneering works, such as the framework that applied CNNs as similarity metrics for driving 3D ultrasound registration, achieved an error below 5 mm in 81% of the samples, proving that deep features have greater potential than conventional intensity-based measures or even manual alignment [17].

Later, popular approaches (e.g., VoxelMorph) established the trend of utilizing encoder-decoder networks for direct prediction of displacement fields [16]. Based on these works, researchers first worked on optimizing registration workflows and basic architectures for addressing the challenges of real-time performance and large deformations intrinsic to prostate interventions. For example, Dupuy et al. proposed an online 2D/3D registration approach based on CNNs, incorporating the registration result of the previous frame as a trajectory prior and substantially improving temporal consistency [11]. To address large deformation, Guo et al. proposed the multi-stage registration architecture with a “coarse-to-fine” approach and a synthesized error-scaled data generation method to address the missing coverage of large-deformation labels in training, reducing the surface registration error to 3.57 mm [18]. In a subsequent study, Guo et al. proposed a deep learning-based ultrasound frame-to-volume registration pipeline for registering 2D ultrasound frames to 3D volumes [19]. With the integration of a frame-to-frame registration network and a slice-correction network, and a similarity filtering mechanism, the pipeline was able to achieve an approximate target navigation error of 1.93 mm at a clinical-grade frame rate of 5–14 frames per second, demonstrating both high accuracy and fast inference speed suitable for intraoperative use in an MRI-ultrasound fusion-guided biopsy system.

Besides improving the network structure and inference process, another important line of work is exploring more detailed anatomical features and prior knowledge in order to facilitate better inter-modality correspondence learning with extra semantic content provided by inputs or additional geometrical regularization imposed explicitly. One common approach is to enrich input semantics via segmentation probability maps. Zeng et al. proposed a hybrid supervision method that first uses two independent fully convolutional networks to produce prostate probability maps on MRI and TRUS, respectively, which are further fed to the registration network as auxiliary channels. By combining Dice loss and surface distance loss, they obtained a mean target registration error (TRE) value of 2.53 mm using this technique in 36 patients [20]. As research deepened, to elevate feature expression from a frequency-domain perspective, Jiang et al. proposed a method based on joint learning and a multi-level wavelet feature pyramid. Through a cyclic mutual-assistance mechanism between the segmentation and registration networks, they achieved superior accuracy across 642 public cases [21]. Additionally, to endow networks with physical interpretability, incorporating physical model constraints has become an important direction. Fu et al. proposed a biomechanically constrained registration method based on 3D point-cloud matching as a proxy task. They used deformation fields obtained from finite element analysis (FEA) as supervision, enabling them to encode these biomechanical constraints in the network and achieve high-precision MR-TRUS registration [22].

2.2 Generative model-based methods (adversarial networks and diffusion models)

A GAN is composed of two competitive networks called a generator and a discriminator. Given an image registration task, the generator usually predicts the deformation field or generates the warped image, whereas the discriminator discriminates the warped image from the ground-truth target image, compelling the Generator to generate more realistic and plausible results. Yan et al. first proposed such an idea for multimodal MI registration using the framework of adversarial image registration network. Using a Wasserstein GAN approach, the generator directly estimates rigid transformation parameters, while the discriminator is used to evaluate the alignment quality, achieving fast (<100 ms) and accurate (TRE 3.48 mm) registration for 763 cases, effectively validating the viability of adversarial learning [23].

To cope with such anatomical complexity, later works focused on maintaining prominent regional information as well as adding some physics-based restrictions. For example, Lian et al. presented an region of interest-guided deformable image-to-image translator, which converts one modality into another while enforcing consistent structural prior knowledge in order to generate plausible TRUS-like images, simplifying later registration challenges and preserving geometric truthfulness [24]. Likewise, Feng et al. proposed a salient region matching approach; while not a pure GAN structure, they integrate cross-modal spatial attention into a generative framework. They use a salient region matching loss to regularize the shape and intensity consistency between the region of interest of the prostate, and they obtained a Dice coefficient of 85.9%, thus improving the robustness of a purely automatic registration [25]. In addition, to address the problem of potential unphysical deformations in pure data-driven generation, we use a variational autoencoder as a component of our biomechanics-informed generative registration framework architecture, which encodes biomechanical priors (computed via FEM) onto a latent manifold. In registration, the network searches within such a manifold for an optimal deformation field fulfilling both image information and biomechanical constraints, effectively suppressing unreasonable local distortion [26].

Although successful for cross-modal registration with GANs, due to its unstable training process as well as the mode collapse issue, it is not powerful enough to model the complicated distribution of large deformations. Recently, diffusion probabilistic models have been introduced. Based on their good distribution coverage and stable training performance, they have been brought into the registration domain to break through such a bottleneck. Yao et al. conspicuously introduced diffusion models to the weakly supervised setting. They designed a label-aware diffusion model, which can correct the initial bias by aligning label centroids, followed by producing a high-quality deformation field through a feature-guided module with label supervision. It is highly robust for large deformations, substantially lowering the TRE to 0.940 mm, with an exceptionally small Jacobian determinant fold-over rate (0.134), and is much more precise than classical unsupervised or supervised registration techniques such as VoxelMorph [27].

2.3 Transformer-based registration methods

Although CNN-based registration approaches are highly efficient with good local texture extraction ability, they are based on a strong inductive bias—i.e., the set of structural assumptions that the model makes for generalization to new data. For CNNs, this bias mainly consists of locality and translation invariance, which means the model assumes the relevant fea-

Table 1. Characteristics of representative public datasets for prostate MRI-TRUS registration

Dataset	Scale	Raw resolution	Preprocessing and characteristics
μ-RegPro [32]	76 cases	MRI: 120×128×128 voxels TRUS: 81×118×88 voxels	Resampled to 0.8 mm^3 isotropic resolution
TCIA [33]	Prostate subset	T2WI: 0.5 mm/plane, 3.6 mm slice DWI: 2.0 mm/plane, 3.6 mm slice DCE: 1.5 mm/plane, 4.0 mm slice	Raw DICOM; Heterogeneous resolutions; Significant slice thickness

Note: MRI, magnetic resonance imaging; TRUS, transrectal ultrasound; TCIA, The Cancer Imaging Archive; T2WI, T2-weighted imaging; DWI, diffusion-weighted imaging; DCE, dynamic contrast-enhanced imaging; DICOM, Digital Imaging and Communications in Medicine.

ture is spatially clustered and its relative position remains unchanged. While this facilitates feature learning, it nevertheless results in the problem of the local receptive field, which brings great difficulties in modeling large-scale non-rigid deformation.

This limitation is especially important for prostate MRI-TRUS registration. In a clinical setting, inserting and manipulating the transrectal probe applies considerable pressure to the gland, inducing large-scale global deformations and non-uniform displacements of the whole organ, which have long-range spatial correlations, where the pressure on the rear end directly determines the shape and location of the front tip. Since the prostate is a soft tissue, all these global variations are distributed across the whole field of view. It is difficult for CNNs to learn such inherent correlations among far-away anatomical structures because of their nature as local window-based operations. In order to address this limitation, equipping them with an attention mechanism, i.e., transformers for global context extraction, is now a mainstream direction of development.

Originally having great success in natural language processing, transformers use their key component, self-attention, which is able to compute association weights between any two positions in an input sequence adaptively. Since vision transformers for unsupervised volumetric medical image registration was introduced into MI registration in 2021, the field quickly moved on, exploring shifts from pure global attention to hierarchical feature representations, such as hierarchical vision transformer [28, 29]. Nevertheless, the pure transformer network usually sacrifices matching precision for minute structures because there is no local texture information. Therefore, the "hybrid architectures" that combine the local feature extraction ability of CNN and the global screening ability of attention mechanisms are becoming mainstream choices.

For instance, Mahmoudi et al. proposed 3D-Wavelet-Deep-Separable-Attention-PMorph, which is a hybrid of Swin Transformer and CNN with an encoder-decoder architecture [30]. They introduced their main contribution as the Wavelet-Deep-Separable-Attention, which relies on wavelet transform to achieve multi-scale space-frequency fusion that overcomes the shortcomings of conventional transformer in representing frequency-domain information. Assuming standardized μ-RegPro dataset and pre-aligned conditions, this approach reached a TRE of less than one millimetre [30]. Parallel to this is another technical route focusing on clinical robustness and salient feature mining. Yu et al. proposed a weakly-supervised framework integrating an Attention-enhanced U-shaped CNN and a Residual-Enhanced Registration Network [31]. This method utilizes a sparse set of expert-annotated anatomical landmarks to compute loss functions, alleviating the annotation burden. In this work, we not only evaluated our algorithms with the publicly available μ-RegPro dataset, but also provided a patient cohort including Prostate Imaging Reporting and Data System scores as an additional test set [32]. Despite the fact that, due to its complexity, the mean TRE increased up to 8.63 mm compared to the public dataset, the study showed that weakly supervised DL can be used to assist biopsy navigation in realistic clinical workflow scenarios [31]. The above two recent works show that either learning the multi-scale frequency perception by transformer blocks or promoting salient feature representation with attention mechanisms, enhancing the model's perception of global topology and important anatomical structures, is the critical path to solving the challenging non-rigid prostate registration problem.

3 DATASETS AND EVALUATION METRICS

Consistent datasets and unified evaluation criteria are essential to objectively evaluate and compare different DL-based registration approaches. In this part, we list typical public resources and basic evaluation criteria used by the prostate MRI-TRUS registration community.

3.1 Public datasets

The size, variety and annotation quality of a dataset directly determine how well a DL model will perform and to what extent it can be generalized. Today, most studies on prostate MRI-TRUS registration are based on a few publicly available datasets and internal private datasets, as patient privacy strictly restricts the collection and distribution of high-quality, large-scale labelled datasets. A comparison between the two most representative public datasets and their peculiarities is provided in **Table 1**.

Datasets used in recent works are based on the following two approaches to algorithm verification. μ-RegPro is a very well-controlled testbed [32]. In this dataset, MRI and TRUS volumes are resampled to the same 0.8 mm^3 isotropic resolution. Although this strict standardization removes the impact of resolution disparity and improves the repeatability of registration outcomes, it can lose some of the finer anatomy that is present in raw data.

On the other hand, The Cancer Imaging Archive maintains the raw clinical state of the imaging data [33]. It comprises diagnostic-grade Digital Imaging and Communications in Medicine sequences with heterogeneous resolutions, with good in-plane resolution (e.g., 0.5 mm, T2-weighted imaging) but large slice thicknesses (up to 4.0 mm). The Cancer Imaging Archive requires stronger preprocessing and adaptation capabilities for the registration models but provides better ecological validity in that it reflects the natural heterogeneity and artifacts of clinical acquisitions.

3.2 Evaluation metrics

To evaluate registration performance in a quantitative manner, several widely used evaluation criteria are employed to measure registration quality based on different aspects, such as volume intersection ratio, surface point-to-surface distance, and inlier correspondences between intrinsic features.

TRE: TRE, which has been regarded as a "gold standard" to assess registration accuracy, quantifies the Euclidean distance between one fixed set of predefined, known corresponding points (landmarks) in the source and target images, which usually correspond to important anatomical landmarks, such as the urethra, prostate calcifications, or cysts. The TRE is defined as:

$$TRE = sqrt\left(\frac{1}{N} * \sum_{i=1}^{N} (\| \mathrm{p}_i^m - T(\mathrm{p}_i^s) \|^2)\right) \tag{1}$$

where p_i^m and p_i^s denote the coordinates of the i-th corresponding landmark in the target and source images, respectively, T represents the learned transformation, and N is the total number of points. A lower TRE indicates higher precision. In clinical practice, a TRE of less than 2.0–3.0 mm is generally considered acceptable for targeted interventions.

Dice similarity coefficient (DSC): DSC measures the volume overlap between two binary masks (e.g., organ segmentations). It has values ranging from 0 to 1, where 1 represents an exact match. This is used to measure the similarity between the warped source mask and the target mask in a registration problem:

$$DSC = 2 * \frac{|\mathrm{A} \cap \mathrm{B}|}{(|\mathrm{A}| + |\mathrm{B}|)} \tag{2}$$

where A and B represent the source and target masks. DSC is a highly effective metric for assessing the overall global alignment of an organ.

Hausdorff distance (HD): HD measures the maximum mismatch between two surfaces:

$$h(\mathrm{A}, \mathrm{B}) = \min_{\{a \in A\}} \min_{\{b \in B\}} \| \mathrm{a} - \mathrm{b} \| \tag{3}$$

$$HD = \max\left(h(\mathrm{A}, \mathrm{B}), h(\mathrm{B}, \mathrm{A})\right) \tag{4}$$

To reduce the impact of noise and outlier points, the 95th percentile HD is usually reported in the literature instead. It picks the maximal distance after removing the top 5 percent of the largest distances and serves as a strong indicator of local misalignment at the boundary.

Mean surface distance (MSD): MSD measures the mean distance between every pair of corresponding points on each surface. Compared with HD, MSD is not as sensitive to local outliers and thus better reflects the overall surface matching quality.

Table 2 outlines the performance of state-of-the-art (SOTA) methods, categorized by their respective evaluation datasets. While direct numerical benchmarking is hindered by varying normalization protocols and dataset scales, these metrics offer a comprehensive trajectory of the field's technical maturation and its ongoing shift toward clinical integration.

While the different origins of data, as well as the lack of standardization for data preprocessing, prevent us from making an absolute quantitative comparison among them, horizontally, the quantitative metrics summarized in **Table 2** highlight a clear technical evolution, as well as the present-day bottlenecks in translating these advances into clinical practice. We also note, however, that performance can only ever be as good as the dataset on which it has been tested, with "best-case scenario" performance typically being obtained when testing on already aligned or curated datasets. Performance remains very limited on raw clinical data.

The first is that developments in model architecture have directly led to a decade's worth of orders-of-magnitude improvements in registration accuracy. Initial CNN and GAN-based approaches (e.g., Zhu et al.; Yan et al.), which were restricted to the local receptive field and simple adversarial strategies, have TREs in a range of approximately 3.0–5.0 mm most of the time [17, 23]. By comparison, works in 2025 using transformers or diffusion models manage to lower the TRE below one millimeter (<1.0 mm) [30]. Note that those SOTA results mainly report on very standard benchmarks, where spatial pre-alignment reduces the initial pose variance. This indicates that, although modeling of complex deformations is important, the quality of the input data is one of the most important factors that determines the degree of reported success.

Table 2. Summary of performance metrics for representative DL-based prostate registration methods

Study	Method	TRE (mm)	DSC	HD95 (mm)	MSD (mm)	Dataset sources
Yan et al. 2018 [23]	Adversarial (AIR-net)	3.480	-	-	-	Private
Zhu et al. 2019 [17]	CNN-based similarity	<5.000 (in 81% cases)	-	-	-	Private
Zeng et al. 2020 [20]	Label-driven weakly supervised	2.530±1.390	0.910±0.020	4.410	0.880	Private
Fu et al. 2021 [22]	Biomechanically constrained	1.570±0.770	0.940±0.020	2.960±1.000	0.900±0.230	Private
Guo et al. 2023 [19]	FVReg (CNN)	1.930	-	-	-	Private
Lian et al. 2025 [24]	Deformation-aware GAN	2.240±1.020	0.876	-	-	Private
Feng et al. 2025 [25]	Salient region match	4.650±1.760	0.859±0.035	-	-	μ-RegPro
Mahmoudi et al. 2025 [30]	3D-WDA-PMorph	0.850	0.941	-	-	TCIA
Jiang et al. 2025 [21]	Joint learning (MWFP)	-	0.811±0.018	-	-	TCIA
Yu et al. 2025 [31]	Weakly supervised (RERN)	7.860 (Pub); 8.630 (Clin)	0.748 (Pub); 0.730 (Clin)	10.180 (Pub); 11.180 (Clin)	-	μ-RegPro
Yao et al. 2025 [27]	Label-aware diffusion	0.940	0.880	-	-	μ-RegPro
Hao et al. 2026 [26]	BGProReg (Phy-GAN)	3.720±1.050	0.924±0.027	3.000±0.700	-	μ-RegPro and TCIA

Note: TRE, target registration error; DSC, Dice similarity coefficient; HD95, 95th percentile Hausdorff distance; MSD, mean surface distance; DL, deep learning; AIR-net, adversarial image registration network; CNN, convolutional neural network; FVReg, frame-to-volume registration; GAN, generative adversarial network; MWFP, multi-level wavelet feature pyramid; RERN, residual-enhanced registration network; BGProReg, biomechanics-informed generative registration framework; Phy-GAN, physics-informed generative adversarial network; 3D-WDA-PMorph, three-dimensional wavelet-deep-separable-attention PMorph; TCIA, The Cancer Imaging Archive; Pub, public dataset; Clin, clinical dataset.

Second, physics-based constraints are important for achieving plausible deformation, although purely data-driven approaches typically achieve lower TRE: the absence of anatomical constraints leads to non-physical topology changes, as shown in works by Fu et al. and Hao et al. [22, 26]. While the TREs (approximately 1.57–3.72 mm) are still slightly higher than those from recent pure DL methods, the use of biomechanical constraints or FEM priors can prevent topological errors, e.g., Jacobian determinant folding. It stresses that forcing a physically realistic deformation field is at the core of ensuring safety for clinical deployment, even with low errors.

Lastly, there is still a large performance gap between controlled experiments and the true clinical setting. As shown in the work of Yu et al., the same algorithm performed significantly worse on more heterogeneous clinical data than it did on very well-controlled public datasets, with an increase in TRE from 7.86 mm to 8.63 mm [31]. This indicates that the good performance obtained by SOTA methods when operating on preprocessed data could be an overestimate of their performance in realistic settings, and future efforts should not focus solely on improving performance on a reference dataset but should also attempt to take into account domain shift, multimodal interference, e.g., artifacts, and multicenter pathological variation to improve the clinical robustness of such models.

4 CLINICAL APPLICATIONS AND CHALLENGES

DL-based MRI-TRUS registration is not only a matter of developing algorithms, but it is also useful when incorporated into clinical diagnosis and treatment processes. It could improve the accuracy and safety of prostate biopsy and other interventions: these tools are changing the paradigm of treatment. Here we highlight some fundamental applications to clinical practice, and discuss various factors that make their clinical adoption difficult.

4.1 Clinical application scenarios

At present, MRI-TRUS fusion-targeted biopsy is the most mature and popular application scenario for fusion technology. Conventional systematic biopsy is usually restricted by the poor specificity of ultrasound, which results in "blind sampling" and a high omission rate of csPCa [34-36]. Fusion-guided systems instead map high-risk regions identified on MRI in real time, e.g., Prostate Imaging Reporting and Data System 4/5 lesions, onto intra-operative TRUS images to guide a precise trajectory of the needle, thereby improving the diagnostic yield of csPCa with a minimal number of cores and related complications [37]. For example, the PROST-Net model implemented on the PROST robotic platform to perform transperineal biopsy with real-time prostate segmentation and tracking, thus minimizing errors due to subjective human factors [38]. Wang et al. developed an E2E pipeline including a preprocessing module, the DeepReg registration network, and a live planning module, which generates both trajectory suggestions and "collision warnings" between the organs at risk closest to each other, showing that it is possible to have completely intelligent surgical assistance [39].

Precise registration technology is a key support for achieving focal therapy, such as high-intensity focused ultrasound or cryoablation [40, 41]. For both treatments, the accurate delivery of energy to the preoperative MRI-targeted plan and real-time MRI-TRUS registration facilitate accurate execution of the treatment plan as well as an adaptive response to tissue deformation due to thermal ablation or ice ball growth [40, 42]. Additionally, for minimally invasive procedures such as brachytherapy, registering the pre-operative MRI and intra-operative ultrasound improves targeting accuracy while sparing organs at risk to decrease morbidity [43]. In order to tackle registration challenges due to the presence of metallic needle artifacts on ultrasound, researchers have suggested weakly supervised approaches for propagating the MRI-based prostate contours onto the treatment-planning ultrasound image, enabling efficient dose allocation and analysis [44].

4.2 Technical and clinical translation challenges

Despite these promising results in isolated experiments, there are still some challenges that need to be addressed before applying a DL model as an evidence-based medical tool in real-life conditions [45].

4.2.1 Generalization and data bottlenecks

However, deep models, especially complicated CNNs or transformers, are prone to highly fitting the particular distribution in the training set, which results in a severe performance drop when facing the heterogeneity of data from different centers, as the data are often unstandardized across device manufacturers or imaging protocols [46]. Additionally, due to strict patient privacy rules, e.g., General Data Protection Regulation and Health Insurance Portability and Accountability Act, and the prohibitive cost of expert annotation, most research is based on small-scale and single-center datasets. Large-scale standardized multicenter data are one of the major limitations for demonstrating generalizability and reaching the maximum achievable performance.

4.2.2 Technical robustness and interpretability in complex scenarios

Although registration error is reported at a millimeter scale, the robustness of algorithms to inevitable "clinical noise", e.g., bowel air-gas artifacts, non-uniform probe pressure, or extreme tissue distortion, remains to be rigorously stress-tested [47]. Furthermore, due to the "black-box" nature of DL, the decision-making process is opaque: if a clinician cannot interpret the registration logic, e.g., which landmarks have been matched, they will not necessarily believe in the outcome of the system. Therefore, designing explainable artificial intelligence (AI) to explain how the model reaches its conclusions can be necessary before building confidence between engineers and clinicians [48].

4.2.3 Workflow integration and regulatory hurdles

A successful algorithm will need to seamlessly integrate into existing hospital information systems, i.e., picture archiving and communication systems/radiology information systems, and match the user interface/user experience preferences of surgical teams. As medical devices, such systems have to be approved by regulatory authorities such as the Food and Drug Administration or National Medical Products Administration, and there must be an built-in fault detection and error handling within the system itself. When the model's confidence is low or an anomaly has been detected, it should warn and hand over control back to the doctor, adhering to the "human-in-the-loop" paradigm as a baseline for surgical safety.

5 FUTURE DIRECTIONS AND PERSPECTIVES

To solve the problems mentioned above, high-precision prostate MRI-TRUS registration technology is developing toward robustness and deep clinical integration. With leading-edge technical trends as well as pressing clinical needs, we envision that future studies will focus on the following five main aspects:

The first is the development of a new generation of hybrid architectures as well as the exploration of more generative frameworks in future registration models. The key point of future registration models should be how to combine global information with local information from different scales. On one hand, transformer-based networks can effectively capture long-range dependencies, which have obvious advantages when dealing with the estimation of large-scale non-rigid deformation. On the other hand, it is believed that hybrid "transformer+CNN"-based networks would be responsible for keeping the global topological structure consistent while enforcing local pixel-wise correspondence. Furthermore, recently emerging generative networks, especially diffusion models, have shown remarkable performance in distribution modeling. Applying them to the task of cross-modal image synthesis or utilizing them as priors for deformation field generation can outperform conventional similarity measures, contributing new mathematical tools to the field of unsupervised registration.

Second, future studies should focus on learning from limited resources. In order to avoid requiring massive amounts of expert supervision, few-shot learning and domain adaptation would be another major focus in future studies. A meta-learning or metric-learning framework could help improve the rapid adaptation capability when facing data from novel clinical centers or imaging devices. Furthermore, exploring efficient transfer learning schemes, where the anatomical representation learned in well-populated domains, e.g., mono-modal MRI registration, is distilled to data-scarce MRI-TRUS tasks, is one such possible route toward alleviating the annotation bottleneck problem.

Third, future studies should focus on deep mining of multi-modal radiomics. The existing methods mainly use single-channel intensity or a binary mask. The paradigm is changing toward multi-parameter and multidimensional information fusion. Fusing multi-sequence MRI information, e.g., T2-weighted imaging, diffusion-weighted imaging, and dynamic contrast-enhanced imaging, and possibly including the metabolic functional signature of prostate-specific membrane antigen positron emission tomography/computed tomography, may be useful. Researchers may build high-dimensional radiomics feature spaces. These types of multi-modal constraints will help disambiguate the registration problem where there is ambiguity due to, for example, low contrast between tissues and/or similar grey values across modalities, thus improving the stability of soft tissue matching.

Fourth, future studies should focus on the incorporation of physical laws and biomechanical constraints. Purely data-driven models tend to suffer from non-physical deformations, e.g., tissue folding. The integration of physics priors is critical for guaranteeing a unique and safe solution. Future work will focus more on explicitly incorporating diffeomorphic mapping, volume preservation, and biomechanical consistency constraints in terms of losses or networks. This guarantees that the learned deformation fields are also physically invertible, respecting the hyperelastic behavior of prostate tissues. FEA-based strategies using FEA results to guide latent space learning would be important baselines in building "physics-aware" neural networks.

Finally, future studies should focus on "perception-planning-execution" closed-loop intelligent system building. The final value of registration algorithms should be their translation into clinical practice. The mainstream direction is deeply integrating AI registration with surgical robotic control and real-time intraoperative imaging systems. The development of smart navigators that are able to track deformations in real time and dynamically correct trajectories, enabling a fully automatic closed loop between imaging diagnosis/pre-operative planning and robotically accurate intervention, is the ultimate way toward bringing PCa management to the age of automation, personalization, and non-invasive precision. Importantly, in order for such a shift to occur, further work is needed to reduce the "black-box" aspect of DL and focus on explainable AI techniques. Providing interpretable models with associated measures of confidence or maps showing important anatomical features would allow clinicians to understand how the model made its prediction, thus enabling the clinical interpretability and trust required for high-risk intervention adoption.

6 CONCLUSION

Non-rigid MRI-TRUS registration is the key technology to achieve precision-targeted biopsy or focal therapy of PCa. Although it plays a very important role, accurate alignment has long been a challenge in the field of MI computing due to different cross-modal imaging physics, complex soft-tissue deformations, and the ubiquitous absence of voxel-level annotations. DL has brought about a revolution in tackling such a non-linear mapping issue. In this survey, we have described clearly how the field has evolved: starting with fully supervised CNN E2E regressors, progressing to mixed CNN architectures using both weak supervision and physics-based constraints; from the use of GANs for bridging modality gaps to high-precision generative registration with diffusion models; and, finally, toward joint modeling of local fine structure and global topology through hybrid architectures and attention mechanisms. To date, it has been shown by the following works that DL techniques are making significant progress in improving registration accuracy, acceleration, and convergence.

Currently, SOTA DL models have achieved sub-millimeter-level accuracy (TRE<1.0 mm) on standard public datasets, with short (<0.5 s) inference times, making them suitable for intraoperative real-time navigation. All these advances provide solid anatomical guidance for accurate biopsy, ablative therapy, and brachytherapy. However, we have to admit that all the above results are built upon idealized experiments; there exists a large "generalization gap" from lab benchmarks to real-world clinical scenarios. The stability and safety of such models against multicenter heterogeneous data, pathologically complex cases, and intraoperative artifacts require additional validation in large-scale, prospective clinical studies.

In future work, we expect to shift our attention away from small improvements in accuracy and toward improving utility for clinical practice by focusing on out-of-sample performance across institutions, developing low-resource training techniques, and strongly incorporating physics and biomechanical prior knowledge into neural network architecture design for plausible and safe deformation predictions. The final goal is to build highly automated, explainable, and effortlessly deployed E2E intelligent interventional systems through the synergistic coupling between advanced algorithms and surgical robots. These tools are expected to bring the field of PCa care into an age of automation, precision, and intelligent minimally invasive interventions.

DECLARATIONS

Author contributions

Peiyu Chen conducted the data analysis and drafted the initial manuscript. Xudong Guo supervised the entire study and provided critical contributions to the writing and revision of the paper.

Funding

This work was supported by the National Natural Science Foundation of China (62373253), the National Key Research and Development Program of China (2023YFF1105002), the Open Project Fund of the State Key Laboratory of Cardiovascular Diseases (2024SKL-TJ012), and the Huzhou Key Laboratory of Precise Diagnosis and Treatment of Urinary Tumor Fund Project (HZMNZL202401).

Data availability

The datasets analyzed in this review are available in the public repositories cited within the article. No new data were created during this study.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

Not applicable.

REFERENCES

[1] Hassan MR, Islam MF, Uddin MZ, Ghoshal G, Hassan MM, Huda S, et al. Prostate cancer classification from ultrasound and MRI images using deep learning based Explainable Artificial Intelligence. Future Generation Computer Systems. 2022;127:462-472. https://doi.org/10.1016/j.future.2021.09.030

[2] Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2024 May-Jun;74(3):229-263. https://doi.org/10.3322/caac.21834

[3] Avolio PP, Lughezzani G, Anidjar M, Hassan T, Rompré-Brodeur A, Buffi NM, et al. The diagnostic accuracy of micro-ultrasound for prostate cancer diagnosis: A review. World J Urol. 2023 Nov;41(11):3267-3276. https://doi.org/10.1007/s00345-023-04521-w

[4] Bektaş M, Chia CM, Burchell GL, Daams F, Bonjer HJ, van der Peet DL. Artificial intelligence-aided ultrasound imaging in hepatopancreatobiliary surgery: Where are we now? Surg Endosc. 2024 Sep;38(9):4869-4879. https://doi.org/10.1007/s00464-024-11130-0

[5] Erbin A, Caglar U, Turkay R. Evaluating the effectiveness and safety of robotic-assisted MRI/TRUS fusion transperineal prostate biopsy systems: A narrative review based on current literature. Eurasian J Med. 2023 Dec;55(Suppl 1):S125-S130. https://doi.org/10.5152/eurasianjmed.2023.23370

[6] Inoue T, Shin T. Current magnetic resonance imaging-based diagnostic strategies for prostate cancer. Int J Urol. 2023 Dec;30(12):1078-1086. https://doi.org/10.1111/iju.15281

[7] Studholme C, Hill DLG, Hawkes DJ. An overlap invariant entropy measure of 3D medical image alignment. Pattern Recognition. 1999;32(1):71-86. https://doi.org/10.1016/s0031-3203(98)00091-0

[8] Wang Z, Bovik AC. Mean squared error: Love it or leave it? A new look at signal fidelity measures. IEEE Signal Processing Magazine. 2009;26(1):98-117. https://doi.org/10.1109/MSP.2008.930649

[9] Heinrich MP, Jenkinson M, Bhushan M, Matin T, Gleeson FV, Brady SM, et al. MIND: Modality independent neighbourhood descriptor for multi-modal deformable registration. Med Image Anal. 2012 Oct;16(7):1423-1435. https://doi.org/10.1016/j.media.2012.05.008

[10] Shakeri S, Le W, Ménard C, Kadoury S. Deformable MRI to transrectal ultrasound registration for prostate interventions with shape-based deep variational auto-encoders. 2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI). 2021; 174-178. https://doi.org/10.1109/ISBI48211.2021.9434101

[11] Dupuy T, Beitone C, Troccaz J, Voros S. 2D/3D deep registration for real-time prostate biopsy navigation. Medical Imaging 2021: Image-Guided Procedures, Robotic Interventions, and Modeling. 2021;58. https://doi.org/10.1117/12.2579874

[12] Mitra J, Martí R, Oliver A, Lladó X, Ghose S, Vilanova JC, et al. Prostate multimodality image registration based on B-splines and quadrature local energy. Int J Comput Assist Radiol Surg. 2012 May;7(3):445-454. https://doi.org/10.1007/s11548-011-0635-8

[13] Zettinig O, Shah A, Hennersperger C, Eiber M, Kroll C, Kübler H, et al. Multimodal image-guided prostate fusion biopsy based on automatic deformable registration. Int J Comput Assist Radiol Surg. 2015 Dec;10(12):1997-2007. https://doi.org/10.1007/s11548-015-1233-y

[14] Hu Y, Gibson E, Ahmed HU, Moore CM, Emberton M, Barratt DC. Population-based prediction of subject-specific prostate deformation for MR-to-ultrasound image registration. Med Image Anal. 2015 Dec;26(1):332-344. https://doi.org/10.1016/j.media.2015.10.006

[15] Yang X, Rossi P, Mao H, Jani AB, Ogunleye T, Curran WJ, et al. A MR-TRUS registration method for ultrasound-guided prostate interventions. Proc SPIE Int Soc Opt Eng. 2015 Feb;9415:94151Y. https://doi.org/10.1117/12.2077825

[16] Balakrishnan G, Zhao A, Sabuncu MR, Guttag J, Dalca AV. VoxelMorph: A learning framework for deformable medical image registration. IEEE Trans Med Imaging. 2019 Aug;38(8): 1788-1800. https://doi.org/10.1109/TMI.2019.2897538

[17] Zhu N, Najafi M, Han B, Hancock S, Hristov D. Feasibility of image registration for ultrasound-guided prostate radiotherapy based on similarity measurement by a convolutional neural network. Technol Cancer Res Treat. 2019 Jan 1;18:1533033818821964. https://doi.org/10.1177/1533033818821964

[18] Guo H, Kruger M, Xu S, Wood BJ, Yan P. Deep adaptive registration of multi-modal prostate images. Comput Med Imaging Graph. 2020 Sep;84:101769. https://doi.org/10.1016/j.compmedimag.2020.101769

[19] Guo H, Xu X, Song X, Xu S, Chao H, Myers J, et al. Ultrasound frame-to-volume registration via deep learning for interventional guidance. IEEE Trans Ultrason Ferroelectr Freq Control. 2023 Sep;70(9):1016-1025. https://doi.org/10.1109/tuffc.2022.3229903

[20] Zeng Q, Fu Y, Tian Z, Lei Y, Zhang Y, Wang T, et al. Label-driven magnetic resonance imaging (MRI)-transrectal ultrasound (TRUS) registration using weakly supervised learning for MRI-guided prostate radiotherapy. Phys Med Biol. 2020 Jun 26;65(13):135002. https://doi.org/10.1088/1361-6560/ab8cd6

[21] Jiang H, Hu J, Qian X, Zheng Y, Zhou Z, Dai Y. Registration method of MRI-TRUS images based on joint learning and multi-level wavelet feature pyramid. Computer Engineering. 2025;51(10):270-283. https://doi.org/10.19678/j.issn.1000-3428.0069093

[22] Fu Y, Lei Y, Wang T, Patel P, Jani AB, Mao H, et al. Biomechanically constrained non-rigid MR-TRUS prostate registration using deep learning based 3D point cloud matching. Med Image Anal. 2021 Jan;67:101845. https://doi.org/10.1016/j.media.2020.101845

[23] Yan P, Xu S, Rastinehad AR, Wood BJ. Adversarial image registration with application for MR and TRUS image fusion. Machine Learning in Medical Imaging. 2018;197-204. https://doi.org/10.1007/978-3-030-00919-9_23

[24] Lian J, Pang Y, Chen X, Huang Y, Lian B, Yap P. Deformation-aware MR-TRUS image translation for prostate cancer brachytherapy. Res Sq. 2025 Nov 18;rs.3.rs-7958252. https://doi.org/10.21203/rs.3.rs-7958252/v1

[25] Feng Z, Ni D, Wang Y. Salient region matching for fully automated MR-TRUS registration. 2025 IEEE 22nd International Symposium on Biomedical Imaging (ISBI). 2025;1-5. https://doi.org/10.1109/isbi60581.2025.10981019

[26] Hao Y, Li S, Zhao J, Zhang C. BGProReg: A biomechanically generative framework for prostate MRI-TRUS deformable image registration. Information Fusion. 2026;127:103911. https://doi.org/10.1016/j.inffus.2025.103911

[27] Yao Z, Chen J, Wen T. A label-aware diffusion model for weakly supervised deformable registration of multimodal MRI–TRUS in prostate cancer. Int J Comput Assist Radiol Surg. 2025 Nov 6. https://doi.org/10.1007/s11548-025-03538-3

[28] Chen J, He Y, Frey EC, Li Y, Du Y. ViT-V-Net: Vision transformer for unsupervised volumetric medical image registration. arXiv. 2021. https://doi.org/10.48550/arXiv.2104.06468

[29] Ghahremani M, Khateri M, Jian B, Wiestler B, Adeli E, Wachinger C. H-ViT: A Hierarchical vision transformer for deformable image registration. 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2024;11513-11523. https://doi.org/10.1109/CVPR52733.2024.01094

[30] Mahmoudi H, Ramadan H, Riffi J, Tairi H. 3D-WDA-PMorph: Efficient 3D MRI/TRUS prostate registration using transformer-CNN network and wavelet-3D-depthwise-attention. J Imaging Inform Med. 2026 Apr;39(2):1474-1493. https://doi.org/10.1007/s10278-025-01615-2

[31] Yu J, Xiao B, Wang J, Mi H, Wang E, Huang N, et al. Weakly supervised deep learning for multimodal MRI-TRUS registration: Toward assisting prostate biopsy guidance. Digit Health. 2025 Sep 8;11:20552076251375870. https://doi.org/10.1177/20552076251375870

[32] Baum ZMC, Saeed SU, Min Z, Hu Y, Barratt DC. MR to ultrasound registration for prostate challenge - Dataset. 2023. https://doi.org/10.5281/zenodo.8004388

[33] Natarajan S, Priester A, Margolis D, Huang J, Marks L. Prostate MRI and ultrasound with pathology and coordinates of tracked biopsy (Prostate-MRI-US-Biopsy) [Data set]. The Cancer Imaging Archive (TCIA). 2020.

[34] Benelli A, Vaccaro C, Guzzo S, Nedbal C, Varca V, Gregori A. The role of MRI/TRUS fusion biopsy in the diagnosis of clinically significant prostate cancer. Ther Adv Urol. 2020 May 18;12:1756287220916613. https://doi.org/10.1177/1756287220916613

[35] Liu W, Kadier A, Shen D, He Y, Dong S, Zhu K, et al. Combined MRI-TRUS fusion targeted and systematic biopsy versus systematic biopsy alone for the detection of prostate cancer: protocol for a prospective single-centre trial. BMJ Open. 2024 Mar 1;14(3):e080593. https://doi.org/10.1136/bmjopen-2023-080593

[36] Wu Y, Lu Q, Liu Z, Wu J, He X, Yang Y, et al. Advances in imaging and artificial intelligence for precision diagnosis and biopsy guidance in prostate cancer. Front Oncol. 2025 Oct 13;15:1614891. https://doi.org/10.3389/fonc.2025.1614891

[37] Soerensen SJC, Fan RE, Seetharaman A, Chen L, Shao W, Bhattacharya I, et al. Deep learning improves speed and accuracy of prostate gland segmentations on magnetic resonance imaging for targeted biopsy. J Urol. 2021 Sep;206(3):604-612. https://doi.org/10.1097/ju.0000000000001783

[38] Palladino L, Maris B, Antonelli A, Fiorini P. PROST-Net: A deep learning approach to support real-time fusion in prostate biopsy. IEEE Trans Med Robot Bionics. 2022;4(2):323-326. https://doi.org/10.1109/tmrb.2022.3145667

[39] Wang B, Yu L, Pan Z, Ni H, Lin Z, Fan Y. MRI/TRUS-guided auxiliary system for transperineal prostate biopsy based on deep learning. 2023 6th International Conference on Information Communication and Signal Processing (ICICSP). 2023;914-921. https://doi.org/10.1109/ICICSP59554.2023.10390795

[40] Alabousi M, Ghai S. Magnetic resonance imaging-guided ultrasound ablation for prostate cancer – A contemporary review of performance. Front Oncol. 2023 Jan 4;12:1069518. https://doi.org/10.3389/fonc.2022.1069518

[41] Karwacki J, Kiełbasa J, Szczepaniak Z, Zagórski K, Kobylański M, Gurwin A, et al. Current status of cryoablation in prostate cancer management. Clin Med Insights Oncol. 2025 Jul 3;19:11795549251350830. https://doi.org/10.1177/11795549251350830

[42] Beek E, Hata N, Tuncali K, Moreira P. Image-guided adaptive cryotherapy for prostate cancer treatment. Ann Biomed Eng. 2025 Nov;53(11):3237-3246. https://doi.org/10.1007/s10439-025-03833-9

[43] van Luijtelaar A, Fütterer JJ, Bomers JGR. Minimally invasive magnetic resonance image-guided prostate interventions. Br J Radiol. 2022 Mar 1;95(1131):20210698. https://doi.org/10.1259/bjr.20210698

[44] Chen Y, Xing L, Yu L, Liu W, Pooya Fahimian B, Niedermayr T, et al. MR to ultrasound image registration with segmentation-based learning for HDR prostate brachytherapy. Med Phys. 2021 Jun;48(6):3074-3083. https://doi.org/10.1002/mp.14901

[45] Yang Y, Zhang H, Gichoya JW, Katabi D, Ghassemi M. The limits of fair medical imaging AI in real-world generalization. Nat Med. 2024 Oct;30(10):2838-2848. https://doi.org/10.1038/s41591-024-03113-4

[46] Kushol R, Wilman AH, Kalra S, Yang Y. DSMRI: Domain shift analyzer for multi-center MRI datasets. Diagnostics (Basel, Switzerland). 2023 Sep 14;13(18):2947. https://doi.org/10.3390/diagnostics13182947

[47] Karnik VV, Fenster A, Bax J, Cool DW, Gardi L, Gyacskov I, et al. Assessment of image registration accuracy in three-dimensional transrectal ultrasound guided prostate biopsy. Med Phys. 2010 Feb;37(2):802-813. https://doi.org/10.1118/1.3298010

[48] Chaddad A, Peng J, Xu J, Bouridane A. Survey of explainable AI techniques in healthcare. Sensors (Basel, Switzerland). 2023 Jan 5;23(2):634. https://doi.org/10.3390/s23020634