International Journal of Computer Assisted Radiology and Surgery

Global Registration of Kidneys in 3D Ultrasound and CT images

--Manuscript Draft--

Manuscript Number:	CARS-D-24-00082
Full Title:	Global Registration of Kidneys in 3D Ultrasound and CT images
Article Type:	Original Article
Keywords:	Ultrasound; registration; landmark localization; fusion imaging; interventional guidance
Corresponding Author:	Toby Collins IRCAD France Clermont-Ferrand, FRANCE
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	IRCAD France
Corresponding Author's Secondary Institution:	
First Author:	William Brice Ndzimbong, M.D.
First Author Secondary Information:	
Order of Authors:	William Brice Ndzimbong, M.D.
	Nicolas Thome
	Cyril Fourniol
	Yvonne Keeza
	Benoît Sauer
	Jacques Marescaux
	Daniel George
	Alexandre Hostettler
	Toby Collins, Ph.D
Order of Authors Secondary Information:	
Funding Information:	Conseil régional du Grand Est — Pr Jacques Marescaux Dr Alexandre Hostettler
Abstract:	Purpose Automatic registration between abdominal ultrasound (US) and computed tomography (CT) images is needed to enhance interventional guidance of renal procedures, but it remains an open research challenge. We propose a novel method that doesn't require an initial registration estimate (a global method) and also handles registration ambiguity caused by the organ's natural symmetry. Combined with a registration refinement algorithm, this method achieves robust and accurate kidney registration while avoiding manual initialization. Methods The global registration method has 3 main steps: 1) Automatic anatomical landmark localization, where two deep neural networks (DNNs) are trained to localize a set of renal landmarks in both modalities. 2) Registration hypothesis generation, where all geometrically compatible registration solutions are computed using a novel algorithm (LESAC), designed to handle label constraints and ambiguous landmarks due to image self-similarity. 3) Hypothesis selection, where the best solution is selected using landmark registration error and physical constraints to break registration symmetries. The selected hypothesis can be used to initialize any suitable registration refinement method. Results are presented with state-of-the-art surface-based matching (BCPD and ICP) and automatic kidney segmentation. Results This automatic approach for registration initialization using direct voxel

	information gives better results than several state-of-the-art methods using organ surface information. The results obtained on 59 pairs of 3D US/CT kidney images show that the average error on a registration target (Target Registration Error: TRE) is 5.45 mm with ICP refinement and 5.52 mm with BCPD refinement, and the average Dice score is 80.7\% with ICP refinement, 83.3\% with BCPD refinement. Conclusion This work presents the first approach for automatic kidney registration in US and CT images, which doesn't require an initial manual registration estimate to be known a priori. The results show a fully automatic registration approach with performances comparable to manual methods is feasible.

Global Registration of Kidneys in 3D Ultrasound and CT images

William Ndzimbong[1,2*], Nicolas Thome[3], Cyril Fourniol[4], Yvonne Keeza[2], Benoît Sauer[5], Jacques Marescaux[6], Daniel George[1], Alexandre Hostettler[2,6†], Toby Collins[2,6*†]

[1]University of Strasbourg, CNRS, ICUBE Laboratory, Strasbourg, France.
[2]Research Institute against Digestive Cancer (IRCAD), Kigali, Rwanda.
[3]Sorbonne University, CNRS, ISIR, Paris, France.
[4]Department of Urology, HEGP, APHP, Paris, France.
[5]Medical Imaging Group MIM, Clinique Sainte Anne, Strasbourg, France.
[6]Research Institute against Digestive Cancer (IRCAD), Strasbourg, France.

*Corresponding author(s). E-mail(s): william.ndzimbong@ircad.fr; toby.collins@ircad.fr;
†These authors share the author position.

Abstract

Purpose Automatic registration between abdominal ultrasound (US) and computed tomography (CT) images is needed to enhance interventional guidance of renal procedures, but it remains an open research challenge. We propose a novel method that doesn't require an initial registration estimate (a global method) and also handles registration ambiguity caused by the organ's natural symmetry. Combined with a registration refinement algorithm, this method achieves robust and accurate kidney registration while avoiding manual initialization.

Methods The global registration method has 3 main steps: 1) Automatic anatomical landmark localization, where two deep neural networks (DNNs) are trained to localize a set of renal landmarks in both modalities. 2) Registration hypothesis generation, where all geometrically compatible registration solutions are computed using a novel algorithm (LESAC), designed to handle label constraints and ambiguous landmarks due to image self-similarity. 3) Hypothesis selection, where the best solution is selected using landmark registration error and physical constraints to break registration symmetries. The selected hypothesis can be used to initialize any suitable registration refinement method. Results are presented with state-of-the-art surface-based matching (BCPD and ICP) and automatic kidney segmentation.

Results The results obtained on 59 pairs of 3D US/CT kidney images show that the average error on a registration target (Target Registration Error: TRE) is 5.45 mm with ICP refinement and 5.52 mm with BCPD refinement, and the average Dice score is 80.7% with ICP refinement, 83.3% with BCPD refinement. This global registration method gives better results than several state-of-the-art global methods.

Conclusion This work presents the first approach for automatic kidney registration in US and CT images, which doesn't require an initial manual registration estimate to be known a priori. The results show a fully automatic registration approach with performances comparable to manual methods is feasible.

Keywords: Ultrasound, registration, landmark localization, fusion imaging, interventional guidance

1 Introduction

Abdominal US is a cost-effective, real-time imaging modality for various procedures. However, it has limitations like poor contrast, acoustic shadows, artifacts, and operator dependence. Image fusion, combining pre-operative CT and intra-operative US, addresses these limitations by spatially aligning (registering) these modalities. Despite decades of research, US inter-modal registration remains challenging. Currently, certified devices, like Siemens Acuson S3000, rely on a manual alignment process, posing issues of operator dependence and time consumption. The need for an accurate and automatic registration algorithm is crucial for reducing time, reducing, operator-dependence, and advancing fusion imaging availability for standard and robotic-assisted percutaneous procedures.

Inter-modal 3D medical image registration is a major and active research topic, where methods can be characterized as either a *global* or *refinement* approaches. Global methods don't require an initial registration estimate to be known *a priori*, and there are two main approaches: *feature-based* and *shape-based*. Feature-based methods use distinct 3D points with discriminative structural information, which can be detected in the source and target images. They are then matched using a descriptor vector that characterizes the keypoint's local structure. A matching step is then followed using descriptor similarity and a geometric alignment algorithm such as RANSAC or TEASER++ [1]. Various methods exist to extract features from 3D surfaces or point clouds, such as FPFH [2] (hand-crafted) and DIP [3] (learning-based). Their main advantage is to reduce the domain gap between modalities using surfaces as a domain bridge. Other methods aim to extract features and descriptors directly from 2D [4] or 3D images [5]. However, due to strong domain differences, they have limited success in inter-modal registration with US/CT (or MR) images. In contrast to feature-based methods, shape-based methods aim to register segmented 3D surfaces. Various methods have been proposed, including spectral graph matching [6], Principal Component Analysis (PCA) [7], and Deep Learning [8]. However, they are limited by the need for accurate segmentation as a pre-processing step. all use

Several global registration methods have been presented specifically to register CT (or MR) and US abdominal organs [5, 9–13]. For example, in [9], a Controlled Random Search (CRS) algorithm is applied in the diaphragm's segmented region for global liver registration in 3D US/MR images. In [5], the authors use two CNNs to extract features in 2D US/MR images of the liver to learn "correspondent keypoints", which in turn are used to estimate a rigid transformation. An approach based on particle filtering was proposed to register the prostate in 3D US/MR images globally [10]. However, there have been few attempts to solve global kidney registration, which is challenging due to the domain difference, US artifacts such as acoustic shadows from ribs and intestinal gas, and low contrast between the kidney and surrounding structures. Another important challenge is that the kidney has strong bilateral symmetry, which can cause existing global methods to fail to find the correct solution. Previously published works don't solve global kidney registration [11–13]. Instead, they solve registration refinement, where an initial registration is provided a priori, which is improved typically with an iterative process. Various refinement methods have

been proposed, including intensity-based (ImFusion Suite - ImFusion GmbH, Munich, Germany), and VoxelMorph), and surface-based [13, 13–16] methods. However, they require a good initial registration estimate for initialization *i.e.* a global registration method. Consequently, new contributions for robust, automatic global registration are important.

Contributions statements.

This work presents a novel feature-based method for the automatic global registration of 3D ultrasound and CT kidney images. Key contributions include:

1. Achieving automatic spatial localization of registration landmarks using two independent landmark localization DNNs, one for each modality. This eliminates the need for paired training data and avoids using an intermediate representation, such as segmented surfaces, to bridge the domain gap.
2. Introducing a landmark matching approach using a small number of semantic labels (default of 4). This not only eliminates the need for feature descriptors but also enables global robust alignment through an exhaustive search method. We propose the LESAC method, a variant of RANSAC without randomization, offering determinism, guaranteed global optimality (important for algorithm documentation and clinical translation), and the ability to find multiple registration solutions due to quasi-symmetry, plus the removal of a key RANSAC parameter (number of search iterations).
3. The presented global registration method may be combined with any registration refinement method. We demonstrate it with Bayesian Coherent Point Drift (BCPD) refinement, known for strong performance in inter-modal and surface-based registration tasks [14–16], and ICP.
4. To the best of our knowledge, this work is the first method for US/CT global registration of the kidney, facilitating fusion imaging for renal procedures without manual registration. Comparative evaluations against various baseline methods from related registration tasks highlight superior performance in accuracy, speed, and handling kidney symmetry.

2 Method

2.1 Dataset

The global registration method here relies on a training dataset of 3D US and CT images of human kidneys, requiring several anatomical landmark annotations. The images don't need to be paired, and organ segmentation isn't necessary. A newly created dataset [14] (under peer-review) includes 3D US and CT images from 48 human patients. We used the image, kidney segmentation, and anatomical landmark annotations of 5 landmarks (Figure 1) for all kidneys that were annotated in both modalities (59 kidneys). While [14] benchmarks several existing registration refinement methods, this work uses the dataset for a new global registration methodology. Please refer to [14] for specific dataset details.

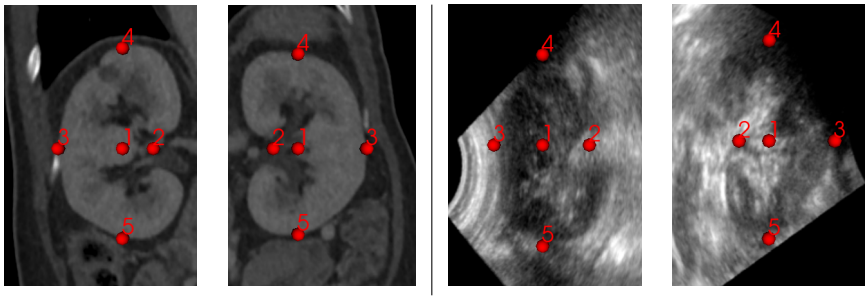


Fig. 1 Dataset used to train and test the proposed method [14]. 2D CT (two left) and 2D US (two right) slices of the left and right kidney in the canonical sagittal plane with 5 anatomical landmarks overlaid.

2.2 Global registration methodology

2.2.1 Step 1: Symmetry-invariant anatomical landmark localization

The DNNs were trained to segment local regions surrounding each landmark, similarly to heat-maps [17, 18], where one binary region (mask) was inferred per landmark. Mask values of non-zero or zero represented voxels within or outside the landmark's local vicinity using a spherical radius in voxels (we use a default radius corresponding to approximately 10 mm). A significant challenge, especially for localizing US landmarks, was due to the limited imaging window associated with 3D US data (the images had sizes of approximately $228 \times 162 \times 222\ mm^3$), and the kidney strong 2-fold (bilateral) symmetry in the transversal plane (Figure 1). Consequently, differentiating rotationally-symmetric landmarks (4 and 5) is highly challenging for state-of-the-art DNNs using only image data. We solved this using label fusion, where landmarks 4 and 5 were merged as one label ('kidney-pole'). Consequently, the DNN's task was to detect the poles but not differentiate between the superior and anterior poles. This differentiation was performed in the downstream registration task, using additional geometric information (Section 2.2.3).

Any state-of-the-art DNN may be used to segment landmarks. We present results using nnU-Net [19], based on its strong general performance for medical image segmentation. The DNNs were trained with supervised learning, using the "3d_fullres" self-configuration and standard parameters. The complete training configuration is provided in the supplementary material. After inference, landmark positions are estimated using region centroids from a connected component analysis. We denote as $\mathcal{P}$ and $\mathcal{Q}$ the set of all detected landmarks in US (moving) and CT (fixed) 3D image coordinates, respectively. They have associated integer labels denoted by $L(\mathbf{p} \in \mathcal{P}) \in \mathbb{Z}$ and $L(\mathbf{q} \in \mathcal{Q}) \in \mathbb{Z}$. In practice, a label may have zero, one, or more points in $\mathcal{P}$ or $\mathcal{Q}$. In the US, for example, a landmark may not be detected due to an acoustic shadow. DNN errors may also cause spurious landmarks. Consequently, estimating the spatial transform from detected landmarks isn't trivial. In the following step, a simple, fast, and deterministic algorithm is used to establish an optimal registration that handles these challenges.

2.2.2 Step 2: Exhaustive registration hypothesis generation

A set of registration hypotheses are generated from $\mathcal{P}$, $\mathcal{Q}$, and L, by adapting RANSAC in two aspects: firstly, to ensure point matches are made only between points with the same labels, secondly, by generating registration hypotheses from all possible minimal sets of M point matches. This not only removes the randomization component and associated hyper-parameters of RANSAC, but also enables multiple registration hypotheses to be found due to organ symmetry (which generates symmetries in the point sets). We refer to this as robust registration hypothesis generation using

Labeled Exhaustive Sampling and Consensus (LESAC), with pseudo-code provided in Algorithm 1.

Similar to RANSAC, a spatial transform function that trades off registration accuracy with complexity must be chosen. Our objective is to have an approximate global registration that can be refined by a registration refinement algorithm such as BCPD. We present results with a similarity transform, requiring a minimum of $M = 3$. Note that the approach generalizes to other transforms.

Data: The moving and fixed point sets, $\mathcal{P}$ and $\mathcal{Q}$, with associated integer labels $L(\mathbf{p} \in \mathcal{P}) \in \mathbb{Z}$ and $L(\mathbf{q} \in \mathcal{Q}) \in \mathbb{Z}$
Result: The set of registration hypotheses, $\mathcal{H}$, with their associated numbers of inlier matches $\mathcal{I}$ and root mean squared errors (RMSEs) $\mathcal{E}$
$\mathcal{H} \leftarrow \emptyset$, $\mathcal{I} \leftarrow \emptyset$, $\mathcal{E} \leftarrow \emptyset$;

forall $\mathbf{p}_1, \ldots, \mathbf{p}_M, \mathbf{q}_1, \ldots, \mathbf{q}_M$ such that $L(\mathbf{p}_1) = L(\mathbf{q}_1), \ldots, L(\mathbf{p}_M) = L(\mathbf{q}_M)$ **do**
  Compute best-fitting transform $\mathbf{H}$ aligning $(\mathbf{p}_1, \ldots, \mathbf{p}_M)$ to $(\mathbf{q}_1, \ldots, \mathbf{q}_M)$
  $s \leftarrow 0$; `// Number of inlier matches associated to` $\mathbf{H}$
  forall $\mathbf{p} \in \mathcal{P}$ **do**
    $\mathbf{p}' \leftarrow \mathbf{H}(\mathbf{p})$;
    $\mathcal{S} \leftarrow \emptyset$; `// Set of inlier matches`
    $\mathcal{Q}' = \{\mathbf{q} \in \mathcal{Q} \mid L(\mathbf{q}) = L(\mathbf{p})\}$;
    $\mathbf{q} = \arg\min_{\mathbf{q}_i \in \mathcal{Q}'} \|\mathbf{q}_i - \mathbf{p}'\|$; `// nearest neighbour match`
    if $\|\mathbf{q} - \mathbf{p}'\| < \tau$ **then**
      $\mathcal{S} \leftarrow \mathcal{S} \cup \{(\mathbf{p}, \mathbf{q})\}$; `// add match to inlier set`
      $s \leftarrow s + 1$
    end
  end
  if $s \geq M$ **then**
    Compute best-fitting transform $\mathbf{H}'$ with corresponding RMSE e, using all inlier matches $\mathcal{S}$;
    $\mathcal{H} \leftarrow \mathcal{H} \cup \{\mathbf{H}'\}$;
    $\mathcal{I} \leftarrow \mathcal{S} \cup \{s\}$;
    $\mathcal{E} \leftarrow \mathcal{E} \cup \{e\}$;
  end
end

return $\mathcal{H}, \mathcal{S}, \mathcal{E}$;

Algorithm 1: LESAC: Registration hypothesis generation using Labeled Exhaustive Sampling and Consensus

2.2.3 Step 3: Registration hypothesis selection

Having run Algorithm 1, the best registration hypotheses are extracted using the following sequential (greedy) selection. First the hypothesis $H_{i*} \in \mathcal{H}$ with highest score $s_{i*} = \max_{s \in \mathcal{I}}$ is selected. In the case of ties, the associated RMSE scores in $\mathcal{E}$ are used to tie-break. Then, all transforms in $\mathcal{H}$ that are geometrically similar to H_{i*} are eliminated, generating a residual hypothesis set $\mathcal{H}' \subset \mathcal{H}$. Similarity was assessed by measuring if the rotational difference in degrees (using Rodrigues' formula) exceeded a threshold (we used as default 30 degrees). Having formed $\mathcal{H}'$, the next best hypothesis in $\mathcal{H}'$ is selected, and the greedy selection process continues until either no more hypotheses can be selected or a maximum number max_H of hypotheses have been

selected. For the kidney, knowing that it normally has strong bilateral symmetry, we imposed the early stopping criteria of $max_H = 2$ in our experiments.

2.3 Initialization hypothesis selection

A process is then required to eliminate the incorrect registration hypothesis from Step 3. Additional geometric information about the patient's posture may be exploited to achieve this goal. There are several possible approaches depending on the hardware setup. For example, consider situations when the 3D US data is acquired using a tracked 2D probe, with an EM sensor (a common scenario in image fusion applications and devices, such as the Acuson S3000). In this situation, the patient's head-feet 3D direction (in US image coordinates) can be approximated by moving the tracked probe from the patient's head to the feet. Combined with the fact that CT images are normally acquired in the supine position (head-feet direction along the positive z-axis), there is enough geometric information to eliminate the incorrect hypothesis (that flips the kidney poles wrongly). While this needs further investigation, the dataset lacks tracking information. Consequently, to assess registration accuracy in the subsequent section, we eliminated the least compatible hypothesis with respect to the ground-truth landmark annotations (the compatibility function is provided in the supplementary material).

3 Experimental Validation

This section describes the experimental validation we conducted using the dataset of [14]. 5-fold cross-validation (CV) was applied using the same splits for all methods. For baseline methods with free hyper-parameters (FPFH and DIP - notably feature scale), the hyper-parameters were optimized on each CV training fold, and then frozen before evaluating performance on the test fold. All DNNs were trained on each CV training fold and tested without optimization on the test fold.

3.1 Baselines and method configurations

Baselines

The method was compared against a range of competitive global registration baselines. For a fair comparison, all baselines were configured to use similarity spatial transforms, which were then refined using the same registration refinement methods. For all methods using feature detection and matching on 3D surfaces (FPFH [2] and DIP [3]), the kidney surfaces were computed automatically in CT and US images using trained nnU-Net models, as described [14].

The global registration baselines were:

- **(FPFH,RANSAC)**. This used surface feature matching with state-of-the-art hand-crafted features (FPFH). RANSAC was used to perform robust feature matching and transform estimation.
- **(DIP,RANSAC)**. This used surface feature matching with state-of-the-art deeply-learned features (DIP), combined with RANSAC.
- **(FPFH,TEASER)**. This used FPFH features combined with TEASER++ [1] to perform robust feature matching and transform estimation.
- **(DIP,TEASER)**. This used DIP features combined with TEASER++
- **PCA**. This used surface segmentation and PCA-based pose estimation, which were implemented as follows. Firstly, PCA aligned the two principal axes of the US and CT surfaces. This alignment had four solutions, from which two solutions with the lowest surface-to-surface error were outputted. Similarly to our method, the solution used for evaluation was the one with the highest ground-truth landmark compatibility.

- **Manual**. This uses manual initialization, using the similarity transform computed from ground-truth landmarks. This was included to assess the performance gap between automatic and manual initialization.

In our proposed approach, there were instances where fewer than 3 landmark matches were found from the landmark localization DNNs. This mainly happened when the US image quality was especially low and where registration may not ultimately be achievable with sufficient reliability for clinical application. To fairly compare the proposed method on all cases in the evaluation, we used a fall-back validation method - which enables all methods to be compared using all cases, even if some methods do not return solutions in all cases. We used PCA as the fall-back method, whose solution was reverted to any case with fewer than 3 matching landmarks. There were 6 such cases (10.1%).

The efficacy of the above baselines and our method was assessed in the context of the intended use *i.e.* as an automatic initialization method that may be used by a registration refinement method that requires an initial registration estimate. Two competitive registration refinement methods were considered in this work:

- **ICP**. Iterative Closest Point, configured with an affine spatial transform. This has been used in many medical image registration tasks and requires segmented surfaces. We used the same surfaces to compute FPFH and DIP features from nnU-Net.
- **BCPD**. Bayesian Coherent Point Drift [20], a recent extension of Coherent Point Drift, which is popular for registering surfaces and point clouds extracted from medical images, e.g. [21].

3.2 Experimental results

3.3 Method configurations and performance metrics

We compared our method in two configurations. The first, denoted as **Proposed-4L**, was the same as described in Section 3.4. As an ablation study, a second configuration (denoted as **Proposed-5L**) was used to demonstrate the limitations if we ignore organ symmetry in the algorithm design. To this end, **Proposed-5L** differed from **Proposed-4L** in two aspects: First, label fusion wasn't applied to the kidney pole training labels. Therefore, the trained landmark localization DNNs in **Proposed-5L** differentiated superior versus inferior poles. Secondly, the maximum number of outputted solutions max_H was set to 1.

3.4 Registration refinement methods and performance metrics

In this study, we applied all global registration methods as registration initializers for two surface-based refinement methods: Bayesian Coherent Point Drift [15, 20]- considered state-of-the-art, as well as ICP (implemented in Open3D). nnU-Net was used to automatically segment the kidney surfaces in US and CT modalities, as described in [14]. Performance was measured with three metrics:

- **Target Registration Error (TRE)**. We used the ground-truth (GT) position of the central landmark ('renal pelvis') as a registration target, from which TRE was assessed. Note that we didn't provide the ground-truth location of this target to any method. We chose not to use other landmarks as registration targets due to the relatively high uncertainty associated with their ground-truth [14].
- **Region overlap (DICE)**. Using the estimated registration, we used the region overlap (DICE score) between the GT organ surface segmentations after spatial alignment.
- **Rotation agreement**. While widely employed for general medical registration assessment, the previous metrics exhibit a notable limitation when applied to structures with strong symmetry, as illustrated in the case of the kidney (Table 1). Good

DICE and TREs may result from registrations incorrectly flipping the kidney's superior/anterior orientation. We introduced the rotation agreement metric to address this issue, quantifying the proportion of cases where the rotational component aligns with the superior/anterior orientation. Specifically, a rotational difference exceeding 90 degrees, as measured via Rodrigues' formula, indicates a disagreement. A relatively high threshold was used so the metric would identify cases where the registration is evidently wrong due to geometric symmetry.

3.5 Main results

The quantitative results are shown in Table 1. Considering rotation agreement, among the automatic methods, **Proposed-4L** achieved 100%, substantially higher than the next best method (PCA). Note that our ablation configuration **Proposed-5L** was in third place, showing the importance of our design choices to handle the kidney's symmetry. The other methods had rotation agreement between 27.12% (FPFH, RANSAC) and 77.96% (DIP, TEASER). The best mean and median TRE was obtained by **Proposed-4L**, with both BCPD and ICP refinements, with only approximately 1mm higher than **Manual+BCPD**. With ICP refinement, **Proposed-4L** had a lower 3rd quartile TRE than **Proposed-5L**, and a greater one with BCPD. Recall that the landmark used to evaluate TRE was approximately in the center of the kidney. Therefore, the TRE metric was not highly sensitive to large errors in rotation about the center (unlike the rotation agreement metric). Concerning Dice, the highest mean, median, and 3rd quartile scores were obtained by **Proposed-4L**. Refinement with BCPD achieved superior performance with a mean DICE score of 83.30%.

	Average rot. agreement in % (↑)	Median TRE over folds in mm (↓)	3rd quartile TRE over folds in mm (↓)	Average TRE (std) over folds in mm (↓)	Median Dice score over folds in % (↑)	3rd quartile Dice score over folds in % (↑)	Average Dice score (std) over folds in % (↑)
Manual + BCPD	100	4.73	4.75	4.53 (0.79)	83.70	83.50	84.20 (1.80)
(Initialization method) + BCPD refinement							
PCA	86.44*	6.13	7.56	6.54 (1.68)*	80.70	79.90	81.40 (2.40)*
(FPFH,RANSAC)	27.12*	11.55	13.28	11.70 (2.72)*	66.60	66.40	67.90 (3.50)*
(FPFH,TEASER)	44.07*	8.73	8.94	9.26 (1.31)*	72.90	71.30	72.70 (1.70)*
(DIP,RANSAC)	55.93*	15.62	20.09	15.46 (4.46)*	65.90	61.90	65.70 (5.40)*
(DIP,TEASER)	77.96*	5.81	7.96	6.61 (1.73)*	79.40	77.60	80.4 (3.10)*
Proposed-5L	77.97*	6.08	**6.30**	5.76 (0.85)	81.10	81.00	82.00 (2.40)*
Proposed-4L	**100**	**5.58**	6.43	**5.52 (1.11)**	**82.00**	**81.80**	**83.30 (2.10)**
(Initialization method) + ICP refinement							
PCA	86.44*	6.50	7.15	6.32 (1.34)*	78.3	77.40	78.90 (2.20)*
(FPFH,RANSAC)	30.51*	12.09	12.58	11.41 (2.73)*	64.90	64.50	65.60 (2.30)*
(FPFH,TEASER)	47.46*	6.78	8.39	7.58 (1.62)*	73.00	72.80	73.10 (1.50)*
(DIP,RANSAC)	52.54*	14.65	15.10	12.67 (3.27)*	64.60	61.30	63.80 (5.90)*
(DIP,TEASER)	76.27*	**5.30**	5.82	6.14 (1.91)*	78.20	78.1	79.2 (2.2)*
Proposed-5L	77.97*	6.22	6.65	6.02 (1.18)	78.30	78.00	79.20 (1.8)*
Proposed-4L	**100**	5.44	**5.46**	**5.45 (0.96)**	**80.60**	**79.5**	**80.70 (1.40)**

Table 1 Quantitative comparison of registration methods. The (↑ / ↓) sign indicates whether higher or lower values are better. The table is divided into three main parts: First, results using **Manual+BCPD** (the reference used against the automatic methods). Second, results of global methods combined with BCPD refinement. Third, results of global methods combined with ICP refinement. The columns show the performance metrics for each configuration. Because cross-validation was used, results are reported using inter-fold averaging. For refinement method, bold shows the best-performing initialization/refinement method combination. Stars indicate if a method's performance is significantly different compared to the best combination (Wilcoxon Signed Rank Test, $p < 0.05$).

Figure 2 shows the distribution of TRE and Dice scores for all global methods, combined with either refinement method. One can see that the curves for **Proposed-4L** are generally closest to the curve of **Manual+BCPD** (the reference) compared to all baselines. Figure 3 shows the predicted landmarks masks in CT (left) and US (right) images. GT masks are shown in light color, with predictions in the same color

with a darker shade. One can see that there is generally closer agreement in CT compared to US, caused by the poorer contrast and shadow artefacts in US. Also note that the bottom pole in US wasn't detected, and there is a significantly high prediction error for one landmark (red - opposite the Renal Hilum). Nevertheless, the algorithm described in 2.2.2 is robust to such problems due to the exhaustive minimal sampling strategy. Such robustness has been demonstrated in Table 1, showing 100% rotation agreement compared to GT.

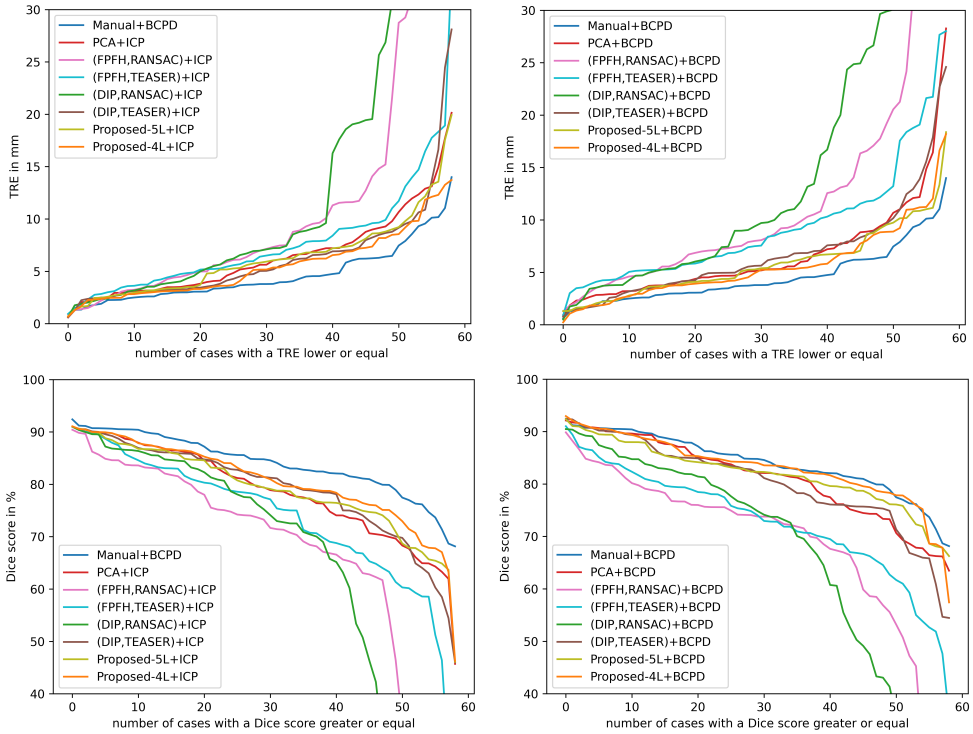


Fig. 2 Performance of global registration methods using the TRE metric (first row) and the Dice score (second row), with ICP (left) and BCPD (right) refinement. The graph shows the number of cases with a TRE (resp. Dice score) lower or equal (resp. greater or equal) to a value for each method on the dataset.

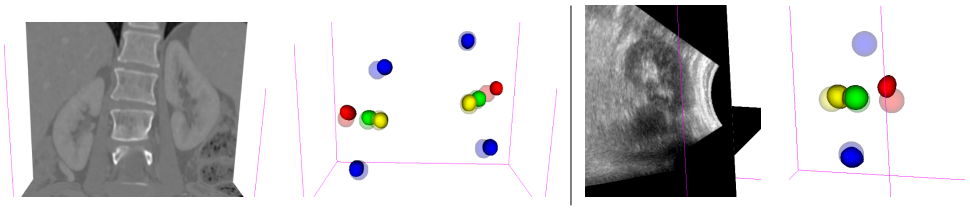


Fig. 3 Visualisations of automatic kidney landmark localization in a CT image (two left) and 3D US volume (two right). The transparent balls represent the label maps corresponding to each ground-truth landmark, and the opaque balls represent the label maps predicted by nnU-Net.

Figure 4 shows two registration examples with each global registration method using BCPD refinement. These images were selected systematically by sorting the kidneys in terms of registration difficulty and then showing the 50^{th} percentile (the median case - top row) and the 75^{th} percentile (bottom row). Difficulty for each case

was assessed by computing the median DICE score from all methods. One can see consistently strong performance with **Proposed-4L**. Global registration computational times ranged from under a second (**Proposed-4L**, **Proposed-5L**, **PCA**, **FPFH**) to over 30 seconds (**DIP**, **TEASER**). Further details and time breakdowns are provided in the supplementary material.

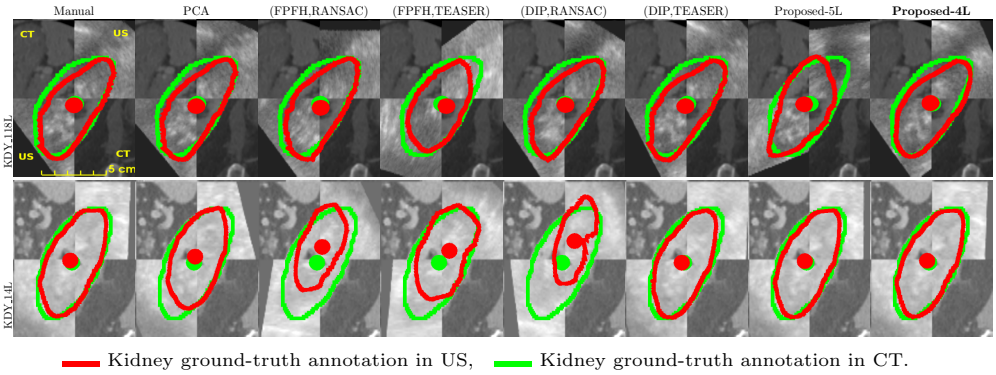


Fig. 4 Qualitative results showing registration performance for examples (KDY_118L) and (KDY_14L) using checker-board visualization. Each column corresponds to an initial registration method + BCPD refinement. Each image shows 4 annotations: green contour - the GT kidney segmentation in CT, green point - the GT position of Landmark 1 in CT, red contour - the GT kidney segmentation in US moved to CT coordinates according to the estimated registration, red point - the GT position of Landmark 1 in US, moved to CT coordinates according to the estimated registration

4 Conclusion

We have introduced a novel global registration method for the kidney, applicable across CT and 3D US modalities. Compatible with various registration refinement methods, our approach demonstrates robustness and superior performance compared to baselines. This method streamlines automatic registration and fusion imaging in renal procedures, eliminating practical inconveniences associated with manual global registration. Future research includes clinical validation with 3D US and freehand 3D ('2D+t') probes and extending its application to deformable Voxelmorph and other image-based refinement techniques, eliminating organ segmentation needs in the refinement step.

References

[1] Yang, H., Shi, J., Carlone, L.: TEASER: Fast and Certifiable Point Cloud Registration. IEEE Trans. Robotics (2020)

[2] Rusu, R.B., Blodow, N., Beetz, M.: Fast point feature histograms (fpfh) for 3d registration. In: IEEE International Conference on Robotics and Automation, pp. 3212–3217 (2009)

[3] Poiesi, F., Boscaini, D.: Distinctive 3d local deep descriptors. In: International Conference on Pattern Recognition, pp. 5720–5727 (2021)

[4] Hu, K., Yuan, X., Chen, S.: Real-time CNN-based keypoint detector with Sobel filter and descriptor trained with keypoint candidates. In: International Conference on Machine Vision, vol. 12701, p. 127010 (2023)

[5] Markova, V., Ronchetti, M., Wein, W., Zettinig, O., Prevost, R.: Global multimodal 2d/3d registration via local descriptors learning. In: Medical Image Computing and Computer-Assisted Intervention, pp. 269–279 (2022). Springer

[6] Zhao, Q., Pizer, S., Niethammer, M., Rosenman, J.: Geometric-feature-based spectral graph matching in pharyngeal surface registration. In: Medical Image Computing and Computer-Assisted Intervention, pp. 259–266 (2014)

[7] Rehman, H.Z.U., Lee, S.: Automatic image alignment using principal component analysis. IEEE Access **6**, 72063–72072 (2018)

[8] Wang, Y., Solomon, J.M.: Deep closest point: Learning representations for point cloud registration. In: IEEE/CVF International Conference on Computer Vision, pp. 3523–3532 (2019)

[9] Müller, M., Helljesen, L.E.S., Prevost, R., Viola, I., Nylund, K., Gilja, O.H., Navab, N., Wein, W.: Deriving anatomical context from 4d ultrasound. In: VCBM, pp. 173–180 (2014)

[10] Gao, Y., Sandhu, R., Fichtinger, G., Tannenbaum, A.R.: A coupled global registration and segmentation framework with application to magnetic resonance prostate imagery. IEEE transactions on Medical Imaging **29**(10), 1781–1794 (2010)

[11] Leroy, A., Mozer, P., Payan, Y., Troccaz, J.: Rigid registration of freehand 3d ultrasound and ct-scan kidney images. In: Medical Image Computing and Computer-Assisted Intervention, pp. 837–844 (2004)

[12] Leroy, A., Mozer, P., Payan, Y., Richard, F., Chartier-Kastler, E., Troccaz, J.: Percutaneous renal puncture: requirements and preliminary results. arXiv preprint physics/0610209 (2006)

[13] Leroy, A., Mozer, P., Payan, Y., Troccaz, J.: Intensity-based registration of freehand 3d ultrasound and ct-scan images of the kidney. International journal of computer assisted radiology and surgery **2**(1), 31–41 (2007)

[14] Ndzimbong, W., Fourniol, C., Themyr, L., Thome, N., Keeza, Y., Sauer, B., Piechaud, P.-T., Mejean, A., Marescaux, J., George, D., et al.: Trusted: The paired 3d transabdominal ultrasound and ct human data for kidney segmentation and registration research. arXiv preprint arXiv:2310.12646 (2023)

[15] Xing, S., Cambranis-Romero, J., Roy, P., Cool, D., Tessier, D., Chen, E., Peters, T., Fenster, A.: 3d us-ct/mri registration for percutaneous focal liver tumor ablations. International journal of computer assisted radiology and surgery **18** (2023)

[16] Joutard, S., Pheiffer, T., Audigier, C., Wohlfahrt, P., Dorent, R., Piat, S., Vercauteren, T., Modat, M., Mansi, T.: A multi-organ point cloud registration algorithm for abdominal ct registration. In: International Workshop on Biomedical Image Registration, pp. 75–84 (2022)

[17] Payer, C., Štern, D., Bischof, H., Urschler, M.: Regressing heatmaps for multiple landmark localization using cnns. In: Medical Image Computing and Computer-Assisted Intervention, pp. 230–238 (2016)

[18] Wang, X., Yang, X., Dou, H., Li, S., Heng, P.-A., Ni, D.: Joint segmentation

and landmark localization of fetal femur in ultrasound volumes. In: IEEE EMBS International Conference on Biomedical & Health Informatics, pp. 1–5 (2019)

[19] Isensee, F., Jaeger, P.F., Kohl, S.A., Petersen, J., Maier-Hein, K.H.: nnu-net: a self-configuring method for deep learning-based biomedical image segmentation. Nature methods **18**(2), 203–211 (2021)

[20] Hirose, O.: A bayesian formulation of coherent point drift. IEEE transactions on pattern analysis and machine intelligence **43**(7), 2269–2286 (2021)

[21] Farnia, P., Ahmadian, A., Khoshnevisan, A., Jaberzadeh, A., Serej, N.D., Kazerooni, A.F.: An efficient point based registration of intra-operative ultrasound images with mr images for computation of brain shift; a phantom study. In: IEEE Engineering in Medicine and Biology Society, pp. 8074–8077 (2011)

Declarations

Funding

This study was funded by France's Région Grand Est under the Project Disrumpere.

Competing interests

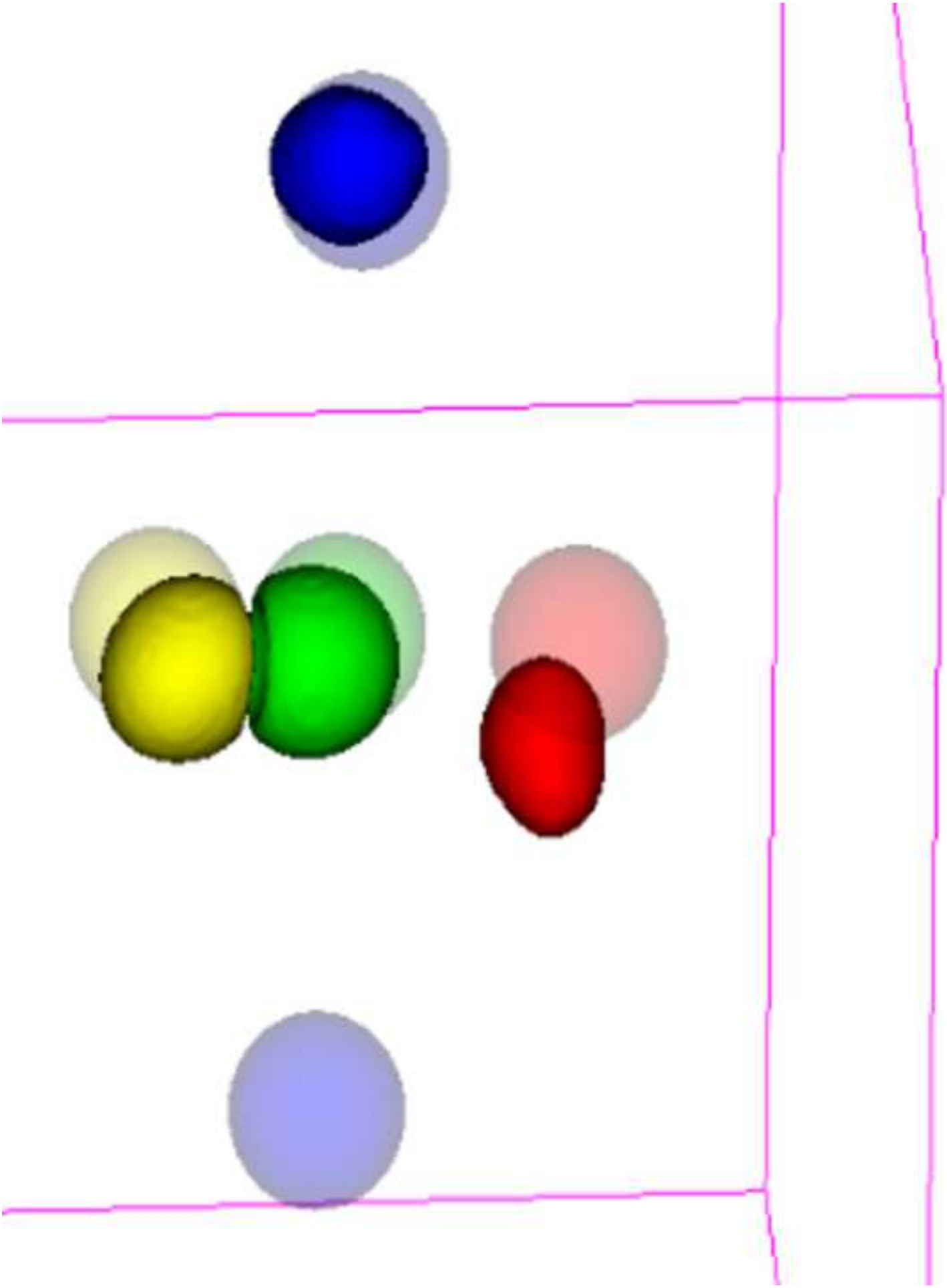
The authors have no relevant financial or non-financial interests to disclose.

Research involving human participants, their data or biological material

The anonymous US and CT image data used in this study was collected in a prior research study [14] (ID RCB: 2020-A01029-30/SI:20.05.01.539110), and ethically approved by the French government's "Comité de Protection des Personnes Sud - Mediterranée II".

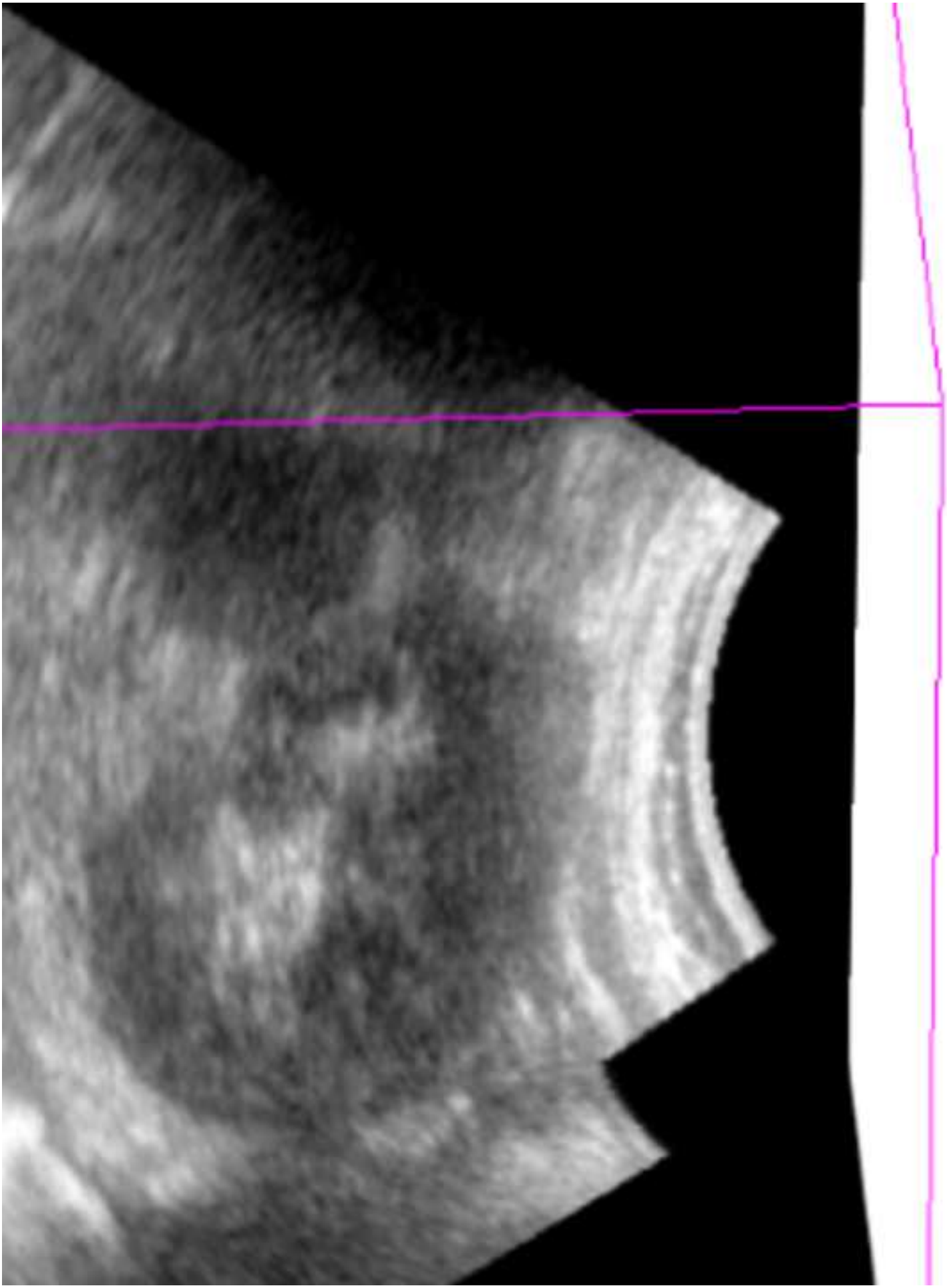
Figure

Click here to access/download;Figure;01Lusldks2.png



Figure

Click here to access/download;Figure;01Lusimg2.png



Journal specific style file