LCNet: A Robust and Accurate Non-Rigid 3D Point Set Registration Approach for Image-Guided Liver Surgery

Mingyang Liu, Zhe Min*, Geng Li, Hao Yu, Rui Song, Yibin Li, and Max Q.-H. Meng, *Fellow, IEEE*

***Abstract*— In this paper, we propose a novel unsupervised learning-based non-rigid 3D point set registration method, Learning Coherent Point Drift Network (LCNet), for image-guided liver surgery (IGLS). We reformulate the classical probabilistic registration approach, i.e., Coherent Point Drift (CPD) into a learning-based paradigm. We first establish reliable correspondences between two point sets using the optimal transport (OT) module by leveraging both original points and learnt features, which are robust to the rigid transformation. Then, rather than directly regressing displacement vectors, we compute the displacements by solving the involved matrix equation in the Transformation module, where the point localisation noise is explicitly considered. In addition, we present three variants of the proposed approach, i.e., LCNet, LCNet-ED and LCNet-WD, where LCNet outperforms the other two, which demonstrates the superiority of the Chamfer loss. We have extensively evaluated LCNet on MedShapeNet, consisting of 615 different liver shapes of real patients. For example, when the deformation magnitude is 12mm, the maximum noise magnitude is 2mm and the rotation angle lies in the range of $[-45°, 45°]$, LCNet achieves a root-mean-square error (rmse) value being 2.18 mm, outperforming CPD and RoITr, which are 6.95 mm ($p < 0.001$) and 2.69 mm ($p < 0.001$) respectively. Extensive experimental results, under different deformation and noise levels, show that LCNet exhibits significant improvements over existing state-of-the-art registration methods and shed light on its promising use in IGLS.**

Index Terms—Image-guided liver surgery, non-rigid point set registration, optimal transport, unsupervised learning scheme.

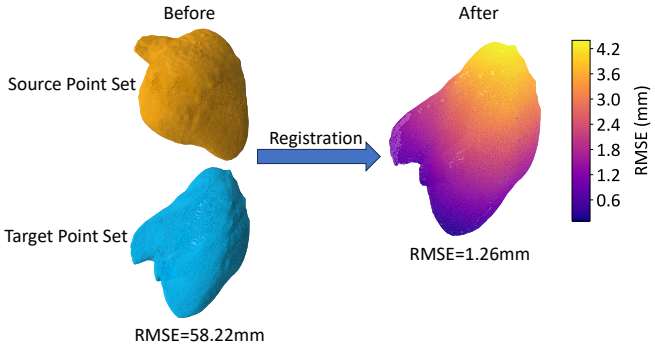


Fig. 1. The error distribution results of one pair of liver point sets using the proposed approach, i.e., Learning Coherent Point Drift Network (LCNet), where the yellow and blue meshes represent the preoperative (i.e., source) and intraoperative (i.e., target) liver shapes respectively. The organ's purple regions indicate lower error, whereas yellow regions have higher error.

I. Introduction

Liver cancer is a prevalent and highly fatal disease worldwide [1]. According to data from the Global Cancer Observatory (GLOBOCAN) 2022 [2], in 2022, the numbers of new liver cancer cases and deaths cases caused by liver cancer all over the world were 865,269 and 757,948 respectively. Surgical removal and ablation of the cancerous areas are common ways of treating diagnosed liver cancer. Image-guided liver surgery (IGLS) can enhance the interventional accuracy and reduce surgical risks significantly, while accurately registering the preoperative and intraoperative spaces serves as the prequisite for IGLS. During liver surgery, the liver organ is subjected to significant deformation caused by compression from other organs, respiratory movements, heartbeat, and other factors [3]. With successful image-to-patient registration, the preoperative surgical planning information can be mapped into the intraoperative space, which enables accurate real-time intraoperative navigation that assists surgeons make correct decisions to avoid damaging critical tissues. Image-to-patient registration in IGLS is often achieved by conducting non-rigidly registering preoperative and intraoperative point sets, which is a fundamental problems across various research fields such as shape matching [4], [5], surgical robotics [6], [7], medical image analysis [8], [9], computer-assisted intervention [10].

There exists several challenges to accurately register the preoperative and intraoperative liver point set. First, the acquired intraoperative liver point sets contain non-negligible localisation noise, which heavily affect the registration accuracy [11]. Second, the preoperative and intraoperative data points exhibit a significant difference in the coordinate reference frames (i.e., the CT and navigation frames), i.e., the rigid transformation including the rotation and translation, which makes finding the correct correspondences more difficult [12].

CPD [13] is one of the classical algorithms in the field of non-rigid registration, which has a wide range of application scenarios, but the convergence speed of CPD [13] is slow and it is easy to fall into local minimum. In order to solve these problems, we reformulate CPD into the paradigm of

*Corresponding author.

Mingyang Liu, Zhe Min, Rui Song and Yinbin Li are with the School of Control Science and Engineering, Shangdong University, Jinan, China. Zhe Min is also with Wellcome/EPSRC Centre for Surgical and Interventional Sciences, University College London, London, UK minzhe@sdu.edu.cn.

Geng Li is with the School of Electrical Engineering, Shandong University, Jinan, China.

Hao Yu is with the School of Information Science and Engineering, Shandong University, Jinan, China.

Max Q.-H. Meng is with the Department of Electronic and Electrical Engineering, the Southern University of Science and Technology, Shenzhen, China.

Manuscript 1990 submitted to 2025 IEEE International Conference on Robotics and Automation (ICRA). Received September 15, 2024.

deep learning. Unlike other deep learning methods based on CPD, which directly predicts the displacement field [14], our method explicitly considers the influence of noise in the step of solving the displacement field, thus enhancing the robustness to noise and improving the registration performance. Finally, in order to apply our method to image-guided liver surgery and solve the above-mentioned challenges, we tested it in these scenes, and the experimental results show that our method has good registration effect. Fig.1 shows registration results of one example case, where the rotational discrepancy between two point sets is 30° and the average deformation magnitude is 12 mm.

Our main contributions are summarised as follows.

- We propose a novel and unsupervised non-rigid point set registration method, i.e., LCNet, which consists of two stages that find the point correspondences with an OT module extracting rotation-invariant features and estimate the non-rigid transformation by solving the involved equation, respectively.
- To enhance the flexibility and applicability of the registration process under different scenarios, we have proposed and compared three variants of the proposed approach, which utilise three different loss functions that are Chamfer distance (i.e., LCNet), the weighted Euclidean Distance by the estimated correspondence probabilities (i.e., LCNet-ED), and the weighting of the above two distances(i.e., LCNet-WD).
- Our method achieves state-of-the-art performance, rivaling and even surpassing some fully-supervised models that rely on ground truth annotations on the datasets consisting of 615 different liver shapes of real patients.

II. Related Work

In this section, we recap the existing non-rigid 3D point set registration methods that can be broadly classified into traditional and deep learning-based categories, which are further classified into correspondence-based and correspondence-free registration approaches.

Traditional Methods The Iterative Closest Point (ICP) [15] algorithm based on the least square minimization method relies heavily on the initialization and tends to converge slowly. Coherent Point Drift (CPD) [13] formulates the registration task as the probability density function estimation problem, which is solved by updating the point correspondences and non-rigid transformation parameters iteratively, using the expectation-maximization technique. CPD has the disadvantages of heavy computational costs and the sensitivity to occlusion and outliers. Bayesian Coherent Point Drift (BCPD) [16] extends CPD by introducing variational Bayesian inference, which ensures theoretical convergence and is capable of handling the rigid transformation between two spaces simultaneously. Nevertheless, BCPD is susceptible to local minima [16]. CluReg [17] innovatively addresses non-rigid registration by reformulating it as an unsupervised clustering problem. It leverages cluster analysis to optimize point set correspondence and employs an improved Nyström method for efficient computational processing.

Correspondence-based Methods Correspondence-based methods first establish correspondences between two spaces, with which the non-rigid transformation is then computed. Lepard [18] first uses the fully convolutional feature extractor KPFCN [19] to extract global features from point sets and performed feature interaction through self-attention and cross-attention mechanisms to obtain point correspondences, and then utilises the non-rigid iterative closest point (N-ICP) method [20] to compute the non-rigid transformation. LiverMatch [21] utilises the encoder to extract features and realise the feature interaction based on a Transformer module, afterwards it leverages the decoder to compute the correspondence confidence matrix and point-wise displacement vectors. Nevertheless, LiverMatch [21] does not perform well with noisy target surfaces. SPFlowNet [22] updates superpoints using the bidirectional flow information, establishes matching relationships between points and superpoints, and then leverages correspondences to compute the scene flow accurately. GraphSCNet [23] calculates correspondences using a local spatial consistency measure on a deformation graph and refines them with an attention mechanism, afterwards the deformation is computed with the refined correspondences and the N-ICP algorithm [20]. RoITr [24] adopts a rotation-invariant attention mechanism based on Point Pair Features within an encoder-decoder framework to compute correspondences.

Correspondence-free Methods Correspondence-free registration approaches utilize neural networks to directly regress the displacement vectors. Pointpwc [25] employs a coarse-to-fine approach to refine the flow prediction, which improves efficiency and captures large-scale motion. As a deep neural network architecture for non-rigid point set registration, Free Point Transformer (FPT) [26] can complete the registration without known correspondence or heuristic constraints. Scoop [27] learns discriminative point features for soft correspondence initialization through a neural network, followed by a direct optimization of the displacement field at runtime for enhanced accuracy without reliance on ground-truth labels.

III. Our Method

A. Problem Setting

We define the source (i.e., preoperative) point set as $\mathbf{X} = \{\mathbf{x}_1, ...\mathbf{x}_i..., \mathbf{x}_M\}$ where $\mathbf{x}_i \in \mathbb{R}^3$ is the i-th point and $M \in \mathbb{N}^+$ is the number of points, the target (i.e., intraoperative) point set as $\mathbf{Y} = \{\mathbf{y}_1, ...\mathbf{y}_j..., \mathbf{y}_N\}$ where $\mathbf{y}_j \in \mathbb{R}^3$ and $N \in \mathbb{N}^+$. The goal of the non-rigid point set registration problem is to estimate the displacement vectors $\mathbf{D}_{\text{pred}} = [\mathbf{d}_1^{\text{pred}}, ...\mathbf{d}_i^{\text{pred}}..., \mathbf{d}_M^{\text{pred}}]^\mathsf{T} \in \mathbb{R}^{M\times 3}$ with $\mathbf{d}_i^{\text{pred}} \in \mathbb{R}^3$ denotes the displacement vector with $\mathbf{x}_i$, which warps $\mathbf{X}$ to $\mathbf{Y}$, so that the warped source point set $\mathbf{X}' = \{\mathbf{x}'_1, ...\mathbf{x}'_i \in \mathbb{R}^3..., \mathbf{x}'_M\}$ and $\mathbf{Y}$ become aligned. Note that we also reload symbols $\mathbf{X}$ and $\mathbf{Y}$ as matrices $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_M]^\mathsf{T} \in \mathbb{R}^{M\times 3}$ and $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_N]^\mathsf{T} \in \mathbb{R}^{N\times 3}$ respectively.

Manuscript 1990 submitted to 2025 IEEE International Conference on Robotics and Automation (ICRA). Received September 15, 2024.

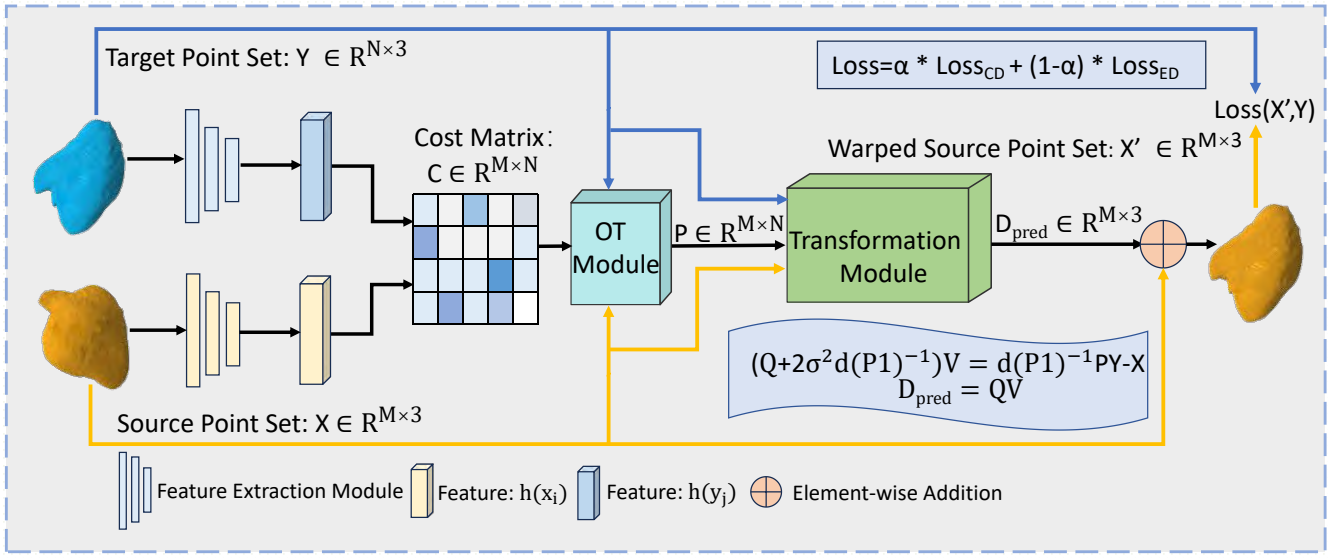


Fig. 2. The architecture of the proposed 3D point cloud registration network, i.e., Learning Coherent Point Drift Network (LCNet). The source and target point sets $\mathbf{X}$ and $\mathbf{Y}$, are fed into the feature extraction module (FEM) that is detailed in Sect. III-B.1. The correspondences $\mathbf{P} \in \mathbf{R}^{M\times N}$ are generated by the optimal transport (OT) module that is detailed in Sect. III-B.2. We solve the equation using estimated point correspondences to compute the displacement $\mathbf{D}_{\text{pred}} = \mathbf{QV}$ through the Transformation module that is detailed in Sect. III-B.3, where $\mathbf{Q}$ is the Gaussian kernel matrix, $\mathbf{V}$ is the weight matrix of the Gaussian kernel, $\oplus$ is denoted as element-wise addition. We have proposed and evaluated three training loss function that is detailed in Sect. III-B.4, where α is the weight of the Chamfer loss in the overall training loss.

B. Network Architecture

Fig. 2 illustrates the schematic of the proposed non-rigid point set registration network, i.e., Learning Coherent Point Drift Network (LCNet), which consists of three components: the feature extraction module (FEM), the optimal transport (OT) module and the Transformation module. As shown in Fig. 2, FEM first extracts features from both $\mathbf{X}$ and $\mathbf{Y}$, and then the OT module leverages the relevant feature information to calculate the optimal correspondences matrix, with which the displacement $\mathbf{D}_{\text{pred}}$ is obtained through the Transformation module.

1) Feature Extraction Module (FEM): FEM consists of three convolutional layers based on PointNet++ [28], and can learn features at different scales and levels. FEM consumes raw point sets $\mathbf{X}$ and $\mathbf{Y}$, and computes point-wise high-dimensional features $\mathbf{h}(\mathbf{x}_i) \in \mathbb{R}^{M\times 128}$ and $\mathbf{h}(\mathbf{y}_j) \in \mathbb{R}^{N\times 128}$, which together with raw point sets will be used in the following OT module.

Algorithm 1 Optimal Transport

Input: Transport cost matrix $\mathbf{C}$, hyperparameter λ, iteration number L_o
Output: Correspondences matrix $\mathbf{P}^\star$

$\mathbf{K} \leftarrow exp(-\mathbf{C}/\lambda)$;
$\boldsymbol{\theta}_x \leftarrow 1/M\mathbf{1}_M, \boldsymbol{\theta}_y \leftarrow 1/N\mathbf{1}_N, \mathbf{a} \leftarrow 1/M\mathbf{1}_M$;
for $l = 1$ **to** L_o **do**
  $\mathbf{b} \leftarrow \boldsymbol{\theta}_y/(\mathbf{K}^\mathsf{T}\mathbf{a})$;
  $\mathbf{a} \leftarrow \boldsymbol{\theta}_x(\mathbf{Kb})$;
end for
return $\mathbf{P}^\star \leftarrow \text{diag}(\mathbf{a})\mathbf{K}\text{diag}(\mathbf{b})$

2) OT Module: The OT module utilises Sinkhorn algorithm [29], [30] to solve the correspondences matrix $\mathbf{P}^\star \in \mathbb{R}^{M\times N}$. We transform the correspondence estimation problem into the optimal transport problem that seeks a transport plan that moves the distribution $\boldsymbol{\theta}_x \in \mathbb{R}^M$ of source point set $\mathbf{X}$ to the distribution $\boldsymbol{\theta}_y \in \mathbb{R}^N$ of target point set $\mathbf{Y}$ with a minimum transport cost. Assuming that each point in $\mathbf{X}$ has a mass $\frac{1}{M}$, each point in $\mathbf{Y}$ receives a mass $\frac{1}{N}$ and no mass is lost during the transportation, the optimal transport problem can be written as:

$$\begin{aligned} \mathbf{P}^\star &= \arg\min_{\mathbf{P}\in\mathbb{R}^{M\times N}} \sum_{ij} C_{ij}P_{ij} + \lambda P_{ij}(\log P_{ij} - 1), \\ \textbf{s.t.} \quad & \mathbf{P}\mathbf{1}_N = \boldsymbol{\theta}_x, \quad \mathbf{P}^\mathsf{T}\mathbf{1}_M = \boldsymbol{\theta}_y, \end{aligned} \tag{1}$$

where $\mathbf{P}^\star$ is the optimal correspondences matrix, C_{ij} is the transport cost of relating $\mathbf{x}_i$ with $\mathbf{y}_j$, P_{ij} is the amount of mass transported between $\mathbf{x}_i$ and $\mathbf{y}_j$, $\lambda \in \mathbb{R}$ is a hyperparameter, $\mathbf{1}_M \in \mathbb{R}^M$ and $\mathbf{1}_N \in \mathbb{R}^N$ are vectors with all elements being 1. We first compute the cosine similarity $S_{ij} \in \mathbb{R}$ between $\mathbf{x}_i$ and $\mathbf{y}_j$ by utilising extracted point features $\mathbf{h}(\mathbf{x}_i)$ and $\mathbf{h}(\mathbf{y}_j)$ computed in Sect. III-B.1 rather than simply taking the Euclidean distance between each pair of points $(\mathbf{x}_i, \mathbf{y}_j)$. The higher the similarity between each pair of points is, the smaller the transport cost between them is. Specifically, the similarity S_{ij} is computed by:

$$S_{ij} = \frac{\mathbf{h}(\mathbf{x}_i)^\mathsf{T}\mathbf{h}(\mathbf{y}_j)}{\|\mathbf{h}(\mathbf{x}_i)\|_2\|\mathbf{h}(\mathbf{y}_j)\|_2}, \tag{2}$$

where $||\star||_2$ denotes the L2 norm of a vector. Then, the transport cost C_{ij} is defined as:

$$C_{ij} = (1 - S_{ij}) \star W_{ij}, \tag{3}$$

where the conditional function $W_{ij} \in \mathbb{R}$ is used to avoid displacement between points too far apart and is defined as

$$W_{ij} = \begin{cases} 1 \text{ if } \|\mathbf{x}_i - \mathbf{y}_j\|_2 \leqslant l_{\max}, \\ +\infty \text{ otherwise} \end{cases}, \tag{4}$$

where the value of l_{max} is set to 10 cm. After defining the transport cost between two point sets, we use the Sinkhorn algorithm to estimate the optimal transport matrix $\mathbf{P}^\star$ which represents the soft correspondence between two point sets. As shown in Algorithm 1, the first step of the OT module is to compute the parameter $\mathbf{K} \in \mathbb{R}^{M\times N}$ that is the element-wise exponential of $-\mathbf{C}/\lambda \in \mathbb{R}^{M\times N}$ and initialise $\boldsymbol{\theta}_x$, $\boldsymbol{\theta}_y$ as $\frac{1}{M}\mathbf{1}_M$ and $\frac{1}{N}\mathbf{1}_N$ respectively. Two non-negative vectors uniquely defined up to a multiplicative factor, i.e., $\mathbf{a} \in \mathbb{R}^M$ and $\mathbf{b} \in \mathbb{R}^N$, are used in the process, where $\mathbf{a}$ is initialised to be $\frac{1}{M}\mathbf{1}_M$. Then the best transport plan is finally obtained as $\mathbf{P}^\star = \text{diag}(\mathbf{a})\mathbf{K}\text{diag}(\mathbf{b})$ where here $\mathbf{a}$ and $\mathbf{b}$ denote their values in the last iteration and diag$(\star)$ converts a vector into a diagonal matrix.

3) Transformation Module: Given the point correspondences $\mathbf{P}$, the Transformation module detailed in the Algorithm 2 is used to compute the displacement $\mathbf{D}_{\text{pred}}$. In the first train epoch, we define the Guassian kernel matrix $\mathbf{Q} \in \mathbb{R}^{M\times M}$ and localisation noise uncertainty σ^2, which are initialised as $Q_{ij} = exp(-1/2\beta \|\mathbf{x}_i - \mathbf{x}_j\|^2)$, and $\sigma^2 = \frac{1}{3MN}\sum_{i,j=1}^{M,N} \|\mathbf{x}_i - \mathbf{y}_j\|_2^2$ respectively. The parameter $\beta \in \mathbb{R}$ in Q_{ij} defines the width of the smoothing Gaussian filter, which is set as β = 2. Then we compute $\mathbf{V} \in \mathbb{R}^{M\times 3}$ by solving the matrix equation:

$$(\mathbf{Q} + 2\sigma^2 \text{d}(\mathbf{P1})^{-1})\mathbf{V} = \text{d}(\mathbf{P1})^{-1}\mathbf{PY} - \mathbf{X}. \tag{5}$$

where d$(\star)$ is short for diag$(\star)$, $(\star)^{-1}$ denotes the inverse of a matrix. Finally, we obtain the displacement $\mathbf{D}_{\text{pred}} = \mathbf{QV}$ and update the localisation noise uncertainty σ^2. In next epoch, the updated parameter σ^2 will be used in the Eq. (5) to get more accurate displacement $\mathbf{D}_{\text{pred}}$. The warped source point set is then obtained as $\mathbf{X}' = \mathbf{X} + \mathbf{D}_{\text{pred}}$.

Algorithm 2 Transformation Module

Input: Correspondences matrix $\mathbf{P}$
Output: Predicted Displacement $\mathbf{D}_{\text{pred}}$
1: **Initialize** $\sigma^2 = \frac{1}{3MN}\sum_{i,j=1}^{M,N} \|\mathbf{x}_i - \mathbf{y}_j\|^2$, $Q_{ij} = \exp(-1/2\beta \|\mathbf{x}_i - \mathbf{x}_j\|^2)$
2: Solve $(\mathbf{Q} + \epsilon\sigma^2 d(\mathbf{P1})^{-1})\mathbf{V} = d(\mathbf{P1})^{-1}\mathbf{PY} - \mathbf{X}$
3: $U = \mathbf{1}^T\mathbf{P1}, \mathbf{X}' = \mathbf{X} + \mathbf{QV}$
4: $\sigma^2 = \frac{1}{3U}(\text{tr}(\mathbf{Y}^T\text{d}(\mathbf{P}^T\mathbf{1})\mathbf{Y}) - 2\text{tr}((\mathbf{PY})^T\mathbf{X}') + \text{tr}((\mathbf{X}')^T\text{d}(\mathbf{P1})\mathbf{X}'))$
5: **return** $\mathbf{D}_{\text{pred}} \leftarrow \mathbf{QV}$

4) Loss Function: We utilise three unsupervised loss functions for training, i.e., $\mathcal{L}_{\text{CD}}$, $\mathcal{L}_{\text{ED}}$ and $\mathcal{L}_{\text{WD}}$, which directly leads to three variants of the proposed registration approach, i.e., LCNet (or LCNet-CD), LCNet-ED, LCNet-WD. First, the Chamfer loss $\mathcal{L}_{\text{CD}}$ in LCNet is

$$\mathcal{L}_{\text{CD}} = \frac{1}{M}\sum_{\mathbf{x}'_i \in \mathbf{X}'} \min_{\mathbf{y}_j \in \mathbf{Y}} ||\mathbf{x}'_i - \mathbf{y}_j||_2 + \frac{1}{N}\sum_{\mathbf{y}_j \in \mathbf{Y}} \min_{\mathbf{x}'_i \in \mathbf{X}'} ||\mathbf{x}'_i - \mathbf{y}_j||_2. \tag{6}$$

Second, $\mathcal{L}_{\text{ED}}$ in LCNet-ED computes the average weighted Euclidean distance between $\mathbf{X}'$ and $\mathbf{Y}$ by the correspondence probability

$$\mathcal{L}_{\text{ED}} = \sum_{i=1}^{M}\sum_{j=1}^{N} P_{ij}^\star \|\mathbf{x}'_i - \mathbf{y}_j\|_2 . \tag{7}$$

Finally, the combined loss function $\mathcal{L}_{\text{WD}}$ in LCNet-WD is

$$\mathcal{L}_{\text{WD}} = \alpha \star \mathcal{L}_{\text{CD}} + (1-\alpha) \star \mathcal{L}_{\text{ED}}, \tag{8}$$

where $\alpha \in \mathbb{R}$ is a balancing factor. We select the best value of α accoridng to the registration performances.

IV. Experiments and Results

A. Dataset and Evaluation Metrics

The utilised datasets are generated by selecting 615 different liver shapes of real patients from the MedShapeNet medical dataset [31], which consists of more than 10,000 medical shapes, including liver, kidney, brain, bone and surgical instruments, etc.. These liver shapes were extracted from 1200 imaging data (CT or MRI) of more than 600 real patients including healthy as well as pathological subjects. The dataset is further divided into 487 training, 64 validation and 64 test samples respectively. The intraoperative liver point sets is obtained through deformation processes using the thin plate spine (TPS) [32] method. Three datasets are generated with different deformation magnitudes whose deformation values are about 8 mm, 12 mm and 17 mm respectively. The preoperative point set is further rotated using angles sampled randomly in the range of $[-45°, 45°]$ and different magnitudes of noise (i.e., 1 mm, 2 mm and 3 mm) are further injected to create the pairs of point set $\mathbf{X}$ and $\mathbf{Y}$.

In order to evaluate registration performances, we utilise three different evaluation metrics, i.e., mean absolute error (MAE), root mean square error (RMSE) and Chamfer distance (CD), which are defined as follows,

$$\begin{cases} \text{MAE} = \frac{1}{M}\sum_{i=1}^{M} ||\mathbf{x}'_i - \mathbf{x}_i^{\text{gt}}||_2 \\ \text{RMSE} = \sqrt{\frac{1}{M}\sum_{i=1}^{M} ||\mathbf{x}'_i - \mathbf{x}_i^{\text{gt}}||_2^2} \\ \text{CD} = \frac{1}{M}\sum_{\mathbf{x}'_i \in \mathbf{X}'} \min_{\mathbf{y}_j \in \mathbf{Y}} ||\mathbf{x}'_i - \mathbf{y}_j||_2 + \\ \quad \frac{1}{N}\sum_{\mathbf{y}_j \in \mathbf{Y}} \min_{\mathbf{x}'_i \in \mathbf{X}'} ||\mathbf{x}'_i - \mathbf{y}_j||_2 \end{cases} \tag{9}$$

where $\mathbf{x}_i^{\text{gt}} = \mathbf{x}_i + \mathbf{d}_i^{\text{gt}}$, $\mathbf{d}_i^{\text{gt}} \in \mathbb{R}^3$ represents the ground-truth of displacement, $||\star||_2$ denotes the L2 norm of a vector.

B. Implementation Details

We compare the proposed LCNet with traditional registration methods including CPD [13], BCPD [16],CluReg [17] and learning-based algorithms including Scoop [27], Pointpwc [25], SPFlowNet [22], FPT [26], Lepard [18], Livermatch [21] and RoITr [24]. Our method was implemented using the Pytorch framework on one single GPU (i.e., Nvidia GeForce RTX 4090, 24GB). The model is trained on M = 1024 points, randomly sampled from original two point sets that consists of 10,000 points. The Adam optimizer was used,

TABLE I

QUANTITATIVE RESULTS OF DIFFERENT METHODS AT DIFFERENT EXPERIMENT CONDITIONS. THE BEST RESULTS ARE DISPLAYED IN BOLD.

Method	Case1				Case2			
	MAE (mm)	RMSE (mm)	CD (mm)	p-value	MAE (mm)	RMSE (mm)	CD (mm)	p-value
CPD [13]	4.86±3.68	6.95±4.94	4.43±1.29	9.90E-06	5.81±4.11	7.65±5.49	4.98±1.24	9.27E-08
BCPD [16]	20.18±10.34	25.02±12.57	20.00±11.37	2.85E-06	18.64±5.94	23.06±6.82	17.72±10.35	8.12E-28
PointPWC-Net [25]	11.34±3.66	23.44±7.86	9.37±1.50	2.65E-06	11.84±4.09	24.38±8.86	12.53±2.83	2.23E-24
FPT [26]	8.36±2.68	16.95±5.73	8.99±1.66	2.15E-06	14.56±6.10	30.72±13.66	12.19±3.46	9.37E-21
Lepard [18]	4.23 ±0.83	8.44±1.65	6.10±0.81	6.61E-36	6.01 ±2.92	12.19±6.26	6.98±1.44	1.38E-12
Scoop [27]	1.67±1.32	3.37±2.70	2.30±0.83	6.34E-06	2.90±0.57	5.83±1.15	3.88±0.37	2.90E-5
SPFlowNet [22]	6.58±1.72	13.28±3.65	7.61±1.10	1.69E-06	6.45±1.57	12.99±3.31	7.59±1.33	1.20E-31
Livermatch [21]	9.60±3.14	22.09±12.45	10.61±5.69	4.08E-12	11.18±4.09	27.01±15.09	12.90±6.81	2.26E-16
RoITr [24]	**1.34±0.09**	2.69±0.18	2.20±0.12	3.61E-06	3.35±0.61	6.71±1.24	4.24±0.47	4.56E-21
CluReg [17]	9.38±2.64	11.95±3.44	11.40±2.83	3.66E-25	9.74±2.97	12.44±4.08	11.73±3.26	6.37E-20
LCNet-ED	2.56±1.74	2.73±1.86	2.55±1.06	1.65E-04	4.58±1.26	4.77±1.32	4.55±1.67	2.17E-05
LCNet-WD	2.11±1.10	2.21±1.23	2.12±0.68	0.71	3.80±0.86	3.84±0.93	3.56±1.14	0.06
LCNet	2.06±0.62	**2.18±0.94**	**1.99±0.37**	/	**2.21±1.18**	**3.46±0.81**	**3.12±0.62**	/

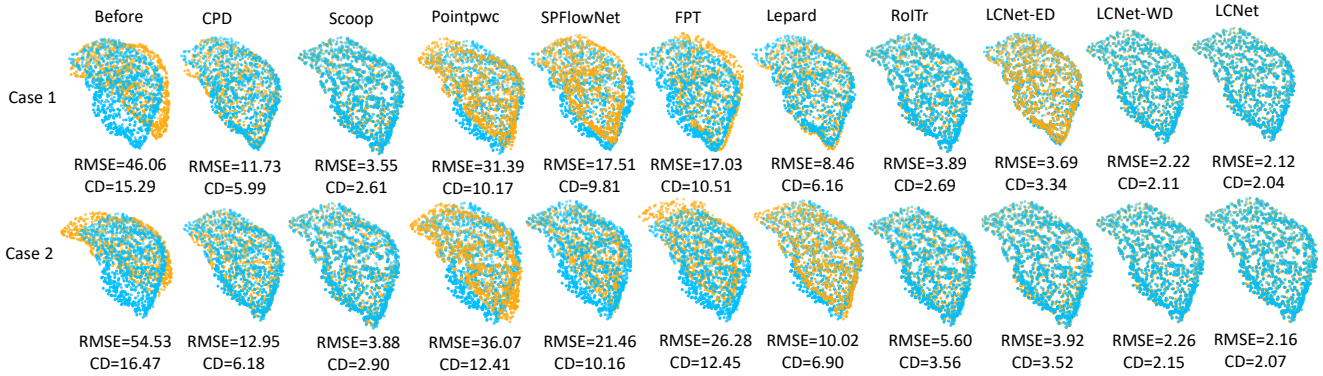


Fig. 3. Qualitative registration results on dataset under different experiment conditions. The yellow and blue point sets represent the preoperative and intraoperative liver shapes respectively. For Case 1, the deformation magnitude is 12 mm, the rotation angle is 32° and the maximum noise magnitude is 2 mm. For Case 2, the rotation angle, deformation magnitude and the maximum noise magnitude are the same as those in Case 1, but there is still a large translation between two points sets, and the range of translation is [20,30] mm.

the batch size was set to be 1, and the network was trained for a total of 2600 epochs. For the compared approaches, we use their publicly released codes. Our approach and other comparative methods have been uniformly trained on the same training dataset and subsequently evaluated on the same test dataset to ensure a fair comparative analysis.

C. Experiment Results

1) Comparions with Other Methods: Table I displays the corresponding quantitative results, including MAE, RMSE and CD as defined in Sect. IV-A. In Table I, Case1 illustrates the experimental scenario where only rigid rotation occurs between two point sets, whereas Case2 depicts the experimental conditions where, in addition to rigid rotation, there is also large displacement between the two point sets. LCNet achieves the rmse value being 2.18 ± 0.94 mm under Case1, which are better than counterparts with CPD [13] and Scoop [27] being 6.95 ± 4.94 mm ($p < 0.001$) and 3.37 ± 2.70 mm ($p < 0.001$) respectively. The rmse value of LCNet is 3.46 ± 0.81 mm under Case2, marginally outperforming RoITr [24] and Scoop [27] whose rmse values are 6.71 ± 1.24 mm ($p < 0.001$) and 5.83 ± 1.15 mm ($p < 0.001$), respectively. In addition, we have conducted the paired t-tests between the rmse values using LCNet and the compared methods. The small p-values ($p < 0.001$) in Table I indicate that LCNet and compared methods exist statistically significant difference. While LCNet and LCNet-ED are statistically significant different, the p-value being 0.71 and 0.06 show that the differences between LCNet and LCNet-WD are not statistically significant.

Fig. 3 shows the qualitative results using different methods, under scenarios of dataset with rotation. As shown in Fig.3, LCNet achieves effective registration for the example cases, while most of compared approaches fail to align the two point sets. For example, the rmse values using LCNet for Case 1 and Case 2 are 2.12 mm and 2.16 mm respectively, while the rmse values of Pointpwc [25] on two samples are 31.39 mm and 36.07 mm respectively.

2) Different Deformation Levels: In IGLS, the intraoperative liver point set often undergoes various levels of deformation due to factors such as respiration, heartbeat and tool-tissue interactions [33]. To explore the influence of the various deformation levels over the LCNet's performances, this sub-experiment compares LCNet and two variants, under three different deformation magnitudes as defined in Sect. IV-A. Specifically, three different deformation levels (low, medium and high) correspond to deformation magnitudes being 8 mm, 12 mm and 17 mm respectively. Fig. 4 (a)

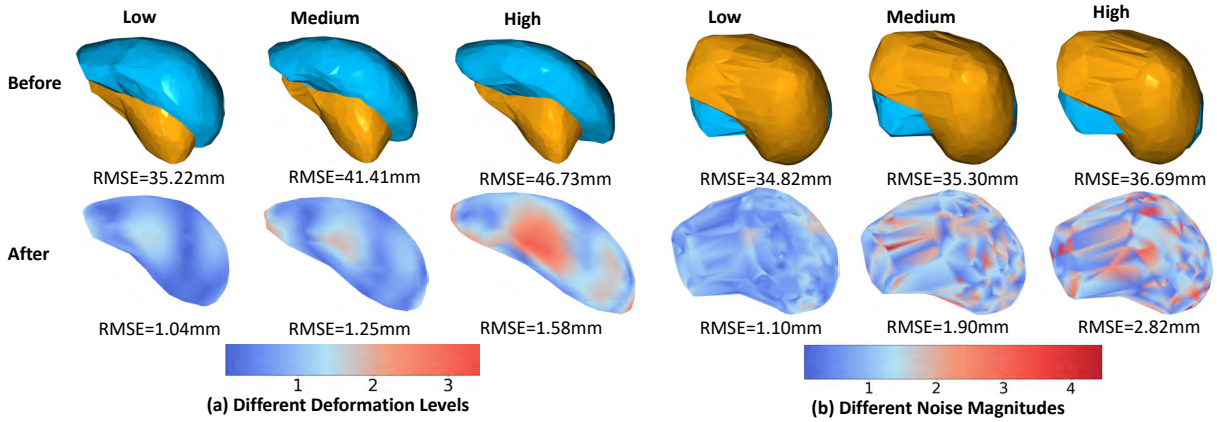
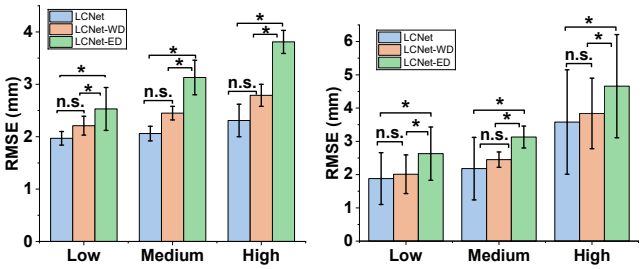


Fig. 4. The error distribution results of point sets (with rotation and with noise being 2 mm) using LCNet under different deformation levels and different noise magnitudes. The yellow and blue meshes represent the preoperative and intraoperative liver shapes respectively. The low, medium and high deformation levels correspond to cases whose deformation magnitude are 8 mm, 12 mm and 17 mm respectively. The low, medium and high noise magnitudes correspond to cases whose maximum noise magnitudes are 1 mm, 2 mm and 3 mm, respectively. The organ's blue regions indicate lower error, whereas red regions have higher error.



(a) Different Deformation Levels (b) Different Noise Magnitudes

Fig. 5. The boxplot of the RMSE under different deformation conditions and different noise conditions. The low, medium and high deformation levels correspond to cases whose deformation magnitudes are 8 mm, 12 mm and 17 mm respectively. The low, medium and high noise magnitudes correspond to cases whose maximum noise magnitudes are 1 mm, 2 mm and 3 mm, respectively. n.s.: no significant, ⋆: p<0.001.

and Fig. 5 (a) show the qualitative and quantitative results respectively. Fig. 4 (a) shows that LCNet is able to successfully register the two point sets, with the rmse being 1.58 mm, even under high deformation magnitude being 17 mm. Two observations can be made from Fig. 5 (a). First, LCNet consistently outperforms both LCNet-WD and LCNet-ED under different deformation magnitudes. Second, the registration error values of all three variants generally increase with larger deformation magnitudes but are still within the acceptable range. For example, LCNet, LCNet-WD and LCNet-ED achieve rmse values being 2.31 ± 0.31 mm, 2.79 ± 0.21 mm and 3.81 ± 0.22 mm respectively, under the high deformation level.

3) Different Noise Levels: This sub-experiment aims to evaluate LCNet, LCNet-WD and LCNet-CD under different magnitudes of noise that are defined in Sect. IV-A. Fig. 4 (b) and Fig. 5 (b) show the corresponding qualitative and quantitative results. Fig. 4 (b) shows that LCNet is able to successfully register the two point sets, with the rmse value being 2.82 mm, even under high noise magnitudes being 3 mm. Two observations can be made from Fig.5 (b). First, LCNet consistently outperforms both LCNet-WD and LCNet-ED under the three various noise magnitudes. Second, the registration error values of all three variants generally increase with larger noise levels but are still within the acceptable range. For example, LCNet, LCNet-WD and LCNet-ED achieve rmse values being 3.58 ± 1.57 mm, 3.84 ± 1.06 mm and 4.66 ± 1.55 mm respectively, under the high noise level.

V. Discussions and Conclusions

As shown in the experimental results in Sect. IV-C.1, it is evident that LCNet outperforms the traditional methods including CPD [13] and BCPD [16]. The improvements can be attributed to the utilisation of the OT module in LCNet, which enables more reliable point correspondence estimation through learning high-dimensional, robust and rotational-invariant features. In contrast, CPD [13] and BCPD [16] leverage only position information in point sets to compute the point correspondences. On the other hand, the improvements of LCNet and variants over learning-based methods can be attributed to the fact that correspondence-free methods rely on the direct regression of displacement vectors, which lacks theoretical explanation and support. Additionally, pure deep-learning-based registration methods tend to work better with larger data samples, while medical data sets are limited in real-world scenarios.

To conclude, this paper has introduced a novel unsupervised non-rigid point set registration approach, which casts the point correspondence estimation task as an optimal transport problem, and computes the displacement vectors by solving the involved mathematical equations numerically. Given the significant superior performances, we believe that the proposed LCNet approach has promising applications in IGLS, particularly in the challenging scenario of rotation, localisation noise and large deformations. Future work plans to incorporate the overlap-estimation module in the proposed registration method, aiming to tackle the challenging partial to full registration problem.

Manuscript 1990 submitted to 2025 IEEE International Conference on Robotics and Automation (ICRA). Received September 15, 2024.

REFERENCES

[1] H. Sung, J. Ferlay, R. L. Siegel, M. Laversanne, I. Soerjomataram, A. Jemal, and F. Bray, "Global cancer statistics 2020: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA: a cancer journal for clinicians*, vol. 71, no. 3, pp. 209–249, 2021.

[2] J. F. ME, R. L. Siegel, M. Isabelle Soerjomataram, and D. Ahmedin Jemal, "Global cancer statistics 2022: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries," 2024.

[3] J. A. Collins, J. A. Weis, J. S. Heiselman, L. W. Clements, A. L. Simpson, W. R. Jarnagin, and M. I. Miga, "Improving registration robustness for image-guided liver surgery in a novel human-to-phantom data framework," *IEEE transactions on medical imaging*, vol. 36, no. 7, pp. 1502–1510, 2017.

[4] A. Fan, J. Ma, X. Tian, X. Mei, and W. Liu, "Coherent point drift revisited for non-rigid shape matching and registration," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, pp. 1424–1434.

[5] X. Chen, N. Ravikumar, Y. Xia, R. Attar, A. Diaz-Pinto, S. K. Piechnik, S. Neubauer, S. E. Petersen, and A. F. Frangi, "Shape registration with learned deformations for 3d shape reconstruction from sparse and incomplete point clouds," *Medical image analysis*, vol. 74, p. 102228, 2021.

[6] Z. Min, J. Wang, J. Pan, and M. Q.-H. Meng, "Generalized 3-d point set registration with hybrid mixture models for computer-assisted orthopedic surgery: From isotropic to anisotropic positional error," *IEEE Transactions on Automation Science and Engineering*, vol. 18, no. 4, pp. 1679–1691, 2020.

[7] J.-X. Zhao, C. Li, H. Ren, M. Hao, L.-C. Zhang, and P.-F. Tang, "Evolution and current applications of robot-assisted fracture reduction: a comprehensive review," *Annals of biomedical engineering*, vol. 48, pp. 203–224, 2020.

[8] J. Luo, K. Kim, T. Delie, S. Esfahani, P. Heidari, B. Saboury, and Q. Li, "Multi-timepoint spect image registration for psma-targeted radiopharmaceutical therapy dosimetry," 2023.

[9] Y. Fu, Y. Lei, T. Wang, P. Patel, A. B. Jani, H. Mao, W. J. Curran, T. Liu, and X. Yang, "Biomechanically constrained non-rigid mr-trus prostate registration using deep learning based 3d point cloud matching," *Medical image analysis*, vol. 67, p. 101845, 2021.

[10] R. H. Taylor, N. Simaan, A. Menciassi, and G.-Z. Yang, "Surgical robotics and computer-integrated interventional medicine [scanning the issue]," *Proceedings of the IEEE*, vol. 110, no. 7, pp. 823–834, 2022.

[11] M. Jia and M. Kyan, "Improving intraoperative liver registration in image-guided surgery with learning-based reconstruction," in *ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2021, pp. 1230–1234.

[12] Z. Min, D. Zhu, H. Ren, and M. Q.-H. Meng, "Feature-guided nonrigid 3-d point set registration framework for image-guided liver surgery: From isotropic positional noise to anisotropic positional noise," *IEEE Transactions on Automation Science and Engineering*, vol. 18, no. 2, pp. 471–483, 2020.

[13] A. Myronenko and X. Song, "Point set registration: Coherent point drift," *IEEE transactions on pattern analysis and machine intelligence*, vol. 32, no. 12, pp. 2262–2275, 2010.

[14] L. Wang, X. Li, J. Chen, and Y. Fang, "Coherent point drift networks: Unsupervised learning of non-rigid point set registration," *arXiv preprint arXiv:1906.03039*, 2019.

[15] P. J. Besl and N. D. McKay, "Method for registration of 3-d shapes," in *Sensor fusion IV: control paradigms and data structures*, vol. 1611. Spie, 1992, pp. 586–606.

[16] O. Hirose, "A bayesian formulation of coherent point drift," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 43, no. 7, pp. 2269–2286, 2021.

[17] M. Zhao, J. Jiang, L. Ma, S. Xin, G. Meng, and D.-M. Yan, "Correspondence-free non-rigid point set registration using unsupervised clustering analysis," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2024, pp. 21 199–21 208.

[18] Y. Li and T. Harada, "Lepard: Learning partial point cloud matching in rigid and deformable scenes," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2022, pp. 5554–5564.

[19] H. Thomas, C. R. Qi, J.-E. Deschaud, B. Marcotegui, F. Goulette, and L. J. Guibas, "Kpconv: Flexible and deformable convolution for point clouds," in *Proceedings of the IEEE/CVF international conference on computer vision*, 2019, pp. 6411–6420.

[20] H. Li, R. W. Sumner, and M. Pauly, "Global correspondence optimization for non-rigid registration of depth scans," in *Computer graphics forum*, vol. 27, no. 5. Wiley Online Library, 2008, pp. 1421–1430.

[21] Z. Yang, R. Simon, and C. A. Linte, "Learning feature descriptors for pre-and intra-operative point cloud matching for laparoscopic liver registration," *International Journal of Computer Assisted Radiology and Surgery*, pp. 1–8, 2023.

[22] Y. Shen, L. Hui, J. Xie, and J. Yang, "Self-supervised 3d scene flow estimation guided by superpoints," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023, pp. 5271–5280.

[23] Z. Qin, H. Yu, C. Wang, Y. Peng, and K. Xu, "Deep graph-based spatial consistency for robust non-rigid point cloud registration," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023, pp. 5394–5403.

[24] H. Yu, Z. Qin, J. Hou, M. Saleh, D. Li, B. Busam, and S. Ilic, "Rotation-invariant transformer for point cloud matching," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2023, pp. 5384–5393.

[25] W. Wu, Z. Y. Wang, Z. Li, W. Liu, and L. Fuxin, "Pointpwc-net: Cost volume on point clouds for (self-) supervised scene flow estimation," in *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part V 16*. Springer, 2020, pp. 88–107.

[26] Z. M. Baum, Y. Hu, and D. C. Barratt, "Real-time multimodal image registration with partial intraoperative point-set data," *Medical image analysis*, vol. 74, p. 102231, 2021.

[27] I. Lang, D. Aiger, F. Cole, S. Avidan, and M. Rubinstein, "Scoop: Self-supervised correspondence and optimization-based scene flow," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023, pp. 5281–5290.

[28] C. R. Qi, L. Yi, H. Su, and L. J. Guibas, "Pointnet++: Deep hierarchical feature learning on point sets in a metric space," *Advances in neural information processing systems*, vol. 30, 2017.

[29] C. M. S. Distances, "Lightspeed computation of optimal transport," *Advances in neural information processing systems*, vol. 26, pp. 2292–2300, 2013.

[30] L. Chizat, G. Peyré, B. Schmitzer, and F.-X. Vialard, "Scaling algorithms for unbalanced transport problems," *arXiv preprint arXiv:1607.05816*, 2016.

[31] M. Reyes, "Medshapenet–a large-scale dataset of 3d medical shapes for computer vision," 2023.

[32] F. L. Bookstein and W. Jaynes, "Thin-plate splines and the analysis of biological shape," *Video tape*, vol. 21, 1990.

[33] J. S. Heiselman, W. R. Jarnagin, and M. I. Miga, "Intraoperative correction of liver deformation using sparse surface and vascular features via linearized iterative boundary reconstruction," *IEEE transactions on medical imaging*, vol. 39, no. 6, pp. 2223–2234, 2020.