Bidirectional Partial-to-Full Non-Rigid Point Set Registration with Non-Overlapping Filtering

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

***Abstract*— In this paper, we introduce Bidirectional Non-Overlapping Filtering Network (Bi-NOFNet), which registers the partial intraoperative point set with full preoperative point set for computer-assisted interventions (CAI). Our contributions are three-folds. First, Bi-NOFNet adopts customised feature extractor to extract distinctive features from both point sets, with which the per-point overlap mask is predicted and the overlapping region is segmented for the preoperative point set. Furthermore, we propose two methods to filter out the non-overlapping regions, at feature-level (i.e., Bi-NOFNet(Feature)) and point-level (i.e., Bi-NOFNet (Point)). For these two methods, we develop supervised registration strategy where the ground-truth overlap mask and displacement vectors are employed, and weakly-supervised registration strategies where only the ground-truth overlap mask is available. Additionally, to fully utilise the information in both space, we propose a bidirectional registration mechanism, which predicts the displacement vectors associated with the intraoperative point set (i.e., the *forward* way) and those warpping the preoperative point set (i.e., the *backward* way). Experiments have been conducted on the proposed DeformMedShapeNet dataset that contains 615 different liver shapes. Extensive results demonstrate that Bi-NOFNet performs well for partial-to-full registration tasks under various scenarios of noise, overlap ratios and deformation levels, outperforming existing non-rigid registration approaches. Code and data will be made publicly available after the review process.**

I. Introduction

Point Set Registration (PSR) is a critical and fundamental issue in fields of medical robotics [1], medical image analysis [2] and image-guided surgery [3]. The goal of surgical navigation is to assist surgeons intraoperatively by providing visualisations of structures of interest (e.g., tumors, vessels) and surgical plan (e.g., planned resection lines). Typically, a detailed information-rich preoperative computed tomography (CT) scan is conducted, which has to be accurately aligned with intraoperative scene to enable successful navigation. In image-guided liver surgery (IGLS), during surgery, soft tissue often undergoes non-rigid deformation due to factors like breathing and squeezing [4]. Non-rigid point set registration is thus needed to warp the preoperative model to match with the intraoperative data. However, the non-rigid registration task still faces several challenges: (1) the intraoperative data is usually noisy, and contains outliers. (2) intraoperative scenes are often only partially overlapping with the full preoperative model due to the limited camera view leading only a partial region of the organ surface visible. (3) the non-rigid deformation of intraoperative organs is extremely complex to accurately model.

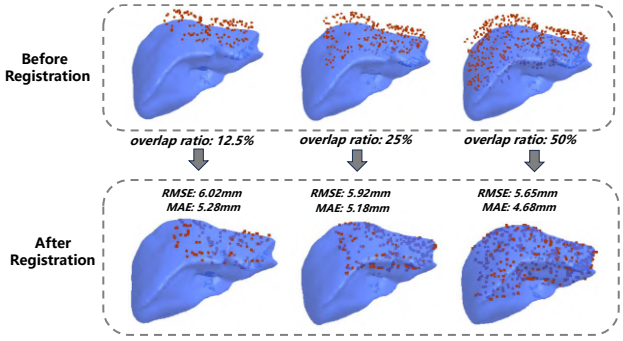

Fig. 1. The registration results using Bi-NOFNet under three different overlap ratios (i.e., 12.5%, 25%, 50%) between the intraoperative and preoperative point sets. The red points represent the partial intraoperative liver and the blue mesh represents the preoperative liver, respectively.

Recently, deep-learning-based methods have been proposed to solve the non-rigid point set registration problem [5][6][7][8][9][10]. However, most of them may not be suitable for the partial-to-full registration problem. To deal with the challenge of partial overlapping, Predator [11] introduces the concept of overlap ratio, based on which utilised points will be more sampled in the overlapping regions. OMNet [12] proposes an end-to-end model and utilises network prediction to generate an overlap mask, which is incorporated in the feature extraction phase to eliminate interference from non-overlapping regions. However, Predator and OMNet were designed specifically for the rigid point set registration problem.

In this paper, we propose the Bidirectional Non-Overlapping Filtering Registration Network (Bi-NOFNet) – an end-to-end non-rigid registration network that can adapt to input point sets with varying degrees of overlap. Bi-NOFNet is designed with three core modules, i.e., Feature Extractor Block, Overlap-Aware Block and Bidirectional Registration Block, combined through two non-overlapping filtering strategies. First, preoperative and intraoperative point sets are input into an feature extractor to learn features that are used

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

Zhe Min is with School of Control Science and Engineering, Shandong University, Jinan, China and Centre for Medical Image Computing and Wellcome/EPSRC Centre for Interventional & Surgical Sciences, University College London, London, United Kingdom z.min@ucl.ac.uk.

Mingyang Liu, Rui Song and Yibin Li are with School of Control Science and Engineering, Shandong University, Jinan, China.

Max Q.-H. Meng is with the Department of Electronic and Electrical Engineering, Southern University of Science and Technology, Shenzhen, China, and also with the Shenzhen Research Institute, The Chinese University of Hong Kong, Shenzhen, China, on leave from the Department of Electronic Engineering, Hong Kong, China.

*Corresponding Author.

Manuscript 427 submitted to 2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Received March 17, 2024.

for predicting the overlap mask. Second, using the overlap mask, we can filter out non-overlapping regions from the preoperative point set at feature-level or point-level, after which filtered features are learned through another feature extractor. Finally, we design a bidirectional registration mechanism which predicts both the displacement vectors associated with the intraoperative point set to match the preoperative point set (i.e., the *forward* way) and those warpping the preoperative point set to match the intraoperative point set (i.e., the *backward* way). Experimental results demonstrate Bi-NOFNet's superior performance on DeformMedShapeNet, significantly outperforming existing non-rigid registration approaches in handling partial-to-full registration tasks.

In summary, our contributions can be outlined as follows:

- We have constructed DeformMedShapeNet, a deformable liver registration dataset from the recently introduced large-scale public medical dataset MedShapeNet [13], which accounts for various levels of noise, overlap ratio and deformation to extensively evaluate the non-rigid registration algorithms for image-guided liver surgery .
- We have proposed the Bidirectional Non-Overlapping-Region Filtering Registration Network (Bi-NOFNet) for partial-to-full non-rigid point set registration, where we have successfully integrated the bidirectional registration mechanism in order to fully utilize the information of both point sets.
- We have proposed an Overlap-Aware module to predict the overlap mask that estimates the overlapping regions in the full preoperative point set compared to the partial intraoperative data, and carefully designated two approaches (i.e., at point-level and feature-level) to filter out non-overlapping regions under the supervised (i.e., with known GT overlap mask and point-wise displacement vectors) and weakly-supervised (i.e., with only known overlap mask) registration paradigms respectively.

II. Related Work

We review the relevant non-rigid point set registration approaches, which are categorized into traditional registration methods, correspondence-based registration methods and correspondence-free registration methods.

A. Traditional Registration Methods

The Non-rigid Iterative Closest Point (NICP) [14] is a classic non-rigid point cloud registration algorithm. It allows optimization through non-rigid deformation models, and ultimately formulates the energy function of template-to-target optimization as a sparse matrix least squares problem, achieving higher registration accuracy and robustness. Coherent Point Drift (CPD) [15] casts the registration task as a probability density estimation problem by representing one point set as Gaussian Mixture Model (GMM), where the coherent motion among neighboring points for registration is promoted. Very recently, BCPD reformulates both the rigid and non-rigid registration problems into the Bayesian framework, where the theoretical convergence is guaranteed [16]. The traditional registration methods mentioned above rely on manually selecting features or designing feature descriptors which makes extracted features susceptible to noise, and are usually slow in the optimization speed.

B. Correspondence-based Registration Methods (learned)

The rationale of methods in this category is that the learned features are first utilised to establish correspondences between source and target point set, with which the non-rigid transformation is computed with traditional non-rigid methods (e.g. NICP [17]). FLOT [6] and RobOT [18] formulates the correspondence estimation as the optimal transport problem. Lepard [19] adds 3D positional information to Transformer [20] and utilises dual-softmax for differentiable matching. The correspondence-based methods have high robustness and good generalization capability, but cannot predict the deformation vectors in an end-to-end manner and thus own limited running speed.

C. Correspondence-free Registration Methods (learned)

Correspondence-free registration methods directly predict displacement vectors between two point sets by leveraging learned features. FlowNet3D [5] introduces new learning layers based on Pointnet++ [21], one for learning the associative encoding of two frames of point clouds, and the other for propagating inter-frame point cloud features, thereby enabling it to predict displacement vectors from sequential point clouds. Very recently, Bi-PointFlowNet [9] and MSBRN [10] use a hierarchical architecture with bidirectional flow embeddings to further learn the associative information between the two point sets. NDP [22] utilize a pyramid to represent non-rigid motion and achieves better non-rigid registration results through hierarchical motion decomposition.Specifically, for surgical scenarios, FPT [7] utilizes PointNet [23] to learn high-dimensional features, and then predicts a non-rigid displacement field from the preoperative to the intraoperative point set using MLP. While [5][9][10][7][8] are designed for non-rigid registration tasks, they mainly focus on full-to-full registration problem, which limits their applications in real surgical scenarios. NDP [22] considers the low overlap of point sets, but it needs to be further combined with Lepard [19] to achieve good non-rigid registration of low-overlap point sets, making it not a correspondence-free method.

Due to the limitations of the aforementioned methods, we aim to design a correspondence-free method that can accurately and robustly align the full preoperative and partial intraoperative point sets.

III. Methods

As illustrated in Fig. 2, we propose two architectures of Bidirectional Non-Overlapping Filtering Network (Bi-NOFNet), i.e., Bi-NOFNet (Feature) and Bi-NOFNet (Point), according to the utilised strategy of filtering out non-overlapping regions in the preoperative point set. First, the

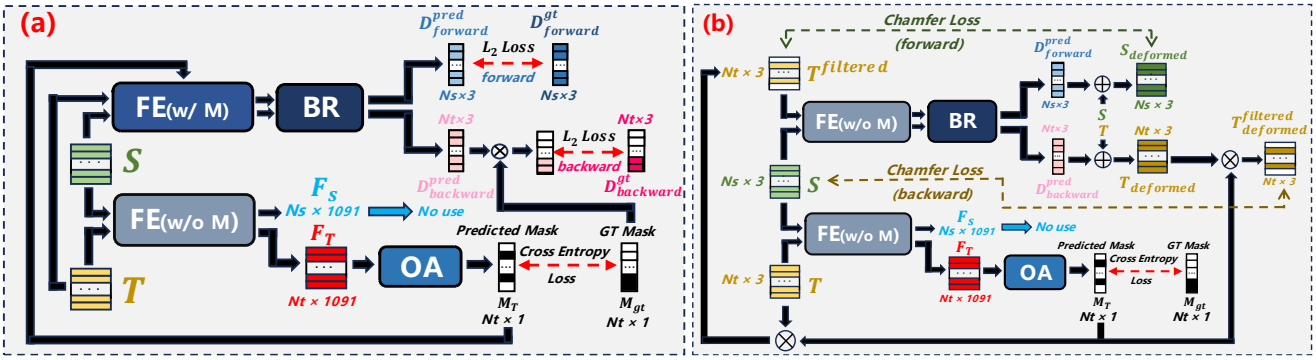

Fig. 2. Overview of the proposed approach. (a) and (b) illustrate the architectures of Bi-NOFNet (Feature) and Bi-NOFNet (Point), respectively. The inputs of our architectures are the source point set $\mathcal{S}$ and the target point set $\mathcal{T}$, and the outputs are $\mathbf{D}^{\text{pred}}_{\text{forward}}$ and $\mathbf{D}^{\text{pred}}_{\text{backward}}$. "FE (w/o M)" and "FE (w/ M)" represent FE-Block (without Mask), FE-Block (with Mask) detailed in Sec. III-B respectively, while "OA" and "BR" denote overlap-aware (OA) and bidirectional registration (BR) blocks detailed in Sec. III-C and Sec. III-E, respectively. In Bi-NOFNet (Feature), we conduct non-overlapping filtering operations at feature-level under the supervised registration paradigm whose forward and backward registration loss terms are defined in Eq. (13) and Eq. (14) respectively. In Bi-NOFNet (Point), we perform non-overlapping filtering operations at point-level under a weakly-supervised registration paradigm whose forward and backward registration loss terms are given in Eq. (19) and Eq. (20) respectively.

features are extracted from input point sets by Feature Extractor (FE) Block (cf. Sec. III-B). Then the learned features are fed into Overlap-Aware (OA) Block to segment the overlapping regions (cf. Sec. III-C) and we propose two ways of filtering out non-overlapping regions using the predicted overlap mask (cf. Sec. III-D). Meanwhile, a Bidirectional Registration (BR) Block that encompasses both forward and backward registration processes is used to estimate the non-rigid displacement vectors (cf. Sec. III-E). Finally, we discuss the loss functions used to optimize each module in Sec. III-F.

A. Non-Rigid Point Set Registration

Given the source point set (e.g., intraoperative) $\mathcal{S} \in \mathbb{R}^{N_s \times 3}$ and the target point set (e.g., preoperative) $\mathcal{T} \in \mathbb{R}^{N_t \times 3}$, where $N_s \in \mathbb{N}^+$ and $N_t \in \mathbb{N}^+$ are the number of points in the two spaces, the aim of the non-rigid point set registration (NRPSR) is to predict the point-wise displacement vectors $\mathbf{D}_{\mathcal{S}} \in \mathbb{R}^{N_s \times 3}$ from $\mathcal{S}$ to $\mathcal{T}$.

B. Feature Extractor (FE) Block

In the proposed approach, as shown in Fig. 2 (a) and Fig. 2 (b), the FE-Block is utilised twice. First, it is used to learn features $\mathbf{F}_{\mathcal{S}} \in \mathbb{R}^{N_s \times 1091}$ and $\mathbf{F}_{\mathcal{T}} \in \mathbb{R}^{N_t \times 1091}$ which are subsequently used for predicting the overlap mask in Sec. III-C. This FE-Block is named FE-Block without Mask and abbreviated as FE (w/o M). Second, it is also used to extract features $\mathbf{F}^{\text{filtered}}_{\mathcal{S}} \in \mathbb{R}^{N_s \times 1091}$ and $\mathbf{F}^{\text{filtered}}_{\mathcal{T}} \in \mathbb{R}^{N_t \times 1091}$ which are further used for predicting displacement vectors after filtering out non-overlapping regions. This FE-Block is named FE-Block (with Mask) and abbreviated as FE (w/ M). In this section, we mainly discuss how to learn initial features $\mathbf{F}_{\mathcal{S}}$ and $\mathbf{F}_{\mathcal{T}}$ with FE (w/o M).

OMNet [12] uses PointNet [23] to learn global features of both source and target point sets. However, we notice that there are two main issues with this way: (1) The local context information within each point set is not explicitly considered. (2) It does not specifically learn the associative relations between two point sets. To address these issues, we propose a new feature extractor, as illustrated in the Fig. 3. Similar to [7][8][12], we also employ PointNet [23] to learn global features $\mathbf{g}_{\mathcal{S}} \in \mathbb{R}^{1024}$, $\mathbf{g}_{\mathcal{T}} \in \mathbb{R}^{1024}$, which are then duplicated as $\mathbf{G}_{\mathcal{S}} \in \mathbb{R}^{N_s \times 1024}$, $\mathbf{G}_{\mathcal{T}} \in \mathbb{R}^{N_t \times 1024}$. The global features $\mathbf{G}_{\mathcal{S}}$, $\mathbf{G}_{\mathcal{T}}$ are further concatenate with local features $l_{\mathcal{S}} \in \mathbb{R}^{N_s \times 64}$, $l_{\mathcal{T}} \in \mathbb{R}^{N_t \times 64}$ (cf. $l_{\mathcal{S}}$ and $l_{\mathcal{T}}$ in Fig. 3) to obtain the fused features $\mathbf{f}_{\mathcal{S}} \in \mathbb{R}^{N_s \times 1088}$ and $\mathbf{f}_{\mathcal{T}} \in \mathbb{R}^{N_t \times 1088}$ as

$$\mathbf{f}_{\mathcal{S}} = \mathbf{G}_{\mathcal{S}} \oplus \mathbf{l}_{\mathcal{S}}, \tag{1}$$

$$\mathbf{f}_{\mathcal{T}} = \mathbf{G}_{\mathcal{T}} \oplus \mathbf{l}_{\mathcal{T}}, \tag{2}$$

where $\oplus$ denotes elementwise addition.

Unlike OMNet [12] that just uses the fused features for predicting overlap mask and displacement vectors, we feed the fused features $\mathbf{f}_{\mathcal{S}}$, $\mathbf{f}_{\mathcal{T}}$ into a transformer-based module [20] to learn more distinctive features. Then we concatenate the output features $\mathbf{C}_{\mathcal{S}} \in \mathbb{R}^{N_s \times 1088}$, $\mathbf{C}_{\mathcal{T}} \in \mathbb{R}^{N_t \times 1088}$ of the transformer with raw point sets $\mathcal{S}$ and $\mathcal{T}$:

$$\mathbf{F}_{\mathcal{S}} = \mathbf{C}_{\mathcal{S}} \oplus \mathcal{S}, \tag{3}$$

$$\mathbf{F}_{\mathcal{T}} = \mathbf{C}_{\mathcal{T}} \oplus \mathcal{T}. \tag{4}$$

By appending points' coordinate information to features, we aim to better guide the feature learning process especially during the initial training stage.

C. Overlap-Aware(OA) Block

After the FE (w/o M), we feed $\mathbf{F}_{\mathcal{T}}$ into the OA-Block, which outputs the per-point overlap mask for preoperative point set $\mathcal{T}$ (cf. Fig. 2). OA-Block processes the input features $\mathbf{F}_{\mathcal{T}}$ with MLP (1024, 512, 256, 128, 64, 1) to create a binary mask $\mathbf{M}_{\mathcal{T}} \in \mathbb{R}^{N_t \times 1}$. To ensure differentiability for backpropagation, the network first outputs a soft mask $\rho \in \mathbb{R}^{N_t \times 1}$ and then compute $\mathbf{M}_{\mathcal{T}}(i)$, for $i \in \{1, \ldots, N_t\}$ by applying a fixed threshold σ:

$$\mathbf{M}_{\mathcal{T}}(i) = \begin{cases} 1, & \text{if } \rho(i) \geq \sigma \\ 0, & \text{otherwise} \end{cases} \tag{5}$$

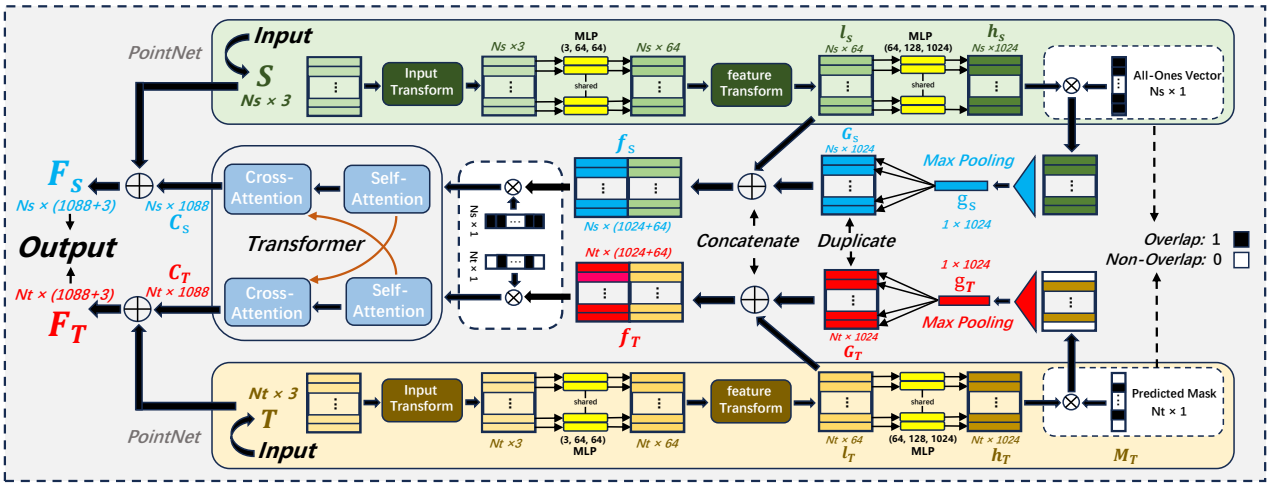

Fig. 3. The overall architecture of Feature Extractor (FE) Block. The FE-Block with and without the use of the overlap mask are labeled as FE (w/ M) and FE (w/o M) in Fig. 2, respectively. A detailed description of the FE-Block is found in Sec. III-B.

where σ is set to be 0.5 in practical implementation, $\mathbf{Z}(i)$ represents taking the i^{th} row of matrix $\mathbf{Z}$. On the other hand, we define the ground-truth overlap mask as $\mathbf{M_{gt}} \in \mathbb{R}^{N_t \times 1}$, and $\mathbf{M_{gt}}(i), i \in [1, \ldots, N_t]$ is:

$$M_{\mathbf{gt}}(i) = \begin{cases} 1 & \text{if } \mathcal{T}(i) \text{ is in the overlapping regions,} \\ 0 & \text{otherwise.} \end{cases} \tag{6}$$

D. Non-Overlapping Filtering Strategy

We propose two methods for filtering out non-overlapping regions of the preoperative point set using $\mathbf{M}_{\mathcal{T}}$, corresponding to the two architectures, i.e., Bi-NOFNet (Feature) and Bi-NOFNet (Point) in Fig. 2 (a) and Fig. 2 (b), respectively.

1) Feature-Level Filtering: In Bi-NOFNet (Feature), we consider that in the FE-Block, non-overlapping points mainly affect global features $\mathbf{g}_{\mathcal{S}}$ and $\mathbf{g}_{\mathcal{T}}$, and the context information, so we design FE (w/ M). For the preoperative point set $\mathcal{T}$, we multiply the output of the layer before max-pooling $\mathbf{h}_{\mathcal{T}} \in \mathbb{R}^{N_t \times 1024}$ by a predicted overlap mask $\mathbf{M}_{\mathcal{T}}$ (cf. Fig. 3), resulting the filtered global feature $\mathbf{g}_{\mathcal{T}}^{\text{filtered}} \in \mathbb{R}^{1024}$:

$$\mathbf{g}_{\mathcal{T}}^{\text{filtered}} = \max\{\mathbf{h}_{\mathcal{T}} \otimes \mathbf{M}_{\mathcal{T}}\}, \tag{7}$$

where max denotes the Max-Pooling Layer, $\otimes$ is the elementwise multiplication operator. By assuming that all points in $\mathcal{S}$ lie in the overlapping region with $\mathcal{T}$, which is reasonable for the image-to-patient registration, we simply multiply the feature $\mathbf{h}_{\mathcal{S}} \in \mathbb{R}^{N_s \times 1024}$ by a all-ones vector $\mathbf{1}_{N_s} \in \mathbb{R}^{N_s \times 1}$ to get the global feature $\mathbf{g}_{\mathcal{S}}^{\text{filtered}} \in \mathbb{R}^{1024}$:

$$\mathbf{g}_{\mathcal{S}}^{\text{filtered}} = \max\{\mathbf{h}_{\mathcal{S}} \otimes \mathbf{1}_{N_s}\}. \tag{8}$$

Additionally, since the transformer calculates the interrelations between all points, it is also necessary to multiply the fused feature $\mathbf{f}_{\mathcal{T}}$ and $\mathbf{f}_{\mathcal{S}}$ by the overlap mask $M_{\mathcal{T}}$ and $\mathbf{1}_{N_s}$ respectively:

$$\mathbf{f}_{\mathcal{T}}^{\text{filtered}} = \mathbf{f}_{\mathcal{T}} \otimes \mathbf{M}_{\mathcal{T}}, \tag{9}$$

$$\mathbf{f}_{\mathcal{S}}^{\text{filtered}} = \mathbf{f}_{\mathcal{S}} \otimes \mathbf{1}_{N_s}, \tag{10}$$

which are then inputted into a transformer module, whose outputs are then concatenated with raw point sets $\mathcal{S}$ and $\mathcal{T}$ to obtain filter features $\mathbf{F}_{\mathcal{S}}^{\text{filtered}}$ and $\mathbf{F}_{\mathcal{T}}^{\text{filtered}}$ (cf. Sec. III-B). The filtered features $\mathbf{F}_{\mathcal{S}}^{\text{filtered}}$ and $\mathbf{F}_{\mathcal{T}}^{\text{filtered}}$, as shown in Fig. 2 (a), will be leveraged by the BR-Block in Sec. III-E to predict displacement vectors.

2) Point-Level Filtering: In Bi-NOFNet (Point), we conduct the non-overlapping filtering before the FE-Block. As shown in Fig. 2 (b), we firstly input $\mathcal{S}$ and $\mathcal{T}$ into the lower FE (w/o M) to obtain initial features $\mathbf{F}_{\mathcal{S}}$, $\mathbf{F}_{\mathcal{T}}$, which are then inputted into the OA-Block in Sec. III-C to obtain a mask $\mathbf{M}_{\mathcal{T}}$ for $\mathcal{T}$. We then directly multiply $\mathbf{M}_{\mathcal{T}}$ with $\mathcal{T}$, which removes points predicted by the OA-Block that are not within the overlapping regions:

$$\mathcal{T}^{\text{filtered}} = \mathcal{T} \otimes \mathbf{M}_{\mathcal{T}}, \tag{11}$$

where $\mathcal{T}^{\text{filtered}} \in \mathbb{R}^{N_t \times 3}$, $\mathbf{M}_{\mathcal{T}}$ is computed in Sec. III-C.

Subsequently, as shown in Fig. 2 (b), we further process $\mathcal{T}_{\text{filtered}}$ and S with the upper FE (w/o M) to obtain filtered features $\mathbf{F}_{\mathcal{S}}^{\text{filtered}}$ and $\mathbf{F}_{\mathcal{T}}^{\text{filtered}}$, which will be utilised by the BR-Block in Sec. III-E to predict displacement vectors. Here, the non-overlapping regions are directly filtered out at point-level, which prevents non-overlapping regions from affecting feature extraction steps.

E. Bidirectional Registration (BR) Block

The BR-Block consists of MLP (1024, 512, 256, 128, 64, 3), whose inputs are the filtered features $\mathbf{F}_{\mathcal{S}}^{\text{filtered}}$ and $\mathbf{F}_{\mathcal{T}}^{\text{filtered}}$ computed in Sec. III-D, and outputs are predicted displacement vectors in two directions, i.e., forward displacement vectors $\mathbf{D}_{\text{forward}}^{\text{pred}} \in \mathbb{R}^{N_s \times 3}$ from $\mathcal{S}$ to $\mathcal{T}$ and backward displacement vectors $\mathbf{D}_{\text{backward}}^{\text{pred}} \in \mathbb{R}^{N_t \times 3}$ from $\mathcal{T}$ to $\mathcal{S}$ (cf. Fig. 2). The bidirectional registration mechanism allows us to comprehensively learn information from both point sets, thereby improving the accuracy of predicting non-rigid displacement vectors.

F. Loss Function

The training loss contains two terms, which are used to optimise the predictions of the overlap mask and the non-rigid displacement vectors, respectively.

1) Mask Prediction Loss: To balance the contributions of points within and outside the overlapping regions, the weighted cross-entropy loss is utilised for training the OA-Block in Sec. III-C,

$$L_{\text{mask}} = -\alpha M_{\text{gt}} \log M_{\mathcal{T}} - (1-\alpha)(1-M_{gt}) \log(1-M_{\mathcal{T}}), \tag{12}$$

where $(1-\alpha) \in \mathbb{R}$ is the overlap ratio of the inputs, $M_{\text{gt}} \in \mathbb{R}^{N_t \times 1}$ and $M_{\mathcal{T}} \in \mathbb{R}^{N_t \times 1}$ are the ground-truth and predicted overlap masks for $\mathcal{T}$ respectively.

2) Bidirectional Registration Loss: During the bidirectional registration process, we utilise different loss functions corresponding to the two non-overlapping filtering strategies in Sec. III-D.

a) L_2 Distance Loss: In Bi-NOFNet (Feature), L_2 distance loss is defined as the distance between predicted and ground-truth non-rigid displacement vectors. The L_2 Distance Loss $L^{L_2}_{\text{forward}}$ in the forward registration process is:

$$L^{L_2}_{\text{forward}} = \frac{\sum_{i=1}^{N_s} \|\mathbf{D}^{\text{gt}}_{\text{forward}}(i) - \mathbf{D}^{\text{pred}}_{\text{forward}}(i)\|}{N_s}, \tag{13}$$

where $\mathbf{D}^{\text{gt}}_{\text{forward}} \in \mathbb{R}^{N_s \times 3}$, $\mathbf{D}^{\text{pred}}_{\text{forward}} \in \mathbb{R}^{N_s \times 3}$ are the ground-truth and predicted *forward* displacement vectors respectively. The L_2 Distance Loss $L^{L_2}_{\text{backward}}$ with the backward registration process is:

$$L^{L_2}_{\text{backward}} = \frac{\sum_{i=1}^{N_s} \|\mathbf{D}^{\text{gt}}_{\text{backward}}(i) - \mathbf{D}^{\text{pred}}_{\text{backward}}(i)\|}{N_t}, \tag{14}$$

where $\mathbf{D}^{\text{gt}}_{\text{backward}} \in \mathbb{R}^{N_s \times 3}$, $\mathbf{D}^{\text{pred}}_{\text{backward}} \in \mathbb{R}^{N_t \times 3}$ are the ground-truth and predicted *backward* displacement vectors respectively. Since $\mathbf{D}^{\text{gt}}_{\text{backward}}$ only contains N_s points, all of which are in the overlapping regions though $\mathbf{D}^{\text{pred}}_{\text{backward}}$ contains N_t points, we only need to consider N_s points in the overlapping region when computing L_{backward}.

b) Chamfer Distance Loss: In Bi-NOFNet (Point), the warped source and the target point sets, i.e., $\mathcal{S}_{\text{deformed}} \in \mathbb{R}^{N_s \times 3}$ and $\mathcal{T}_{\text{deformed}} \in \mathbb{R}^{N_t \times 3}$, are defined as,

$$\mathcal{S}_{\text{deformed}} = \mathcal{S} + \mathbf{D}^{\text{pred}}_{\text{forward}}, \tag{15}$$

$$\mathcal{T}_{\text{deformed}} = \mathcal{T} + \mathbf{D}^{\text{pred}}_{\text{backward}}. \tag{16}$$

Then we use the predicted overlap mask $\mathbf{M}_{\mathcal{T}}$ to filter out the non-overlapping regions in $\mathcal{T}$ and $\mathcal{T}^{\text{filtered}}$:

$$\mathcal{T}^{\text{filtered}} = \mathcal{T} \otimes \mathbf{M}_{\mathcal{T}}, \tag{17}$$

$$\mathcal{T}^{\text{filtered}}_{\text{deformed}} = \mathcal{T}_{\text{deformed}} \otimes \mathbf{M}_{\mathcal{T}}. \tag{18}$$

On one hand, the Chamfer distance loss $L^{\text{Chamfer}}_{\text{forward}}$ with the forward registration process is defined as:

$$\begin{aligned} L^{\text{Chamfer}}_{\text{forward}} = &\frac{1}{N_s} \sum_{i \in N_s^{\text{set}}} \min_{j \in N_o^{\text{set}}} \|\mathcal{S}_{\text{deformed}}(i) - \mathcal{T}^{\text{filtered}}(j)\|^2 \\ &+ \frac{1}{N_o} \sum_{j \in N_o^{\text{set}}} \min_{i \in N_s^{\text{set}}} \|\mathcal{S}_{\text{deformed}}(i) - \mathcal{T}^{\text{filtered}}(j)\|^2, \end{aligned} \tag{19}$$

where the involved two sets are defined as $N_s^{\text{set}} = \{1, ..., N_s\}$ and $N_o^{\text{set}} = \{k \in \{1, ..., N_t\} | \mathcal{T}(k) \neq 0\}$, N_o is the cardinality of N_o^{set} and can also be calculated as $\sum_{k=1}^{N_t} \mathbb{1}(\mathbf{M}_{\mathcal{T}}(k))$, where $\mathbb{1}$ is the indicator function being one and zero when the input is non-zero and zero respectively. The rationale is that we only consider points in $\mathcal{T}$ predicted to be within the overlapping region with $\mathcal{S}$. On the other hand, the Chamfer distance loss $L^{\text{Chamfer}}_{\text{backward}}$ with the backward registration process is defined as:

$$\begin{aligned} L^{\text{Chamfer}}_{\text{backward}} = &\frac{1}{N_s} \sum_{i \in N_s^{\text{set}}} \min_{j \in N_o^{\text{set}}} \|\mathcal{S}(i) - \mathcal{T}^{\text{filtered}}_{\text{deformed}}(j)\|^2 \\ &+ \frac{1}{N_o} \sum_{j \in N_o^{\text{set}}} \min_{i \in N_s^{\text{set}}} \|\mathcal{S}(i) - \mathcal{T}^{\text{filtered}}_{\text{deformed}}(j)\|^2, \end{aligned} \tag{20}$$

where N_s^{set}, N_o^{set} and N_o are the same as those defined in the forward registration loss, and the inherent rationale is also to consider the predicted overlapping points in $\mathcal{T}_{\text{deformed}}$ or $\mathcal{T}$.

3) The Overall Loss: The overall training loss is defined as the weighted sum of the mask prediction Loss in Sec. III-F.1 and the bidirectional registration Loss in Sec. III-F.2:

$$L_{\text{total}} = \alpha_1 \cdot L_{mask} + \alpha_2 \cdot (L^{\star}_{\text{forward}} + L^{\star}_{\text{backward}}). \tag{21}$$

where $\alpha_1 \in \mathbb{R}$ and $\alpha_2 \in \mathbb{R}$ are empirically set to be 0.4 and 0.6 respectively, $L^{\star}_{\text{forward}}$ and $L^{\star}_{\text{backward}}$ are $L^{L_2}_{\text{forward}}$ in Eq. (13) and $L^{L_2}_{\text{backward}}$ in Eq. (14) for the supervised registration paradigm, $L^{\text{Chamfer}}_{\text{forward}}$ in Eq. (19) and $L^{\text{Chamfer}}_{\text{backward}}$ in Eq. (20) for weakly-supervised registration paradigm.

IV. Experiments And Results

A. Dataset

There are many challenges in acquiring large scale of medical point sets (e.g., the liver organ that we are interested in this study), which includes privacy, ethics, etc. Very recently, MedShapeNet [13], a large-scale medical dataset has been proposed, which contains over 100,000 medical shapes, including bones, organs, vessels, muscles, etc., as well as surgical instruments. We propose DeformMedShapeNet, using $N_{\text{patient}} = 615$ liver point sets in MedShapeNet as the full preoperative model data in the image-to-patient registration task. $N^{\text{train}}_{\text{patient}} = 551$, $N^{\text{val}}_{\text{patient}} = 32$ and $N^{\text{test}}_{\text{patient}} = 32$ liver shapes are utilised as the training, validation and test preoperative data respectively. Similar to those in [8][24], we use Thin-Plate Spline (TPS) interpolation[25] to apply controllable deformations over preoperative liver point sets and generate intraoperative counterparts. The details of TPS deformation can be found in [26]. We first normalize the $N_{\text{patient}} = 615$ preoperative liver point sets and apply the TPS technique to generate the intraoperative data while the extent of deformation can be adjusted by TPS Deformation Level (TDL).

1) Various Deformation Levels: In image-to-patient liver regisration, the organ deformation can frequently reach several centimeters in magnitude [27]. In open surgery, the surface displacements occur with magnitudes from 0.5 to 2.0 cm and maximum closest point distance magitudes from approximately 1.0 to 2.0 cm [28]. The deformation between

TABLE I

QUANTITATIVE RESULTS UNDER DIFFERENT OVERLAP RATIOS AND NOISE LEVELS WITH THE DEFORMATION LEVEL BEING 0.2.

Noise Level	Overlap Ratio	12.5%			25%			50%		
	Metrics / Methods	RMSE(mm)	MAE(mm)	RR(%)	RMSE(mm)	MAE(mm)	RR(%)	RMSE(mm)	MAE(mm)	RR(%)
Low Noise	CPD [15]	85.72	79.80	0.0	71.31	65.39	0.0	47.17	41.99	0.0
	FPT(unsupervised) [7]	111.62	101.29	0.0	106.98	92.96	0.0	71.49	49.07	0.0
	FPT(supervised) [7]	13.85	13.24	24.1	13.74	13.06	21.9	12.71	11.92	30.0
	FLOT [6]	88.25	82.51	3.1	9.85	8.56	89.1	8.58	7.57	89.4
	NDP [22]	94.66	84.38	0.0	74.75	63.37	0.0	34.10	20.54	14.4
	Bi-PointFlowNet [9]	15.56	14.36	22.8	13.15	11.50	41.3	10.37	8.49	64.1
	MSBRN [10]	9.98	8.66	62.5	8.98	6.97	75.3	10.26	8.03	68.1
	Bi-NOFNet(Point)	13.49	11.91	35.3	13.52	11.73	32.5	13.89	12.01	23.4
	Bi-NOFNet(Feature)	**8.97**	**8.07**	**76.9**	**8.13**	**6.87**	**87.2**	**7.61**	**6.40**	**93.4**
Medium Noise	CPD [15]	85.80	79.89	0.0	71.40	65.46	0.0	47.20	42.02	0.0
	FPT(unsupervised) [7]	112.19	101.87	0.0	107.42	93.37	0.0	71.80	49.34	0.0
	FPT(supervised) [7]	13.97	13.35	23.8	13.90	13.20	19.7	12.87	12.06	29.1
	FLOT [6]	91.94	86.41	1.6	12.89	11.34	84.1	10.88	9.61	87.8
	NDP [22]	94.56	84.54	0.0	73.36	62.10	0.0	34.99	22.00	6.9
	Bi-PointFlowNet [9]	15.46	14.57	22.8	13.08	11.63	36.9	11.18	9.59	57.2
	MSBRN [10]	10.19	8.87	21.6	9.95	8.01	31.6	10.61	8.55	46.9
	Bi-NOFNet(Point)	13.72	12.17	32.5	13.77	12.04	26.9	14.06	12.25	20.3
	Bi-NOFNet(Feature)	**9.30**	**8.41**	**76.3**	**8.47**	**7.27**	**85.6**	**7.94**	**6.79**	**91.9**
High Noise	CPD [15]	85.85	79.97	0.0	71.43	65.52	0.0	47.25	42.11	0.0
	FPT(unsupervised) [7]	113.01	102.71	0.0	108.08	94.01	0.0	72.26	49.78	0.0
	FPT(supervised) [7]	14.24	13.57	20.9	14.19	13.45	16.3	13.18	12.35	22.2
	FLOT [6]	94.65	89.13	0.3	26.92	23.99	49.1	32.77	29.03	27.5
	NDP [22]	93.55	83.49	0.0	74.61	63.98	0.0	36.54	24.58	2.2
	Bi-PointFlowNet [9]	15.40	14.86	21.6	13.23	11.84	31.6	11.12	11.24	46.9
	MSBRN [10]	10.87	9.63	60.9	11.35	9.58	58.4	11.65	9.79	60.0
	Bi-NOFNet(Point)	14.07	12.57	26.3	14.15	12.50	20.0	14.42	12.69	13.4
	Bi-NOFNet(Feature)	**9.85**	**8.91**	**70.9**	**9.10**	**7.96**	**82.8**	**8.55**	**7.47**	**89.1**

the preoperative and laparoscopic(14mmHg) is 10.1±5.9 mm, between the preoperative and laparoscopic(7mmHg) is 9.0±7.0 mm [29]. To ensure the created data sets reflect real surgical scenarios, during the training phase, TDL are set to lie in the range $[0.1, 0.2]$. In the test phase, the data' TDLs are 0.1, 0.15, and 0.2, respectively. It is noted that TDL of 0.1, 0.15, and 0.2 approximately corresponds to physical deformation magnitudes being 10mm, 12mm, and 15mm respectively, which resembles those real cases in open and laparoscopic surgeries [27][28][29].

2) Various Overlap Ratios and Noise Levels: The following two steps were used to further process the full intraoperative point set $\mathcal{S}_{\text{Full}} \in \mathbb{R}^{N_f \times 3}$ where $N_f = 1024$ to generate the partial counterparts $\mathcal{S} \in \mathbb{R}^{N_s \times 3}$. First, to mimic different intraoperative visibility, we randomly cropped the simulated intraoperative point set $\mathcal{S}_{\text{Full}}$ with three different overlap ratios, i.e., 12.5%, 25% and 50%, which represent low, medium, and high levels of overlap ratios respectively. In these three cases, $N_s = 128$, $N_s = 256$, and $N_s = 512$ respectively. Specifically, we randomly generated a direction vector to represent a 3D infinite line with the centroid of $\mathcal{S}_{\text{Full}}$ as the origin, after which the shortest distances from each point in $\mathcal{S}_{\text{Full}}$ to this line are calculated and the closest points are sampled. Second, random noise was applied on both the cropped intraoperative point set $\mathcal{S}$ and preoperative point set $\mathcal{T}$ with low, medium and high noise whose mean magnitudes are 0.63mm, 1.90mm, 3.16mm respectively.

3) DeformMedShapeNet: We have conducted 10 simulated deformations for each preoperative liver model, so in each case of overlap ratio (i.e., 12.5%, 25%, 50%) and noise (i.e., medium), we totally have $N_{\text{total}}^{\text{train}} = 5510$ intraoperative point sets for training, $N_{\text{total}}^{\text{val}} = 320$ intraoperative point sets for validation, and we test our model under 9 different experimental settings (i.e., three different TDLs and noise levels according to Sec. IV-A.1 and Sec. IV-A.2), with $N_{\text{total}}^{\text{test}} = 320$ deformations for the test in each experimental setting.

B. Implementation Details

Bi-NOFNet was implemented using Pytorch on a single GPU (i.e., Nvidia GeForce RTX 4090, 24GB). During training, we used the Adam as the optimizer, the initial learning rate was set to 0.001 and the learning rate decreased by half every 100 epochs. The batch size was 16 and the network was trained with a total of 300 epochs.

C. Evaluation Metrics

For evaluation, we use root mean squared error (RMSE) and mean absolute error (MAE). $\text{RMSE}^{(m)}$ and $\text{MAE}^{(m)}$ are calculated for the m^{th} liver in the test set, where $\text{RMSE}^{(m)} = \sqrt{\frac{\sum_{i=1}^{N_s} \|\mathbf{D}_{\text{forward}}^{\text{gt}}(i) - \mathbf{D}_{\text{forward}}^{\text{pred}}(i)\|^2}{N_s}}$ and $\text{MAE}^{(m)} = \frac{\sum_{i=1}^{N_s} \|\mathbf{D}_{\text{forward}}^{\text{gt}}(i) - \mathbf{D}_{\text{forward}}^{\text{pred}}(i)\|}{N_s}$, respectively. Though we use a bidirectional registration loss function during training, we only need to compute the forward registration error from the intraoperative point set $\mathcal{S}$ to the preoperative point set $\mathcal{T}$ in the test stage. The recorded RMSE and MAE in Table I are the average values computed over liver pairs in the test set. Additionally, we also compute the Registration Recall (RR) defined as $\text{RR} = \frac{\sum_{i=1}^{N_{\text{total}}^{\text{test}}} \mathbb{1}(\text{MAE}^{(m)} < \tau)}{N_{\text{total}}^{\text{test}}}$, where $\tau \in \mathbb{R}$ is the threshold and set to be 10 mm.

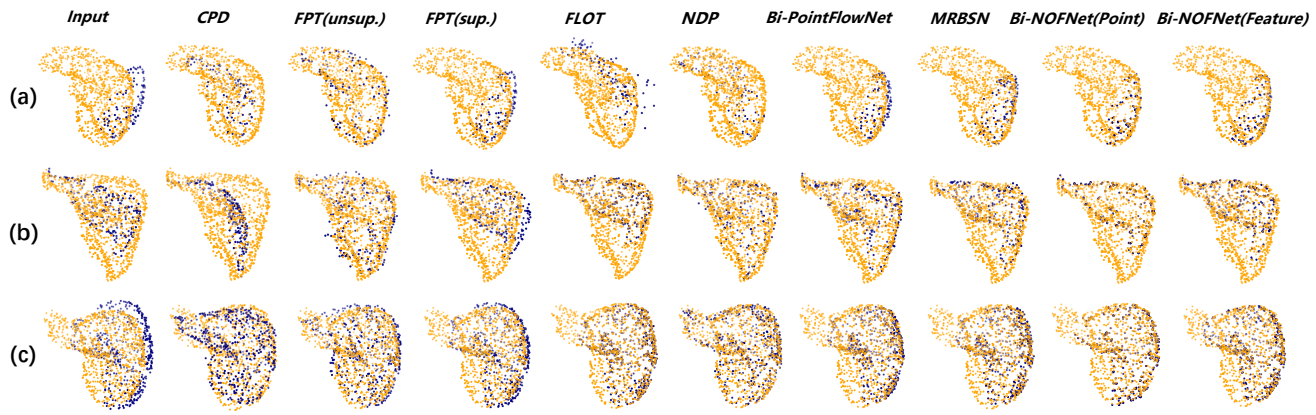

Fig. 4. Qualitative results under three overlap ratios of 12.5%, 25%, and 50% are presented in the upper (a), middle (b), and lower (c) rows respectively. The full preoperative and partial intraoperative point sets, i.e., $\mathcal{T}$ and $\mathcal{S}$, are denoted with yellow and blue dots respectively. The first column represents the case before registration while other columns denote the two point sets after registration. Bi-NOFNet (Point) and Bi-NOF (Feature) are the two proposed approaches.

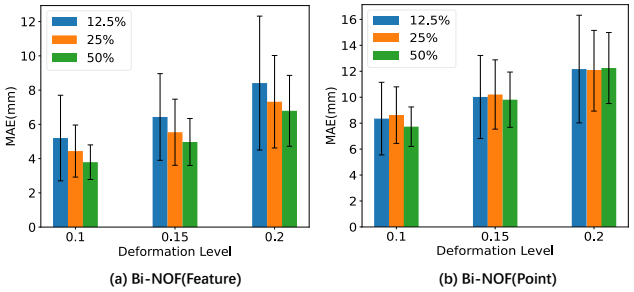

Fig. 5. Quantitative results (i.e., MAE) by Bi-NOFNet (Feature) and Bi-NOFNet (Point) under varying deformation levels and overlap ratios with medium noise conditions.

TABLE II
QUANTITATIVE RESULTS OF ABLATION STUDIES WITH OVERLAP RATIO BEING 25%, MEDIUM NOISE AND TDL BEING 0.2.

Methods	RMSE(mm)	MAE(mm)	RR(%)
FPT(unsupervised)	107.42±10.72	93.37 ±10.20	0.0
FPT(unsupervised) + NOF(Point)	16.27±4.15	15.30±4.05	6.6
Bi-NOFNet(Point)	$\mathbf{13.77 \pm 3.21}$	$\mathbf{12.04 \pm 3.11}$	**26.9**
FPT(supervised)	13.90±3.60	13.20±3.58	19.7
FPT(supervised) + NOF(Feature)	11.60±3.11	10.88±3.00	41.9
Bi-NOFNet(Feature)	$\mathbf{8.47 \pm 2.97}$	$\mathbf{7.27 \pm 2.74}$	**85.6**

D. Results

We benchmark Bi-NOFNet with related non-rigid point set registration methods, including CPD [15], FPT [7], Flot [6], NDP [22], Bi-PointFlowNet [9] and MSBRN [10].

1) Various Noise and Overlap Ratios: Table I shows quantitative results under three different noise levels and overlap ratios, where TDL is fixed as 0.2. Fig. 4 shows corresponding qualitative results. Experimental results demonstrate that Bi-NOFNet(Feature) outperforms other methods under various experimental settings. For example, in the case of "Overlap Ratio:12.5%, TDL:0.2, Noise Level:Low", Bi-NOFNet (Feature) achieves RMSE, MAE and RR of 8.97mm, 8.07mm and 76.9%, respectively, outperforming other methods. Despite utilising a weakly-supervised registration paradigm, Bi-NOFNet(Point) outperforms some supervised learning methods(e.g., FPT(supervised)[7], Flot[6], NDP[22], Bi-PointFlowNet[9]).

2) Various Deformation Levels and Overlap Ratios: To investigate the influence of overlap ratios and deformation levels over registration performances, we validated Bi-NOFNet and its variant under three overlap ratios (i.e., 12.5%, 25%, 50%) and three TDLs (i.e., 0.1, 0.15, 0.2). Fig. 5 shows the corresponding results, from which we have two observations. First, for both Bi-NOFNet (Feature) and Bi-NOFNet (Point), registration error values increase with larger deformations. Second, for Bi-NOFNet (Feature), smaller overlap ratios lead to an evident increase in both RMSE and MAE, while Bi-NOFNet (Point)'s performance is less affected by varying overlap ratios. We suspect that the lower the overlap between preoperative and intraoperative point sets is, the more difficult it is to effectively filter out non-overlapping regions in the feature space. In contrast, in Bi-NOFNet(Point), the capability of directly filtering non-overlapping regions at point-level does not heavily depend on the overlap ratio.

3) Ablation Study: We investigate the contributions of two non-overlapping filtering strategies, i.e., NOF (Feature) and NOF (Point) in Sec. III-D, and the bidirectional registration mechanism in Sec. III-E under the experimental setting being "Overlap Ratio: 25%, Medium Noise, TDL: 0.2". Table II shows the corresponding results. We compare with two baselines: FPT (unsupervised) [7] and FPT (supervised) [7].

Using FPT (unsupervised) [7], the RMSE, MAE and RR are 107.42±10.72 mm(mean±std), 93.37±10.20 mm and 0.0%, respectively, which indicates that FPT (unsupervised) is fails to handle the partial-full liver registration. FPT(supervised)[7] achieves RMSE of 13.90±3.60 mm, MAE of 13.20±3.58 mm and RR of 19.7%, which is better than FPT (unsupervised) but the error is still relatively large. FPT(unsupervised)[7] combined with NOF(Point) achieves RMSE of 16.27±4.15 mm, MAE of 15.30±4.05

mm and RR of 6.6%. FPT(supervised)[7] combined with NOF(Feature) can achieve RMSE of 11.60±3.11 mm, MAE of 10.88±3.00 mm and RR of 41.9%. These significant improvements demonstrate the effectiveness of the proposed non-overlapping filtering strategies for both registration paradigms, particularly for FPT(unsupervised)[7]. Furthermore, registration errors further decrease with the bidirectional mechanism, especially with a significant improvement in RR(cf. Bi-NOFNet(Point) and Bi-NOFNet(Feature) in Table II). The significance of the proposed bidirectional registration mechanism is thus validated.

V. Conclusions

In this paper, we have introduced Bi-NOFNet, a learning-based bidirectional partial-to-full non-rigid point set registration, for image-guided liver surgery (IGLS). We have first proposed a DeformMedShapeNet data set to validate image-to-patient registration algorithms for IGLS. Two non-overlapping region filtering strategies have then been formulated, and the bidirectional registration mechanism is formally defined in the partial-to-fll non-rigid registration. Additionally, extensive experimental results demonstrate the superiority of the Bi-NOFNet over exsiting approaches, and the ablation studies further validate the significance of the proposed modules. Furture work will explore the use of biomechanical simulation approach to generate more data in DeformMeshapeNet, and will consider the biomechanical constraints in the registration approach.

References

[1] R. H. Taylor, N. Simaan, A. Menciassi, and G.-Z. Yang, "Surgical robotics and computer-integrated interventional medicine [scanning the issue]," *Proc. IEEE*, vol. 110, pp. 823–834, 2022.

[2] Z. Min, Z. M. Baum, S. U. Saeed, M. Emberton, D. C. Barratt, Z. A. Taylor, and Y. Hu, "Non-rigid medical image registration using physics-informed neural networks," in *International Conference on Information Processing in Medical Imaging*. Springer, 2023, pp. 601–613.

[3] Z. R. Yaniv, "Registration for orthopaedic interventions," 2016.

[4] M. Pfeiffer, C. Riediger, S. Leger, J.-P. Kühn, D. Seppelt, R.-T. Hoffmann, J. Weitz, and S. Speidel, "Non-rigid volume to surface registration using a data-driven biomechanical model," in *Medical Image Computing and Computer Assisted Intervention–MICCAI 2020: 23rd International Conference, Lima, Peru, October 4–8, 2020, Proceedings, Part IV 23*. Springer, 2020, pp. 724–734.

[5] X. Liu, C. R. Qi, and L. J. Guibas, "Flownet3d: Learning scene flow in 3d point clouds," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2019, pp. 529–537.

[6] G. Puy, A. Boulch, and R. Marlet, "Flot: Scene flow on point clouds guided by optimal transport," in *European conference on computer vision*. Springer, 2020, pp. 527–544.

[7] 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.

[8] J. Shi, P. Wan, and F. Chen, "An unsupervised non-rigid registration network for fast medical shape alignment," in *2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. IEEE, 2021, pp. 1887–1890.

[9] W. Cheng and J. H. Ko, "Bi-pointflownet: Bidirectional learning for point cloud based scene flow estimation," in *European Conference on Computer Vision*. Springer, 2022, pp. 108–124.

[10] Wencan Cheng and Jong Hwan Ko, "Multi-scale bidirectional recurrent network with hybrid correlation for point cloud based scene flow estimation," *2023 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 10 007–10 016, 2023.

[11] S. Huang, Z. Gojcic, M. Usvyatsov, A. Wieser, and K. Schindler, "Predator: Registration of 3d point clouds with low overlap," in *Proceedings of the IEEE/CVF Conference on computer vision and pattern recognition*, 2021, pp. 4267–4276.

[12] H. Xu, S. Liu, G. Wang, G. Liu, and B. Zeng, "Omnet: Learning overlapping mask for partial-to-partial point cloud registration," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2021, pp. 3132–3141.

[13] J. Li, A. Pepe, C. Gsaxner, G. Luijten, Y. Jin, N. Ambigapathy, E. Nasca, N. Solak, G. M. Melito, A. R. Memon, *et al.*, "Medshapenet–a large-scale dataset of 3d medical shapes for computer vision," *arXiv preprint arXiv:2308.16139*, 2023.

[14] B. Amberg, S. Romdhani, and T. Vetter, "Optimal step nonrigid icp algorithms for surface registration," *2007 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–8, 2007.

[15] A. Myronenko and X. B. Song, "Point set registration: Coherent point drift," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, pp. 2262–2275, 2009.

[16] O. Hirose, "A bayesian formulation of coherent point drift (vol 43, pg 2269, 2021)," *IEEE transactions on pattern analysis and machine intelligence*, vol. 43, no. 9, pp. 3273–3273, 2021.

[17] H. Li, R. W. Sumner, and M. Pauly, "Global correspondence optimization for non-rigid registration of depth scans," *Computer Graphics Forum*, vol. 27, 2008.

[18] Z. Shen, J. Feydy, P. Liu, A. H. Curiale, R. S. J. Estépar, R. S. J. Estépar, and M. Niethammer, "Accurate point cloud registration with robust optimal transport," in *Neural Information Processing Systems*, 2021.

[19] Y. Li and T. Harada, "Lepard: Learning partial point cloud matching in rigid and deformable scenes," *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5544–5554, 2021.

[20] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is all you need," *Advances in neural information processing systems*, vol. 30, 2017.

[21] 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.

[22] Li, Yang and Harada, Tatsuya, "Non-rigid point cloud registration with neural deformation pyramid," *Advances in Neural Information Processing Systems*, vol. 35, pp. 27 757–27 768, 2022.

[23] C. R. Qi, H. Su, K. Mo, and L. J. Guibas, "Pointnet: Deep learning on point sets for 3d classification and segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 652–660.

[24] Z. Shen, J. Feydy, P. Liu, A. H. Curiale, R. San Jose Estepar, R. San Jose Estepar, and M. Niethammer, "Accurate point cloud registration with robust optimal transport," *Advances in Neural Information Processing Systems*, vol. 34, pp. 5373–5389, 2021.

[25] F. L. Bookstein, "Principal warps: Thin-plate splines and the decomposition of deformations," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 11, no. 6, pp. 567–585, 1989.

[26] 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.

[27] J. S. Heiselman, J. A. Collins, M. J. Ringel, T. Peter Kingham, W. R. Jarnagin, and M. I. Miga, "The image-to-physical liver registration sparse data challenge: comparison of state-of-the-art using a common dataset," *Journal of Medical Imaging*, vol. 11, no. 1, pp. 015 001–015 001, 2024.

[28] L. W. Clements, P. Dumpuri, W. C. Chapman, B. M. Dawant, R. L. Galloway, and M. I. Miga, "Organ surface deformation measurement and analysis in open hepatic surgery: method and preliminary results from 12 clinical cases," *IEEE transactions on biomedical engineering*, vol. 58, no. 8, pp. 2280–2289, 2011.

[29] J. S. Heiselman, L. W. Clements, J. A. Collins, J. A. Weis, A. L. Simpson, S. K. Geevarghese, T. P. Kingham, W. R. Jarnagin, and M. I. Miga, "Characterization and correction of intraoperative soft tissue deformation in image-guided laparoscopic liver surgery," *Journal of Medical Imaging*, vol. 5, 2017.