2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)
October 14-18, 2024. Abu Dhabi, UAE

DeepBHMR: Learning Bidirectional Hybrid Mixture Models for Generalized Rigid Point Set Registration

Zhe Min[†⋆], Zhengyan Zhang[†], Ang Zhang, Rui Song, Yibin Li, Max Q.-H. Meng, *IEEE Fellow*

***Abstract*— In this paper, we introduce a novel normal-assisted learning-based rigid registration approach, i.e., Deep Bi-directional Hybrid Mixture Registration (DeepBHMR). Our approach utilises helpful normal vectors explicitly in both correspondence and transformation stages and formulates the optimization objective of registration in a bi-directional way that considers noise in both point sets. DeepBHMR consists of three modules: (1) the correspondence network that estimates the correspondence probability relating points within one generalized point set (i.e., positional and normal vectors) with components of Hybrid Mixture Models (HMMs) representing the other generalized point set; (2) the posterior module that computes HMMs parameters; (3) the transformation module that computes the rotation matrix and the translation vector given the estimated generalized-point to hybrid-distribution correspondences and HMMs parameters. DeepBHMR has been validated on 291 human femur and 260 hip models, and extensive experimental results demonstrate that DeepBHMR outperforms the state-of-the-art registration methods (p-value < 0.01). In the circumstance of femur bones, the mean rotation and translation error values are around 1° (i.e., 1.01°) and less than 1 mm (i.e., 0.36mm), respectively. Furthermore, even under the large transformation (i.e., in the range of $[0, 180]^\circ$ and [0, 100] mm), the mean RMSE values being 3.05 mm is still satisfactory. Additionally, the results demonstrate the DeepBHMR's favorable generalizability from femur shapes to hip shapes. We have carefully validated the significant benefits of incorporating normal vectors and the bidirectional mechanism. DeepBHMR can successfully handle the challenging scenario of large transformation and partial registration. The codes are available at https://github.com/zzyrobot/DeepBHMR.git.**

I. INTRODUCTION

Registration is a fundamental problem that has wild applications in fields of medical imaging and computer-assisted interventions (CAI) [1]–[3], etc. For instance, in CAI, it is vital to fuse information from multiple modalities (e.g., computed tomography (CT) and optical tracking system (OTS)) at different stages to enable accurate disease diagnosis and further interventions [1]. Point set registration aims to find the transformation between two spaces represented by point sets, where the transformation can be either parameterized or parameter-free [4], [5]. Particularly, the rigid registration is to estimate the rigid transformation matrix that best aligns two point sets given certain metrics (e.g., Euclidean distance) [6]. In CAI, point sets are reconstructed from volumetric medical images (with limited resolutions) pre-operatively and localised using optically tracked surgical pointers intra-operatively, which may introduce noise and outliers in two spaces. On one hand, we identify that the core of registration is data association, for example at point-level, which obtains one-to-one hard correspondence (i.e.,$\{0, 1\}$) or one-to-all soft correspondence probabilities (i.e, $[0, 1]$), etc. On the other hand, the above-mentioned adverse factors including noise and outliers will cause accurate point-level correspondences to not exist in real-world scenarios [7], [8], which makes it very difficult to solve the data association problem accurately.

Motivations. In this paper, we adopt the hybrid learning paradigm that combines probabilistic mixture models and end-to-end learning. The motivations of this paper are two-fold. *On one hand*, most learning-based registration methods only use positional vectors [9]–[11]. To improve registration robustness and accuracy, in this paper, our approach leverages the additional useful information (i.e., normal vectors). *On the other hand*, most probabilistic registration methods (e.g., CPD [12], BCPD [13]) assume that one point set is perfect and the other contains noise and outliers, while noise and outliers exist in both point sets in real-world scenarios [14]. To consider noise and outliers in both point sets, the bidirectional registration process is carefully formulated.

Contributions. The contributions of this study are three-fold. *First*, we propose the Deep Bi-directional Hybrid Mixture Registration (DeepBHMR), the first learning-based generalized point set registration method that explicitly leverages the accompanying normal vectors (cf. Fig. 1), which directly contribute to both correspondence and transformation steps. This is done through formulating the registration as the minimization of KL-divergence between two (i.e., in two spaces) probabilistic distributions modeled as hybrid mixtures consisting of Gaussian [10] and Fisher [15] distributions. *Second*, the bi-directional mechanism is utilized in DeepBHMR, where the optimization objective contains the KL divergence between the rigidly transformed source and latent target distributions, and that between the rigidly transformed target and source distributions. *Third*, we evaluate DeepBHMR

[†]Equal contribution. [⋆]Corresponding Author. This work was supported in part by the National Natural Science Foundation of China under Grant 62303275. This work was also supported by the National Natural Science Fund for Excellent Young Scientists Fund Program (Overseas) under Grant 22IAA01849, Jinan Science and Technology Bureau under Grant 202333011.

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

Zhengyan Zhang is with the Department of Aeronautical and Aviation Engineering, The Hong Kong Polytechnic University.

Ang Zhang is with Yuanhua Robotics, Perception and AI Technologies Ltd, Shenzhen, China.

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

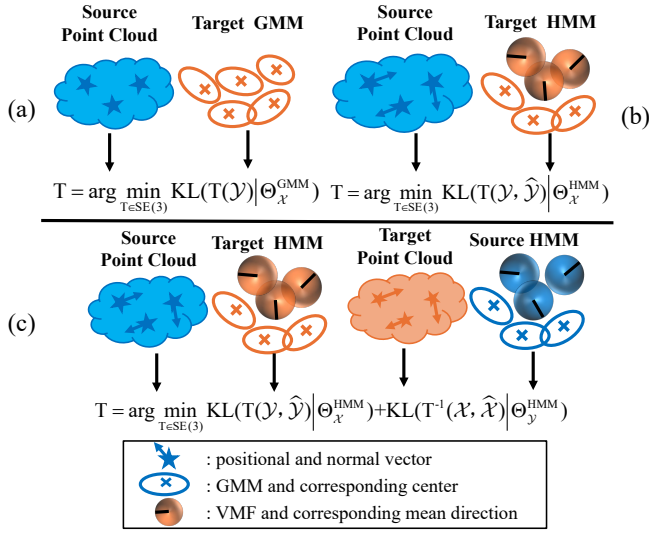

Fig. 1. Visualization of different hybrid registration methods. In (a) and (b), the registration task is in a forward way (i.e., from source to target point sets). As shown in (a), DeepGMR, the objective function consists of positional vectors with the source point set $\mathcal{Y} \in \mathbb{R}^{3\times M}$ and the target Gaussian Mixture Model (GMM) Θ_x^{GMM}. In (b), DeepBHMR (*forward*), the objective function is formulated as the KL divergence between the transformed generalized source point set $\mathbf{T}(\mathcal{Y}, \widehat{\mathcal{Y}})$ and the target HMM Θ_x^{HMM} consisting of GMM and Fisher Mixture Model (FMM), where $\mathbf{T}$ is the rigid transformation operator. As shown in (c), DeepBHMR, the proposed bidirectional registration approach, DeepBHMR, is defined as the sum of KL divergence between the $\mathbf{T}(\mathcal{Y}, \widehat{\mathcal{Y}})$ and Θ_x^{HMM}, and that between the inversely transformed generalized target point set $\mathbf{T}(\mathcal{X}, \widehat{\mathcal{X}})$ and the source HMM Θ_y^{HMM}.

against state-of-the-art registration methods, using point sets of 291 femur and 260 hip shapes[16]. The extensive results validate DeepBHMR's strong generalizability to new data, and capability to handle the challenging global and partial-to-full registration problems.

Organizations. The remainder of this paper is organized as follows. Sect. II introduces relevant rigid point set registration methods in the literature. Sect. III describes the generalized rigid point set registration problem. Sect. IV formally formulates the registration process as a two-step optimization problem. Sect. V describes the proposed DeepBHMR approach in detail. Sect. VI introduces experiments and results. Sect. VII concludes the paper.

II. Related Work

In this section, we first review the most related rigid registration approaches categorized into conventional, deep-learning-based, and hybrid registration methods. Afterwards, we highlight two remarkable differences of DeepBHMR from the state-of-the-art probabilistic learning-based registration approach.

Conventional Registration Methods. Iterative Closest Point (ICP) [17] iterates over updating one-to-one point correspondences and estimating the rigid transformation matrix, which is susceptible to noise and outliers. Coherent Point Drift [12] regards points in one point set as centroids of one Gaussian Mixture Model (GMM), from which the other point set is generated. Bayesian CPD (BCPD) [13] reformulates CPD according to the Bayesian setting, where the theoretical convergence is guaranteed. In ICP [17], CPD [12] and BCPD [13], one point set is considered to be perfect while the other contains noise and outliers. In contrast to most probabilistic methods.

Deep-learning-based Registration Methods. FMR [9] solves the registration problem by minimizing the projection error within the feature space in a semi-supervised manner. PointNetLK++ [18] circumvents the limitations of PointNetLK, by analytically computing the Jacobian of PointNet features for the exponential map *twist* parameters. Predator[19], is designed for feature matching of two point sets with low overlap, comprising three key modules: an encoder that extracts superpoints and their associated features, an overlap-attention module that captures co-contextual information between the two point sets, and a decoder that predicts per-point features, overlap scores, and matchability scores. In RegTr [20], the KPConv backbone [21] is employed to extract superpoints along with their features, the transformer cross-encoder together with output layer are then utilized to predict the transformed superpoints and overlap scores.

Hybrid Registration Methods. DeepGMR [10] is the milestone hybrid learning-based registration method that explicitly leverages the probabilistic registration paradigm. DeepGMR [10] consists of three consecutive components: *one* learnable correspondence network that estimates point-to-distribution correspondence probabilities, and *two* parameter-free compute blocks that estimate GMM parameters and rigid transformations. Based on the DeepGMR's framework, OGMM [22] further uses a clustering-based transformer module to predict the overlap score of each point in two point sets, and then establish the overlap-guided GMM to solve the partial registration problem. UGMM [11] considering two point sets to be samples generated from one unified GMM, first learns the point-to-distribution correspondences, then estimates the posterior GMM parameters, afterward computes the point-to-GMM transformations using estimated correspondences and GMM parameters. UGMM is an unsupervised learning method guided by the Chamfer distance loss. It is noticed that only positional information is utilized in hybrid registration methods such as DeepGMR [10], UGMM [11], and OGMM [22].

Differences from Hybrid Registration Methods. Our method is inspired by the hybrid registration methods but is different from them significantly in two aspects. First, DeepGMR [10] and UGMM [11] only use positional vectors from which GMMs are estimated, while our proposed DeepBHMR utilises generalized point sets with additional normal information, from which HMMs are computed. Second, in the transformation computation block of DeepGMR [10] and OGMM [22], the KL divergence between the transformed source and target distributions (i.e., the *forward* way) is minimized while in DeepBHMR the optimization objective contains the *additional* KL-divergence between the transformed (using the inverse transformation matrix) target

and source distributions (i.e., the *backward* way).

III. The Generalized Rigid Point Set Registration

We first formally formulate the point set registration problem, where both positional and normal information is leveraged. One target and one source generalized points are defined as $\mathbf{d}_n^x = [\mathbf{x}_n^\mathsf{T}, \widehat{\mathbf{x}}_n^\mathsf{T}]^\mathsf{T} \in \mathbb{R}^6$ where $\mathbf{x}_n \in \mathbb{R}^3$ and $\widehat{\mathbf{x}}_n \in \mathbb{R}^3$ with $||\widehat{\mathbf{x}}_n|| = 1$ denote the positional and normal vectors respectively, and $\mathbf{d}_m^y = [\mathbf{y}_m, \widehat{\mathbf{y}}_m]^\mathsf{T} \in \mathbb{R}^6$ where $\mathbf{y}_m \in \mathbb{R}^3$ and $\widehat{\mathbf{y}}_m \in \mathbb{R}^3$ with $||\widehat{\mathbf{y}}_m|| = 1$. We also define $\mathcal{X} = \{\mathbf{x}_n\}_{n=1}^N$ and $\mathcal{Y} = \{\mathbf{y}_m\}_{m=1}^M$ for positional vector sets, $\widehat{\mathcal{X}} = \{\mathbf{x}_n\}_{n=1}^N$ and $\widehat{\mathcal{Y}} = \{\mathbf{y}_m\}_{m=1}^M$ for normal vector sets. For clarity, we further denote the generalized point sets in two spaces as $\mathcal{D}_\mathcal{X} = \{\mathcal{X}, \widehat{\mathcal{X}}\}$ and $\mathcal{D}_\mathcal{Y} = \{\mathcal{Y}, \widehat{\mathcal{Y}}\}$. Given $\mathcal{D}_\mathcal{X}$ and $\mathcal{D}_\mathcal{Y}$, the *generalized Rigid Point Set Registration (GRPSR)* is to estimate the rigid transformation including the rotation matrix $\mathbf{R} \in SO(3)$ and the translation vector $\mathbf{t} \in \mathbb{R}^3$ that best align the two spaces.

IV. HMM-Based generalized Point Set Registration

Hybrid Mixture Models. The Hybrid Mixture model (HMM) consists Gaussian Mixture Model (GMM) and Fisher Mixture Model (FMM), which establishes the probability distribution over 6D space as a weighted sum of $J \in N^+$ multiplications of Gaussian and Fisher distributions,

$$p(\mathbf{d}_n^x|\mathbf{\Theta}_\mathcal{X}) = \sum_{j=1}^{J} \pi_j \underbrace{\mathcal{N}(\mathbf{x}_n|\boldsymbol{\mu}_j^x, (\sigma_j^x)^2)}_{\text{Gaussian Distribution}} \underbrace{\mathcal{F}(\widehat{\mathbf{x}}_n|\widehat{\boldsymbol{\mu}}_j^x, \kappa_j^x)}_{\text{Fisher Distribution}}, \quad (1)$$

where HMM parameters $\mathbf{\Theta}_\mathcal{X} = \{\boldsymbol{\pi}_j^x, \boldsymbol{\mu}_j^x, \boldsymbol{\sigma}_j^x, \widehat{\boldsymbol{\mu}}_j^x, \boldsymbol{\kappa}_j^x\}_{j=1}^J$ consist of the mixture weight π_j ($0 < \pi_j < 1$ and $\sum_{j=1}^J \pi_j = 1$), the mean positional vector $\boldsymbol{\mu}_j^x \in \mathcal{R}^3$ and the variance $(\sigma_j^x)^2 \in \mathcal{R}$ associated with the j-the component, the mean normal vector $\widehat{\boldsymbol{\mu}}_j^x \in \mathcal{R}^3(||\widehat{\boldsymbol{\mu}}_j^x|| = 1)$ and concentration parameter $\kappa_j \in \mathcal{R}$. The probability distribution $p(\mathbf{d}_m^y|\mathbf{\Theta}_y)$ where $\mathbf{\Theta}_\mathcal{Y} = \{\boldsymbol{\pi}_j^y, \boldsymbol{\mu}_j^y, \boldsymbol{\sigma}_j^y, \widehat{\boldsymbol{\mu}}_j^y, \boldsymbol{\kappa}_j^y\}_{j=1}^J$ can be defined similarly as (1).

HMM-based Point Set Registration. The registration process is formulated as a two-step optimization. First, the generalized point-set pairs $\mathcal{D}_\mathcal{X}$ and $\mathcal{D}_\mathcal{Y}$ are represented by HMMs. Second, we are maximizing the probability of the transformed source (or target) generalized point set being generated by the HMMs estimated from the target (or source) generalized point set. Given $\mathcal{D}_\mathcal{X}$ and $\mathcal{D}_\mathcal{Y}$, and the space of permitted HMM parameterizations $\mathcal{H}$, similar to that in DeepGMR, the registration from $\mathcal{D}_\mathcal{Y}$ to $\mathcal{D}_\mathcal{X}$ can be formulated as a two-step optimization

$$\textbf{Fitting: } \mathbf{\Theta}_\mathcal{X}^\star = \arg\max_{\mathbf{\Theta}_\mathcal{X} \in \mathcal{H}} p(\mathcal{D}_\mathcal{X}|\mathbf{\Theta}_\mathcal{X}), \quad (2)$$

$$\textbf{Registration: } \arg\max_{\mathsf{T} \in \mathrm{SE}(3)} p(\mathsf{T}(\mathcal{D}_\mathcal{Y})|\mathbf{\Theta}_\mathcal{X}^\star). \quad (3)$$

where SE(3) is the special Euclidean group. The fitting step fits HMM $\mathbf{\Theta}_\mathcal{X}^\star$ to the target generalized point set $\mathcal{D}_\mathcal{X}$, while the registration step estimates the optimal transformation $\mathsf{T}^\star$ that aligns the source point set $\mathcal{D}_\mathcal{Y}$ to $\mathbf{\Theta}_\mathcal{X}^\star$. With the general assumption of point independence between $\mathbf{d}_m^y$, we have the follows,

$$p(\mathsf{T}(\mathcal{D}_\mathcal{Y})|\mathbf{\Theta}_\mathcal{X}^\star) = \prod_{m=1}^{M} p(\mathbf{d}_m^y|\mathbf{\Theta}_\mathcal{X}^\star). \quad (4)$$

A. The Fitting of HMMs to the Generalized Point Sets

Generally speaking, directly maximizing $p(\mathsf{T}(\mathcal{D}_\mathcal{Y})|\mathbf{\Theta}_\mathcal{X}^\star)$ in Eq. (4) is intractable. By introducing a set of point-to-distribution correspondence variable $\mathcal{C}_\mathcal{X} = \{c_{nj}\}$ with $c_{nj} \in \{0, 1\}$, the lower bound on p can be maximized using the expectation maximization procedure. The E-step updates the lower bound p to the posterior over $\mathcal{C}_\mathcal{X}$ given the guessed parameters $\mathbf{\Theta}_\mathcal{X}^k$ while the M-step estimates the parameters by maximizing the expected joint likelihood under q.

The EM updates for the fitting step in Eq. (2) are

$$\mathbf{E}_{\mathbf{\Theta}_\mathcal{X}} : q(\mathcal{C}) := p(\mathcal{C}|\mathcal{D}_\mathcal{X}, \mathbf{\Theta}_\mathcal{X}^k), \quad (5)$$

$$\mathbf{M}_{\mathbf{\Theta}_\mathcal{X}} : \mathbf{\Theta}_\mathcal{X}^{k+1} := \arg\max_{\mathbf{\Theta}_\mathcal{X} \in \mathcal{H}} \mathbb{E}_q[\ln p(\mathcal{D}_\mathcal{X}, \mathcal{C}|\mathbf{\Theta}_\mathcal{X})], \quad (6)$$

where $k \in \mathbb{N}^+$ is the index of the EM step. With the generalized point set $\mathcal{D}_\mathcal{X}$ and HMM parameter set $\mathbf{\Theta}_\mathcal{X}^k = \{\pi_j, \boldsymbol{\mu}_j^x, (\sigma_j^x)^2, \widehat{\boldsymbol{\mu}}_j^x, \kappa_j^x\}$, the correspondence set $\mathcal{C}$ composed of latent variables $\{c_{nj}\}_{n,j=1,1}^{N,J}$ whose posteriors are computed as

$$p(c_{nj} = 1|\mathbf{d}_n^x, \mathbf{\Theta}_\mathcal{X}^k) = \frac{\pi_j \mathcal{N}(\mathbf{x}_n|\boldsymbol{\mu}_j^x, (\sigma_j^x)^2)\mathcal{F}(\widehat{\mathbf{x}}_n|\widehat{\boldsymbol{\mu}}_j^x, \kappa_j^x)}{p(\mathbf{d}_n^x|\mathbf{\Theta}_\mathcal{X}^k)}, \quad (7)$$

where $p(\mathbf{d}_n|\mathbf{\Theta}_\mathcal{X}^k)$ is defined in Eq. (1), $p(c_{nj} = 1|\mathbf{d}_n^x, \mathbf{\Theta}_\mathcal{X}^k)$ can be seen as the correspondence probability between the point $\mathbf{d}_n^x$ and the j-th HMM component whose centers are $\boldsymbol{\mu}_j^x$ and $\widehat{\boldsymbol{\mu}}_j^x$.

By defining $\gamma_{nj} = \mathbb{E}_q[c_{nj}]$, the solution for Eq. (6) is

$$\boldsymbol{\mu}_j^x = \frac{\sum_{n=1}^N \gamma_{nj}\mathbf{x}_n}{\sum_{n=1}^N \gamma_{nj}}, \ \widehat{\boldsymbol{\mu}}_j^x = \frac{\sum_{n=1}^N \gamma_{nj}\widehat{\mathbf{x}}_n}{||\sum_{n=1}^N \gamma_{nj}\widehat{\mathbf{x}}_n||}, \quad (8)$$

$$(\sigma_j^x)^2 = \frac{\sum_{j=1}^J (\mathbf{x}_n - \boldsymbol{\mu}_j^x)^\mathsf{T}(\mathbf{x}_n - \boldsymbol{\mu}_j^x)}{3\sum_{j=1}^J \gamma_{nj}}, \quad (9)$$

$$r_j = \frac{\sum_{j=1}^J \gamma_{nj}\widehat{\mathbf{x}}_n^\mathsf{T}\widehat{\boldsymbol{\mu}}_j^x}{\sum_{j=1}^J \gamma_{nj}}, \ \kappa_j = \frac{r_j(3 - r_j^2)}{(1 - r_j^2)}. \quad (10)$$

B. The Registration under the HMM

The registration procedure in Eq. (3) is solved with respect to the rigid transformation T again with the EM steps,

$$\mathbf{E}_\mathsf{T} : q(\mathcal{C}_\mathcal{X}) := p(\mathcal{C}_\mathcal{X}|\mathsf{T}^k(\mathcal{D}_\mathcal{Y}), \mathbf{\Theta}_\mathcal{X}^\star), \quad (11)$$

$$\mathbf{M}_\mathsf{T} : \mathsf{T}^{k+1} := \arg\max_{\mathsf{T} \in \mathrm{SE}(3)} \mathbb{E}_q[\ln p(\mathsf{T}(\mathcal{D}_\mathcal{Y}), \mathcal{C}_\mathcal{X}|\mathbf{\Theta}_\mathcal{X}^\star)]. \quad (12)$$

V. The Deep Bi-Directional Hybrid Mixture Registration (DeepBHMR)

Bidirectional HMM-based Registration. In the forward HMM-based registration, the optimization objectives in Eq. (11) and Eq. (12) are utilized. In the bidirectional HMM-based registration, both the *forward* process and the *backward* process are considered. In the *forward* process, we are aligning $\mathcal{D}_\mathcal{Y}$ using the forward transformation matrix T with $\mathbf{\Theta}_\mathcal{X}$, where T contains the rotation matrix $\mathbf{R} \in SO(3)$ and the translation vector $\mathbf{t} \in \mathbb{R}^3$. In the *backward* process, we are aligning $\mathcal{D}_\mathcal{X}$ using the backward transformation matrix T^{-1} with $\mathbf{\Theta}_\mathcal{Y}$. The EM steps that optimize over the transformation parameters T in a bi-directional manner are as follows.

$$\begin{aligned}\mathbf{E}_\mathsf{T}: q(\mathcal{C}_\mathcal{X}, q(\mathcal{C}_\mathcal{Y}) :=& p(\mathcal{C}_\mathcal{X}|\mathsf{T}^k(\mathcal{D}_\mathcal{Y}), \mathbf{\Theta}^\star_\mathcal{X}) \\ =& p(\mathcal{C}_\mathcal{Y}|(\mathsf{T}^k)^{-1}(\mathcal{D}_\mathcal{X}), \mathbf{\Theta}^\star_\mathcal{Y}),\end{aligned} \tag{13}$$

$$\begin{aligned}\mathbf{M}_\mathsf{T}: \mathsf{T}^{k+1} := \arg\max_{\mathsf{T}\in \mathrm{SE}(3)} \mathbb{E}_q[\ln p(\mathsf{T}(\mathcal{D}_\mathcal{Y}), \mathcal{C}_\mathcal{X}|\mathbf{\Theta}^\star_\mathcal{X}) + \\ \ln p(\mathsf{T}^{-1}(\mathcal{D}_\mathcal{X}), \mathcal{C}_\mathcal{Y}|\mathbf{\Theta}^\star_\mathcal{Y})].\end{aligned} \tag{14}$$

DeepBHMR iteratively estimates the generalized-point to hybrid-mixture-model correspondence, and updates the model and transformation parameters, while the objective function in the transformation module is defined in a bi-directional manner. In what follows, we introduce the correspondence network that estimates point-to-distribution correspondence probabilities in Sect. V-A, the posteriors computing block that estimates the HMM parameters from the generalized point sets in Sect. V-B, the transformation computing block that updates the rotation and translation in Sect. V-C.

A. Correspondence Network

The correspondence network f_ϕ takes one generalized point set (e.g., the source generalized point set $\mathcal{D}_\mathcal{Y}$), and outputs the correspondence probabilities $[\gamma_{mj}]$, where $\sum_{j=1}^J \gamma_{mj} = 1$ for all m. Similarly, with the generalized target point set $\mathcal{D}_\mathcal{X}$ and another correspondence network f_ϕ, the correspondence probabilities $[\gamma_{nj}]$ are computed where $\sum_{j=1}^J \gamma_{nj} = 1$ for all n. Each γ_{mj} or γ_{nj} represents the latent correspondence probability between $\mathbf{d}^x_n$(or $\mathbf{d}^y_m$) and the j-th component of the corresponding latent HMMs.

B. Learning Posteriors

With the simplified assumptions that all mixture components share the common weights, all positional variance and concentration parameters are the same across the data points $\mathcal{D}_\mathcal{X}$, summarised as $\pi^x_j = \frac{1}{J_x}$, for $j = 1, ..., J_x$, $(\sigma^x_1)^2 = ...(\sigma^x_j)^2... = (\sigma^x_{J_x})^2 = \sigma^2$, and $\kappa^x_1 = ...\kappa^x_j... = \kappa^x_{J_x} = \kappa$. Take the generalized target point set $\mathcal{D}_\mathcal{X}$ as an example,, the estimation of the mean positional vector $\boldsymbol{\mu}^x_j$ and normal vector $\widehat{\boldsymbol{\mu}}^x_j$ are similar to those in Eq. (8) while the other two parameters σ^2 and κ in $\mathbf{\Theta}_\mathcal{X}$ are estimated with

$$\sigma^2 = \frac{\sum_{j=1}^J \sum_{n=1}^N \gamma_{nj}(\mathbf{x}_n - \boldsymbol{\mu}^x_j)^\mathsf{T}(\mathbf{x}_n - \boldsymbol{\mu}^x_j)}{3\sum_{j=1}^J \sum_{n=1}^N \gamma_{nj}}, \tag{15}$$

$$\kappa = \frac{\sum_{j=1}^J \sum_{n=1}^N \gamma_{nj} \widehat{\mathbf{x}}^\mathsf{T}_n \widehat{\boldsymbol{\mu}}^x_j}{\sum_{j=1}^J \sum_{n=1}^N \gamma_{nj}}. \tag{16}$$

The computation of $\mathbf{\Theta}_\mathcal{Y}$ from $\mathcal{D}_\mathcal{Y}$ is conducted in a similar manner.

C. Transformation Module

The bi-directional transformation module is given as follows. At the beginning, the maximum likelihood objective is to minimize the sum of KL-divergence between the transformed latent source distribution $\mathsf{T}(\mathbf{\Theta}_\mathcal{Y})$ and latent target distribution $\mathbf{\Theta}_\mathcal{X}$ (i.e., forward), and that between the inversely transformed latent target distribution $\mathsf{T}^{-1}(\mathbf{\Theta}_\mathcal{X})$ and latent source distribution $\mathbf{\Theta}_\mathcal{Y}$ (i.e., backward),

$$\mathsf{T}^\star = \arg\min_\mathsf{T} \underbrace{\mathrm{KL}(\mathsf{T}(\mathbf{\Theta}_\mathcal{Y})|\mathbf{\Theta}_\mathcal{X})}_{\text{forward}} + \underbrace{\mathrm{KL}(\mathsf{T}^{-1}(\mathbf{\Theta}_\mathcal{X})|\mathbf{\Theta}_\mathcal{Y})}_{\text{backward}}, \tag{17}$$

where the forward transformation operation $\mathsf{T}(\bullet)$ with rotation $\mathbf{R}$ and $\mathbf{t}$ is $\mathsf{T}(\mathbf{y}_m) = \mathbf{R}\mathbf{y}_m + \mathbf{t}$ and $\mathsf{T}(\widehat{\mathbf{y}}_m) = \mathbf{R}\widehat{\mathbf{y}}_m$, $\mathsf{T}(\boldsymbol{\mu}^y_j)$ and $\mathsf{T}(\widehat{\boldsymbol{\mu}}^y_j)$ can computed similarly. On the other hand, the operation of the inverse transformation matrix $\mathsf{T}^{-1}(\bullet)$ with the corresponding rotation $\mathbf{R}^\mathsf{T}$ and translation $-\mathbf{R}^\mathsf{T}\mathbf{t}$ is defined as $\mathsf{T}^{-1}(\mathbf{x}_n) = \mathbf{R}^\mathsf{T}\mathbf{x}_n - \mathbf{R}^\mathsf{T}\mathbf{t}$ and $\mathsf{T}^{-1}(\widehat{\mathbf{x}}_n) = \mathbf{R}^\mathsf{T}\widehat{\mathbf{x}}_n$ with which $\mathsf{T}^{-1}(\boldsymbol{\mu}^x_j)$ and $\mathsf{T}^{-1}(\widehat{\boldsymbol{\mu}}^x_j)$ can be computed in the similar manner. After manipulations that are similar to those in [10], the *forward* and *backward* objective functions $\mathcal{L}_\text{forward}$ and $\mathcal{L}_\text{backward}$ to be minimized are respectively

$$\mathcal{L}_\text{forward} = \sum_{m,j=1}^{M,J} \gamma_{mj}\Big(\frac{1}{2\sigma^2}||\mathsf{T}(\mathbf{y}_m) - \boldsymbol{\mu}^x_j||^2 - \kappa(\mathbf{R}\widehat{\mathbf{y}}_m)^\mathsf{T}\widehat{\boldsymbol{\mu}}^x_j\Big), \tag{18}$$

and $\mathcal{L}_\text{backward}$ is defined as

$$\begin{aligned}&\sum_{n,j=1}^{N,J} \gamma_{nj}\Big(\frac{1}{2\sigma^2}||\mathsf{T}^{-1}(\mathbf{x}_n) - \boldsymbol{\mu}^y_j||_2^2 - \kappa(\mathbf{R}^\mathsf{T}\widehat{\mathbf{x}}_n)^\mathsf{T}\widehat{\boldsymbol{\mu}}^y_j\Big) \\ &= \sum_{n,j=1}^{N,J} \gamma_{nj}\Big(\frac{1}{2\sigma^2}||\mathsf{T}(\boldsymbol{\mu}^y_j) - \mathbf{x}_n||_2^2 - \kappa(\mathbf{R}\widehat{\boldsymbol{\mu}}^y_j)^\mathsf{T}\widehat{\mathbf{x}}_n\Big),\end{aligned} \tag{19}$$

where from the first to the second line we have adopted the distance-preserving property of the rigid transformation. Then the overall objective function $\mathcal{L}_\text{bidirectional}$ is

$$\mathcal{L}_\text{bidirectional} = \mathcal{L}_\text{forward} + \mathcal{L}_\text{backward}. \tag{20}$$

Translation Vector. *1) Bidirectional Registration* In the bi-directional HMM-based registration scenario, the optimal translation $\mathbf{t}^\star$ is computed by solving $\frac{\partial \mathcal{L}_\text{bidirectional}}{\partial \mathbf{t}} = \mathbf{0}$ which results in

$$\mathbf{t}^\star = \boldsymbol{\mu}^\text{bidirectional}_{V_x} - \mathbf{R}\boldsymbol{\mu}^\text{bidirectional}_y, \tag{21}$$

where $\boldsymbol{\mu}^\text{bidirectional}_{V_x}$ and $\boldsymbol{\mu}^\text{bidirectional}_y$ are

$$\begin{aligned}\boldsymbol{\mu}^\text{bi}_{V_x} &= \frac{\sum_{m=1}^M \sum_{j=1}^J \gamma_{mj}\boldsymbol{\mu}^x_j + \sum_{n=1}^N \sum_{j=1}^J \gamma_{nj}\mathbf{x}_n}{\sum_{m=1}^M \sum_{j=1}^J \gamma_{mj} + \sum_{n=1}^N \sum_{j=1}^J \gamma_{nj}}, \\ \boldsymbol{\mu}^\text{bi}_y &= \frac{\sum_{m=1}^M \sum_{j=1}^J \gamma_{mj}\mathbf{y}_m + \sum_{n=1}^N \sum_{j=1}^J \gamma_{nj}\boldsymbol{\mu}^y_j}{\sum_{m=1}^M \sum_{j=1}^J \gamma_{mj} + \sum_{n=1}^N \sum_{j=1}^J \gamma_{nj}}.\end{aligned} \tag{22}$$

2) Forward Registration In the *forward* HMM-based registration, the optimal translation $\mathbf{t}^\star$ is computed by solving $\frac{\partial \mathcal{L}_{\text{forward}}}{\partial \mathbf{t}} = \mathbf{0}$, resulting

$$\mathbf{t}^\star = \boldsymbol{\mu}_{V_x} - \mathbf{R}\boldsymbol{\mu}_y, \tag{23}$$

where $\boldsymbol{\mu}_{V_x} = \frac{\sum_{m=1}^{M}\sum_{j=1}^{J}\gamma_{mj}\boldsymbol{\mu}_j^x}{\sum_{m=1}^{M}\sum_{j=1}^{J}\gamma_{mj}}$ and $\boldsymbol{\mu}_y = \frac{\sum_{m=1}^{M}\sum_{j=1}^{J}\gamma_{mj}\mathbf{y}_m}{\sum_{m=1}^{M}\sum_{j=1}^{J}\gamma_{mj}}$.

The Rotation Matrix. *1) Bidirectional Registration* The objective function related with $\mathbf{R}$, after substituting Eq. (21) into Eq. (20), is simplified as

$$\mathbf{R}^\star = \arg\max_{\mathbf{R}} \operatorname{Tr}\Big(\mathbf{R}\mathbf{H}\Big), \tag{24}$$

where $\mathbf{H}$ is defined as

$$\begin{aligned}\mathbf{H} = \frac{1}{\sigma^2}\Big(\sum_{m=1}^{M}\sum_{j=1}^{J}\gamma_{mj}\mathbf{y}'_m(\boldsymbol{\mu}_j^{x'})^\mathsf{T} + \sum_{n=1}^{N}\sum_{j=1}^{J}\gamma_{nj}\boldsymbol{\mu}_j^{y'}(\mathbf{x}'_n)^\mathsf{T}\Big) \\ + \kappa\Big(\sum_{m=1}^{M}\sum_{j=1}^{J}\gamma_{mj}\widehat{\mathbf{y}}_m(\widehat{\boldsymbol{\mu}}_j^{x})^\mathsf{T} + \sum_{n=1}^{N}\sum_{j=1}^{J}\gamma_{nj}\widehat{\boldsymbol{\mu}}_j^{y}(\widehat{\mathbf{x}}_n)^\mathsf{T}\Big),\end{aligned} \tag{25}$$

with

$$\mathbf{y}'_m = \mathbf{y}_m - \boldsymbol{\mu}_y^{\text{bi}},\ \boldsymbol{\mu}_j^{x'} = \boldsymbol{\mu}_j^x - \boldsymbol{\mu}_{V_x}^{\text{bi}}, \tag{26}$$

and

$$\boldsymbol{\mu}_j^{y'} = \boldsymbol{\mu}_j^y - \boldsymbol{\mu}_y^{\text{bi}},\ \mathbf{x}'_n = \mathbf{x}_n - \boldsymbol{\mu}_{V_x}^{\text{bi}}. \tag{27}$$

By conducting the singular value decomposition (SVD) of $\mathbf{H}$ as $\mathbf{H} = \mathbf{U}\mathbf{S}\mathbf{V}^\mathsf{T}$, we can get the updated rotation matrix $\mathbf{R}^\star$ as

$$\mathbf{R}^\star = \mathbf{V}\text{diag}([1, 1, \det(\mathbf{V}\mathbf{U}^\mathsf{T})])\mathbf{U}^\mathsf{T}. \tag{28}$$

2) Forward Registration In the *forward* HMM-based registration, let us define

$$\mathbf{y}'_m = \mathbf{y}_m - \boldsymbol{\mu}_y,\ \boldsymbol{\mu}_j^{x'} = \boldsymbol{\mu}_j^x - \boldsymbol{\mu}_{V_x}. \tag{29}$$

Similar to the bidirectional scenario, let us redefine $\mathbf{H}$

$$\mathbf{H} = \frac{1}{\sigma_x^2}\sum_{m=1}^{M}\sum_{j=1}^{J}\gamma_{mj}\mathbf{y}'_m(\boldsymbol{\mu}_j^{x'})^\mathsf{T} + \kappa_x\sum_{m=1}^{M}\sum_{j=1}^{J}\gamma_{mj}\widehat{\mathbf{y}}_m(\widehat{\boldsymbol{\mu}}_j^{x})^\mathsf{T}. \tag{30}$$

whose matrix form is

$$\mathbf{H} = \frac{1}{\sigma^2}\mathcal{Y}'\Gamma_{\mathcal{Y}}(\boldsymbol{\mu}^{x'})^\mathsf{T} + \kappa\widehat{\mathcal{Y}}\Gamma_{\mathcal{Y}}(\widehat{\boldsymbol{\mu}}^{x})^\mathsf{T}, \tag{31}$$

where $\mathcal{Y}' = [\mathbf{y}'_1, ..., \mathbf{y}'_M] \in \mathbb{R}^{3\times M}$ and $\boldsymbol{\mu}^{x'} = [\boldsymbol{\mu}_1^{x'}, ..., \boldsymbol{\mu}_J^{x'}] \in \mathbb{R}^{3\times J}$ denote the demeaned positional vectors computed with Eq. (29), $\widehat{\mathcal{Y}}$ is reloaded as $\widehat{\mathcal{Y}} = [\widehat{\mathbf{y}}_1, ..., \widehat{\mathbf{y}}_M] \in \mathbb{R}^{3\times M}$ and $\widehat{\boldsymbol{\mu}}^x = [\widehat{\boldsymbol{\mu}}_1^x, ..., \widehat{\boldsymbol{\mu}}_N^x] \in \mathbb{R}^{3\times N}$. The $\Gamma_{\mathcal{X}}$ and $\Gamma_{\mathcal{Y}}$ are correspondence probability matrices estimated by the correspondence network. The rotation matrix is computed from $\mathbf{H}$ in a similar way as Eq. (28) in the bi-directional case.

D. Loss function

Given the ground-truth transformation matrix $\mathbf{T}_{\text{gt}} \in \mathbb{R}^{4\times4}$ that transforms $\mathcal{D}_{\mathcal{Y}}$ to $\mathcal{D}_{\mathcal{X}}$, the mean-squared error is minimized as

$$||\mathsf{T}_{yx}\mathsf{T}_{\text{gt}}^{-1} - \mathbf{I}||^2 + ||\mathsf{T}_{xy}^{-1}\mathsf{T}_{\text{gt}} - \mathbf{I}||^2, \tag{32}$$

where T_{yx} and T_{xy} are transformation matrixes that transform $\mathcal{D}_{\mathcal{Y}}$ to $\mathcal{D}_{\mathcal{X}}$ and transform $\mathcal{D}_{\mathcal{X}}$ to $\mathcal{D}_{\mathcal{Y}}$ respectively, and $\mathbf{I} \in \mathbb{R}^{4\times4}$ is the identity matrix.

VI. Experiments and Results

We utilize the medical 3D shape dataset to compare the registration performance of various methods, i.e., MedshapeNet [16], which contains more than 100,000 medical shapes including bones, livers, lungs, medical instruments, etc. We have conducted two sets of experiments, i.e., full-to-full and partial-to-full point set registrations. For full-to-full registration experiments, we evaluate DeepBHMR(i.e., *Ours*) against conventional methods including ICP [17], CPD [12], BCPD [13], deep-learning-based methods including FMR [9], PointNetLK++ [18], hybrid registration methods including DeepGMR [10], UGMM [11]. For partial-to-full registration experiments, we add other three methods, i.e., Predator[19], RegTR[20], OGMM[22].

Implementation Details. In this paper, we choose 291 femur and 260 hip shapes for the targeted computer-assisted orthopedic surgery. We first normalize each point set within a unit box $[-1, 1]^3$ and then sample 1024 points from the model surface. During the training process, three rotation angles in the range $[0, 45]^\circ$ and translation distance magnitudes in the range $[0, 50]$mm are sampled on each axis independently. During the test process, we denormalize point sets by multiplying the scale of each point set to measure the accuracy of registration in millimeter units. In the correspondence network f_ϕ introduced in Sect. V-A, PointNet with TNet is first utilized as the backbone to extract high-dimensional features from the generalized point sets. Taking $\mathcal{D}_{\mathcal{X}} \in \mathbb{R}^{6\times1024}$ as an example, the global high-dimensional features $\mathbf{g}_x \in \mathbb{R}^{1024\times1}$ are concatenated with point-wise local features $\mathbf{l}_x \in \mathbb{R}^{N\times1024}$, and together with the Softmax layer the point-to-distribution correspondences can be estimated. All learning-based registration methods are implemented in Pytorch and trained on NVidia GeForce RTX 3090 GPU.

Evaluation Metrics. We use the rotation error $\text{Error}_{\text{Rot}} = \arccos\left[\frac{\text{tr}\left(\mathbf{R}_{\text{true}}\mathbf{R}_{\text{cal}}^\top\right)-1}{2}\right] \times \frac{180^\circ}{\pi}$ and the translation error $\text{Error}_{\text{Trans}} = \|\mathbf{t}_{\text{cal}} - \mathbf{t}_{\text{true}}\|_2$ to evaluate registration accuracy, which is measured in degrees and millimeters respectively. Following DeepGMR, we also assess the average root-mean-square-error (RMSE) and Registration recall, which is the fraction of registration with RMSE less than a threshold.

$$\text{RMSE} \approx \frac{1}{\mathcal{N}}\sqrt{\sum_{i=1}^{\mathcal{N}}\|\mathsf{T}_{\text{gt}}(\mathbf{y}_i) - \mathsf{T}_{yx}(\mathbf{y}_i)\|^2}, \tag{33}$$

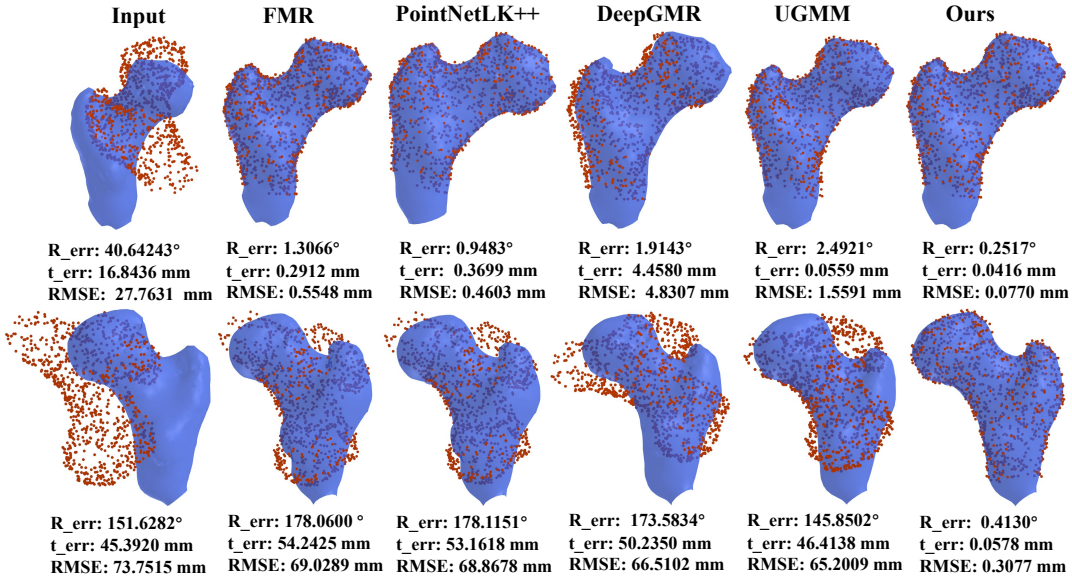

Fig. 2. Qualitative full-to-full registration results on the noisy femur bones, where the source and target point sets are shown in red points and blue surface respectively. The registration error values of each trial are labeled below the plot.

TABLE I
COMPARISON ON NOISY POINT CLOUDS ON MEDSHAPENET FEMUR DATASET[16].

Method	Rot (°)	Trans (mm)	RMSE (mm)	Recall (%)
ICP[17]	22.61	11.73	21.75*	11.00
CPD[12]	7.57	4.69	5.30*	87.00
BCPD[13]	22.68	12.48	23.13*	10.00
FMR[9]	**0.70**	4.34	4.37*	91.75
PointNetLK++[18]	1.39	0.86	1.11*	99.66
DeepGMR[10]	1.28	0.60	0.99*	**100.00**
UGMM[11]	2.85	1.36	1.22*	95.38
DeepBHMR(*Ours*)	1.01	**0.31**	**0.51**	**100.00**

TABLE II
GENERALIZATION PERFORMANCE ON MEDSHAPENET FEMUR DATASET [16] WITH LARGER TRANSFORMATION.

Method	Rot (°)	Trans (mm)	RMSE (mm)	Recall (%)
ICP[17]	90.32	25.73	81.97*	0.69
CPD[12]	73.99	35.18	34.38*	47.00
BCPD[13]	85.28	38.87	68.45*	0.00
FMR[9]	75.96	48.55	57.85*	25.09
PointNetLK++[18]	79.33	38.28	42.74*	50.17
DeepGMR[10]	25.26	13.23	17.73*	84.19
UGMM[11]	32.39	22.38	14.00*	42.31
DeepBHMR(*Ours*)	**3.47**	**2.08**	**3.05**	**96.91**

where $\mathbf{y}_i$ denotes a point sampled from the source point set. In our evaluation, we use $\mathcal{N} = 500$. To verify the statistical significance of the improvement, p-values are computed by paired test between the RMSE of DeepBHMR and the RMSE of compared methods at $\alpha = 0.01$ significance level. Significant differences with respect to RMSE are denoted with an asterisk (*) in tables.

A. Evaluation on the Femur Dataset

In this experiment, we divided the Femur Dataset into two subsets, allocating 1,166 femur models for training and 291 for testing. Each pair of point sets was subject to isotropic Gaussian noise, with the absolute value of noise varying between 0 mm and 3.81 mm. The localisation error of common optical tracking systems, e.g., the NDI Polaris P4, is usually smaller than 1 mm [23]. Consequently, the utilised Gaussian noise setting not only reflects the actual tracking error magnitude typical in real-world scenarios but also indeed poses an even greater challenge for registration approaches. According to [24], a registration is deemed successful when the RMSE is below 10 mm, which is set as the threshold to compute Recall in this experiment.

Table I compares performances of various methods on noise-affected femur bone shapes. DeepBHMR significantly outperforms other approaches in terms of both RMSE and RR, demonstrating the significance of incorporated normal vector information and the utilised bidirectional mechanism. Although FMR exhibits a slightly lower rotational error value than ours, it demonstrates a significantly higher translational error, leading to an inferior overall registration outcome. Thus, our algorithm proves to be more effective than the existing registration methods. The mean rotation error of DeepBHMR is about 1° (i.e., 1.01°) and the mean translation error is less than 1 mm (i.e., 0.31mm), which satisfies the clinical criteria for surgical navigation [25]. Furthermore, we compute p-values of paired t-tests using RMSE values of DeepBHMR and those using other methods. Additionally, the p-value of the t-test on RMSE of DeepBHMR and FMR is less than 0.01, which also validates the improvement of DeepBHMR. Fig. 2 (cf., the first row) shows the qualitative registration results on the noisy femur model.

B. Generalization to Larger Transformations

Table II includes the results on the noisy femur point sets with larger rotation angles and translation magnitudes

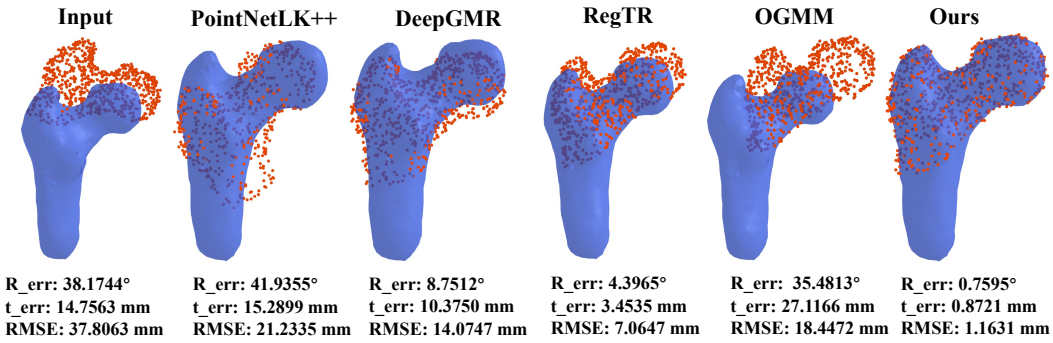

Fig. 3. Qualitative results of partial-to-full registrations on the noisy femur bones, where source and target point sets are shown in red and blue respectively. The registration error values of each trial are also shown.

TABLE III
PARTIAL-TO-FULL REGISTRATION RESULTS ON MEDSHAPENET FEMUR DATASET[16].

Method	Rot (°)	Trans (mm)	RMSE (mm)	Recall (%)
ICP[17]	24.45	14.00	22.19*	10.31
CPD[12]	16.82	16.90	17.03*	35.00
BCPD[13]	21.42	12.08	21.17*	10.00
DeepGMR[10]	9.50	18.31	19.88*	43.99
FMR[9]	14.04	14.51	15.77*	25.09
PointNetLK++[18]	20.86	14.50	14.85	52.58
UGMM[11]	59.70	26.88	33.66*	0.00
Predator[19]	8.71	11.27	13.22*	62.30
RegTR[20]	4.29	4.11	7.79*	81.44
OGMM[22]	11.19	7.07	19.05*	44.67
DeepBHMR(*Ours*)	**3.73**	**4.04**	**4.01**	**95.53**

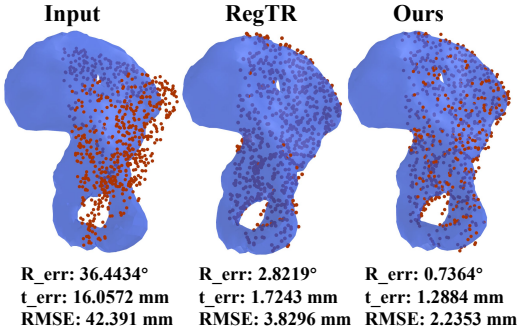

Fig. 4. Qualitative results of partial-to-full registrations on the noisy femur bones, where source and target point sets are shown in red and blue respectively. The error values of each registration trial are also shown.

sampled in the range of $[0, 180]^\circ$ and in $[0, 100]$ mm, respectively. In these challenging circumstances, DeepBHMR still achieves remarkable performances (e.g., $3.47^\circ/2.08$mm) while other approaches fail to align the two point sets (e.g., $25.26^\circ/13.23$mm for DeepGMR). The results demonstrate DeepBHMR's strong ability to handle the global registration problem, which are very desirable in image-to-patient registration, where the cumbersome additional landmark-based registration step is still needed to coarsely align the two point sets in the current clinical routine. DeepBHMR sheds light on the possibility of eliminating the coarse registration procedure.

C. Partial-to-Full Registration

As a typical clinical application scenario of DeepBHMR, computer-assisted orthopedic surgery involves collecting intraoperative points from a bone surface using an optically tracked pointer. Usually, the intraoperative point set denotes the partial subset of the full model reconstructed from preoperative CT or MRI. We generate the partial point sets, in a similar way as those in RegTR [20] and Predator [19]. Specifically, we randomly select a point, preserve its K nearest neighbors, and discard other points located far from the randomly selected point. The overlap ratio between the partial point set and the full point set is set as 70%. Table III includes the mean values of evaluative metrics, which clearly shows that DeepBHMR outperforms both RegTR [20] and Predator [19]. Unlike ReTR, Predator, and OGMM, which calculate an overlap score for each point, DeepBHMR bypasses this step, showcasing the advantages of integrating the normal vector information and the bidirectional mechanism. All p-values are smaller than 0.01, which demonstrates that improvements of DeepBHMR over existing registration approaches are statistically significant at $\alpha = 0.01$ level. Fig. 3 shows the qualitative results, which demonstrate the superiority of DeepBHMR in partial-to-full femur registration.

D. Generalization on the Hip Dataset

This sub-experiment aims to evaluate the generalizability of registration methods to different bone shapes in the scenario of partial-to-full registration. More specifically, the trained models in Sect. VI-C using the femur data are now leveraged to test on the hip shapes. Fig. 4 shows the qualitative results of RegTR [20] and DeepBHMR, which again validates DeepBHMR's superior performance (i.e., $0.74^\circ/1.29$ mm) and also the better generalizability to distinct shapes than RegTR [20] (i.e., $2.82^\circ/1.72$ mm). We posit that the superiority stems from the incorporation of normal vectors that act as valuable and effective features and the utilization of a bidirectional registration mechanism that imposes additional strong constraints.

E. Ablation Study

Table IV includes the results of ablation studies, conducted under the same experimental conditions described in Sect.

TABLE IV

ABLATION STUDIES ON THE FEMUR BONE MODELS.

Method	Rot (°)	Trans (mm)	RMSE (mm)	Recall (%)
CPD[12]	16.82	16.90	17.03	35.00
DeepGMR[10]	9.50	18.31	19.88	43.99
DeepGMR (bi)	9.95	12.59	12.56	49.83
DeepBHMR (*forward*)	6.29	9.46	9.45	83.51
DeepBHMR	**3.73**	**4.04**	**4.01**	**95.53**

VI-C. We observe from Table IV (1) the performance enhancement of DeepBHMR (*forward*) over DeepGMR [10] underscores the utility of incorporating additional normal information; (2) the improvements of DeepGMR (bi) over DeepGMR, and DeepBHMR over DeepBHMR (*forward*), validate the benefits of the bidirectional mechanism; (3) the superior performance of DeepBHMR compared to DeepGMR highlights the effectiveness of both employing additional normal vectors and a bidirectional mechanism within the network.

VII. CONCLUSIONS

This paper proposes DeepBHMR, a novel and effective learning-based generalized point set registration approach that leverages normal vectors and considers both the forward and backward registration processes. Extensive full-to-full registration experiments on the human bone dataset validate the remarkable registration accuracy and strong capability to handle global registration. In addition, DeepBHMR exhibits significantly improved performances in partial-to-full registration. The effectiveness of incorporating normal vectors and the bidirectional registration mechanism has been validated in the ablation study. Although DeepBHMR has shown utility in partial-to-full registration, further optimization is still needed for tasks with low overlap rates. This can be improved through the prediction of overlap scores. Future work plans to adopt the overlap prediction module to improve partial-to-full registration performance. We believe DeepBHMR is useful for computer-assisted interventions that require accurate, robust, and global registrations.

REFERENCES

[1] Z. Min, A. Zhang, Z. Zhang, J. Wang, S. Song, H. Ren, and M. Q.-H. Meng, "3d rigid point set registration for computer-assisted orthopedic surgery (caos): A review from the algorithmic perspective," *IEEE Transactions on Medical Robotics and Bionics*, 2023.

[2] A. Zhang, Z. Min, Z. Zhang, and M. Q.-H. Meng, "Generalized point set registration with fuzzy correspondences based on variational bayesian inference," *IEEE Transactions on Fuzzy Systems*, vol. 30, no. 6, pp. 1529–1540, 2022.

[3] Z. Zhang, Z. Min, A. Zhang, J. Wang, S. Song, and M. Q.-H. Meng, "Reliable hybrid mixture model for generalized point set registration," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–10, 2021.

[4] Z. Min, A. Zhang, D. Zhu, J. Pan, Z. Zhang, and M. Q.-H. Meng, "Anisotropic 6-d rigid point set registration," *IEEE Transactions on Instrumentation and Measurement*, vol. 72, pp. 1–19, 2023.

[5] X. Jiang, J. Ma, G. Xiao, Z. Shao, and X. Guo, "A review of multimodal image matching: Methods and applications," *Information Fusion*, vol. 73, pp. 22–71, 2021.

[6] A. Zhang, Z. Min, Z. Zhang, X. Yang, and M. Q.-H. Meng, "Anisotropic generalized bayesian coherent point drift for point set registration," *IEEE Transactions on Automation Science and Engineering*, vol. 20, no. 1, pp. 495–505, 2023.

[7] A. Zhang, Z. Min, L. Liu, and M. Q.-H. Meng, "Bidirectional generalised rigid point set registration," in *2023 IEEE International Conference on Robotics and Automation (ICRA)*, 2023, pp. 6873–6879.

[8] Z. Zhang, A. Zhang, J. Lai, H. Ren, R. Song, Y. Li, M. Q.-H. Meng, and Z. Min, "Ghmm: Learning generative hybrid mixture models for generalized point set registration in computer-assisted orthopedic surgery," *IEEE Transactions on Medical Robotics and Bionics*, vol. 6, no. 3, pp. 1285–1295, 2024.

[9] X. Huang, G. Mei, and J. Zhang, "Feature-metric registration: A fast semi-supervised approach for robust point cloud registration without correspondences," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 11 366–11 374.

[10] W. Yuan, B. Eckart, K. Kim, V. Jampani, D. Fox, and J. Kautz, "Deepgmr: Learning latent gaussian mixture models for registration," in *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part V 16*. Springer, 2020, pp. 733–750.

[11] X. Huang, S. Li, Y. Zuo, Y. Fang, J. Zhang, and X. Zhao, "Unsupervised point cloud registration by learning unified gaussian mixture models," *IEEE Robotics and Automation Letters*, vol. 7, no. 3, pp. 7028–7035, 2022.

[12] 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.

[13] O. Hirose, "A bayesian formulation of coherent point drift," *IEEE transactions on pattern analysis and machine intelligence*, vol. 43, no. 7, pp. 2269–2286, 2020.

[14] Z. Zhang, E. Lyu, Z. Min, A. Zhang, Y. Yu, and M. Q.-H. Meng, "Robust semi-supervised point cloud registration via latent gmm-based correspondence," *Remote Sensing*, vol. 15, no. 18, p. 4493, 2023.

[15] A. Banerjee, I. S. Dhillon, J. Ghosh, S. Sra, and G. Ridgeway, "Clustering on the unit hypersphere using von mises-fisher distributions." *Journal of Machine Learning Research*, vol. 6, no. 9, 2005.

[16] 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.

[17] 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.

[18] X. Li, J. K. Pontes, and S. Lucey, "Pointnetlk revisited," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2021, pp. 12 763–12 772.

[19] 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.

[20] Z. J. Yew and G. H. Lee, "Regtr: End-to-end point cloud correspondences with transformers," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2022, pp. 6677–6686.

[21] 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.

[22] G. Mei, F. Poiesi, C. Saltori, J. Zhang, E. Ricci, and N. Sebe, "Overlap-guided gaussian mixture models for point cloud registration," in *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, 2023, pp. 4511–4520.

[23] R. Elfring, M. de la Fuente, and K. Radermacher, "Accuracy of optical localizers for computer aided surgery," in *World Congress on Medical Physics and Biomedical Engineering, September 7-12, 2009, Munich, Germany: Vol. 25/6 Surgery, Nimimal Invasive Interventions, Endoscopy and Image Guided Therapy*. Springer, 2009, pp. 328–330.

[24] S. D. Billings, E. M. Boctor, and R. H. Taylor, "Iterative most-likely point registration (imlp): A robust algorithm for computing optimal shape alignment," *PloS one*, vol. 10, no. 3, p. e0117688, 2015.

[25] L. Xu, H. Zhang, J. Wang, A. Li, S. Song, H. Ren, L. Qi, J. J. Gu, and M. Q.-H. Meng, "Information loss challenges in surgical navigation systems: From information fusion to ai-based approaches," *Information Fusion*, vol. 92, pp. 13–36, 2023.