EG-Net: An Edge-Guided Network for Rigid Registration of Laparoscopic Low-Overlap Point Clouds

Wenbin Wu[1,2], Yifan Gao[1,3], Yixiu Wang[4], Jiayi Zhang[1,2,5], Yiming Zhao[4], and Xin Gao[2,5(✉)]

[1] School of Biomedical Engineering, Division of Life Sciences and Medicine, University of Science and Technology of China, Hefei, Anhui, 230026, P.R.China
wenbinwu@mail.ustc.edu.cn, yifangao@mail.ustc.edu.cn

[2] Suzhou Institute of Biomedical Engineering and Technology, Chinese Academy of Sciences, Suzhou, China
zhangjy@sibet.ac.cn

[3] Shanghai Innovation Institute, Shanghai, China

[4] Department of Hepatic Surgery, Fudan University Shanghai Cancer Center, Shanghai,China
ikkyuwang@foxmail.com, gomas1711@163.com

[5] Jinan Guoke Medical and Technology Development Co., Ltd., Jinan, China
xingaosam@163.com

Abstract. Laparoscopic augmented reality (LAR) enables real-time visualization of internal organ anatomy, effectively reducing surgical risks. Rigid point cloud registration aligns the spatial position of the preoperative image point cloud with the intraoperative laparoscopic video point cloud, playing a pivotal role in the virtual-real fusion visualization for LAR. However, the limited field of view in laparoscopic surgery results in only partial visibility of the organ. This leads to an incomplete video point cloud that exhibits low overlap with the image point cloud, rendering registration highly susceptible to local optima. Moreover, the smooth and texture-deficient organ surface makes popular superpoint matching methods based on feature similarity ineffective. Inspired by the highly consistent morphology of the video and image point clouds at organ bottom edges, we propose an edge guidance (EG) mechanism to address the challenge of sparse surface features in laparoscopic scenes. The EG mechanism identifies edge points by calculating the standard deviation of correlations among neighboring points, prioritizes edge alignment, and subsequently guides the matching of other points. We leverage this mechanism to develop an edge-guided rigid point cloud registration network, EG-Net. Compared with the state-of-the-art method PARE-Net, EG-Net achieves at least a 7% improvement in accuracy and an 11% increase in speed across three laparoscopic datasets: the public DePoLL dataset, a pig liver surgery dataset, and a human liver surgery dataset. With its high accuracy, fast speed, and strong generalization, EG-Net holds significant potential for clinical applications in laparoscopic surgery. The code is available at: https://github.com/FDC-WuWeb/EG-Net.

Keywords: Rigid registration · Low-overlap point clouds · Laparoscopic augmented reality · Edge Guidance.

1 Introduction

As a representative of minimally invasive surgery, laparoscopic surgery offers benefits with minimal trauma, rapid recovery, and fewer complications. However, in surgeries involving internal organs, such as liver and kidney tumor resection, laparoscopic surgery faces limitations due to its inability to visualize the internal structures of organs, resulting in high surgical risks. Laparoscopic augmented reality (LAR) [1] addresses this issue by reconstructing the organ's anatomical structures from preoperative images into a virtual 3D model. The model is superimposed onto the intraoperative laparoscopic view to achieve real-time internal organ visualization. The key technology of LAR is rigid registration between the virtual 3D surface model (image point cloud) and the laparoscopic video (video point cloud). However, the field of view is limited in laparoscopic scenes, and only a portion of the organ is visible. Consequently, the video point cloud is incomplete and is much smaller than the image point cloud (Fig. 1(a)). After registration, the overlap region is less than 30% of the image point cloud (Fig. 1(b)). This low-overlap characteristic, combined with the organ's smooth surface and lack of texture features, makes registration prone to falling into a local optimum (Fig. 1(c)).

The traditional methods iterative closest point (ICP) [2] and global iterative closest point (Go-ICP) [3] rely on the nearest neighbor search to force the matching of closest points. Due to inaccuracies in forced correspondences, these methods are prone to get stuck in local minima and suffer from time costs for iteration. The deep learning method PCRNet [4] directly solves the transformation matrix, offering faster speed but with limited accuracy due to neglecting the low-overlap scenes. To address this, Predator [5] introduces the concept of superpoint matching. Point clouds are downsampled into superpoints containing features of local regions. Correspondences between superpoints are calculated based on feature similarity to match the local overlap regions. Nevertheless, Predator overlooks the geometric structure of local point clouds, resulting in matching errors. P2PNet [6], based on kernel point convolution (KPConv) [7], encodes the

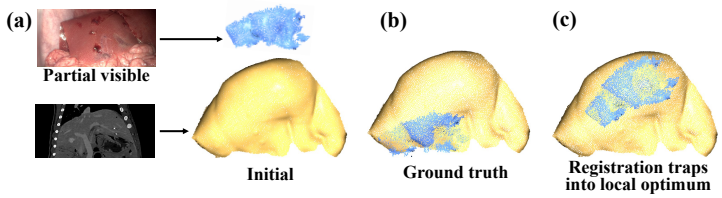

Fig. 1. Illustration of low-overlap point clouds in a laparoscopic scene. The blue represents the video point cloud, and the yellow represents the image point cloud.

geometric information from each point's neighborhood to capture local geometric structures. PARE-Net [8] incorporates rotation-invariant convolutions, improving the network's robustness to point cloud rotations. However, in laparoscopic scenes, organ surfaces are smooth and texture features are sparse, resulting in extremely high feature similarity among local regions and superpoints. Existing superpoint matching methods struggle to accurately match high-similarity superpoints, failing to guarantee globally optimal registration.

Inspired by the consistent morphology of the video and image point clouds at the organ bottom edges, we propose an edge guidance (EG) mechanism to address the challenge in laparoscopic scenes. EG mechanism identifies edge points by calculating the standard deviation of correlations between neighboring points, prioritizes edge points aligning, and subsequently guides the matching of non-edge points. Based on this mechanism, we construct the EG-Net to align the overlapping regions precisely. We compared EG-Net with state-of-the-art methods using three (one public and two private) datasets, i.e., the public laparoscopic porcine liver dataset DePoLL [9], the laparoscopic porcine liver dataset from People's Hospital of Suzhou High-tech District (referred to as Suzhou), and the laparoscopic human liver surgery dataset from Fudan University Shanghai Cancer Center (referred to as Fudan) to validate its performance and generalizability.

2 Method

2.1 EG-Net architecture

As shown in Fig. 2, EG-Net consists of a point cloud feature extraction and a point matching part. The source point cloud P and target point cloud Q

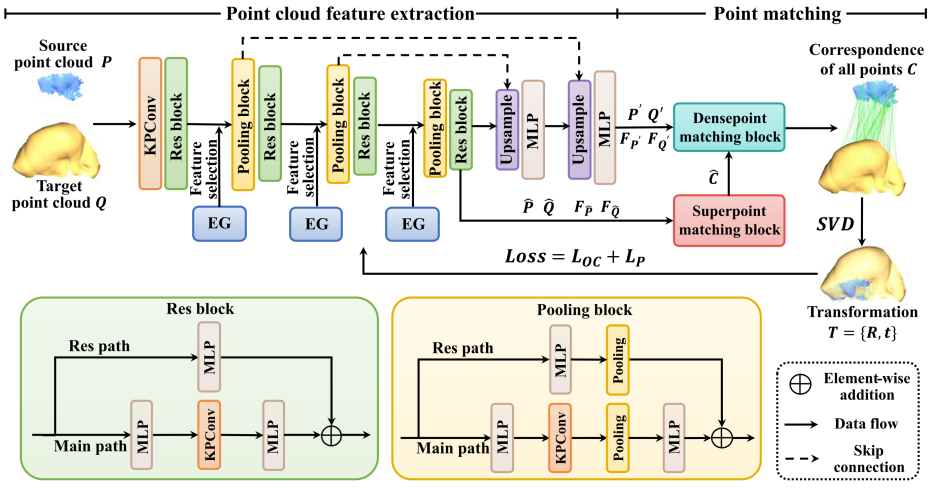

Fig. 2. The architecture of EG-Net for rigid low-overlap point cloud registration.

are fed into KPConv to extract shallow topological features, such as neighbor relationships. Then, four Res blocks and three Pooling blocks are alternately stacked, gradually extracting deeper semantic features while downsampling point clouds into superpoints. The Res block shown in the green box uses the Res path to preserve shallow topological features while learning deeper semantic features. The Pooling block shown in the yellow box is based on the Res block, with max-pooling layers added to both paths to perform downsampling and feature aggregation. The EG mechanism is applied before each Pooling block to identify edge points and select their features, guiding the matching of high-similarity superpoints in non-edge regions. The superpoints $\hat{P}$ and $\hat{Q}$ output by the last Res block, along with their corresponding local features, $F_{\hat{P}}$ and $F_{\hat{Q}}$, are sent to the upsampling layer and the superpoint matching block. An MLP layer follows each upsampling layer, and the final one outputs the upsampled point clouds $P^{'}$ and $Q^{'}$ as well as their features $F_{P'}$ and $F_{Q'}$. Along with correspondences $\hat{C}$ calculated by the superpoint matching block, they are fed into the dense point matching block to generate correspondences C for all points. Finally, singular value decomposition (SVD) is used to solve the rigid transformation matrix $T = \{R, t\}$ from C. The overlap-aware circular loss L_{OC} [6] and point matching loss L_P [10], which have been explained in [6] and [8], are used to train the EG-Net.

2.2 Point Cloud Edge Guidance Mechanism

To address the challenge of sparse texture features in laparoscopic scenes, we propose an EG mechanism shown in the blue box of Fig. 3. For a given candidate point $p_i \in P$, K-nearest neighbors (KNN) is used to search for its neighbors $N(p_i) = \{p_1, p_2, \ldots, p_K\}$. Two 1*1 convolutions are applied to calculate the features of p_i and $N(p_i)$. By performing matrix multiplication and Softmax

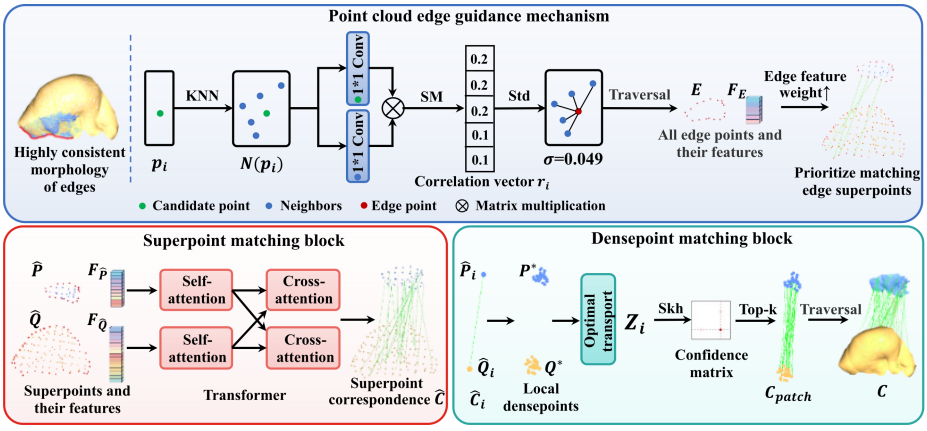

Fig. 3. The point cloud edge guidance (EG) mechanism (blue box), superpoint matching (red box) and densepoint matching (cyan box) blocks.

normalization (SM), correlation vector r_i between p_i and each of its neighbors is obtained. r_i and its standard deviation (Std) σ is calculated as follows:

$$r_{i,k} = SM(\frac{(W_i \cdot F(p_i))^T (W_{i,k} \cdot F(p_{i,k}))}{\sqrt{d}}), \sigma = \sqrt{\frac{1}{K}\sum_{k=1}^{K}(r_{i,k} - \frac{1}{K}\sum_{k=1}^{K} r_{i,k})^2} \tag{1}$$

where, $p_{(i,k)} \in N(p_i)$, $W \cdot F$ represents the convolution and d is the feature dimension. A lower σ indicates the values in r_i are similar, meaning the neighbors of p_i are evenly distributed, corresponding to the smooth non-edge region of the point cloud. Conversely, a higher σ suggests significant differences in the values of r_i, meaning the neighbors of p_i are unevenly distributed, corresponding to the edge. Points with a σ above a certain threshold (empirically set at 0.04) are considered edge points. Finally, P and Q are traversed to identify all edge points and select edge features. By increasing the weight of edge features in $F_{\hat{P}}$ and $F_{\hat{Q}}$, the superpoint matching block is trained to prioritize the precise matching of edge superpoints, facilitating initial coarse localization of point clouds and guiding the matching of highly similar non-edge regions.

As shown in the red box, the superpoint matching block is built upon a Transformer-based self- and cross-attention structure [6]. This structure enables interaction between two feature groups, enabling global matching. As shown in the cyan box, for each superpoint correspondence $\hat{C}_i$, the dense point matching block expands it to local dense points P^* and Q^* based on $P^{'}$ and $Q^{'}$, subsequently sends them to optimal transport layer [11], which calculates the matching cost Z_i of each pair and uses Sinkhorn (Skh) algorithm [12] to generate a confidence matrix. The Top-k algorithm selects the highest-probability correspondences, generating C_{patch}. This process is repeated for each pair of superpoints, producing the final correspondences C for all points.

3 Experiments

DePoLL, Suzhou and Fudan used in this study contain 13, 30, and 10 cases, each comprising a preoperative liver CT sequence and an intraoperative laparoscopic video (Data will be made available on request). The liver is segmented by a senior physician and reconstructed into a 3D model, from which surface points are uniformly sampled to create the image point cloud. The video point cloud is generated from the video using MetaShape software. The physician performs rigid registration to align two point clouds for each case. Point clouds are uniformly resized to below 25mm, with a minimum point distance of 0.015mm to reduce computational load. Data augmentation is applied to DePoLL Case 1(Fig. 4), generating 11,200 pairs split into a 9:1 training-validation set. The remaining 52 cases are used for testing.

EG-Net is trained using the Adam optimizer in a PyTorch 1.7.1 and CUDA 11.0 environment. The batch size, learning rate, and training epochs are set to 1, 1e-4, and 80, respectively. Relative rotation error (RRE) measures the geodesic

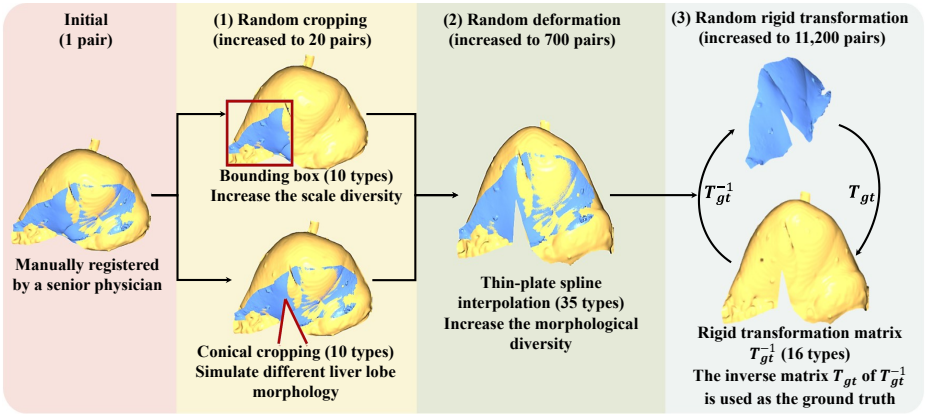

Fig. 4. Data augmentation flowchart.

distance between the predicted and ground truth rotation matrices. The relative translation error (RTE) quantifies the Euclidean distance between the predicted and ground truth translation vectors. The DePoLL dataset provides 15 liver surface landmarks for calculating the target registration error (TRE). Lower values of all three metrics indicate better performance.

4 Results and discussion

We perform comparative and ablation studies to evaluate the performance of EG-Net. Accuracy is assessed using RRE, RTE, and TRE, while registration speed is measured by time. The ablation method (EG-Net w/o EG) is excluded from the metric rankings. Results are presented in Table 1.

A comprehensive comparison of all methods reveals that superpoint matching methods (Predator, etc.) outperform traditional (ICP, etc.) and direct-solving (PCR-Net) methods. EG-Net achieves the highest accuracy and the second-fastest registration speed across three datasets. Traditional methods perform poorly due to their reliance on nearest neighbor search, leading to inaccurate correspondences and transformations. The direct-solving method, PCR-Net, neglects low-overlap scenes and mistakenly uses global features of non-overlapping regions to calculate the transformation, resulting in registration errors. Although PCR-Net is the fastest, the insufficient accuracy limits its practical applicability.

EG-Net outperforms the other superpoint matching methods, reducing RRE, RTE, TRE, and Time by 12%, 14%, 7%, and 11% on the DePoLL dataset compared with the second-best PARE-Net. This improvement stems from the proposed EG mechanism, enabling EG-Net to match edge superpoints accurately, facilitating initial coarse localization of point clouds and guiding the matching of highly similar non-edge superpoints. When the EG mechanism is removed, the EG-Net w/o EG shows a significant increase in RRE, RTE, and TRE by 80%, 117%, and 50% on the DePoLL dataset, demonstrating the effectiveness of it.

Table 1. Results of the comparative and ablation studies. Bold represents the best, underline indicates second best.

Methods	**RRE**(deg)↓	**RTE**(mm)↓	**TRE**(mm)↓	**Time**(s)↓
DePoLL				
ICP$_{(\text{TPAMI'1992})}$	48.38±50.06	27.36±3.11	41.71±28.57	0.10±0.04
Go-ICP$_{(\text{TPAMI'15})}$	45.58±54.37	22.15±23.61	35.13±38.14	0.21±0.08
PCR-Net$_{(\text{arXiv'19})}$	15.19±6.59	8.58±4.15	24.48±8.95	**0.04±0.07**
Predator$_{(\text{CVPR'21})}$	8.68±4.99	5.23±3.34	7.08±4.44	0.13±0.01
P2PNet$_{(\text{TPAMI'23})}$	6.20±3.51	3.73±2.17	5.43±2.62	0.10±0.02
PARE-Net$_{(\text{ECCV'24})}$	4.78±2.76	2.74±1.37	4.62±1.73	0.09±0.01
EG-Net$_{(\text{Ours})}$	**4.19±1.75**	**2.35±0.91**	**4.29±1.55**	0.08±0.02
EG-Net w/o EG	7.58±5.42	5.11±4.57	6.44±6.61	0.07±0.01
Suzhou				
ICP$_{(\text{TPAMI'1992})}$	22.20±38.99	13.17±23.24	-	0.09±0.04
Go-ICP$_{(\text{TPAMI'15})}$	16.25±29.79	8.40±15.18	-	0.16±0.04
PCR-Net$_{(\text{arXiv'19})}$	46.05±10.38	28.81±5.40	-	**0.03±0.05**
Predator$_{(\text{CVPR'21})}$	9.94±8.65	6.30±6.64	-	0.10±0.02
P2PNet$_{(\text{TPAMI'23})}$	10.01±8.32	5.95±5.53	-	0.10±0.02
PARE-Net$_{(\text{ECCV'24})}$	9.11±18.44	5.66±13.13	-	0.09±0.01
EG-Net$_{(\text{Ours})}$	**7.36±4.59**	**4.65±3.62**	-	0.08±0.01
EG-Net w/o EG	9.58±8.82	6.27±6.09	-	0.07±0.01
Fudan				
ICP$_{(\text{TPAMI'1992})}$	34.15±38.70	16.76±18.36	-	0.09±0.06
Go-ICP$_{(\text{TPAMI'15})}$	62.14±58.37	30.80±26.25	-	0.17±0.01
PCR-Net$_{(\text{arXiv'19})}$	44.43±24.48	21.88±14.51	-	**0.05±0.09**
Predator$_{(\text{CVPR'21})}$	54.06±47.19	29.28±18.04	-	0.10±0.02
P2PNet$_{(\text{TPAMI'23})}$	33.09±19.49	19.54±12.73	-	0.11±0.02
PARE-Net$_{(\text{ECCV'24})}$	10.73±6.86	6.39±4.57	-	0.09±0.02
EG-Net$_{(\text{Ours})}$	**8.55±4.08**	**4.98±2.61**	-	0.08±0.01
EG-Net w/o EG	21.11±12.65	11.56±6.67	-	0.07±0.01

EG-Net also outperforms other methods on Datasets Suzhou and Fudan, showing exceptional generalization. This is particularly evident on the more challenging human laparoscopic surgery dataset (Fudan), where Predator and P2PNet struggle with local optima, resulting in significant performance declines. Compared to PARE-Net, EG-Net further extends its advantage, reducing RRE and RTE by 20% and 22%. These results highlight EG-Net's remarkable suitability for human data and its strong potential for applications in laparoscopic surgery.

To visually compare the registration results, we randomly select one case from each dataset for testing. The registration results and corresponding error distance

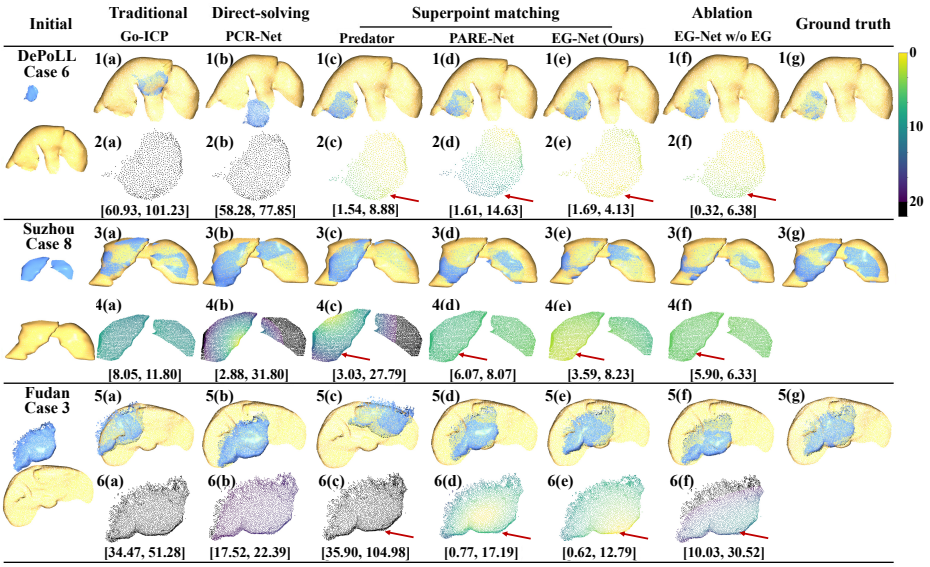

Fig. 5. Visualization of rigid point cloud registration results. The first column shows initial point clouds, while the last column displays the ground truth. Odd rows present the results of each method, while even rows show the distance heatmaps between the registered point clouds and the ground truth, along with the error range (in mm).

heatmaps relative to the ground truth are shown in Fig. 5. In heatmaps, lighter colors indicate lower error. Taking DePoLL Case 6 as an example, the similarity between registration results and the ground truth progressively improves from left to right across all methods. The distance heatmap colors gradually lighten at the same time, indicating a step-wise accuracy improvement from traditional methods to the direct-solving method and superpoint matching methods. The distance heatmap of EG-Net (2(e)) shows the lightest overall color, with the smallest maximum error distance of 4.13mm, indicating the highest accuracy. Compared with 2(c) and 2(d), 2(e) shows a notably lighter color at the edges (marked by red arrows). In contrast, 2(f) of the ablation method lacks this advantage, demonstrating the effectiveness of the EG mechanism. Visualization results on the other datasets reveal similar regularities, confirming that EG-Net is also applicable to the human laparoscopic surgery data and highlighting its strong generalization.

Although the datasets used in this study contain visible organ edges, situations where no true edge information is available may still arise in more complex scenes or during certain short-term phases of surgical procedures, which could lead to a failure of EG-Net's edge guidance. Our future work will focus on registration under such conditions.

5 Conclusion

We develop EG-Net, an edge-guided network for rigid registration of laparoscopic low-overlap point clouds. This network incorporates the proposed edge guidance mechanism to address the sparse texture features in laparoscopic low-overlap scenes. We use three laparoscopic datasets to compare EG-Net with state-of-the-art rigid point cloud registration methods. The results demonstrate EG-Net's superior performance, robust generalization, and potential for LAR applications.

Acknowledgments. This work was supported in part by National Natural Science Foundation of China under Grant 82372052.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Ali, S., Espinel, Y., Jin, Y., Liu, P., Güttner, B., Zhang, L., Dowrick, T., Clarkson, J. M., Xiao, S., Wu, Y., Yang, Y., Zhu, L., Sun, D., Li, L., Pfeiffer, M., Farid, S., Maier-Hein, L., Buc, E., Bartoli, A.: An objective comparison of methods for augmented reality in laparoscopic liver resection by preoperative-to-intraoperative image fusion from the MICCAI 2022 challenge. Med. Image Anal. **99**, 103371 (2025)
2. Besl, P., McKay, N.: A method for registration of 3D shapes. IEEE Trans. Pattern Anal. Mach. Intell. **14**(2), 239-256 (1992)
3. Yang, J., Li, H., Campbell, D., Jia, Y.: Go-icp: Aglobally optimal solution to 3d icp pointset registration. IEEE Trans. Pattern Anal. Mach. Intell. **38**(11), 2241-2254 (2015)
4. Sarode, V., Li, X., Goforth, H., Aoki, Y., Srivatsan, R. A., Lucey, S., Choset, H.: PCRNet: Point cloud registration network using pointnet encoding. arXiv preprint arXiv:1908.07906 (2019)
5. Huang, S., Gojcic, Z., Usvyatsov, M., Wieser, A., Schindler, K.: Predator: Registration of 3D point clouds with low overlap. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 4267-4276 (2021)
6. Qin, Z., Yu, H., Wang, C., Guo, Y., Peng, Y, Ilic, S., Hu, D., Xu, K.: Geotransformer: Fast and robust point cloud registration with geometric transformer. IEEE Trans. Pattern Anal. Mach. Intell. **45**(8), 9806-9821 (2023)
7. Thomas, H., Qi, R. C., Deschaud, J. E., Marcotegui, B., Goulette, F., Guibas, J. L.: Kpconv: Flexible and deformable convolution for point clouds. In: Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 6411-6420. Seoul, South Korea (2019)
8. Yao, R., Du, S., Cui, W., Tang, C., Yang, C.: PARE-Net: Position-Aware Rotation-Equivariant Networks for Robust Point Cloud Registration. In: European Conference on Computer Vision. pp. 287-303. Springer (2024)
9. Modrzejewski, R., Collins, T., Seeliger, B., Bartoli, A., Hostettler, A., Marescaux, J.: An in vivo porcine dataset and evaluation methodology to measure soft-body laparoscopic liver registration accuracy with an extended algorithm that handles collisions. Int. J. Comput. Ass. Rad. **14**, 1237-1245 (2019)

10. Sarlin, P. E., DeTone, D., Malisiewicz, T., Rabinovich, A.: SuperGlue: Learning feature matching with graph neural networks. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 4938-4947 (2020)
11. Yu. H., Li, F., Saleh, M., Busam, B., Ilic, S.: CoFiNet: Reliable coarse-to-fine correspondences for robust pointcloud registration. Advances in neural information processing systems **34**, 23872-23884 (2021)
12. Cuturi, M.: Sinkhorn distances: Lightspeed computation of optimal transport. Advances in neural information processing systems **26**, 2292-2300 (2013)