Full Title of Article

Author name(s) withheld EMAIL(S) WITHHELD
Address withheld

Editors: Under Review for MIDL 2023

Abstract

This is a great paper and it has a concise abstract.
Keywords: List of keywords, comma separated.

1. Introduction

As the population ages in many countries, the number of orthopedic consultations has increased significantly, particularly for lumbar diseases such as spinal stenosis, disc herniation, and spinal fractures (Compston et al., 2019). Surgeries represented by pedicle screw fixation and laminectomy decompression are the important operations for treating these diseases (Ghogawala et al., 2016). To reduce the operational risk and release the burden on surgeons, researchers have developed some surgical robot systems in recent years. At the same time, it is an important task to plan the robotic trajectory based on medical images.

Since manual labeling and measurement of bone topography is a time-consuming work, image registration can be considered as an ideal solution. It can be used to map a segmented and labeled image onto an new image. It typically only requires manual labeling on one or a few image templates, and has the potential to map the planned robotic trajectories. However, in orthopedic surgery, surgical planning is often based only on the bone surface (Li et al., 2021). Previous image registration algorithms focus on the correspondence of each voxel between two images and perform poorly in the surface registration tasks. Therefore, a method focused on registering bone surfaces is urgently expected to be developed.

Medical image registration plays an important role in many practical applications, such as tumor contour drawing, surgery planning and navigation for cardiovascular diseases. In recent years, many learning-based methods for image registration have been developed due to the development of deep learning technology. Some of them rely on the manually labelled ground truth (GT) warp fields (Krebs et al., 2017; Sokooti et al., 2017). Although they performed well in the registration experiments, GT is usuall difficult to obtain in many tasks. To avoid this drawback, (Vos et al., 2017) proposed a CNN-based model DIRNet which is trained by unsupervised optimization of a similarity metric between input image pairs. Another successful approach is to use auxiliary anatomical information for weakly supervised registration (Hu et al., 2018a,b). These works usually output a motion vector for each voxel point. When facing the bone surface registration, it may mean a lot of wasted computational effort and ineffective learning.

Graph networks can be viewed as a good structure for dealing with irregular structures like surfaces. In some works, researchers have achieved notable results by extracting features from images and using graph networks to process them. In (Hansen and Heinrich, 2021), the authors leverage discrete dense displacement maps in the registration of Lung CTs.

© 2023 Anonymous.

Based on the key points, a geometric network GraphRegNet was presented. (Fu et al., 2021) construct a graph from the point cloud which was processed and registered by the proposed graph network. In these works, models were designed to process the graphs in images for a few specific tasks and not suitable for the bone surface registration.

Mesh is a data structure for representing the surface of an object. It is made up of basic elements called vertices, edges, and faces. It contains the connection relationships between vertices and can be easily described and processed using a graph. Recent research has reported the exciting performance of extracting surface meshes from images. Pixel2Mesh (Wang et al., 2020) generates 3D mesh for a single perspective image. (Wickramasinghe et al., 2020) used a encoding and decoding structure and extended the work to 3D volumetric medical images. (Kong and Shadden, 2021) and (Kong et al., 2021) modified it to predict multiple anatomical structures and generate whole heart geometries. (Bongratz et al., 2022) used application-specific templates to predict the cortex. These models combine CNNs and GNNs to deform specific templates gradually. However, the purpose of these works is to extract object surface meshes which is similar to segmentation. Important correspondences before and after the template mesh deformation are ignored. Therefore, in this paper, we use a similar approach and apply it to the surface registration task.

2. Method

As mentioned earlier, in orthopedic surgery, the bone surface often plays a more important role. Some anatomical parameter measurements and surgical planning are based on the bone surface. Registration of the labeled template mesh to a new image is expected to reduce surgical planning time. However, surfaces are usually described by meshes and most of the existing medical image acquisition devices only provide voxel images. To this end, this section proposes a network based on CNN and GNN to achieve the surface registration from the template mesh to the voxel image.

2.1. Registration architecture

Since the input image and the template mesh have different data structure and information component. As shown in Figure 1, two kinds of modules are used to process them prospectively. The CNN module is for the voxel image and the GNN module takes the template mesh as input and outputs the displacement of vertices. However, learning the vertices movements from a random initial feature is difficult and it is easy to cause divergence in the network training process. Therefore, the CNN module is designed to learn the meaningful features from the segmentation task. The proposed approach utilizes the information of the input image to move the vertices of the template mesh and generate a new mesh for the target vertebra. Since there is relationship between the template vertices and the moved vertices and the two mesh shares the vertices connection, template mesh is registered with the vertebra surface of the input image.

The voxel segmentation module is a Unet-like CNN. Residual convolution blocks are introduced in the encoder and decoder. Since the input image often contains the upper and lower vertebrae besides the target one, and the two adjacent vertebrae are very close, this module is trained to segment the central vertebra only. At this time, the learned features can be considered to be more likely to contain information on the surface of the

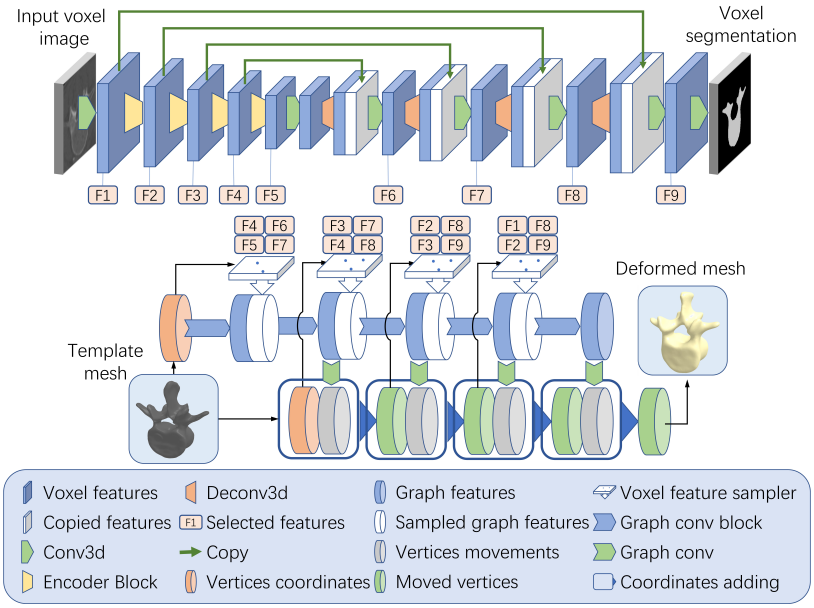

Figure 1: The proposed surface registration architecture

target vertebral segment. The network takes a vertebral template mesh and a spinal CT image as input for the two modules. The input image is manually cropped to a size of $64mm \times 96mm \times 96mm$ around the center of a specific vertebra. An image of this size can usually include a whole vertebra and parts of adjacent vertebrae.

A GNN-based module is used to process the template mesh. The coordinate values of the vertices are regarded as the initial nodes features and mesh edges describe the node connections. 4 graph blocks are designed. They are expected to learn the vertices movements from the nodes features and based on them deform the template mesh iteratively. During this, voxel features from the CNN module are sampled based on the vertices coordinates as the extra graph node features. In order to provide the network more spatial information, points on the normal direction of the vertices are also involved in voxel feature sampling. And then, all of them are concatenated with the initial node features. Graph layers are designed to process them and output displacement vector for each vertex.

In the proposed method, the template mesh was randomly chosen from the mesh label of the training dataset. It was also applied 30 times of Laplacian smoothing algorithm to remove those unsmooth and personalized details. It has realistic shape of the vertebra, rather than the simple shape mesh in previous studies (Kong et al., 2021; Wickramasinghe et al., 2020) or the highly smoothed mesh in (Bongratz et al., 2022). Such a template has actual anatomical information. When registering it to the vertebra surface of the input image, it can map the anatomical information or marked surgical plan to the new image.

2.2. Loss functions

The loss of the network consists of two parts, $\mathcal{L}_{\text{vox}}$ is used to enable the network to learn voxel segmentation tasks, and $\mathcal{L}_{\text{mesh}}$ is used to learn mesh deformation.

$$\mathcal{L} = \lambda_{\text{vox}}\mathcal{L}_{\text{vox}} + \lambda_{\text{mesh}}\mathcal{L}_{\text{mesh}} \tag{1}$$

The former is not the main purpose of the task, but it helps the network to learn the meaningful features. At the beginning of training, λ_{mesh} is set to 0 to help the network learn more valuable features from the segmentation task. And in the later stage of training, λ_{vox} is gradually reduced to 0, in order to make the network more focused on the mesh deformation task.

For the predicted surface meshes, it needs to be compared with the mesh labels. However, due to the lack of vertices correspondence between them and the difference of the vertices connection, it is difficult to define a loss function to describe the difference between the two meshes directly. In this paper, the point chamfer loss $\mathcal{L}_{\text{point}}$ and normal loss $\mathcal{L}_{\text{norm}}$ are used to describe the geometric difference of the mesh. To regularize the predicted vertices displacements, edge length loss $\mathcal{L}_{\text{edge}}$ and squared L2-norm derivatives of the displacement L_{disp} is introduced. It is expected to limit the predicted vertices displacements and smooth the predicted mesh. Therefore, the loss function for mesh deformation can be expressed as follows.

$$\mathcal{L}_{\text{mesh}} = \lambda_{point}\mathcal{L}_{\text{point}} + \lambda_{\text{norm}}\mathcal{L}_{\text{norm}} + \lambda_{\text{edge}}\mathcal{L}_{\text{edge}} + \lambda_{\text{disp}}\mathcal{L}_{\text{disp}} \tag{2}$$

2.2.1. Voxel segmentation loss

In the voxel segmentation task, the input image is trained as a binary segmentation problem for individual vertebral segments. It is expected to help the mesh deformation network to learn the boundary of the intermediate vertebral segment. We used a BCE loss in this paper to describe the difference between the predicted vertebra segmentation I_{p} and the segmentation label I_{gt}.

$$\mathcal{L}_{\text{vox}} = BCE(I_{\text{p}}, I_{\text{gt}}) \tag{3}$$

2.2.2. Weighted point distance loss

Due to the lack of correspondence between vertices in the predicted mesh and those in the GT, when calculating the point loss, the connections between the vertices are ignored. Therefore, the task is similar to a point cloud registration problem, so the chamfer loss is introduced. For a vertex from one mesh, chamfer distance calculates the distance between it and the nearest vertex in the other mesh. Because of the lack of connection relationship among vertices when calculating the point distance loss, it may cause unexpected movement or edge intersections (Gupta, 2020). It is recommended to combine other regular terms when using chamfer loss for mesh registration. Different from the practice in previous studies, the proposed network uses a specific template which was random selected from training datasets. Since the template has a realistic initial shape, it may affect the learning progress of different vertices. Furthermore, the registration results of different vertices also have different effects on the network output. Inspired by (Bongratz et al., 2022), weights defined

by curvatures are introduced to guide the network to focus on the complex structures. Since the curvature of the template at each vertex before and after the registration will not change greatly, which is also what the work pursues, the curvatures of the both two meshes were used to weight chamfer distances. For a point $\mathbf{p}$ on the surface, assuming its discrete mean curvature is $\kappa(\mathbf{p})$, the defined weight can be written as

$$k(\mathbf{p}) = min\{1 + \kappa(\mathbf{p}), k_{max}\} \tag{4}$$

Therefore, based on the definition of chamfer loss,the proposed curvature weighted point distance loss can be described as

$$\begin{aligned} \mathcal{L}_{\text{point}} = & \frac{1}{|\mathcal{P}_{\text{gt}}|} \sum_{\mathbf{u} \in \mathcal{P}_{\text{gt}}} k_{\text{gt}}(\mathbf{u}) \min_{\mathbf{v} \in \mathcal{P}_{\text{p}}} \|\mathbf{u} - \mathbf{v}\|^2 \\ & + \frac{1}{|\mathcal{P}_{\text{p}}|} \sum_{\mathbf{v} \in \mathcal{P}_{\text{p}}} k_{\text{p}}(\mathbf{v}) \min_{\mathbf{u} \in \mathcal{P}_{\text{gt}}} \|\mathbf{v} - \mathbf{u}\|^2, \end{aligned} \tag{5}$$

where, the $\mathcal{P}_{\text{p}}$ denotes the predicted mesh vertices and the $\mathcal{P}_{\text{gt}}$ the vertices of the GT mesh.

2.2.3. Normal distance loss

In addition to the point position, the normal direction is also an important information describing the surface shape. Referring to the chamfer point loss, for a vertex from the prediction mesh or target mesh, the normal distance loss calculates the cosine distance between its normal vector and that of the nearest point in another mesh.

$$\begin{aligned} \mathcal{L}_{\text{ norm}} = & \frac{1}{|\mathcal{P}_{\text{gt}}|} \sum_{\mathbf{u} \in \mathcal{P}_{\text{gt}}} 1 - \cos(\mathbf{n}(\mathbf{u}), \mathbf{n}(\tilde{\mathbf{v}})) \\ & + \frac{1}{|\mathcal{P}_{\text{p}}|} \sum_{\mathbf{v} \in \mathcal{P}_{\text{p}}} 1 - \cos(\mathbf{n}(\mathbf{v}), \mathbf{n}(\tilde{\mathbf{u}})), \end{aligned} \tag{6}$$

where $\tilde{\mathbf{v}} = \arg\min_{\mathbf{r} \in \mathcal{P}_{\text{p}}} \|\mathbf{u} - \mathbf{r}\|^2$ and $\tilde{\mathbf{u}} = \arg\min_{\mathbf{r} \in \mathcal{P}_{\text{gt}}} \|\mathbf{v} - \mathbf{r}\|^2$

2.2.4. Edge length loss

Edge length loss can be used to limit the undesired vertices movements and constrain the generated mesh. This is at the beginning of training, it can be helpful to converge the training process. However, it also limits the mesh deformation, and inhibit the network from generating accurate registration results. To deal with this dilemma, the edge length loss weight λ_{edge} is reduced gradually in the experiments. The loss is defined as follows

$$\mathcal{L}_{\text{edge}} = \sum_{\mathbf{p} \in \mathbf{P}_{\text{p}} \mathbf{k}_p \in \mathcal{N}(\mathbf{p})} \|\mathbf{p} - \mathbf{k}_{\mathbf{p}}\|_2^2, \tag{7}$$

2.2.5. Displacement regularization loss

When the network gradually deforms the template mesh, it is expected that each vertex can move together with other adjacent vertex around it. This is beneficial to maintain the same anatomical position for vertices in registration. And those non-smooth movements also need to be penalized as they usually cause wrong local protrusion or collapse of the surface. Therefore, an L2 norm displacement regularization loss is designed. For each vertex p, loss is defined as the sum of squares of the derivatives of the displacement (Rueckert et al., 1999). Obviously the displacements of those closer neighbors should have a greater influence on the computation. In this paper, normalized weights defined by the edge length were used.

$$w(\mathbf{p},\mathbf{q}) = \frac{1}{\|\mathbf{p}-\mathbf{q}\|}\left(\sum_{\mathbf{q}\in\mathcal{N}(\mathbf{p})}\frac{1}{\|\mathbf{p}-\mathbf{q}\|}\right)^{-1} \tag{8}$$

Therefore, the displacement regularization loss can be defined as

$$\mathcal{L}_{disp} = \sum_{\mathbf{p}\in\mathbf{P}_{\mathrm{p}}}\left\|disp(\mathbf{p}) - \sum_{\mathbf{q}\in\mathcal{N}(\mathbf{p})} w(p,q)disp(\mathbf{q})\right\|_2^2 \tag{9}$$

3. Experiments

The proposed network generates the surface mesh by gradually deforming the template mesh. In this process, the corresponding relationship of the mesh vertices is preserved, which makes it possible to simultaneously obtain the accurate vertebrae surface and the registration results between the template mesh and the bone surface of the input image. In this chapter, we compare the proposed method and baseline models in terms of image segmentation performance, bone surface registration performance and target registration results.

3.1. Datasets

Three online published spine CT image segmentation datasets were used for training and validation, namely Lumbar vertebra segmentation CT image datasets (LumSeg), Spine and Vertebrae Segmentation Datasets (SpiSeg), and xVertSeg datasets (xVertSeg). In these datasets, lumbar segment data are all included, so they were used in the experiments. The mixed data contain a total of 35 subjects, and for each one, the original CT image and the voxel-labeled masks for each lumbar vertebrae are given. The data were randomly divided into a training set of 24 subjects and a validation set of 11 subjects.

3.2. Pre-processing

Since the resolution of each image is different, to facilitate network processing, all images and segmentation GTs were resampled at a resolution of $0.5mm \times 0.5mm \times 0.5mm$. During the training and validation of the mesh generating network, images and segmentation labels with a size of 128*192*192 are randomly cropped around the vertebral centers.

All the vertebrae surface meshes were obtained based on the segmentation labels using the marching cube algorithm. They also went through 20 times of Laplacian smoothing to eliminate the staircase effect caused by the low resolution of the image. These generated meshes were used as the GT during the mesh generating network training and validation. Note that the vertices coordinates were adjusted according to the random selection of the vertebrae centers.

Bone surface registration can easily map measurements, surgical planning marked on the template to a new image. Therefore, considering the clinical application, we manually labelled the template mesh and the validation mesh GT. Based on the selected 9 planes, vertices were divided into 10 categories, respectively representing spinous process (SP), left lamina (LL), right lamina (RL), left articular process (LAP), right articular process (RAP), left transverse process (LTP), right transverse process (RTP), left pedicle (LP), right pedicle (RP) and vertebral body (VB). They were used to validate the registration results.

3.3. Results

The experimental results were compared with two image registration baselines, the traditional method NiftyReg and the learning based method DeepReg. The latter one was retrained with the training dataset. The labeled template mesh and the mesh GT of the validation dataset were transferred to voxel images. Baselines were applied to the template image and validation image to generate DDFs. Based on them, the marked information on the template image was mapped on these validation data.

3.3.1. Vertebra segmentation results

In the task of vertebra surface registration, a key application is the segmentation of the bone, which is necessary in some surgical planning. The segmentation mask or the surface mesh of the template vertebra were registered to the voxel images and compared with the corresponding GT. Hausdorff distance (HD), the average symmetric surface distance (ASSD), and Dice score are shown in Table 1. Some segmentation examples can be seen in Figure 2. It indicates that the performance of the proposed method is significantly better than the baselines on the vertebra segmentation task.

Table 1: Vertebra segmentation results

Models	**HD (mm)**	**ASSD (mm)**	**Dice (%)**
Ours	**1.14±0.49**	**0.54± 0.11**	**93.25± 1.47**
DeepReg	3.22±1.76	1.19± 0.46	86.50± 5.21
NiftyReg	7.97± 4.65	2.52± 1.48	71.82±15.50

The data in bold indicates that the result is significantly different from others in the T test (p value < 0.05).

3.3.2. Multi-classes segmentation results

One important advantage of image registration is the ability to establish correspondence between the images. This allows for the transfer of any markup made on the template to

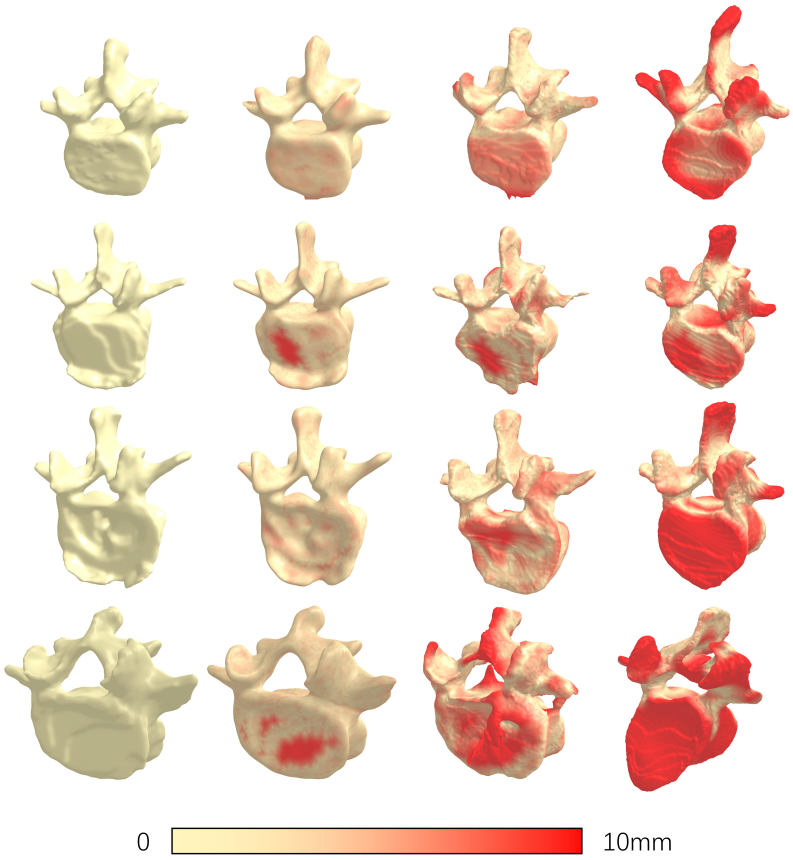

Figure 2: Example of vertebrae segmentation
The surface distance between the predicted mesh and GT mesh is illustrated with red color as shown in the color bar. The figure shows the four vertebral segmentation results of the validation dataset. The first column is the GT mesh, the second one is the prediction of the proposed model, and the third and fourth are the segmentation results of DeepReg and NiftyReg respectively.

new images, although there may be some error introduced. For vertebrae images, different operations may focus on different anatomical features. For example, in the case of pedicle screw fixation, more attention may be paid to the position, radius, and inclination of the pedicles, while in decompressive laminectomy, the position and thickness of the laminae may be more important. Therefore, the multi-class labels manually marked on the template mesh were mapped to those generated meshes based on the vertices correspondence. Figure 3 shows the predicted results and the manually marked GT.

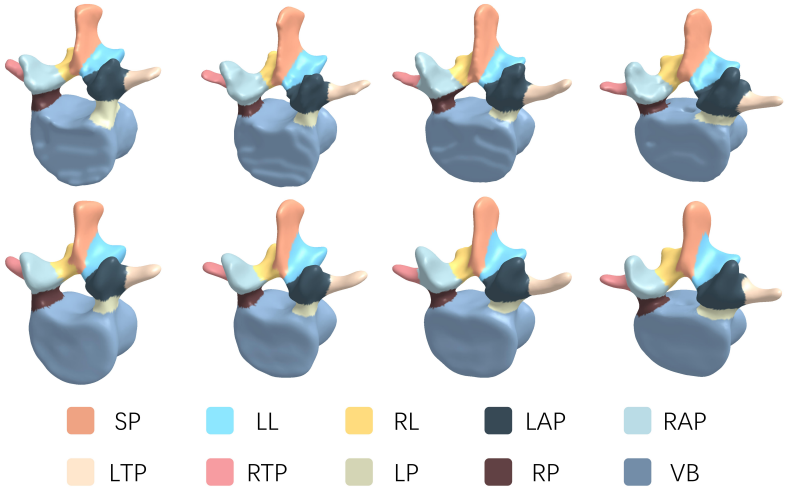

Figure 3: Example of multi-classes segmentation results
Four registration results in the validation are shown. The first row is the manually labeled classification, and the second row is the result predicted by the proposed network.

To compare the registration results and multi-class segmentation results obtained by the three methods, distances between the predicted surface and the GT mesh surface were calculated for each category. To eliminate the effect caused by inconsistencies in manual labeling, points near the boundary between two adjacent categories were excluded from the calculation. They can cause larger distance errors and obscure the expected surface distance when acceptably classified into the other category. Table 2 compares the HD and ASSD for the three methods.

3.3.3. Target registration results

In image registration, the target registration error is often used as a metric for evaluating the results. Locating the different parts of the vertebra can be considered a basic requirement for spinal surgery planning. In this paper, the center of each vertebral category was treated as the target point. Based on the manually labelled surface mesh, the average coordinates of the vertices was calculated for each category as the target GTs. In the predicted mesh, the target point was also obtained based on the predicted multi-class results. Figure 4 illustrates the target registration error for the proposed and baseline methods.

Table 2: Multi-classes segmentation results

Models	**SP**		**LL**		**RL**	
	HD	**ASSD**	**HD**	**ASSD**	**HD**	**ASSD**
Ours	**0.97±1.29**	**0.39±0.45**	**0.84±0.55**	**0.32±0.12**	**0.78±0.56**	**0.31±0.16**
DeepReg	2.46±1.29	0.96±0.52	2.58±0.90	1.07±0.39	2.73±1.09	1.09±0.47
NiftyReg	4.98±2.92	1.98±1.58	5.67±4.24	2.27±2.20	5.59±3.69	2.02±1.50
	LAP		**RAP**		**LTP**	
	HD	**ASSD**	**HD**	**ASSD**	**HD**	**ASSD**
Ours	**0.95±0.67**	**0.35±0.13**	**1.37±2.61**	**0.41±0.28**	**1.44±2.27**	**0.54±0.73**
DeepReg	2.54±0.99	1.04±0.40	3.12±2.66	1.15±0.64	3.65±2.48	1.65±1.12
NiftyReg	4.68±2.84	1.97±1.46	5.11±3.30	2.09±1.41	8.46±4.51	5.02±3.53
	RTP		**LP**		**RP**	
	HD	**ASSD**	**HD**	**ASSD**	**HD**	**ASSD**
Ours	**1.06±1.23**	**0.43±0.50**	**1.52±0.82**	**0.49±0.24**	**1.62±0.97**	**0.57±0.45**
DeepReg	4.04±3.98	2.20±3.00	2.57±1.36	1.07±0.58	2.77±1.42	1.25±0.88
NiftyReg	6.78±4.78	4.05±4.05	3.81±3.36	1.68±1.87	3.69±2.96	1.63±1.58
	VB					
	HD	**ASSD**				
Ours	**0.96±0.40**	**0.45±0.10**				
DeepReg	3.32±2.34	1.22±0.68				
NiftyReg	8.88±5.98	3.59±2.66				

Data are shown in mm. Those data in bold indicates that the result is significantly different from others in the T test (p value < 0.05).

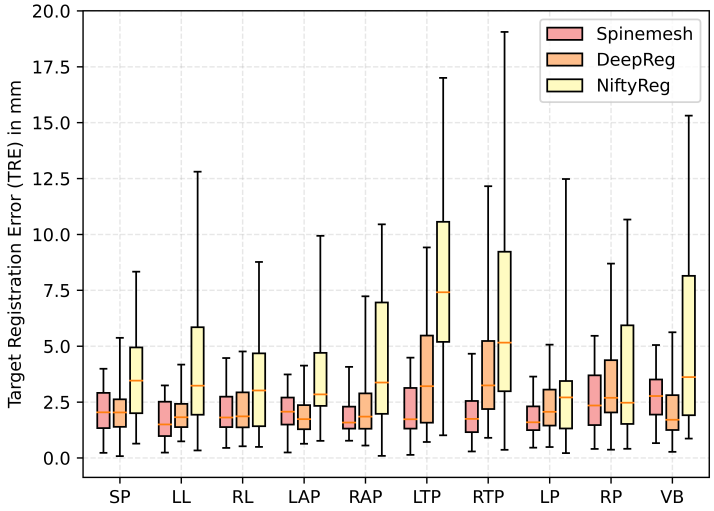

Figure 4: Example of multi-classes segmentation results
Four registration results in the validation are shown. The first row is the manually labeled classification, and the second row is the result predicted by the proposed network.

3.4. Disccusion

In the planning of robot-assisted orthopedic surgery, the bone surface information contained in medical images is often one of the key pieces of information. Surface shape is more worthy of attention when registering these bone images. The proposed network uses a mesh to describe the bone surface and predicts the vertices movement between the template surface and the target image to achieve mesh registration. This model only focuses on the surface registration, rather than outputting the movement of each voxel. This may be one reason why the proposed network performs better on these tasks.

Acknowledgments

Acknowledgments withheld.

References

Fabian Bongratz, Anne-Marie Rickmann, Sebastian Pölsterl, and Christian Wachinger. Vox2cortex: Fast explicit reconstruction of cortical surfaces from 3d mri scans with geometric deep neural networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 20773–20783, 2022.

Juliet E Compston, Michael R McClung, and William D Leslie. Osteoporosis, 2019.

Kexue Fu, Shaolei Liu, Xiaoyuan Luo, and Manning Wang. Robust point cloud registration framework based on deep graph matching. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 8893–8902, June 2021.

Zoher Ghogawala, James Dziura, William E Butler, Feng Dai, Norma Terrin, Subu N Magge, Jean-Valery CE Coumans, J Fred Harrington, Sepideh Amin-Hanjani, J Sanford Schwartz, et al. Laminectomy plus fusion versus laminectomy alone for lumbar spondylolisthesis. *New England Journal of Medicine*, 374(15):1424–1434, 2016.

Kunal Gupta. *Neural mesh flow: 3d manifold mesh generation via diffeomorphic flows*. University of California, San Diego, 2020.

Lasse Hansen and Mattias P. Heinrich. Graphregnet: Deep graph regularisation networks on sparse keypoints for dense registration of 3d lung cts. *IEEE Transactions on Medical Imaging*, 40(9):2246–2257, 2021. doi: 10.1109/TMI.2021.3073986.

Yipeng Hu, Marc Modat, Eli Gibson, Nooshin Ghavami, Ester Bonmati, Caroline M Moore, Mark Emberton, J Alison Noble, Dean C Barratt, and Tom Vercauteren. Label-driven weakly-supervised learning for multimodal deformable image registration. In *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, pages 1070–1074. IEEE, 2018a.

Yipeng Hu, Marc Modat, Eli Gibson, Wenqi Li, Nooshin Ghavami, Ester Bonmati, Guotai Wang, Steven Bandula, Caroline M Moore, Mark Emberton, et al. Weakly-supervised convolutional neural networks for multimodal image registration. *Medical image analysis*, 49:1–13, 2018b.

Fanwei Kong and Shawn C. Shadden. Whole heart mesh generation for image-based computational simulations by learning free-from deformations, 2021. URL https://arxiv.org/abs/2107.10839.

Fanwei Kong, Nathan Wilson, and Shawn Shadden. A deep-learning approach for direct whole-heart mesh reconstruction. *Medical image analysis*, 74:102222, 2021.

Julian Krebs, Tommaso Mansi, Hervé Delingette, Li Zhang, Florin C Ghesu, Shun Miao, Andreas K Maier, Nicholas Ayache, Rui Liao, and Ali Kamen. Robust non-rigid registration through agent-based action learning. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 344–352. Springer, 2017.

Qian Li, Zhijiang Du, and Hongjian Yu. Grinding trajectory generator in robot-assisted laminectomy surgery. *International Journal of Computer Assisted Radiology and Surgery*, 16(3):485–494, 2021.

Daniel Rueckert, Luke I Sonoda, Carmel Hayes, Derek LG Hill, Martin O Leach, and David J Hawkes. Nonrigid registration using free-form deformations: application to breast mr images. *IEEE transactions on medical imaging*, 18(8):712–721, 1999.

Hessam Sokooti, Bob de Vos, Floris Berendsen, Boudewijn PF Lelieveldt, Ivana Išgum, and Marius Staring. Nonrigid image registration using multi-scale 3d convolutional neural

networks. In *International conference on medical image computing and computer-assisted intervention*, pages 232–239. Springer, 2017.

Bob D de Vos, Floris F Berendsen, Max A Viergever, Marius Staring, and Ivana Išgum. End-to-end unsupervised deformable image registration with a convolutional neural network. In *Deep learning in medical image analysis and multimodal learning for clinical decision support*, pages 204–212. Springer, 2017.

Nanyang Wang, Yinda Zhang, Zhuwen Li, Yanwei Fu, Hang Yu, Wei Liu, Xiangyang Xue, and Yu-Gang Jiang. Pixel2mesh: 3d mesh model generation via image guided deformation. *IEEE transactions on pattern analysis and machine intelligence*, 43(10):3600–3613, 2020.

Udaranga Wickramasinghe, Edoardo Remelli, Graham Knott, and Pascal Fua. Voxel2mesh: 3d mesh model generation from volumetric data. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 299–308. Springer, 2020.