2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

Shape Reconstruction for Abdominal Organs based on a Graph Convolutional Network

Zijie Wang[1], *Student Member, IEEE*, Megumi Nakao[1], *Member, IEEE*,
Mitsuhiro Nakamura[2], and Tetsuya Matsuda[1], *Member, IEEE*

***Abstract*— Computed tomography and magnetic resonance imaging produce high-resolution images; however, during surgery or radiotherapy, only low-resolution cone-beam CT and low-dimensional X-ray images can be obtained. Furthermore, because the duodenum and stomach are filled with air, even in high-resolution CT images, it is hard to accurately segment their contours. In this paper, we propose a method that is based on a graph convolutional network (GCN) to reconstruct organs that are hard to detect in medical images. The method uses surrounding detectable-organ features to determine the shape and location of the target organ and learns mesh deformation parameters, which are applied to a target organ template. The role of the template is to establish an initial topological structure for the target organ. We conducted experiments with both single and multiple organ meshes to verify the performance of our proposed method.**

I. Introduction

Three-dimensional (3D) medical imaging technologies, such as computed tomography (CT) and magnetic resonance imaging (MRI), play a role in visual observation of human structure for diagnosis, surgery, and radiotherapy. However, patients may undergo multiple treatments over the course of several weeks, during which time changes to their abdominal anatomy may occur. Because CT equipment is not portable, and repeated imaging increases patient radiation exposure, CT images are only obtained for initial planning [1]. Daily images can only be obtained using low-dimensional X-ray or low-resolution cone-beam CT (CBCT), but some organs, such as the stomach, duodenum, and pancreas, cannot be clearly resolved because missing values or artifacts often appear in CBCT images [2]. Furthermore, because the stomach and duodenum are filled with air, images of these organs are not very clear, even in CT. Organ shape and location vary from person to person [3], further complicating organ detection. In order to obtain a clear image of internal organs from the CBCT image, an experienced doctor needs to spend a lot of time to manually define. If this work can be performed by machines, it could greatly reduce the workload of doctors. Some researchers have proposed methods to reconstruct an organ's 3D shape from a single image [4] [5] [6] For example, X-ray2Shape reconstructs the 3D shape of an organ from a single-view two-dimensional (2D) X-ray [5]; however, in 2D X-rays, it is not easy to segment the stomach and the duodenum. Nakao et al. [2] suggest reconstructing organs that have poorly defined outlines using the contours of surrounding organs that are easily segmented. The surrounding organs of the stomach and duodenum are shown in Fig. 1. The outlines of the liver and kidneys are clearly visible in CT images. Though Nakao et al. proposed a method using multiorgan features to localize pancreatic cancer, it was a kernel-based method that mainly focused on the intrapatient deformations caused by respiratory motion. Organ shape reconstruction from interpatient data remains challenging. The purpose of our study was to develop a method based on a graph convolutional network (GCN) [7] to reconstruct the 3D shape of an undetectable organ from a patient's more easily detectable organs. GCN has mainly been used for classification tasks in non-Euclidean space [8]. Herein, we extend the application of GCN to the processing of mesh data to calculate the deformation from a target organ template, derived from a 3D-CT dataset, to the unique shape of the organ using the patient's 3D-CBCT data of surrounding easily detectable organs.

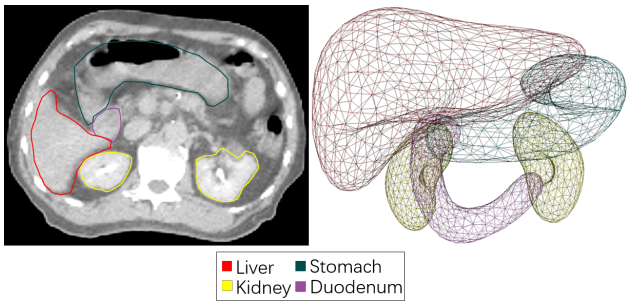


Fig. 1. The relative positions of the stomach, duodenum, liver, and kidneys: CT slice image with blurred outline of stomach and duodenum and original mesh which was generated on the basis of CT images

II. Methods

A. Dataset and Processing

The dataset comprised 3D-CT images from 124 patients and 10-phase 4D-CT from 35 patients who underwent intensity-modulated radiotherapy for pancreatic cancer at Kyoto University Hospital [3]. This study was performed in accordance with the Declaration of Helsinki. In each image, the 3D contours of the stomach, duodenum, and pancreas had been manually defined by board-certified radiation oncologists. The original mesh data were generated on the basis of these 3D contours. Deformable mesh registration [2] was applied to obtain meshes with point-to-point correspondence.

[1]Z. Wang, M. Nakao and T. Matsuda are with the Graduate School of Informatics, Kyoto University, Kyoto, Japan. (e-mail: wangzijie@sys.i.kyoto-u.ac.jp)
[2]M. Nakamura is with Human Health Sciences, Graduate School of Medicine, Kyoto University, Kyoto, Japan.

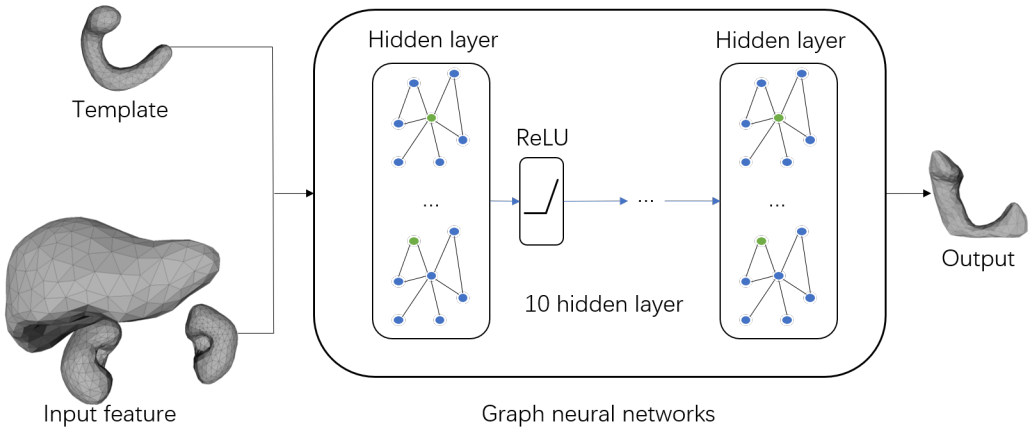


Fig. 2. Framework architecture. The organ template is derived from the mean shape (from a multiple patient dataset) of the duodenum, and the input is a set of features of the patient's surrounding organs (liver, left kidney, and right kidney). The GCN transforms the template, and the output is the patient's duodenum.

The liver model comprised 500 vertices and 996 faces, the right kidney and duodenum model comprised 400 vertices and 796 faces. Using the same number of vertices for each patient's data in the dataset made it possible to calculate a mean target organ shape. We randomly selected data from 100 patients for the interpatient training dataset and used the remaining data (from 24 patients) for the interpatient test dataset. For the intrapatient experiments, we randomly selected 300 images for the training dataset, and the remaining 50 images were used as the test dataset.

B. Duodenum-targeted GCN Framework

The template is the mean shape of target organ. The framework (Fig. 2, with the duodenum as the target organ) uses the features of organs that surround the target organ that are easily detected (liver, right kidney, and left kidney) as input. The network learns how to deform the mean shape of duodenum according to the unique features of the patient's liver, right kidney, and left kidney outlines. The output is a reconstruction of the patient's target organ. In this section, to facilitate description of the network, we use the example in which the duodenum is the target organ and the liver is the input. For the template, we calculated the mean shape of the duodenum from the dataset. The GCN module uses deep learning to obtain features in non-Euclidean space. Our mesh model is a graph $G(\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ is the set of nodes and $\mathcal{E}$ is the set of edges. The mesh is deformed by updating the coordinates of vertices. In our experiments, the GCN network contained 10 graph convolutional layers, which can be defined as

$$X^{(l+1)} = \sigma(\hat{D}^{-\frac{1}{2}}(A+I)\hat{D}^{-\frac{1}{2}}X^{(l)}W^{(l)}), \tag{1}$$

where $X^{(l)}$ is the feature representing the vertices before convolution, $X^{(l+1)}$ is the feature representing the vertices after convolution, $A \in \mathbb{R}^{n\times n}$ is the adjacency matrix, $I \in \mathbb{R}^{n\times n}$ is an identity matrix, $\hat{D} \in \mathbb{R}^{n\times n}$ is the degree matrix of $A+I$, n is the total number of vertices, W is the trainable weight, and σ is the activation function. To combine the duodenum template and liver input feature, we preprocess the input data as

$$X_1 = \begin{pmatrix} v_1 & u_1 & \dots & u_{500} \\ \vdots & \vdots & \ddots & \vdots \\ v_{400} & u_1 & \dots & u_{500} \end{pmatrix}, \tag{2}$$

where v_1 is the first coordinate of the duodenum template, v_{400} is the 400th coordinate of the duodenum template, u_1 is the first coordinate of the liver input feature, u_{500} is the 500th coordinate of the liver input feature, and 400 and 500 represent the number of vertices of the duodenum template and liver input, respectively. When the input feature uses both the liver and right kidney, we preprocess the input data as

$$X_2 = \begin{pmatrix} & w_1 & \dots & w_{400} \\ X_1 & \vdots & \ddots & \vdots \\ & w_1 & \dots & w_{400} \end{pmatrix}, \tag{3}$$

where X_1 is the input matrix from (2), w_1 is the first coordinate of the right kidney input, w_{400} is the 400th coordinate of input right kidney, and 400 represents the number of vertices of the input.

C. Loss Function

To generate accurate prediction results, we used three loss functions: positional loss, discrete Laplacian loss [2], and translation loss. Positional loss is used to minimize the distance between corresponding predicted shape and ground truth shape vertices and is defined as

$$\mathcal{L}_{pos} = \frac{1}{n}\sum_{i=0}^{n} \|v_i - \hat{v}_i\|_2^2, \tag{4}$$

where n is the number of vertices, v_i represents the ground truth vertices, and $\hat{v}_i$ represents the predicted shape vertices. Equation (4) defines the function that moves the vertices of the template (generalized mean shape of the duodenum shape) to the corresponding positions of the ground truth shape (unique characteristics of the patient's duodenum

based on those of the surrounding organs). The function causes the overall shape to stretch, translate, rotate, and twist, which results in the model shape having a rough surface.

To maintain a smooth surface, we used a discrete Laplacian loss function. The discrete Laplacian is defined below for vertex v_i,

$$L(v_i) = \frac{1}{N(v_i)} \sum_{j \in N(v_i)} (v_i - v_j) \tag{5}$$

where $N(v_i)$ is the number of adjacent vertices of one ring connected by vertex v_i and the edge, and v_j is the neighboring vertex of v_i; therefore, the discrete Laplacian loss function is

$$\mathcal{L}_{laplacian} = \frac{1}{n} \sum_{i=0}^{n} \|L(v_i) - L(\hat{v}_i)\|_2^2, \tag{6}$$

where $L(v_i)$ is the discrete Laplacian, as defined in (5), of the template, and $L(\hat{v})$ is the discrete Laplacian of ground truth. The Laplacian loss function restricts the deformation permitted by the positional loss function. To control the predicted location of the organ (i.e., where the duodenum will be located with respect to the surrounding organs such as the liver and right kidney), we used a translation loss function based on the center of gravity of the mesh structure:

$$\mathcal{L}_{trans} = \|G(\mathcal{V}) - G(\hat{\mathcal{V}})\|_2^2 \tag{7}$$

where $G(\mathcal{V}) = \sum_{i=1}^{n} v_i/n$ is the center of gravity before deformation and $G(\hat{v}_i)$ is the center of gravity after deformation. The total loss is the weighted sum of the three loss functions

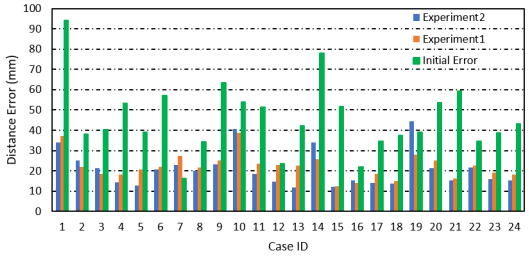
$$\mathcal{L}_{total} = \mathcal{L}_{pos} + \lambda_{laplacian}\mathcal{L}_{laplacian} + \lambda_{trans}\mathcal{L}_{trans} \tag{8}$$

III. Experiments

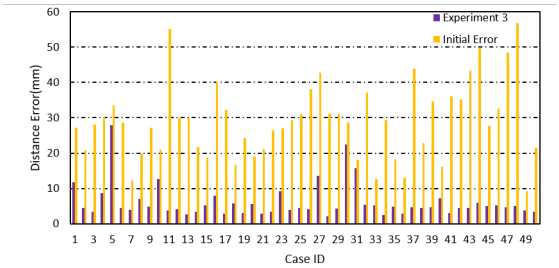
A. Performance Evaluation

We conducted three experiments. Experiments 1 and 2 were interpatient experiments, and Experiment 3 was an intrapatient experiment. Experiments 1 and 3 used only the liver as input, whereas Experiment 2 used both the liver and right kidney as input. All experiments used the duodenum as the target organ. The framework was implemented in TensorFlow-GPU, with a single NVIDIA Geforce RTX 2070 GPU. The network was trained using an Adam optimizer with 2×10^{-4} learning rate and 5×10^{-6} weight decay. Given that there was a limited number of samples, the batch size was set to 1. The total loss converged at approximately 4500 epochs; therefore, we set the number of training epochs to 5000. Training in Experiments 1, 2, and 3 lasted 14 hours, 32 hours, and 61 hours, respectively. We set the hyperparameters $\lambda_{laplacian} = 1$ and $\lambda_{trans} = 1$ to balance the three loss functions, and we used Euclidean distance

$$\mathcal{D}_{Euclidean} = \frac{1}{n} \sum_{i=0}^{n} \sqrt{(v_i - \hat{v}_i)^2}, \tag{9}$$



(a)



(b)

Fig. 3. Euclidean distance error. (a) interpatient test set: Experiment 1 final, Experiment 2 final, and initial template; (b) intrapatient image test set: Experiment 3 final and initial template.

TABLE I
PERFORMANCE RESULTS FOR EXPERIMENTS 1 AND 2

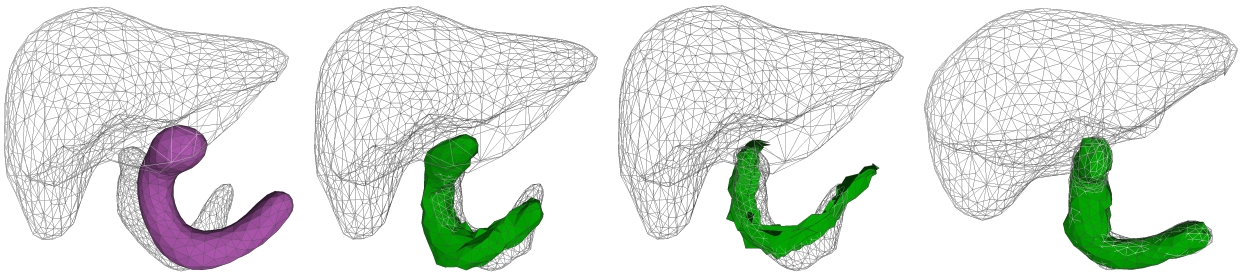
	Initial	Experiment 1	Experiment 2
ED[1](mean) [mm]	45.86	22.29	20.98
ED[1](min) [mm]	16.35	12.55	11.93
MD[2](mean) [mm]	23.40	9.35	8.52
MD[2](min) [mm]	8.29	5.81	5.15

[1] Euclidean distance (ED)
[2] Mean distance (MD)

where n is the number of vertices, v_i represents ground truth vertices, and $\hat{v}_i$ represents predicted shape vertices and mean distance $\mathcal{D}_{Mean}$, which is the mean value of the nearest bidirectional point-to-surface distance [9] to evaluate the difference between ground truth and predicted shapes. We also compared the difference between the ground truth and predicted shape to the difference between the ground truth and template shape to ascertain how much deformation the network had processed. Because distance measures do not reflect the quality of surface smoothness, we visually inspected the output.

B. Results

Euclidean distance error are shown for the interpatient test set (Fig. 3(a)) and the intrapatient test set (Fig. 3(b)). The blue bars show Euclidean distance error between ground truth and estimated shape in Experiment 2, with the mean value being 20.98mm, the orange bars show the error between ground truth and estimated shape in Experiment 1, with the mean value being 22.29mm and the green bars show the error between ground truth and the initial template in Experiment 1, 2, with the mean value being 45.86mm.



(a) (b) (c) (d)

Fig. 4. Visualization result of duodenum reconstruction. The meshes without faces are ground truth, the mesh with faces is estimated shape: (a) the purple shape is initial template, (b) the green shape is $\lambda_{laplacian} = 1$, estimated shape with a smooth surface, (c) the green shape is $\lambda_{laplacian} = 0$, estimated shape with a wrinkled surface, (d) the green shape is the estimated intrapatient shape.

TABLE II
PERFORMANCE RESULTS FOR EXPERIMENT 3

	Initial	Experiment 3
ED[1](mean) [mm]	28.98	6.13
ED[1](min) [mm]	9.22	2.11
MD[2](mean) [mm]	13.13	1.94
MD[2](min)[mm]	3.75	0.83

[1] Euclidean distance (ED)
[2] Mean distance (MD)

The purple bars show the Euclidean distance error between ground truth and estimated shape in Experiment 3,with the mean value being 6.13mm and the yellow bars show the error between ground truth and the initial template in Experiment 3, with the mean value being 28.98mm. The average and minimum values of Euclidean distance and mean distance for experiments are shown in Table I and Table II. In Experiments 1 and 2, compared with the initial error, all evaluation values are smaller. Based on this, we conclude that it is feasible for the framework to learn mesh data deformation and predict the shape of the duodenum from data about the surrounding organs. The results of experiment 2 were better than those of Experiment 1. The resulting algorithm from Experiment 3 will produce more accurate results if using intrapatient organ deformation data. The visual result is shown in Fig. 4. Fig. 4(a) shows the difference between template and one ground truth. To clearly show the position of the estimated result compared with that of the ground truth, we added the ground truth of the liver. In experiments 1 and 2, the estimated shape is smooth (Fig. 4(b)). To show how the wrinkled surface would appear, we show an image for which $\lambda_{laplacian} = 0$ (Fig. 4(c)). Fig. 4(d) shows the final result of experiment 3.

IV. CONCLUSION

This method estimates the shape and position of difficult or undetectable organs from easily detectable organs using a GCN. To obtain an accurate position and a smooth shape, the network was trained with positional loss, discrete Laplacian loss, and translation loss functions. We used both interpatient data and intrapatient 3D-CT data. Compared with single-organ input, using multi-organ input required more time but resulted in better performance. In future work, we will improve the framework and reduce prediction error to a clinically acceptable range.

ACKNOWLEDGMENT

This research was supported by the Japan Society for the Promotion of Science Grant-in-Aid for Scientific Research (B) (grant number 18H02766) and for challenging Exploratory Research (grant number 18K19918). We thank Coren Walters-Stewart, PhD, and Edanz Group (https://en-author-services.edanz.com/ac) for editing a draft of this manuscript.

REFERENCES

[1] S. Miao, Z. J. Wang, and R. Liao, "A CNN Regression Approach for Real-Time 2D/3D Registration", *IEEE Trans Med Imaging*, vol. 35, no. 5, pp. 1352-1363, 2016.
Phys. Med, vol.59, pp. 13-21, 2019.
[2] M. Nakao, M. Nakamura, T. Mizowaki, and T. Matsuda, "Statistical deformation reconstruction using multi-organ shape features for pancreatic cancer localization", *Medical image analysis*, vol. 67, p. 101829, 2021.
[3] M. Nakamura, M. Nakao, N. Mukumoto, R. Ashida, H. Hirashima, M. Yoshimura et al., "Statistical shape model-based planning organ-at-risk volume: application to pancreatic cancer patients", *Phys Med Biol*, vol. 66, 014001, 2021.
[4] S. Wu, M. Nakao, J. Tokuno, T. Chen-Yoshikawa, and T. Matsuda, "Reconstructing 3D lung shape from a single 2D image during the deaeration deformation process using model-based data augmentation", *IEEE Int. Conf. on Biomedical and Health Informatics (BHI)*, pp. 1-4, 2019.
[5] F. Tong, M. Nakao, S. Wu, M. Nakamura and T. Matsuda, "X-ray2Shape: Reconstruction of 3D Liver Shape from a Single 2D Projection Image", *IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 1608-1611, 2020.
[6] Y. Wang, Z. Zhong, and J. Hua, "DeepOrganNet: On-the-Fly Reconstruction and Visualization of 3D / 4D Lung Models from Single-View Projections by Deep Deformation Network", *IEEE Transactions on Visualization and Computer Graphics*, vol. 26, no. 1, pp. 960-970, 2020.
[7] T. N. Kipf and M. Welling, "Semi-Supervised Classification with Graph Convolutional Networks", *Proceedings of the 5th International Conference on Learning Representations (ICLR)*, 2017.
[8] Jie Zhou, Ganqu Cui, Zhengyan Zhang, Cheng Yang et al., "Graph Neural Networks: A Review of Methods and Applications", *arXiv preprint*, Art. no. arXiv:1812.08434, 2019.
[9] J. Kim, C. Valdes-Hernandez Mdel, N. A. Royle, and J. Park, "Hippocampal shape modeling based on a progressive template surface deformation and its verification", *IEEE Trans. Med. Imag.*,vol. 34, pp. 1242-1261, 2015.