

Shape Reconstruction for Undetectable Regions of Abdominal Organs Based on a Graph Convolutional Network

Zijie Wang^a (wangzijie@sys.i.kyoto-u.ac.jp), Megumi Nakao^a
(megumi@i.kyoto-u.ac.jp), Mitsuhiro Nakamura^b
(m_nkmr@kuhp.kyoto-u.ac.jp), Tetsuya Matsuda^a (tetsu@i.kyoto-u.ac.jp)

^a Graduate School of Informatics Kyoto University, Yoshida-honmachi, Sakyo-ku,
Kyoto-shi, Kyoto, JAPAN

^b Graduate School of Medicine Kyoto University, 53, Shogoin Kwaracho, Sakyo-ku,
Kyoto, JAPAN

Corresponding Author:

Zijie Wang

Graduate School of Informatics Kyoto University, Yoshida-honmachi, Sakyo-ku,
Kyoto-shi, Kyoto, JAPAN

Tel: 090-5496-7733

Email: wangzijie@sys.i.kyoto-u.ac.jp

Shape Reconstruction for Undetectable Regions of Abdominal Organs Based on a Graph Convolutional Network

Zijie Wang^a, Megumi Nakao^a, Mitsuhiro Nakamura^{b,*}, Tetsuya Matsuda^a

^a*Graduate School of Informatics Kyoto University, Yoshida-honmachi, Sakyo-ku, Kyoto-shi, Kyoto, JAPAN*

^b*Graduate School of Medicine Kyoto University, 53, Shogoin Kawaracho, Sakyo-ku, Kyoto, JAPAN*

Abstract

Although computed tomography (CT) and magnetic resonance imaging (MRI) produce high-resolution images, during surgery or radiotherapy, only low-resolution cone-beam CT and low-dimensional X-ray images are usually obtained. Furthermore, because the duodenum and stomach are filled with air, it may be hard to accurately segment their contours, even on high-resolution images. Although denoising and image enhancement can be used to assist in the segmentation and reconstruction of 3D images, some regions of organs may remain difficult to detect. To approach this issue, we propose a graph convolutional network (GCN) to reconstruct organs that are hard to detect on medical images. The method uses a framework to transform an initial template into an estimated shape of the patient's organs based on previous information attained from different detectable organs from different patients. To overcome the inaccurate estimation results of triangular surface mesh data with the GCN, we propose a method to add internal vertices and edges, which has the effect of improving the calculation accuracy by enhancing the topological structure of the mesh data. In this study, the organ data of 124 patients who underwent pancreatic tumor treat-

*Corresponding author.

Email addresses: wangzijie@sys.i.kyoto-u.ac.jp (Zijie Wang), megumi@i.kyoto-u.ac.jp (Megumi Nakao), m_nkmr@kuhp.kyoto-u.ac.jp (Mitsuhiro Nakamura), tetsu@i.kyoto-u.ac.jp (Tetsuya Matsuda)

ment were used to verify the predictive effect of the proposed method. The 3D organ contours were defined by board-certified radiation oncologists and used as the ground truth surface meshes. For 100 randomly selected detectable feature vertices, the average Euclidean distance error of the liver was 3.72 mm. The topological enhancement proposed in this study reduced the prediction error of the liver by an average of 0.84 mm.

Keywords: abdominal organs, mesh data, GCN, shape reconstruction

1. Introduction

Three-dimensional (3D) medical imaging technologies such as computed tomography (CT) and magnetic resonance imaging (MRI) play a major role in the visualization of human structures for diagnosis, surgery, and radiotherapy. However, for radiotherapy, patients may undergo multiple treatments over the course of several weeks, and changes in their abdominal anatomy may occur over such periods. Because CT equipment is not portable and repeated CT imaging increases the patient’s radiation exposure (Islam et al., 2006), CT images are generally only obtained for the initial planning (Miao et al., 2016; Kan et al., 2008), with the daily images being obtained using low-dimensional X-ray or low-resolution cone-beam CT (CBCT). However, some organs such as the stomach, duodenum, and pancreas cannot be clearly resolved because of missing sections or artifacts that often appear in CBCT images (Nakao et al., 2021). Furthermore, because the stomach and duodenum are filled with air, images of these organs may not be very clear, even on CT. The shape and location of these organs also vary from person to person (Nakamura et al., 2021) (Rigaud et al., 2018), further complicating the process of organ detection. To obtain a precise contour of such internal organs from the CBCT image, an experienced doctor needs to spend considerable time on manual contouring. If this work could be assisted by computers, it would greatly reduce the workload. Occasionally, certain regions of the target organ are detected but the entire region is difficult to segment. Therefore, we consider it meaningful to estimate the shape

of the undetected region from the detectable one. Statistical shape modeling based on prior knowledge has been widely used to obtain accurate organ shapes from medical images (Nakamura et al., 2019; Jia, 2019). Unlike physics-based modeling (Fuerst et al., 2014; Nakao & Minato, 2010), statistical modeling is a data-driven approach that does not explicitly describe the elasticity and physical conditions of organs. Statistical modeling methods such as principal component analysis (PCA) and other PCA-based modeling methods have been proposed for parameterizing organs and analyzing their movements and changes in shape (Rigaud et al., 2018; Cootes & Taylor, 2001; Heitz et al., 2005). However, PCA is a linear model, which results in considerable errors in shape reconstruction (Nakao et al., 2021). There are also shape reconstruction methods based on partially observed data that do not suffer from inter-patient limitations. Some researchers have proposed methods to reconstruct an organ’s 3D shape from a single image (Tong et al., 2020; Wang et al., 2019; Wu et al., 2019). However, it is not easy to segment the stomach and the duodenum on 2D X-rays, and the methods have limitations for organs that cannot be completely detected on 2D images. Although kernel-based statistical modeling (Nakao et al., 2021; Yamamoto et al., 2021) has been proposed for estimating the shape of flexible organs that are easily deformed, if there is no method to create associations between the detectable and undetectable regions it is difficult to complete the prediction of the undetectable ones. Therefore, it remains a challenge to estimate the shape of the remaining part from the limited organ data provided. This study proposes a framework combining a Graph Convolutional Network (GCN) with a topological mesh update to calculate the undetectable shape of organs based on partially-detectable features. We use surface meshes to represent the 3D shapes of abdominal organs. Compared with other 3D representations such as point clouds, surface meshes define 3D geometry and topological structures that are important for calculating deformations (Holzinger, 2014). The GCN is a deep learning framework for graphs with nodes and edges (Kipf & Welling, 2016), and has mainly been used for classification tasks in non-Euclidean space (Zhou et al., 2020), although there are also studies applying GCNs to mesh-

based shape reconstruction (Wang et al., 2020) and surface mesh segmentation (Tang & Qiu, 2021). Herein, we extend the application of a GCN to the processing of mesh data to calculate the deformation from a template mesh to the patient-specific shape of an organ according to its detectable part. In the GCN, the information transfer between points is particularly critical. We overcome the problem of weak correlation information between distant vertices by adding internal vertices. We incorporate a k-means clustering algorithm into the GCN to enhance the connection between distant vertices to further reduce the Euclidean distance error of the shape prediction.

In this study, the 3D-CT images of 124 patients with pancreatic cancer who underwent radiotherapy at Kyoto University Hospital were used to verify the accuracy of the proposed method.

2. Related Work

There are three main aspects to the work. The first subsection covers the application of the GCN to 3D reconstruction. The second part is related to research on the 3D reconstruction of organs. The last subsection explains the contribution of this research.

2.1. Graph convolutional network for 3D reconstruction

The GCN is a neural network used mainly in non-Euclidean space and for classification tasks (Yao et al., 2019; Zhang et al., 2019). Because 3D mesh data are also a kind of graph, there has been recent research on applying GCNs to 3D reconstruction. One such application of GCNs involves the reconstruction of 3D shapes from 3D point clouds. Luo et al. proposed a method that took the point cloud and triangular section as input and decided whether the tetrahedron was inside or outside the output to create the surface of an object (Luo et al., 2021). However, a premise of such a method is to have information on all the vertices, whereas, for undetectable organs, the corresponding vertex information is unknown. Another application of GCNs to 3D reconstruction involves

reconstructing an object in three dimensions from the original 2D images. Ci et al. proposed a GCN-based method for 3D human pose estimation from a single image. Their method included two steps: the first was to estimate the 2D pose from the image, and the second was to estimate the corresponding 3D pose. (Ci et al., 2020). However, predicting an entire graph from a part of a 3D graph is still a challenge.

2.2. 3D reconstruction of organs

In this subsection, we give an overview of some related medical 3D reconstruction research. A large number of methods are used to generate two-dimensional medical images, including CT, X-ray, ultrasound, and MRI. Three-dimensional reconstruction is used to visualize the anatomical information obtained by these digital imaging technologies, and computer-generated three-dimensional visualization effects have high clinical value. There are already methods proposed to reconstruct abdominal organs, such as PCA-based and kernel-based statistical modeling. Rigaud et al. (Rigaud et al., 2018) generated a PCA model of the cervix and bladder to extract the dominant deformation modes. Heitz et al. (Heitz et al., 2005) also proposed a method to produce a statistical model using PCA, and created a model of the lower cervical vertebrae to verify the accuracy, compactness, and generalization ability of their method. However, PCA may not be a good solution for complex and non-linear inter-patient differences in organ shape. When large differences in shape are present, such as between patients with large individual differences or with soft organs with large deformation potential, PCA-based approaches are subject to limitations. Ehrhardt et al. (Ehrhardt et al., 2010) constructed a four-dimensional statistical model of the lung, which is a spatiotemporal model, using a statistical mean displacement model. Nakao et al. (Nakao & Minato, 2010; Nakao et al., 2021) and Nakamura et al. (Nakamura et al., 2021) extracted the mean shapes and displacements of organs from multiple patients' 3D images and built a statistical deformation modeling to estimate the movement of the pancreatic tumor with respiratory motion. There has also been research on 3D reconstruc-

tion from a single X-ray image. Tong et al. (Tong et al., 2020) proposed a framework combining a Convolutional Neural Network (CNN) and GCN to reconstruct a patient’s liver from a single X-ray image. However, some abdominal organs, such as the stomach and duodenum are unclear on X-ray images.

2.3. Contributions of this study

Medical 3D reconstruction is mainly based on tasks completed on multiple 2D medical images, but there are generally unclear parts of organs on various two-dimensional images. Therefore, it makes sense to estimate the undetectable part from the detectable part. However, limited data availability is an unavoidable problem in medical research; it is difficult for a traditional GCN to obtain accurate estimation results when dealing with triangular surface meshes with a small number of samples and simple topological structures. This study, therefore, proposes a method that can complete the estimation of the undetectable regions from the detectable ones. There are two main contributions of this paper.

- A proposed method to improve the accuracy of GCN shape estimation by enhancing the topology of mesh data.
- Analysis of shape reconstruction performance in the estimation of undetectable regions from detectable regions under different conditions and for different organs.

3. Methods

3.1. Problem setting

X-ray, ultrasound, CT, and MRI images are widely used in clinical medicine. However, no medical imaging modality can clearly detect all organs. For example, it is usually difficult to completely determine the contour of the stomach on CT images because of the presence of air, although the partial shape may be easily defined. Figure 1 (a) shows the contours of four abdominal organs

on CT images. It can be seen that part of the contour of the stomach is clear, but another part is difficult to distinguish. If the shape of the undetectable regions can be estimated from the detectable regions, as shown in Figure 1 (b), it will not be necessary to manually segment the organ contour, improving the efficiency of organ detection.

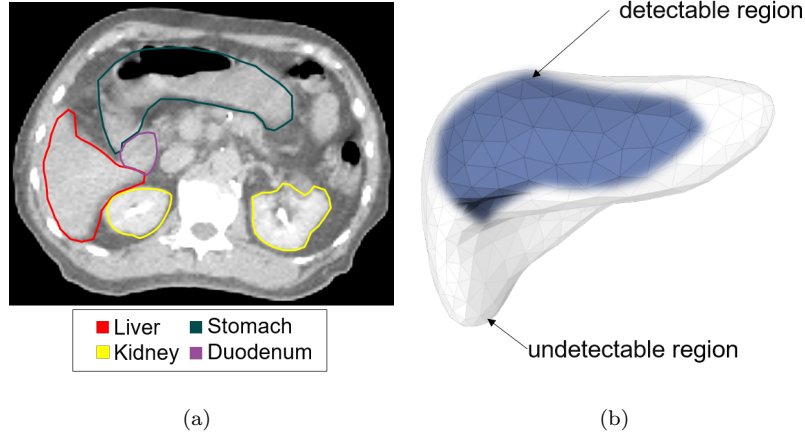


Figure 1: Shape reconstruction for undetectable regions of organs. (a) Abdominal CT slice with manually-defined organ contours. A Complete definition of the stomach contour is difficult because of air, (b) detectable and undetectable regions of an organ in our problem setting.

3.2. Dataset and preprocessing

The dataset consisted of 3D-CT images from 124 patients who underwent intensity-modulated radiotherapy for pancreatic cancer at Kyoto University Hospital. The clinical data were collected with the approval of the Kyoto University Graduate School and Faculty of Medicine Ethics Committee, and with the consent of the patient regarding the use of their anonymized data for academic research. In each image, the 3D contours of the liver, stomach, duodenum, and pancreas were manually defined by board-certified radiation oncologists, and the original tetrahedral mesh data were generated from these 3D contours. The original tetrahedral meshes differ in their number of vertices and

structure because they were independently generated from different CT images. One case was randomly selected to be the initial template mesh and its surface was resampled to a new mesh with the same number of vertices and triangles. Then, we used deformable mesh registration (DMR) (Nakao et al., 2021) to deform the template to the organ of an individual patient to obtain registered meshes with point-to-point correspondence. The DMR process is shown in Figure 2. In this way, each organ from the different patients had the same number of vertices and faces. The liver model comprised 500 vertices and 996 faces, and the stomach model of 400 vertices and 796 faces. Figure 3 shows the registered surfaces and templates of the liver and stomach. A standardized number of vertices and faces resulted in a similar topology, allowing the shape to be easily calculated within our framework.

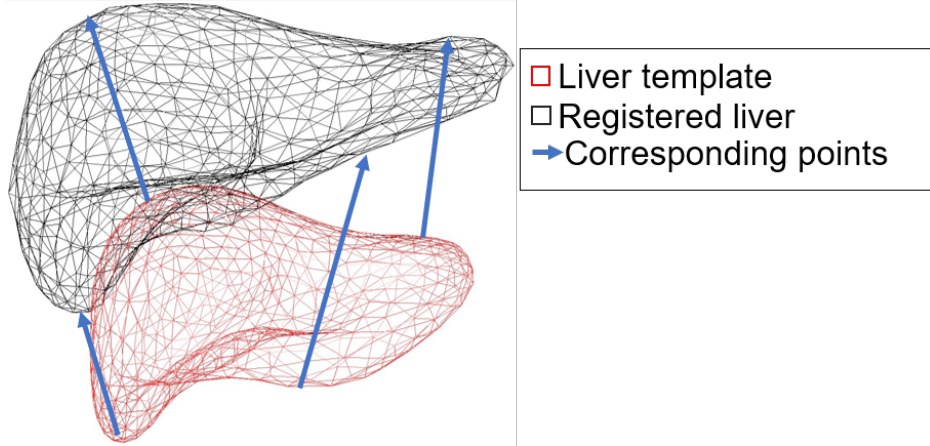


Figure 2: The DMR process. The liver template with 500 vertices and 966 faces is deformed through registration to the target shape with different topologies.

3.3. Proposed framework

The framework proposed in this study can reconstruct the organ shape of different patients (inter-patient) using a mean organ shape as the initial template. After deformable mesh registration (DMR) processing, the same organ

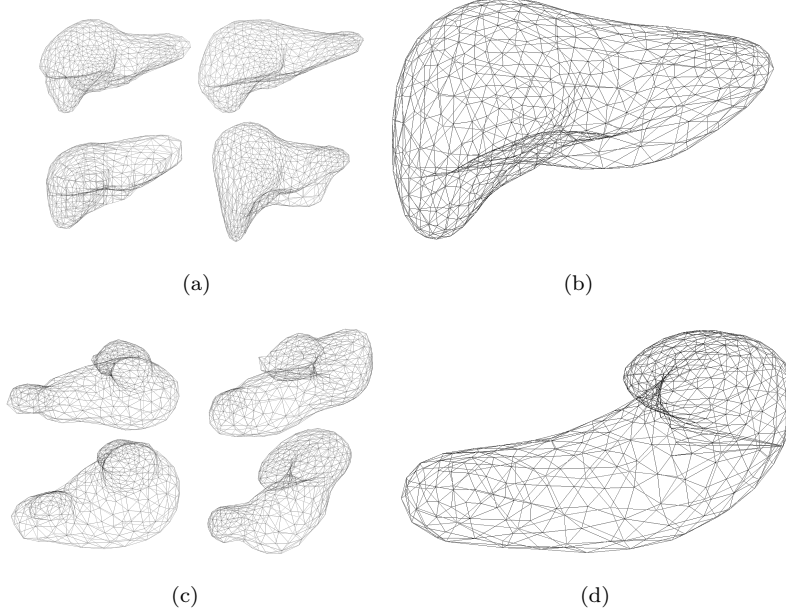


Figure 3: Registered surfaces and templates of abdominal organs, (a) patient-specific liver surfaces, (b) liver template, (c) patient-specific stomach surfaces, and (d) stomach template

in different patients is represented by the same number of vertices and faces. The standardized topology makes it possible to calculate the mean shape. We define the coordinates of the vertices of the template as:

$$\bar{v}_i = (\bar{v}_{xi}, \bar{v}_{yi}, \bar{v}_{zi}), i \in n. \quad (1)$$

where n is the number of vertices.

Figure 4 shows the architecture of the proposed framework. The liver is an abdominal organ with a large volume and a special shape, and it is therefore relatively easy to detect more features, so we set the liver as the first research object to analyze the influence of different factors on the prediction results. In Figure 4, the initial template is a mean shape of the liver and the input feature is the detectable region of an individual patient’s liver. After adding internal vertices automatically, the GCN deforms the mesh to the final output.

The output is the liver shape corresponding to the individual patient. Each component of the framework will be explained in detail.

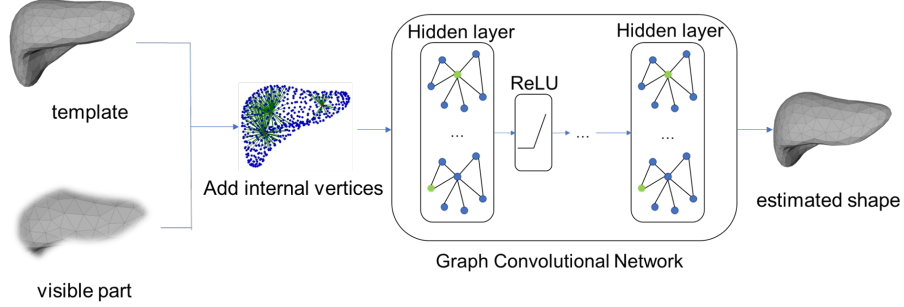


Figure 4: The architecture of the proposed framework.

3.3.1. Topological update for surface meshes

The information transfer of the GCN is delivered along the edge, which will be described in detail in 3.3.2. Therefore, there will be a problem if there are too many edges between the detectable vertices and the target vertices of the prediction, which will cause a substantial drop in the correlation between the two, greatly reducing the prediction capabilities. To approach this issue, our framework can add internal vertices and build new edges among the surface vertices and internal vertices. We use the k-means algorithm (MacQueen et al., 1967) to calculate cluster centers on the vertices of the template and these cluster centers become the added internal vertices. Then, new edges are built between the added internal vertices and surface vertices belonging to the same cluster. As new edges are generated, the adjacency matrix in the GCN is also updated. Figure 5 shows the addition of vertices and edges with different numbers of clusters.

3.3.2. GCN-based shape reconstruction

The detectable region of the proposed framework is defined by the coordinates of the vertices corresponding to the clinically detectable region, which

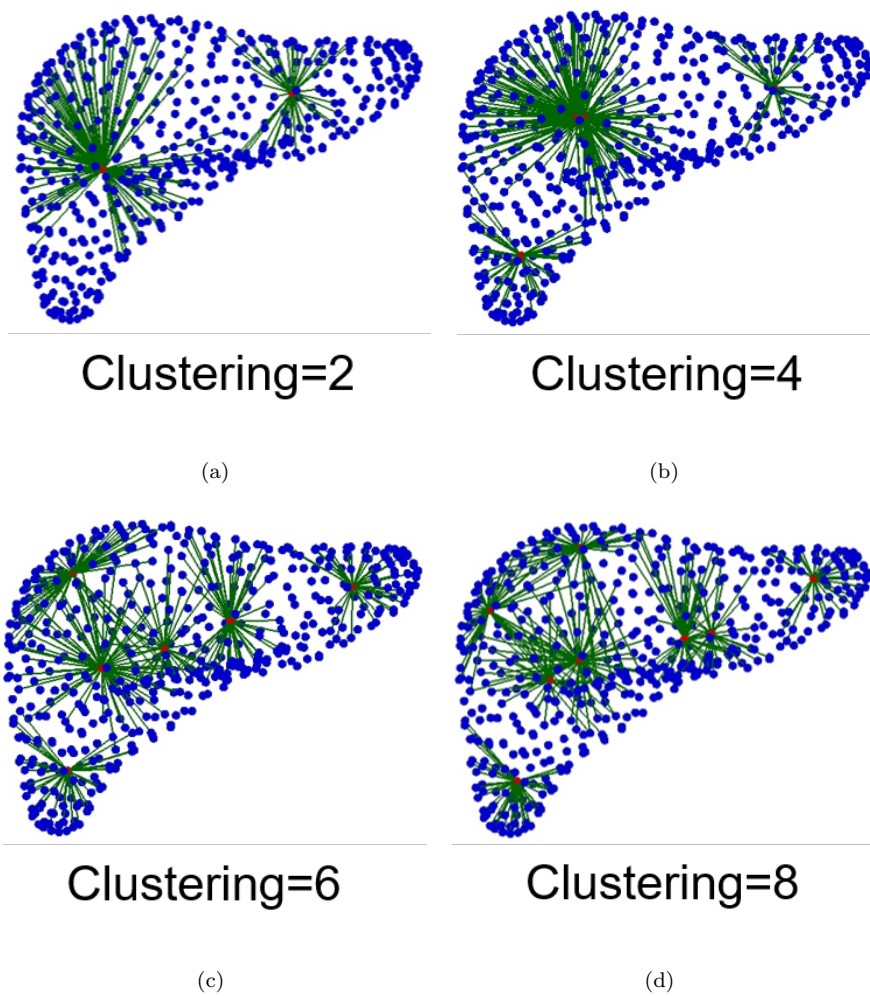


Figure 5: Added vertices and edges with different numbers of clusters. (a) clusters=2, (b) clusters=4, (c) clusters=6, (d) clusters=8.

are defined as

$$v_i = (v_{xi}, v_{yi}, v_{zi}), i \in n, \quad (2)$$

the coordinates of the remaining undetectable region are set as $(0, 0, 0)$. The final input fed into the framework for calculation is the relative displacement between the detectable region feature and the template, which is described as

$$X^{(1)} = [u_1, u_2, \dots, u_n]^T, \quad (3)$$

where $X^{(1)}$ is the feature of vertices before the first convolution layer, $u_i = v_i - \bar{v}_i, i \in n$, which is indicated by the blue arrows in Figure 6 (a).

The GCN generalizes the convolution operations from traditional data (such as images) to graph data. The core idea is to learn a function through which a node in the graph can aggregate its features and its neighbors' features to generate a new representation of the node. Our mesh model is a graph $G(V, E)$, where V is the set of vertices and E is the set of edges. The mesh is deformed by updating the coordinates of the vertices. The information that caused the deformation is delivered along the edges. In our framework, each layer is defined following the description by Kipf and Welling (*Kipf&Welling*, 2016)

$$X^{(l+1)} = \sigma(\hat{D}^{-\frac{1}{2}}(A + I)\hat{D}^{-\frac{1}{2}}X^{(l)}W^{(l)}), \quad (4)$$

where $X^{(l)}$ is the feature of vertices before convolution, $X^{(l+1)}$ is the feature of vertices after convolution, and l is the number of layers. $A \in \mathbb{R}^{n \times n}$ 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. Black arrows in Figure 6 (a) indicate the displacement caused by the GCN, which makes undetectable vertices move close to the target position. Figure 6 (b) shows how information is delivered into the GCN. In the first layer, detectable vertices deliver their information to their one-hop neighbors. In the second layer, one-hop neighbors deliver information to their one-hop neighbors, which is equivalent to the detectable vertices delivering information to their two-hop neighbors. As the number of layers increases,

the delivery range is extended. However, with a triangular surface mesh, it is difficult to complete the information transfer between the farthest vertices by increasing the number of layers, and we therefore add internal vertices to the triangular surface mesh to enhance connectivity of distant vertex pairs, consequently turning the mesh into a tetrahedral mesh.

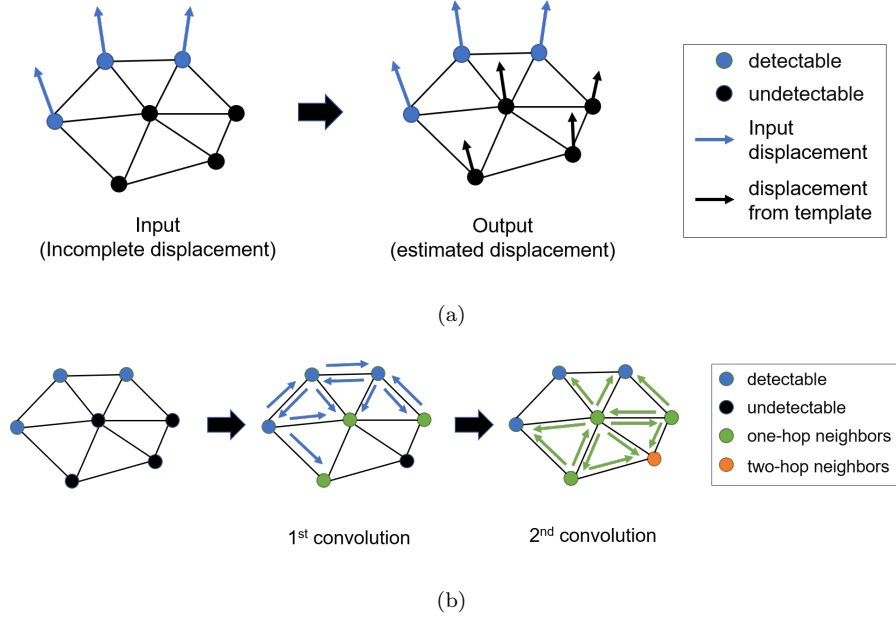


Figure 6: Deformation process in the GCN. (a) Input displacement and displacement caused by the GCN, (b) information delivery path in the GCN

3.3.3. Loss Function

This section will introduce two loss functions used in the proposed framework to obtain accurate and smooth organ shapes.

Positional loss

Positional loss is used to minimize the distance between the corresponding predicted shape and the ground truth shape vertices and is defined as:

$$\mathcal{L}_{pos} = \frac{1}{n} \sum_{i=1}^n \|v_i - \hat{v}_i\|_2^2, \quad (5)$$

where n is the number of vertices, v_i is the ground truth vertex, and $\hat{v}_i$ is the predicted shape vertex. This function acts to move the vertices on the template to the corresponding position on the ground truth. This will cause the overall shape to stretch, translate, rotate, and twist. These kinds of movements will result in a rough surface, so we also need a loss function to maintain the smoothness of the surface.

Laplacian loss

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

$$L(v_i) = \frac{1}{N(v_i)} \sum_{j \in N(v_i)} (v_i - v_j), \quad (6)$$

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 . The discrete Laplacian value of a vertex represents the curvature of the surface at that vertex. Keeping the curvature consistent can ensure that the surface of the calculated shape is smooth. Therefore, the discrete Laplacian loss is defined as,

$$\mathcal{L}_{laplacian} = \frac{1}{n} \sum_{i=1}^n \|L(v_i) - L(\hat{v}_i)\|_2^2, \quad (7)$$

where $L(v_i)$ is the discrete Laplacian of the ground truth, and $L(\hat{v})$ is the discrete Laplacian of the predicted shape. The Laplacian loss function restricts the deformation permitted by the positional loss function.

The total loss is the weighted sum of the two loss functions

$$\mathcal{L}_{total} = \mathcal{L}_{pos} + \lambda_{laplacian} \mathcal{L}_{laplacian}, \quad (8)$$

3.3.4. Evaluation value

We introduce the Euclidean distance error $\mathcal{D}_{Euclidean}$ (ED) to evaluate the distance error between the prediction result and the ground truth. The Euclidean distance error is defined as

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

where n is the number of vertices, v_i represents ground truth vertices, and $\hat{v}_i$ represents predicted shape vertices. $\mathcal{D}_{Euclidean}$ represents the distance error between the corresponding vertices.

To determine the weight of the Laplacian loss, we used cosine similarity (CS) to measure the surface difference between the output and the ground truth. CS calculates the cosine value of the angle between the normal vectors of the two corresponding triangle planes. The closer the value is to 1, the higher the similarity of the planes. The average CS of all planes is used to measure the surface quality of the entire shape. We compared the performance using 0, 0.001, 0.01, 0.1, and 1 as Laplacian loss weights, and obtained the results shown in Table 1. When the weight was 0.1, the Euclidean distance error was the smallest and the surface cosine similarity was the highest. Therefore, in the experiment in this paper, we used: $\lambda_{laplacian} = 0.1$.

$\lambda_{laplacian}$	ED(mm)	CS
0	8.7	0.957
0.001	8.7	0.957
0.01	8.6	0.957
0.1	8.6	0.960
1	9.4	0.963

Table 1: Evaluation of outputs under different Laplacian loss weights

In order to intuitively observe the surface smoothness, we also visually inspected the output.

4. Experiment

The following four experiments were conducted in this study. The framework was implemented in TensorFlow-GPU with a single NVIDIA GeForce RTX 2070 GPU. The network was trained using an Adam optimizer with 1×10^{-4} learning rate, 5×10^{-6} weight decay. We visualized the average training er-

ror and validation error for the 1000-step training, as shown in Figure 7. The validation error stabilized around 500 epochs, so we set the number of training epochs to 500. Given the limited number of samples, the batch size was set to 1. In our experiments, data from 100 patients were used for training, and data from 24 patients were used for testing.

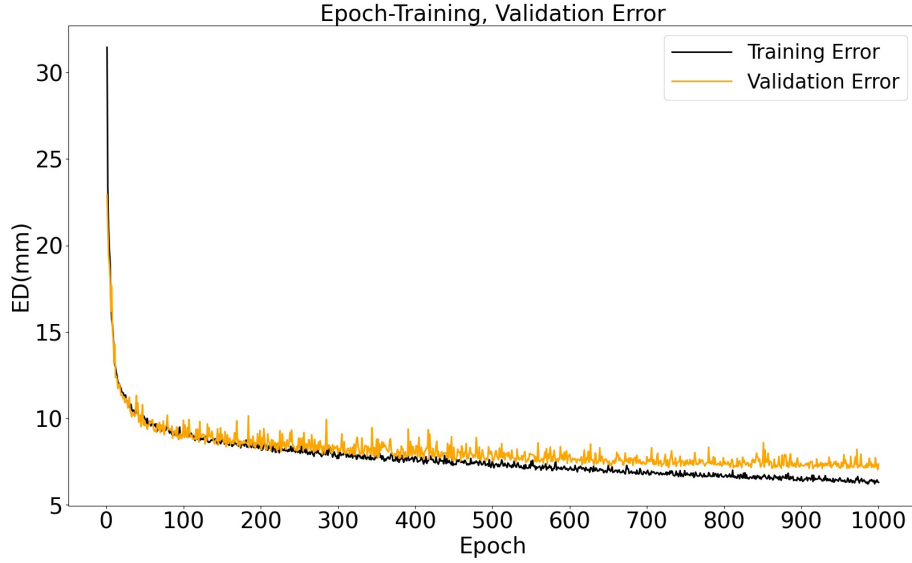


Figure 7: Learning curve. Euclidean error became stable at around 500 epochs.

4.1. Experiment 1: Number of detectable vertices and layers

To evaluate the effectiveness of the GCN for organ shape estimation within the proposed framework, we set the number of layers of the GCN to 0, 2, 4, 6, and 8 for repeated experiments. The 0-layer GCN consists of a simple fully connected network that also deforms the initial mesh based on detectable features. Detectable features are coordinates of detectable vertices. Trials were performed using from 10 to 500 detectable features in step 10. The detectable features were randomly selected and incrementally added to ensure that the previously selected random vertices were all the same and that only the newly added ones differed. The experimental results are shown in Figure 8. Compared

with the black curve representing no GCN layer, the test errors represented by the other curves are smaller. Within the selected range, as the number of layers increased the test error showed a decreasing trend, and the smaller the number of detectable features, the greater the error difference. In practical applications, the most important thing to focus on is those cases where the number of detectable features is low. Therefore, from the results of this experiment, we selected the test results with 100 detectable features for visual analysis, as shown in Figure 9. We chose the estimated shapes with the smallest and largest errors with the number of layers set to 8, and also show the results for the same cases with the number of layers set to 0. Both the error data and the visualization results demonstrate that the GCN layer played a major role in the shape estimation task. As the number of layers was increased, the difference between the estimation errors became smaller, but when the number of detectable features was in the range of 10 to 150, the performance of the 8-layer GCN was better than the other settings. Therefore, we chose eight layers for the following GCN experiments.

4.2. Experiment 2: Detection of the region in different directions

In reality, a centralized distribution in one location is more common than a random distribution of detectable features. Therefore, in this experiment, we set the centrally distributed regions as detectable features. We took detectable features in order along the z-axis and divided them into ascending and descending order for detecting features. We define the following coordinate axes for the human body, x: left-right, y: anterior-posterior, and z: head-foot. Correspondingly, the detectable parts were the bottom of the liver or the top of the liver. Figure 10 shows the error change curves for random addition, ascending order, and descending order of detectable features when the number of GCN layers was 0 and 8, respectively. It can be concluded from the graph that the detectable features from different directions do not affect the GCN. The estimation effect with GCN layers was much better than without GCN layers. The estimation accuracy was the highest when the detectable features were random, followed

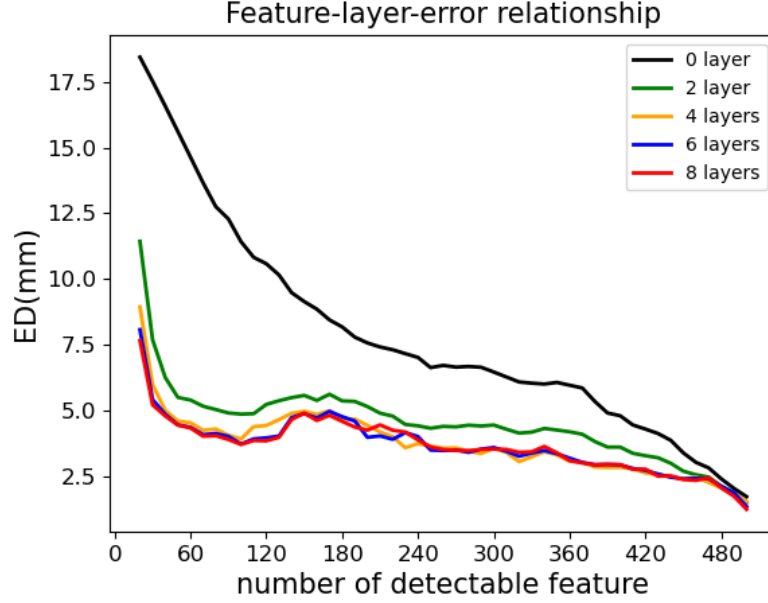


Figure 8: The relationships between test ED error, number of detectable features, and number of layers.

by descending order, then ascending order, which showed the lowest accuracy. When the detectable features were in ascending order, the feature vertices were clustered together and the feature information was relatively concentrated because the shape of the liver was wide at the top and narrow at the bottom, which is potentially the reason for the large error. In contrast, the randomly selected detectable features had the widest distribution, and the detectable features obtained in descending order had a wider distribution range on the x- and y-axes.

4.3. Experiment 3: Topology enhancement

From the results of Experiment 2, under the condition that the detectable features are taken in ascending and descending order along the z-axis, the shape estimation error is relatively large, and it is, therefore, necessary to reduce it further. In the experiments in this section, we selected 100 detectable features

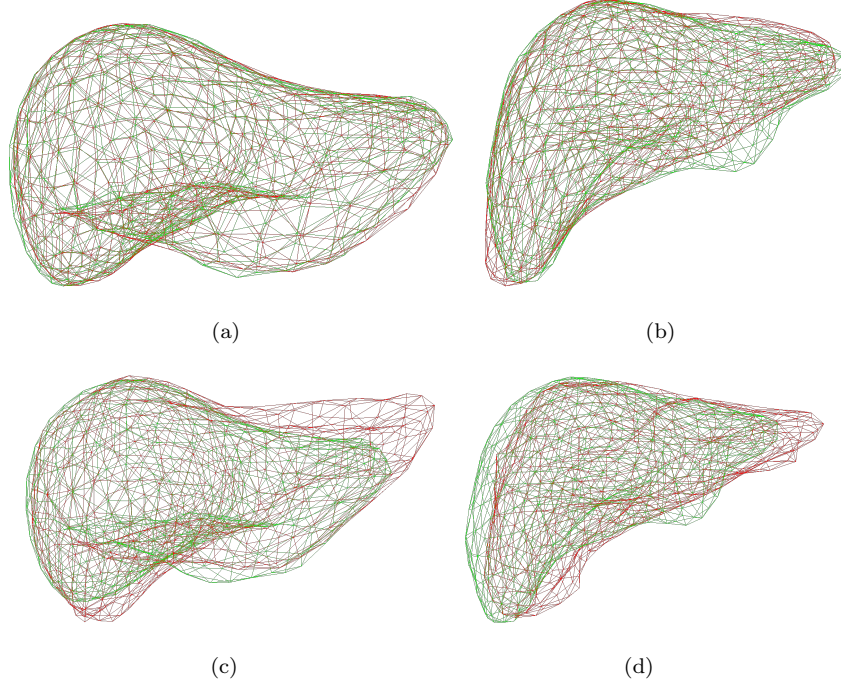


Figure 9: Visualization of the best and worst prediction results for different network layers. Green mesh represents the ground truth, red mesh the estimated shape (a) Case with minimum ED error and layer=8, (b) case with maximum ED error and layer=8, (c) the same case ID as (a) with layer=0 and (d) the same case ID as (b) with layer=0.

in descending order and added different numbers of internal vertices to verify the feasibility of the method to enhance the topology. We set the number of clusters from 0 to 20 for the experiments and the number of GCN layers to 4 as a control. The results are shown in Figure 11. Compared with the case where the number of clusters was 0, the error did not decrease after adding internal vertices. After the number of vertices added internally was greater than 6, the estimation error started to oscillate. When the number of clusters was 16, the average ED error reached a minimum of 8.32 mm. Compared with the average error with 0 clusters for the same 24 cases, which was 9.16 mm, the average error

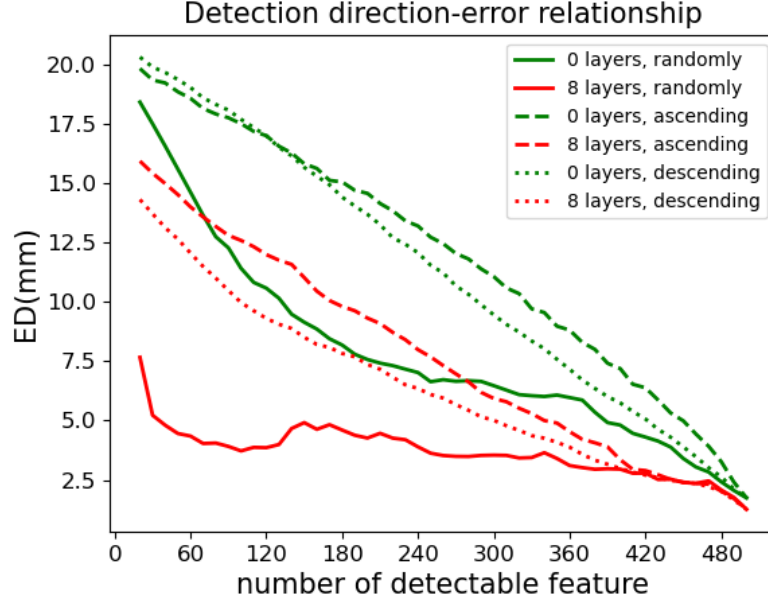


Figure 10: Influence of detectable features from different directions on the estimation results.

was reduced by 0.84 mm. At the same time, considering the calculation time, we set the number of clusters to 6 for specific experimental analysis. Figure 12 shows the ED error of the test data of the 24 cases under the above experimental conditions. The mean ED error of all test cases was 8.57 mm. Compared with the average error of 9.16 mm for these 24 cases with 0 clusters, the average error was reduced by 0.59 mm. Case 17 had the lowest error of 5.89 mm, whereas case 20 had the highest error of 17.59 mm. These cases are shown in Figure 13. At the same time, because the number of experimental samples was limited, we adopted a 5-fold cross-validation method, which divided 124 patients into five groups of 25, 25, 25, 25, and 24 to verify the stability of the method by cross-validation. The result of the 5-fold cross-validation was 8.18 ± 2.23 mm (mean ED $\pm$ standard deviation).

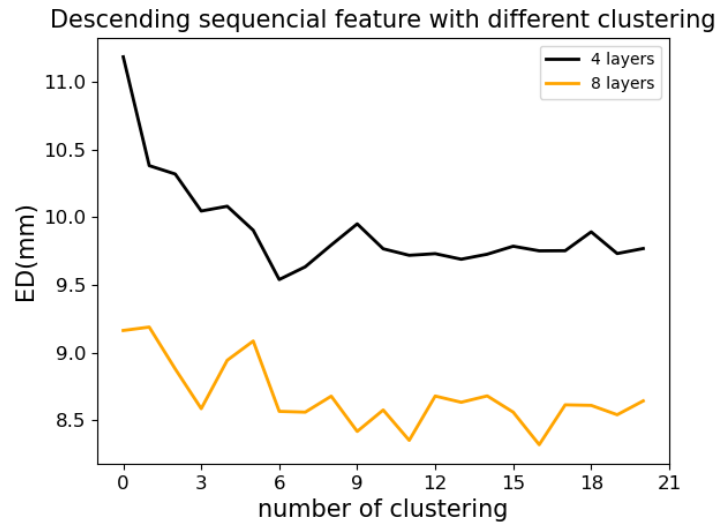


Figure 11: Relationship between ED error and the number of clusters.

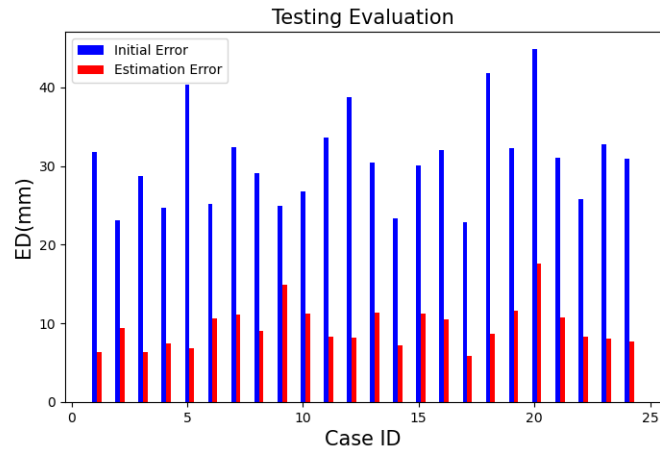


Figure 12: ED error for the liver in the 24 test cases

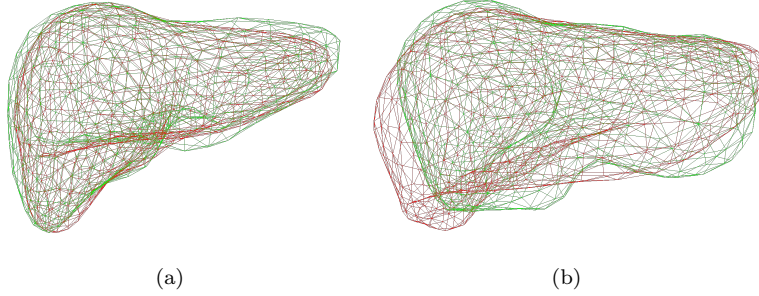


Figure 13: Visualization of the liver. Green mesh is the ground truth, and the red mesh is the estimated shape (a) Case 17 with the lowest error, (b) case 20 with the highest error

4.4. Experiment 4: Another organ

The liver has a special shape and a larger volume than the stomach, and therefore the estimated effects are likely to differ between the two organs. In this section, the stomach was set as the experimental object so that the difference in estimation results could be compared between the stomach and the liver under the same conditions. Section 4.2.3 specifically analyzed the situation with 8 network layers, the detectable features arranged in descending order along the z-axis, 100 detectable features, and 6 clusters. These parameter settings were maintained for the experiment on the stomach, and the following results shown in Figure 14 were obtained. The mean Euclidean distance error of all test cases was 12.77 mm. Case 4 had the lowest error of 6.12 mm, whereas case 20 had the highest error of 25.47 mm. A visualization of the cases with the highest and lowest errors is shown in Figure 15. The large error with case 20 may be due to two reasons. On the one hand, the detectable features were concentrated in the top part and had a little constraint on the lower end, whereas, on the other hand, the shape of the stomach was quite different from that in the other cases. Compared with the liver, the estimation error for the stomach was much larger under the same conditions, and shape detection for the stomach remains a challenging problem.

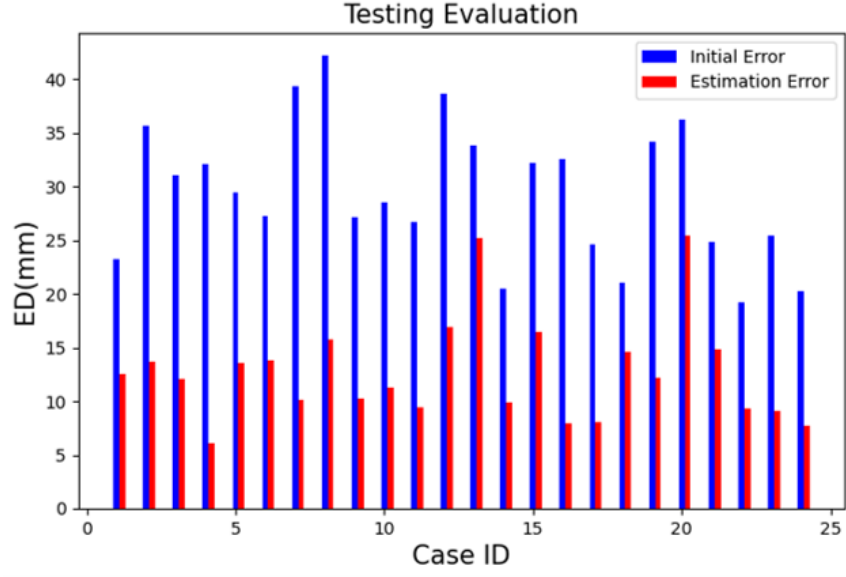


Figure 14: ED error for the stomach for the 24 test cases

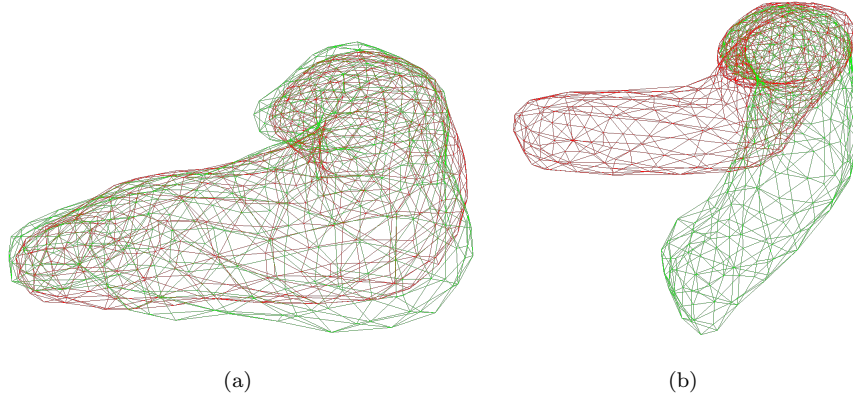


Figure 15: Visualization of the stomach. Green mesh is the ground truth, red mesh is the estimated shape. (a) Case 4 with the lowest error, (b) case 20 with the highest error

5. Conclusions and future work

This paper proposes a GCN-based method for estimating undetected organ regions using the information from detectable organ regions. A GCN-based method to enhance the data topology is proposed to obtain more accurate estimation results. We evaluated the effectiveness of the proposed method under different conditions using abdominal 3D-CT images from 124 patients who received radiation therapy. This study theoretically challenges the prediction of organ shape between patients with large individual differences. In practice, it could assist doctors in their work, providing reference information and improving work efficiency.

In the experiments, we demonstrated the role of the GCN in accomplishing the estimation task. With 100 randomly set detectable vertices, the estimation error has an average Euclidean distance as small as 3.72 mm. Through the experiments, we confirmed that the methods for setting the detectable features can greatly affect the estimation results and that the method for enhancing the topology structure proposed in this study can reduce the estimation error to a small value. We also showed that although the performance of our method differed between the liver and stomach, it was feasible for both organs.

Because the estimation accuracy did not reach clinically acceptable accuracy (which is within 5 mm) under many conditions in the experiment, it is necessary to continue to improve the method in the future, so that it can achieve higher accuracy. The experimental samples were all from Kyoto University Hospital, and because of this we cannot study the influence of differences between different populations. However, despite this homogenous dataset, the error still failed to meet the clinical requirements. We need to further pursue the generalizability of our method and improve its robustness. In addition, we intend to extend this intra-organ estimation framework to an inter-organ estimation framework in the future.

Acknowledgements

We thank Edanz (<https://jp.edanz.com/ac>) for editing a draft of this manuscript.

References

- Ci, H., Ma, X., Wang, C., & Wang, Y. (2020). Locally connected network for monocular 3d human pose estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, .
- Cootes, T. F., & Taylor, C. J. (2001). Statistical models of appearance for medical image analysis and computer vision. In *Medical Imaging 2001: Image Processing* (pp. 236–248). International Society for Optics and Photonics volume 4322.
- Ehrhardt, J., Werner, R., Schmidt-Richberg, A., & Handels, H. (2010). Statistical modeling of 4d respiratory lung motion using diffeomorphic image registration. *IEEE transactions on medical imaging*, *30*, 251–265.
- Fuerst, B., Mansi, T., Carnis, F., & Sälzle, M. (2014). Patient-specific biomechanical model for the prediction of lung motion from 4-d ct images. *IEEE transactions on medical imaging*, *34*, 599–607.
- Heitz, G., Rohlfing, T., & Maurer Jr, C. R. (2005). Statistical shape model generation using nonrigid deformation of a template mesh. In *Medical Imaging 2005: Image Processing* (pp. 1411–1421). International Society for Optics and Photonics volume 5747.
- Holzinger, A. (2014). On topological data mining. In L. Yu (Ed.), *Interactive Knowledge Discovery and Data Mining in Biomedical Informatics* (pp. 331–356). Springer volume 8401 of *Lecture Notes in Computer Science LNCS*. (1st ed.). doi:10.1007/978-3-662-43968-5_19.
- Islam, M. K., Purdie, T. G., Norrlinger, B. D., & Alasti, H. (2006). Patient dose from kilovoltage cone beam computed tomography imaging in radiation therapy. *Medical physics*, *33*, 1573–1582.

- Jia, S. (2019). *Population-based models of shape, structure, and deformation in atrial fibrillation*. Ph.D. thesis Université Côte d’Azur.
- Kan, M. W., Leung, L. H., Wong, W., & Lam, N. (2008). Radiation dose from cone beam computed tomography for image-guided radiation therapy. *International Journal of Radiation Oncology* Biology* Physics*, 70, 272–279.
- Kipf, T. N., & Welling, M. (2016). Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907*, .
- Luo, Y., Mi, Z., & Tao, W. (2021). Deepdt: Learning geometry from delaunay triangulation for surface reconstruction. *arXiv preprint arXiv:2101.10353*, .
- MacQueen, J. et al. (1967). Some methods for classification and analysis of multivariate observations. In *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability* (pp. 281–297). Oakland, CA, USA volume 1.
- Miao, S., Wang, Z. J., & Liao, R. (2016). A cnn regression approach for real-time 2d/3d registration. *IEEE transactions on medical imaging*, 35, 1352–1363.
- Nakamura, M., Nakao, M., Hirashima, H., & Iramina, H. (2019). Performance evaluation of a newly developed three-dimensional model-based global-to-local registration in prostate cancer. *Journal of Radiation Research*, 60, 595–602.
- Nakamura, M., Nakao, M., Mukumoto, N., & Ashida, R. (2021). Statistical shape model-based planning organ-at-risk volume: application to pancreatic cancer patients. *Physics in Medicine & Biology*, 66, 014001.
- Nakao, M., & Minato, K. (2010). Physics-based interactive volume manipulation for sharing surgical process. *IEEE Transactions on Information Technology in Biomedicine*, 14, 809–816.

- Nakao, M., Nakamura, M., Mizowaki, T., & Matsuda, T. (2021). Statistical deformation reconstruction using multi-organ shape features for pancreatic cancer localization. *Medical image analysis*, 67, 101829.
- Rigaud, B., Simon, A., Gobeli, M., & Leseur, J. (2018). Statistical shape model to generate a planning library for cervical adaptive radiotherapy. *IEEE transactions on medical imaging*, 38, 406–416.
- Tang, W., & Qiu, G. (2021). Dense graph convolutional neural networks on 3d meshes for 3d object segmentation and classification. *Image and Vision Computing*, 114, 104265.
- Tong, F., Nakao, M., Wu, S., & Nakamura, M. (2020). X-ray2shape: reconstruction of 3d liver shape from a single 2d projection image. In *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)* (pp. 1608–1611). IEEE.
- Wang, Y., Zhong, Z., & Hua, J. (2019). 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*, 26, 960–970.
- Wang, Z.-Y., Zhou, X.-Y., Li, P., & Theodoreli-Riga, C. (2020). Instantiation-net: 3d mesh reconstruction from single 2d image for right ventricle. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 680–691). Springer.
- Wu, S., Nakao, M., Tokuno, J., & Chen-Yoshikawa, T. (2019). Reconstructing 3d lung shape from a single 2d image during the deaeration deformation process using model-based data augmentation. In *2019 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI)* (pp. 1–4). IEEE.
- Yamamoto, U., Nakao, M., Ohzeki, M., & Tokuno, J. (2021). Kernel-based framework to estimate deformations of pneumothorax lung using relative

position of anatomical landmarks. *Expert Systems with Applications*, 183, 115288.

Yao, L., Mao, C., & Luo, Y. (2019). Graph convolutional networks for text classification. In *Proceedings of the AAAI conference on artificial intelligence* (pp. 7370–7377). volume 33.

Zhang, K., Hao, M., Wang, J., & de Silva, C. W. (2019). Linked dynamic graph cnn: Learning on point cloud via linking hierarchical features. *arXiv preprint arXiv:1904.10014*, .

Zhou, J., Cui, G., Hu, S., & Zhang, Z. (2020). Graph neural networks: A review of methods and applications. *AI Open*, 1, 57–81.