

3次元メッシュ変分オートエンコーダーを用いた 臓器形状アトラスの構築

梅原 隆一[†] 中村 光宏^{††} 中尾 恵^{††}

[†] 京都大学大学院情報学研究科 〒606-8501 京都市左京区吉田本町
^{††} 京都大学大学院医学研究科 〒606-8507 京都市左京区聖護院川原町 54
E-mail: †{u-ryuichi,megumi}@i.kyoto-u.ac.jp

あらまし 生体が有する臓器や骨格の形状と位置を少ないパラメーターで表現するモデルは放射線治療や外科手術支援など幅広い臨床応用が期待できる。しかし、軟臓器は患者間で多様な形状と位置を取るため、線形モデルでは局所的に変化が大きい形状を再構成することは難しく、非線形モデルでは特にデータ数が少ないときに過学習に陥りやすい。本研究では高精度で汎化性能が高い形状モデルの構築を目指し、3次元メッシュデータを入出力として、形状だけではなく位置も含めて高精度に再構成可能なメッシュ変分オートエンコーダーを構築した。125例からなる肝臓の臓器形状メッシュデータを用いて学習したモデルの検証を行い、内19症例のテストデータに対して、平均4.3mmの精度で位置と形状の再構成が可能であることを確認したので報告する。

キーワード 変分オートエンコーダー, メッシュ, 臓器形状アトラス, グラフ畳み込みネットワーク

Construction of an organ shape and position atlas using 3D Mesh Variational Autoencoder

Ryuichi UMEHARA[†], Mitsuhiro NAKAMURA^{††}, and Megumi NAKAO^{††}

[†] Graduate School of informatics, Kyoto University
Yoshida Honmachi, Sakyo-ku, Kyoto, 606-8501, Japan
^{††} Graduate School of Medicine, Kyoto University
54 Kawahara-chou, Shougoin, Sakyo-ku, Kyoto, 606-8507, Japan
E-mail: †{u-ryuichi,megumi}@i.kyoto-u.ac.jp

Abstract A model that represents the shapes and positions of organ or skeletal structures with a small number of parameters may be expected to have a wide range of clinical applications, such as radiotherapy and surgical guidance. However, because soft organs vary in shape and position between patients, it is difficult for linear models to reconstruct locally variable shapes, and nonlinear models are prone to overfitting, particularly when the quantity of data is small. The aim of this study was to construct a shape atlas with high accuracy and good generalization performance. We designed a mesh variational autoencoder that can reconstruct both nonlinear shape and position with high accuracy. We validated the trained model for liver meshes of 125 cases, and found that it was possible to reconstruct the positions and shapes with an average accuracy of 4.3 mm for the test data of 19 cases.

Key words Variational AutoEncoder, mesh, organ shape atlas, Graph Convolutional Network

1. はじめに

臓器形状アトラスは生体内の臓器の位置や形状などの幾何学的情報とその個人間の変動を少ないパラメータによって表現するモデルとして知られ、医用画像解析や生体医工学分野において幅広く利用されている。これまでに、臓器形状アトラスは統

計的形状モデリング (Statistical Shape Model, SSM) [1][2] の考え方に基づいて構築され、医療画像の位置合わせ [3] や手術計画支援 [4]、放射線治療における腫瘍位置の予測などに応用がなされている。SSM [5] は主に画像ベースの方法 [6] とメッシュベースの方法 [7] に大別される。画像ベースの方法は医用画像を直接利用して SSM を構築でき、また画像データは格子

状のデータ構造であるため、特徴量の抽出やフィルタリングなどの計算機上での処理が容易である。一方、データ量が画像サイズの3乗に比例して大きくなるため、高解像度なモデルの構築が難しく、複雑な臓器形状や臓器間の非線形かつ不連続な変形を表現することが難しい。一方、メッシュベースの方法では、データの作成に様々な前処理が必要である。またメッシュデータはグラフデータ構造であるため、統一的な特徴量の計算や変換などの計算機上の処理が画像に比べて難しいが、画像に比べて非常に小さなデータ量で形状を表現でき、変形も扱いやすいという特徴を有する。本研究では後者のメッシュベースのモデルを採用し、その形状表現と汎化性能の向上に着目している。

メッシュベースのSSMでは、頂点对応が取られた臓器形状メッシュ群から主成分分析 (Principal Component Analysis, PCA) や独立成分分析 (Independent Component Analysis, ICA) に基づいて特徴量を抽出し、平均形状との荷重和を取ることによって任意の臓器形状を表現する。臓器の形状データベースは多くがデータ数が数百例程度の少標本データである。一方、メッシュデータは3次元空間であれば頂点数の3倍の自由度を有する高次元データであるため、高次元かつ少標本なデータが処理対象となる。また、臓器の形状と位置は患者間で非常に多様な形状と位置をとる。このような複雑な高次元少標本データにおいては、PCAやICAに代表される線形モデルでは局所的に大きな変形を表現することが難しいという課題が存在する。この課題を解決するために近年では非線形モデルを使用した研究も行われている [7]。しかし、非線形モデルに限られたデータ数では過学習に陥りやすいという課題が存在する。そこで本研究では非線形モデルの中でも近年、医療分野を含む様々な分野で使用されている深層学習の枠組みを用いてこれらの課題の解決を試みる。深層学習を用いた研究例はすでに、[8]で報告されており、深層学習モデルの1つである、オートエンコーダー (Autoencoder, AE) を用いて実装されている。モデルの入力としては点群データを用いているが、臓器形状を考慮した場合、生体内において臓器は多臓器間で隣接しており、小腸や大腸といった臓器においては臓器間でも隣接しているため、頂点情報のみを使用したデータ表現では境界面を表現することは困難である。そこで本研究では、メッシュデータを用いた変分オートエンコーダー (Variational Autoencoder, VAE) である MeshVAE に注目した。VAE は画像データに対して広く利用されており、入力データを低次元に圧縮した潜在変数空間の分布として表現することから AE よりも更に高い汎化性能を持つと考えることができる。近年 MeshVAE を用いたメッシュ再構成方法が広く研究されており様々な枠組みが提案されている [9] [10]。しかし、それらのモデルでは変形勾配ベースの局所特徴量が用いられており頂点の絶対座標の位置を含めた再構築については対象とされていない。

本研究では生体臓器の複雑な形状と個人間の非線形な形状差を表現し、かつ、高い解釈性を有する深層形状アトラスの構築を目指している。本発表では最初の試みとして、形状再構成

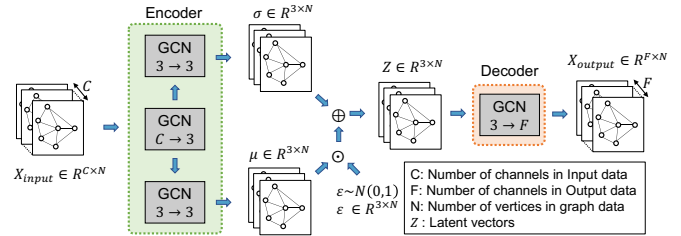


図1 提案モデル. GCN を用いて Encoder では3チャンネルの潜在変数を学習し、Decoder では潜在変数から出力への変換を学習する

に有効な特徴量の解析を行い、臓器の3次元メッシュデータを入力とし、形状だけではなく位置も含めて高精度に再構成可能なメッシュ変分オートエンコーダーの枠組みを提案する。125名の患者の3次元CT画像から生成した肝臓形状のメッシュ群を用いてモデルの学習を行い、再構成性能を確認する。

2. 提案方法

図1に提案するメッシュ変分オートエンコーダーのモデルを示す。本モデルはEncoderとDecoderから構成され、Encoderにグラフ畳み込みネットワーク (Graph Convolutional Network, GCN) を利用している。メッシュ再構成時には、最初に臓器メッシュが持つ頂点と辺の情報から、モデルに入力する特徴量 X_{input} を算出する。次に、Encoderは入力された特徴量からGCNに基づいて各頂点単位に定義する潜在変数 Z を計算する。Decoderは得られた潜在変数 Z からメッシュの特徴量を再構成し、モデルの出力 X_{output} とする。本モデルにおける重みの学習はモデルの入出力から定義される損失関数の最小化によって達成される。

なお、本研究において再構成対象とする臓器メッシュは頂点数と辺の数が同じで、かつ、頂点对応が取れたメッシュ群を想定している。患者固有の3次元CT画像から得られるメッシュはそのままでは患者間で頂点や辺が統一されていない。そこで、前処理として平均的な形状を有するテンプレートメッシュを患者固有の臓器メッシュに対して事前にメッシュ変形位置合わせ [7] [11] を行うことによって頂点对応が取れたメッシュを得る。

2.1 臓器メッシュの特徴量

提案モデルの入力とする特徴量にはメッシュの頂点位置が利用可能であるが、臓器は複雑な3次元形状を有する場合があります。GCNを用いても頂点情報のみから高精度な形状再構成が達成されるとは限らない。そこで本研究では、メッシュから事前に算出可能な以下の3種類の特徴量とその組み合わせを考える。

- 頂点の位置座標 (3次元)
- 変形勾配 (9次元)
- 離散ラプラシアン (3次元)

2.1.1 変形勾配

変形勾配 (deformation gradient, DG) は図2(a)に示すように、2物体間の対応する点同士的位置関係の変化を記述する量

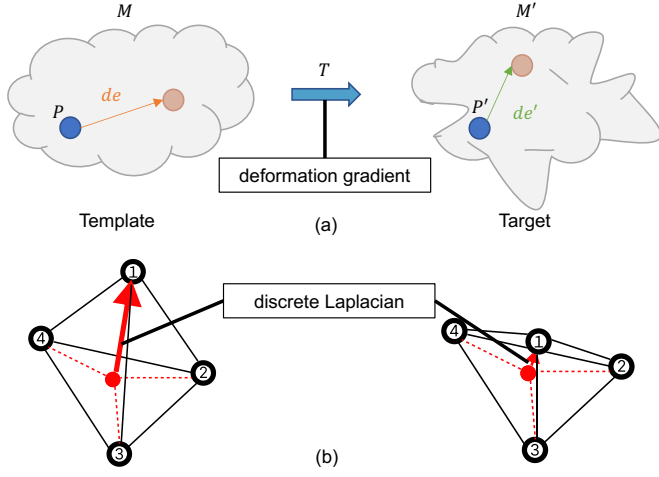


図 2 本研究で使用されている特徴量。(a): DG の概念図。(b): DL の概念図。左の場合は曲率が大きく、右の場合は曲率が小さくなっている。

である [9]。ある物体 M が伸縮や曲げ、ねじれなどの局所的な変形によって M' になったとする。 M おいて微小距離離れた点 P と $P + de$ が変形の影響によって、 P' と $P' + de'$ にそれぞれ移動したとする。このとき、 P と $P + de$ の間における DG、 T は式 (1) のように定義される。

$$de' = Tde \quad (1)$$

DG は頂点座標が 3 次元ベクトルであるとする、 3×3 の行列である。 T に特異値分解の一種である極分解を適用することで、以下のように正定値対称行列 S と直行列 R に分解される。

$$T = RS \quad (2)$$

ここで、 R は変形における回転や曲げを表現する回転行列であり、 S は歪テンソルである。例えば、物体 M が全く変形せず、移動のみが生じたとする、式 (1) から T は単位行列 I となる。平行移動では形状に回転や歪みは生じておらず、式 (2) においても $R = S = I$ となって物体が示す特性と一致する。

次に、DG の算出方法を説明する。メッシュ変形位置合わせ前の平均形状テンプレートの形状を M とし、各患者 i のメッシュ $M^{(i)}$ への変形を表す DG、 $T^{(i)}$ を求め、特徴量 $f^{(i)}$ を作成する。まず、メッシュ $M^{(i)}$ における頂点 j に対して DG、 $T_j^{(i)}$ を用いた式 (3) のエネルギー $E(T_j^{(i)})$ を考える。

$$E(T_j^{(i)}) = \sum_{k \in N_j} c_{jk} |e'_{jk} - T_j^{(i)} e_{jk}|^2 \quad (3)$$

ここで、

$$e_{jk} = v_j - v_k,$$

$$e'_{jk} = v'_j - v'_k$$

であり、 v'_j は $M^{(i)}$ の頂点位置、 v_j は M の頂点位置である。また、 N_j は頂点 v_j に隣接する頂点の集合である。 c_{jk} は以下のように定義される。

$$c_{jk} = \cot \alpha + \cot \beta$$

α と β は辺 e_{jk} を共有する 2 つの三角形の対角に位置する角度を指す。本係数は式 (3) の最適化において、メッシュの表面の滑らかさを維持する働きがある。次に、

$$dR_{jk} = R_j^T R_k \quad (4)$$

によって定義される回転差分を定義する。この回転差分は、メッシュ全体における回転成分を除いた後の隣接する頂点を結ぶ辺同士の回転を表す。さらに回転差分の自然対数 $\ln(dR_{jk})$ を計算することで、要素の値が回転角となり、回転角を歪テンソルと同様に線形に扱うことが可能になる。

$$\ln(dR_{jk}) = \begin{pmatrix} 0 & \omega_x & \omega_y \\ \omega_x & 0 & \omega_z \\ \omega_y & \omega_z & 0 \end{pmatrix} \quad (5)$$

また、歪テンソル S_j は正定値対称行列であるため、頂点 j における DG ベースの特徴量 f_j^M は dR_{jk} の独立な 3 成分と、 S_j の独立な 6 成分を用いて以下のように表現される。

$$f_j^M = \left(\frac{1}{|N_j|} \ln(dR_{jk}); S_j \right) (\forall k \in N_j) \quad (6)$$

2.1.2 離散ラプラシアン

離散ラプラシアン (discrete Laplacian, DL) は各頂点に定義される 3 次元特徴量であり、対象頂点の位置とその周辺頂点の位置との差分ベクトルとして定義される。辺の長さや隣接頂点数がメッシュ全体で比較的均一であるメッシュに対しては各頂点における平均曲率を近似的に表現する特徴量である [11]。図 2(b) に示すように各頂点において曲率が大きいほど、DL の大きさも大きくなる。

次に、DL の算出の流れについて説明する。DL は式 (7)

$$L = D - A \quad (7)$$

で定義されるグラフラプラシアン (D 、 A はそれぞれグラフにおいて各頂点の隣接頂点数を表現する次元行列と、頂点同士の接続関係を表現する隣接行列である。) を用いて、頂点 v_j の 3 次元位置座標を並べた行列 V に対して式 (8) のように定義される。

$$f^{DL} = LV \quad (8)$$

頂点単位では DL 特徴量 f_i^{DL} として以下のように計算される。

$$f_i^{DL} = |N_i| v'_i - \sum_{j \in N_i} v'_j \quad (9)$$

2.2 ネットワーク構成

本研究ではネットワーク構成として、MeshVAE [9] で用いられている全結合層 (FC) ベースのネットワーク構成とグラフ畳み込み (GCN) を使用したネットワーク構成の 2 種類を検討した。

FC では入力ベクトルの全要素から出力成分を決定する一方で、GCN では隣接頂点のみの情報から出力成分を計算するので、効率的に学習を進めることが可能であると考えられる。

GCN は画像分野における DeepLearning 手法で広く用いられている畳み込み演算をグラフデータに対して応用した手法であり、隣接頂点間で情報をやり取りするため FC よりも効率的に学習を進めることが可能であると考えられる。

2.3 損失関数

本研究では汎化性能を担保しながら、高精度の再構成を行うために、潜在変数の分布と入出力の再構成誤差に関する損失関数を導入する。

前者に関しては、潜在変数 Z の分布を多重正規分布に近づけるために式 (10) で定義される \mathcal{L}_{KL} を導入する。

$$\mathcal{L}_{KL} = \frac{1}{N} \sum_{i=0}^N DK[N(\mu_i, \sigma_i^2) || N(0, 1)] \quad (10)$$

ここで $N(\mu, \sigma^2)$ は平均 μ 、標準偏差 σ の正規分布を表す。

後者に関しては、特徴量 F に対して再構成誤差 \mathcal{L}_F を式 (11) のように定義する。

$$\mathcal{L}_F = MSE(F^{label}, F^{pred}) \quad (11)$$

ここから全体の損失関数 \mathcal{L} は式 (12) のように計算される。

$$\mathcal{L} = \mathcal{L}_{KL} + \alpha\mathcal{L}_{Pos} + \beta\mathcal{L}_{DG} + \gamma\mathcal{L}_{DL}, \quad (12)$$

ここで \mathcal{L}_{Pos} 、 \mathcal{L}_{DG} 、 \mathcal{L}_{DL} はそれぞれ、位置、DG、DL の再構成誤差を表す。ただし、式 (12) はすべての特徴量を用いる場合の損失関数であり、実験において利用しない特徴量に関しては係数を 0 とする。例えば、特徴量として頂点位置と DG を用いた場合は $\gamma = 0$ とする。

3. 実験

提案手法の有効性を確認するために、提案する特徴量とモデルに関していくつかの組み合わせで実験を行った。提案モデルの実装には Python3.9, Pytorch, Pytorch-Geometric を用いた。学習におけるバッチサイズを 32、トレーニング epoch は最大を 1000 として、validation 用データの \mathcal{L}_{Pos} を監視する EarlyStopping を用いて学習を進めた。EarlyStopping の patience は 50epoch に設定した。またネットワークの最適化には学習率を 1×10^{-2} に設定した Adam を用いた。損失関数のハイパーパラメータは幾つかの組み合わせについてモデルの性能を確認し、最も性能が高かった $\alpha = 1.0 \times 10^{12}$, $\beta = 1.0 \times 10^{12}$, $\gamma = 1.0 \times 10^6$ を用いることとした。

3.1 データセット及び前処理

京都大学医学部附属病院放射線治療科において、膵がんの放射線治療を受けた 125 名の患者の腹部 3D-CT データから得られた肝臓の臓器メッシュを対象に実験を行った。放射線腫瘍専門医によって定義された輪郭情報から各臓器領域のサーフェスマッシュを得た後、テンプレートベースのメッシュ変形位置合わせによって位置合わせ済みの肝臓メッシュを得た。これら前処理は過去の研究 [7] [11] において実施された手続きと同じである。なお、一連の実験は京都大学における医の倫理委員会によって承認された (承認番号: R1446)。

表 1 モデルと特徴量の組み合わせと学習結果. Epoch は EarlyStopping で Val データにおける再構成誤差が最良の Epoch を示す. Train, Val, Test の列は EarlyStopping を用いて、それぞれのデータに対して推論を行った際の、正解データと推論データの頂点の平均距離を示す。距離の単位は [mm]

model	feature	epoch	Train	Val	Test
FC	Position	230	6.7509	11.929	12.158
GCN	Position	995	5.3704	5.4113	5.3337
GCN	Position + DG	676	5.1983	5.1691	5.2139
GCN	Position + DL	981	4.2451	4.2300	4.3023
GCN	Position + DL + DG	991	4.2126	4.2298	4.2643

メッシュ頂点は放射線治療における治療用ビームの照射対象である膵癌の位置を原点として、左右: x 軸、背腹: y 軸、頭尾: z 軸のそれぞれにおいて $[-256, 256]$ (単位: mm) の範囲に存在している。利用する特徴量間のオーダーを合わせるために、それぞれの特徴量に対して以下の定数で除算を行うことで、値の範囲が $[-1, 1]$ になるように正規化した。125 名のメッシュに対して、87 を学習用に、19 を検証用に、19 をテスト用に分割して実験を行った。

- 位置の特徴量に対して 256
- DG の局所回転 dR_{ij} に対して π
- DG のストレッチテンソル S_i に対して 20
- DL 特徴量に対して 100

3.2 メッシュ再構成性能の比較

本実験では臓器メッシュの再構成に有用な特徴量とモデルの組み合わせを探るべく、表 1 の組み合わせを比較対象にモデルの学習を行い、肝臓メッシュの再構成結果を得た。結果の定量評価には式 (13) に基づいて再構成結果 v' と正解データ v の対応する頂点同士の距離を計算し、症例ごとに平均を算出することで推論結果の再構成誤差 E を計算した。

$$E = \frac{1}{NM} \sum_i^M \sum_j^N \|v_i^{(j)} - v'_i{}^{(j)}\|_2 \quad (13)$$

ここで N は 1 つのメッシュデータの頂点数を表し、 M はデータセットに含まれるメッシュデータの数を表す。

表 1 から頂点位置 + DL + DG の特徴量を用いた GCN ベースのモデルが、Test データにおける精度が最も良いことが確認された。次に精度が良い頂点位置 + DL の特徴量を用いた GCN ベースのモデルとの比較のために、Wilcoxon の符号順位検定を用いて有意水準 0.05 で両側検定を行ったところ p 値が 0.007145 となり、両者の結果には有意差があることが確認された。

再構成結果を 2 例可視化した図を図 3 に示す。図から FC を用いたモデルは全体的に誤差が大きく、GCN を用いたモデルは全体的に精度良く再構成できていることが分かる。特に図中矢印で示す左葉の先端部分において、FC を用いたモデルでは位置を大きく外してしまっている一方、GCN を用いたモデルでは精度良く再構成できていることから GCN を用いたモデルは FC を用いたモデルよりも優れていることが確認できる。

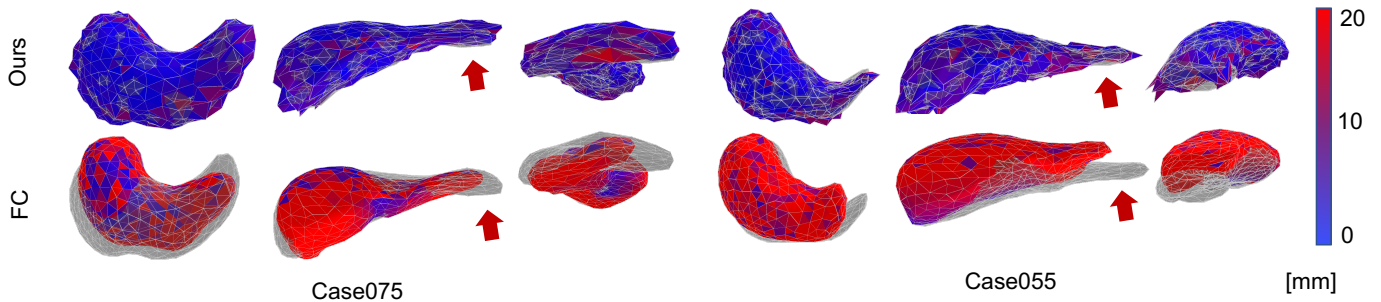


図 3 テストデータ内で最も再構成誤差が大きかった 2 症例の推論結果. 左から axial, coronal, sagittal. 灰色の部分は正解データのメッシュであり、着色されたメッシュは推論結果を表す. 赤色の部分ほど誤差が大きく、青色になればなるほど誤差が小さいことを表す. [FC] 特徴量: 頂点の位置, モデル: FC ベース [Ours] 特徴量: 頂点位置 + DL + DG モデル: GCN ベース

4. おわりに

本研究では高精度で汎化性能が高い形状モデルの構築を目指し、3次元メッシュデータを入出力として、形状だけではなく位置も含めて高精度に再構成可能なメッシュ変分オートエンコーダーを構築した。125例からなる肝臓の臓器形状メッシュデータを用いて学習したモデルの検証を行い、内19症例のテストデータに対して、平均4.3mmの精度で位置と形状の再構成が可能であることを確認した。今後は今回の研究で得られた結果を元に、モデルのアーキテクチャを工夫し解釈可能性を向上させ、人体における臓器形状を構成する少数のパラメーターの解明に取り組む。

謝辞

本研究は日本学術振興会 科研費補助金 基盤研究 (B) 22H03021 及び 19H04484 の支援による。

文 献

- [1] F. Stefan, B. Mario, "Example-driven deformations based on discrete shells," *Computer Graphics Forum*, vol.30, pp. 2246-2257, 2011.
- [2] R. W. Sumner, M. Zwicker, C. Gotsman, J. Popović, "Mesh-based inverse kinematics," *Special Interest Group on Computer Graphics (SIGGRAPH)*, pp. 488 - 495, 2005.
- [3] B. Koo, E. Özgür, B. Le Roy, E. Buc, A. Bartoli, "Deformable registration of a preoperative 3D liver volume to a laparoscopy image using contour and shading cues," in *Medical Image Computing and Computer Assisted Intervention (MICCAI)*, pp. 326 - 334, 2017.
- [4] S. Raith et al., "Planning of mandibular reconstructions based on statistical shape models," *International Journal of Computer Assisted Radiology and Surgery*, vol. 12, no. 1, pp. 99 - 112, 2017.
- [5] T. Heimann, H. Meinzer, "Statistical shape models for 3D medical image segmentation: A review," *Medical Image Analysis*, vol. 13, no. 4, pp. 543 - 563, 2009.
- [6] S. Oh, S. Kim, "Deformable image registration in radiation therapy," *Radiation Oncology Journal*, vol. 35, no. 2, pp. 101 - 111, 2017.
- [7] M. Nakao, M. Nakamura, T. Mizowaki, T. Matsuda, "Statistical deformation reconstruction using multi-organ shape features for pancreatic cancer localization", *Medical Image Analysis*, vol. 67, p. 101829, 2021.
- [8] H. Dai, L. Shao, "PointAE: point auto-encoder for 3D statistical shape and texture modelling," *International Conference on Computer Vision (ICCV)*, pp. 5409-5418, 2019.
- [9] Q. Tan, L. Gao, Y. K. Lai, S. Xia, "Variational autoencoders for deforming 3D mesh models," *Computer Vision and Pattern Recognition (CVPR)*, pp. 5841-5850, 2018.
- [10] Y. J. Yuan, Y. K. Lai, J. Yang, Q. Duan, H. Fu, L. Gao, "Mesh variational autoencoders with edge contraction pooling," *Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 1105-1112, 2020.
- [11] M. Nakao, M. Nakamura, T. Matsuda, "Image-to-graph convolutional network for 2D/3D deformable model registration of low-contrast organs," *IEEE Trans. on Medical Imaging*, Vol. 41, No. 12, pp. 3747-3761, 2022.