

胸腔鏡画像における教師なし画像変換を用いた肺領域抽出

新田 潤平[†] 中尾 恵[†] 今西 勁峰^{††} 松田 哲也[†]

[†] 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

^{††} イーグロース株式会社 〒604-8006 京都市中京区下丸屋町 403

E-mail: [†]njum@sys.i.kyoto-u.ac.jp

あらまし 内視鏡手術において安全性向上の観点から関心領域の三次元構造の把握が必要である。術中の変形が小さい臓器については事前に撮影した CT 画像を三次元構造の把握のために用いることができるが、術中変形の大きい臓器については変形推定が必要である。術中臓器の変形推定に関する研究は広く行われているが、変形推定を行うためにはカメラ画像内の二次元臓器領域の抽出が必要である。本研究では、術中変形が大きい臓器である肺を対象に、U-net を用いて胸腔鏡画像に対する肺領域抽出を試みた。胸腔鏡画像のうち喫煙者の画像については抽出精度が低下する問題に対し、正則化項を追加した損失関数に基づいた CycleGAN を用いて教師なし画像変換を行い、抽出精度の向上を試みたので報告する。

キーワード 教師なし学習, 肺領域抽出, 敵対的生成ネットワーク, 胸腔鏡画像

Lung region segmentation of thoracoscopic image with unsupervised image translation

Jumpei NITTA[†], Megumi NAKAO[†], Keiho IMANISHI^{††}, and Tetsuya MATSUDA[†]

[†] Graduate School of Informatics, Kyoto University

Yoshida Honmachi, Sakyo-ku, Kyoto, 606-8501, Japan

^{††} e-Growth Co., Ltd., 403, Shimo-Maruya-cho, Nakagyo-ku, Kyoto, 604-8006, Japan

E-mail: [†]njum@sys.i.kyoto-u.ac.jp

Abstract In endoscopic surgery, it is necessary to understand the three-dimensional structure of the target region to improve safety. For organs that do not deform much during surgery, preoperative CT images can be used to understand their three-dimensional structure, however, deformation estimation is necessary for organs that deform substantially. Even though the intraoperative deformation estimation of organs has been widely studied, two-dimensional organ region segmentations from camera images are necessary to perform this estimation. In this paper, we performed lung region segmentation method using U-net for the thoracoscopic image, which is an organ that deforms substantially during surgery. To solve the problem of low segmentation accuracy of smoker thoracoscopic images, we performed unsupervised image translation using a CycleGAN, which we added a regularization term to the loss function. This presentation reports the image translation results and its effect of lung region segmentation.

Key words Unsupervised learning, lung region segmentation, generative adversarial network, thoracoscopic images

1. はじめに

近年医療技術および手術を支援する機器などの進歩により、内視鏡を用いた手術が広く行われるようになっていく。内視鏡手術は従来の開腹手術と比べて痛みが少なく、術後の回復も早い。術後創も開腹手術に比べて小さいという利点もある反面、手術の難易度が高く手術時間も長くなることが課題となっている。内視鏡手術ではカメラを体内に挿入し、モニタ

を通して手術を行うため安全性向上の観点から関心領域の三次元構造の把握が必要である。しかし、内視鏡の狭い術野から臓器の三次元構造を捉えることは容易ではなく、これを支援する研究が進められている [1] [2]。

例えば肝臓は術中の体積変化が小さい臓器であり、事前に撮影した Computed Tomography (CT) 画像から作成した臓器の血管や腫瘍のバーチャル画像を手術のガイドとして腹腔鏡画像に重畳する研究 [3] が報告されている。一方で術中の変形が

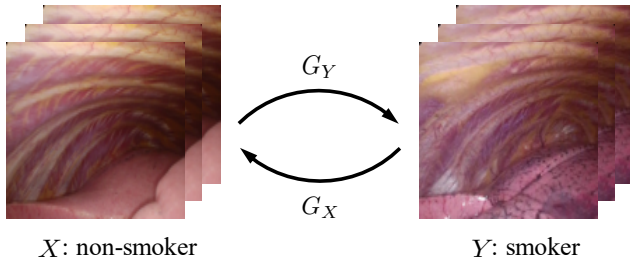


図1 肺画像に対する敵対的生成ネットワークによる画像変換, X:非喫煙者肺画像群, Y:喫煙者肺画像群.

大きい臓器は、3次元CT画像から生成された臓器モデルと術中の臓器形状に差が大きく、単純にバーチャル画像を重ね合わせることはできない。特に肺は術中に胸郭内部の圧変化に伴って肺実質が大きく虚脱し、その変形量も50%以上と大きい[4]、腫瘍位置の同定には変形推定が必要不可欠である。従来研究において、術中カメラ画像における臓器形状に基づいて変形推定が可能であることが示されているが[5]、カメラ画像内の肺領域の抽出が前提となる。Wuらは手術動画における臓器変形の連続性に着目し、グラフカットに基づく肺領域抽出方法[6]を提案している。また今西らはU-net[7]を手術動画に応用し、連続するフレーム画像に対して高精度な抽出を達成している[8]。

本研究では前実験として、胸腔鏡画像における静止画像にU-netを適用し、肺領域の抽出を試みた。その結果、抽出の精度にばらつきがあり、特に喫煙者の肺では非喫煙者の肺と比べて抽出精度が低下することが確認された。喫煙者の肺は非喫煙者と比較して、肺や胸腔が紫色や黒色に変色しており、このバリエーションの多さが抽出精度の低下の原因となっていると考えられた。この問題に対し、本研究では教師なし画像変換によって喫煙者肺の画像を非喫煙者肺の見た目に近くなるように変換する前処理を施すことによってU-netによる抽出精度が向上するか否かを調査することを目的とした。

深層学習を用いた画像のスタイル変換に関する研究は多く報告されており[9][10][11]、中でも画像の大域的特徴を維持したまま局所的な物体の描き方のみを変える手法としてCycleGAN[12]が提案されている。CycleGANはGenerative Adversarial Networks (GAN)[13]の派生手法の一つであり、二つの画像群の特徴を学習し一方の画像群に属する画像を他方の画像群に属する画像に変換する教師なし学習の手法である。

学習および評価には、胸腔鏡下肺がん切除術の手術動画のうち術具が含まれていないシーンから複数フレームを抽出した静止画像群(図1)を用いる。まず静止画像群に対してU-netの学習を行い、次に非喫煙者肺画像群と喫煙者肺画像群のそれぞれに対して抽出精度を確認する。また2群の画像間の変換をCycleGANによって学習し、学習済みモデルを用いた画像のスタイル変換がU-netによる領域抽出精度の向上に貢献するか否かを確認する。

表1 モデルの学習に用いた胸腔鏡画像データ

	U-net		CycleGAN	
	非喫煙者	喫煙者	非喫煙者	喫煙者
手術動画	12 例	13 例	1 例	1 例
抽出画像枚数	666 枚	1,126 枚	55 枚	80 枚
拡張後枚数	9,324 枚	15,764 枚	770 枚	1,120 枚
拡張後合計枚数	25,088 枚		1,890 枚	

2. 手 法

2.1 胸腔鏡画像データ

本研究で用いる胸腔鏡画像データについて説明する。本研究では京都大学医学部附属病院呼吸器外科から提供を受けた胸腔鏡下肺がん切除術の手術動画25例から、術具が含まれていないシーンを20から100フレーム程度抽出した静止画像群を用いる。表1に枚数の内訳を示す。データを構成する25例の手術動画の内訳は非喫煙者が12例、喫煙者が13例である。さらに抽出した静止画像に対して、平行移動、アフィン変換、拡大縮小、コントラスト変更等の処理によってデータ拡張を行った。また肺領域抽出の際の正解データとして、各静止画像に対応する肺領域を示す画像を手作業によって作成した。U-netの学習には25例分のデータを全て用い、CycleGANの学習には2例分のデータを用いた。テストデータには、6例の手術動画から2シーンずつ静止画像を抽出した計12枚の胸腔鏡画像を用いた。ここで6例のうち5例の手術動画は学習に用いた25例に含まれないものであり、残りの1例はCycleGANの学習に用いたデータのうち喫煙者のデータと同じ手術動画である。ただし、学習データに用いたシーンとは異なるシーンを抽出することとした。

2.2 肺領域抽出モデル

本研究では教師なし画像変換を用いて胸腔鏡画像を変換し、得られた画像に対して肺領域抽出法を適用することを提案する。喫煙者の胸腔鏡画像を非喫煙者の見た目をもつ画像に変換することで抽出精度の向上を試みる。図2に提案する枠組みを示す。まず喫煙者肺画像をCycleGANによって非喫煙者肺画像の見た目をもつ画像に変換する。次に得られた変換後画像をU-netに入力し、肺領域を抽出する。この手続きによって喫煙者肺画像に対する肺領域抽出の難易度が低下し、抽出精度が向上することが期待される。評価実験においてCycleGANによる変換を行わない場合との比較を行い、抽出精度の向上を確認する。

肺領域抽出モデルは[8]で提案されているモデル1を用いる。これはU-net[7]をそのまま用いたモデルであり、1枚の静止画像を対象として肺領域ラベルを出力する。ここで入力である静止画像のサイズは256×256であり、チャンネル数はRGBの3チャンネルである。また出力はサイズ256×256で1チャンネルの2値ラベル画像であり、白が肺領域を示し、黒が背景を示す。

2.3 胸腔鏡画像に対する教師なし画像変換

CycleGANでは2つのGANを用いてネットワークを構成し、

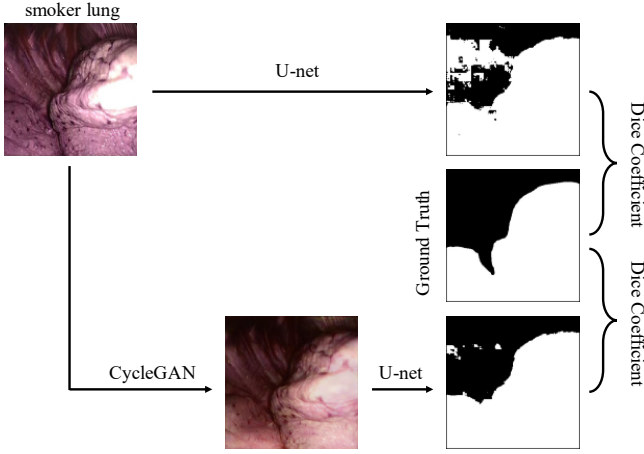


図2 教師なし画像変換を用いた肺領域抽出の概要図. U-net: 肺領域抽出, CycleGAN: スタイル変換, Ground Truth: 真値となる肺領域ラベル.

それぞれの GAN が一方の画像群に属する画像から他方の画像群に属する画像を生成するように学習を進める. このとき片方の生成器で生成した画像をもう片方の生成器に入力して再構成画像を生成することで, 生成画像が実画像の特徴を保った画像となることを要請する. CycleGAN では GAN の損失関数である Adversarial loss に Cycle consistency loss を加えたものを損失関数としている. CycleGAN で扱う 2 つの画像群を X, Y とすると Adversarial loss は式 (1) のように定義される.

$$\begin{aligned} \mathcal{L}_{\text{adv}}(G_X, D_X) = & \mathbb{E}_x[\log D_X(x)] \\ & + \mathbb{E}_y[\log(1 - D_X(G_X(y)))]. \end{aligned} \quad (1)$$

ここで G_X は Y に属する画像 y から X に属する画像 x を生成する生成器であり, D_X は画像が実画像であるか G_X によって生成された生成画像 $G_X(y)$ であるかを識別する識別器である. また G_Y, D_Y についても同様である. また Cycle consistency loss は式 (2) のように定義される.

$$\begin{aligned} \mathcal{L}_{\text{cyc}}(G_X, G_Y) = & \mathbb{E}_x[\|G_X(G_Y(x)) - x\|_1] \\ & + \mathbb{E}_y[\|G_Y(G_X(y)) - y\|_1]. \end{aligned} \quad (2)$$

Cycle consistency loss によって再構成画像が実画像と同じ画像になるように学習を進めることができ, これによって生成画像が実画像の特徴を保持したままの画像となることを要請できる. CycleGAN の損失関数をまとめると,

$$\begin{aligned} \mathcal{L}_{\text{cgan}} = & \mathcal{L}_{\text{adv}}(G_X, D_X) + \mathcal{L}_{\text{adv}}(G_Y, D_Y) \\ & + \lambda_{\text{cyc}} \times \mathcal{L}_{\text{cyc}}(G_X, G_Y), \end{aligned} \quad (3)$$

となる. ここで λ_{cyc} は Adversarial loss と Cycle consistency loss のバランスを決めるパラメータである.

胸腔鏡画像から肺領域抽出を行う際に, 肺の形状が変形したり, 崩れたりしてしまうと正しい領域を抽出できなくなる. そのため肺画像のスタイル変換では肺表面のテクスチャのみを変換し, 変換後の画像と元の画像とで肺の形状が保持され

る必要がある. しかし CycleGAN の損失関数は, 肺の形状を保持したままの変換を行うには制約として不十分である. そのため本研究では Nakao らによる金属アーチファクト低減の研究 [14] において提案された Feature loss を式 (3) の CycleGAN の損失関数に加える.

深層特徴の抽出には ResNet [15] を用い, 画像 $x, y, G_X(y), G_Y(x), G_X(G_Y(x)), G_Y(G_X(y))$ のそれぞれの特徴を抽出した. Feature loss は, これらの抽出した各画像の特徴の差分をとったものの総和である. f を ResNet による画像の特徴抽出の関数として, Feature loss を式 (4) のように定義する.

$$\begin{aligned} \mathcal{L}_{\text{feat}} = & \mathbb{E}_x[\|f(G_Y(x)) - f(x)\|_2] \\ & + \mathbb{E}_y[\|f(G_X(y)) - f(y)\|_2] \\ & + \mathbb{E}_x[\|f(x) - f(G_X(G_Y(x)))\|_2] \\ & + \mathbb{E}_y[\|f(y) - f(G_Y(G_X(y)))\|_2] \\ & + \mathbb{E}_x[\|f(G_X(G_Y(x))) - f(G_Y(x))\|_2] \\ & + \mathbb{E}_y[\|f(G_Y(G_X(y))) - f(G_X(y))\|_2]. \end{aligned} \quad (4)$$

Feature loss によって実画像と同じ特徴を保持した生成画像および再構成画像を生成するように学習を進めることができる.

以上を用いて本研究では以下の式 (5) を損失関数として用いる. ここで λ_{feat} は Feature loss の相対的な重みを決定するパラメータである.

$$\mathcal{L} = \mathcal{L}_{\text{cgan}} + \lambda_{\text{feat}} \times \mathcal{L}_{\text{feat}}. \quad (5)$$

喫煙者肺画像を非喫煙者肺画像に変換する学習済み画像変換モデル G_X^* は以下の最適化問題を解くことで得られる.

$$G_X^*, G_Y^* = \arg \min_{G_X, G_Y} \max_{D_X, D_Y} \mathcal{L}(G_X, G_Y, D_X, D_Y). \quad (6)$$

3. 実験

提案する教師なし画像変換を用いた肺領域抽出の性能を検証するために 2.1 節の胸腔鏡画像データを用いて実験を行った. まず U-net を用いた肺領域抽出において, 喫煙者肺画像と非喫煙者肺画像それぞれに対する抽出精度を確認する前実験を行った. 次に喫煙者肺画像を非喫煙者の見た目に変換するために, CycleGAN を用いた教師なし画像変換のモデルを学習した. 最後に喫煙者肺画像に対して肺領域抽出の前処理としてスタイル変換を施すことについて, 抽出精度への影響を確認する実験を行った.

3.1 肺領域抽出

U-net を用いた肺領域抽出モデルを学習し, 得られたモデルを用いて肺画像に対する肺領域ラベルを生成した. モデルの学習には 2.1 節のデータを用い, モデルの評価には 12 枚の胸腔鏡画像に対して出力された肺領域ラベルと真値となる肺領域ラベルのダイス係数を算出して用いた.

表 2 の 2 列目と 4 列目に求めたダイス係数を示す. 非喫煙者肺および喫煙者肺のダイス係数の平均値はそれぞれ 0.9793, 0.9002 となり, 喫煙者肺画像に対する肺領域抽出の精度が非

表2 非喫煙者、喫煙者および変換後喫煙者肺画像に対する肺領域抽出精度 (ダイス係数)

非喫煙者		喫煙者		
		Original	Translated	
Case 1a	0.9832	Case 4a	0.9095	0.9738
Case 1b	0.9937	Case 4b	0.7515	0.8402
Case 2a	0.9911	Case 5a	0.9164	0.8824
Case 2b	0.9550	Case 5b	0.9083	0.8470
Case 3a	0.9741	Case 6a	0.9800	0.9823
Case 3b	0.9785	Case 6b	0.9354	0.9054
Mean	0.9793	Mean	0.9002	0.9052

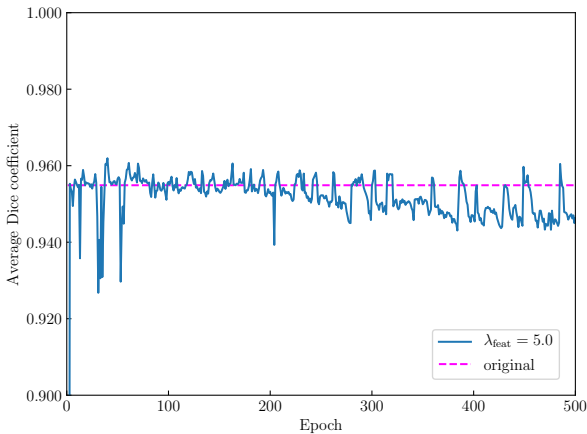


図3 画像変換モデルのエポックと肺領域抽出精度の関係、青：変換後画像の抽出精度、紫：変換前画像の抽出精度

喫煙者肺画像に対する精度よりも低いことが確認される。一方で Case 6a は喫煙者肺画像であるが、ダイス係数の値がいくつかの非喫煙者肺画像のものよりも高い。これは喫煙者肺画像の中には境界線が明瞭なものがあり、これによって肺と胸腔とを区別しやすくなっているためと考えられる。

3.2 教師なし画像変換と肺領域抽出

CycleGAN を用いて喫煙者肺画像を非喫煙者肺画像に変換する教師なし画像変換のモデルを学習した。学習によって得られたモデルを用いて喫煙者肺画像のスタイル変換を行い、変換後の画像を肺領域抽出モデル (U-net) に入力して変換後の画像に対する肺領域ラベルを生成した。出力された肺領域ラベルと真値となる肺領域ラベルのダイス係数を算出し、3.1 節の結果と比較することで CycleGAN による教師なし画像変換が肺領域抽出の精度向上に寄与するか否かを確認した。

CycleGAN の学習では X として 770 枚の非喫煙者肺画像を用い、 Y として 1120 枚の喫煙者肺画像を用いた。これらの肺画像は 2.1 節で述べた 25 例の手術動画のうちそれぞれ 1 例の動画から静止画像を抽出しデータ拡張を行い作成したものである。生成器 G_X と G_Y には共に U-net を用い、識別器 D_X と D_Y は CycleGAN と同じとした。損失関数 $\mathcal{L}$ のパラメータはいくつかのパラメータセットを試行し、 $\lambda_{cyc} = 10.0$ 、 $\lambda_{feat} = 5.0$ とした。

肺領域抽出の精度が最も高いモデルを用いて評価を行うた

め、1 エポック毎にモデルを保存し、それらのモデルを用いて喫煙者肺画像に対してスタイル変換を施した。変換後の画像に対して U-net による領域抽出を行い、真値となる肺領域ラベルとのダイス係数を算出した。ここで学習は 500 エポックまで行い、ダイス係数の算出に用いた画像は学習に用いた喫煙者肺画像のうちデータ拡張によって作成したデータを除いたものとした。

図3に画像変換モデルのエポックと肺領域抽出精度の関係を示す。図よりおおそ 50 エポックから 200 エポックの間のモデルでは、変換後の肺画像の平均ダイス係数が変換前の平均ダイス係数を上回っている。エポックが増加していくと徐々に平均ダイス係数は減少しているため、50 エポックから 200 エポックの間のモデルが形状を維持したスタイル変換が可能なモデルであると考えられる。これらのうち 163 エポック目のモデルがダイス係数の平均値が最大となっていたため、以降の実験ではこのモデルを用いる。

図4に 163 エポック目のモデルを用いて喫煙者肺画像をスタイル変換した結果と、変換後画像に対して肺領域抽出モデルにより抽出された肺領域画像を示す。図は各行にそれぞれ別の手術動画より抽出した喫煙者の静止画像の結果を示しており、1 列目の y は変換前の静止画像、2 列目は U-net を用いて y から抽出した肺領域結果である。3 列目の $G_X(y)$ は y を CycleGAN によって変換した画像であり、4 列目はこの画像より抽出した肺領域である。また 5 列目は肺領域画像の真値である。図より Case 4a では画像のスタイル変換によって抽出された肺領域が真値に近づいていることがわかる。一方で Case 5a では抽出された肺領域は真値よりも小さくなっている。また Case 6a では変換の前後で抽出結果はほとんど変わらなかった。

表2に静止画像に対する領域抽出の結果を示す。表は非喫煙者の結果と喫煙者の結果を分けて示しており、喫煙者の結果については教師なし画像変換を行う前後で結果を分けて示した。表の喫煙者の列において、Original が元の静止画像に対する抽出結果であり、Translated がスタイル変換を施した後に肺領域抽出を行った結果である。ダイス係数の平均値は Original よりも Translated の方が高くなったが、一方で有意差は見られなかった。また、6 症例中 3 症例でダイス係数が減少した。

3.3 考察

3.2 節の結果より、喫煙者肺画像に対して肺領域抽出の前に教師なし画像変換を適用することで、一部の症例では領域抽出精度が向上されることを確認した。具体的に、今回の結果では Case 4a, 4b, 6a においてダイス係数が増加した。ここで Case 6a の肺画像は CycleGAN の学習に用いた手術動画と同じ動画から学習に用いていないシーンを抽出したものである。これらの症例ではスタイル変換によって喫煙者肺画像が非喫煙者肺画像にうまく変換され、肺領域抽出の精度が向上したと考えられる。Case 6a に関しては CycleGAN の学習に用いたデータと同じ肺画像であるという点で学習データとの類似性が認められるため、変換が正しく行われたと考えられる。Case 4a と 4b に関しては学習データに含まれる肺画像と肺表面の色が似ており、妥当な変換が達成されたことが要因と考えられる。

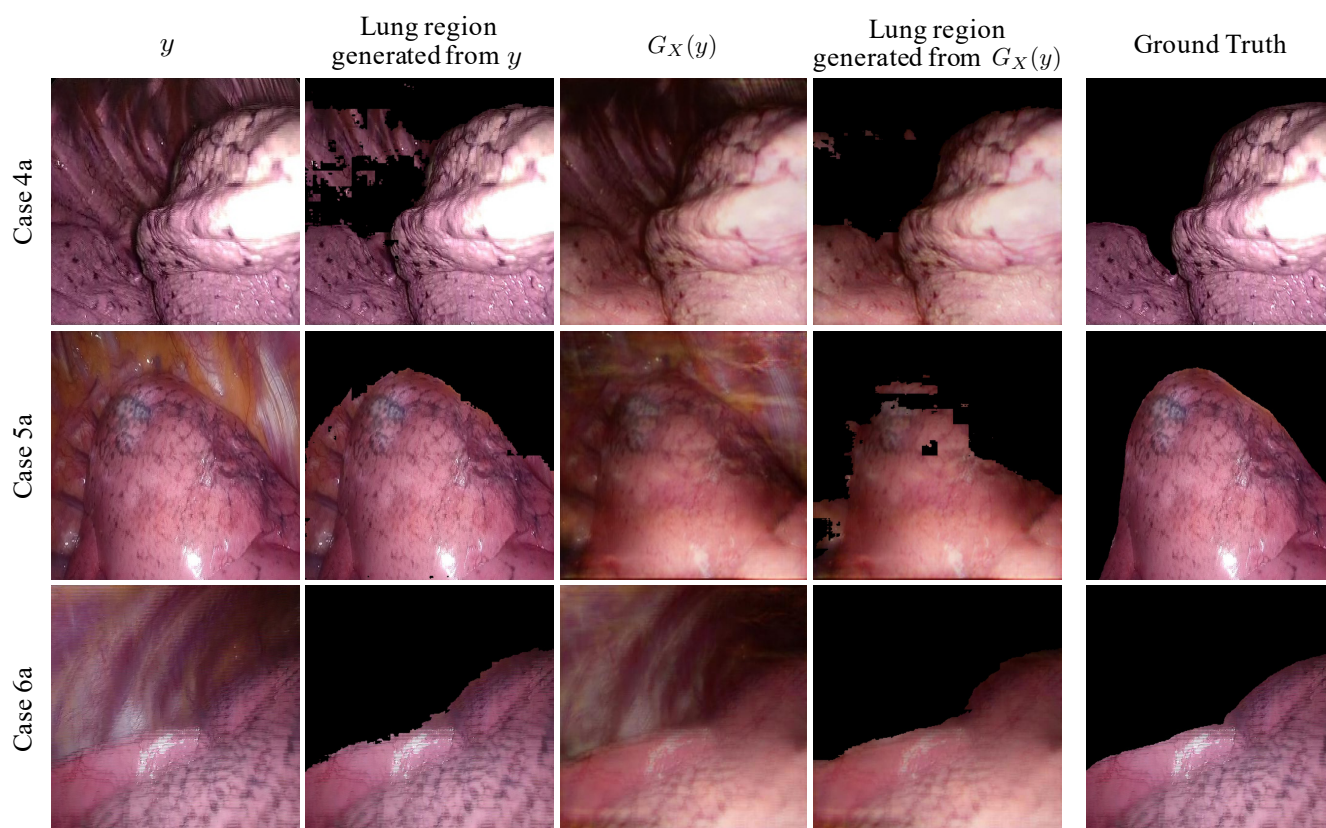


図4 喫煙者肺画像3例の画像変換による肺領域抽出結果. 1列目: 喫煙者肺の元画像, 2列目: 元画像の抽出結果, 3列目: 変換後画像, 4列目: 変換後画像の抽出結果, 5列目: 肺領域画像の真値

一方で Case 5a, 5b, 6b においてはダイス係数が減少した. これらの症例では CycleGAN によるスタイル変換で, 肺の形状を保持しない変換や肺と胸腔との境界線が不明瞭になる変換になったことが抽出精度低下の原因だと考えられる. 例として図4の Case 5a の $G_X(y)$ が挙げられ, ここでは変換前の画像と比べて肺と胸腔との境界線が不明瞭になっている. Case 5a と 5b において変換が不十分であったのは, 肺の形状や表面の色等が学習データと異なっているためと考えられる. また Case 6b は学習データと同じ手術動画から抽出された肺画像であるが, 大きさや角度が異なるシーンであり, 抽出精度の向上に寄与する変換とならなかった.

以上の結果より肺の形状や胸腔との境界線を維持した変換が可能なモデルを学習することで, 肺領域抽出の精度を向上させることが示された. 一方で多様な肺画像に対して適切な変換を行うモデルを学習することには課題が残されている.

4. おわりに

本稿では深層学習を用いた肺領域抽出法と敵対的生成ネットワークを用いた教師なし画像変換の手法を紹介し, その性能評価結果について報告した. 胸腔鏡下肺がん切除術の手術動画より抽出した静止画像群を用いて肺領域抽出モデルおよび教師なし画像変換のモデルを学習した. 結果より U-net を用いた肺領域抽出法は高い精度で胸腔鏡画像中の肺領域を抽出できることが確認された.

また喫煙者肺画像に対するダイス係数が非喫煙者肺画像に対するものよりも小さい問題に対して, 一部の肺画像においては教師なし画像変換によってダイス係数が増大することが確認された. しかし画像変換モデルは不完全なものであり, 一部の肺画像においてはダイス係数を減少させる結果となった. 今後は多様な肺画像に対する画像変換を可能とするモデルを構築するとともに, 変換の頑健性を向上させることで肺領域抽出の精度向上を目指していきたい.

謝辞

本研究は, AMED 産学連携医療イノベーション創出プログラム (ACT-M) 「脱気変形に対応した微小結節の術中同定法」及び挑戦的研究 (萌芽) (課題番号: 18K19918) の支援による.

文献

- [1] M. Sato, M. Omasa, F. Chen, T. Sato, M. Sonobe, T. Bando, and H. Date, "Use of virtual assisted lung mapping (val-map), a bronchoscopic multispot dye-marking technique using virtual images, for precise navigation of thoracoscopic sublobar lung resection," The Journal of thoracic and cardiovascular surgery, vol.147, no.6, pp.1813–1819, 2014.
- [2] B. Koo, E. Özgür, B. Le Roy, E. Buc, and A. Bartoli, "Deformable registration of a preoperative 3D liver volume to a laparoscopy image using contour and shading cues," International Conference on Medical Image Computing and Computer-Assisted Intervention, pp.326–334, 2017.
- [3] N. Haouchine, S. Cotin, I. Peterlik, J. Dequidt, M.S. Lopez, E. Kerrien,

- and M.-O. Berger, "Impact of soft tissue heterogeneity on augmented reality for liver surgery," *IEEE transactions on visualization and computer graphics*, vol.21, no.5, pp.584–597, 2014.
- [4] M. Nakao, J. Tokuno, T.F. Chen-Yoshikawa, H. Date, and T. Matsuda, "Surface deformation analysis of collapsed lungs using model-based shape matching," *Int. J. Computer Assisted Radiology and Surgery*, vol.14, no.10, pp.1763–1774, 2019.
 - [5] M. Nakao, A. Saito, and T. Matsuda, "A simulation study on deformation estimation of elastic materials using monocular images," *Int. J. Computer Assisted Radiology and Surgery*, 12(1), S257–258, 2017.
 - [6] S. Wu, M. Nakao, and T. Matsuda, "Continuous lung region segmentation from endoscopic images for intra-operative navigation," *Computers in Biology and Medicine*, vol.87, no.1, pp.200–210, 2017.
 - [7] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," *International Conference on Medical image computing and computer-assisted intervention*, pp.234–241, 2015.
 - [8] 今西 勁峰, 武 淑瓊, 中尾 恵, 松田 哲也, 深層学習を用いた術野映像内肺領域抽出, 第 38 回日本医用画像工学会大会, OP2-12, 2019.
 - [9] L.A. Gatys, A.S. Ecker, and M. Bethge, "Image style transfer using convolutional neural networks," *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp.2414–2423, 2016.
 - [10] J. Johnson, A. Alahi, and L. Fei-Fei, "Perceptual losses for real-time style transfer and super-resolution," *European conference on computer vision*, pp.694–711, 2016.
 - [11] Y. Li, C. Fang, J. Yang, Z. Wang, X. Lu, and M.-H. Yang, "Universal style transfer via feature transforms," *Advances in neural information processing systems*, pp.386–396, 2017.
 - [12] J.-Y. Zhu, T. Park, P. Isola, and A.A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," *Proceedings of the IEEE international conference on computer vision*, pp.2223–2232, 2017.
 - [13] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," *Advances in neural information processing systems*, pp.2672–2680, 2014.
 - [14] M. Nakao, K. Imanishi, N. Ueda, Y. Imai, T. Kirita, and T. Matsuda, "Regularized three-dimensional generative adversarial nets for unsupervised metal artifact reduction in head and neck ct images," *IEEE Access*, vol.8, pp.109453–109465, 2020.
 - [15] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp.770–778, 2016.