

CT 画像間の変位場を用いた敵対的データ拡張の試み

A Preliminary Study on Adversarial Data Augmentation Methods Using Displacement Fields between CT Images

栗山 由也[†] 中尾 恵^{††} 中村光宏^{††}

Yuya Kuriyama Megumi Nakao Mitsuhiro Nakamura

1. はじめに

近年、機械学習や深層学習は目覚ましい発展を遂げており、あらゆる分野で利用されている。深層学習において、高い学習効果を得るためにはデータの質と量の確保が重要であり、特に画像認識タスクにおいては、認識対象のバリエーションを豊富に含むデータ群が望まれる。医用画像分野においても、Computed Tomography (CT) や Magnetic Resonance Imaging (MRI) を対象に、深層学習を用いた画像認識の研究は盛んに行われている [1]。しかし、医用画像分野では疾患例を対象とした画像認識タスクが求められる一方で、疾患によっては症例数が少なく、豊富なバリエーションを有するデータを集積することが困難である [2]。特に胃や脾臓など CT 画像上で輪郭が不鮮明な臓器の場合には医療従事者による手動での輪郭抽出、教師データ作成の際に個人差が生じやすく、信頼性の高い学習用データベースの構築が課題となっている。

大規模な教師データベースの準備が難しい場合に、既存のデータに対して何らかの加工や変換を行うことで新たなデータを生成するデータ拡張 (Data Augmentation) 手法が考えられている。簡便な方法に、元画像に対して回転、反転、クロップやアフィン変換、さらには色やコントラストの変換を適用する方法 [3] があり、画像同士を混合する Mixup [4] 等も挙げられる。しかしながら、複数の臓器が密接に存在する医用画像に対して単純にアフィン変換等を適用すると臓器の形状や位置関係が不自然になる場合があり、生成可能なバリエーションも限定的である。

非線形変換によって、より柔軟で自然な画像生成の実現を目的として、近年敵対的生成ネットワーク (Generative Adversarial Network, GAN) [5] のデータ拡張への応用が試みられている。特に GAN を顔画像生成に拡張した枠組みとして StyleGAN [6] が知られており、約 70000 枚の顔画像データベースを用いて有効性が示されている。より効率的に高品質な画像生成を可能とした StyleGAN2 [7] も報告されている。医用画像分野においては脳の 3 次元画像に StyleGAN を適用した 3D StyleGAN [8] や、臓器の臓器領域ラベルを追加的に受け付けることで、より自然な医用画像の生成を試みた MDGAN [9] などの報告例があるが、いずれも大規模に集積されたパブリックドメインの医用画像データベースを対象としている。しかし、これまでに挙げた枠組みではデータ拡張時に画素値が変化するため、CT 画像の定量性が維持されづらい。特に疾患を有する画像で症例数が十分でない場合には、不自然な構造が新たに生み出

される、解剖学的に不自然な形状になりやすい等、医用画像におけるデータ拡張には依然課題が残っている。

本研究では StyleGAN2 を CT 画像に応用し、深層学習における拡張データとして利用可能な疑似 CT 画像を生成することを目的とする。上で述べたように、自由度の高い非線形変換によるデータ拡張は、CT 値の定量性が維持されづらく、特に症例数が十分でない場合に有効とは言い難い。そこで、拡張対象を CT 画像そのものではなく、CT 画像間の変形を定義した変位場とし、変位場を再構成する潜在変数を抽出する。一部の潜在変数の値の変更によって新たに生成した疑似変位場を CT 画像に適用することによって変形 CT 画像を生成する枠組みを提案する。変位場はデータ拡張時の画像変換における制約の役割を果たすため、不自然な構造が生み出されることを防ぎながら、疾患を有する患者の臓器形状や配置のバリエーションを多数生成可能になると期待される。

提案する変位場に基づくデータ拡張方法の有効性を確認するために、脾臓に対する放射線治療を受けた患者 115 名から得られた CT 画像群 [10, 11] から疑似 CT 画像を生成するデータ拡張を行う。既存手法と提案手法によって作成された異なる拡張データを用いて学習された U-net [12, 13] による臓器領域の抽出性能を比較する。

2. 関連研究

2.1 StyleGAN

従来の GAN ではノイズ $z \in Z$ (Z : 潜在空間) がそのまま Generator に入力されるが、StyleGAN [6] では、画像の特徴量の一部の組み合わせが存在しないケースにおいて、ノイズ z が存在する潜在空間 Z は線形ではなく entanglement (もつれ) のある歪んだ空間となる。そこで StyleGAN では z を Mapping Network f という 8 層の多層パーセプトロンで構成されるネットワークの入力とし、別の潜在空間 W へと写像して潜在変数 $w \in W$ を獲得して Generator の入力とする。この Mapping Network f の学習を進めることにより、潜在変数 w は画像の特徴量のもつれが少ない状態で最適化されることが期待される。

StyleGAN2 [7] は StyleGAN を改良した枠組みである。まず、従来の StyleGAN では実際の統計量を用いて正規化を行っていたが、推定の統計量で正規化を行うことにより、不要なモードの顕在化を防ぎ、生成画像のノイズの出現を抑えることが可能となった。さらに、Generator に skip connection を導入することで、生成画像の特徴の一部が不自然になるという問題を低減している。潜在空間の知覚的な滑らかさを表す Perceptual Path Length をモデルの正則化に組み込むことにより、画像品質の向上を実現している。

[†] 京都大学大学院 情報学研究科 Graduate School of Informatics, Kyoto University

^{††} 京都大学大学院 医学研究科 Graduate School of Medicine, Kyoto University

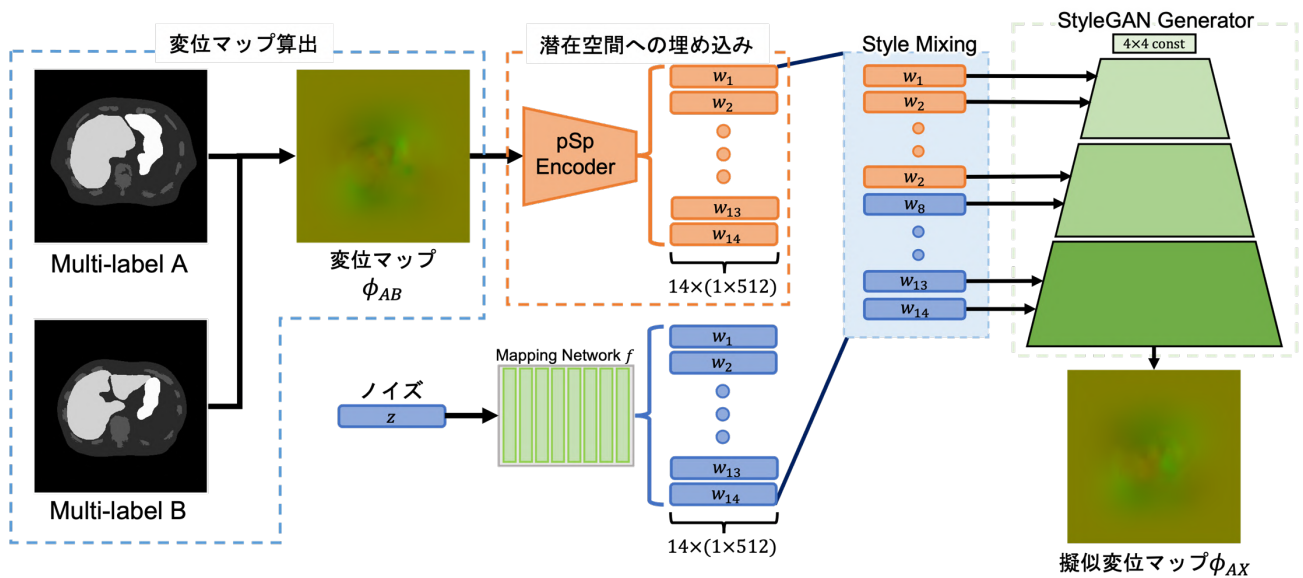


図 1 提案する疑似 CT 画像生成の枠組み

2 つのノイズパラメータ $z_1, z_2 \in Z$ 、および、それぞれに対応する潜在変数 $w_1, w_2 \in W$ をパラメータとして Generator に入力する際、ある解像度までは w_1 を入力し、それ以降の解像度では w_2 を入力することで、それぞれの画像の特徴を各解像度の層に局所化させることが可能になる。人間の顔画像においては、潜在変数の低解像度成分を操作すると顔の向きや髪型、顔の形、メガネの有無といった帯域的な違いを制御でき、中解像度成分を操作すると目、鼻、口等の小規模な顔の特徴が変化し、さらに高解像度成分を操作した場合には髪の色や背景の色のみを制御できることが示されている。この処理を Style Mixing と呼ぶ。

2.2 pix2Style2pix

StyleGAN では、ノイズ z から高品質な疑似画像が生成され、潜在変数を変更する Style Mixing によって生成画像を制御できるが、潜在変数と顔画像内の特徴との対応が事前には分からず、所望の特徴の制御を行うためにはノイズからランダムに生成された画像群の中から適した画像を手動で選び出す必要があった。そこで、現実画像を効率的に直接 StyleGAN の潜在空間へと埋め込み、全く同じような画像を生成させる潜在変数を算出することによって、容易に所望の特徴に対応する潜在変数を得る事を可能にしたネットワークが pix2Style2pix (pSp) [14] である。

pSp では、最初に訓練画像を Feature Pyramid Network と呼ばれる ResNet をベースとした特徴抽出ネットワークに入力し、StyleGAN Generator の低・中・高解像度それぞれに対応する特徴マップを取得する。次に、それぞれの特徴マップを map2style というストライド 2 の畳み込みネットワークに入力し、潜在変数を取得する。得られた潜在変数をあらかじめ用意された学習済みの StyleGAN Generator の対応する部分に入力し、画像を出力する。この出力画像と正解画像との差を最小化するように Encoder の学習を進める。ここで、StyleGAN Generator はあらかじめ用意された学習済みのネットワークを用い、パラメータの更新等は行われない。学習が完了した pSp Encoder は正解画像を生成

する潜在変数を入力画像から算出することができるエンコーダとなる。これにより、StyleGAN では課題であった、所望の画像特徴に対応する潜在変数の取得が容易に達成される。

3. 提案手法

図 1 に提案する疑似 CT 画像生成の枠組みを示す。本枠組みにおける大まかな流れとして、まず VoxelMorph [15] により、患者 A の臓器領域ラベルを別の患者 B の臓器領域ラベルへと変形させるような変位場 ϕ_{AB} を取得し、pSp Encoder を用いて StyleGAN の潜在空間へと埋め込む。この潜在変数の高解像度側の成分をランダムな値に入れ替え、学習済みの StyleGAN Generator へと入力することで、患者 A の CT スライス画像を全く別の存在しない患者 X の CT スライス画像へと変形させるような疑似変位場 ϕ_{AX} を出力する。得られた疑似変位場を患者 A の画像に適用することで、患者 X の CT スライス画像を疑似 CT 画像として得ることができる。次節以降において各処理の詳細を説明する。

3.1 変位場の算出

変位場は二つの画像 A, B 間で定義され、図 2 (a) のように画像 A 内の画素に対応する画像 B 内の画素への写像を示す変位ベクトル群として表現される。図 2 (b) は変位ベクトルの x, y 成分をそれぞれ R, G として変位場を表現した 2 チャネルのカラー画像であり、本研究ではこれを変位マップと呼ぶ。

変位場の算出には幾つかの方法が知られているが、本研究では、深層学習をベースとした医用画像間の位置合わせ方法として知られる VoxelMorph [15] を用いて変位場の算出を行う。ここで、腹部領域の CT 画像は多くの臓器が隣接し、かつ、全体的に低コントラストであるため、CT スライス画像間で正確な変位場を算出すること自体が困難である。そこで、本研究では VoxelMorph への入力画像には CT スライス画像ではなく、放射線腫瘍専門医によって事

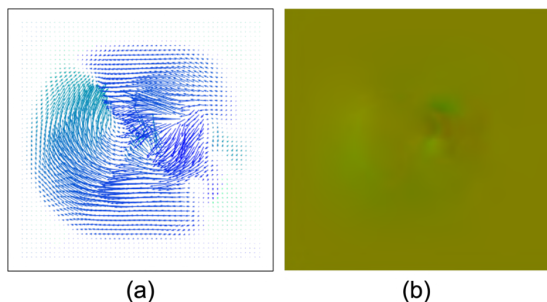


図 2 2次元変位場の例, (a) 変位ベクトルによる表現, (b) 変位ベクトルを RGB 値で表現した変位マップ

前に定義された臓器入力ラベルを用い、より正確な変位マップを得ることとした。なお、N 例の医用画像データを対象とする場合には、変位元と変位先の組み合わせ数である最大 $N \times (N - 1)$ の変位場を得ることができる。

3.2 変位場の拡張による疑似 CT 画像の生成

StyleGAN2 は多段にネットワークが構成され、 4×4 pixel の低解像度の画像から、最大 1024×1024 pixel までの高解像度の画像まで段階的に処理が行われる。本研究においても、大きく以下のようなステップで処理を行う。

- Step 1. 4×4 pixel の低解像度において、従来の GAN 構造に従って学習が安定化するまで十分に Generator と Discriminator の学習を行う。
- Step 2. Generator にアップサンプリングを行う nearest neighbor フィルタリングと、2つの逆畳み込み層を追加したネットワークで 8×8 pixel の解像度で学習が安定化するまで再度 Generator と Discriminator の学習を行う。
- Step 3. Step 1, 2 の処理を最高解像度まで段階的に行う。

ただし、本手法では得られた変位マップを訓練データとして StyleGAN2 の学習を行う。なお損失関数やパラメータは [7] における設定のまま用いている。StyleGAN2 によってノイズ z からランダムに疑似変位マップを生成可能となるが、そのままでは生成された疑似変位マップがどの患者の画像の変形に適用すべきものなのかは判断できない。そこで学習済みの StyleGAN2 に pSp Encoder を導入することにより、入力画像に基づく変位元画像の指定を可能にする。

本研究では、ある患者 A の CT スライス画像を変位させる変位マップを pSp Encoder によって StyleGAN Generator の潜在空間へと埋め込み、潜在変数を得る。ここで、潜在変数の数 n は入力/出力する画像の解像度 res に対応しており、 $n = 2 \times (\log_2(res) - 1)$ となる。例えば、今回扱う予定の 256×256 pixel の画像であれば $n = 14$ となる。

従来の StyleGAN で確認されている顔画像に対する Style Mixing の挙動を考慮すると、CT スライス画像に適用した場合、図 3 の様に得られた潜在変数の低解像度成分を操作すれば胴体の大きさや概形といった大域的な違いを、中解像度成分を操作すれば肝臓や胃といった内部の臓器の形状を、そして高解像度成分を操作すればさらに局所的な臓器の形状をコントロール可能になることが期待できる。pSp Encoder によって得られた 14 次元の潜在変数のうち、1~7

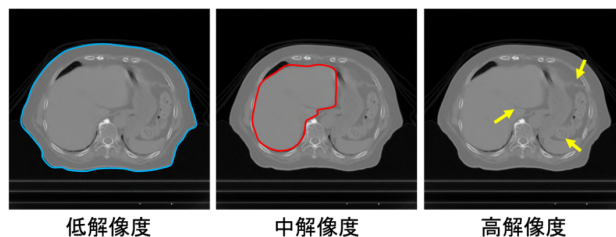


図 3 潜在変数の各成分を操作した際に影響を受ける画像特徴の違い

番目の潜在変数はそのまま用い、8~14 番目の潜在変数は乱数によって上書きする。すなわち、体の概形などの大域的な要素に対応すると考えられる潜在変数は残し、内部の臓器形状や細部の特徴、テクスチャ等に対応すると考えられる潜在変数をランダムに更新する。

このようにして得た潜在変数を StyleGAN の Generator の各解像度の層へと入力することにより、患者 A を実際には存在しない架空の患者 X の CT 画像へと変位させる疑似変位マップ ϕ_{AX} を生成することができる。最後にこの変位マップを患者 A の CT 画像に適用し、疑似 CT 画像を生成する。

4. 実験

提案方法の有効性を確認するために腹部領域の3次元CTデータを対象にデータ拡張を行い、疑似CTデータを作成する実験を行った。また、臓器領域抽出タスクを対象として疑似CTデータの有効性を確認する実験を行った。実験には Python 3.7 及び TensorFlow 2.4 によって実装された VoxelMorph, PyTorch 1.7 によって実装された StyleGAN2, pSp, U-Net [13] を用い、GPU は NVIDIA TITAN RTX を使用した。また、パラメータはそれぞれデフォルト設定のまま用いた。

4.1 データセット及び前処理

本実験では、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 115 名の患者の腹部 3D-CT データを用いた。本研究は 京都大学医学研究科 医の倫理委員会によって承認された (承認番号: R1446)。放射線腫瘍専門医によって胴体、肝臓、胃、腎臓の臓器領域ラベルが定義済み [10, 11] となっており、追加で実験者が手動で骨に相当する部位を抽出し、新たなラベルとして保存した。115 名のデータはデータ拡張の対象とする 90 名のデータと評価実験の対象とする 25 名のデータに分割した。また、評価対象を腹部領域の一定範囲内に限定するために、肝臓の臓器領域ラベルから重心位置を算出し、その重心位置の上下 5 スライスずつ、合計 11 スライスをデータ拡張の対象とした。図 4 は用いた CT 画像と臓器ラベルの一例である。

最初に VoxelMorph を用いて異なる患者の CT 画像間の 2 次元画像位置合わせを行い、変位場を取得した。データ数が過剰になることを避けつつ、できるだけ多様な変位場のバリエーションを獲得するために、各患者の 11 枚の CT スライス画像のうち、上から 1, 3, 5, 7, 9 枚目の 5 枚を学習対象とすることとした。結果として変位元の画像と変位先の

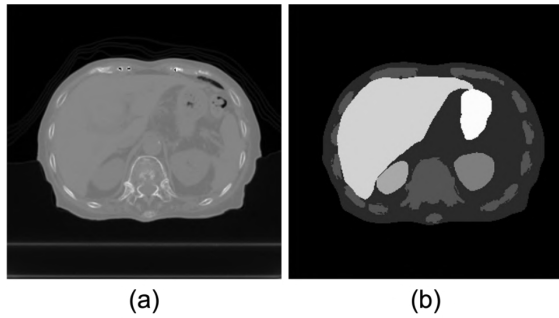


図 4 学習単位となるデータの例, (a) CT スライス画像, (b) 臓器領域ラベル

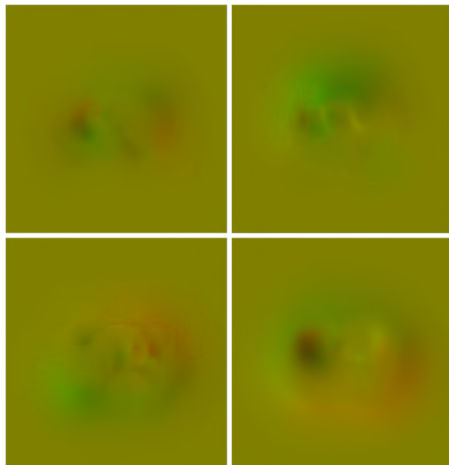


図 5 ノイズ z の値の変更によって生成した擬似変位マップの例

画像の組み合わせ数である $(90 \times 5) \times (89 \times 5) = 200,250$ 枚の変位マップを取得し、StyleGAN2 へ入力する学習データとして用いることとした。

4.2 学習及び擬似 CT スライス画像の生成

取得した 200,250 枚の変位マップを訓練データとして、バッチサイズを 32, ステップ数を 80,000 に設定して StyleGAN2 の学習を行った。なお、学習時間の短縮のため、変位場は 256×256 pixel にリサイズした。学習終了後の Generator によって乱数によって生成されたノイズ z から生成された 4 枚の擬似変位マップの例を図 5 に示す。異なる擬似変位マップが生成されているが、前述の通り、このままではどの患者の CT 画像の変形に適した変位場かは事前には分からない。

次に、学習時間の短縮のため、200,250 枚の変位マップのうち、ランダムな 990 枚の変位マップを訓練データとして、バッチサイズを 8, ステップ数を 50,000 に設定して pSp Encoder の学習を行った。ここで、StyleGAN Generator には上記で学習した Generator を用いた。入力した変位マップと同じ変位マップを出力できるように pSp Encoder の学習を進めた。図 6 に pSp Encoder に入力した変位マップと再構成された変位マップを示す。入力した変位マップに類似した変位マップが出力されており、潜在空間への埋め込みが正しく行われていることを確認した。

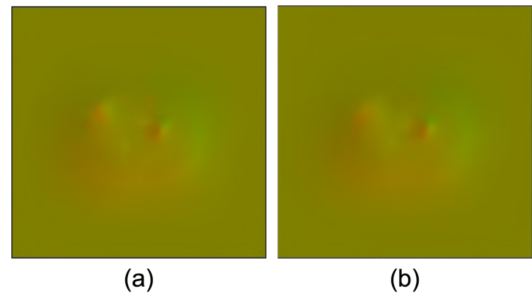


図 6 pSp Encoder による擬似変位場生成, (a) 実変位マップ, (b) 再構成された変位マップ

続いて、pSp Encoder により得られた潜在変数に対して Style Mixing を適用し、中間潜在変数の 8~14 行目を複数の異なるノイズ z から生成されたものと入れ替えることにより、さまざまなパターンの臓器形状を持つ擬似 CT 画像を生成できるか否かを確認した。また、比較として、標準的なデータ拡張方法として採用されてきたランダムな Affine 変換によって得られる疑似 CT 画像も生成した。

図 7 に Affine 変換によるデータ拡張と提案するデータ拡張によって得られた疑似 CT 画像をそれぞれ 2 例ずつ示す。Affine 変換による拡張 (図 7 (b)) では、胴体の形や傾きなどにバリエーションが見られるが、線形的かつ変化量を抑えた変形であるため、内部の臓器形状の細かな違いや臓器間の位置関係の変化は見られない。それに対し、提案手法による拡張 (図 7 (c)) では、胴体の概形の変化は微小でありながら、肝臓や胃が様々な形状へと変化している事が分かる。また、不自然な臓器の変形や画素値の変化も見られなかった。

4.3 提案手法の臓器領域抽出における有効性

次に、提案するデータ拡張方法の有効性を評価するために、データ拡張の違いによる U-net の臓器領域の抽出性能を比較する実験を行った。比較する学習条件は次の 3 通りとした。

拡張なし : データ拡張を行わず、90 例 11 スライスからなる計 990 枚の CT 画像を用いて学習

Affine 変換 : 990 枚の CT 画像に簡単なランダムな Affine 変換を適用した $990 + 990 \times 10 = 10,890$ 枚を用いて学習。なお、0.9~1.1 倍のスケール、xy 方向にそれぞれ最大 ± 16 ピクセルの並行移動、 ± 5 度の回転をランダムに適用することとした。

提案手法 : 990 枚の CT 画像に提案手法によってデータ拡張を行い、10,890 枚を用いて学習

U-net はいずれの条件においても胴体、肝臓、胃の各単一領域を抽出するタスクとして学習を行い、25 例 11 スライスからなる計 275 枚をテストデータとして Dice 係数と IoU の平均値を算出した。

表 1 に各手法を用いて実験を行った時の算出結果を示す。肝臓と胃の領域抽出において、どちらも提案手法によって拡張されたデータセットを用いて学習した場合が最も良い性能を示した。Affine 変換を用いた場合と比較して、提案手法を用いた場合には肝臓の Dice 係数は 0.0009, IoU は 0.0015 上昇しているのに対して、胃の Dice 係数は 0.0167, IoU は 0.0241 上昇しており、提案手法が特に胃のような複雑な臓器に対して有効であることが分かった。

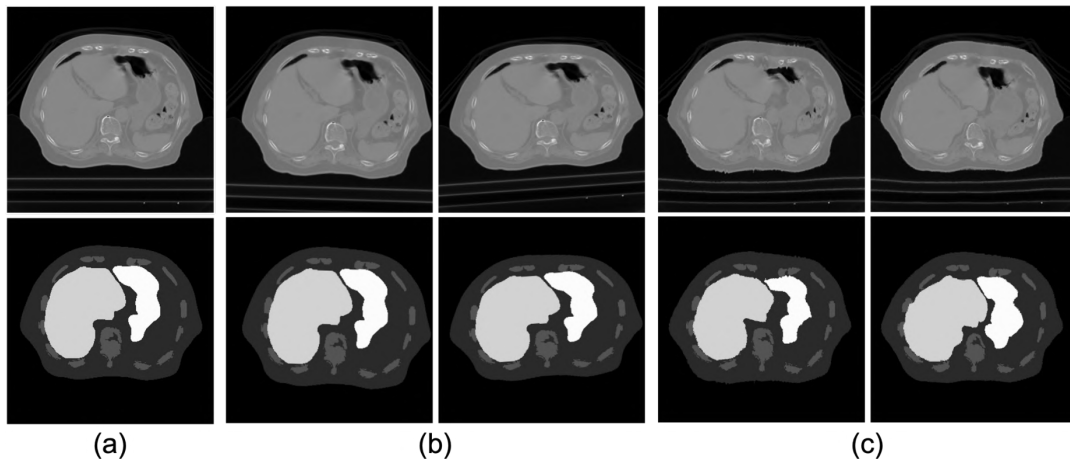


図 7 疑似 CT 画像と対応するラベル画像の例, (a) 元画像, (b) Affine 変換による拡張 2 例, (c) 提案手法による拡張 2 例

表 1 異なる拡張データセットによって学習された U-net の臓器領域抽出性能の比較

		拡張なし	Affine変換	提案手法
胴体	Dice係数	0.9991	0.9992	0.9990
	IoU	0.9983	0.9984	0.9981
肝臓	Dice係数	0.9505	0.9576	0.9585
	IoU	0.9080	0.9208	0.9223
胃	Dice係数	0.8244	0.8481	0.8648
	IoU	0.7380	0.7727	0.7968

図 8 にそれぞれの拡張データセットによって学習された U-net による臓器領域抽出結果の例を示す。対象臓器の検出に成功している領域を白、過検出域を青、未検出の領域を赤で示している。図 8 上段の例では、Affine 変換を用いた場合には肝臓、胃の両臓器において過検出領域が見られるが、提案手法を用いた場合には大きく改善された。中段の例では、特に胃に関して、未検出領域が小さくなっている事が確認できた。

一方、図 8 の下段に示すように、胃が存在していない CT スライス画像において、大きく誤識別する結果が見られた。これは訓練データに胃が存在しないデータの数が極端に少ないためと考えられる。そのような数の少ないパターンの画像に対して偏りを持たせてデータ拡張を施した場合や、Single-label ではなく Multi-label で他の臓器の位置も考慮して学習を行った場合に関して再度実験を行い検証する必要がある。また、今回は潜在変数の半分を編集して疑似変位場の生成を行ったが、よりさまざまな編集パターンを試し、検証を行う必要がある。今回は学習時間の短縮のため、本来の 512×512 pixel よりもダウンサイズして 256×256 pixel の画像を用いて実験を行ったが、最終的には 512×512 pixel のサイズで実験を行うことができればより精度が高まる事が期待できる。

5. おわりに

本研究では、CT 画像間に定義される変位場を StyleGAN2 と pSp Encoder によって学習することによる疑似 CT 画像生成方法を提案した。実験によって、疑似変位場を用いたデータ拡張が臓器領域抽出タスクにおいて有効であること、特に U-net による高い抽出精度が得られにくい胃の領域抽出タスクにおいて提案するデータ拡張が有効であることを確認した。

謝辞

本研究は 日本学術振興会 科研費補助金 基盤研究(B) (課題番号: 22H03021 及び 19H04484) の助成による。

参考文献

- [1] D. Shen, G. Wu, H.I. Suk, Deep learning in medical image analysis. Annual review of biomedical engineering, 19, pp. 221–248, 2017.
- [2] 大矢 友貴, 中尾 恵, 松田 哲也, 生成型バーチャル学習を用いた高遮蔽カメラ画像からの形状再構成, 情報処理学会 コンピュータビジョンとイメージメディア研究会(CVIM), 2022-CVIM-229, No. 10, pp. 1-8, 2022.
- [3] C. Shorten, T.M. Khoshgoftaar, "A survey on image data augmentation for deep learning," J Big Data 6, 60, 2019.
- [4] H. Zhang, M. Cisse, Y.N. Dauphin, D. Lopez-Paz. "mixup: Beyond empirical risk minimization." The International Conference on Learning Representations (ICLR), 2018.
- [5] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio. Generative adversarial networks. Neural Information Processing Systems (NeurIPS), 2014.
- [6] T. Karras, S. Laine, T. Aila, "A style-based generator architecture for generative adversarial networks," In Proc. The IEEE / CVF Computer Vision and Pattern Recognition Conference (CVPR), 2018.
- [7] Tero Karras, Samuli Laine, Miika Aittala Janne Hellsten, Jaakko Lehtinen, and Timo Aila. Analyzing and improving the image quality of stylegan. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 8110–8119, 2020.
- [8] S. Hong, 3D-StyleGAN: A style-based generative adversarial network for generative modeling of three-dimensional medical images. In: Engelhardt S. et al. (eds) Deep Generative Models, and Data Augmentation, Labelling, and Imperfections. DGM4MICCAI 2021, DALI 2021. Lecture Notes in Computer Science, Vol. 13003. Springer, Cham, 2021.
- [9] L. Siyu, D. Jason, E. Craig, G. Peter, C. Stuart, C. Shekhar,

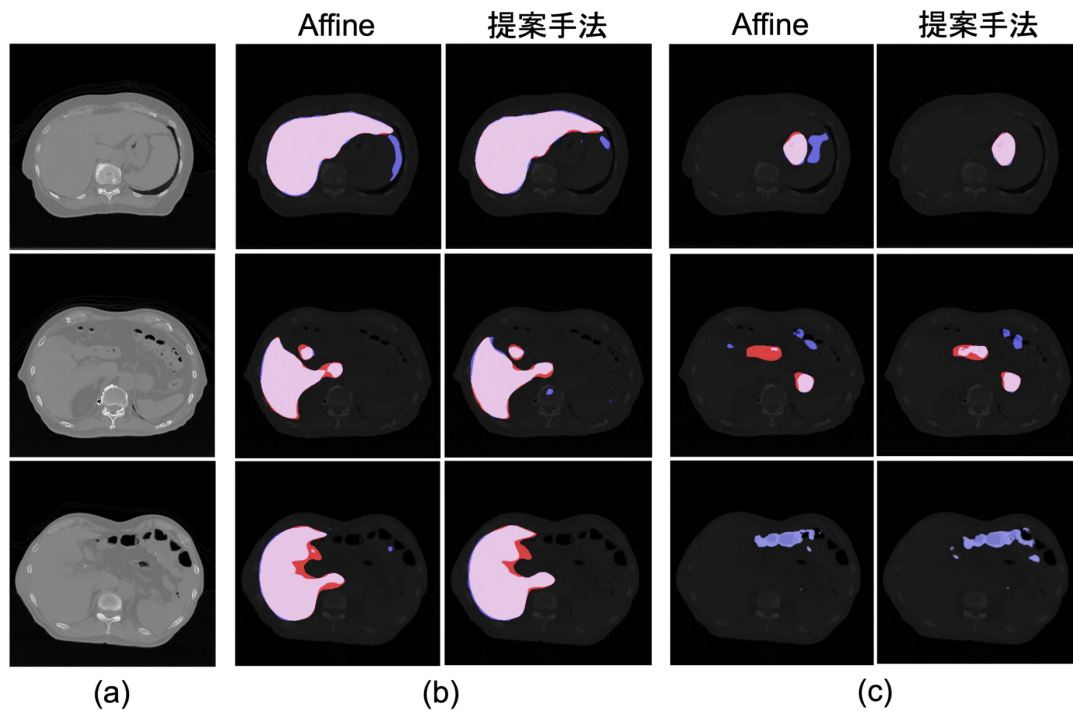


図 8 異なる拡張データセットによって学習された U-net による臓器領域抽出結果の例, 白 : 検出に成功した領域, 青 : 過検出領域, 赤 : 未検出領域, (a) 元画像, (b) 肝臓領域の抽出結果, (c) 胃領域の抽出結果

Manipulating medical image translation with manifold disentanglement, 2020.

- [10] 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, 101829, 2021.
- [11] M. Nakao, M. Nakamura, T. Matsuda, "Image-to-Graph Convolutional Network for Deformable Shape Reconstruction from a Single Projection Image", *International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI)*, pp. 259-268, 2021.
- [12] O. Ronneberger, P. Fischer, T. Brox, "U-net: Convolutional networks for biomedical image segmentation," In *International Conference on Medical image computing and computer-assisted intervention*, Springer, Cham, pp. 234-241, 2015.
- [13] 4uiiurz1, pytorch-nested-unet, <https://github.com/4uiiurz1/pytorch-nested-unet>, 2020.
- [14] T. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, S. Belongie. Feature pyramid networks for object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2117-2125, 2017.
- [15] G. Balakrishnan, A. Zhao, M.R. Sabuncu, J. Guttag, A.V. Dalca, "Voxelmorph: a learning framework for deformable medical image registration," *IEEE transactions on medical imaging*, 38(8), 1788-1800, 2019.