

CT 画像間の変位場を用いた Style Mixing の評価と解析

栗山 由也¹ 中村 光宏² 中尾 恵²

概要: 医用画像におけるデータ拡張に敵対的学習を利用することが考えられているが、極端にデータの少ない医用画像群に適用すると解剖学的に不自然な構造や不正確な CT 値を持つ画像が生成される可能性がある。本研究は CT 画像間で生成可能な変位場をデータ拡張の対象として活用し、画像の特徴を階層的に制御可能な StyleGAN2 及びそれに対応したエンコーダである pixel2Style2pixel を用い、一部の潜在変数を変更する Style Mixing によって擬似変位場を生成してデータ拡張を行う枠組みを開発した。提案手法を用いて生成された擬似 CT 画像の評価と、操作する潜在変数との関係性の解析を行った。従来手法と比較して高品質な擬似 CT 画像の生成、及び画像特徴のコントロールが可能であることが確認したので報告する。

キーワード: データ拡張, 敵対的学習, 変位場

Evaluation and analysis on Style Mixing using displacement fields between CT images

YUYA KURIYAMA¹ MITSUHIRO NAKAMURA² MEGUMI NAKAO²

Abstract: Adversarial learning has been considered for data augmentation in medical images, but when applied to extremely small sets of medical images, it may produce images with anatomically unnatural structures or inaccurate CT values. This study utilizes the displacement field that can be generated between CT images as a target for data expansion, and develops a Style Mixing framework for data expansion by generating a pseudo-displacement field by changing some latent variables using StyleGAN2 and its corresponding encoder pixel2Style2pixel, which can control image features in a hierarchical manner. We evaluated the pseudo-CT images generated by the proposed method and analyzed the relationship between the latent variables to be manipulated. We report on the results, which show that the proposed method is able to generate high-quality pseudo-CT images and control image features in comparison with conventional methods.

Keywords: Data Augmentation, Adversarial Learning, Displacement Field

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]。特に胃や膵臓など 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。

¹ 京都大学大学院情報学研究科
Graduate School of Informatics, Kyoto University
² 京都大学大学院医学研究科
Graduate School of Medicine, Kyoto University

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] や一方の画像のオブジェクトのインスタンスを別の画像にコピーする Copy-Paste [5] 等の方法が知られている。しかし、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、人体が有する自然な解剖学的特徴から逸脱した臓器形状や位置関係を示すデータが生成されることが多く、また生成可能なバリエーションも限定的である。

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) [6] による非線形変換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 [7] が知られており、約 70,000 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 [8] 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 [9] や、臓器の領域ラベルを追加的に受け付けることで、より自然かつ所望の特徴を持つ医用画像の生成を試みた MDGAN [10] などの報告例があるが、いずれも大規模に集積されたデータベースを訓練対象としている。さらに、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、深層学習におけるデータ拡張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。前述したように、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、StyleGAN 用のエンコーダである pix2Style2pix (pSp) Encoder [11] を用いて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り新たに生成した疑似変位場を CT 画像に適用して空間変換を行うことによって、変形 CT 画像を生成する Style Mixing の枠組みを提案する。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。

本研究では、膵がん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 [12], [13] を用いて疑似 CT 画像を生成する。我々は [14] において提案する変位場の Style Mixing によるデータ拡張が肝臓と胃の領域抽出精度の向上に有効であることを確認した。本稿では、提案方法によって生成された疑似 CT 画像と従来の枠組みから生成された疑似 CT 画像の比較評価を行う。さらに、Style Mixing の対象とする潜在変数と、生成される疑似 CT 画像の関係性を調査するための解析を行ったので報告する。

2. 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 StyleGAN2 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 画像を実在しない患者 X の CT 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 CT 画像に適用することで、新たな患者 X の CT 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。

2.1 変位マップ算出

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称する。

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 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、全体的に非常に低コントラストであるため、CT 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 画像ではなく、放射線腫瘍専門医によって事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 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 $N \times (N - 1)$ の変位場を取得することができ、元のデータ数が少ない場合でも、十分な数のデータを得ることができるというメリットがある。

2.2 潜在空間への埋め込み

pSp Encoder を用いて、得られた変位マップ群を Style-

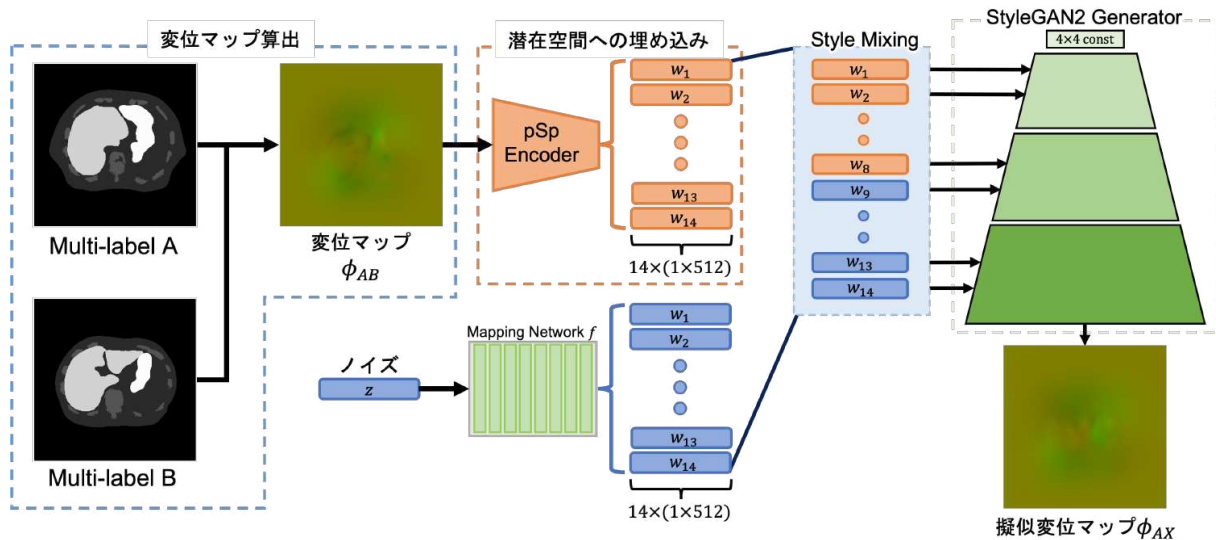


図 1: 提案する擬似 CT 画像生成の枠組み

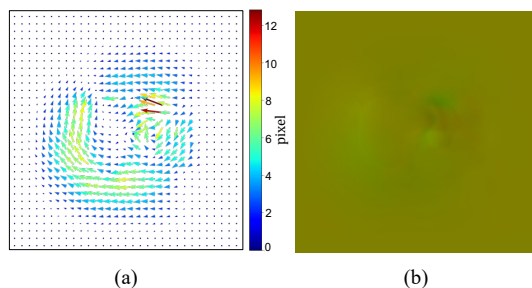


図 2: 2次元変位場の例, (a) 変位ベクトルによる表現, (b) 変位ベクトルを RG 値で表現した変位マップ

GAN2 Generator の潜在空間へと埋め込み, 潜在変数を取得する. なお, pSp Encoder の学習には学習済みの StyleGAN2 Generator が必要であるため, まず StyleGAN2 の処理の説明を行う.

StyleGAN2 は多段にネットワークが構成され, 4×4 pixel の解像度から, 最大 1024×1024 pixel の高解像度の画像まで段階的に学習が行われる. 本研究においても, 大きく以下のようなステップで学習を行う.

Step 1. 4×4 pixel の低解像度において, 従来の GAN の学習と同様に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.

本手法では得られた変位マップ群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 [8] における設定のまま用いる. 学習が完了すると, StyleGAN2 によってノイズ z からランダムに擬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擬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.

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量抽出ネットワークに入力し, StyleGAN2 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 StyleGAN2 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, StyleGAN2 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.

以上の処理に基づいて, 本研究では 2.1 節の処理によって得られた変位マップ群を 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.

2.3 Style Mixing

StyleGAN2 では, 2つのノイズパラメータ $z_1, z_2 \in Z$, 及び, 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。例えば、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大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。

Style Mixing を CT 画像の変位マップに対して適用した場合、図 3 のような挙動を示すと考えられる。変位マップは、画像としての特性上、人の顔画像等と比較すると複雑な特徴や、見た目のバリエーションが少ない。そのため、pSp Encoder から得られた変位マップの潜在変数の低解像度成分を操作しても大きな変化は現れないと考えられる。一方で、中解像度成分を操作した場合には胴体や内部臓器を大きく変化させ、高解像度成分を操作した場合には各臓器の特に細かい形状の変化を操作可能になることが期待できる。例えば、1 層から 11 層までの潜在変数はそのまま、12 層から 14 層までの潜在変数は、ノイズ z から StyleGAN2 の Mapping Network f を通して得られた潜在変数によって上書きする。これにより、胴体の形状はそのまま、内部臓器の形状を少しだけ変化させるような擬似変位マップを生成する潜在変数を取得できると考えられる。ただし、本研究が対象とする CT 画像間の変位場の Style Mixing において、どの層の潜在変数を変更すると自然な疑似 CT 画像の生成が可能になるかは自明ではない。潜在変数の各成分と生成される変位場及び疑似 CT 画像への影響については実験によって解析を行う。

2.4 擬似変位マップ及び擬似 CT 画像の生成

Style Mixing によって得られた潜在変数を StyleGAN2 Generator の対応する各解像度の層へと入力することにより、任意の患者 A の CT 画像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擬似変位マップ 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 画像へと適用し、患者 X の疑似 CT 画像を生成する。なお、CT 画像の変形には VoxelMorph の SpatialTransformer を用いる。

3. 実験

本研究では、京都大学医学部附属附属病院 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 データを対象に疑似 CT 画像を生成する実験を行なった。3D-CT データの収集及び本実験における利用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)。提案手法を用いて生成さ

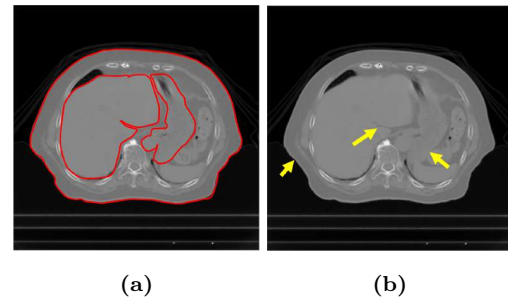


図 3: 潜在変数の各成分を操作した際に影響を受ける画像特徴の違い、(a) 中解像度成分を操作した場合、(b) 高解像度成分を操作した場合

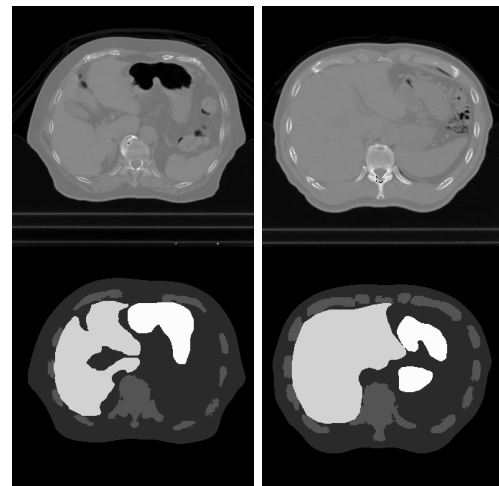


図 4: 学習単位となる実 CT 画像及びその臓器領域ラベルの例

れた疑似 CT 画像の解析を行い、従来法との比較を行った。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 を用い、GPU は NVIDIA TITAN RTX を使用した。また、各パラメータはそれぞれデフォルト設定のまま用いた。

3.1 前処理と学習の流れ

今回用いる 3D-CT データは放射線腫瘍専門医によって胴体、肝臓、胃、腎臓等の臓器領域ラベルが定義済み [12], [13] 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 ずつ、合計 11 スライスデータをデータ拡張の対象とした。図 4 に用いた CT 画像と臓器領域ラベルの例を示す。

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

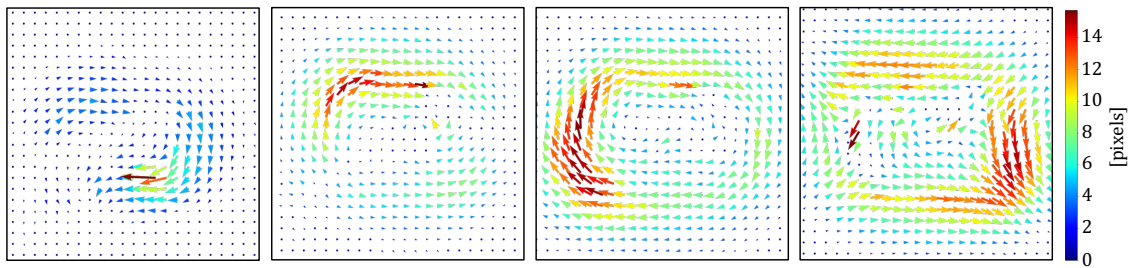


図 5: 学習終了後の StyleGAN2 Generator によってランダムに生成された変位場の例

CT 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た。

取得した 20,250 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 512×512 pixel から 256×256 pixel にリサイズした。学習終了後の Generator に乱数によって生成されたノイズ z を入力し、生成された擬似変位場 (ベクトル表現) を図 5 に示す。異なるバリエーションの擬似変位場を生成することに成功しているが、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を事前に判別することは不可能である。

次に、ランダムな 990 枚の実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、StyleGAN2 Generator には上記で学習した Generator を用いており、pSp Encoder の学習中には Generator のパラメータの更新は行われていない。入力した擬似変位マップと近い画像を生成する潜在変数を推定できるように pSp Encoder の学習を進めた。

3.2 実験 1: 擬似 CT 画像の比較

以下の従来手法を用いた場合と、提案手法を用いた場合とで、生成された擬似 CT 画像の比較を行った。擬似 CT 画像の生成は以下のような条件の元を行った。

StyleGAN2: 990 枚の実 CT 画像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 に設定して学習を行い、乱数 z から生成された潜在変数から生成された擬似 CT 画像。

StyleGAN2 + pSp: 990 枚の実 CT 画像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、上記で学習した StyleGAN2 の Generator を用いて入力した実 CT 画像と近い画像を生成する潜在変数を推定できるように pSp Encoder の学習を行った。その後、実 CT 画像の潜在変数を推定し、9 層から 14 層に対して Style Mixing

を行った潜在変数から生成された擬似 CT 画像。

提案手法: 提案手法において 9 層から 14 層に対して Style Mixing を行った潜在変数から生成された擬似変位場を実 CT 画像に適用し、生成された擬似 CT 画像。

これらの条件でそれぞれ 9,900 枚の擬似 CT 画像群を用意し、実 CT 画像群 990 枚との Fréchet Inception Distance (FID) [16] を算出した。ここで、FID は、 μ_A, Σ_A をそれぞれ画像群 A の Inception-v3 モデル埋め込み表現の平均と共分散として、 $FID = \|\mu_X - \mu_Y\|^2 + \text{Tr}(\Sigma_X + \Sigma_Y - 2\sqrt{\Sigma_X \Sigma_Y})$ で表される。2 つの画像群間の距離を表す指標であり、今回の場合は数値が小さければ小さいほど、擬似 CT 画像群の特徴が実 CT 画像群と近いことを示している。

算出結果を表 1 に、それぞれの手法によって生成された擬似 CT 画像例を図 6 に示す。提案手法を用いた場合の FID が最も小さく、従来手法を用いた場合には FID の値が大きくなった。従来手法では、全体的に不鮮明な擬似 CT 画像が生成され、図中の矢印部分に示すような片側のみの不自然な骨の位置・形状や胴体表面に穴が空いてしまっているようなもの、さらに StyleGAN2 の 1 列目や 3 列目の画像のように、本来左側にあるはずの肝臓らしき臓器が右側に見られるなど、異常な構造が多数生成されることが確認された。一方で、提案手法による擬似 CT 画像にはそのような異常は見られなかった。

表 1: 各手法によって生成された擬似 CT 画像群の FID 値

	StyleGAN2	StyleGAN2 + pSp	提案手法
FID	51.52	83.78	16.76

3.3 実験 2: 変位場を対象とした Style Mixing の解析

提案手法による変位場に対する Style Mixing において、操作対象の潜在変数の層と生成される擬似 CT 画像との関係性の解析を行った。操作する潜在変数の層の選び方は、連続する n 層 ($n = 1, 2, \dots, 14$) の全通りとした。図 7 に同一変位場の潜在変数の各層を操作した場合の、変位後の胴体、肝臓、胃の面積 [pixel] を各臓器領域ラベルから算出し、標準偏差を計算した結果を示す。

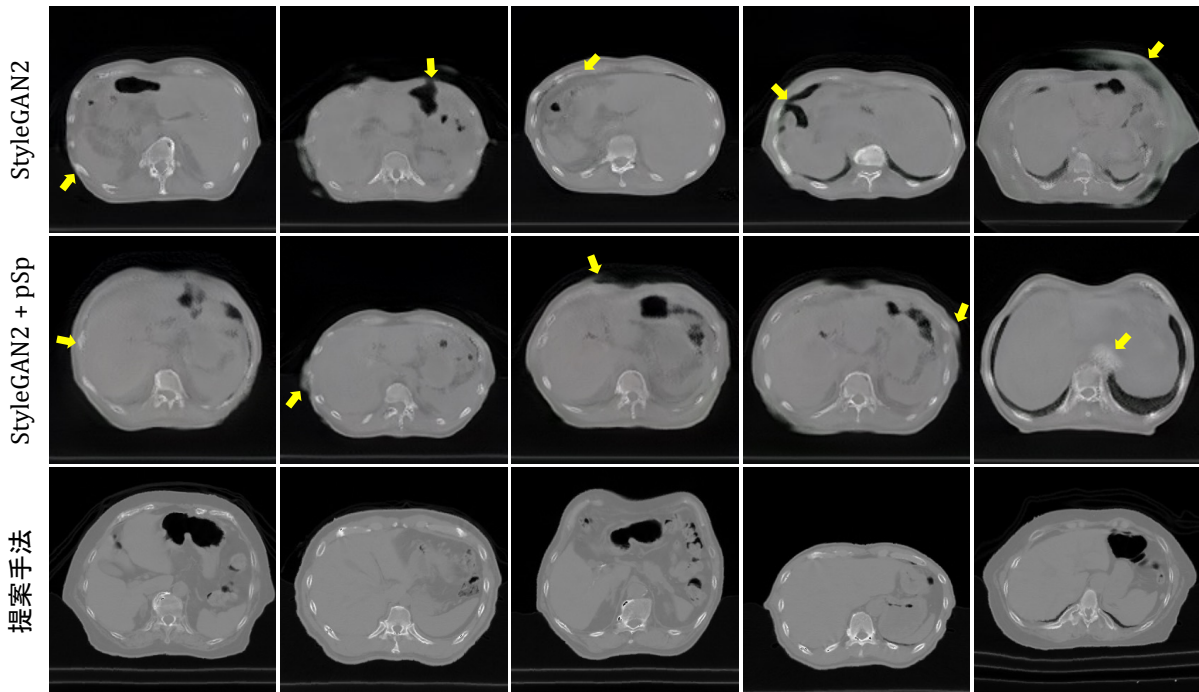


図 6: 各手法によって生成された疑似 CT 画像例

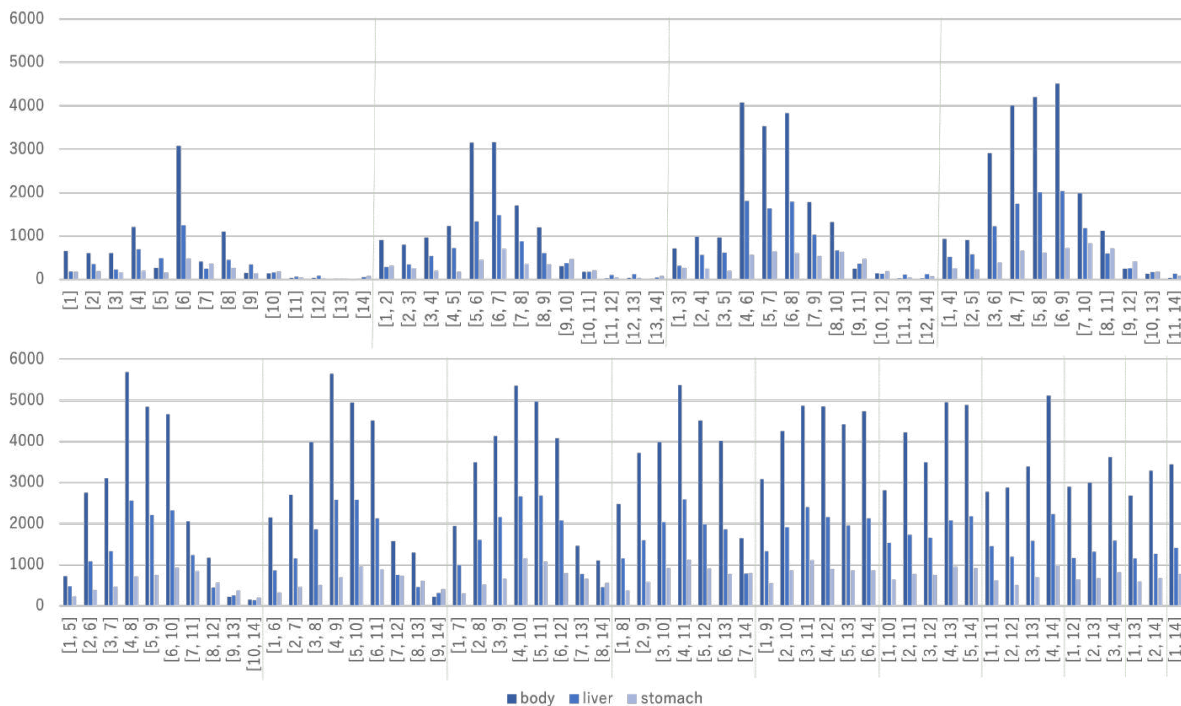


図 7: 各臓器の面積 [pixel] の標準偏差

標準偏差が大きければ大きいほど、その臓器の形状のバリエーションが豊富であると捉えることができる。中解像度成分を操作した時、どの臓器も変化が大きくなっており、特に第6層を含んだ時に顕著になっている。また、高解像度成分を操作した場合に注目すると、最も小さい臓器である胃の標準偏差が他の臓器よりも大きくなっている例が多く見られる。このことから、高解像度成分の Style Mixing

は、画像のより細かい部分に影響を与えるということを確認した。以上の潜在変数の階層と得られる疑似 CT 画像の関係を踏まえ、図 8 に潜在変数 z のうち [1,3] を低解像度帯、[4,7] を中低解像度帯、[8,11] を中高解像度帯、[12,14] を高解像度帯としてそれぞれを操作した場合に生成される疑似 CT 画像及びその臓器領域ラベルの例を示す。

低解像度帯の潜在変数を操作した場合には、比較的微小

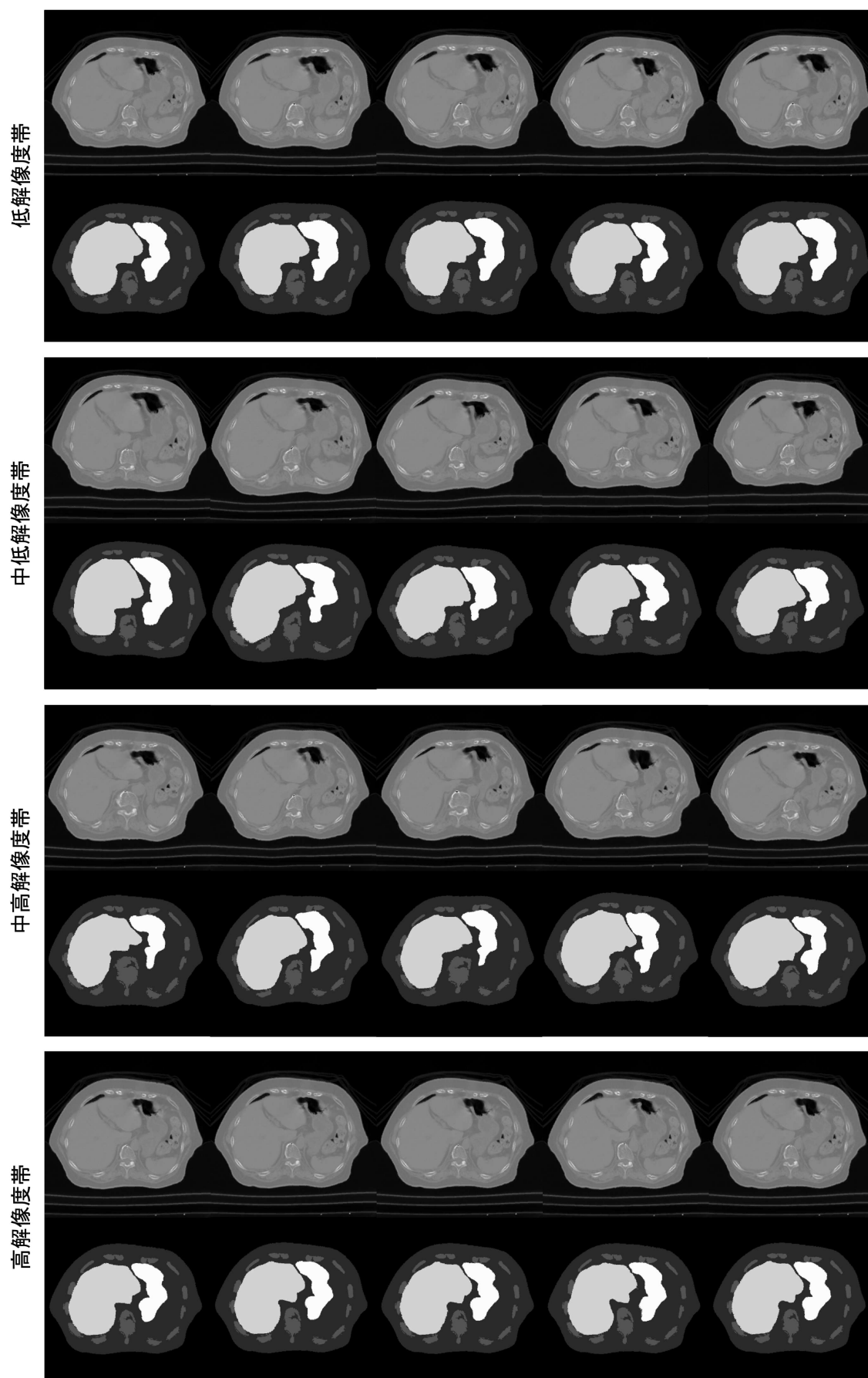


図 8: 潜在変数の各解像度帯を操作した場合の疑似 CT 画像及び臓器領域ラベルの例

な変形が画像全体に渡って及んでいることが確認される。中低解像度帯の潜在変数を操作した場合には胴体、肝臓、胃全てが大きく変形している。一方で、中高解像度帯の潜在変数を操作した場合には、胴体の変形が抑えられながら、内部臓器が様々に変形している。さらに高解像度帯の潜在変数を操作すると、胴体の概形はほとんど変化せずに、内部臓器の輪郭が少しずつ変形している様子が確認される。

4. おわりに

本研究では、CT 画像間で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データ数が十分でない医用画像に対しても適用可能なデータ拡張手法を提案し、生成された疑似 CT 画像の解析と評価を行った。実験の結果、従来手法のみを用いた場合よりも高品質な疑似 CT 画像が得られることを確認した。さらに、Style Mixing の操作対象となる潜在変数と疑似 CT 画像との関係性を解析し、形状特徴を異なる解像度で制御可能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] G. Ghiasi, Y. Cui, A. Srinivas, R. Qian, T. Lin, E. D. Cubuk, Q. V. Le, B. Zoph, simple copy-paste is a strong data augmentation method for instance segmentation, 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2917-2927, 2021.
- [6] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio. Generative adversarial networks. Neural Information Processing Systems (NeurIPS), 2014.
- [7] 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.
- [8] T. Karras, S. Laine, M. Aittala, J. Hellsten, J. Lehtinen, T. 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.
- [9] 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.
- [10] L. Siyu, D. Jason, E. Craig, G. Peter, C. Stuart, C. Shekhar, manipulating medical image translation with manifold disentanglement, 2020.
- [11] 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.
- [12] 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.
- [13] 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.
- [14] 栗山 由也, 中尾 恵, 中村光宏, CT 画像間の変位場を用いた敵対的データ拡張の試み, 第 21 回情報科学技術フォーラム (FIT), pp.535-540, 2022.
- [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.
- [16] M. Heusel, H. Ramsauer, T. Unterthiner, B. Nessler, and S. Hochreiter, GANs trained by a two time-scale update rule converge to a local Nash equilibrium. In Proc. NIPS, pages 6626–6637, 2017.