

令和 6 年度修士論文

三次元医療データ超解像への 3D-SRCGAN の応用

指導教員 椋木 雅之

宮崎大学大学院 工学研究科 工学専攻

機械・情報系コース 情報システム工学分野

T2303219 倉永将宏

概要

本論文では、本研究室で提案された三次元ボクセルデータ超解像のためのモデルである、3D-Super Resolution Conditional Generative Adversarial Networks (3D-SRCGAN) [1]を以下の3つのポイントを考慮し、ボリュームデータである三次元医療データへ応用できるか検証する。

1つ目は、損失関数である。損失関数は、深層学習において教師データと出力を近づけるための評価を行う関数であり、その値が小さいほど出力が正解に近いことを示す。ボクセルデータ超解像のための3D-SRCGANで、損失関数はMSEのみが採用されていた。本研究で扱うデータはボリュームデータであり、構造的な特徴が重要だと考えられる。そこで、MSE Loss、SSIM Loss、MSE Loss + SSIM Lossの3つの損失関数が超解像精度の向上に与える影響を評価する。

2つ目は、パッチングである。パッチングとは、データを部分に分割してその単位で超解像を行う手法である。パッチングは、画像の超解像時に一般的に使用されるが、ボクセルデータの超解像では有効でなかった。これらを受けて、パッチングがボリュームデータに対して有効か評価する。

最後は、クラス識別情報である。クラス識別情報は、データに内在するカテゴリごとの特徴を利用して識別を行うための情報であり、ボクセルデータの超解像では有効とされている。一方、画像データの超解像では、多様な被写体(クラス)を含んだ画像を混ぜて学習しており、クラス識別情報は有効でない。しかし、本研究で使用するボリュームデータは、臓器ごとに特徴が大きく異なるため、ボクセルデータと同様にクラス識別情報が有効である可能性がある。これを踏まえ、クラス識別情報がボリュームデータ超解像において有効かを評価する。

これらの評価の結果、ボリュームデータ超解像において、損失関数にSSIMを用いることとパッチングを用いることの有効性が示された。一方で、クラス識別情報はボリュームデータの超解像精度に対して顕著な影響を与えないことが明らかとなった。

目次

1. はじめに	1
2. 超解像に関する従来研究	4
2.1. 超解像の定義	4
2.2. 二次元画像超解像の従来研究	5
2.3. 三次元ボリュームデータ超解像の従来研究	7
2.4. 3D-SRCGAN による三次元ボクセルデータの超解像	9
3. 三次元医療データ超解像への 3D-SRCGAN の応用	10
3.1. 三次元医療データへの応用	10
3.2. 損失関数	11
3.3. パッチング	12
3.4. 3D-SRCGAN のネットワーク構造	13
3.5. 3D-SRCGAN の Generator の構造	14
3.6. 3D-SRCGAN の Discriminator の構造	17
3.7. Generator の学習更新	20
3.8. Discriminator の学習更新	22
4. 実験	23
4.1. 実験設定	23
4.2. 損失関数の違いによる精度評価	25
4.2.1. 実験内容	25
4.2.2. 実験結果	25
4.3. パッチングの有無による精度評価	28
4.3.1. 実験内容	28

4.3.2. 実験結果	28
4.4. 3D-SRGAN と 3D-SRCGAN の精度評価	31
4.4.1. 実験内容	31
4.4.2. 実験結果	31
5. おわりに.....	34
謝辞	36
参考文献.....	37
付録 A. 超解像結果の画像例.....	39
付録 B. 数値評価結果	44

1. はじめに

医療診断や治療計画において、MRI や CT スキャンといった三次元医療データは重要な役割を果たしている。データの解像度が低い場合、微細な病変の検出や正確な解析が困難になることがある。しかし、診断や治療に必要な高解像度データを取得しようとする、MRI では高価な機器を購入する必要がありコストがかかる。また、CT スキャンでは長時間 X 線を照射する必要があり放射線被爆のリスクが高まるという問題がある。これに対して、近年、超解像によってデータの解像度を上げる研究が盛んになっている。

超解像の研究は二次元画像を対象とするものが多い。代表的な例として、SRGAN[3]が挙げられる。この手法は、敵対的生成ネットワーク (GAN) を利用して、低解像度画像から高品質な高解像度画像を生成することに成功したものである。しかし、三次元医療データはボリュームデータであるため、二次元画像に対する超解像技術とは異なるアプローチが求められる。

三次元医療データの超解像に関する研究は、近年注目を集めている。三次元ボリュームデータは、複数のスライス画像から構成されるため、画像間の関係を保持しながら解像度を向上させることが求められる。Li ら[4]は、3D 畳み込みニューラルネットワーク (CNN) を用いた軽量な並列ネットワーク「VolumeNet」を提案し、脳 MRI や腹部 CT 画像の高精度な超解像を実現した。彼らの手法では、ボリュームデータ全体を一度に処理することで、スライス間の情報を適切に補完しており、精度の高い三次元超解像を達成している。

ボリュームデータに類似した三次元データとして、ボクセルデータがある。ボリュームデータは三次元空間を賽の目状に分割したボクセルにその位置での観測データを保持することで、三次元的なデータの分布を記録する。ボクセルデータも、ボクセル値を保持するが、その位置が物体の内部であるか外部であるかを表す 2 値データを保持することで、三次元形状を記録するもので、性質が異なる。ボクセルデータについても、超解像の研究が行われている。3D-SRCGAN[1]は、画像の超解像手法である SRGAN[3]を三次元用に改良したものである。この手法は、特にボクセルレベルでの高解像度化を実現するために設計されており、例えば家具などの三次元形状を表すボクセルデータを超解像するために有効である。ボクセルデータは三次元空間内で格子状に配置された小さな単位から構成されており、これを超解像す

するためには空間的な構造を保持しながら解像度を向上させる必要がある。空間的な構造の特徴は、超解像の対象となる物体の種類(クラス)により大きく異なる。3D-SRCGAN は入力データにクラスに識別情報を付加することにより、複数クラスのデータを同時に学習する際に発生する異なるクラス間の特徴の混合問題を解消し、精度向上を実現している。

画像データでは分布が重要で、ボクセルデータは形状を表現することが重要でありそのデータ特性が異なる。一方、ボリュームデータは内部の分布が重要な要素となるため、画像データの特性と似ている。本研究は、画像の超解像を目的として開発された SRGAN[3]をボクセルデータに適用した 3D-SRCGAN[1]の手法を、さらにボリュームデータに適用する。その際、ボクセルデータ、ボリュームデータ、画像データの特徴の違いを考慮して、3つの観点から、3D-SRCGAN のボリュームデータへの適用可能性を検証する。

1つ目の観点は損失関数である。損失関数は、深層学習において教師データと出力を近づけるための評価を行う関数であり、その値が小さいほど出力が正解に近いことを示す。一般に深層学習を用いた超解像では、損失関数として平均二乗誤差 (Mean Squared Error, MSE) が広く使用されているが、近年では視覚的な品質を向上させるために構造的類似度 (Structural Similarity Index Measure, SSIM) を組み込んだ損失関数も提案されている。本研究では、ボリュームデータにおいても画像データと同様に構造情報が重要であることに着目し、SSIM を損失関数として使用することの有効性を調査する。

2つ目は、パッチング手法の有効性である。パッチングとは、データを部分に分割してその単位で超解像を行う手法で、画像の超解像時に一般的に使用される。しかし、ボクセルデータの超解像のための 3D-SRGAN[2]の研究では精度が下がる結果になっていたため、3D-SRCGAN の研究では採用されていなかった。本研究ではボリュームデータに対してパッチングが有効かどうかを調査する。

3つ目の観点は、クラス識別情報である。3D-SRCGAN[1]によるとクラス識別情報はボクセルデータの超解像では有効とされている。一方、画像データの超解像では多様な被写体(クラス)を含んだ画像を混ぜて学習しており、クラス識別情報はあまり有効ではない。しかし、本研究で使用するボリュームデータは、臓器ごとに特徴が大きく異なるため、ボクセルデータと同様にクラス識別情報が有効かどうかを調査する。

以下、2章では、従来行われていた画像やボクセルデータの超解像について述べる。3章では、本研究で使用する3D-SRCGANを説明するとともに、その三次元医療データへの適用において考慮する3つの観点について述べる。4章では、損失関数、パッチング、クラス識別情報が3D-SRCGANを用いた三次元医療データの超解像の精度向上に有効であるか、実験により評価する。最後に5章で、結論と今後の課題を述べる。

2. 超解像に関する従来研究

2.1. 超解像の定義

超解像は、低解像度データを基に高解像度データを生成および補完する技術である。解像度とは、画素やボクセルなどのデータ構成単位が一定範囲内にどの程度密に存在するかを示す指標である。本研究では、超解像の対象として三次元医療データを扱っている（図1参照）。超解像技術によって生成された高解像度データを「超解像三次元データ」と呼ぶ。

超解像は、低解像度データを高解像度化する手法であるが、この変換は不良設定問題（ill-posed problem）[4]に該当する。不良設定問題とは、解を導き出すのに必要な情報が一部欠如している状態を指す。低解像度データから高解像度データを得るためには、低解像度データに存在しない部分の情報を生成する必要があるが、その生成方法には無数の可能性が考えられる。また、実際のデータにはノイズが含まれていることが多く、さらにデジタルデータには量子化誤差も含まれるため、低解像度データから正確な高解像度データを生成することは一般的に非常に困難な課題である。

入力：低解像度三次元データ

出力：超解像三次元データ

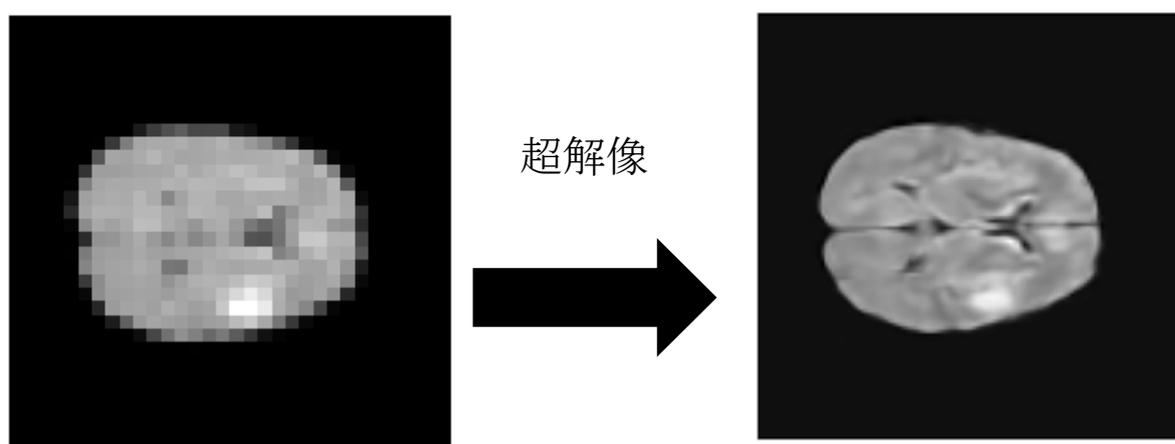


図1. 三次元医療データの超解像(スライス画像)

2.2. 二次元画像超解像の従来研究

二次元画像の超解像技術には、大きく分けて複数の画像を利用する手法と単一の画像を用いる手法の2種類がある。複数画像を活用する方法では、同一の対象を撮影した複数の画像を組み合わせることで、高精細な画像を生成する。このアプローチは実際の画像から情報を引き出すため、精度の高い結果が得られる点が利点である。しかしながら、適用には条件があり、撮影環境や画像の取得方法が限定される場合がある。

一方で、単一画像を基にした超解像では、補間に基づく信号処理アプローチ[5][6]と、学習や事例ベースの復元アプローチ[7]が存在する。補間ベースの手法は、元画像の高周波成分が失われている場合に十分な改善が期待しづらい。一方、学習や事例ベースの手法は、データセットや学習した情報を基にして画像に欠けている高周波成分を補い、高解像度化を可能にする。しかし、この方法では、元の画像に存在しない情報を補う必要があるため、生成される画像が元画像の正確な高解像版とは限らない。

近年では、深層学習を利用した超解像技術の進展が著しい。深層学習は多層構造を持つニューラルネットワークを用いた手法であり、特に畳み込みニューラルネットワーク (Convolutional Neural Network: CNN) が広く採用されている。CNNは、カーネルと呼ばれる小さな数値行列を用いて入力画像を畳み込み処理し、特徴を抽出する。具体的には、カーネルが画像の一部をスライドしながら画素ごとの積の和を計算し、出力として特徴マップを生成する。この処理を繰り返し行うことで、画像の重要な情報を抽出する。

CNNを活用した手法の代表例として、SRCNN[8]が挙げられる。このモデルは、大量の画像データを学習し、一般的な画像の超解像に適した特徴を習得することで、自然な高解像度画像を生成する。

また、近年ではGAN (敵対的生成ネットワーク) を活用した超解像手法も提案されている。GANベースの技術は、SRCNNに比べてさらに高い精度での超解像を可能にする点が特長である。GANベースの手法として、SRGAN[3]が挙げられる。SRGAN[3]とは、GANを活用した画像の超解像手法である。生成器と判別器の敵対的学習を通じて、従来の平均二乗誤差 (MSE) を最小化する手法では得られない、鮮明で視覚的に優れた画像を生成する。SRCGAN[9]は、SRGANに条件付きの要素

を加えたもので、識別器に高解像画像を条件として使用している。この条件付きのアプローチにより、識別器が本物の画像と超解像画像を区別できるようになり、精度の向上が実現されている。

深層学習を用いた画像の超解像では、正解の高解像度画像と類似した超解像画像を生成するために、損失関数の一部として平均二乗誤差(MSE)が利用されている。一方で、超解像結果の評価には、人間の視覚的評価により近いとされる構造類似性指数(SSIM)も使われている。本研究では、SSIMを損失関数として直接利用して学習を行うことで、視覚的により良い超解像結果が得られるか検証する。また、画像の超解像では一般に、画像を小さな領域に分割し、それぞれの領域毎に超解像を行う「パッチング」が使用されている。パッチングにより、超解像の際に参照するデータサイズを限定でき、問題を簡素化して超解像を実現できる。本研究では、画像に似た性質を持つボリュームデータの超解像において、パッチングが有効か検証する。

2.3. 三次元ボリュームデータ超解像の従来研究

三次元超解像は、二次元画像の超解像技術を三次元ボリュームデータへと応用することで実現される。この技術は主に医療画像の分野で利用されている。高解像度な医療画像は、診断の正確性や治療計画の最適化に寄与するだけでなく、新しい知見や治療法の開発にもつながり、医療分野の進展において極めて重要な役割を果たしている。

医療画像分野における三次元超解像は、主に磁気共鳴画像法 (Magnetic Resonance Imaging: MRI) やコンピュータ断層撮影 (Computed Tomography: CT) で得られる三次元データに適用されている [10][11]。MRI は、磁場と電磁波を利用して体内の水素原子を振動させ、その過程で発生する微弱な電磁波を測定して多値表現の断層画像を生成する。一方、CT は X 線を人体に照射し、透過した X 線量を基に断層画像を得る。これらの断層画像を積み重ねることで三次元データを構築し、人体内部の臓器の状態や分布を視覚化することが可能である。しかし、これらの技術で高解像度な画像を直接取得する場合、患者への負担が増大するという課題がある。

低解像度の三次元画像から超解像技術を用いて高解像度画像を生成することで、患者の被曝を最小限に抑えつつ、精度の高い医療画像を提供できる。このようなアプローチにより、診断や治療における安全性と有用性を両立させることが可能となる。

Li ら [4] が提案した、三次元畳み込みニューラルネットワーク (CNN) を用いた軽量の並列ネットワーク「VolumeNet」は、脳 MRI および腹部 CT 画像における高精度な超解像を実現した。彼らの手法では、ボリュームデータ全体を一度に処理することで、スライス間の情報を適切に補完し、精度の高い三次元超解像を達成している。VolumeNet は、軽量かつ並列処理を可能にする設計により、計算資源を効率的に利用し、リアルタイムで高精度な結果を得ることができる。さらに、三次元畳み込みを採用することで、スライス間の関係性を捉え、詳細な画像再構成を実現しており、医療画像診断における有用性が示された。

このように、医療分野における三次元超解像では、ボリュームデータの特徴である多値表現されたデータを基に、対象物の分布や状態を的確に把握することが重視されている。ボリュームデータは、解像度の向上により、より詳細な構造解析が可

能となり、診断精度の向上に貢献する。特に、脳 MRI や腹部 CT などの複雑な内部構造を有する画像において、この技術は重要な役割を果たす。

2.4. 3D-SRCGAN による三次元ボクセルデータの超解像

三次元ボリュームデータに類似したものとして三次元ボクセルデータが挙げられる。ボクセルデータは、三次元空間を賽の目状に分割したボクセルに、その位置が物体の内部か外部かを表す 2 値データを保持することで三次元形状を記録するものである。三次元ボクセルデータの超解像技術の手法として、3D-SRCGAN (3D-Super Resolution Conditional Generative Adversarial Networks) [1]がある。従来の 3D-SRGAN では、複数クラスのデータを同時に学習する際に発生する異なるクラス間の特徴の混合問題があった。3D-SRCGAN は 3D-SRGAN にクラス識別情報を組み込むことによって、それらを解消し精度向上を図ったものである。実験結果においては、3D-SRCGAN は 3D-SRGAN と比べ、異なるクラスのデータを用いた場合でもクラス間の特徴の混合問題が起きず、超解像精度が顕著に向上することが確認された。本研究は、この 3D-SRCGAN を三次元医療データに適用する。

3. 三次元医療データ超解像への 3D-SRCGAN の応用

3.1. 三次元医療データへの応用

3D-SRCGAN が提案された研究[1]ではボクセルデータを対象に超解像を行っていた。一方、本研究で扱う三次元医療データは、ボリュームデータである。

ボクセルデータは、物体の形状を表現するためのデータ形式であり、各ボクセルが存在 (1) または非存在 (0) を示す 2 値データとして構成される。このデータにおいては、局所的な形状よりも全体に対する位置関係が重要視される。また、物体の内部構造よりも内外の区別が重要となる。さらに、椅子と机のようにクラス間で形状の特性が大きく異なり、椅子同士や机同士の形状の違いよりも顕著な場合が多いため、クラス識別情報を利用することで精度の向上が期待できる。

一方、ボリュームデータは物質の濃度分布を表現する多値データとして構成される。この特性により、局所的な領域を切り出しても、濃度分布として有用な情報を保持することが可能である。また、ボリュームデータにおいては物体の内部構造が重要な要素となるため、構造的類似性を評価する指標である SSIM が有効と考えられる。さらに、脳や胸部といった部位ごとに濃度分布が異なる特徴を持つことから、クラス識別情報を利用することでさらなる性能向上が期待できる。

このような違いから、ボクセルデータ超解像に有効だった 3D-SRCGAN がボリュームデータに対しても有効か調査する。

3.2. 損失関数

損失関数とは、深層学習において教師データと出力との間の差を定量的に評価するための関数である。損失関数の値が小さいほど、モデルの出力が正解に近いことを示しており、この値を最小化するようにモデルが学習される。

3D-SRCGAN を提案した研究[1]では、対象データとして主に家具のボクセルデータが使用されており、損失関数は、個々のボクセル値の誤差を最小化することを目的として、平均二乗誤差 (Mean Squared Error, MSE) が用いられていた。MSE は計算が容易であり、個々のボクセル値の違いを直接的に評価できるため、ボクセルデータを対象とした研究において広く採用されてきた。

しかし、本研究では、対象データとして三次元医療データを扱っており、連続的な空間的構造や特徴が重要な意味を持つ。このため、単純なボクセル単位の誤差評価では不十分であると判断した。医療データの空間的な連続性や構造的特徴をより正確に捉えるため、構造類似性指数 (Structural Similarity Index, SSIM) も損失関数として使用する。SSIM は、画像の輝度、コントラスト、構造といった視覚的特性を基に、2つの画像間の類似度を評価する指標である。これにより、単純なピクセルごとの差ではなく、人間の視覚に基づいたより直感的な評価が可能となる。数値は-1 から 1 の範囲を取り、1 に近いほど画像の類似度が高いことを示す。

本研究では、損失関数として以下の3つを比較検証する。

1. MSE Loss: 個々のボクセル値の誤差を最小化する
2. SSIM Loss: 空間的な連続性と全体の構造的な一貫性を向上させる
3. MSE Loss + SSIM Loss: ボクセル単位の精度と空間的一貫性を同時に最適化

また、これらの損失関数に加え、超解像タスク特有の敵対的ロス (Adversarial Loss) を導入し、生成結果の精度を向上させる。

3.3. パッチング

パッチングとは、データを小さな領域に分割し、それぞれの領域ごとに超解像を行う手法である。一般に、超解像の学習ではデータサイズが大きくなると計算負荷が増加し、学習が困難になる傾向がある。そのため、データを小さい単位に分割して処理することで、問題を簡素化し、効率的に超解像を実現するという背景から、パッチングの手法が用いられている。

パッチングは、ボクセルデータのように形状情報を主とするデータに対しては効果が限定的であるが、画像のような分布を表すデータには適しており、その特性からボリュームデータにおいても有効であると考えられる。また、パッチングはデータ拡張の効果を持ち、学習の多様性を高めることで精度の向上が期待できる。これらの利点を踏まえ、本研究ではボリュームデータの超解像におけるパッチングの有効性を検証する。

本研究では、学習の際にはパッチングを行うが、学習した生成器を用いて超解像を行う際には、パッチング処理を行わない手法を採用している。具体的には、学習時に $16 \times 16 \times 15$ から $64 \times 64 \times 60$ の超解像をパッチングで学習した場合でも、超解像時には $32 \times 32 \times 30$ から $128 \times 128 \times 120$ の超解像を行うことで、パッチ間の境界が生じないようにしている。このアプローチにより、不自然なアーティファクトの発生を防ぎ、空間的な整合性を維持することを可能としている。

3.4. 3D-SRCGAN のネットワーク構造

3D-SRCGAN は、Generative Adversarial Network (GAN) と同様に、Generator と Discriminator の 2 つのネットワークから成り立っており、ボリュームデータを入出力データとして扱う (図 2)。学習データには、低解像度のボリュームデータと高解像度のボリュームデータのペアが使用される。Generator は、低解像度のボリュームデータから高解像度のボリュームデータを生成し、Discriminator は、入力されたボリュームデータが学習データの高解像度データであるか、Generator が生成したデータであるかを識別する。

この 2 つのネットワークは敵対的な関係にあり、それぞれの目的は以下の通りである。Generator の目的は、Discriminator を欺いて、学習データに似たボリュームデータを生成することであり、Discriminator の目的は、学習データと Generator が生成したボリュームデータを区別することにある。最終的に、3D-SRCGAN の目的は、Discriminator が識別できないような高精度なボリュームデータを Generator が生成することである。そして、学習を通じて最終的に得られた Generator により、超解像を行う。

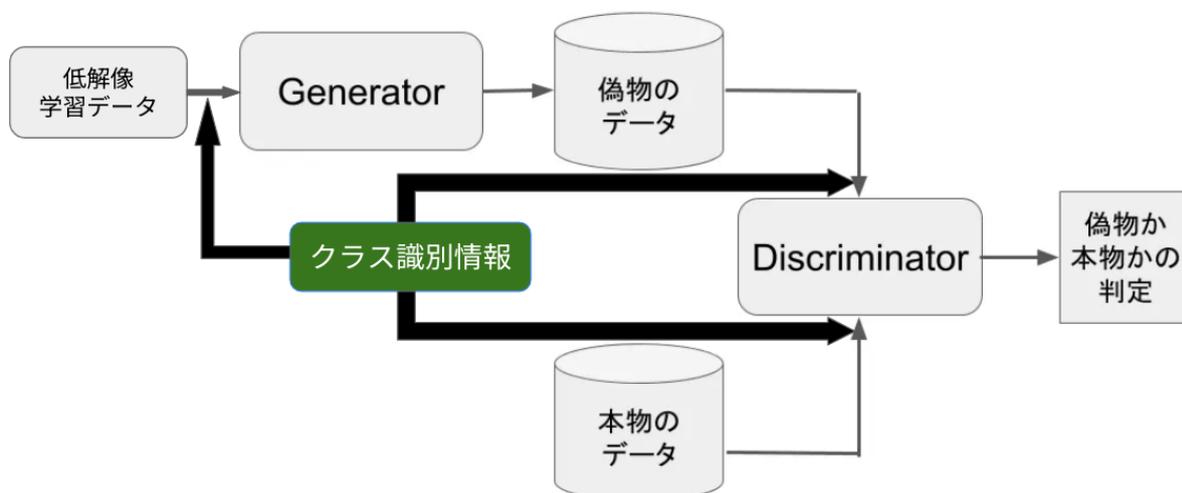


図 2 3D-SRCGAN の概要図

3.5. 3D-SRCGAN の Generator の構造

3D-SRCGAN の Generator の全体構造は図 3 に示す。入力チャンネル数を C_{in} 、出力チャンネル数を C_{out} 、ストライドを、カーネルのサイズを k と定義する。ここで、ストライドは畳み込み処理を行う際にカーネルをずらしていく間隔を指す。3D-SRCGAN における三次元畳み込み層のストライドは全て $s = 1$ である。

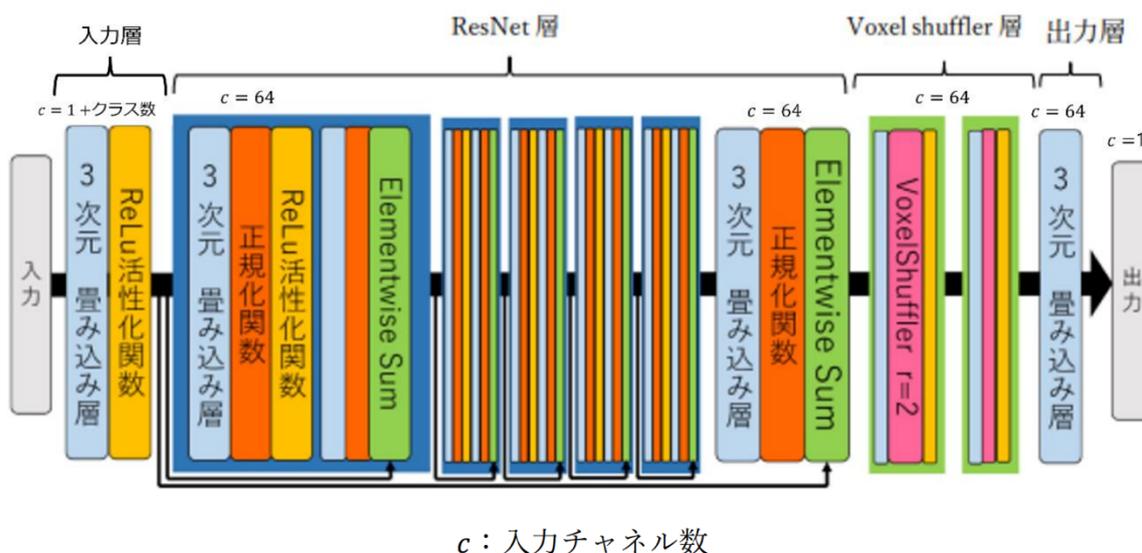


図 3 Generator の構造[1]

Generator は、低解像度ボリュームデータを入力として、高解像度ボリュームデータを出力する。構造は、入力層、ResNet 層、Voxel Shuffler 層、出力層の 4 つの部分に分かれている。入力層では、低解像度のボリュームデータが与えられ、それを三次元畳み込み層を通して複数のチャンネルに変換し、ResNet 層に渡す。ResNet 層では、ボリュームデータの特徴を抽出し、Voxel Shuffler 層でボリュームデータの拡大を行う。その後、出力層において入力の特徴マップから高解像度のボリュームデータを生成して出力する。

Generator では、三次元畳み込み層を使用する。各畳み込み層のカーネルサイズは $k \times k \times k$ である。三次元畳み込み層の後には、正規化関数と活性化関数が適用される。正規化関数は、三次元畳み込み層の出力を正規化し、学習速度の向上と学習の安定化を実現する [17]。活性化関数には式(1)の ReLU 関数(図 4)が使用される。

$$f(x) = \max(0, x) \quad (1)$$

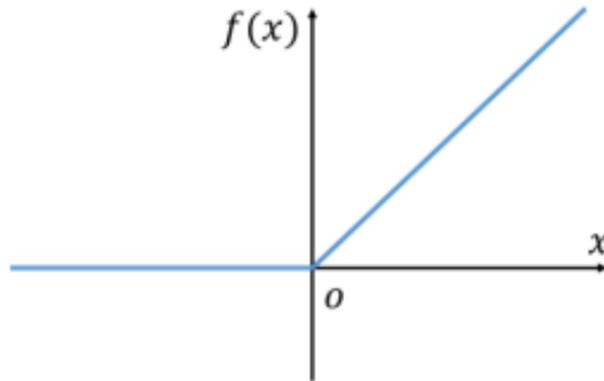


図4 ReLu 活性化関数[1]

Generator の各層における処理は以下の通りである。まず、入力層では、 $C_{in} = 1 +$ クラス数、 $C_{out} = 64$ 、カーネルサイズ $k = 9$ の三次元畳み込み層、続いて ReLU 活性化関数が順次適用される。クラス識別情報は、one-hot 表現を採用している。one-hot 表現とは、クラス数分の次元を持ったベクトルの各クラスに対応する値だけを 1、それ以外を 0 とすることで、クラスの情報を与える手法である。クラス識別情報の付加は、まず各クラスにクラス番号を割り振ることで開始する。次に、入力されるボリュームデータに対して、クラス数分のチャンネルを持つ空のボリュームデータを用意し、割り振られたクラス番号に対応するチャンネルのボクセル値をすべて 1 で埋める形で実現する。

次に、ResNet 層は 5 つの Residual Block で構成されており、各 Residual Block 内では、まず $C_{in} = 64$ 、 $C_{out} = 64$ 、 $k = 3$ の三次元畳み込み層、正規化関数、ReLU 活性化関数が適用され、その後再度、 $C_{in} = 64$ 、 $C_{out} = 64$ 、 $k = 3$ の三次元畳み込み層と正規化関数が続く。各 Residual Block では、入力と出力が加算される。この加算により、各 Residual Block は入力と出力の差異を学習することが可能となる。5 つの Residual Block の処理が完了した後、再度 $C_{in} = 64$ 、 $C_{out} = 64$ 、 $k = 3$ の三次元畳み込み層と正規化関数が適用され、最終的に ResNet 層の入力が加算されて出力される。

Voxel Shuffler 層は、ボリュームデータのサイズを拡大するための重要な役割を担っている。図 5 に示されるように、この層は入力された特徴マップを並び替えることで、ボリュームデータの解像度を効果的に拡大する。Voxel Shuffler 層は、2 つのブロックで構成されており、各ブロックは入力の特徴マップから 2 倍の解像度に

拡大する。各ブロック内の処理は、最初に $C_{in} = 64$ 、 $C_{out} = 512$ 、ストライド $s = 1$ 、カーネルサイズ $k = 3$ の三次元畳み込み層が適用され、その後、 $C_{in} = 512$ 、 $C_{out} = 64$ の Voxel Shuffler 層が続く。最後に、ReLU 活性化関数が適用され、各ブロックを通過した後、出力は 4 倍に拡大される。

最終的に、高解像度のボリュームデータを生成するために、 $C_{in} = 64$ 、 $C_{out} = 1$ ストライド $s = 1$ 、カーネルサイズ $k = 9$ の三次元畳み込み層が適用される。この処理により、低解像度の入力から高解像度のボリュームデータを得ることができる。

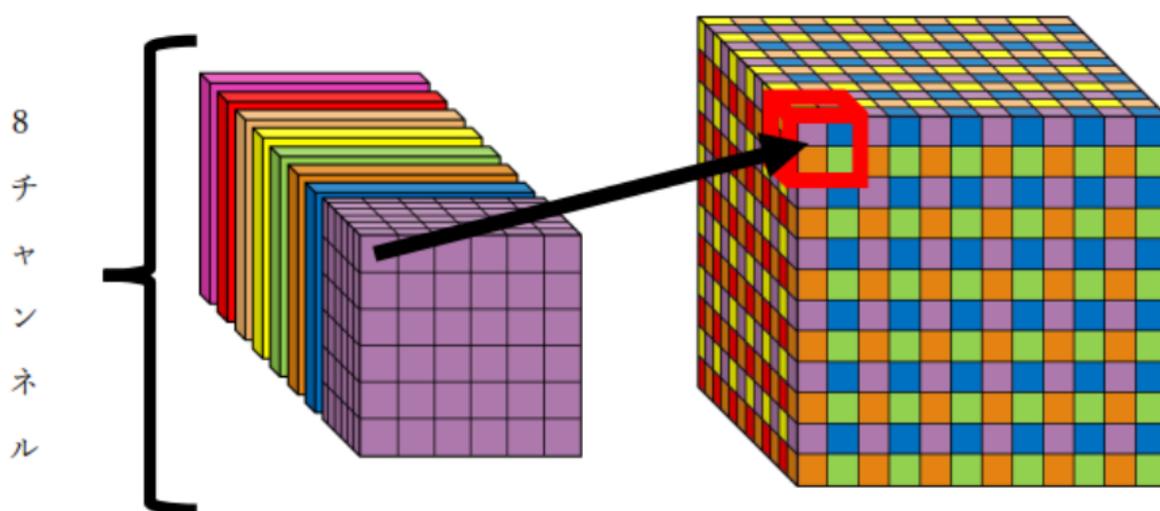
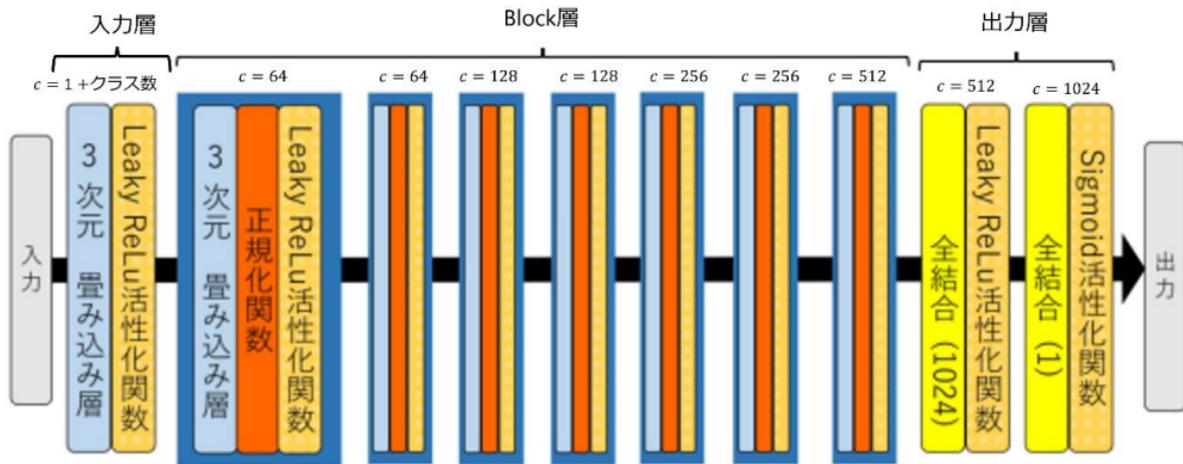


図 5 Voxel Shuffler

3.6. 3D-SRCGAN の Discriminator の構造

3D-SRCGAN の Discriminator の構造を図 6 に示す。Discriminator は入力層、Block 層、出力層の 3 つの部分に分かれている。入力層では、入力されたボリュームデータから複数のチャンネルを抽出し、その情報を Block 層に供給する。Block 層では、ボリュームデータの特徴を効率的に抽出し、出力層では、与えられた ボリュームデータが学習データである確率を出力する。



c : 入力チャンネル数

図 6 Discriminator の構造

Discriminator でも Generator と同様に、三次元畳み込み層を利用して ボリュームデータに対する畳み込み処理が行われる。これにより、Discriminator は入力された ボリュームデータが本物 (学習データ) か、それとも Generator によって生成された偽物であるかを識別できるようになる。

さらに、Discriminator では、活性化関数として Leaky ReLU 活性化関数 (図 7) と Sigmoid 活性化関数 (図 8) を使用しており、これらはそれぞれ式 (2) と式 (3) で表現できる。

$$f(x) = \max(\alpha x, x), 0 < \alpha < 1 \quad (2)$$

$$f(x) = \frac{1}{1 + \exp(-\beta x)}, 0 < \beta \quad (3)$$

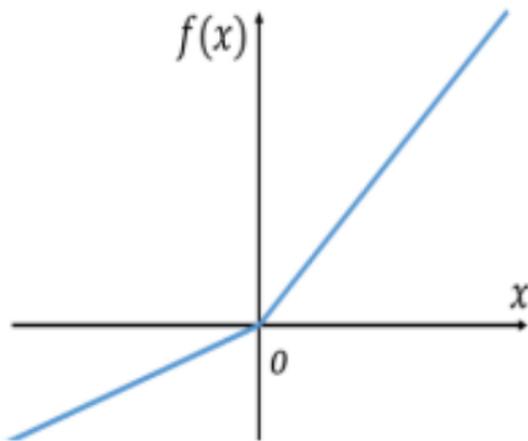


図7 Leaky ReLU 活性化関数

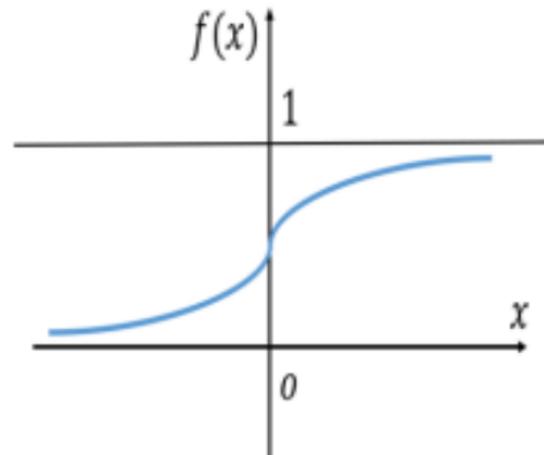


図8 sigmoid 活性化関数

このように、Discriminator のネットワークは、ボリュームデータの識別を行うために、三次元畳み込み層と適切な活性化関数を組み合わせて構成されている。

次に、Discriminator の各層で行われる処理について説明する。まず、最初の層では、入力チャンネル数 $C_{in} = 1 + \text{クラス数}$ 、出力チャンネル数 $C_{out} = 64$ 、ストライド $s = 1$ の三次元畳み込み層が処理され、その後に Leaky ReLU 活性化関数が適用される。クラス識別情報の付加は Generator と同様である。この処理により、入力されたボリュームデータから特徴が抽出される。

次の Block 層は 7 つのブロックで構成されている。それぞれのブロックは、以下の順序で処理を行う。まず 三次元畳み込み層、次に正規化関数、そして最後に Leaky ReLU 活性化関数が適用される。各ブロックにおける 三次元畳み込み層のパラメータは、次のように設定されている。

- 1 番目のブロック： $C_{in} = 64$ 、 $C_{out} = 64$ 、ストライド $s = 2$
- 2 番目のブロック： $C_{in} = 64$ 、 $C_{out} = 128$ 、ストライド $s = 1$
- 3 番目のブロック： $C_{in} = 128$ 、 $C_{out} = 128$ 、ストライド $s = 2$
- 4 番目のブロック： $C_{in} = 128$ 、 $C_{out} = 256$ 、ストライド $s = 1$
- 5 番目のブロック： $C_{in} = 256$ 、 $C_{out} = 256$ 、ストライド $s = 2$
- 6 番目のブロック： $C_{in} = 256$ 、 $C_{out} = 512$ 、ストライド $s = 1$
- 7 番目のブロック： $C_{in} = 512$ 、 $C_{out} = 512$ 、ストライド $s = 2$

各 三次元畳み込み層のカーネルサイズはすべて $k = 3$ であり、特徴量を効率的に抽出する役割を果たす。

7つのブロックを通過した後、次に全結合層が $C_{out} = 1024$ で適用され、さらに Leaky ReLU 活性化関数が続く。その後、出力層で $C_{out} = 1$ の全結合層が適用され、Sigmoid 活性化関数を通して最終的な出力が得られる。この出力は、0 と 1 の間の値となり、入力された ボリュームデータが学習データである確率を表す。

3.7. Generator の学習更新

3D-SRCGAN の損失関数は、MSE Loss の L_{MSE} 、SSIM Loss の L_{SSIM} 、Adversarial Loss の L_{ADV} の三種類があり、次のように定義される。

$$L_{MSE} = \sum_{m=1}^m \frac{1}{r^3 W H D} \sum_{x=1}^{rW} \sum_{y=1}^{rH} \sum_{z=1}^{rD} (l_{x,y,z}^{HR} - G(l^{LR})_{x,y,z})^2 \quad (4)$$

$$L_{SSIM} = \sum_{m=1}^m \frac{((2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2))}{((\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2))} \quad (5)$$

$$L_{ADV} = \sum_{m=1}^m -\log D(G(l^{LR})) \quad (6)$$

MSE Loss は式 (4) によって計算される。ここで、 m はバッチサイズ、 W 、 H 、 D はそれぞれ三次元空間の幅、高さ、奥行きを表す。式(4)では、学習データの高解像度ボリュームデータ l^{HR} と Generator が生成した超解像三次元ボリュームデータ $G(l^{LR})$ の平均二乗誤差が計算される。MSE Loss は、元のデータの見た目がある程度維持したまま異なる操作を加えることができる特徴を有している。このため超解像処理を行っても元のデータの見た目が大きく変化することはない。

SSIM Loss は式(5)によって計算される。ここで、 μ_x と μ_y はそれぞれ元データと生成データの平均値を表し、 σ_x^2 と σ_y^2 は元データおよび生成データの分散を示す。また、 σ_{xy} は元データと生成データの共分散であり、 C_1 と C_2 は数値計算を安定化させるための小さな定数である。SSIM Loss は、元データと生成データの構造的類似性を測定する指標であり、人間の視覚特性をモデル化している。このため、視覚的品質を向上させる効果があり、特に局所的な構造の一貫性を保持する点で優れている。さらに、単なるピクセル単位の誤差を評価する MSE Loss とは異なり、生成されたボリュームデータの形状や質感を正確に維持しながら超解像処理を行うことが可能である。

Adversarial Loss は式 (6) により計算される。ここで、 $G(l^{LR})$ は Generator が生成した超解像ボリュームデータで、 $D(\cdot)$ は Discriminator の出力である。Adversarial Loss は GAN において通常使用される損失関数であり、生成した超解像データが学習データ(高解像度データ)と見分けがつかないように学習を進める特徴を有する。

これらの損失関数を用いて、Generator の損失関数 L_G として次の三種類を定義し

実験により比較する。

$$L_G = \frac{1}{m} (L_{MSE} + \lambda L_{ADV}) \quad (7)$$

$$L_G = \frac{1}{m} (L_{SSIM} + \lambda L_{ADV}) \quad (8)$$

$$L_G = \frac{1}{m} (L_{MSE} + L_{SSIM} + \lambda L_{ADV}) \quad (9)$$

Adversarial Loss の係数は、さまざまな値を試してみた結果、一番良かった $\lambda = 10^{-3}$ を使用する。

MSE Loss のみでは、見た目の変化が解決されても明確なボリュームデータが生成されることはない。一方で、Adversarial Loss のみでは学習データの性質は保持されるが、入力とは異なるボリュームデータが生成されることになる。したがって、式(7)のようにこれら 2 つの Loss を適切に組み合わせることにより超解像が実現される。

式(8)のように SSIM Loss と Adversarial Loss を組み合わせた場合、生成されたボリュームデータは、見た目の視覚的品質をさらに向上させる特徴を持つ。SSIM Loss は視覚的な構造的類似性を高める一方、Adversarial Loss によって学習データと見分けがつかないようなモデルが生成される。この組み合わせにより、視覚的に自然でありながら元データの特性を保持した超解像が可能となる。

さらに、式(9)のように MSE Loss、SSIM Loss、および Adversarial Loss を全て組み合わせた場合、それぞれの特長を活かしつつ、ボリュームモデルの生成を最適化することが可能である。MSE Loss はデータのピクセル単位での正確性を維持し、SSIM Loss は視覚的な構造的類似性を向上させ、Adversarial Loss は生成されたモデルを学習データに近づける役割を果たす。このように、3 種類の Loss を統合することで、異なる観点から超解像の精度向上を図ることができる。

Generator の学習では、損失関数の値が小さくなるようにネットワークのパラメータを逐次更新する。最適化アルゴリズムとして Adam を用いて収束計算を行う。

3.8. Discriminator の学習更新

3D-SRCGAN の Discriminator の損失関数 L_D は、式(10)により定義される。

$$L_D = \frac{1}{m} \sum^m \left[\log D(l^{HR}) - \log(1 - D(G(l^{LR}))) \right] \quad (10)$$

ここで、 m はバッチサイズを示す。 $D(x)$ は Discriminator の出力であり、入力データ x が学習データの ボリュームデータである確率を表す。従って、Discriminator の識別が正確であれば、 $D(l^{HR})$ の値は大きくなり、生成されたボリュームデータ $G(l^{HR})$ に対する $D(G(l^{HR}))$ の値は小さくなることが期待される。

Discriminator の学習は、Generator の学習と同様に、Adam を用いてパラメータを更新し、収束計算が行われる。

4. 実験

4.1. 実験設定

3D-SRCGAN をボリュームデータに適用し、有効性を調査する。本研究で、学習、評価に用いた三次元医療データは、MedicalDecathlon[11]の Brain データセット、Cancer Imaging Archive[12]から Breast、Lung の合計三種類の部位を使用した。Brain と Breast は MRI データで、Lung は CT データである。データ数は、Brain が 484、Breast が 1510、Lung が 553 となっており、学習と評価のデータの割合は 8:2 で行った。

データの解像度は、パッチングを使う学習の場合、三次元データの 8 分の 1 の領域をランダムに抽出し、 $16 \times 16 \times 15$ の低解像度データと $64 \times 64 \times 60$ の高解像度データのペアを使用した。8 つに均等に抽出しなかったのは、データの中央に臓器が位置していたため、ランダムに抽出したほうが 1 データあたりの臓器の情報量が増えると考えたためである。パッチングを使わない学習の場合は、 $32 \times 32 \times 30$ の低解像度データと $128 \times 128 \times 120$ の高解像度データのペアで学習を行った。超解像時は、 $32 \times 32 \times 30$ の低解像度データを $128 \times 128 \times 120$ へ超解像した。

学習のパラメータは予備実験の結果から、学習率が 6×10^{-5} 、Epoch 数が 200 で、70Epoch ごとに学習率を半分になっている。バッチサイズは、実験ごとに異なるので、実験の節でそれぞれ述べる。

3D-SRCGAN を用いた三次元医療データの超解像の結果は、以下の 3 項目で定量評価した。

1. MSE (Mean Squared Error)

$$MSE = \frac{1}{r^3 WHD} \sum_{x=1}^{rW} \sum_{y=1}^{rH} \sum_{z=1}^{rD} (l_{x,y,z}^{HR} - G(l^{LR})_{x,y,z})^2$$

高解像データと超解像データとの差異を二乗して平均した値で、値が小さいほど精度が高い。

2. SSIM (Structural Similarity Index Measure)

$$SSIM = \frac{((2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2))}{((\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2))}$$

構造的な類似度を測定する指標で、1に近いほど元画像と類似している。

3. PSNR (Peak Signal-to-Noise Ratio)

$$PSNR = 10 \log_{10} \left(\frac{L^2}{MSE} \right)$$

信号対雑音比を示す指標で、高い値ほど品質が良い。 L は画素値の最大値を表す。

4.2. 損失関数の違いによる精度評価

ボクセルデータ超解像のために提案された 3D-SRCGAN の研究では、損失関数は、MSE のみを使用していた。本研究はボリュームデータを扱っており、ボクセルデータとは性質が異なる。ボクセルデータ超解像は、個々のボクセル値の誤差を最小化することが目的なので、MSE を利用していたが、ボリュームデータは、空間的な連続性や構造的特徴が重要な意味を持つため、単純なボクセル単位の誤差評価では不十分であると判断した。SSIM を損失関数として使用することで超解像精度の改善が期待できる。また、MSE と SSIM の併用が超解像精度向上に有効かも評価する。

4.2.1. 実験内容

本実験では、3D-SRCGAN を使用して MSE、SSIM、MSE と SSIM の併用、の 3 パターンの損失関数を用いた時の超解像精度の比較を行う。バッチサイズは、14 としたが、メモリ容量の制約によりオーバーフローが発生した場合には、バッチサイズを減らした。

4.2.2. 実験結果

表 1 に損失関数の違いによる超解像の精度を示す。3 種類の損失関数の内、各評価項目で最良の結果を緑字で示している。

表 1 から、MSE Loss で学習すると、MSE の評価値が良い傾向にあった。これは、MSE Loss を最適化するための学習なため、想定通りであった。一方、SSIM の評価値は、SSIM Loss で学習したものよりも、MSE Loss + SSIM Loss で学習した方が優れている傾向にあった。SSIM Loss を良くするための学習より、MSE Loss も加えた学習の方が結果的に SSIM の評価値がよくなるということがわかった。また、PSNR の評価値については、Brain と Breast は MSE Loss での学習が一番よく、Lung は MSE Loss + SSIM Loss の方が良かった。

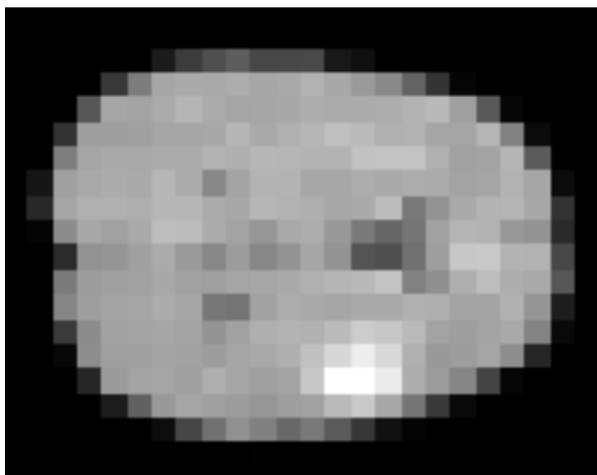
いずれの損失関数も、Lung の結果が Brain と Breast に比べると全ての評価値でかなり悪かった。原因としては、Brain と Breast は MRI データで、Lung だけ CT データであり、かつ、Lung の全体的なデータ数の割合が低かったことが考えられる。

図 9 に Brain の低解像度画像、高解像度画像、およびそれぞれの損失関数での超解像画像を示す。数値評価においては、MSE Loss + SSIM Loss の組み合わせが他の損失関数と比較して最も優れた精度を示していたが、視覚的な評価では、SSIM Loss を

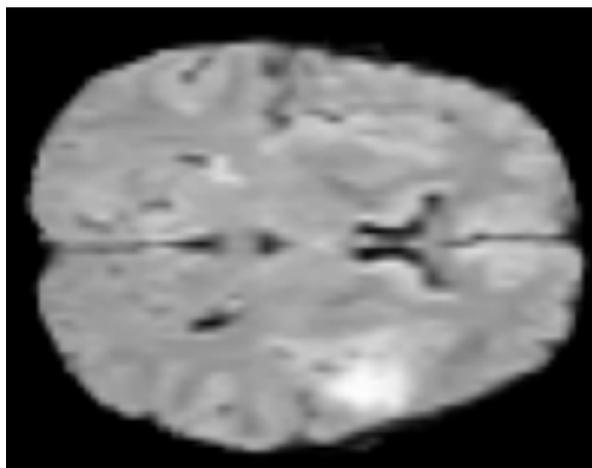
用いた方が他の損失関数に比べて明瞭な濃淡を持ち、ぼやけの表現も改善されていた。また、MSE Loss を用いた画像に比べて、他の損失関数を用いた画像の方が細部の再現がより優れていた。これにより、数値評価と視覚的評価が必ずしも一致しないことがわかった。

表 1 損失関数の違いによる超解像の定量評価

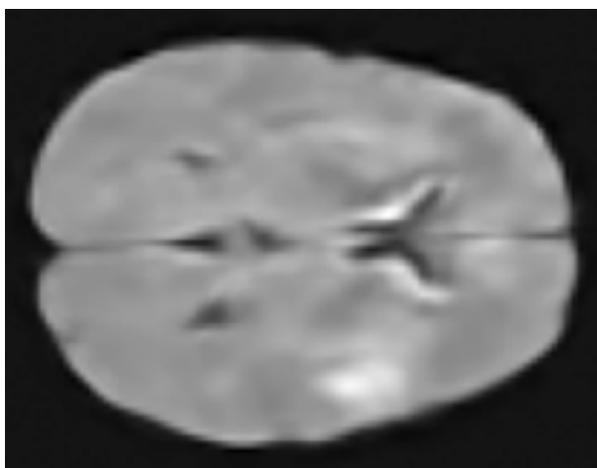
	Brain		
評価項目	MSE Loss	SSIM Loss	MSE Loss + SSIM Loss
MSE	0.0003482	0.0004301	0.0003778
PSNR	34.861	33.915	34.447
SSIM	0.9700	0.9711	0.9716
	Breast		
	MSE Loss	SSIM Loss	MSE Loss + SSIM Loss
MSE	0.002140	0.003357	0.002024
PSNR	30.898	30.076	30.532
SSIM	0.87034	0.8691	0.8836
	Lung		
	MSE Loss	SSIM Loss	MSE Loss + SSIM Loss
MSE	0.07750	0.07773	0.07613
PSNR	13.868	13.623	14.281
SSIM	0.5994	0.6027	0.6626



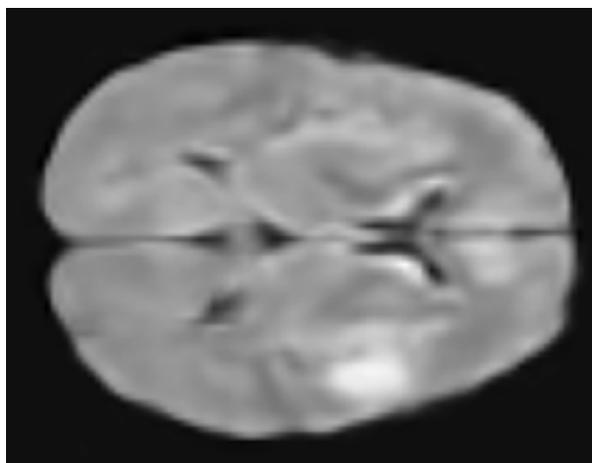
低解度像画像



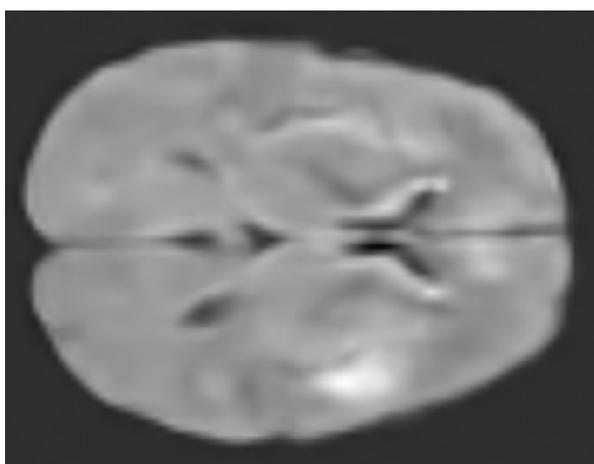
高解像度画像



MSE Loss



SSIM Loss



MSE Loss + SSIM Loss

図 9 Brain の抜粋画像

4.3. パッチングの有無による精度評価

パッチングは主に形状情報を扱うボクセルデータにはあまり効果的ではないが、画像のような分布を表現するデータには適している。その性質を踏まえると、ボリュームデータにも有効である可能性が高い。また、パッチングにはデータ拡張の効果があり、学習の多様性を高めることでモデルの精度向上が期待される。このため、パッチングの有効性を実験により評価する。

4.3.1. 実験内容

パッチングを使用する学習では、 $16 \times 16 \times 15$ の低解像度データと $64 \times 64 \times 60$ の高解像度データのペア、パッチングを使わない学習では、 $32 \times 32 \times 30$ の低解像度データと $128 \times 128 \times 120$ の高解像度データのペアで学習を行った。超解像時は、 $32 \times 32 \times 30$ の低解像度データを $128 \times 128 \times 120$ へ超解像した。バッチサイズは、11 で学習したが、メモリ容量の制約によりオーバーフローが発生した場合には、バッチサイズを減らした。

4.3.2. 実験結果

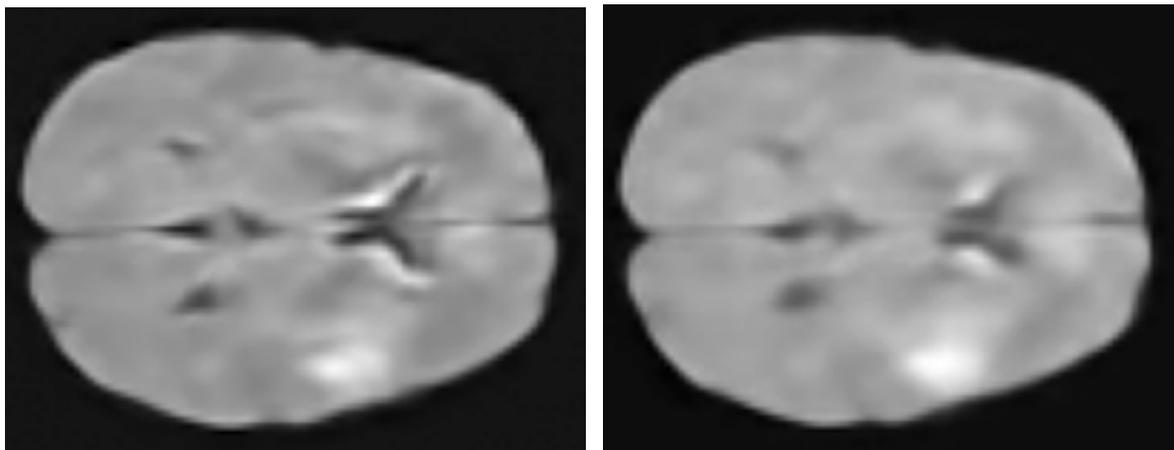
表3に、パッチングの有無の違いによる超解像の精度を示す。表3から、Brainはどの損失関数を使用したものでもパッチングがない方の超解像精度が優れている傾向だった。BreastとLungはパッチングをした方の超解像精度が良い傾向にあった。一方、Lungの数値は、4.2.2.節同様、他の部位より低くなっていた。

図10にBrainパッチングを使用した画像と使用しなかった画像をそれぞれ示す。図10より、パッチングを使用した方が明瞭な濃淡を持ち、ぼやけの表現が優れていることが示された。また、細部の再現性も向上していた。

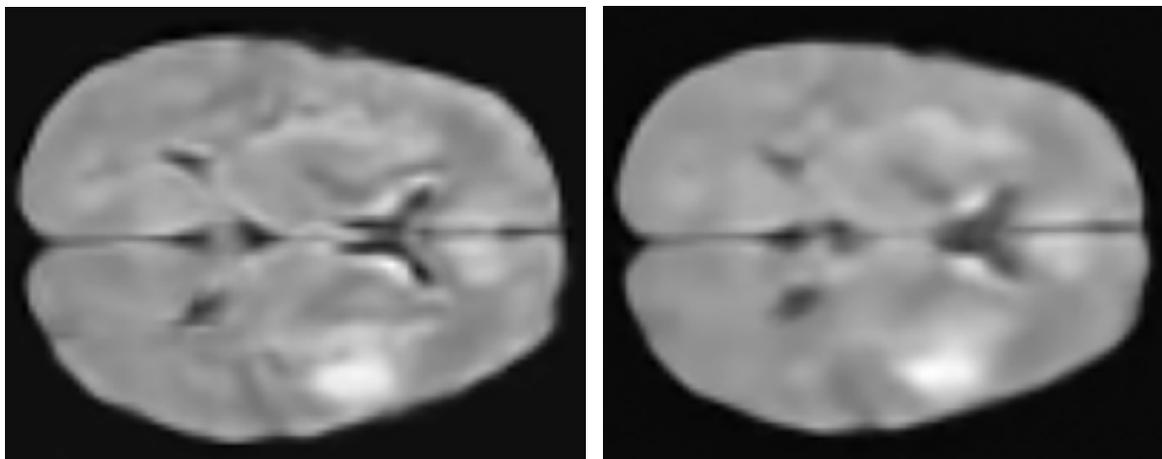
数値評価では、Brainはパッチングを使用しない方の精度が良かったが、視覚評価では、明らかにパッチング使用した方の精度が良かった。他の部位もパッチングを使用した方が数値評価と視覚評価の超解像精度が優れていることがわかった。

表3 パッチングの有無をまとめた定量評価

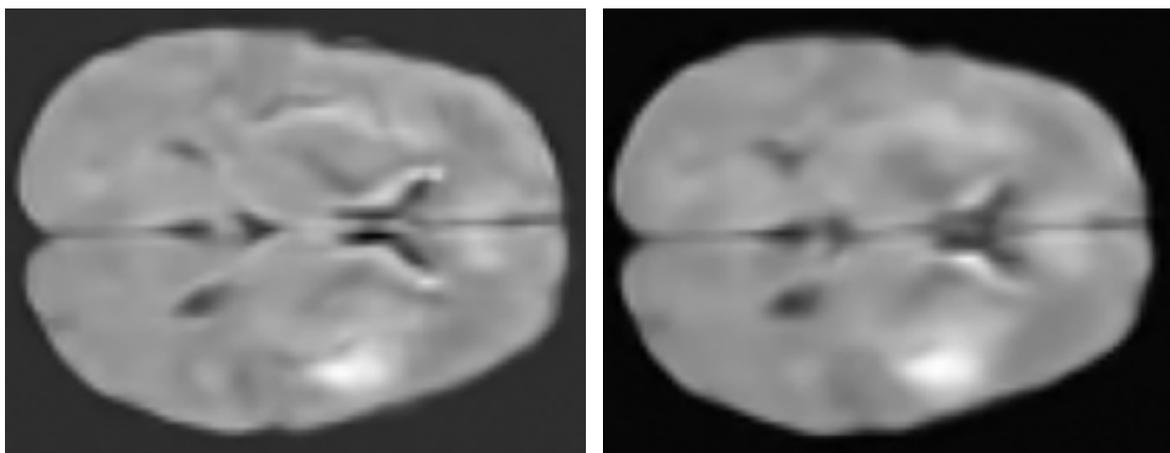
Brain			
評価項目	MSE Loss	SSIM Loss	MSE Loss + SSIM Loss
パッチング	あり/なし	あり/なし	あり/なし
MSE	0.0003482 / 0.0003261	0.0004301 / 0.0003701	0.0003778 / 0.0003648
PSNR	34.8617 / 35.167	33.915 / 34.618	34.447 / 34.647
SSIM	0.9700 / 0.9685	0.9711 / 0.9711	0.9716 / 0.9707
Breast			
	MSE Loss	SSIM Loss	MSE Loss + SSIM Loss
MSE	0.002140 / 0.002229	0.003357 / 0.002605	0.002024 / 0.002139
PSNR	30.898 / 31.225	30.076 / 29.605	30.532 / 29.248
SSIM	0.8703 / 0.8561	0.8691 / 0.8620	0.8836 / 0.8633
Lung			
	MSE Loss	SSIM Loss	MSE Loss + SSIM Loss
MSE	0.07750 / 0.08744	0.07773 / 0.07857	0.07613 / 0.07683
PSNR	13.868 / 12.169	13.623 / 13.310	14.281 / 13.863
SSIM	0.5994 / 0.4545	0.6027 / 0.5812	0.6626 / 0.6288



MSE Loss



SSIM Loss



MSE Loss + SSIM Loss

図 10 Brain の抜粋画像(左:パッチングあり画像、右:パッチングなし画像)

4.4. 3D-SRGAN と 3D-SRCGAN の精度評価

3D-SRGAN にクラス識別情報を加えたものが 3D-SRCGAN である。クラス識別情報はボクセルデータの超解像には効果的である一方、画像データの超解像にはあまり効果を発揮しない。しかし、本研究で取り扱うボリュームデータは Brain、Breast、Lung のそれぞれの臓器ごとに特徴の違いが顕著であるため、ボクセルデータと同様にクラス識別情報が有効か実験により評価する。

4.4.1. 実験内容

3D-SRCGAN の学習と、3D-SRGAN での学習では、3 クラスまとめて学習したものと、1 クラスずつ学習したものの合計 3 パターンがある。バッチサイズは、3D-SRGAN で 3 クラスまとめた学習が 16 で、3D-SRGAN で 1 クラスずつでの学習が 14 である。学習時には、パッチングを使用した。また、損失関数は SSIM Loss を使用した。

4.4.2. 実験結果

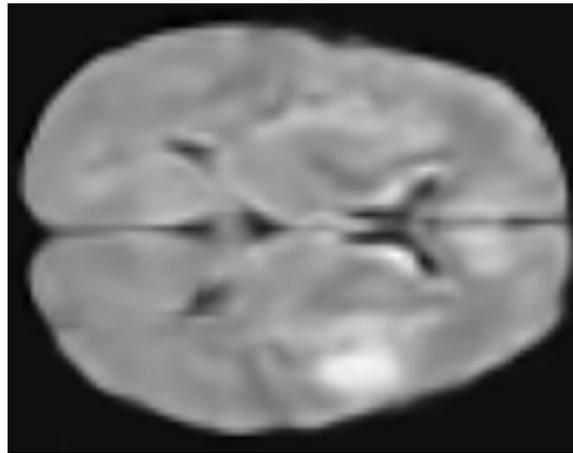
表 4 に、3D-SRGAN(3 クラスと 1 クラス)と 3D-SRCGAN の精度評価を示す。表 4 より、Brain は、3D-SRCGAN の方が他のモデルに比べて超解像精度が良かった。一方、Breast と Lung は、3 クラスまとめて学習した 3D-SRGAN の方の超解像精度が良かった。1 クラスずつ学習した 3D-SRGAN は全体的に超解像精度が低かった。

図 11 にパッチングを使用した Brain の 3D-SRCGAN と 3D-SRGAN の抜粋画像を示す。図 11 より、どのモデルでも、濃淡や細部の再現、ぼやけなどの見た目の優位差が見られなかった。

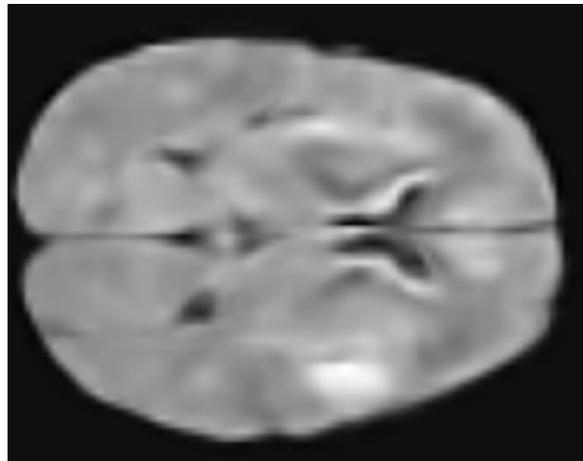
視覚評価より、クラス識別情報が有効ではないことが明らかになった。このことより、ボリュームデータはボクセルデータよりも画像の特徴に近い性質を持つことが示唆された。

表 4 パッチングを使用した 3D-SRGAN と 3D-SRCGAN の定量評価(SSIM Loss)

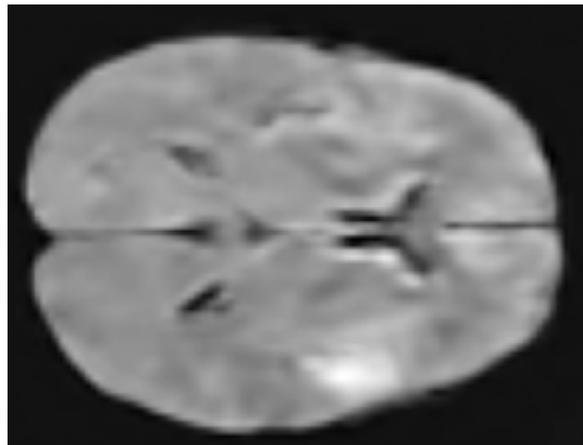
	Brain		
評価項目	3D-SRCGAN	3 クラス 3D-SRGAN	1 クラス 3D-SRGAN
MSE	0.0004301	0.0004341	0.001094
PSNR	33.915	33.848	29.683
SSIM	0.9711	0.9709	0.9338
	Breast		
MSE	0.003357	0.003094	0.005116
PSNR	30.076	30.053	23.111
SSIM	0.8691	0.8719	0.4717
	Lung		
MSE	0.07773	0.07627	0.07616
PSNR	13.623	14.243	14.005
SSIM	0.6027	0.6597	0.6648



3D-SRCGAN



3D-SRGAN(3クラスまとめて)



3D-SRGAN(1クラスずつ)

図 11 パッチングありの Brain の各モデル超解像の抜粋画像(SSIM Loss)

5. おわりに

本研究は、ボクセルデータの超解像を目的として提案された 3D-SRCGAN を、損失関数、パッチング、クラス識別情報の 3 つの観点から分析し、ボリュームデータへの適用可能性を検証した。

まず、ボクセルデータ超解像では、損失関数は MSE のみで使用されていたという背景から、ボリュームデータに対して、MSE Loss、SSIM Loss、MSE Loss + SSIM Loss の 3 つの損失関数を用いて超解像の精度向上に関係するか評価した。その結果、数値的な評価としては、MSE Loss + SSIM Loss の組み合わせが他の損失関数と比較して最も優れた精度を示していた。視覚的な評価では、SSIM Loss を用いた方が他の損失関数に比べて明瞭な濃淡を持ち、ぼやけの表現も改善されていた。一方、MSE Loss を用いた画像に比べ、他の損失関数を用いた画像の方が細部の再現がより優れていた。これらより、SSIM を損失関数として使用することへの有効性が示された。

次に、パッチングは、画像の超解像時に一般的に使用されるが、ボクセルデータの超解像では有効でなかった。これらを受けて、パッチングがボリュームデータに対して有効か評価した。その結果、数値的な評価としては、Brain を除いて、Breast と Lung ではパッチングをした方の超解像精度が良かった。また、視覚的な評価では、全ての部位でパッチングをした方が明瞭な濃淡を持ち、ぼやけの表現が優れていることが示された。また、細部の再現性も向上していた。これらより、ボリュームデータの超解像ではパッチングが有効だと示された。また、ボリュームデータは、パッチングが有効だったため、ボクセルデータより画像に近い特性があると考えられる。

最後に、クラス識別情報は、ボクセルデータの超解像では有効で、画像データでは有効でなかった。本研究では、ボリュームデータが臓器の部位ごとに特徴が大きく異なることから、クラス識別情報が有効かを評価した。その結果、数値的な評価では、Brain のみクラス識別情報ありの超解像精度が高く、Breast、Lung は、クラス識別情報がない方の超解像精度が高かった。視覚的な評価においては、クラス識別情報の有無では差がなかった。これらより、ボリュームデータにおいてクラス識別情報は大きくは影響しないことがわかった。

今後の課題として、医師の協力のもと超解像結果を確認してもらうことが挙げられる。超解像が数値的によくなったとしても、実際に医師がデータを見て読影や診断に有効かを確認する必要がある。また、本研究では Lung の超解像精度が Brain や Breast に比べて低い結果となった。これは、Lung が CT データであり、使用したデータ量が他と比較して少なかったことが原因と考えられる。一方、Brain や Breast は MRI データで構成されており、十分なデータ量が確保されていた。そのため、CT データを十分に収集し、超解像精度の向上を図ることが今後の課題である。

謝辞

本研究を進めるにあたり、指導教員である棕木雅之教授に大変感謝致します。お忙しいなか、研究の実験のアイデアや論文に沢山のアドバイスをいただきました。また、棕木研究室の皆様には、本研究を進めるにあたり多くのご助言、ご協力を頂いたことを感謝致します。

参考文献

- [1] 野村淳也, 椋木雅之, “クラス識別を導入した 3D-SRGAN によるボクセル超解像”, 令和 5 年度宮崎大学大学院工学研究科 修士論文(2024)
- [2] 岡和寿, 椋木雅之, “SRGAN の 3 次元ボクセルモデル超解像への適用”, 画像電子学会 論文誌, vol. 48, no. 4, pp. 448-496 (2019)
- [3] C Ledig, L Theis, F Huszar, J Caballero, A Aitken, A Tejani, J Totz, Z Wang, W Shi, “Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network,” IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 4681-4690 (2017).
- [4] Li, Y., Iwamoto, Y., Lin, L., Xu, R., Tong, R., & Chen, Y. W. “VolumeNet: A lightweight parallel network for super-resolution of MR and CT volumetric data”, IEEE Transactions on Image Processing, vol. 30, pp. 4840-4854.(2021)
- [4] 池谷彰彦, 広明敏彦, “超解像ソリューション (組込みソフトウェア・ソリューション特集)--(イメージ/音声処理コンポーネントソリューション)”, NEC 技報, vol. 60, no.2, pp. 24-26 (2007)
- [5] H Chang, D-Y Yeung, Y Xiong, “Super-resolution through neighbor embedding”, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 275-282 (2004)
- [6] D Khaledyan, A Amirany, K Jafari, MH Moaiyeri, AZ Khuzani, N Mashhadi, “Lowcost implementation of bilinear and bicubic image interpolation for real-time image super-resolution”, IEEE Global Humanitarian Technology Conference (GHTC), pp.1-5 (2020)
- [7] WT Freeman, TR Jones, EC Pasztor, "Example-based super-resolution", IEEE Computer Graphics and Applications (CGA), vol. 22, no. 2, pp. 56-65 (2002)
- [8] C Dong, CC Loy, K He, X Tang, “Image super-resolution using deep convolutional networks”, IEEE Transactions on Pattern Analysis and Machine Intelligence(TPAMI), vol. 38, no. 2, pp. 295-307(2015)
- [9] Qiao, J., Song, H., Zhang, K., Zhang, X., & Liu, Q. Image super - resolution using conditional generative adversarial network. IET Image Processing, vol.13 no. 14, pp.

2673-2679. (2019).

[10]J Wang, Y Chen, Y Wu, J Shi, J Gee, “Enhanced generative adversarial network for 3D brain MRI super-resolution”, IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), pp. 3627-3636 (2020)

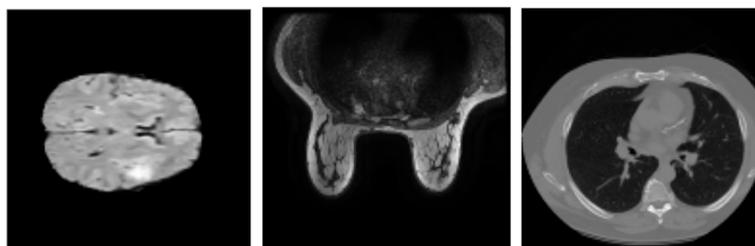
[11]J Hatvani, A Basarab, JY Tournieret, M Gyöngy, D Kouamé, “A tensor factorization method for 3-D super resolution with application to dental CT”, IEEE Transactions on Medical Imaging (TMI), vol. 38, no. 6, pp. 1524-1531 (2018)

[12] Medical Decathlon, <http://medicaldecathlon.com/dataaws/>

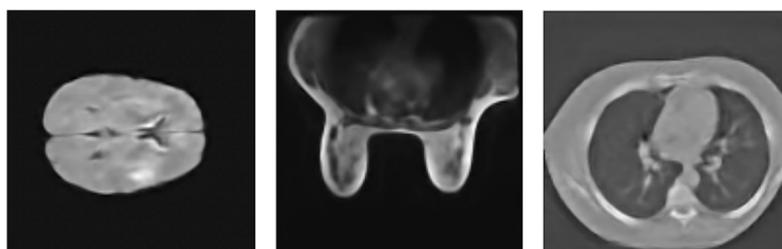
[13] Cancer Imaging Archive, <https://www.cancerimagingarchive.net/>

付録 A. 超解像結果の画像例

図 A.1~A.5 に各手法で超解像した画像の例を示す。



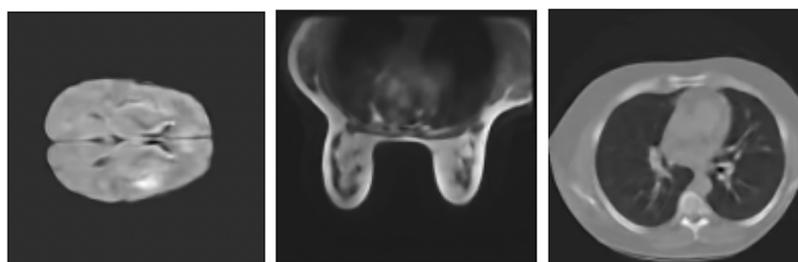
高解像度画像



MSE Loss

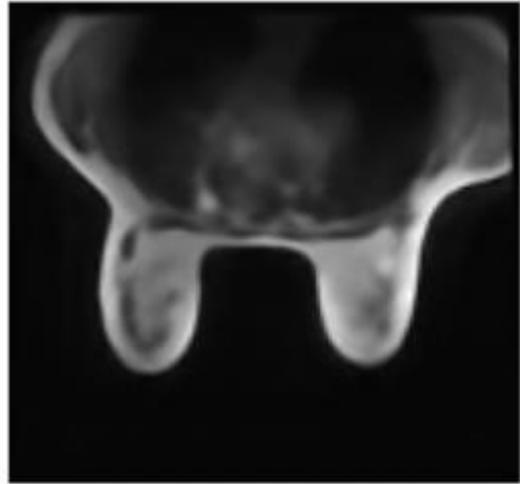
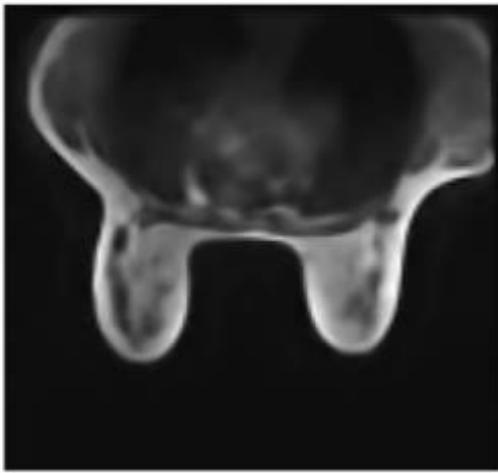


SSIM Loss

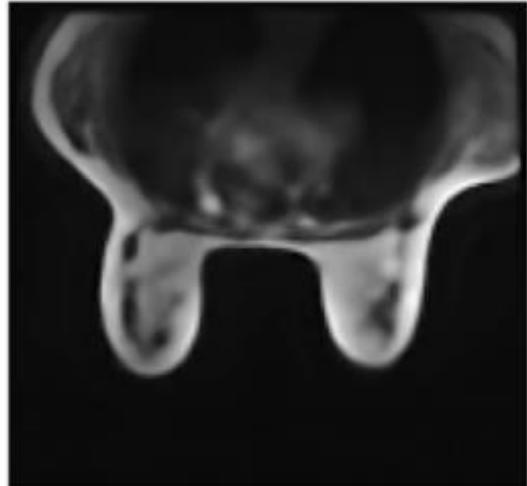
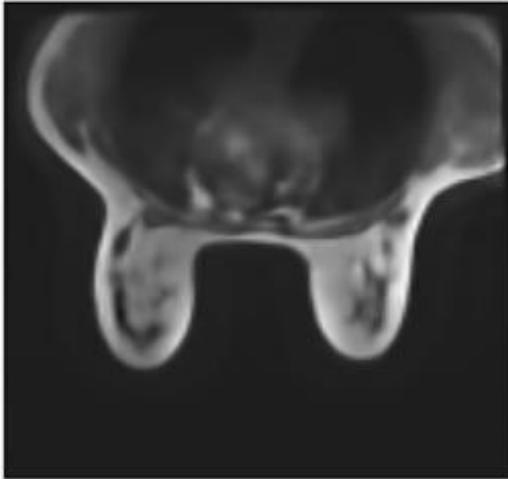


MSE Loss + SSIM Loss

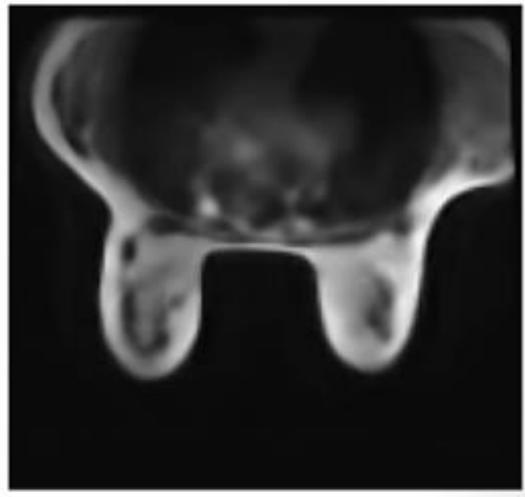
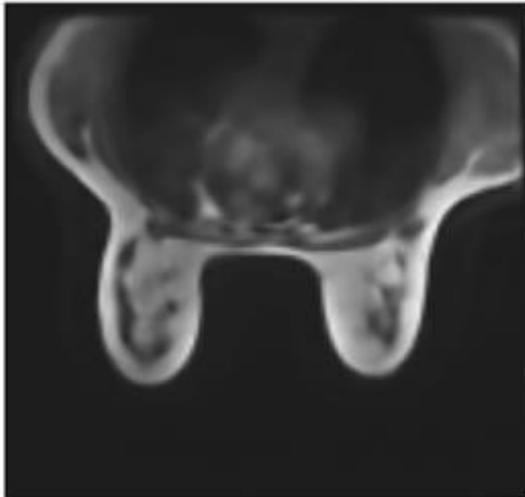
A.1 Brain、Breast、Lung の抜粋画像(パッチングあり)



MSE Loss

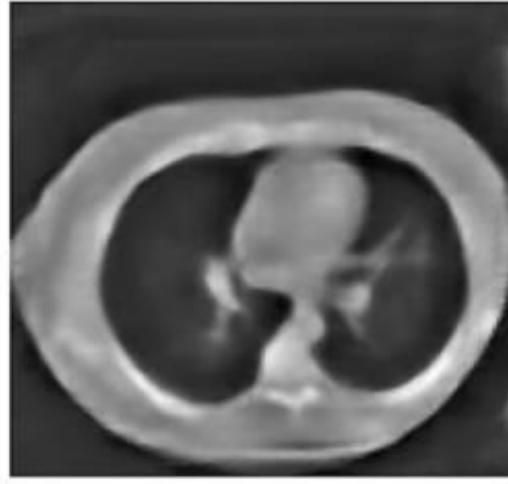
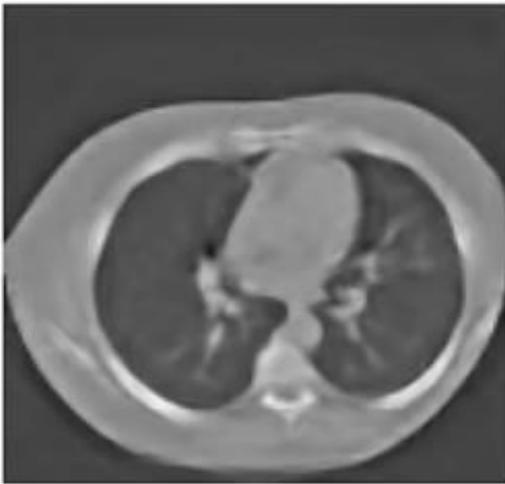


SSIM Loss

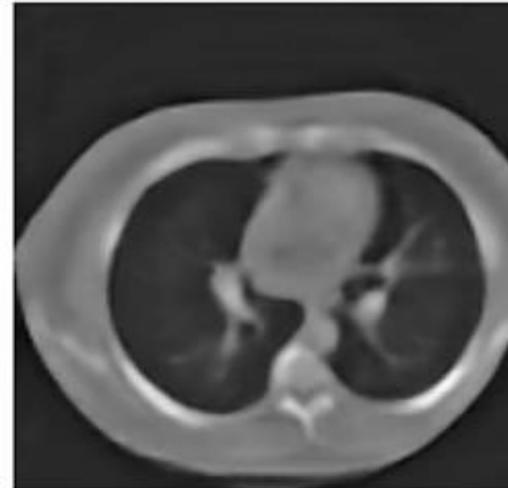
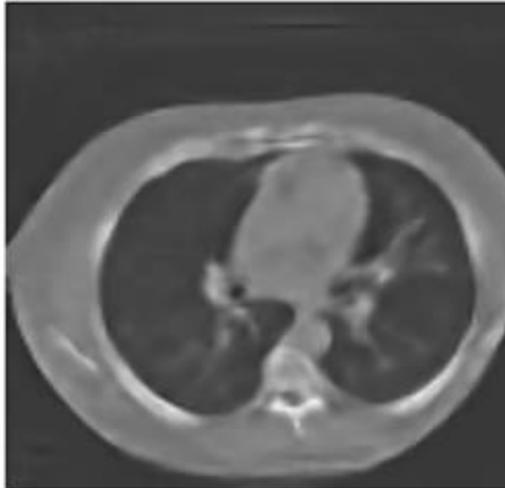


MSE Loss + SSIM Loss

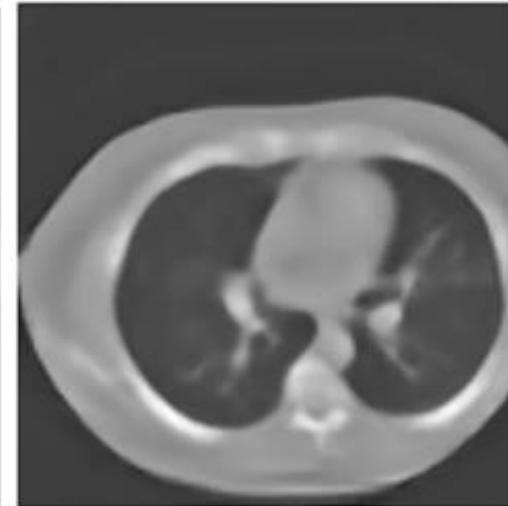
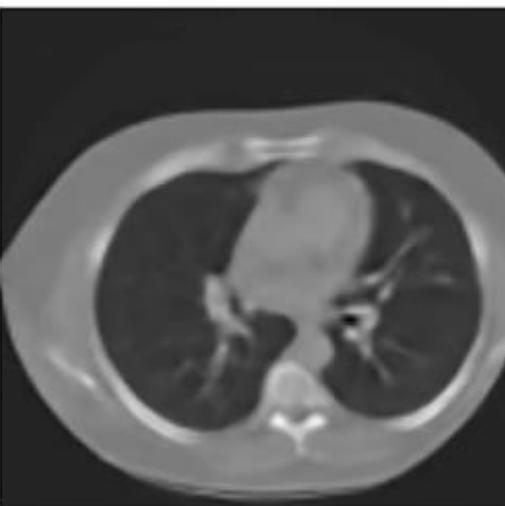
A.2 Breast の画像(左:パッチングあり、右:パッチングなし)



MSE Loss



SSIM Loss



MSE Loss + SSIM Loss

図 A.3 Lung の画像(左:パッチングあり、右:パッチングなし)

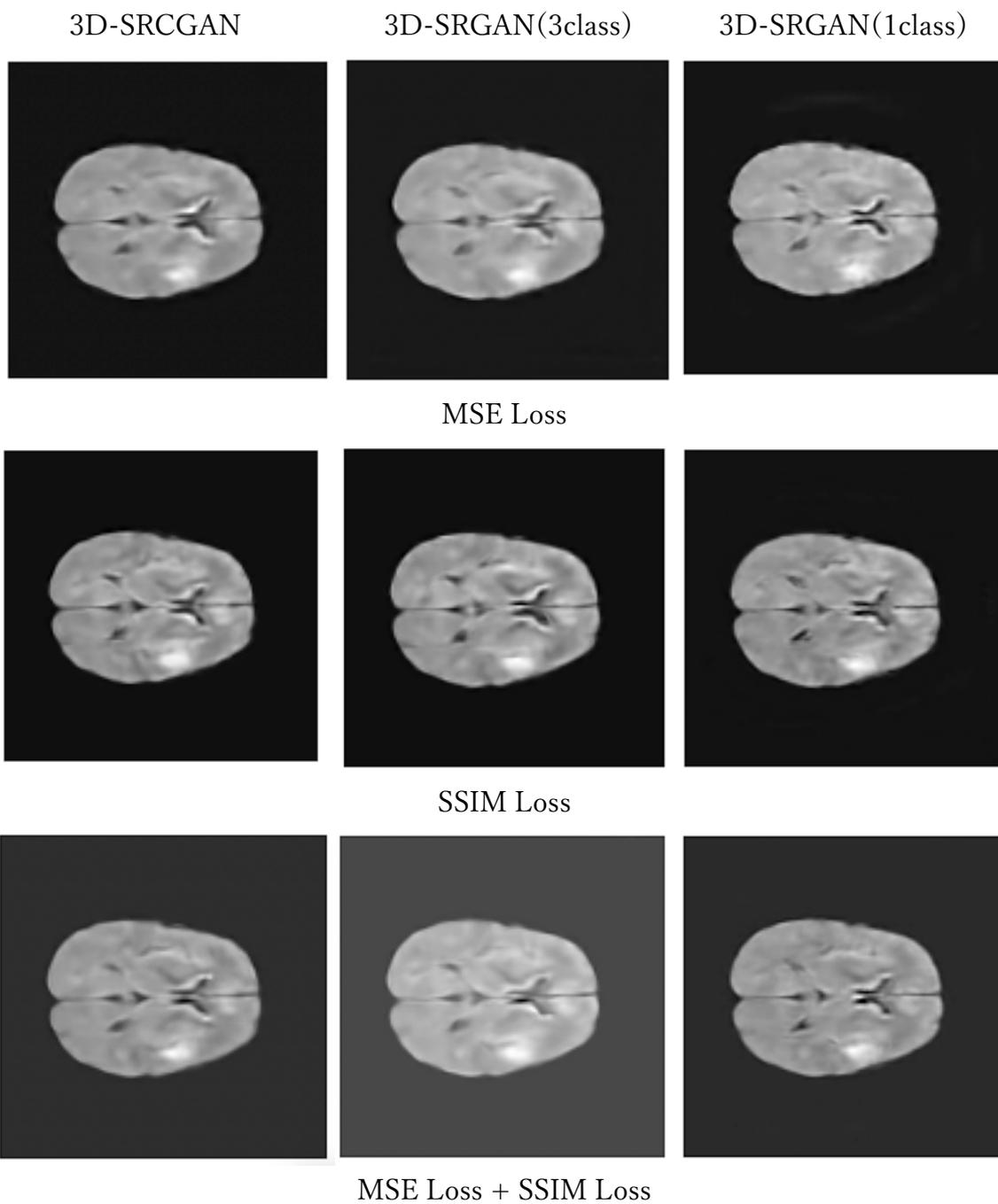


図 A.4 パッチングを使用した Brain の 3D-SRCGAN と 3D-SRGAN の抜粋画像

3D-SRCGAN 3D-SRGAN(3class) 3D-SRGAN(1class)

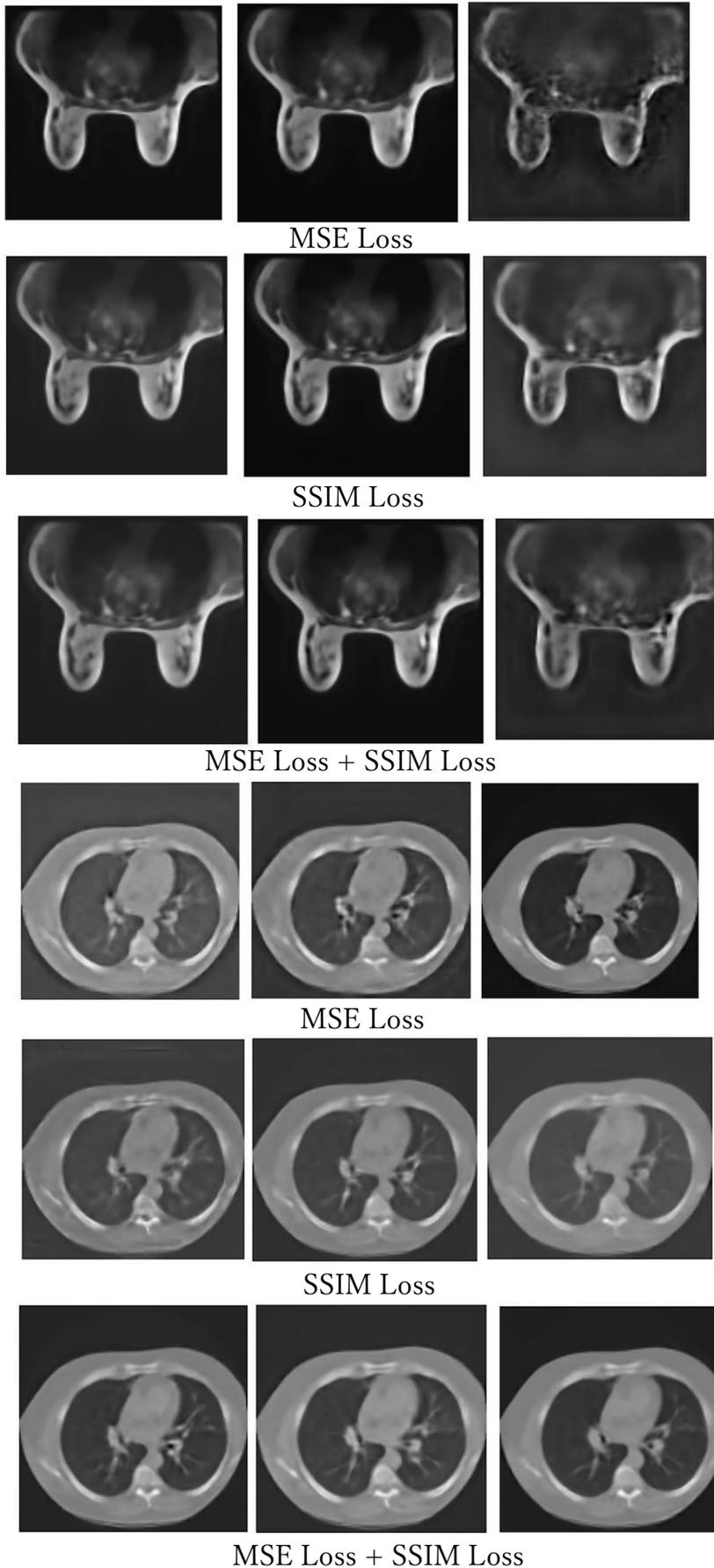


図 A.5 パッチングを使用した Breast、Lung の画像

付録 B. 数値評価結果

表 B.1~B.7 に各手法の数値評価結果の詳細を示す。

表 B.1 パッチングを使用した損失関数の違いによる超解像の定量評価

	Brain		
評価項目	MSE Loss	SSIM Loss	MSE Loss + SSIM Loss
MSE	0.0003482	0.0004301	0.0003778
PSNR	34.8617	33.915	34.4473
SSIM	0.9700	0.9711	0.9716
	Breast		
	MSE Loss	SSIM Loss	MSE Loss + SSIM Loss
MSE	0.002140	0.003357	0.002024
PSNR	30.898	30.076	30.5322
SSIM	0.87034	0.8691	0.88366
	Lung		
	MSE Loss	SSIM Loss	MSE Loss + SSIM Loss
MSE	0.07750	0.07773	0.076134
PSNR	13.868	13.623	14.28143
SSIM	0.59948	0.6027	0.66267

表 B.2 パッチングを使用しない損失関数の違いによる超解像の定量評価

	Brain		
評価項目	MSE Loss	SSIM Loss	MSE Loss + SSIM Loss
MSE	0.00032610	0.00037014	0.0003648
PSNR	35.1678	34.6189	34.647
SSIM	0.9685	0.9711	0.9707
	Breast		
	MSE Loss	SSIM Loss	MSE Loss + SSIM Loss
MSE	0.002229	0.0026054	0.002139
PSNR	31.225	29.6051	29.248
SSIM	0.8561	0.8620	0.8633
	Lung		
	MSE Loss	SSIM Loss	MSE Loss + SSIM Loss
MSE	0.087442	0.078578	0.07683
PSNR	12.169	13.3101	13.863
SSIM	0.4545	0.5812	0.6288

表 B.3 パッチングを使用した 3D-SRGAN と 3D-SRCGAN の定量評価(MSE Loss)

	Brain		
評価項目	3D-SRCGAN	3 クラス 3D-SRGAN	1 クラス 3D-SRGAN
MSE	0.0003482	0.0003367	0.0003545
PSNR	34.8617	35.0189	34.799
SSIM	0.9700	0.9697	0.94522
	Breast		
MSE	0.002140	0.0019863	0.015024
PSNR	30.898	31.1086	18.3310
SSIM	0.87034	0.87569	0.31102
	Lung		
MSE	0.07750	0.077203	0.07569
PSNR	13.868	13.920	14.6544
SSIM	0.59948	0.60675	0.68441

表 B.4 パッチングを使用した 3D-SRGAN と 3D-SRCGAN の定量評価(MSE Loss + SSIM Loss)

	Brain		
評価項目	3D-SRCGAN	3 クラス 3D-SRGAN	1 クラス 3D-SRGAN
MSE	0.0003778	0.0003933	0.0008578
PSNR	34.4473	34.2694	30.762
SSIM	0.9716	0.97174	0.9622
	Breast		
MSE	0.002024	0.002050	0.005233
PSNR	30.5322	30.2175	23.8654
SSIM	0.88366	0.88279	0.4661
	Lung		
MSE	0.076134	0.07628	0.07625
PSNR	14.28143	14.2402	14.2703
SSIM	0.66267	0.6556	0.66554

表 B.5 パッチングを使用しない 3D-SRGAN と 3D-SRCGAN の定量評価(MSE Loss)

	Brain		
評価項目	3D-SRCGAN	3 クラス 3D-SRGAN	1 クラス 3D-SRGAN
MSE	0.00032610	0.0003716	0.00031858
PSNR	35.1678	34.517	35.2947
SSIM	0.9685	0.8996	0.96654
	Breast		
MSE	0.002229	0.001790	0.0018697
PSNR	31.225	31.271	31.3072
SSIM	0.8561	0.8629	0.85908116
	Lung		
MSE	0.087442	0.07662	0.077996
PSNR	12.169	14.276	13.69254
SSIM	0.4545	0.6160	0.601042

表 B.6 パッチングを使用しない 3D-SRGAN と 3D-SRCGAN の定量評価(SSIM Loss)

	Brain		
評価項目	3D-SRCGAN	3 クラス 3D-SRGAN	1 クラス 3D-SRGAN
MSE	0.00037014	0.0008982	0.013757
PSNR	34.6189	30.557	19.2115
SSIM	0.9711	0.9456	0.84737
	Breast		
MSE	0.0026054	0.002706	0.0031593
PSNR	29.6051	29.102	28.9910
SSIM	0.8620	0.8634	0.84187
	Lung		
MSE	0.078578	0.07604	0.079350
PSNR	13.3101	14.423	12.46325
SSIM	0.5812	0.6616	0.58735

表 B.7 パッチングを使用しない 3D-SRGAN と 3D-SRCGAN の定量評価(MSE Loss + SSIM Loss)

	Brain		
評価項目	3D-SRCGAN	3 クラス 3D-SRGAN	1 クラス 3D-SRGAN
MSE	0.0003648	0.0004692	0.00037216
PSNR	34.647	33.402	34.6179
SSIM	0.9707	0.9662	0.96807
	Breast		
MSE	0.002139	0.002038	0.002198
PSNR	29.248	30.494	30.01875
SSIM	0.8633	0.8710	0.86169
	Lung		
MSE	0.07683	0.07601	0.077726
PSNR	13.863	14.469	13.0268
SSIM	0.6288	0.6637	0.608761