

2023年度 修士論文

論文題目

生成モデルによるリンパ節転移有無の分類精度向上手法  
(Improvement Method for Classification of Lymph Node  
Metastasis using Generative Model)

指導教員

舟橋 健司 准教授

名古屋工業大学大学院 工学研究科 工学専攻  
情報工学系プログラム 2022年度入学 34414060 番

鈴木 康介

# 目次

<b>第1章</b>	<b>はじめに</b>	<b>1</b>
<b>第2章</b>	<b>背景知識</b>	<b>4</b>
2.1	造影 CT 画像	4
2.2	DICOM 画像	4
2.3	リンパ節のがん転移有無による特徴の変化	5
<b>第3章</b>	<b>データセット</b>	<b>7</b>
3.1	データの概要	7
3.2	マスク画像の作成	7
3.3	データセットの前処理	9
<b>第4章</b>	<b>超解像処理</b>	<b>11</b>
4.1	予備実験	11
4.1.1	追加情報	11
4.1.2	リンパ節分類実験	15
4.2	超解像モデルの作成と評価	16
4.3	リンパ節分類実験	17
<b>第5章</b>	<b>画像生成によるデータ拡張</b>	<b>22</b>
5.1	画像生成モデルの作成と評価	22
5.2	リンパ節分類実験	23
5.3	Vision Transformer でのリンパ節分類実験	26
<b>第6章</b>	<b>むすび</b>	<b>31</b>
	謝辞	32
	参考文献	33
	発表論文リスト	37

## 第1章 はじめに

がん診断の現場では、X線CTをはじめとする画像診断モダリティが用いられている。X線CTとは、人体の周囲からX線を照射し、X線の透過率の違いを利用して人体の断面を画像化する技術である。近年ではCT技術の発展によりCT画像のスライス間隔が小さくなったことで、一度の検査で大量のデータが得られるようになった。それによって微小な病変や早期の病変の発見が可能となった。一方で、診断する医師の負担の増加や、従来の診断方法では大量のデータを活用しきれないという問題が発生している。また、画像診断は専門性が高いため、医師の経験量によって正診率にばらつきが生じる。例えばリンパ節矩形を測定するという診断では、医師の経験量の違いによって正診率が51-86%と大きくばらつきがある。これらの問題の解決策として、機械学習による診断支援について盛んに研究されており、CT画像からリンパ節を検出する研究 [1-4] や、CT画像をMRI画像へ変換する研究 [5] などがある。

本研究では、直腸がんのリンパ節転移について取り扱う。直腸がんとは、直腸の組織内に悪性がん細胞が認められる病気である。1975年以降、日本では直腸がんの罹患数・死亡数ともに増加を続けている [6]。そのため、特に機械学習による診断支援手法の確立が期待されている病気の1つである。がんはリンパの流れに沿ってリンパ節に転移を生じる可能性があるため、直腸がんと診断された場合はX線CT等での画像診断によって、他の部位に転移していないかどうかを調べる病期診断という診断を行う。転移の有無が特に重要となるのが、側方リンパ節と呼ばれる内腸骨動脈に沿うリンパ節および閉鎖神経と閉鎖動脈周囲のリンパ節である。直腸がんの側方リンパ節転移が陽性の場合の5年生存割合は40%前後であり、転移が陰性の場合の80%と比べて予後不良であるとされている [7,8]。直腸がんは治癒切除後でも再発する危険性が高いが、がん周辺のリンパ節だけでなく側方リンパ節も切除 (側方

リンパ節郭清) することで骨盤内再発のリスクが50%減少し、5年生存割合が8-9%改善すると予測されている。しかし、すべての症例に対して側方リンパ節郭清を行えば良いというわけではない。側方リンパ節郭清には排尿機能障害や性機能障害などの後遺症が残る可能性があるためである。そこで、側方リンパ節への転移が陰性である場合には、再発のリスクよりも側方リンパ節郭清による後遺症のリスクを考慮し、側方リンパ節郭清を省略してもよいとされている [9]。したがって、手術前に側方リンパ節への転移有無を正確に診断し、側方リンパ節郭清が必要であるかどうかを検討することが重要である。しかし、側方リンパ節には確立した診断手法がなく、新たな診断手法の研究が望まれている。

近年、機械学習によるCT画像からのリンパ節分類の研究は多く行われている。Wangらの研究では非小細胞肺癌の縦隔リンパ節転移分類の性能についてCNNと古典的手法、医師の診断を比較しており、その結果ではCNNと古典的手法や医師の診断の間に有意な差はなかったものの、便利さ等の観点でCNNの将来性があるとされている [10]。また、Hartensteinらは前立腺がんの病期分類のためのCNNを提案し、専門の医師と同等の性能であることを示している [11]。このような研究において課題となるのが、データセットの入手である。CT画像を学習に使用するには専門の医師によるアノテーションが必要であるため、大量のデータセットを構築することが困難である。そのような場合、複雑な構造のモデルの学習が十分に進まないことが先行研究で示唆されている [12]。したがって、少量のデータセットから効率よく学習を行う工夫が重要となる。そこで、近年急速に発展をしている生成モデルを用いてデータセットの量の少なさを補う手法について検討した。生成モデルとは、入力に対して何らかの確率を出力する識別モデルと異なり、新たなデータを出力することができるニューラルネットワークモデルである。本研究では敵対的生成ネットワーク (GAN) [13] をベースとした超解像手法と、Diffusionモデルを用いたデータ拡張手法を提案する。

CT画像の超解像は微小な病変の発見に役立つことから注目されている分野である。また、低解像度で撮影したCT画像から診断に十分な解像度のCT画像を復元することができれば、患者の放射線の被曝量を減らすことにも繋がる。例えば、Youらの研究ではGAN-CIRCLEというGANベースのモデルを用いることで、従来の

手法よりも精度の高い超解像が可能であることを、画像評価指標および専門家による視覚評価により示している [14]. しかし、従来の CT 画像の超解像の研究は最終的に医師が画像診断をすることを目的としており、超解像後の画像を別の分類用モデルに入力した場合の影響に言及したものは少ない. そこで、4 章では従来の超解像手法と生成モデルによる超解像の性能とその分類精度への影響を比較し、生成モデルによる高精度の超解像が分類精度向上に効果的であることを確認した.

画像生成は生成モデルの活用方法として特に注目されており、医用画像分野では症例が少ないデータの拡張への活用が期待されている. 例えば、Mann らは GAN を用いて胸部 CT スキャン画像を生成し、COVID-19 の学習のためのデータを生成する手法を提案している [15]. また、Liu らは Mask Attention GAN というモデルを用いて、画像内の肝腫瘍のピクセルをアテンションマップとしてマスクすることで、腫瘍のある合成 CT 画像を生成する手法を提案している. また、それらを腫瘍セグメンテーションタスクの学習に適用することで、データ拡張に有効であることを示している [16]. これらの研究のように画像生成には GAN が用いられることが多いが、近年では Diffusion モデルを用いることで GAN よりも多様性のある画像生成が可能であることが注目されている. しかし、現時点では Diffusion モデルによる画像生成を医用画像に適用して、そのデータ拡張としての有効性を確認するところまで一貫して行った研究は少ない. 5 章では Diffusion モデルを用いたリンパ節 CT 画像の生成手法を提案した. さらに、生成画像をデータ拡張として用いることによる効果を、CNN と Vision Transformer の学習を実際に行うことにより確認した.

## 第2章 背景知識

本章ではCT画像とリンパ節に関する背景知識として、造影CT画像とDICOM画像、リンパ節のがん転移有無による特徴の変化について説明する。

### 2.1 造影CT画像

CT画像を撮影する際に、医師が診断をしやすくするために造影剤と呼ばれる薬剤を使用する場合がある。造影剤を撮影前に静脈に注射することでCT画像の血管や臓器の部分にコントラストが生じるため、そのままではわかりにくい小さな病変も見つけやすくする効果がある。造影剤を使わずに撮影する方法を単純CTと呼び、造影剤を使用して撮影する方法を造影CTと呼ぶ。造影剤を使うことによるCT画像の違いを図2.1に示す。造影剤を使用していなければリンパ節を見つけ出すことは困難とされており、本研究で使用するデータセットも造影CTで撮影されたものである。

### 2.2 DICOM画像

DICOMとは医用画像で一般的に使用されている形式である。デジタル画像の形式の多くが8ビットの階調で表現するのに対し、DICOMは10ビット以上の階調で表現することができる。医師がCT画像による診断を行う際には、ウィンドウレベル変換という処理を行う。ウィンドウレベル変換とは、10ビット以上の画素値を持つCT画像から、特定の値域のみを抽出して8ビットのグレースケール画像に変換する処理である。ウィンドウレベル変換で切り出す階調の中心値と幅をそれぞれウィンドウレベルとウィンドウ幅という。ウィンドウレベルとウィンドウ幅を適切に変更することで特定の組織のみを表示することができ、診断がしやすくなる。

### 2.3 リンパ節のがん転移有無による特徴の変化

リンパ節にがん細胞が転移しているときとしていないときで変化する特徴がある。それは模様・形状・大きさである。通常のリンパ節は一様な模様で、アーモンドのような楕円体の形状をしており、基本的に10mm以下の大きさである。それに対し、転移したリンパ節はまばらな模様になり、腫れ上がることで球形になり、10mm以上の大きさになるとされている [17,18]。模様の違いは造影剤によるものであり、転移したがん細胞が染まる際の障害となってまばらな模様になる。図2.2, 2.3にそれぞれの転移状態のリンパ節を示す。

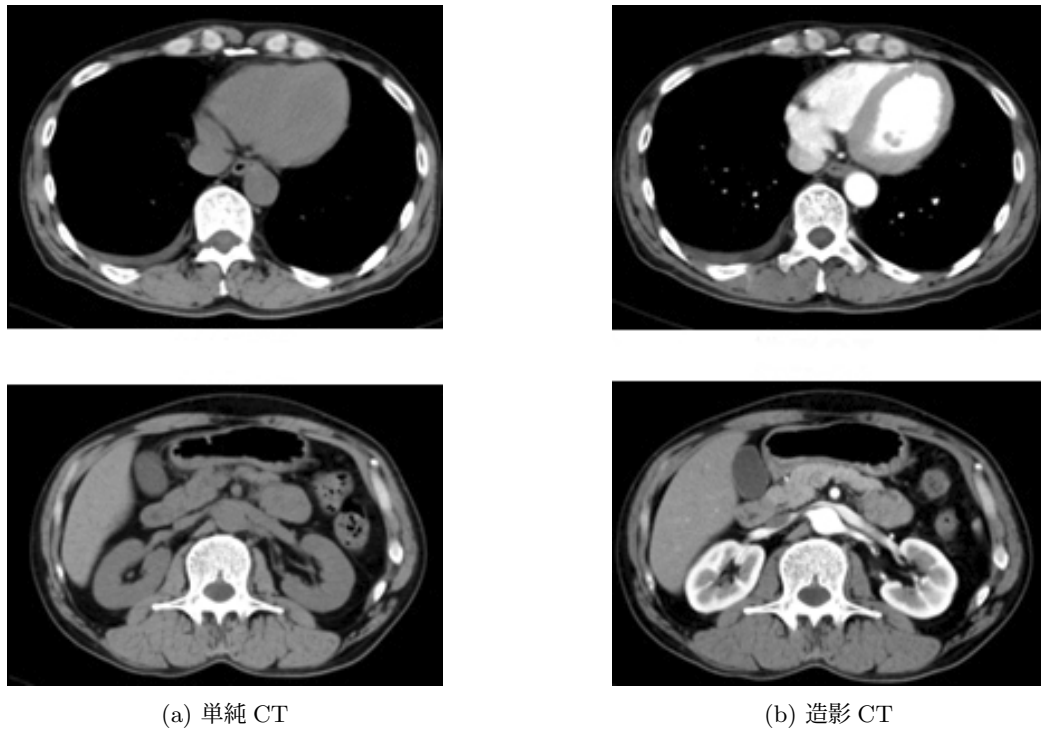


図 2.1: 造影剤による CT 画像の違い



図 2.2: 転移なしリンパ節の例 (画像中央)

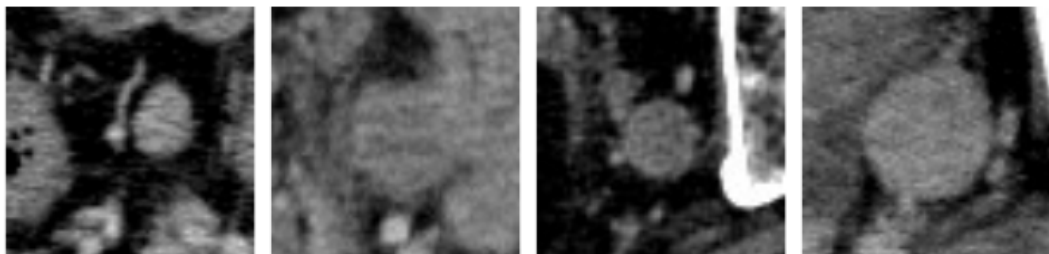


図 2.3: 転移ありリンパ節の例 (画像中央)



## 第3章 データセット

本研究では愛知県がんセンターから提供された54人分のCT画像を使用する。そのデータの概要と、前処理について説明する。

### 3.1 データの概要

提供されたデータは、患者一人当たり500から800枚のスライス、 $512 \times 512$ ピクセルのDICOM形式CT画像と、1人の医師によってリンパ節の位置とがん転移の有無がアノテーションされたPowerPointプレゼンテーションファイル（以下、pptxファイルと呼ぶ）の組になっている。リンパ節は複数のスライスにわたって写っていることが多いが、リンパ節の中心が写っている1スライスのみのアノテーションされている。このpptxファイルの例を図3.1に示す。合計で712個のリンパ節がアノテーションされており、内訳はがん転移なしリンパ節（以下、G1と呼ぶ）が660個、がん転移ありリンパ節（以下、G3と呼ぶ）が52個である。

### 3.2 マスク画像の作成

前述のpptxファイルはプログラムでの処理に不向きであるため、pptxファイルを元にマスク画像の作成を行う。マスク画像作成の手順は以下の通りである。CT画像と対応するマスク画像の例を図3.2, 3.3に示す。

1. CT画像をDICOM形式からTIFF形式に変換する。
2. 画像編集ソフトを用いてpptxファイルでアノテーションされているリンパ節を、G1は $(R,G,B)=(0,255,0)$ 、G3は $(R,G,B)=(255,0,0)$ で塗りつぶす。1スライスに複数のリンパ節がある場合、G1であれば2つ目のリンパ節は $(R,G,B)=(0,200,0)$ 、3つ目は $(R,G,B)=(0,150,0)$ というように色を変えて区別する。

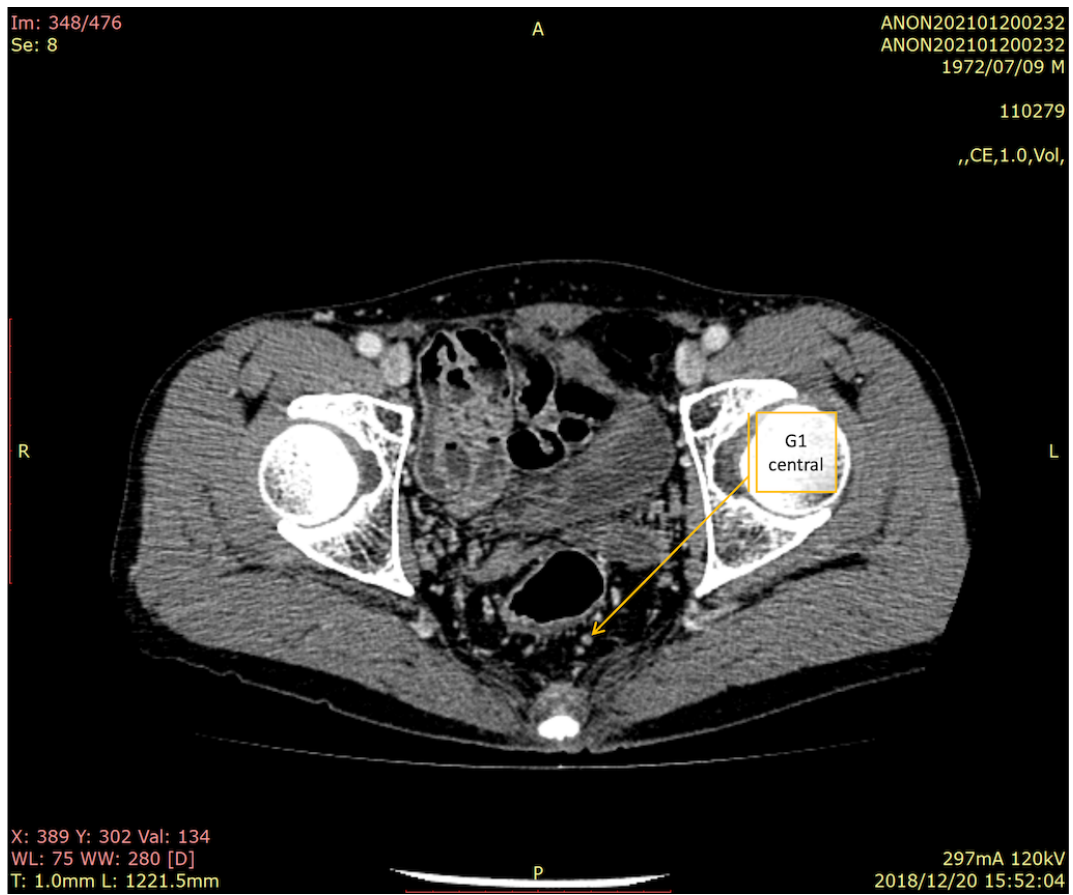


図 3.1: 医師によるアノテーション pptx ファイル

3. それ以外の部分を  $(R,G,B)=(0,0,0)$  で塗りつぶす.

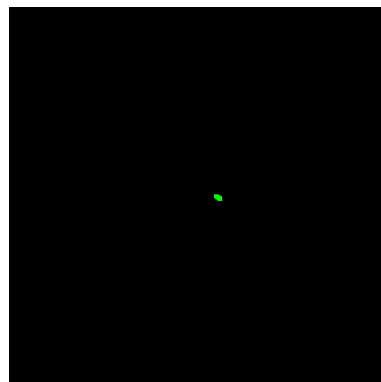
### 3.3 データセットの前処理

データセットの前処理として、実際の画像診断と同様にウィンドウレベル変換を行う。本研究では実際のリンパ節診断に用いられる値を参考にして、ウィンドウレベルを 63, ウィンドウ幅を 256 としてウィンドウレベル変換を行った。すなわち、画素値-65 から 190 を 0 から 255 にマッピングしてグレースケール画像に変換する。

次に、分類用モデルへの入力用にリンパ節領域のクロップを行う。先行研究において、リンパ節の周辺領域もモデルに入力すると分類精度が向上することが示されている [19]。そのため、マスク画像からリンパ節の座標を取得し、リンパ節を中心に  $64 \times 64$  ピクセルにクロップする。本研究では分類用モデルとして 3DCNN を使用する可能性がある。その場合は、前後 16 枚のスライスも同じエリアでクロップして 3 次元的なデータにする。

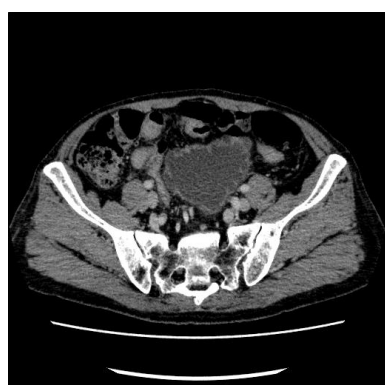


(a) CT 画像

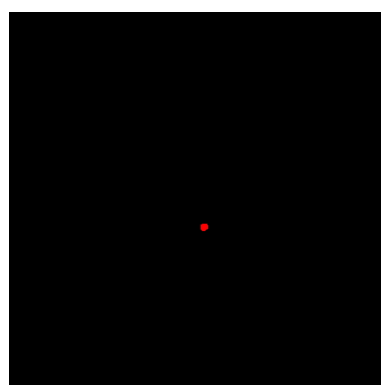


(b) マスク画像

図 3.2: G1 の例



(a) CT 画像



(b) マスク画像

図 3.3: G3 の例

## 第4章 超解像処理

本章では生成モデルを用いた超解像処理によるリンパ節分類精度の向上手法を提案する。初めに、超解像手法がリンパ節分類手法に有効であることを示した予備実験について述べる。次に、超解像モデルの作成手法を説明し、そのCT画像への超解像性能とリンパ節分類への効果について述べる。

### 4.1 予備実験

筆者は少量のデータセットでも分類精度を向上させるためには、データセットに対して超解像や部位のマスキングといった処理を行うことで得られた情報を、CT画像に追加してCNNに入力することが有効であると考えた。本研究ではそのような情報を“追加情報”と定義して、様々な追加情報を使用した場合の分類精度を比較する予備実験を行った。

#### 4.1.1 追加情報

予備実験において検討した追加情報は、超解像CT画像、リンパ節形状ボクセル、楕円体近似ボクセル、楕円体係数の4つである。本項では各追加情報について説明する。

##### 超解像CT画像

超解像CT画像は、リンパ節周辺 $64 \times 64$ ピクセルを $128 \times 128$ ピクセルに超解像した画像である。現在普及しているCTスキャナの分解能は1スライスあたり $512 \times 512$ ピクセルである。解像度の制約により失われた詳細な形状の情報ががん転移分類に必要であると考え、超解像による復元を検討した。超解像手法にはSRCNN [20]を用いる。詳しい学習の手順については4.2節で述べる。

### リンパ節形状ボクセル

リンパ節形状ボクセルは、リンパ節である部分を1、その他の部分を0とした $64 \times 64 \times 16$ のボクセルデータである。リンパ節の3次元形状の取得には画像処理ソフトウェア3D Slicer [21, 22]のGrow from seedセグメンテーション機能を用いる。これはCT画像のマスキングしたい部位の一部に印をつけると、印を元に自動で部位全体を3Dマスキングする機能である。この機能を用いることで、データセットに含まれているリンパ節の位置情報を元に印をつければ、リンパ節の3次元形状データを生成することができる。3D SlicerにはPythonスクリプトを用いて操作するためのAPIが用意されているため、前述の手順を自動化した。3D Slicerを用いてリンパ節につけた印から3次元形状を生成した様子を図4.1に示す。医師がリンパ節へのがん転移の診断を行う場合は3次元的な形状にも注目しているため、CT画像だけではなくリンパ節の3次元形状を表すデータを同時に学習することを検討した。

### 楕円体近似ボクセル

楕円体近似ボクセルは、前述のボクセルデータを最小二乗法によって楕円体近似を行った後、その楕円体を $128 \times 128 \times 16$ でサンプリングし直して構築したボクセルである。G3はG1よりも膨らんだ形状であるという特徴を踏まえ、リンパ節形状ボクセルよりもリンパ節の膨らみに着目したデータとして検討した。

### 楕円体係数

楕円体係数は、楕円体近似で得られた楕円体方程式の係数からなる行列である。楕円体を $Ax^2 + By^2 + Cz^2 + Dxy + Eyz + Fxz + Gx + Hy + Iz - 1 = 0$ と表したときの $A \sim I$ の行列を追加の入力として学習に用いる。楕円体近似ボクセルよりも少ない情報量で、リンパ節の膨らみ方を表す表現形式として検討した。楕円体係数と楕円体近似ボクセルを取得する手順を図4.2に示す。

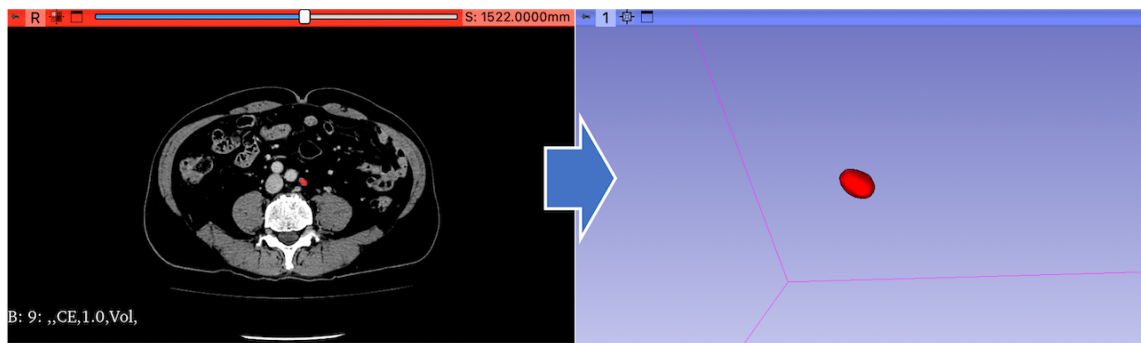


図 4.1: 3D Slicer によるボクセル生成の例

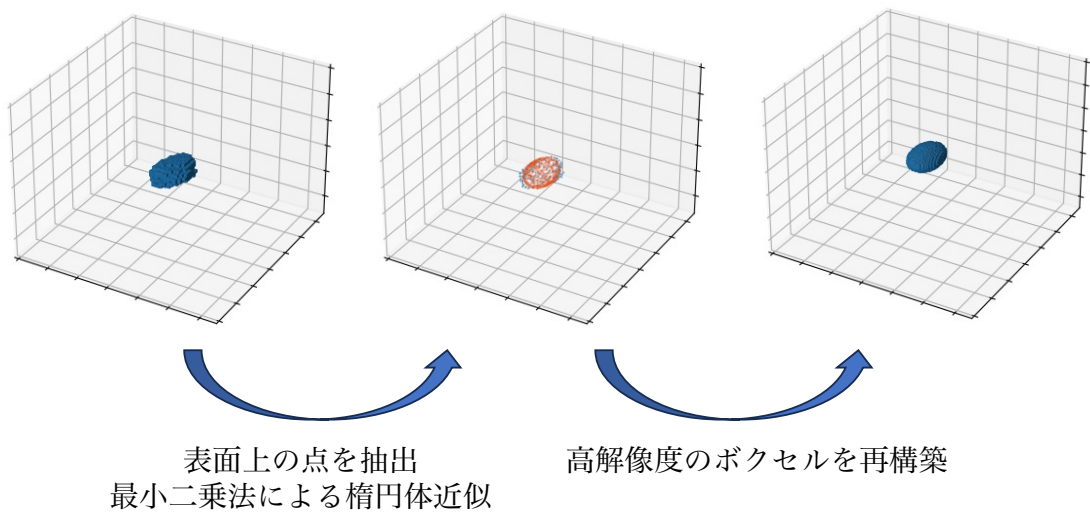


図 4.2: 楕円体係数と楕円体近似ボクセルを取得する手順

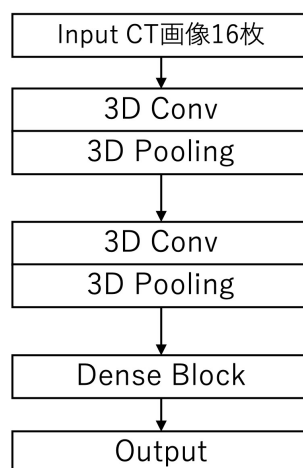


図 4.3: 基準モデル・超解像 CT 用の 3DCNN

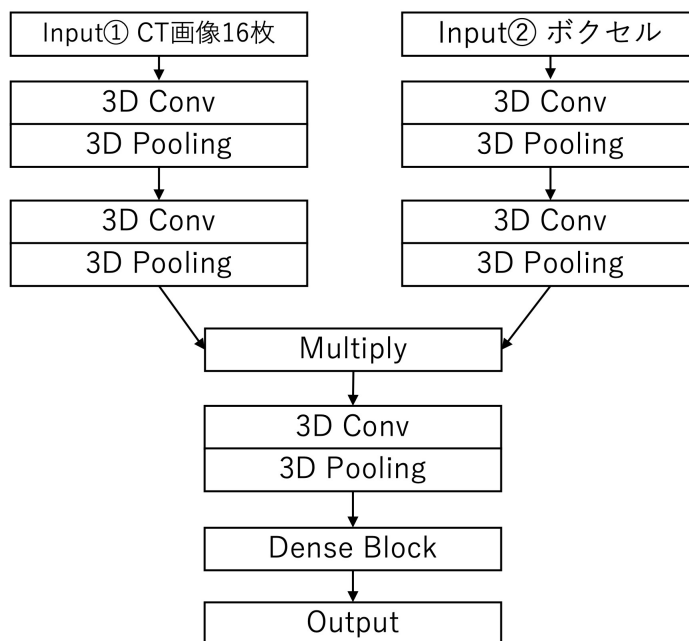


図 4.4: リンパ節形状ボクセル・楕円体近似ボクセル用の 3DCNN

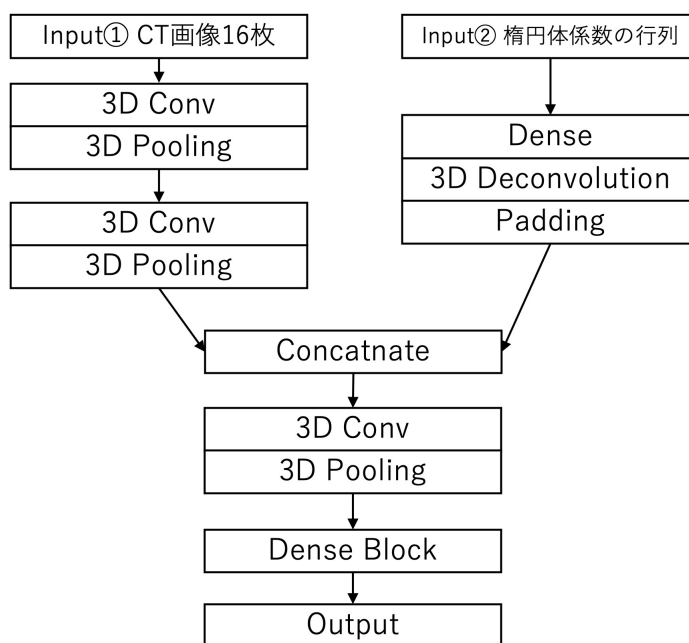


図 4.5: 楕円体係数用の 3DCNN



#### 4.1.2 リンパ節分類実験

各追加情報を用いたときの分類精度を比較する実験を行った。モデルにはCT画像16スライスを入力とする3DCNNと、3DCNNに追加情報のため入力層を追加したネットワークを使用した。ネットワークの概形を図4.3から4.5に示す。なお、G3を正しくG3と分類した場合をTP、G3を誤ってG1と分類した場合をFN、G1を正しくG1と分類した場合をTN、G1を誤ってG3と分類した場合をFPとして評価指標の計算を行った。評価は層化4分割交差検証で行った。データセットを4分割したうちの1つをテストデータとし、残りの3つを訓練データとした。テストデータはクラス不均衡を解消するためG1のアンダーサンプリングとG3のオーバーサンプリングを行い、各クラス44個ずつとした。ただし、テストデータに対してオーバーサンプリングを行うことは一般的には望ましくないため、次節以降の実験では異なる手順で評価することにした。結果の平均値を表4.1に示す。

分類精度のバランスを表すF値に注目すると、いずれの追加情報を使用した場合も向上しており、追加情報を使用することで正しく分類できたリンパ節があることを示している。G3の誤分類の少なさを表すRecallに注目すると、超解像CTを使用した場合に大きく向上している。基準モデルでは誤分類したが、超解像CTでは正しく分類できたデータを確認すると、リンパ節周辺に組織が密集しているようなデータであった。超解像を行い組織同士の配置が鮮明になったことで正しく分類できたと考えられる。また、リンパ節形状ボクセルと楕円体係数でもRecallが向上しており、リンパ節の形状を表す情報がG3の正しい分類に寄与することが確認できる。一方で、楕円体近似ボクセルではRecall以外が低下しており、近似の際に失われる形状情報がG1の正しい分類に必要であったことが示唆された。このように、追加情報の種類によって性能の変化に特徴があることが確認できた。これらの結果から、分類の前処理として超解像を行うことが、リンパ節分類に対して特に効果的であることが示された。そのため、生成モデルを用いたさらに高精度の超解像手法との比較を行うことにした。

表 4.1: 分類結果

追加情報	Accuracy	Precision	Recall	F 値
なし (基準モデル)	0.910	0.941	0.884	0.912
超解像 CT	0.962	0.993	0.931	0.961
リンパ節形状ボクセル	0.921	0.933	0.907	0.919
楕円体係数	0.934	0.975	0.890	0.931
楕円体近似ボクセル	0.886	0.907	0.926	0.916

## 4.2 超解像モデルの作成と評価

本研究では超解像手法として、CNNをベースとした手法であるSRCNNと、GANをベースとした手法であるESRGAN [23]を使用する。SRCNNは超解像分野に機械学習を導入した最初期のモデルであり、パッチ抽出と表現学習を行う畳み込み層、非線形写像を行う活性化関数、および最終的な高解像度の画像を再構築する層から構成されている。SRCNNのアーキテクチャを図4.6に示す。ESRGANはGANを超解像分野に応用した最初期のモデルであるSRGAN [24]を改良したモデルで、視覚的に高品質な超解像が可能である。ESRGANではより現実的な超解像画像を生成するように学習するGeneratorと、本物の画像が偽物の画像よりどれだけ現実的であるかの確率を予測するDiscriminatorによって敵対的学習を行う。ESRGANのGeneratorのアーキテクチャを図4.7に示す。

ここで、超解像モデルの作成手法について説明する。本研究における超解像モデルの目的は、全体が $512 \times 512$ ピクセルのCT画像中の $64 \times 64$ ピクセルの領域を、 $256 \times 256$ ピクセルに超解像することである。そのような学習を行う際は、 $2048 \times 2048$ ピクセルのCT画像中の $256 \times 256$ ピクセルの領域をクロップし、それを $64 \times 64$ ピクセルに低解像度化して、前者を正解画像、後者を入力画像としたデータセットで学習を行うのが望ましい。しかし、 $1024 \times 1024$ ピクセル以上のCT画像はUltra-High Resolution CTと呼ばれ、撮影可能なCTスキャナが普及していないため入手が困難である。そのため $512 \times 512$ ピクセルデータのみを用いて学習を行うために、次の手順で学習を行った。

1.  $512 \times 512$ ピクセルのCT画像から、 $64 \times 64$ ピクセルまたは $128 \times 128$ ピクセルでクロップする。

表 4.2: 超解像モデルの性能評価結果

モデル	PSNR	SSIM
SRCNN	22.50	0.676
ESRGAN	69.04	0.998

2. クロップした画像を縦横それぞれ4分の1に低解像度化する.
3. 2の画像を入力データ, 1の画像を正解データとして学習を行う.

以上の手順を図で示したものが図 4.8 である. 超解像モデルでは入力画像のサイズによらず学習時と同じ倍率での推論を行うため, このような学習が可能である. クロップする解像度は SRCNN では  $128 \times 128$  ピクセル, ESRGAN では  $64 \times 64$  ピクセルとした.

作成したモデルの学習と同じ解像度の組み合わせのテストデータで, PSNR と SSIM [25] を算出した結果を表 4.2 に示す. PSNR は 2 つの画像間の画素ごとの誤差の大きさをデシベルで表す指標であり, 値が大きいほど誤差の画像であることを表す. SSIM は 2 つの画像の明るさ, コントラスト, 構造を比較する指標であり, 1 に近いほど類似した画像であることを表す. SSIM は PSNR よりも人間の視覚特性に近い類似性の評価が可能であるとされている. 実際にリンパ節領域に対して超解像を行った結果を図 4.9 に示す.

結果から, ESRGAN の方が両方の評価指標が高く, CT 画像の超解像に適した手法であることが確認できた. PSNR は 40 以上, SSIM は 0.98 以上の場合には元画像の区別がつかないとされており [26], ESRGAN での超解像は十分に良い精度であるといえる. ただし, これは学習時の解像度の組み合わせでの評価に基づいていることに注意が必要である. 実際に超解像の例を視覚評価すると, 組織の模様や, 背景と組織の境界部分に不自然さが見られる. これは推論時と異なる解像度の組み合わせで学習を行った影響であると考えられる.

### 4.3 リンパ節分類実験

超解像手法の違いによるリンパ節分類への影響を確認するため, 実験を行った. モデルは CT スライス 16 枚を入力とする 3DCNN を用いる. 予備実験よりも超解像の

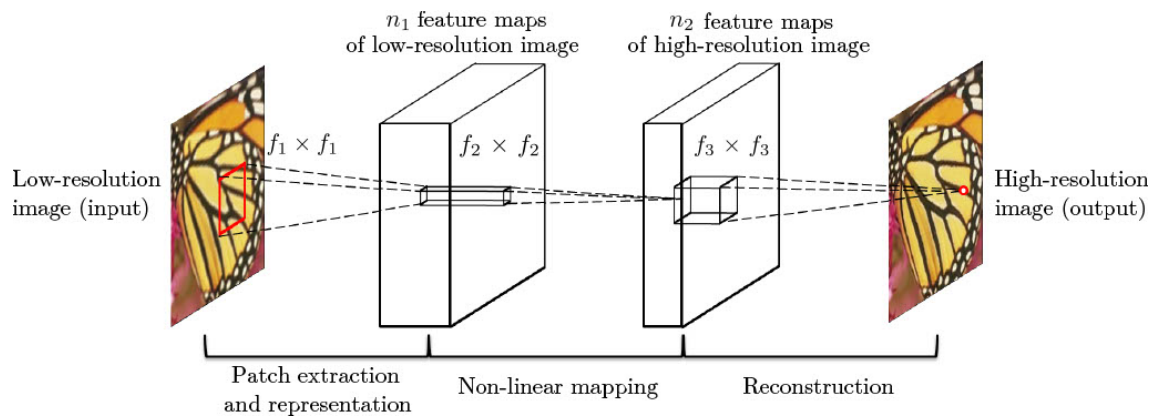


図 4.6: SRCNN のアーキテクチャ (文献 [20] より引用)

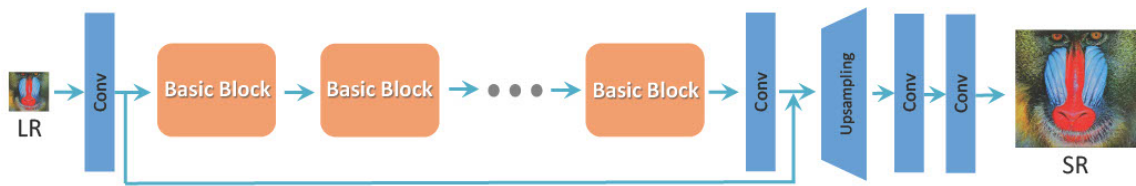


図 4.7: ESRGAN の Generator のアーキテクチャ (文献 [23] より引用)

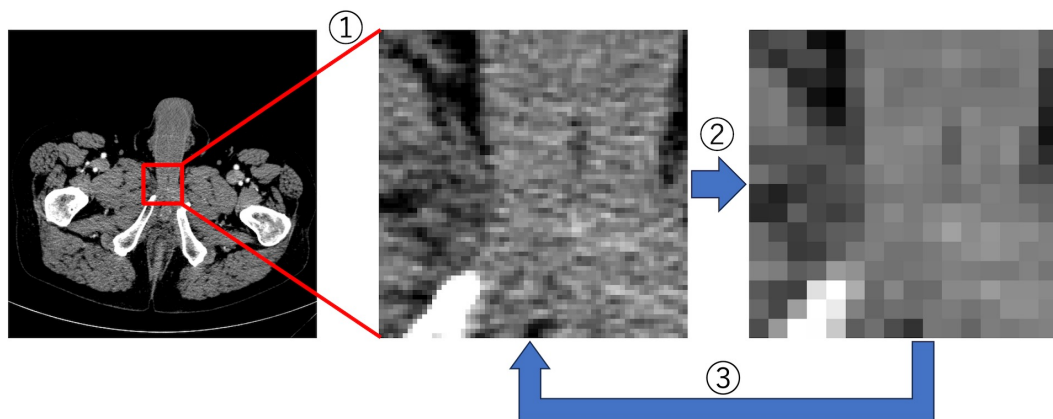
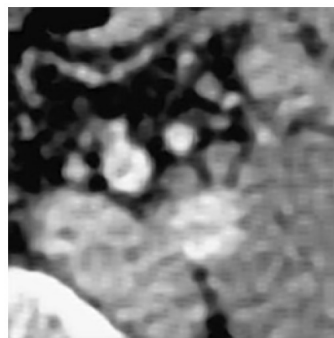


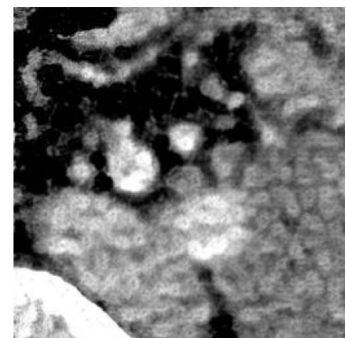
図 4.8: 超解像モデルの学習手順



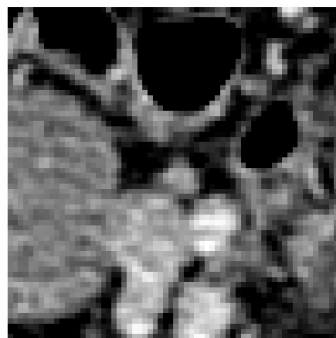
(1a) 入力



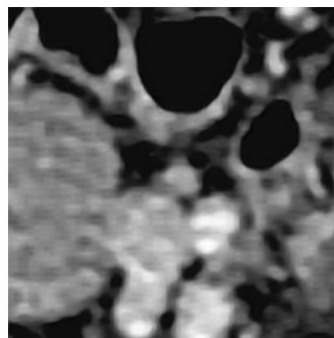
(1b) SRCNN



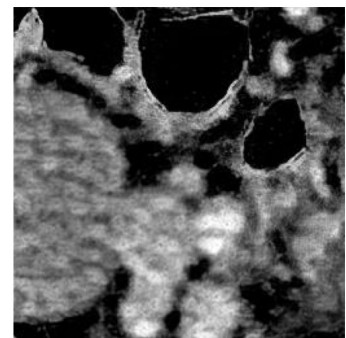
(1c) ESRGAN



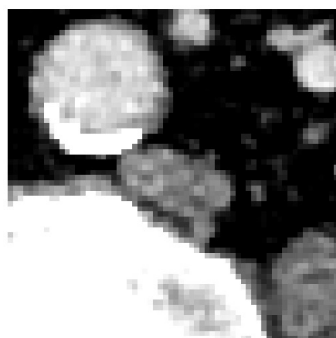
(2a) 入力



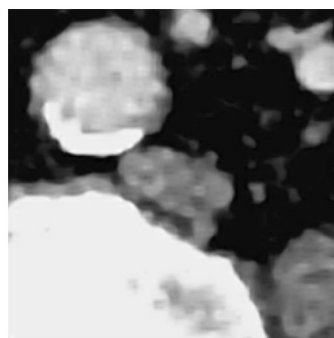
(2b) SRCNN



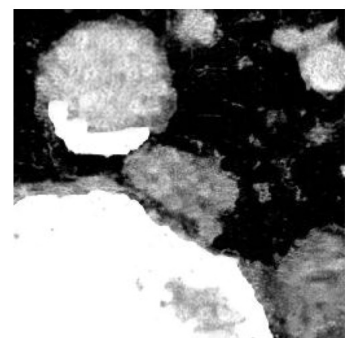
(2c) ESRGAN



(3a) 入力



(3b) SRCNN



(3c) ESRGAN

図 4.9: 超解像の例

倍率を増やしたことで入力する画像のサイズが大きくなったため、予備実験よりも畳み込み層を増やしている。この3DCNNのアーキテクチャを図4.10に示す。超解像処理を行っていない画像、単に拡大をした画像、SRCNNで超解像した画像、ESRGANで超解像した画像を入力データとして学習とテストを行い、分類精度を比較した。評価は層化4分割交差検証で行った。各クラス15個ずつを抽出してテストデータとし、残りのデータセットを4分割したうちの1つを検証データ、残りの3つを訓練データとした。結果の平均値を表4.3に示す。

結果から、より精度の良い超解像手法を用いるほどRecallとF値が向上していることが確認できる。ただし、単に拡大した場合でも向上していることから、入力画像のサイズを大きくすること自体にも効果があることが示唆された。しかし、超解像手法を用いることでさらにRecallとF値が向上しているため、超解像処理はリンパ節分類精度の向上に効果的であることが示された。ESRGANを用いた場合のF値は85.57%であり、文献[11]と比較すると専門医師の85%に匹敵する性能である。一方で、AccuracyとPrecisionは単に拡大した場合が最も良い結果となっている。これらの結果から、超解像によって詳細な部分の情報が追加されることでG3の分類においては良い影響があるが、G1を正しく分類するためにはノイズとなるような情報が追加されていると考えられる。ただし、がん診断においてはG3の見落としを防ぐことが重要となるため、本手法は実際の診断の観点から有効であるといえる。

表 4.3: 分類結果

超解像手法	Accuracy	Precision	Recall	F 値
なし	0.8000	0.9567	0.6333	0.7551
単に拡大	0.8715	1.000	0.7299	0.8391
SRCNN	0.8538	0.9215	0.7824	0.8444
ESRGAN	0.8688	0.9180	0.8018	0.8557



## 第5章 画像生成によるデータ拡張

次に、データセットが少ないという課題の解決方法として、Diffusion モデルで生成した画像をデータ拡張に使用することを考える。これが可能であれば課題解決の糸口となるが、そもそも少量のデータセットでデータ拡張に使用できるクオリティの画像生成モデルを作成できるのかを検証する必要がある。そこで、本章ではリンパ節分類のデータ拡張に使用できる Diffusion モデルの学習手法を提案し、そのリンパ節分類精度への効果を述べる。

### 5.1 画像生成モデルの作成と評価

生成画像をリンパ節分類のデータ拡張を行うためには、任意のがん転移状況のリンパ節周辺画像を生成できる必要がある。まずは、G3 のデータを生成することを目的とする。Diffusion モデルの実装として Denoising Diffusion Implicit Models [27] を使用した。Diffusion モデルはデータに対して段階的にノイズを加える Forward process と、完全なノイズから徐々にノイズを除去することでデータを再構成する Reverse process によって画像を生成するモデルである。この手順を図で表したものが図 5.1 である。G3 はデータセットに含まれる数が少ないため、学習を 2 段階に分ける手法を提案する。手順を以下に示す。

1. G1 と G3 両方のデータを訓練データとして 100 エポック学習する。
2. G3 のデータのみを訓練データとして 50 エポック追加学習する。

なお、手順 2 においては G3 のデータ数が少ないため、リンパ節の中心の前後のスライスも訓練データに追加した。

手順 1 が完了した時点と、手順 2 が完了した時点での生成結果の例を図 5.2, 5.3 に示す。図 5.2 から、G1 と G3 の両方を学習させることで、多様なリンパ節のような



形状が中心に存在する画像が生成できることが確認できる。図 5.3 から、さらに G3 のみで追加学習を行うことで、G3 の特徴を満たすリンパ節のような形状が存在する画像が生成できることが確認できる。これらの結果から、Diffusion モデルによって任意のがん転移状況のリンパ節画像を生成できることが示された。

ここで注意する必要があるのは、少量のデータセットで生成モデルを学習させたため、過学習により訓練データに含まれる画像と酷似した画像が生成される可能性がある点である。生成した画像をデータ拡張として使用することを目的としているため、既存のデータとの類似度の低い画像が生成できていることが重要である。そこで、ランダムに抽出した生成画像と、各訓練データとの PSNR と SSIM を計算して類似度を評価した。生成画像と、それとの PSNR または SSIM が最大であった訓練データ中の画像の組を図 5.4, 5.5 に示す。最大の PSNR は 16.00、最大の SSIM は 0.2257 であった。これらの数値は低く、生成画像が訓練データ中の画像と類似した画像ではないことを示している。また、目視での評価でも 2 つの画像は類似していないことが確認できる。このことから、生成画像をデータ拡張として用いるのは妥当であると考え、リンパ節分類への利用を行った。

## 5.2 リンパ節分類実験

次に、Diffusion モデルで作成した G3 画像が、がん転移分類のためのデータ拡張として有効であるかどうかを確かめる実験を行った。この実験では、単純なデータ拡張のみを行った場合と、それに加えて Diffusion モデルによる生成 G3 画像を使用した場合を比較する。今回は単純なデータ拡張として、-10 から 10 度のランダム回転、水平フリップ、ガウシアンノイズの付与を行った。これらのデータ拡張は、リンパ節の CT 画像として不自然な結果にならないように考慮されている。単純なデータ拡張を行った例を図 5.6 に示す。生成 G3 画像は 150 枚を使用した。このとき、生成 G3 画像は訓練データにのみ使用し、検証データとテストデータには含まれないようにした。ネットワークには大量のデータセットでなくても十分に学習が完了するように、簡単なアーキテクチャの CNN を使用した。そのアーキテクチャを図 5.7 に示す。評価は層化 4 分割交差検証で行った。データの分割方法は 4.2 節と同じである。

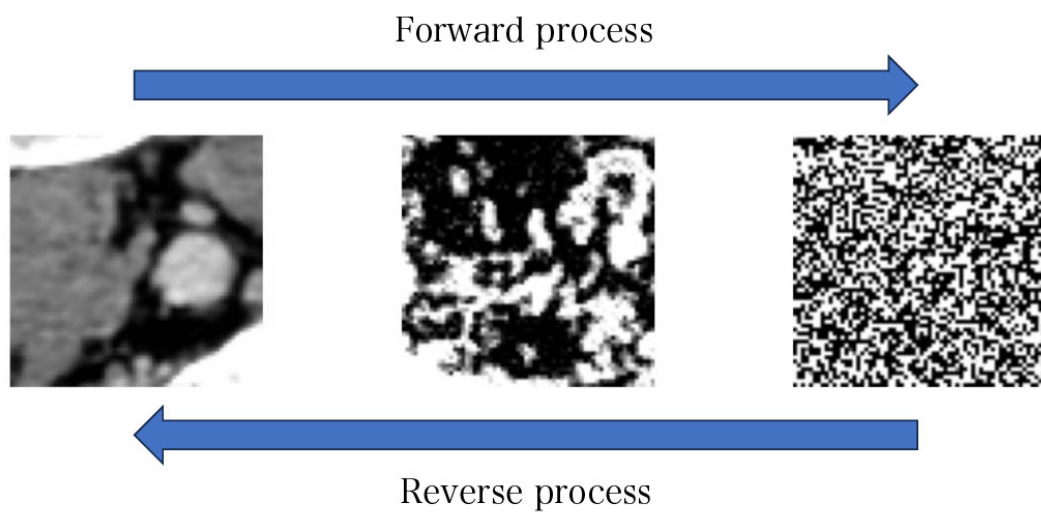


図 5.1: Diffusion モデルの学習プロセス

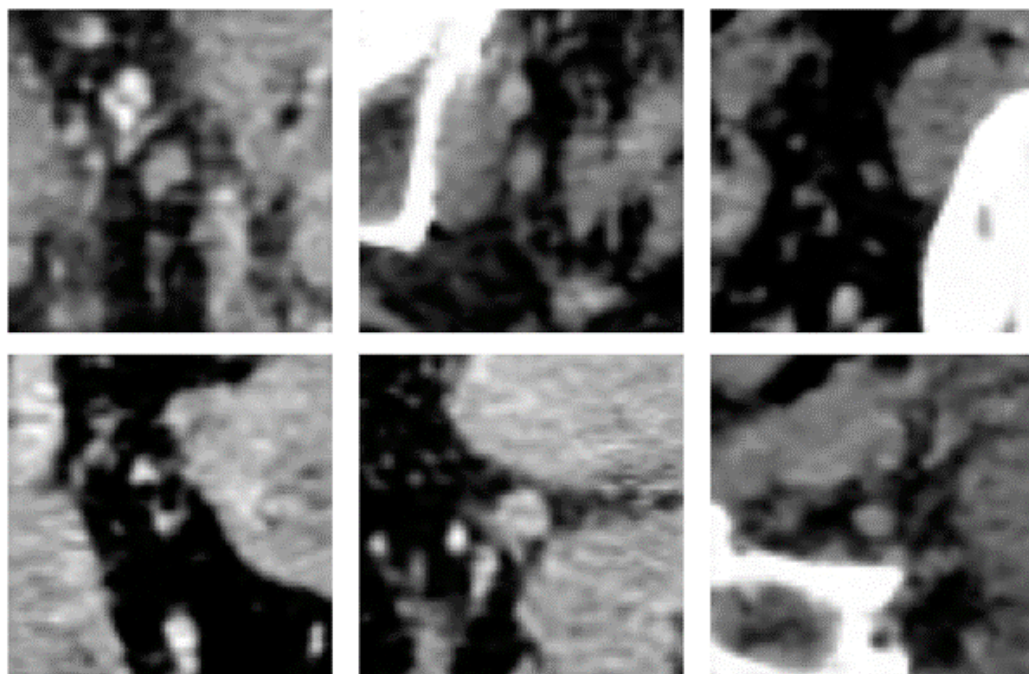


図 5.2: 手順 1 完了時点での生成例

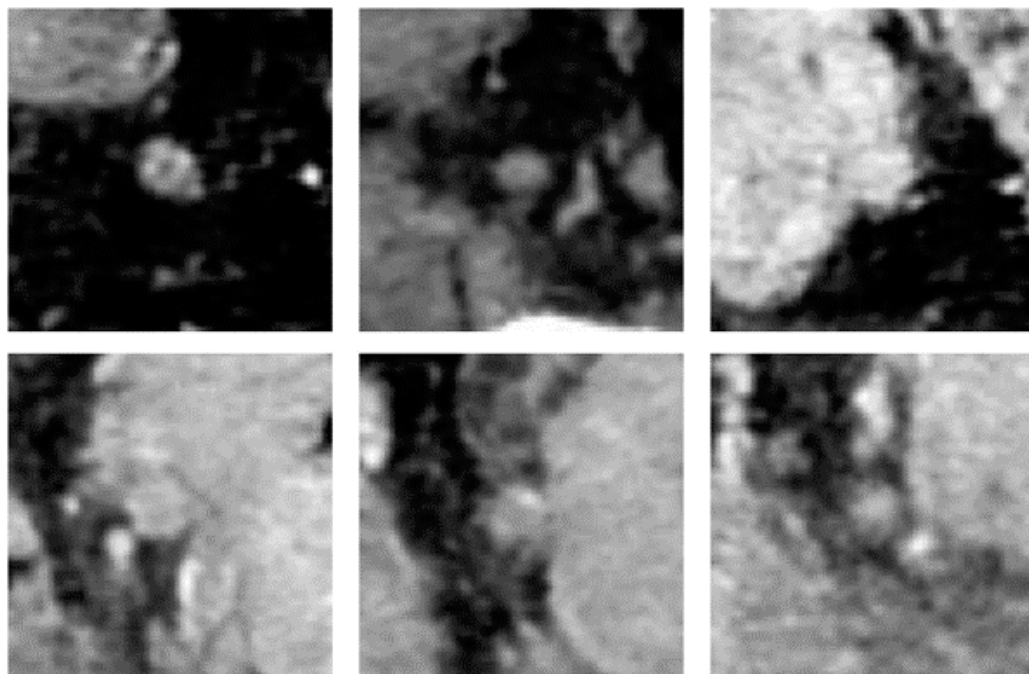
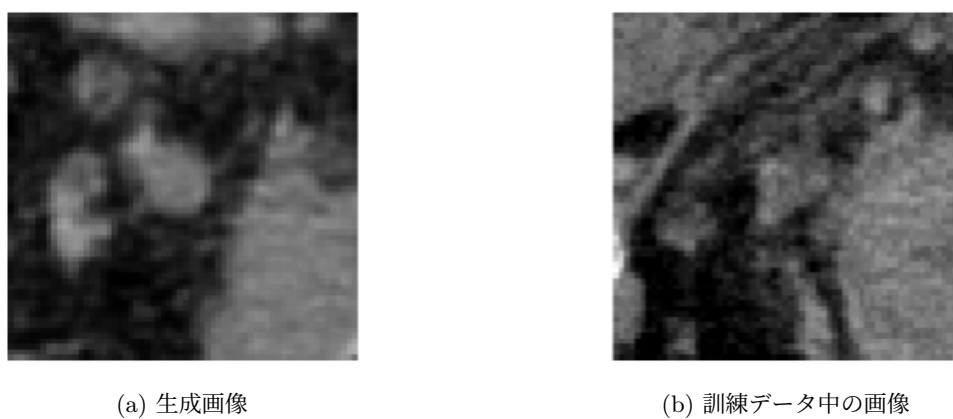


図 5.3: 手順 2 完了時点での生成例



(a) 生成画像

(b) 訓練データ中の画像

図 5.4: PSNR が最大 (16.00) であった組

表 5.1: 分類結果

手法	Accuracy	Precision	Recall	F 値
単純データ拡張	0.8417	0.9372	0.7333	0.8218
生成データ拡張	0.8511	0.9064	0.7865	0.8416

結果の平均値を表 5.1 に示す.

結果から, 生成データ拡張によって Recall が向上しており, G3 の誤分類が減少したことがわかる. Diffusion モデルでの生成によって, 単純なデータ拡張よりも多様性のあるデータの学習が可能となり, より多くの G3 を正しく分類できるようになったことを示している. 一方で, 生成データ拡張では, Precision が下がっており, G1 の誤分類が増えていることがわかる. 現在の手法では, 生成モデルによって図 5.8 のような G1 として分類されるべき特徴を持った画像が生成されたとしても, それも G3 として学習してしまう. このことが G1 の分類に悪影響を及ぼした可能性が考えられる. しかし, そのようなデータを検出するのは現時点では困難である.

### 5.3 Vision Transformer でのリンパ節分類実験

生成データ拡張がより複雑なネットワークの学習にも使えるかを確認するため, Vision Transformer (ViT) [28] の学習を行なった. ViT は自然言語処理分野で高い性能を発揮する Transformer アーキテクチャ[29]を, 画像処理分野に応用したモデルである. ViT では画像を小さなパッチに分割し, それらを単語のように扱い Transformer のエンコーダに入力することで, 画像の局所および全体的な特徴を捉えることができる. ViT は画像分類において高い精度を発揮するが, 学習には大量のデータセットが必要である. 例えば, ViT の事前学習には 3 億枚もの画像が使用されている [28]. そのため, 提供されたデータセットに単純なデータ拡張を実施しただけでは, 十分に学習を進めることができないことが予想される. そこで, 単純なデータ拡張のみを行った場合と, 生成データ拡張を行った場合の分類精度を比較した. 本実験では G3 だけでなく G1 でも生成データ拡張を行った. 5.1 節と同様の手順で作成した G1 の生成例を図 5.9 に示す. 生成データの枚数は各クラス 500 から 4000 枚で変化させて実験した. 評価は 5.2 節と同じテストデータを用いてホールドアウト法で行った.

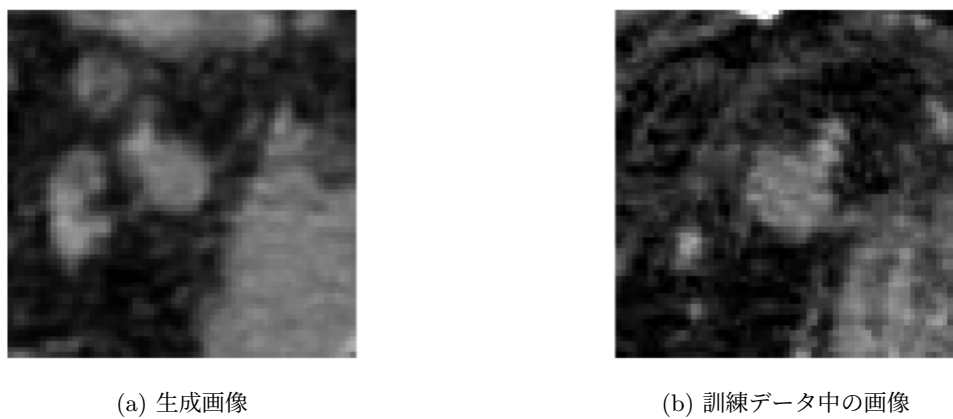


図 5.5: SSIM が最大 (0.2257) であった組

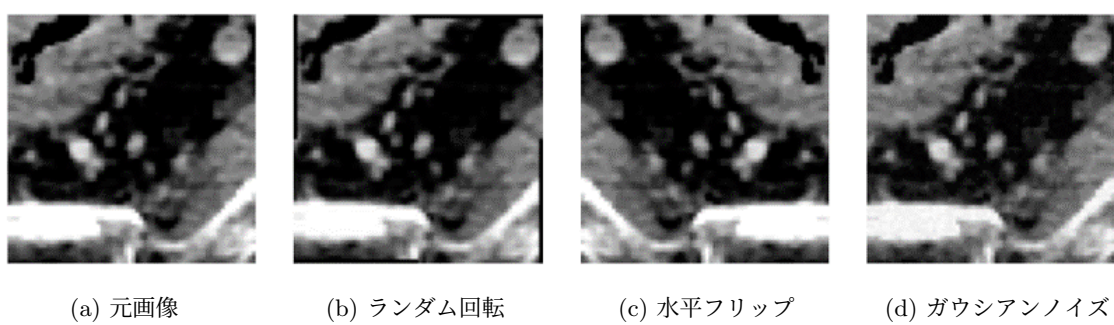


図 5.6: 単純なデータ拡張の例

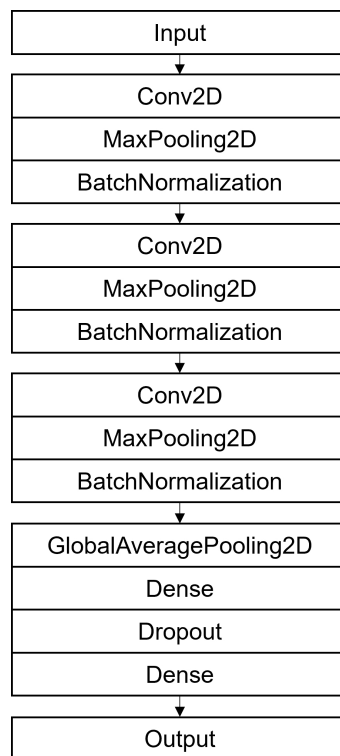


図 5.7: CNN のアーキテクチャ

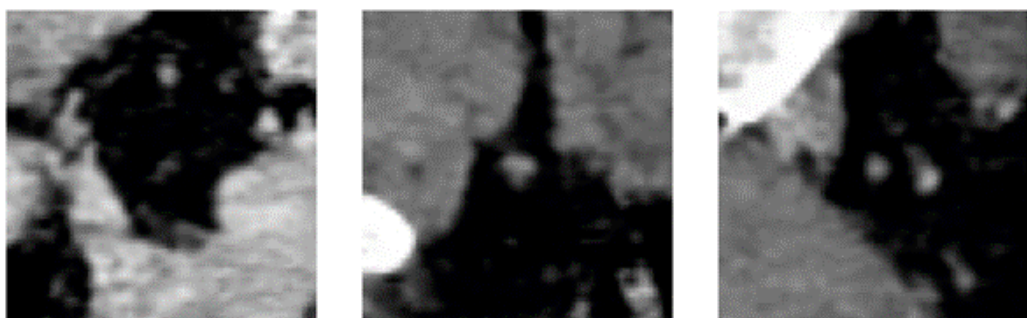


図 5.8: G3 としてふさわしくない生成例

表 5.2: 分類結果

手法	Accuracy	Precision	Recall	F 値
各クラス 1,000 枚, 単純データ拡張	0.7000	1.000	0.4000	0.5714
各クラス 500 枚, 生成データ拡張	0.8333	1.000	0.6667	0.8000
各クラス 1,000 枚, 生成データ拡張	0.8333	1.000	0.6667	0.8000
各クラス 2,000 枚, 生成データ拡張	0.9000	1.000	0.8000	0.8889
各クラス 4,000 枚, 生成データ拡張	0.9000	1.000	0.8000	0.8889

結果を表 5.2 に示す.

結果から, 単純なデータ拡張を行った場合では Recall が極端に低く, 学習が十分に進まずに G1 に偏った出力をするモデルになってしまったことがわかる. 生成データ拡張ではデータ拡張の枚数が単純データ拡張よりも少ない場合でも良い結果が得られた. これは生成データ拡張では, 単純なデータ拡張よりもデータに多様性が生まれるため, 少ない枚数でも効果的なデータ拡張となることを示している. また, 各クラス 2000 枚までは生成データ拡張の枚数を増やすほど性能が向上している. 各クラス 2000 枚の場合での F 値は 88.89% であり, 文献 [11] と比較すると専門医師の 85% を超え, 多量のデータセットで学習した CNN の 88% に匹敵している. 図 5.10 に, 生成データの枚数を増やすことで正しく分類できるようになった例を示す. しかし, これらの間には目視でわかるような特徴の違いは確認できない. 元々のデータセットが少ないため, わずかな違いで誤分類するように学習してしまっていたが, 生成データ拡張によってその違いに対応できるようになったのだと考えられる. 一方で, 生成の枚数を 2000 枚以上に増やしても, それ以上結果が向上していない. これは, 今回作成した生成モデルが出力する画像の多様性の限界を示していると考えられる. これらの結果から, ViT のような複雑なネットワークの学習を行う際に, 用意できるデータセットが少数で単純なデータ拡張では十分に学習ができない場合であっても, 生成データ拡張を用いることによって学習を進められることが示された.

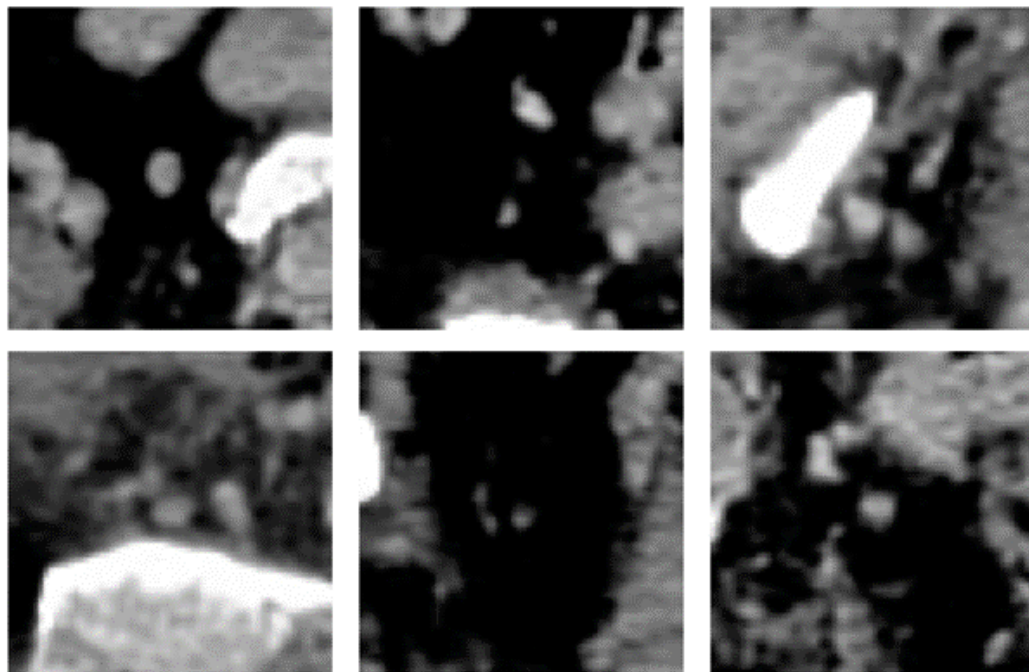


図 5.9: G1 の生成例



(a) 単純データ拡張

(b) 1,000 枚単純データ拡張

(c) 4,000 枚単純データ拡張

図 5.10: 正しく分類できた例 (全て G3)



## 第6章 むすび

本論文では、分類モデルの学習に使用できるデータセットが少ない場合における生成モデルの活用方法を2つ提案した。まず、ESRGANによるCT画像の超解像手法を提案し、分類用モデルに入力する前処理として超解像処理をすることで、G3の分類精度が向上することを示した。F値は85.57%であり、医師の診断と同程度の性能であった。次に、Diffusionモデルによる特定のクラスの画像生成手法を提案し、生成画像をデータ拡張として用いることで分類精度が向上することを示した。さらにViTのように学習に大量のデータが必要とされるネットワークの学習にも応用できることを確認した。ViTで各クラス2000枚の生成データ拡張を行った場合のF値は88.89%であり、先行研究で多量のデータセットでCNNを学習させた場合に匹敵する性能であった。どちらの手法においても、生成モデルを用いることにより、データセットに元々含まれている数が少ないG3についての学習を改善できることがわかった。機械学習による診断支援ではデータの少ない病変を正しく分類できることは重要であるため、本手法は実際の診断支援にも有用であると考えられる。

ここで今後の課題について述べる。超解像手法ではアキシャル断面の超解像しか行っていないため、さらに断面に垂直な方向の超解像も行うことが挙げられる。画像生成手法では複数のスライスを入力する3DCNNのデータ拡張にも対応できるように、連続する複数スライスを生成できるモデルを作成することが挙げられる。今後の展望としては、生成した画像に対して超解像処理を行うことや、超解像モデルの学習に生成した画像を用いることといった、提案手法を組み合わせることによるさらなる分類精度の向上が挙げられる。

## 謝辞

本研究を行うにあたり、終始熱心にご指導を頂いた名古屋工業大学 舟橋健司 准教授、伊藤宏隆 助教に深く感謝いたします。また、日頃よりご指導いただきました中部大学 岩堀祐之 教授に感謝いたします。なお、本研究において、CT 画像の提供をはじめ、ご協力を頂きました愛知県がんセンター 大内晶 先生、清水 泰博 先生に感謝いたします。さらに、本研究を進めるにあたり多くの助言を頂き、御協力いただいた名古屋工業大学 舟橋研究室の皆様にも深く感謝し、厚く御礼申し上げます。

## 参考文献

- [1] Y. Manjunatha, V. Sharma, Y. Iwahori, M. Bhuyan, A. Wang, A. Ouchi, and Y. Shimizu, “Lymph node detection in CT scans using modified U-Net with residual learning and 3D deep network,” *International Journal of Computer Assisted Radiology and Surgery*, vol.18, no.4, pp.723–732, 2023.
- [2] A.-A. Nayan, B. Kijirikul, and Y. Iwahori, “Mediastinal lymph node detection and segmentation using deep learning,” *IEEE Access*, vol.10, pp.89289–89307, 2022.
- [3] T. Ono, Y. Iwahori, H. Usami, B. Kijirikul, M.K. Bhuyan, T. Oshiro, and Y. Shimizu, “Detection of Lymph Nodes using CNN from Contrast-Enhanced CT Images,” *2020 9th International Congress on Advanced Applied Informatics (IIAI-AAI)IEEE*, pp.471–474 2020.
- [4] J.P. Singh, Y. Iwahori, M.K. Bhuyan, H. Usami, T. Oshiro, and Y. Shimizu, “Mediastinal Lymph Node Detection using Deep Learning.,” *ICPRAM*, pp.159–166, 2020.
- [5] S. Kundu, Y. Iwahori, M.K. Bhuyan, M. Bhatt, A. Ouchi, and Y. Shimizu, “Improved Residual Vision Transformer for CT to MRI Translation,” *2023 Fifth International Conference on Transdisciplinary AI (TransAI)IEEE*, pp.58–65 2023.
- [6] “直腸：[国立がん研究センター がん統計],” [https://ganjoho.jp/reg\\_stat/statistics/stat/cancer/7\\_rectal.html](https://ganjoho.jp/reg_stat/statistics/stat/cancer/7_rectal.html). (Accessed on 01/31/2022).
- [7] T. Akiyoshi, T. Watanabe, S. Miyata, K. Kotake, T. Muto, and K. Sugihara, “Results of a Japanese Nationwide Multi-Institutional Study on Lateral Pelvic

- Lymph Node Metastasis in Low Rectal Cancer: Is it regional or distant disease?," *Annals of surgery*, vol.255, no.6, pp.1129–1134, 2012.
- [8] S. Fujita, S. Yamamoto, T. Akasu, and Y. Moriya, "Prognostic Factors of Rectal Cancer Patients with Lateral Pelvic Lymph Node Eetastasis.," *Hepato-gastroenterology*, vol.59, no.120, pp.2494–2497, 2012.
- [9] "JSCCR — 大腸癌研究会," <http://www.jsccr.jp/guideline/2019/cq.html>. (Accessed on 01/12/2022).
- [10] H. Wang, Z. Zhou, Y. Li, Z. Chen, P. Lu, W. Wang, W. Liu, and L. Yu, "Comparison of machine learning methods for classifying mediastinal lymph node metastasis of non-small cell lung cancer from 18 F-FDG PET/CT images," *EJNMMI research*, vol.7, pp.1–11, 2017.
- [11] A. Hartenstein, F. Lübbe, A.D. Baur, M.M. Rudolph, C. Furth, W. Brenner, H. Amthauer, B. Hamm, M. Makowski, and T. Penzkofer, "Prostate cancer nodal staging: using deep learning to predict 68Ga-PSMA-positivity from CT imaging alone," *Scientific reports*, vol.10, no.1, p.3398, 2020.
- [12] K. Suzuki, Y. Iwahori, K. Funahashi, M. Bhuyan, A. Ouchi, and Y. Shimizu, "Lymph-node Detection and Metastasis Classification from CT Images using a Single U-Net Model," *Procedia Computer Science*, vol.207, pp.583–592, 2022.
- [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*, vol.27, 2014.
- [14] C. You, G. Li, Y. Zhang, X. Zhang, H. Shan, M. Li, S. Ju, Z. Zhao, Z. Zhang, W. Cong, M.W. Vannier, P.K. Saha, E.A. Hoffman, and G. Wang, "CT Super-Resolution GAN Constrained by the Identical, Residual, and Cycle Learning Ensemble (GAN-CIRCLE)," *IEEE Transactions on Medical Imaging*, vol.39, no.1, pp.188–203, 2020.

- [15] P. Mann, S. Jain, S. Mittal, and A. Bhat, “Generation of COVID-19 Chest CT Scan Images using Generative Adversarial Networks,” 2021 International Conference on Intelligent Technologies (CONIT), pp.1–5, 2021.
- [16] Y. Liu, L. Meng, and J. Zhong, “MAGAN: mask attention generative adversarial network for liver tumor CT image synthesis,” *Journal of Healthcare Engineering*, vol.2021, pp.1–11, 2021.
- [17] G. Brown, C.J. Richards, M.W. Bourne, R.G. Newcombe, A.G. Radcliffe, N.S. Dallimore, and G.T. Williams, “Morphologic Predictors of Lymph Node Status in Rectal Cancer with Use of High-Spatial-Resolution MR Imaging with Histopathologic Comparison,” *Radiology*, vol.227, no.2, pp.371–377, 2003.
- [18] T. Akasu, G. Iinuma, M. Takawa, S. Yamamoto, Y. Muramatsu, and N. Moriyama, “Accuracy of High-Resolution Magnetic Resonance Imaging in Pre-operative Staging of Rectal Cancer,” *Annals of surgical oncology*, vol.16, no.10, pp.2787–2794, 2009.
- [19] 岡田佳朗, “CNN を用いた造影 CT 画像からのリンパ節転移の有無分類,” 中部大学修士論文, 2021.
- [20] C. Dong, C.C. Loy, K. He, and X. Tang, “Image super-resolution using deep convolutional networks,” *IEEE transactions on pattern analysis and machine intelligence*, vol.38, no.2, pp.295–307, 2015.
- [21] A. Fedorov, R. Beichel, J. Kalpathy-Cramer, J. Finet, J.-C. Fillion-Robin, S. Pujol, C. Bauer, D. Jennings, F. Fennessy, M. Sonka, et al., “3D Slicer as an image computing platform for the Quantitative Imaging Network,” *Magnetic resonance imaging*, vol.30, no.9, pp.1323–1341, 2012.
- [22] “3D Slicer image computing platform,” <https://www.slicer.org/>. (Accessed on 12/31/2023).

- [23] X. Wang, K. Yu, S. Wu, J. Gu, Y. Liu, C. Dong, Y. Qiao, and C. Change Loy, “ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks,” Proceedings of the European Conference on Computer Vision (ECCV) Workshops, Sept. 2018.
- [24] C. Ledig, L. Theis, F. Huszár, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, et al., “Photo-realistic single image super-resolution using a generative adversarial network,” Proceedings of the IEEE conference on computer vision and pattern recognition, pp.4681–4690, 2017.
- [25] Z. Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli, “Image quality assessment: from error visibility to structural similarity,” IEEE transactions on image processing, vol.13, no.4, pp.600–612, 2004.
- [26] 小箱雅彦, “電子化文書の画像圧縮ガイドライン,” 第 50 巻, pp.21–24, 日本文書情報マネジメント協会, 2011.
- [27] J. Ho, A. Jain, and P. Abbeel, “Denoising diffusion probabilistic models,” Advances in neural information processing systems, vol.33, pp.6840–6851, 2020.
- [28] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, et al., “An image is worth 16x16 words: Transformers for image recognition at scale,” arXiv preprint arXiv:2010.11929, 2020.
- [29] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A.N. Gomez, Ł. Kaiser, and I. Polosukhin, “Attention is all you need,” Advances in neural information processing systems, pp.5998–6008, 2017.

## 発表論文リスト

1. 鈴木 康介, 岩堀 祐之, 舟橋 健司, M. K. Bhuyan, 大内 晶, 清水 泰博, “単一 U-Net モデルでの CT 画像からのリンパ節検出と転移有無分類,” 第 25 回画像の認識・理解シンポジウム (MIRU) 予稿集, IS2-73, 2022.
2. K. Suzuki, Y. Iwahori, K. Funahashi, M. K. Bhuyan, A. Ouchi and Y. Shimizu, “Lymph-node Detection and Metastasis Classification from CT Images using a Single U-Net Model,” Proceedings of the International Conference on Knowledge-Based and Intelligent Information & Engineering Systems, pp.583-592, 2022.
3. 鈴木 康介, 岩堀 祐之, 舟橋 健司, 大内 晶, 清水 泰博, “超解像を用いたリンパ節のがん転移分類の精度向上手法,” 2023 年電子情報通信学会総合大会 (情報・システム) 講演論文集, D-16-5, 2023.
4. 鈴木 康介, 岩堀 祐之, 舟橋 健司, 大内 晶, 清水 泰博, “超解像と 3 次元形状を用いたリンパ節の転移有無の分類精度改善,” 第 53 回東海フエジィ研究会予稿集, P1-04, 2023.
5. 鈴木 康介, 岩堀 祐之, 舟橋 健司, 大内 晶, 清水 泰博, “リンパ節のがん転移分類の精度向上のための追加情報の検討,” 令和 5 年度電気・電子・情報関係学会東海支部連合大会講演論文集, F4-3, 2023.
6. 鈴木 康介, 岩堀 祐之, 舟橋 健司, 大内 晶, 清水 泰博, “リンパ節のがん転移分類に有効な超解像手法の検討,” 第 21 回情報学ワークショップ (WiNF) 抄録集, 1A-24, 2023.
7. K. Suzuki, Y. Iwahori, K. Funahashi, M. K. Bhuyan, A. Ouchi and Y. Shimizu,

“Data Augmentation Method for Classification of Lymph Node Metastasis using Diffusion Model,” Proceedings of the International Conference on Image Processing and Robotics, 2024.