

修士論文

金属アーチファクトを含む人工股関節全置換術 術後 CT 画像のセグメンテーション

阪本 充輝

奈良先端科学技術大学院大学

先端科学技術研究科

データサイエンスプログラム

主指導教員: 佐藤 嘉伸 教授

生体医用画像研究室（情報科学領域）

令和 2 年 3 月 13 日提出

本論文は奈良先端科学技術大学院大学先端科学技術研究科に
修士(工学) 授与の要件として提出した修士論文である。

阪本 充輝

審査委員：

佐藤 嘉伸 教授	(主指導教員, 情報科学領域)
向川 康博 教授	(副指導教員, 情報科学領域)
別所 康全 教授	(副指導教員, バイオサイエンス領域)
大竹 義人 准教授	(副指導教員, 情報科学領域)
スーフィー マーゼン 助教	(副指導教員, 情報科学領域)
日朝 祐太 助教	(副指導教員, 情報科学領域)

金属アーチファクトを含む人工股関節全置換術 術後 CT 画像のセグメンテーション*

阪本 充輝

内容梗概

人工股関節全置換術において、患者固有の筋骨格形状を知る事は術前の手術計画立案及び術後の予後評価において非常に重要である。特に術後評価においては筋骨格形状だけでなく、挿入されたインプラントの3次元形状が必要となる。これまでに術前 CT 画像を対象とした筋骨格セグメンテーション手法は提案されているが、術後 CT 画像ではインプラントによって生じる金属アーチファクトが患部周辺の解析精度を低下させる問題がある。そこで本研究では、より定量的な術後評価に向け、術後 CT 画像からの 1) 筋骨格セグメンテーション 2) インプラントセグメンテーションの2つを提案する。

筋骨格セグメンテーションでは、金属アーチファクト低減手法とセグメンテーション手法を組み合わせた手法を提案した。金属アーチファクトシミュレーションを用いたシミュレーション実験と実画像実験を行い、提案法の有効性を示した。

インプラントセグメンテーションでは CAD (Computer-Aided Design) モデルの位置合わせで正解データを作成し、ネットワークの学習を行った。実験では特に、未知のインプラント機種への適用可能性の初期的検討のため、学習データに含まれないインプラント機種を使用して実験を行った。

キーワード

金属アーチファクト, セグメンテーション, 筋骨格, インプラント, 人工股関節全置換術

* 奈良先端科学技術大学院大学 先端科学技術研究科 修士論文, 令和2年3月13日.

Segmentation of postoperative CT images of total hip arthroplasty contaminated with metal artifact^{*}

Mitsuki Sakamoto

Abstract

In total hip arthroplasty, patient-specific musculoskeletal structure is important. Especially, in postoperative assessment, the relative positions of the implant, muscles and bones are required. A method for musculoskeletal segmentation in preoperative CT images has been proposed, but in postoperative CT images, the segmentation accuracy deteriorates due to the metal artifact caused by the implant. Hence, I aim to separately segment 1) musculoskeletal structures and 2) implant shape from postoperative CT images for quantitative postoperative evaluation.

In this thesis, I propose a method that combines a metal artifact reduction method and a segmentation method for musculoskeletal segmentation. Its effectiveness was shown in the simulation and real image experiments. For the implant segmentation, a convolutional neural network was trained with a registered CAD (Computer-Aided Design) model. The experimental results showed the possibility of application to unknown implant models.

Keywords:

Metal Artifact Reduction, Segmentation, Musculoskeletal, Implant, Total Hip Arthroplasty

^{*}Master's Thesis, Graduate School of Science and Technology, Nara Institute of Science and Technology, March 13, 2020.

目次

図目次	v
第 1 章 はじめに	1
第 2 章 術後 CT 画像の金属アーチファクト低減と筋骨格セグメンテーション	3
2.1. 背景	3
2.2. 関連研究	4
2.2.1. 金属アーチファクト低減	4
2.2.2. 筋骨格セグメンテーション	5
2.2.3. 本研究の目的	5
2.3. 手法	6
2.3.1. 金属アーチファクト低減	6
2.3.2. 筋骨格セグメンテーション	8
2.3.3. シミュレーション	8
2.4. 実験	12
2.4.1. 実験概要	12
2.4.2. データセット	12
2.4.3. refinement U-net	12
2.4.4. segmentation U-net	14
2.4.5. シミュレーション実験	15
2.4.6. 実画像実験	16
2.4.7. 評価指標	16
2.5. 実験結果	17
2.5.1. シミュレーション実験	17
2.5.2. 実画像実験	19
2.6. 考察	26
2.6.1. セグメンテーション精度	26
2.6.2. セグメンテーションの不確実性	26
2.6.3. 今後の課題及び展望	27

2.7.	小括	28
第 3 章	術後 CT 画像からのインプラントセグメンテーション	29
3.1.	はじめに	29
3.2.	手法	30
	3.2.1. CAD モデルの位置合わせによる正解データ作成	30
3.3.	実験	31
	3.3.1. データセット	31
	3.3.2. U-net	34
	3.3.3. 実験概要	34
	3.3.4. 評価指標	35
3.4.	実験結果	36
	3.4.1. 実験 1 (学習データ : Centpillar, テストデータ : Centpillar)	36
	3.4.2. 実験 2 (学習データ : Centpillar, テストデータ : Accolade)	38
3.5.	考察	41
	3.5.1. セグメンテーション精度	41
	3.5.2. 今後の課題及び展望	41
3.6.	小括	42
第 4 章	総括	44
謝辞		45
参考文献		46

図目次

2.1	金属アーチファクトを含む術後 CT 画像の例 ([WW, WL]=[500, 100]HU)	4
2.2	提案法の概要図. (a) 術後 CT 画像を用いた推論 (b) 金属アーチファクトシミュレーションを用いた refinement U-net の学習 (c) 術前 CT 画像を用いた segmentation U-net の学習	6
2.3	大腿骨頭骨切のシミュレーション	9
2.4	骨盤臼蓋掘削のシミュレーション	10
2.5	金属アーチファクトシミュレーションの流れ	10
2.6	筋肉ラベルと対応する色 [5] © 2019 IEEE.	13
2.7	3D Slicer [§] を用いた術前 CT 画像への CAD モデルの挿入	15
2.8	シミュレーション実験での各筋肉・骨の Dice 係数	18
2.9	シミュレーション実験での各筋肉・骨の ASD	18
2.10	骨盤・大腿骨の Dice 係数と PSV の関係	19
2.11	臀部の筋肉 (図 2.6) の Dice 係数と PSV の関係. 各プロットの左の図は対象の筋肉 (赤) とインプラント (緑) の位置関係を示す.	20
2.12	大腿部の筋肉 (図 2.6) の Dice 係数と PSV の関係. 各プロットの左の図は対象の筋肉 (赤) とインプラント (緑) の位置関係を示す.	21
2.13	シミュレーション実験でのセグメンテーション結果及び不確実性の可視化	22
2.14	シミュレーション実験での不確実性のプロット (20 症例)	22
2.15	実画像における大臀筋の精度評価	23
2.16	実画像における中臀筋の精度評価	23
2.17	実画像実験でのセグメンテーション結果及び不確実性の可視化	24
2.18	実画像での不確実性のプロット (132 症例)	24

2.19	実画像への提案法の適用結果（左から，術後 CT 画像のボリュームレンダリング，術後 CT 画像，NMAR 適用後の画像，さらに refinement U-net を適用した画像でのセグメンテーション結果及び不確実性）	25
3.1	カップの位置合わせの失敗例（黒のワイヤーフレーム: $p_{\text{threshold}}$, 緑: CAD モデル p_{CAD})	31
3.2	インプラントの CAD モデル（赤: カップ, 緑: ヘッド, 青: ステム）の構造と座標系（いずれも右手系）	33
3.3	インプラントの CAD モデルのバリエーション	33
3.4	ステムに Centpillar を使用した症例のセグメンテーション精度 . . .	36
3.5	ステム（Centpillar）の Dice 係数が最高値であった症例（カップ: 0.896, ヘッド: 0.962, ステム: 0.965). [WW,WL]=[1750, 4500]HU	37
3.6	ステム（Centpillar）の Dice 係数が中央値であった症例（カップ: 0.863, ヘッド: 0.945, ステム: 0.942) [WW,WL]=[1750, 4500]HU	37
3.7	ステム（Centpillar）の Dice 係数が最低値であった症例（カップ: 0.859, ヘッド: 0.898, ステム: 0.831) [WW,WL]=[1750, 4500]HU	37
3.8	位置合わせされたステムの CAD モデルからの表面間距離誤差. 左は Centpillar, 右は Accolade の症例での結果を示す. 青: 閾値 (2000 HU) セグメンテーション結果との位置合わせ誤差 橙: 提案法によるセグメンテーション結果との位置合わせ誤差	38
3.9	セグメンテーションされたインプラントと CAD モデルとの位置合わせ結果の比較. 左: 閾値セグメンテーションとの位置合わせ結果 右: 提案法でのセグメンテーション結果との位置合わせ結果. .	39
3.10	ステム機種が Accolade である症例のセグメンテーション精度. 左: Dice 係数, 右: ASD.	39
3.11	ステム（Accolade）の Dice 係数が最高値であった症例（カップ: 0.915, ヘッド: 0.965, ステム: 0.907) [WW,WL]=[1750, 4500]HU	40
3.12	ステム（Accolade）の Dice 係数が中央値であった症例（カップ: 0.850, ヘッド: 0.960, ステム: 0.848) [WW,WL]=[1750, 4500]HU	40

3.13 ステム（Accolade）の Dice 係数が最低値であった症例（カップ:
0.901, ヘッド: 0.949, ステム: 0.750）[WW,WL]=[1750, 4500]HU 40

第 1 章 はじめに

整形外科領域では，術前の診断・手術計画立案や術後評価のため，CT 画像，MR 画像，X 線画像などの医用画像が用いられる．中でも，CT 画像から得られる 3 次元像は骨形態，骨病変の立体的把握に有用である [1] という特徴がある．また近年では，このような医師の診断を計算機によって支援する技術の研究が進められている．

本研究では，整形外科領域で広く行われている人工股関節全置換術を対象とする．人工股関節全置換術は痛みを伴う股関節をカップ・ステム・ヘッドからなる人工股関節（以下インプラント）に置き換える手術である．この手術では，より優れた術後の股関節機能を得るため，術前にインプラントの形状・サイズ・設置位置・設置角度などを計画することが重要である [2]．医師は術前に撮像した CT 画像を用いた 3 次元テンプレティングによって，股関節の可動域や脚長差などの指標を考慮して手術計画を立案している．またコンピュータ支援の観点では，Otomaru ら [3] はこの術前計画立案を自動で行う手法を提案しており，術前 CT 画像からの筋骨格領域の自動セグメンテーション手法もこれまでに報告されている [4, 5, 6, 7, 8]．

一方，術後 CT 画像では，置換されたインプラントによって金属アーチファクトと呼ばれる偽像が生じ，特に患部周辺の解析が困難になる．筋肉・骨とインプラントの位置関係は脱臼などの合併症に繋がる要因であるため，インプラント設置位置・角度の評価 [9] やインピンジメント（衝突）についての研究 [10] が行われているが，手動もしくは半手動の解析であり，解析可能な症例数には限りがある．また，インプラントの緩みなどが原因で患者が再度手術を受ける際，計画立案に使用する CT 画像には金属アーチファクトが発生し，医師の大きな負担となる．

そこで本研究では，自動での人工股関節全置換術術後解析に向け，金属アーチファクトを含む人工股関節全置換術の術後 CT 画像からの筋骨格及びインプラントのセグメンテーションを目的とする．

本論文の構成は以下の通りである．第 2 章では術後 CT 画像の金属アーチファクト低減と筋骨格セグメンテーションについて，シミュレーション画像と実画像を用いた評価を述べる．続いて第 3 章では術後 CT 画像からのインプラントセグメンテーションについて述べ，特に，未知のインプラント機種への汎用性についての初

期的検討を行う．最後に，第 4 章では総括を行う．

第 2 章 術後 CT 画像の金属アーチファクト低減と筋骨格セグメンテーション

2.1. 背景

人工股関節全置換術において、患者固有の筋骨格の解剖及びそれを用いた診断や術前計画、術後評価は非常に重要である。Uemura ら [11] は筋萎縮が生じた後の筋肉量の回復について、CT 画像の横断面上での筋肉の面積を用いて評価しており、松下ら [47] は術後 CT 画像上で最大筋幅を計測する事によって術後に内閉鎖筋が温存されているかを検証している。このように、筋肉については従来では CT 画像や MR 画像の断面像による 2 次元的な評価が行われてきた。Ogawa ら [12] は筋骨格の手動セグメンテーションを用いて筋萎縮の評価を行っており、3 次元での計測結果と 2 次元での計測結果では相関があるが、正確な評価には 3 次元での計測が必要であると報告している。しかしながら患者固有の筋骨格形状の手動セグメンテーションには膨大な時間を要するため、3 次元での評価が可能な症例数には限りがあり、また臨床現場での患者固有の筋骨格モデルの応用は困難であった。

この問題に対し、CT 画像や MR 画像から筋骨格形状を自動認識する手法がこれまでに提案されてきた [4, 5, 6, 7, 8]。特に深層学習の一種である Convolutional Neural Networks (CNN) を用いた手法 [5] は精度の向上と計算時間の大幅な短縮を実現し、臨床現場での応用可能性を大きく高めた。

ただし、これらの手法は術前に撮影された CT 画像を対象としており、術後 CT 画像を用いた場合の評価はなされていない。人工股関節全置換術の術後 CT 画像では置換されたインプラントによって金属アーチファクトが発生し（図 2.1）、特に観察したいインプラント周辺のセグメンテーション精度が低下する問題がある。術後 CT 画像からの筋骨格セグメンテーションが実現できれば、患者個別に、より定量的な術後評価を行うことが可能となる。また、再置換術のように術前 CT 画像で金属アーチファクトが発生してしまう場合にも、より正確な手術計画を立案できるという利点がある。

本章は著者らがこれまでに報告した内容 [13] から、実験条件の変更・データセットの拡充を行ったものである。

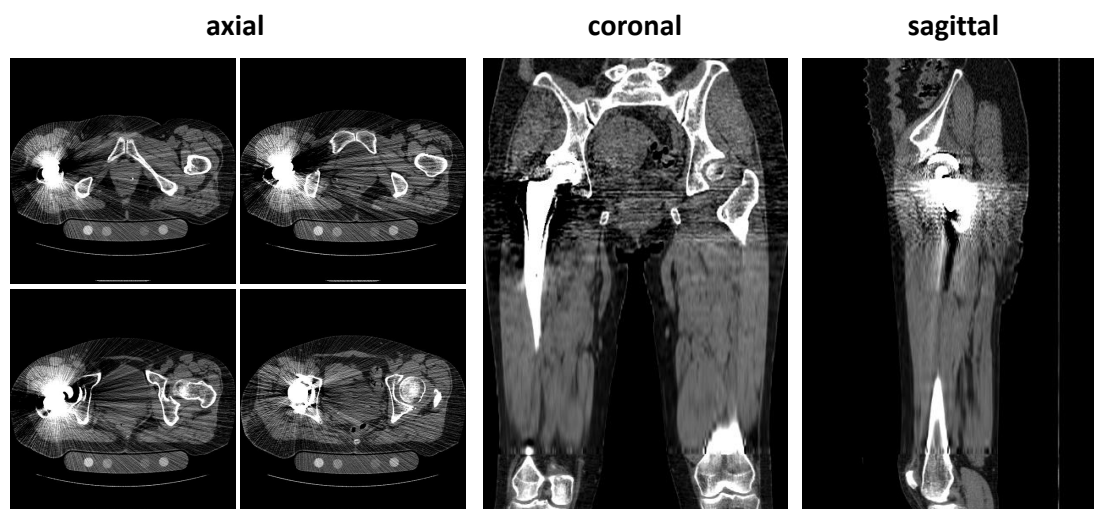


図 2.1 金属アーチファクトを含む術後 CT 画像の例 ([WW, WL]=[500, 100]HU)

2.2. 関連研究

本研究の関連研究を 1) 金属アーチファクト低減手法, 2) 筋骨格セグメンテーション手法の 2 つに分けて述べる.

2.2.1. 金属アーチファクト低減

金属アーチファクト低減手法は大きく以下の 6 種類に大別される [14].

1. インプラントの材質の最適化
2. 撮影法の改善
3. 物理現象に基づく補正
4. サイノグラムの補正
5. 逐次的再構成
6. 再構成された画像の後処理

中でも, 4 に分類される Normalized Metal Artifact Reduction (NMAR)[15] は高精度に金属アーチファクトを低減できる解析的手法として知られている. NMAR では, 骨・水・空気の減弱係数で表された画像による正規化を行う事で従来法より

も正確にサイノグラムを補間し、金属アーチファクトの低減精度を向上させている。

また、4 と 6 を組み合わせた手法として、Gjesteby ら [16] は NMAR によって低減しきれなかった金属アーチファクトの低減を目的とし、CNN を用いて NMAR の出力画像を補正する手法を提案している。

2.2.2. 筋骨格セグメンテーション

術前 CT 画像を対象とした筋骨格セグメンテーション手法として、Yokota ら [4] は統計的アトラスを用いた手法を提案している。また近年では CNN を用いて筋肉・脂肪組織をセグメンテーションする手法が提案されている [5, 6, 7, 8]。その中でも Hiasa ら [5] は、CT 画像の横断面画像から下肢筋骨格形状を高精度かつ短時間で自動認識する Bayesian U-net を提案した。さらに、Bayesian U-net で得られる不確実性とセグメンテーション精度に相関があり、不確実性によって正解データなしに精度評価できる可能性を示した。

一方、術後 CT 画像からの筋骨格セグメンテーションの先行研究としては、Ranzini ら [17] による手法が挙げられる。[17] は統計的アトラスを用いた手法に分類され、同一患者の金属アーチファクト低減後の術後 CT 画像と術後 MR 画像を相補的に使用する事でセグメンテーションを行う手法である。ただし CT 画像と MR 画像の両方を必要とするため、撮像コストが大きいという問題がある。また、CT 画像と MR 画像のどちらかが不足している過去症例の解析には適用できず、CT 画像のみを用いたセグメンテーション精度の向上を目的とする本研究とは異なる。

2.2.3. 本研究の目的

本研究では、金属アーチファクト低減手法と筋骨格セグメンテーション手法を組み合わせ、金属アーチファクトを含む人工股関節全置換術の術後 CT 画像のみから筋骨格セグメンテーションを行うことを目的とする。また筋骨格セグメンテーションの不確実性を推定し、金属アーチファクトを含む術後 CT 画像におけるセグメンテーション精度と不確実性の関係について調査を行う。

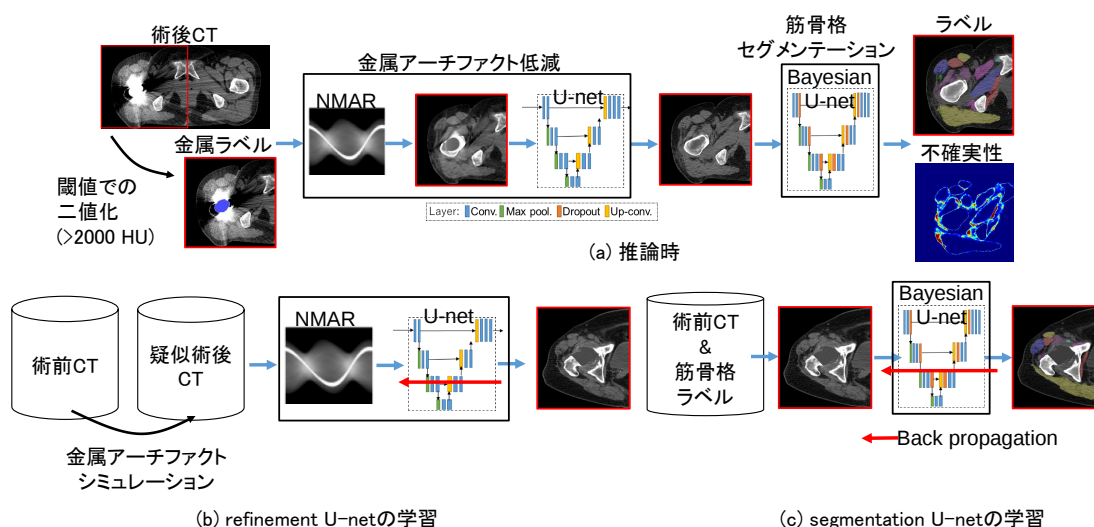


図 2.2 提案法の概要図． (a) 術後 CT 画像を用いた推論 (b) 金属アーチファクトシミュレーションを用いた refinement U-net の学習 (c) 術前 CT 画像を用いた segmentation U-net の学習

2.3. 手法

提案法の全体像を図 2.2 に示す．本研究では，術後 CT 画像に金属アーチファクト低減を適用したのち，術前 CT 画像で学習したネットワークによって筋骨格セグメンテーションを行う．金属アーチファクト低減および筋骨格セグメンテーションについては以降で詳述する．

2.3.1. 金属アーチファクト低減

本研究では金属アーチファクト低減として，以下の 2 つの手法を用いる．

1. NMAR のみ
2. NMAR の出力を U-net[18] により補正

ここで NMAR[15] について述べる．NMAR を含む，サイノグラムの補間によって金属アーチファクトを低減する手法（2.2.1 節）では，まず再構成後の画像から

金属領域のセグメンテーションを行う．この金属領域のセグメンテーションは，計算効率を考慮し，式 (2.3.1) で示される閾値 T を用いた手法で行われる事が多い [15, 19]．本研究では [19] と同様に，閾値 T を 2000 HU とした．

$$f^{metal} = \begin{cases} 0, & f^{image} < T \\ 1, & f^{image} \geq T \end{cases} \quad (2.3.1)$$

ここで f^{image} は金属アーチファクト低減前の画像であり， f^{metal} はセグメンテーションされた金属ラベルである．そしてセグメンテーションされた金属ラベルを CT 画像の撮影時と同様の条件で再投影し，再構成前のサイノグラム上での金属領域を決定する．NMAR の特徴はサイノグラム上の金属領域の補間 [20] の前に，prior image と呼ばれる画像を用いてサイノグラムの除算正規化を行う点である．prior image は水・空気・骨の減弱係数で表される画像であり，NMAR では prior image から得られるサイノグラム p^{prior} で元画像のサイノグラム p を除算正規化し，補間精度を向上させている．補正後のサイノグラム p^{corr} は以下の式で表される ([15] を一部変更)．

$$p^{corr} = R(f^{prior} M(\frac{p}{R(f^{prior})}, R(f^{metal}))) \quad (2.3.2)$$

ここで， f^{prior} は prior image， $R(\cdot)$ はラドン変換などの順投影， $M(\cdot)$ はサイノグラムの補間を意味する．ただし，本研究では再構成前のサイノグラム p が得られなかったため，計算機上での 2 次元再投影によって得られるサイノグラム $R(f^{image})$ を使用した．

NMAR は高精度に金属アーチファクトを低減可能であるが，サイノグラムの補間を行うため，再構成結果の特にインプラント周辺においてボケが生じることがある [21]．また，補間に用いる金属のセグメンテーションの精度によっては，副次的な金属アーチファクトが発生する．

そこで本研究では CNN を用いて，NMAR の出力を補正する生成器 G^{Refine} を求め，筋骨格セグメンテーション精度の向上を試みる (式 (2.3.3))．

$$\hat{I} \cong G^{Refine}(NMAR(I^{post})) \quad (2.3.3)$$

ここで， $\hat{I}$ は所望の金属アーチファクト低減後画像， I^{post} は術後 CT 画像である．生成器 G^{Refine} には，U-net[18] を採用する．U-net は学習データの高周波成分を

捉えるという特徴から、画像変換で用いられることが多い [22]. Gjestebj ら [16] は同様に NMAR の出力を補正する CNN を提案しているが、元画像から切り出したパッチ画像（画像サイズ: 64×64 ）を学習データとしており、解剖構造を考慮した学習にはなっていない。本研究の最終目的は筋骨格形状のセグメンテーションであるため、ネットワークへの入力には CT 画像の横断面全体（画像サイズ: 512×512 ）とした。

以降、この NMAR の出力を補正する U-net を refinement U-net と呼ぶ。

2.3.2. 筋骨格セグメンテーション

本研究では、筋骨格セグメンテーションには Hiasa ら [5] の Bayesian U-net を用いる。Bayesian U-net は Gal ら [23] が提案した MC Dropout を取り入れ、セグメンテーション結果に加えて不確実性を出力する。MC Dropout によって得られる不確実性は epistemic uncertainty と呼ばれ、学習データのバリエーションの不十分さが要因となってモデルに生じる不確実性を指す [24]。

本研究では、金属アーチファクトを含まない術前 CT 画像のみを用いて Bayesian U-net を学習する事で、金属アーチファクトによる不確実性を捉える事を試みる。

以降、筋骨格セグメンテーションに用いるネットワークを segmentation U-net と呼ぶ。

2.3.3. シミュレーション

本研究では、学習及び評価用データセット作成のため、金属アーチファクトのシミュレーションを行った。実際の人工股関節全置換術では、大腿骨頭の骨切や骨盤臼蓋の掘削によって一部の骨を除去する。本研究では術中の骨除去を想定し、以下で求められる骨領域を軟組織の CT 値によって置換する事で、擬似的に大腿骨頭・骨盤臼蓋の一部を除去した画像を作成した。各ボクセルにおいて $[-10, 10]$ HU の範囲から一様分布に従って値をサンプルし、それを軟組織の CT 値として置換した。

- 1) ステム設置位置から決定される骨切面、2) 転子窩と膝中心を通り、CT 座標系の x-z 平面に垂直な平面 の 2 平面から決定される領域と大腿骨ラベル

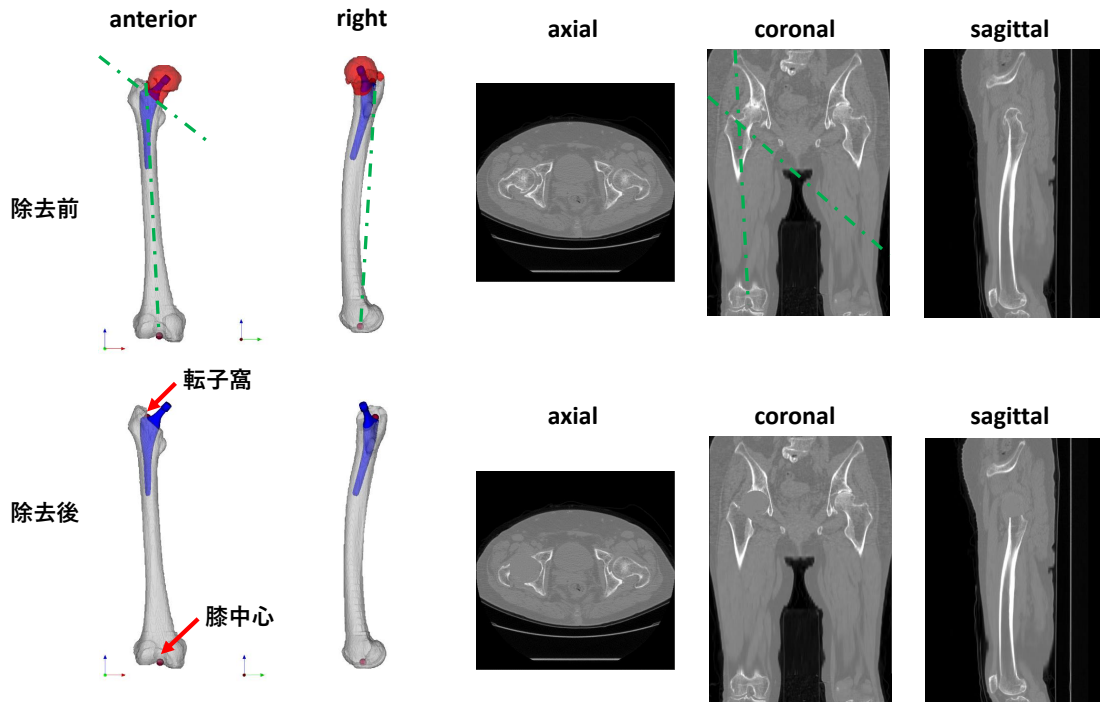


図 2.3 大腿骨頭骨切のシミュレーション

の重畳部分 (図 2.3)

2. 計画されたカップ設置位置と骨盤ラベルの重畳部分 (図 2.4)

学習用及び評価用データの作成に使用した金属アーチファクトのシミュレーションについて述べる．主に Zhang ら [25] による手法を使用しているが，再構成結果を本研究で使用する実際の CT 画像と近づけるため，[26] で提案されたビームハードニング補正の手法を導入している．1) 金属アーチファクトを含まない術前 CT 画像 2) 金属領域を示すラベル画像 の 2 つが入力となり，出力として金属アーチファクトを含むシミュレーション CT 画像が得られる．以下にシミュレーションの流れを述べる．

使用したシミュレーションの流れを図 2.5 に示す．まず，ある X 線のエネルギー E_0 を仮定し，術前 CT 画像の CT 値を線減弱係数に変換する．この得られた画像を $\mathbf{x}$ とする．一方，エネルギー E_0 での線減弱係数値で金属領域を埋めた画像を $\mathbf{x}^m$ とする．ここで本研究では E_0 は 40 keV とした．CT 値 HU から線減弱係数

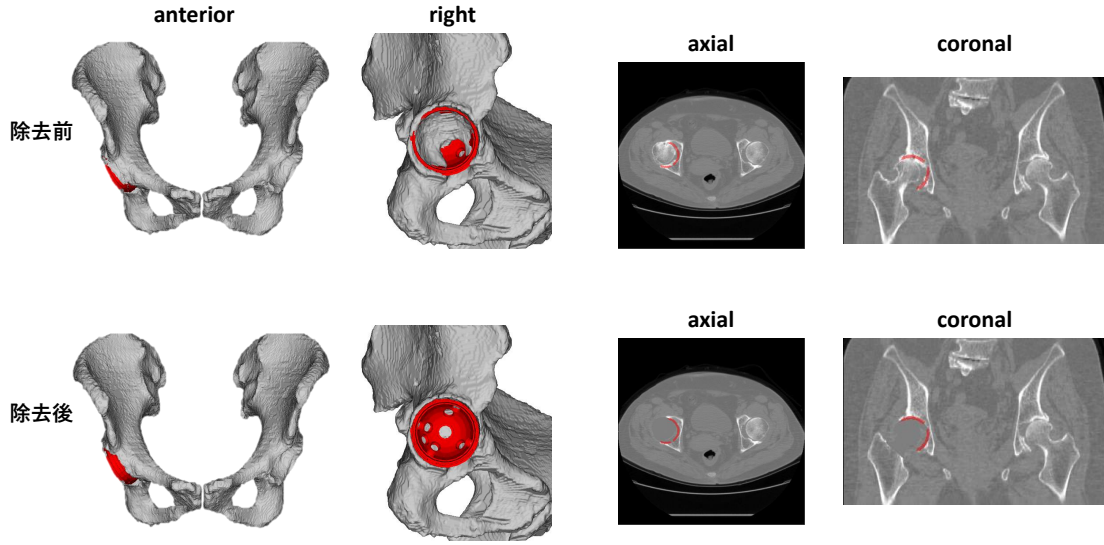


図 2.4 骨盤白蓋掘削のシミュレーション

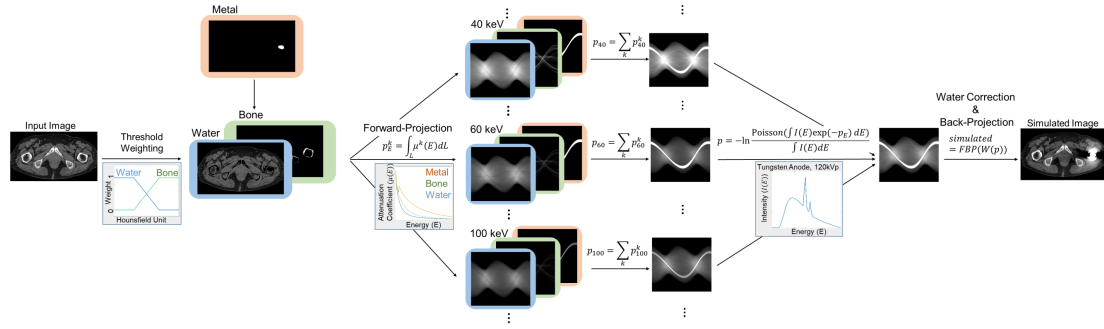


図 2.5 金属アーチファクトシミュレーションの流れ

μ への変換は、水の線減弱係数 $\mu_{water}(E)$ を用いて式 (2.3.4) で表される.

$$\mu = \mu_{water}(E) + \frac{\mu_{water}(E) \cdot HU}{1000} \quad (2.3.4)$$

続いて式 (2.3.5) に従い, $\mathbf{x}$ を軟組織 $\mathbf{x}^w$ および骨組織 $\mathbf{x}^b$ に分離する ([25] を一部変更).

$$\begin{aligned} x_i^b &= \min(\max(\frac{T_2(x_i - T_1)}{T_2 - T_1}, 0), T_2) \\ x_i^w &= x_i - x_i^b \end{aligned} \quad (2.3.5)$$

ただし x_i , x_i^w , x_i^b はそれぞれ $\mathbf{x}$, $\mathbf{x}^w$, $\mathbf{x}^b$ の i 番目の画素を意味する．本研究では [27] に従い，閾値 $T1$, $T2$ をそれぞれ 100HU, 1500HU に対応する線減弱係数値とした．

さらに，物質 k のあるエネルギー E における線減弱係数値は， E 及び E_0 での質量減弱係数値 $m^k(E)$ と $m^k(E_0)$ を用いて式 (2.3.6) によって算出される ([25] を一部変更)．

$$\mathbf{x}^k(E) = \frac{m^k(E)}{m^k(E_0)} \mathbf{x}^k(E_0) \quad (2.3.6)$$

以上で得られた $\mathbf{x}^w$, $\mathbf{x}^b$, $\mathbf{x}^m$ を用いると，X 線パス L_j に沿った波長 E での投影は式 (2.3.7) で表される ([25] を一部変更)．

$$p_j(E) = \sum_{k \in \{w, b, m\}} \int_{L_j} \mathbf{x}^k(E) dl \quad (2.3.7)$$

各エネルギーにおける物質の質量減弱係数値は XCOM dataset* から取得し，X 線のスペクトルはオンラインツール† (Siemens, Germany) を用いて算出した．また，得られたスペクトルを 1keV から 120keV を均等に 120 分割してサンプルしている．

各エネルギー E でのサイノグラム $p(E)$ と X 線量 $I(E)$ を用いて，多波長を仮定したサイノグラム p は式 (2.3.8) で得られる [25]．

$$p = -\ln \frac{\text{Poisson} \left(\int I(E) \exp(-p(E)) dE \right)}{\int I(E) dE} \quad (2.3.8)$$

以上で得られたサイノグラム p に対してビームハードニングを考慮した補正 [26] を行い，Hann フィルタを用いて再構成しシミュレーション画像を得る．このシミュレーションは MATLAB‡ (The Mathworks, Natic, MA) を用いて実装した．

*<http://www.nist.gov/pml/data/xcom/index.cfm>

†<https://www.oem-xray-components.siemens.com/x-ray-spectra-simulation>

‡<https://jp.mathworks.com/products/matlab.html>

2.4. 実験

2.4.1. 実験概要

以降で述べる条件の下で refinement U-net 及び segmentation U-net の学習を行った後、シミュレーション画像を用いたシミュレーション実験と、実術後 CT 画像を用いた実画像実験を行った。

シミュレーション実験での目的は、筋肉 19 種・骨 3 種の正解データが全て得られる術前 CT 画像を使用して提案法の定量評価を行う事である。一方、実画像実験では大臀筋と中臀筋のみの正解データを用いた精度評価と不確実性の比較によって、実画像への提案法の適用可能性について検討する。

2.4.2. データセット

本研究では、大阪大学医学部にて人工股関節全置換術（片側置換術のみ）を受けた患者の術前 CT 画像および術後 CT 画像を使用した。それらを表 2.1 に示す 4 つのデータセットに分類する。これらのデータセットには同一症例は含まれていない。

データセット A の筋骨格ラベルは整形外科医によって作成されている [12] が、データセット C の大臀筋・中臀筋のラベルは筆者が作成し、整形外科医による確認を受けたものである。また、データセット C の術後 CT 画像は金属アーチファクトを含むため、NMAR を適用したアーチファクト低減後の画像を用いて筋骨格ラベルを作成している。筋肉ラベルと、その表示に用いる色の対応は図 2.6 に示す通りである。また、データセット A・C・D の症例においては、体表のセグメンテーション [5] を前処理として適用している。

2.4.3. refinement U-net

refinement U-net の学習には、データセット B を使用した。専門医が作成した手術計画に従ってインプラントを挿入し (2.3.3 節)、金属アーチファクトのシミュレーションを行った。ただし、データセット B では手動で作成した骨ラベルが得られないため、データセット A で事前に学習した Bayesian U-net [5] によって推定された

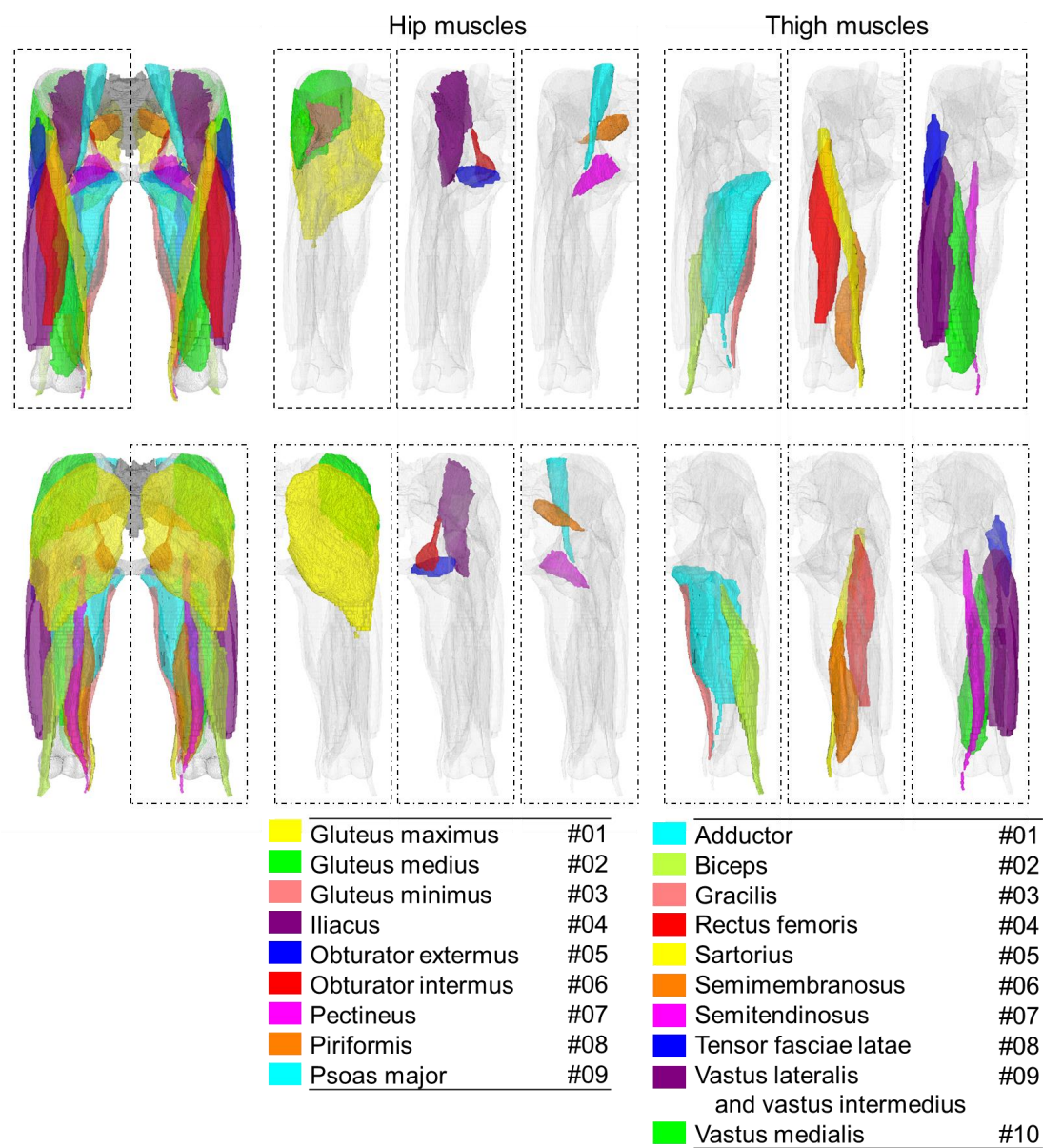


図 2.6 筋肉ラベルと対応する色 [5] © 2019 IEEE.

表 2.1 使用したデータセット

データセット	症例数	時相	筋骨格ラベル	手術計画	体表ラベル
A	20	術前	筋肉 19 種・骨 3 種	なし	あり
B	122	術前	なし	あり	なし
C	3	術後	筋肉 2 種 (大臀筋・中臀筋)	なし	あり
D	129	術後	なし	なし	あり

ラベルを使用した。得られたシミュレーション画像に NMAR を適用した結果と術前 CT 画像をペアとして、refinement U-net の学習データを作成した (図 2.2(b))。ただし refinement U-net の学習にはインプラントが含まれるスライスのみを使用し、また前処理として、 $[-150, 350]$ HU の範囲の輝度値を $[-1, 1]$ に線形に変換した。ネットワークの出力層の活性化関数は $\tanh$ とし、重みの初期化には He ら [28] の手法を用いた。損失関数は平均二乗誤差とし、バッチサイズは 2、学習係数を 0.0001 として Adam[29] による最適化を行った。さらに data augmentation として水平方向の反転、平行移動 ($-25\% \sim 25\%$)、せん断 ($-\pi/8 \sim \pi/8$ rad)、回転 ($-10 \sim 10$ 度) を行った。

2.4.4. segmentation U-net

データセット A を使用し、segmentation U-net を学習した。2.3.3 節で述べた手法でインプラントの CAD モデルを挿入し、値を置換した骨領域は正解ラベルを削除した。ただし、データセット A では各症例に対応する手術計画が得られなかったため、3D Slicer[§]を使用し、手動で CAD モデルを挿入している (図 2.7)。入力画像の前処理として、 $[-150, 350]$ HU の範囲の輝度値を $[0, 1]$ に線形に変換した。ネットワークの出力層の活性化関数は Softmax であり、重みの初期化には He ら [28] の手法を用いた。また、プーリング層は 5 層使用し、ReLU 層の前に BatchNormalization[30] を適用した。Dropout 層の確率は 0.5 とし、10 回の推論

[§]<https://www.slicer.org/>

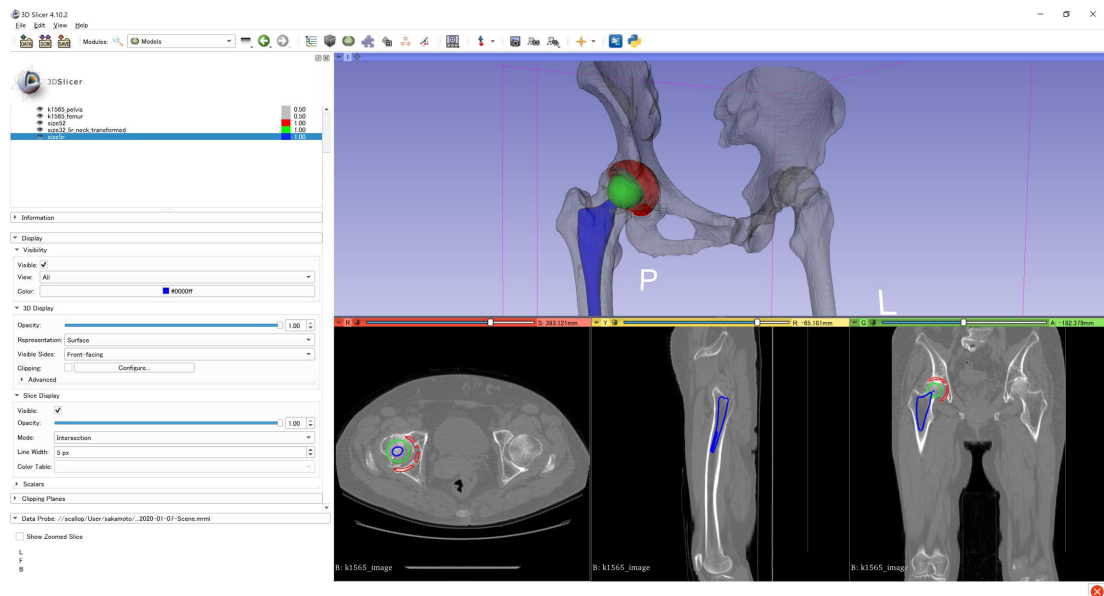


図 2.7 3D Slicer⁸ を用いた術前 CT 画像への CAD モデルの挿入

結果の平均・分散をそれぞれ推定されたラベルの確率値および不確実性として最終出力を得た。損失関数は Softmax cross entropy とし、バッチサイズは 2、学習係数を 0.0001 とし Adam[29] による最適化を行った。data augmentation では、水平方向の反転、平行移動 ($-25\% \sim 25\%$)、せん断 ($-\pi/8 \sim \pi/8$ rad)、回転 ($-10 \sim 10$ 度)、random erasing[31] を行った。また、本稿では最大連結成分の抽出などの後処理は行わず、segmentation U-net から得られた結果を最終出力として評価に使用している。

2.4.5. シミュレーション実験

segmentation U-net で 2 分割交差検証を行い、以下の 4 種類の画像について、セグメンテーション精度及び不確実性を比較した。ただし refinement U-net については、データセット B の症例全てを用いて学習したモデルを使用した。

1. 術前 CT 画像
2. シミュレーション画像

3. 2. に NMAR を適用した画像
4. 2. に NMAR と refinement U-net を適用した画像

2.4.6. 実画像実験

実画像実験では，以下の 3 種類の画像について，データセット C と D を用いて評価した．

1. 術後 CT 画像
2. 1. に NMAR を適用した画像
3. 1. に NMAR と refinement U-net を適用した画像

データセット C では，大臀筋及び中臀筋のセグメンテーション精度と不確実性について評価した．データセット D では筋骨格の正解ラベルが作成されていないため，不確実性についての評価を行った．

2.4.7. 評価指標

セグメンテーション精度の評価指標として，Dice 係数 (Dice Coefficient: DC) [32](式 (2.4.1)) 及び平均表面間距離誤差 (Average Symmetric Surface Distance: ASD) [33](式 (2.4.2)) を使用した．2 つのラベル A , B での Dice 係数と ASD は以下の式で求められる．

$$DC(A = c, B = c) = \frac{2|A \cap B|}{|A| + |B|} \quad (2.4.1)$$

$$ASD(A = c, B = c) = \frac{1}{2}(AAD(A, B) + AAD(B, A)) \quad (2.4.2)$$

$$AAD(A, B) = \frac{1}{|\partial\Omega_A|} \sum_{\mathbf{p}_a \in \partial\Omega_A} \min_{\mathbf{p}_b \in \partial\Omega_B} \|\mathbf{p}_a - \mathbf{p}_b\|^2 \quad (2.4.3)$$

ここで， $\partial\Omega_A$ と $\partial\Omega_B$ はそれぞれラベル A と B でクラス c に属する点の集合である．

一方, 不確実性の指標としては, Predictive Variance (PV) と, Predictive Structure wise Variance (PSV)[5] を使用した. PV は出力された各ラベルの確率値の分散を全体に渡って平均したものである. 一方, PSV は各ラベルにおいてそのラベルであると推定された領域内の不確実性の平均値であり, 未知の画像 $\mathbf{x}^*$ におけるクラス c の PSV は式 (2.4.4) で表される [5]. ここで $\mathbf{s}^*$ は $\mathbf{x}^*$ においてクラス c であると分類された画素の集合であり, $\text{Var}(y^* = k|\mathbf{x}^*)$ は dropout samplings で得られた, クラス k の Softmax 値の分散を示す.

$$PSV(\mathbf{s}^*|\mathbf{x}^*) = \frac{1}{|\mathbf{s}^*|} \sum_{y^* \in \mathbf{s}^*} \sum_k \text{Var}(y^* = k|\mathbf{x}^*) \quad (2.4.4)$$

2.5. 実験結果

2.5.1. シミュレーション実験

2.4.5 節の 4 種類の各画像において, 全筋肉での Dice 係数の平均及び標準偏差は 0.880 ± 0.0619 , 0.775 ± 0.179 , 0.854 ± 0.0812 , 0.870 ± 0.0677 であり, また ASD は 1.10 ± 0.681 mm, 2.46 ± 2.29 mm, 1.37 ± 0.884 mm, 1.20 ± 0.798 mm であった. wilcoxon の符号付順位和検定 (Bonferroni 法で補正) を行った結果, NMAR のみを使用した場合と refinement U-net を適用した場合で, Dice 係数では 19 種類の筋肉のうち 15 種類で, ASD では 14 種類で有意差が見られた ($p < 0.05$).

図 2.8, 図 2.9 には各筋肉のセグメンテーション精度を示す. また, 図 2.10, 図 2.11, 図 2.12 には骨 (骨盤・大腿骨)・臀部の筋肉・大腿部の筋肉に分けて, Dice 係数と PSV の相関を示す. 図 2.13 はセグメンテーション結果と PV の 3 次元可視化である. 図 2.14 はデータセット A の 20 症例について, インプラントを含むスライスでの PV の平均及び標準偏差をプロットしたものであり, エラーバーは標準偏差を示す. ただしインプラントを含むスライス数は症例によって異なるため, B-spline 法によって全て 100 スライスに補間し, それらの平均と標準偏差を計算している.

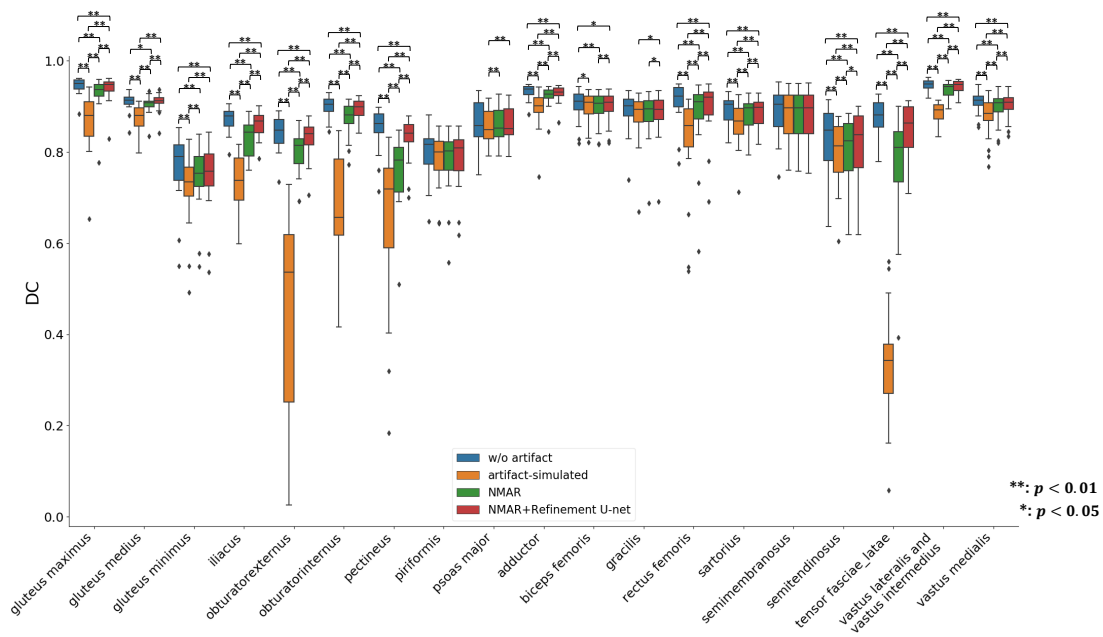


図 2.8 シミュレーション実験での各筋肉・骨の Dice 係数

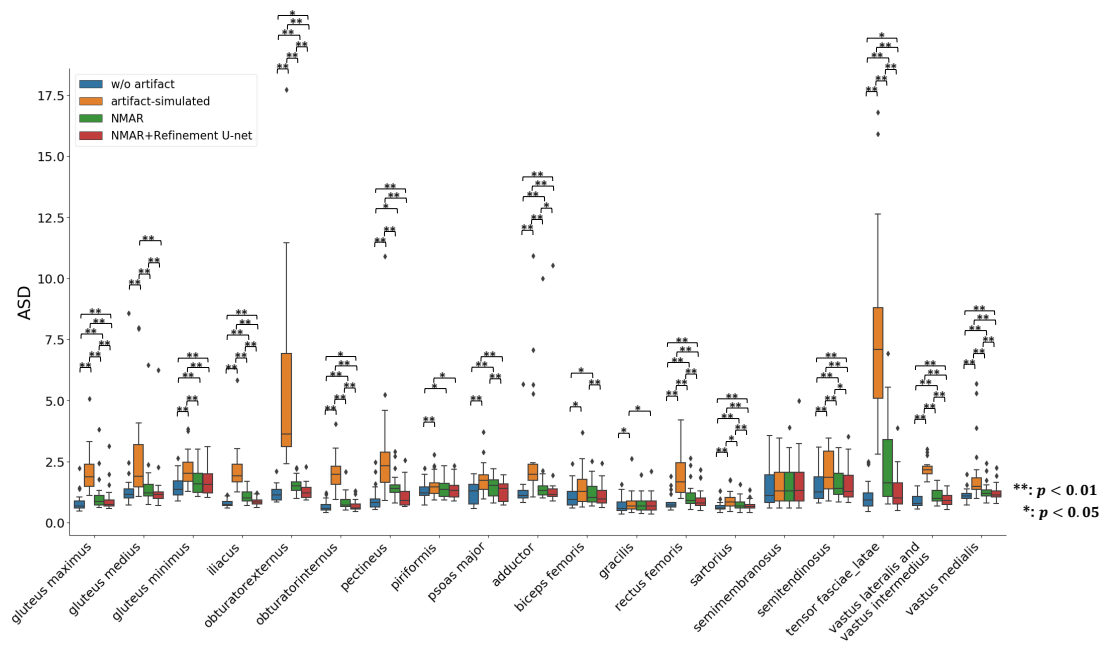


図 2.9 シミュレーション実験での各筋肉・骨の ASD

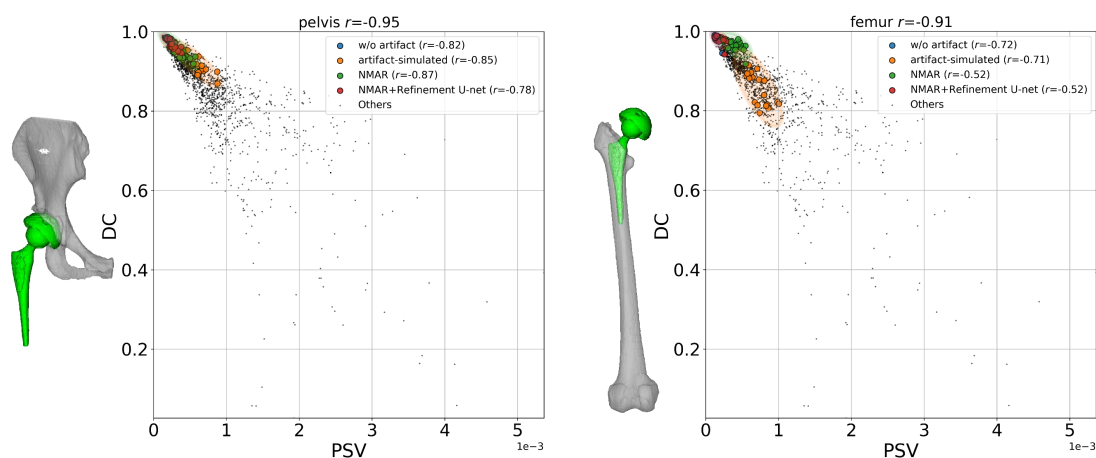


図 2.10 骨盤・大腿骨の Dice 係数と PSV の関係

2.5.2. 実画像実験

図 2.15, 図 2.16 に, データセット C の 3 症例での大臀筋・中臀筋の結果を示す. 図 2.15, 図 2.16 では正解ラベルと推定結果の表面間距離誤差を正解ラベルにマッピングして表示している. 2.4.6 節の 3 種類の各画像において, 大臀筋・中臀筋の Dice 係数の平均および標準偏差は 0.814 ± 0.0524 , 0.871 ± 0.0318 , 0.873 ± 0.0335 であり, また ASD は 3.90 ± 3.90 mm, 1.55 ± 0.253 mm, 1.55 ± 0.249 mm であった. 図 2.17 はシミュレーション実験同様に, セグメンテーション結果と不確実性を 3 次元及び 2 次元で可視化したものである. 図 2.18 はデータセット C・D の計 132 症例について, シミュレーション実験同様にインプラントを含むスライスでの PV の平均及び標準偏差をプロットしたものである. さらに, データセット D に含まれる症例での結果の一部を図 2.19 に示す.

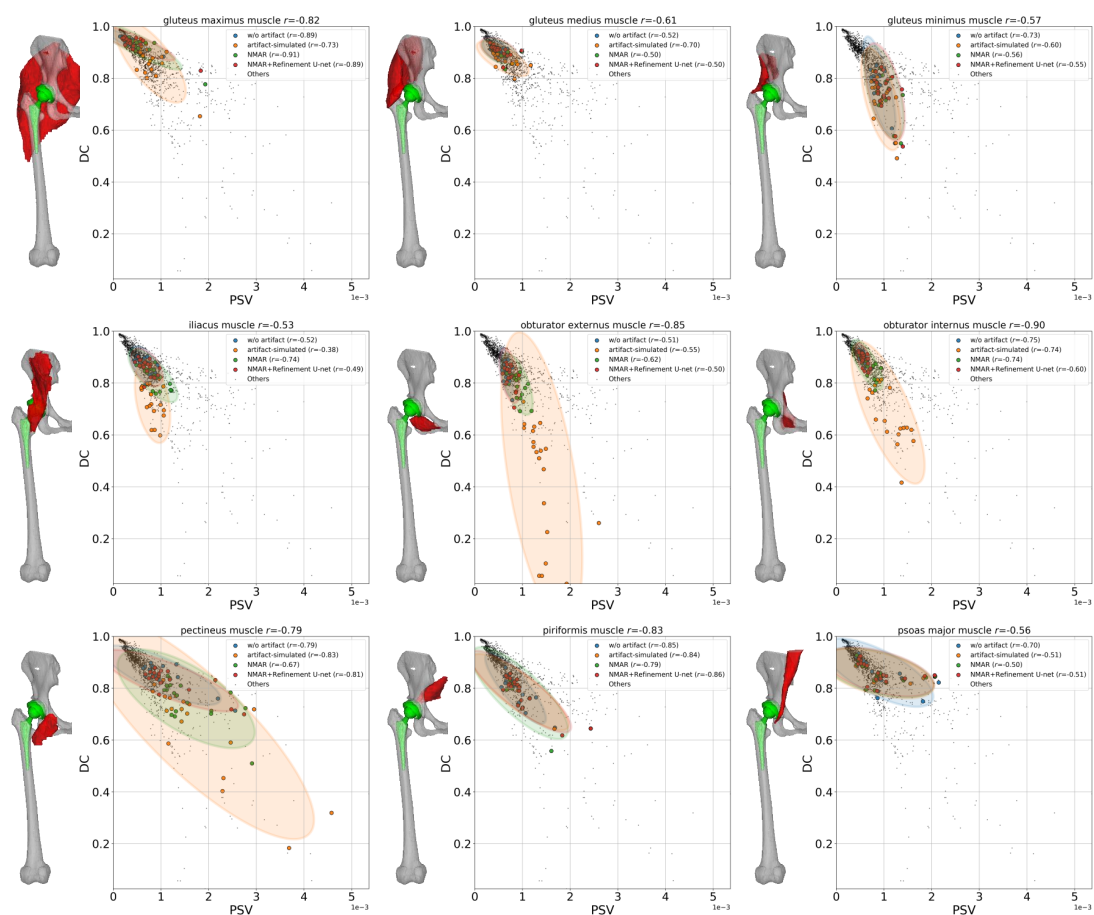


図 2.11 臀部の筋肉（図 2.6）の Dice 係数と PSV の関係．各プロットの左の図は対象の筋肉（赤）とインプラント（緑）の位置関係を示す．

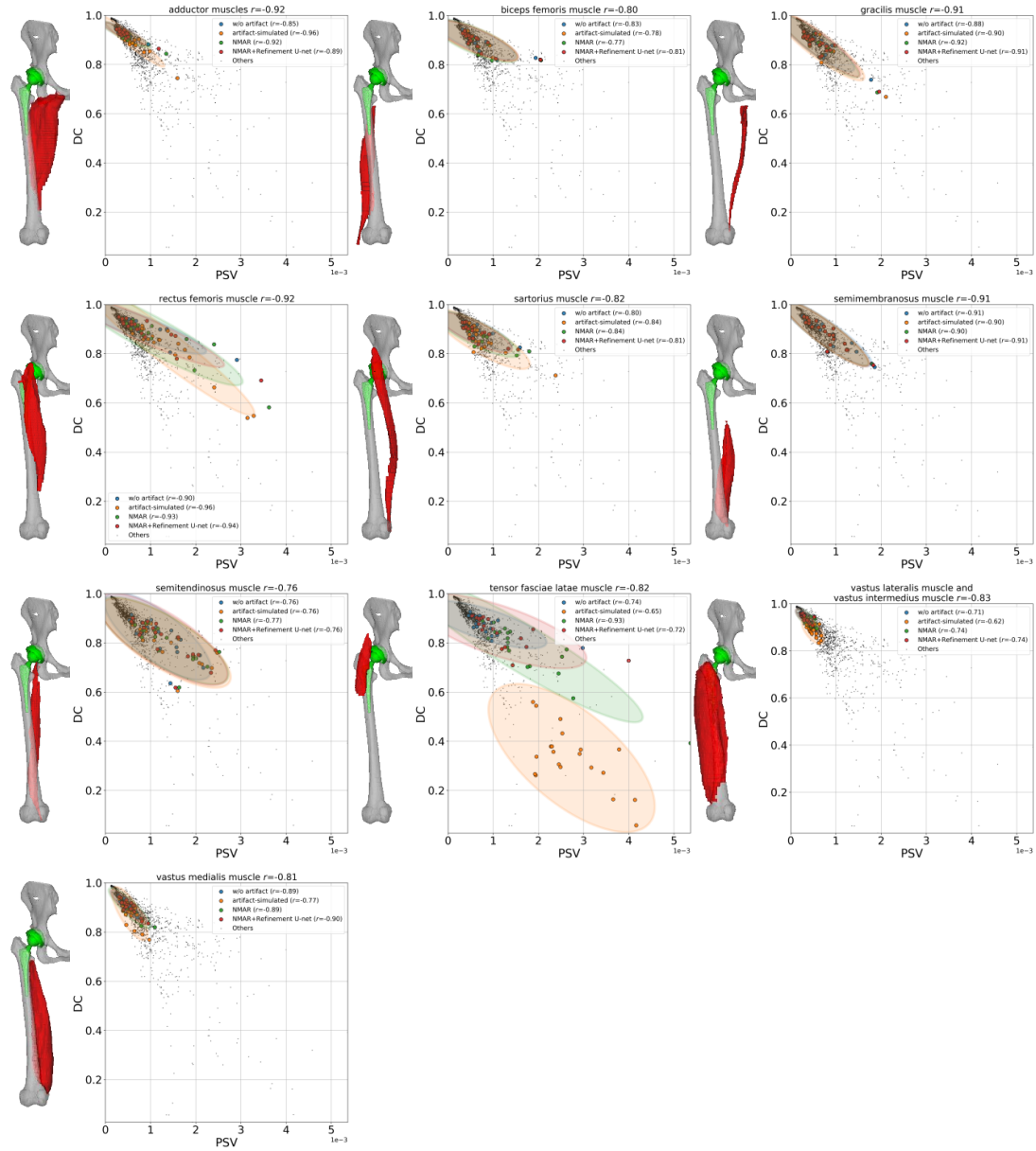


図 2.12 大腿部の筋肉（図 2.6）の Dice 係数と PSV の関係。各プロットの左の図は対象の筋肉（赤）とインプラント（緑）の位置関係を示す。

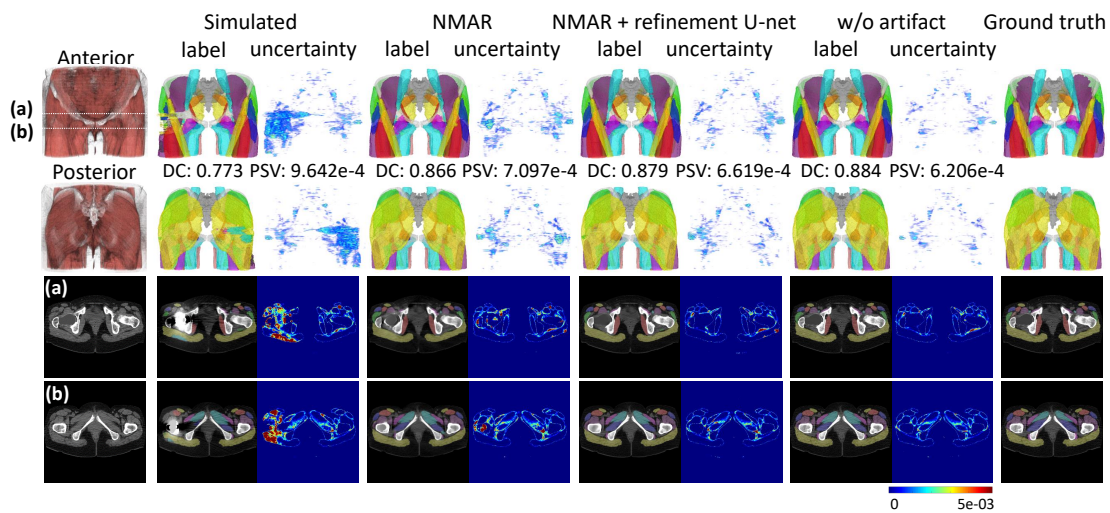


図 2.13 シミュレーション実験でのセグメンテーション結果及び不確実性の可視化

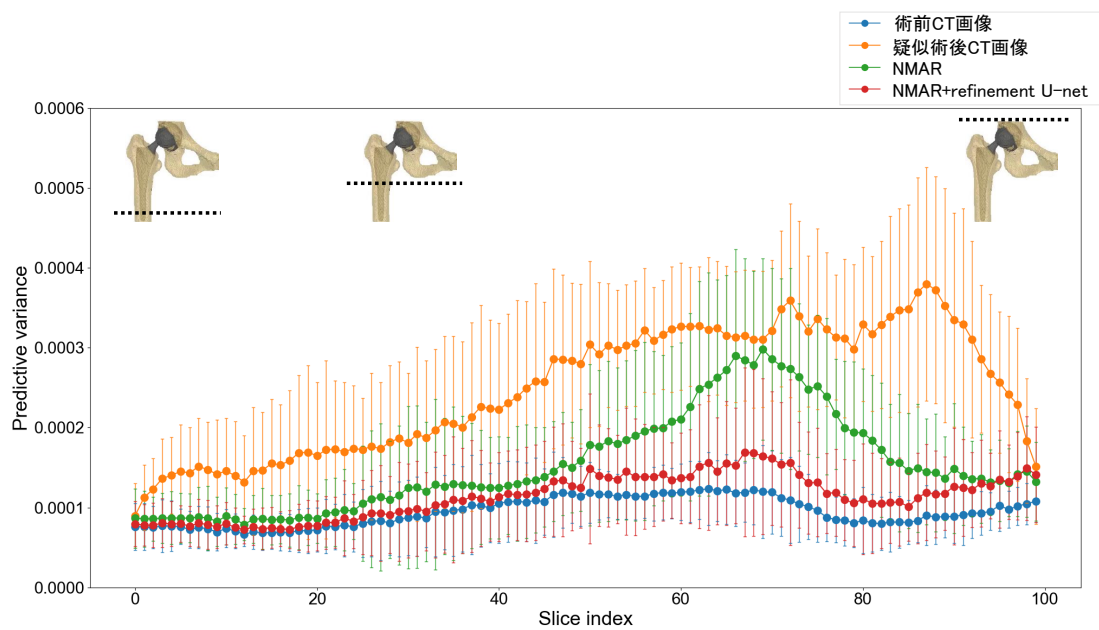


図 2.14 シミュレーション実験での不確実性のプロット（20 症例）

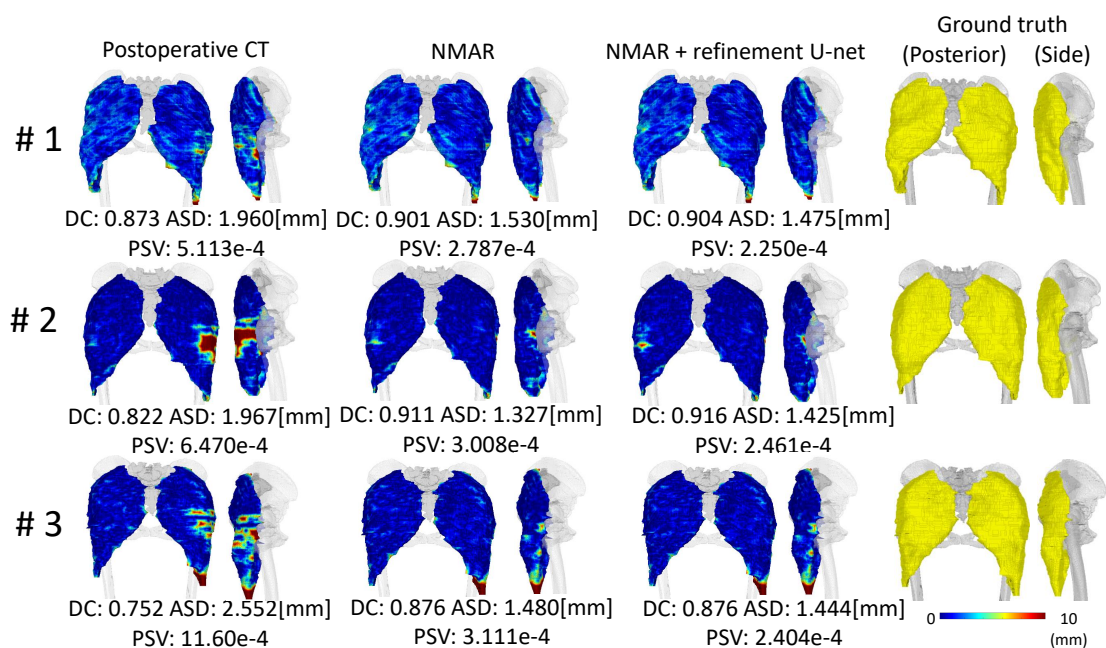


図 2.15 実画像における大臀筋の精度評価

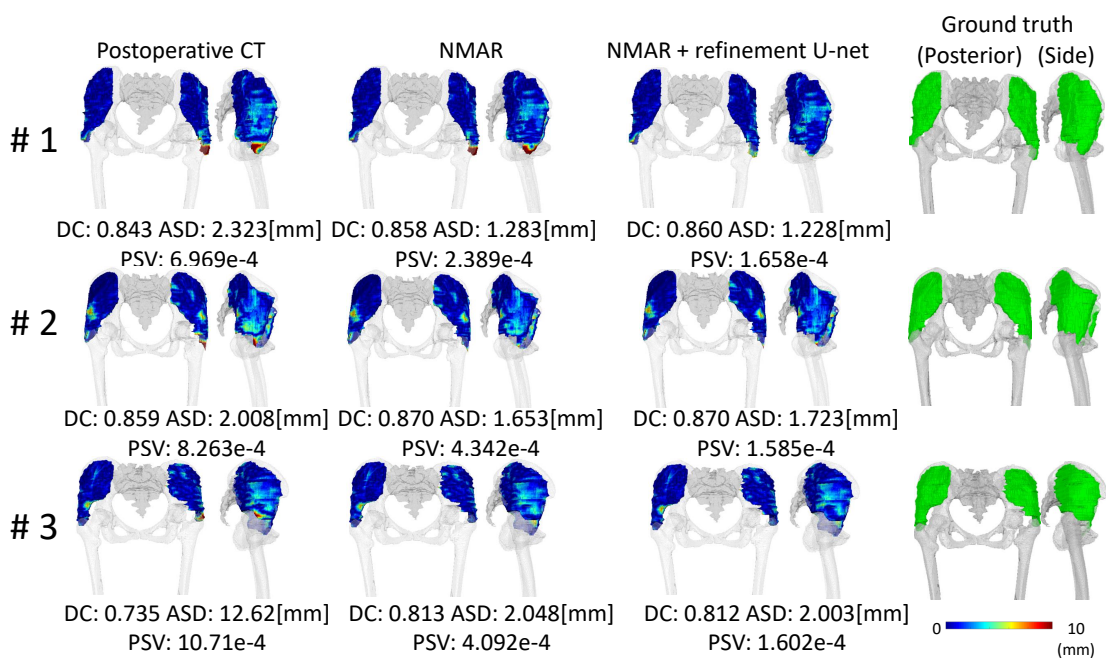


図 2.16 実画像における中臀筋の精度評価

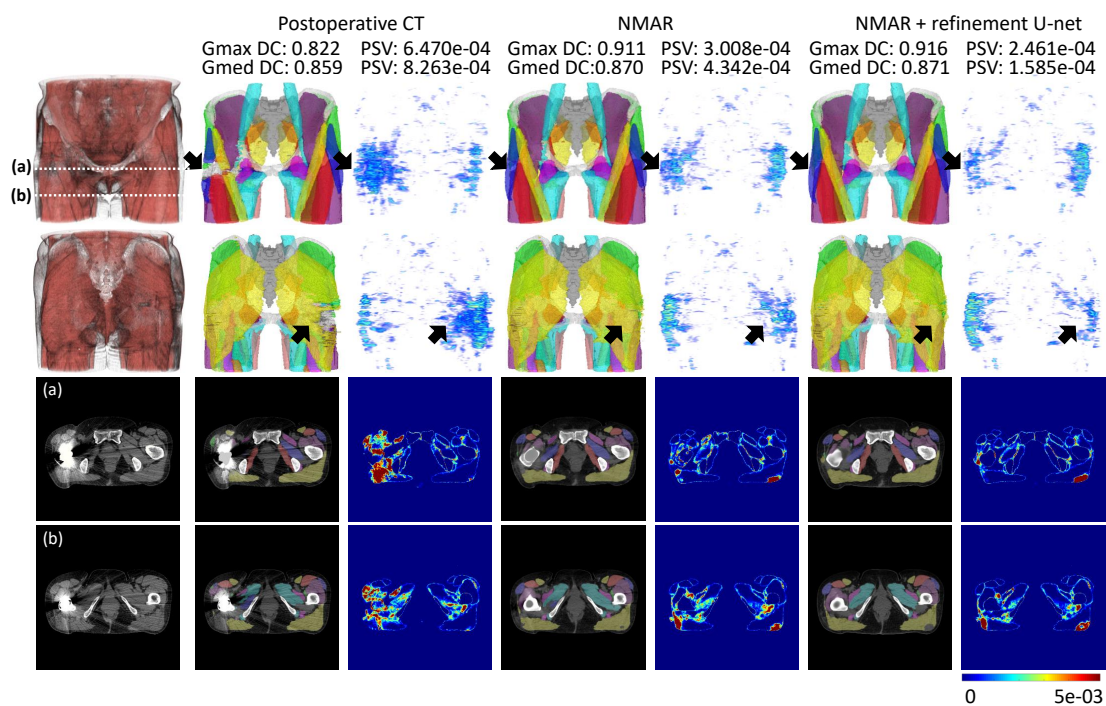


図 2.17 実画像実験でのセグメンテーション結果及び不確実性の可視化

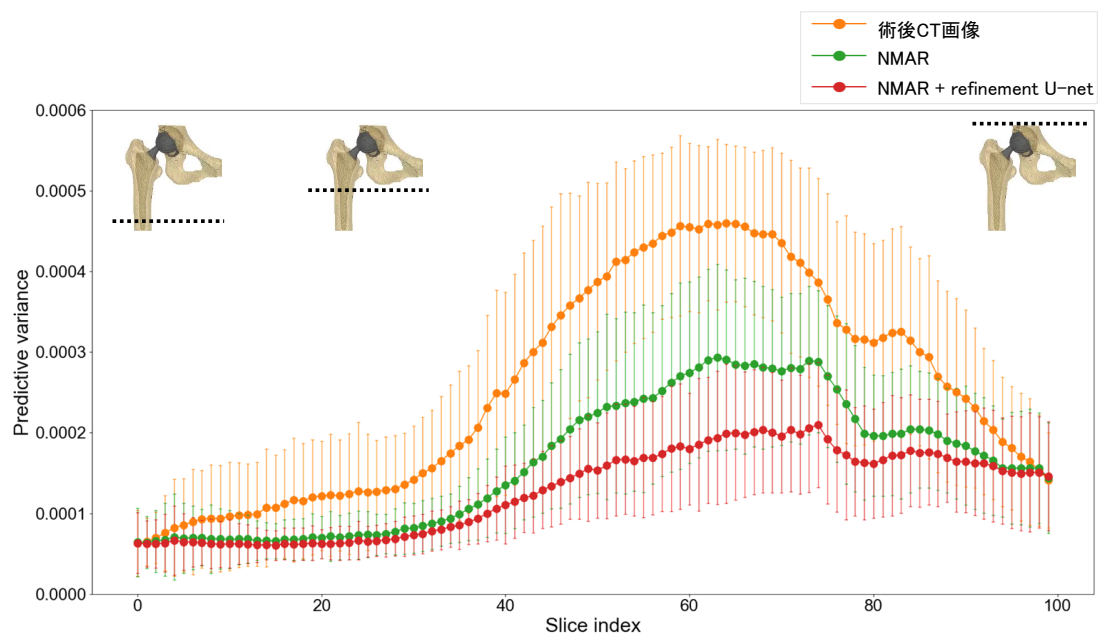


図 2.18 実画像での不確実性のプロット (132 症例)

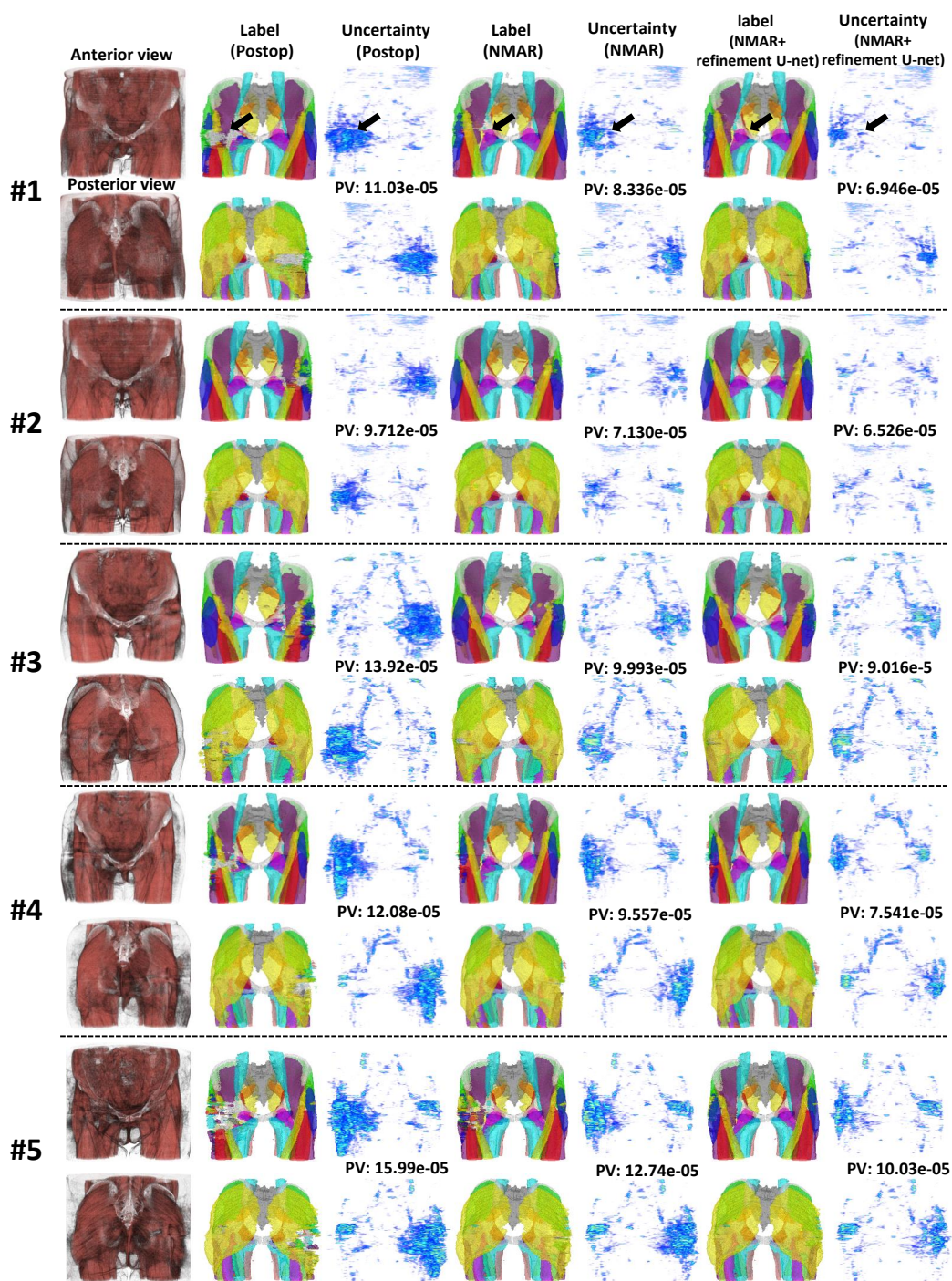


図 2.19 実画像への提案法の適用結果（左から，術後 CT 画像のボリュームレンダリング，術後 CT 画像，NMAR 適用後の画像，さらに refinement U-net を適用した画像でのセグメンテーション結果及び不確実性）

2.6. 考察

2.6.1. セグメンテーション精度

シミュレーション実験では 19 種類の筋肉のうち、15 種類の筋肉で有意に Dice 係数が向上した。精度が向上しなかった筋肉はインプラントから離れた位置にあり、金属アーチファクトによる影響が小さかったと考えられる。例として大腿二頭筋 (biceps femoris muscle) や半膜様筋 (semimembranosus muscle) が挙げられる (図 2.6)。

一方、実画像実験においても、大臀筋及び中臀筋において、金属アーチファクト低減によるセグメンテーション精度の向上が定量的かつ定性的に確認できた (図 2.15, 図 2.16)。ただし、NMAR を適用した場合とさらに refinement U-net を適用した場合で、大きな差は確認できなかった。この点について、実画像実験ではアーチファクト低減後の画像を用いて正解データを作成したために精度の差が小さくなっていると考えられ、アーチファクトなしの画像で正解データを作成したシミュレーション実験との実験条件の違いが問題であると考えられる。

実画像での正解データ作成について、Wang ら [34] は内耳部分の金属アーチファクト低減において同一患者の術前 CT 画像を術後 CT 画像に位置合わせし、変形した術前 CT 画像を正解データとして使用している。しかし、術前後での非剛体変形が大きい股関節領域においては、位置合わせした術前 CT 画像が必ずしも正確であるとは限らないという問題がある。

2.6.2. セグメンテーションの不確実性

シミュレーション実験と実画像実験の両方において、金属アーチファクト低減に伴い、セグメンテーションの不確実性が低下している事が確認できた。さらにシミュレーション実験からは、金属アーチファクトを含む画像においてもセグメンテーション精度と不確実性 (PSV) に相関がある事が示された (図 2.10, 図 2.11, 図 2.12)。この結果から、金属アーチファクトを含む画像においても、不確実性をセグメンテーション精度の指標として利用できる可能性が示唆された。

ただし実画像実験では、NMAR を適用した画像とさらに refinement U-net を

適用した画像において、不確実性に差がある一方で精度には差が見られず、セグメンテーション精度と不確実性に明らかな相関があるとは言えない結果であった。その原因としては、2.6.1 節と同様の、実画像実験における正解データ作成の問題が考えられる。一方、図 2.14 及び図 2.18 に示すように、シミュレーション実験と実画像実験のどちらでも、金属アーチファクトを含む画像で最も不確実性が高く、refinement U-net を適用した場合に最も不確実性が低い結果になった。このことから金属アーチファクト低減によって不確実性が低下していると考えられ、金属アーチファクト低減の定量的指標としてセグメンテーションの不確実性が利用できる可能性が示唆された。

2.6.3. 今後の課題及び展望

提案法は、現時点では金属アーチファクト低減と筋骨格セグメンテーションを独立に学習しているが、シミュレーション画像を用いてこれらを end-to-end に学習させることも可能である。end-to-end での学習によって最終目的とするタスクの精度を向上させた例として、Wu ら [35] は CT 画像の再構成と病変検出を end-to-end で学習する事で、病変検出のみを学習した場合よりも病変検出精度を向上させている。ただし、end-to-end での学習のために金属アーチファクトを含む画像を学習データに加えた場合、金属アーチファクトによる誤差を不確実性として捉えることが困難になる可能性があるため、end-to-end での学習のメリット・デメリットに関してのさらなる調査が必要である。

また、本研究で使用しているシミュレーションは、実画像における refinement U-net の精度に影響を及ぼす。シミュレーション画像と実画像の間に乖離が生じる場合、refinement U-net は実画像に適用すると精度が低下する可能性がある。このような乖離はドメインシフトと呼ばれ、Shrivastava ら [36] は敵対的学習によって、生成画像と実画像のドメインシフトを解消する手法を提案している。著者が所属するグループでも、シミュレーション画像と実画像の金属アーチファクト低減後が見分けられなくするような敵対的学習を採用し、実画像を学習に取り入れた金属アーチファクト低減手法について検討してきた [46]。しかしシミュレーション実験での定量評価に留まっており、実画像での定量評価やセグメンテーションの精度に与え

る影響の調査が今後の課題である。

今後の展望としては、金属アーチファクトに由来するセグメンテーションの不確実性を利用した、金属アーチファクト低減の定量評価が挙げられる。従来、金属アーチファクト低減において、実画像ではその正解データを得られないため定量評価が難しく、定性評価 [25] や画像の一部のみを使用した定量評価 [37] に留まっており、定量評価は主にシミュレーション画像によって行われてきた [25]。一方でセグメンテーションの不確実性はセグメンテーションの正解データ無しに計測可能な値である。そのため金属アーチファクト低減精度の指標としてセグメンテーションの不確実性が利用可能であれば、実画像においてアーチファクトなしの画像なしに金属アーチファクト低減の定量評価が可能であり、その意義は高いと考える。ただし、このためには金属アーチファクトに由来する不確実性のみを取り出す必要があるが、本研究では不確実性を分離するには至っていない。学習データに含まれる筋骨格形状のバリエーションが不十分である場合、未知の症例での推論時には金属アーチファクトだけでなく筋骨格形状に由来する不確実性が含まれる。今後は、このような複数の要因からなる不確実性をそれぞれに分離する手法の開発が求められる。

2.7. 小括

本章では、金属アーチファクト低減手法と筋骨格セグメンテーション手法を組み合わせ、人工股関節全置換術の術後 CT 画像からの筋骨格セグメンテーション手法を提案した。また提案法の有効性をシミュレーション実験及び実画像実験から定量的に評価した。

シミュレーション実験では術前 CT 画像から得られた正解データを使用し、提案法によってセグメンテーション精度が向上するだけでなく、金属アーチファクトを含む術後 CT 画像においても精度と不確実性に相関がある事が示された。また、実画像実験においてもシミュレーション実験同様に金属アーチファクト低減に伴って不確実性が低下する傾向が見られ、提案法の実画像における有効性が示唆された。

最後に今後の展望として、金属アーチファクトに由来するセグメンテーションの不確実性を用いた金属アーチファクト低減の定量評価について述べた。

第 3 章 術後 CT 画像からのインプラントセグメンテーション

3.1. はじめに

人工股関節全置換術において、インプラント同士やインプラント周辺の骨のインピンジメント（衝突）が原因となり、術後に痛みや脱臼が生じる事がある [38, 9]. これまでに Shoji ら [10] は骨・インプラントのインピンジメントについて、術前 CT 画像から得られた骨形状を用いて解析を行なっている. また Iwana ら [9] は術後 CT 画像上で手動でカップ設置角を計測し、術前に計画した角度との違いについて検証している. このように、術前・術後でのインプラントの設置位置・角度の評価は非常に重要であり、これまでに手動もしくは半手動での解析が多く行われてきた. 術後 CT 画像から自動で、カップ・ヘッド・ステムの正確な 3 次元形状・設置位置・角度を得ることができれば、より多くの症例での解析が可能となる.

これまで、術後 CT 画像からの自動での金属セグメンテーションは主に閾値を用いた手法によって行われてきた [19]. 閾値によるセグメンテーションで概ね良好なインプラント形状が得られる場合もあるが、カップ・ヘッド・ステムの各部位に分けて形状が得られない点が課題である. 閾値以外での手法としては Hegazy ら [39] は U-net[18] を用いた投影像からの歯科用インプラントのセグメンテーション手法を提案し、閾値でのセグメンテーションよりも正確に金属領域が得られることを示したが、3 次元形状での評価はなされていない.

股関節領域のインプラントは歯科用のインプラントなどと比較するとそのバリエーションには限りがあり、インプラントの CAD モデル（図 3.2）を学習データとすることで学習可能であると考え. そこで本研究では、過去症例の術後 CT 画像と対応する CAD モデルから正解データを作成し、その正解データで学習した CNN によって術後 CT 画像からインプラントセグメンテーションを行う.

ただし、CT 画像と対応する正確な CAD モデルが常に利用可能であるとは限らず、推論時には CAD モデルなしにインプラント形状が得られる手法が求められる. また、特にステムには 6 種類の形状のバリエーション (single wedge 型, double wedge metaphyseal filling 型, tapered 型, cylindrical fully coated 型, modular

型, anatomical 型) があり [44], 学習データに含まれない形状への汎用性も重要である. そのため, 後述する実験では学習データに含まれないインプラント機種にも提案法を適用し, 未知の機種に対する汎用性の初期的検討を行う. インプラントの自動セグメンテーションが実現できれば, より大規模なデータベースでの解析が可能となると考えられる.

3.2. 手法

3.2.1. CAD モデルの位置合わせによる正解データ作成

CAD モデルを用いた正解データ作成について述べる. CT 画像から閾値で抽出したインプラントと, 使用された CAD モデルを剛体位置合わせすることで学習データを作成する. 本研究では, 2000HU を閾値として抽出したインプラント形状と CAD モデルとの表面間距離誤差を CMA-ES (Covariance Matrix Adaption Evolution Strategy)[40] によって姿勢パラメータ Θ に関して最小化し, 位置合わせを行う. この最適化は式 (3.2.1) で定式化される.

$$\hat{\Theta} = \underset{\Theta \in R^{10}}{\operatorname{argmin}} AAD(p_{\text{CAD}}(\Theta), p_{\text{threshold}}) \quad (3.2.1)$$

ここで, Θ はインプラントの姿勢 (10 自由度: カップ・ヘッドの並進 (x, y, z), カップの回転 (x, y, z), ステムの回転 (x, y, z), ステム並進 (z)) を表す. $p_{\text{threshold}}$ は閾値で取得したインプラント形状であり, p_{CAD} はカップ・ヘッド・ステム (p_{cup} , p_{head} , p_{stem}) で構成されるインプラントの CAD モデルである.

また, カップ・ヘッドの並進を表す変換を T_{global} とすると, カップ・ヘッド・ステムそれぞれの変換は以下で表される.

$$T_{\text{cup}} = T_{\text{cup_rotation}} * T_{\text{global}} \quad (3.2.2)$$

$$T_{\text{head}} = T_{\text{global}} \quad (3.2.3)$$

$$T_{\text{stem}} = T_{\text{stem_offset}} * T_{\text{stem_rotation}} * T_{\text{global}} \quad (3.2.4)$$

ここで, $T_{\text{cup_rotation}}$ と $T_{\text{stem_rotation}}$ はそれぞれカップとステムの回転を表し, $T_{\text{stem_offset}}$ はステムの z 軸方向の並進を表す. 本研究では使用したインプラント

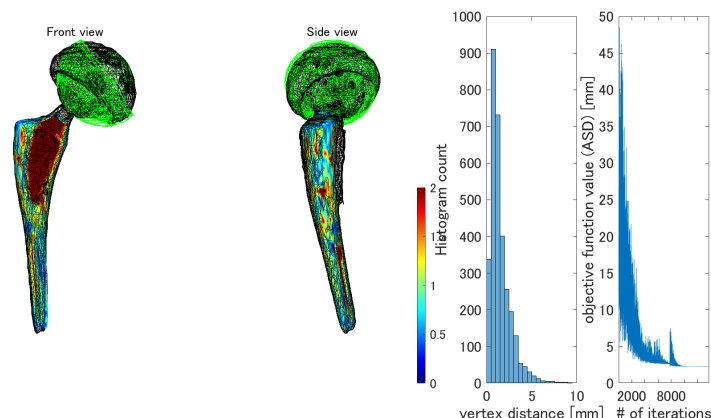


図 3.1 カップの位置合わせの失敗例（黒のワイヤフレーム: $p_{\text{threshold}}$, 緑: CAD モデル p_{CAD} ）

の形状及びサイズは既知であると仮定し、インプラントの姿勢についてのみ最適化を行う。また、計算時間の短縮を図り 2 段階での最適化を行っている。表 3.1 に、位置合わせに用いた最適化パラメータを示す。本研究では [49] の実装を元に実験を行っている。

位置合わせ後のカップの CAD モデルから閾値で抽出した表面形状までの表面間距離誤差が 1.2mm 以上であった場合、位置合わせを最大 20 回まで繰り返す。また、位置合わせの結果を目視で確認し、カップの位置合わせで局所解に陥った症例（図 3.1）は学習データから除いた。

3.3. 実験

3.3.1. データセット

大阪大学医学部にて人工股関節全置換術（片側置換術のみ）を受けた患者の術後 CT 画像計 111 症例を使用した。撮像範囲は約 $360 \times 360 \text{mm}^2$ 、マトリックスサイズは 512×512 、スライス間隔は 1mm である。片側全置換術を受けた症例のみを対象としており、表面置換や両側置換の症例は含まれていない。本研究ではカップ・ヘッド・ステムから構成されるインプラントを想定し（図 3.2）、特に形状のバリエー

表 3.1 最適化において使用したパラメータ

		1st stage	2nd stage
population size		300	200
ポリゴン数	p_{cup}	2000	2000
	p_{head}	500	500
	p_{stem}	1000	4000
	$p_{\text{threshold}}$	3000	15000
initial sigma	カップ・ヘッド並進 (x, y, z)	(10, 10, 10)	(1, 1, 1)
	カップ回転 (x, y, z)	(60, 60, 60)	(6, 6, 6)
	ステム回転 (x, y, z)	(10, 60, 60)	(1, 6, 6)
	ステム並進 (z)	1	3
	カップ・ヘッド並進 (x, y, z)	$p_{\text{threshold}}$ と p_{CAD} の バウンディングボックスの 中心間の変位ベクトル	
	カップ回転 (x, y, z)	(180, 0, 0)	
initial guess	ステム回転 (x, y, z)	(180, 180, -90)	
	ステム並進 (z)	-10	

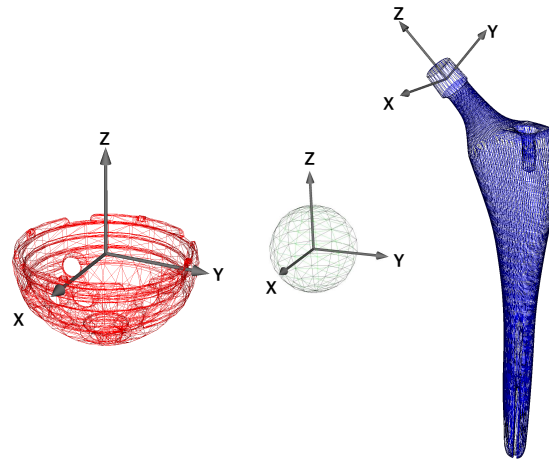


図 3.2 インプラントの CAD モデル（赤：カップ，緑：ヘッド，青：ステム）の構造と座標系（いずれも右手系）

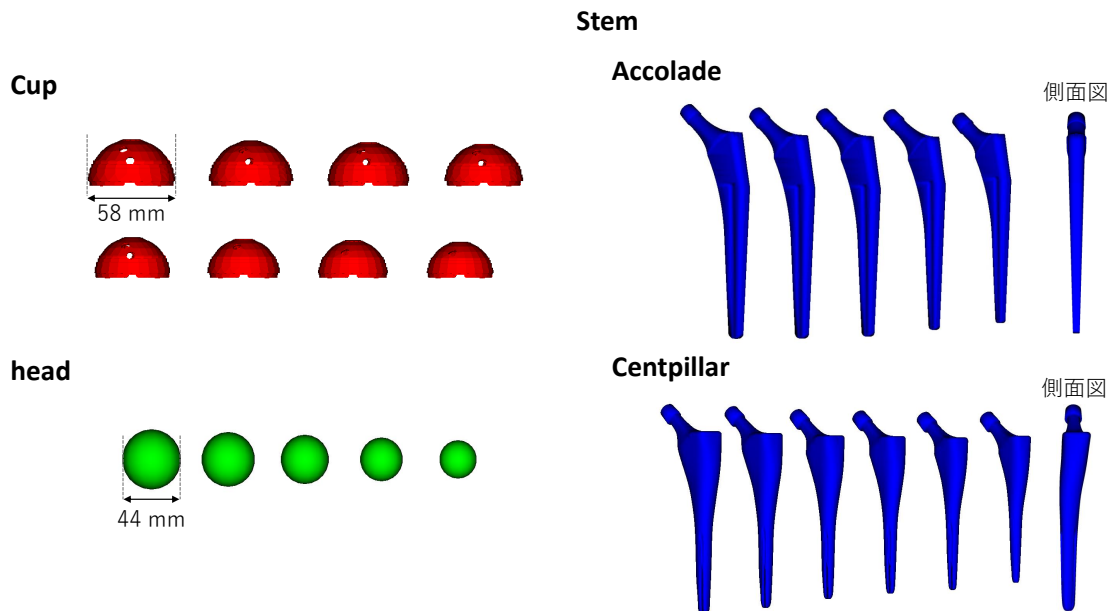


図 3.3 インプラントの CAD モデルのバリエーション

ションが大きいステムでは，異なる形状に分類される 2 機種（Centpillar TMZF (Stryker), Accolade TMZF (Stryker)）を使用している．ステム形状のデザインは straight 型と anatomical 型に大別される [48]．Centpillar はその anatomical 型に属す機種で，Accolade は straight 型のサブカテゴリである single wedge 型に分類される機種である．Centpillar は大腿骨髓腔形状に適合するようにデザインされており，Accolade と比較すると特に側面から見た形状に違いが見られる（図 3.3）．カップは表面の穴の位置や数に違いがあるが，全体の形状は半球形状で類似している．ヘッドの CAD モデルはいずれのサイズも球形状で近似している．図 3.3 にインプラントの CAD モデルのバリエーションを示す．ただしステムは全て右大腿骨に挿入する形状を表示している．

3.3.2. U-net

インプラントセグメンテーションには，U-net[18]（プーリング層は 4 層使用）を用いる．画像の切り出しは行わず，横断面全体（画像サイズ：512×512）を入力とする．入力する CT 画像は $[0, 4000]$ HU の輝度値を $[0, 1]$ に線形に正規化し，水平方向の反転，平行移動（ $-25\% \sim 25\%$ ），せん断（ $-\pi/8 \sim \pi/8$ rad），回転（ $-10 \sim 10$ 度）の data augmentation を加えた．損失関数は Softmax cross entropy とし，バッチサイズを 2，学習係数を 0.0001 として Adam[29] による最適化を行い，初期化には [28] の手法を用いた．また，テスト時には U-net で得られたカップ，ステム，ヘッドそれぞれについて，最大連結成分のみを抽出し最終出力とした．

3.3.3. 実験概要

3.2.1 節の手法で作成した学習データを用いて U-net[18] の学習を行う．また，特に形状にバリエーションがあるステムについて着目し，提案法を実験 1 と実験 2 の 2 つの実験によって定量的に評価した．

実験 1 ではステムに Centpillar を使用した症例のみで学習を行い，Centpillar を用いた症例でテストを行う．実験 1 では，Centpillar を使用した 100 症例を使用し，4 分割交差検証によって精度評価を行なう．

実験 2 では Centpillar を使用した症例のみで学習を行い, Accolade を用いた症例でテストを行う. Centpillar を使用した 100 症例すべてを用いて学習を行い, Accolade が挿入されている 11 症例でテストを行う. 本実験での目的は, 学習データに含まれない未知の機種に対する提案法の適用可能性についての検討である.

3.3.4. 評価指標

3.2.1 節の手法で得られた正解データを用いて, Dice 係数 (式 (2.4.1)) 及び ASD (式 (2.4.2)) によって評価を行う. 各評価指標は 3 次元での計測を行う. さらに, ステムについて, U-net によって得られた形状と CAD モデルを用いて位置合わせを再度行い, 位置合わせ後の表面間距離誤差 (式 (3.3.1)) で評価する. この指標によって, セグメンテーションされたインプラント形状と CAD モデルの形状がどれだけ近いかを評価する.

$$\textit{Registration Error} = \textit{AAD}(p_{\text{stem}}(\Theta_{\text{stem}}), p_{\text{threshold}}) \quad (3.3.1)$$

ここで, Θ_{stem} は最適化で得られたステムの姿勢パラメータである.

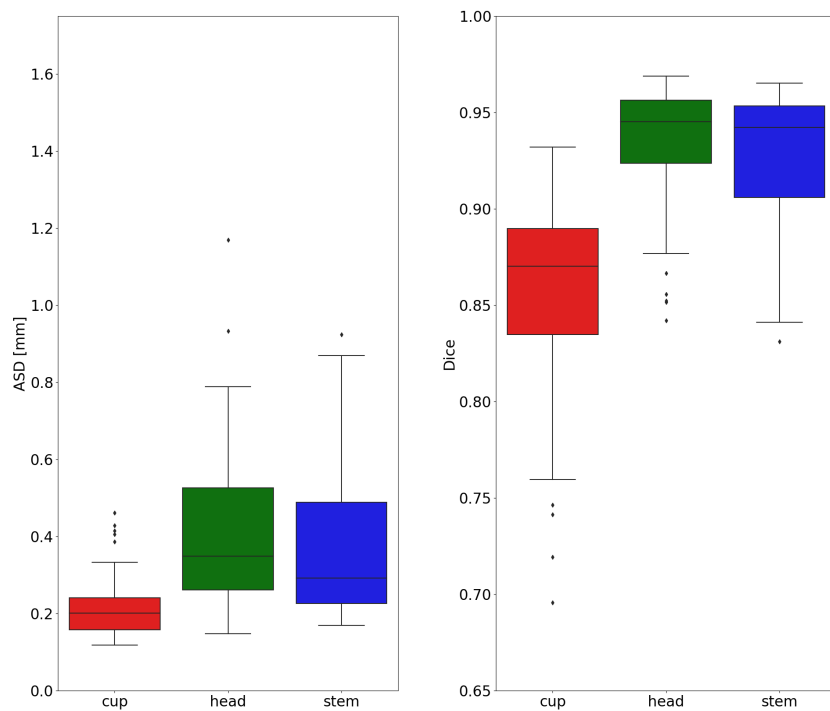


図 3.4 ステムに Centpillar を使用した症例のセグメンテーション精度

3.4. 実験結果

3.4.1. 実験 1 (学習データ : Centpillar, テストデータ : Centpillar)

ステムに Centpillar を用した症例の、カップ・ヘッド・ステムのセグメンテーション精度を図 3.4 に示す。また、ステムの Dice 係数が最高値、最低値であった症例のセグメンテーション結果をそれぞれ図 3.5, 図 3.7 に示す。実験 1 では症例数が偶数 (100 症例) であるため、ステムの Dice 係数が 51 番目に高い症例を中央値の症例としている (図 3.6)。

さらに、提案法で得られたセグメンテーション結果に対して CAD モデルを位置合わせした際の表面間距離誤差を図 3.8 左に示す。比較として、閾値で得られたインプラント形状に CAD モデルを位置合わせした結果を示す。ただし、図 3.8 ではステムの CAD モデルのみからセグメンテーション結果への表面間距離誤差のみを示している。

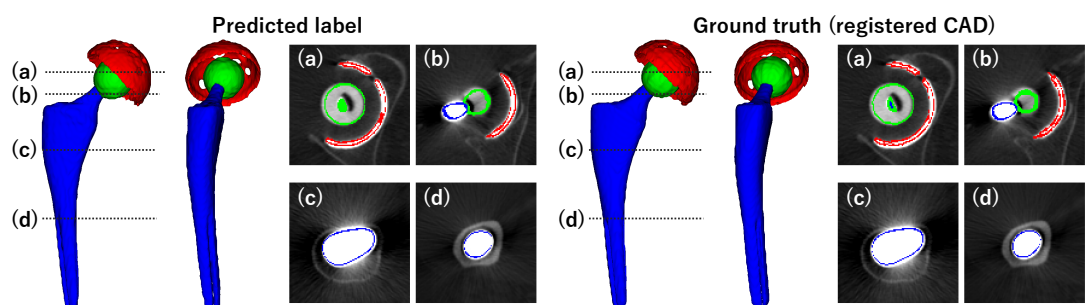


図 3.5 ステム（Centpillar）の Dice 係数が最高値であった症例（カップ: 0.896, ヘッド: 0.962, ステム: 0.965). [WW,WL]=[1750, 4500]HU

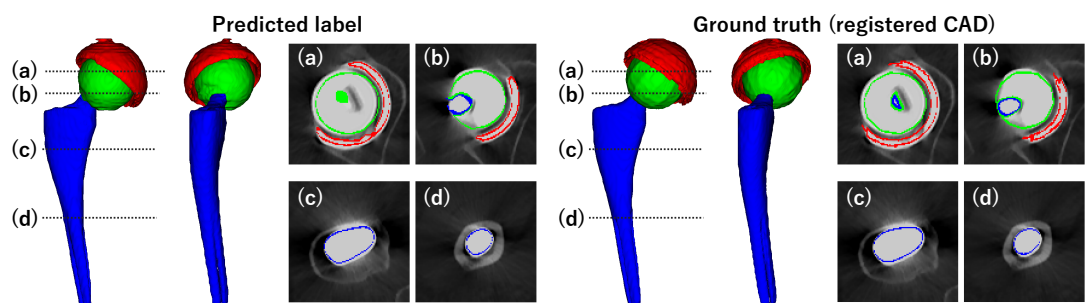


図 3.6 ステム（Centpillar）の Dice 係数が中央値であった症例（カップ: 0.863, ヘッド: 0.945, ステム: 0.942) [WW,WL]=[1750, 4500]HU

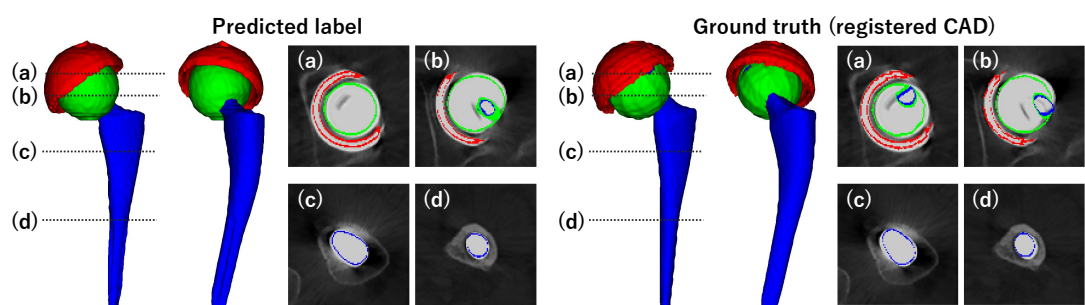


図 3.7 ステム（Centpillar）の Dice 係数が最低値であった症例（カップ: 0.859, ヘッド: 0.898, ステム: 0.831) [WW,WL]=[1750, 4500]HU

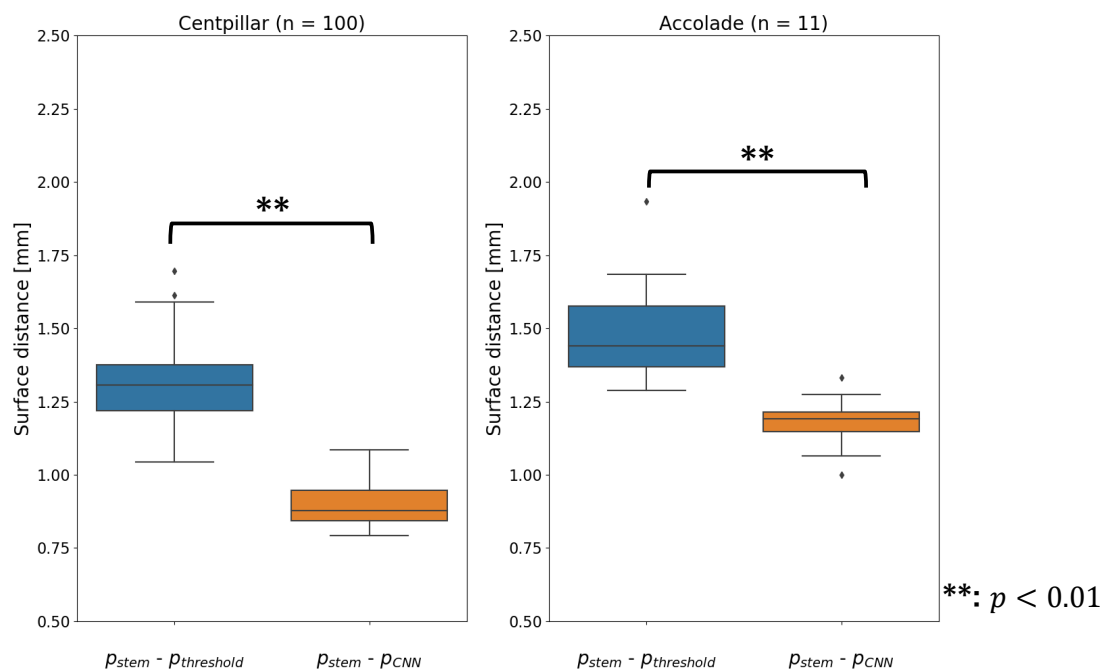


図 3.8 位置合わせされたステムの CAD モデルからの表面間距離誤差．左は Centpillar，右は Accolade の症例での結果を示す．青：閾値 (2000 HU) セグメンテーション結果との位置合わせ誤差 橙：提案法によるセグメンテーション結果との位置合わせ誤差

3.4.2. 実験 2 (学習データ : Centpillar, テストデータ : Accolade)

ステム機種が Accolade の症例のカップ・ヘッド・ステムのセグメンテーション精度を図 3.10 に示す．また，ステムの Dice 係数が最高値，中央値，最低値であった症例のセグメンテーション結果をそれぞれ図 3.11，図 3.12，図 3.13 に示す．さらに，得られたセグメンテーション結果に対して CAD モデルを位置合わせした際の表面間距離誤差を図 3.8 右に示す．

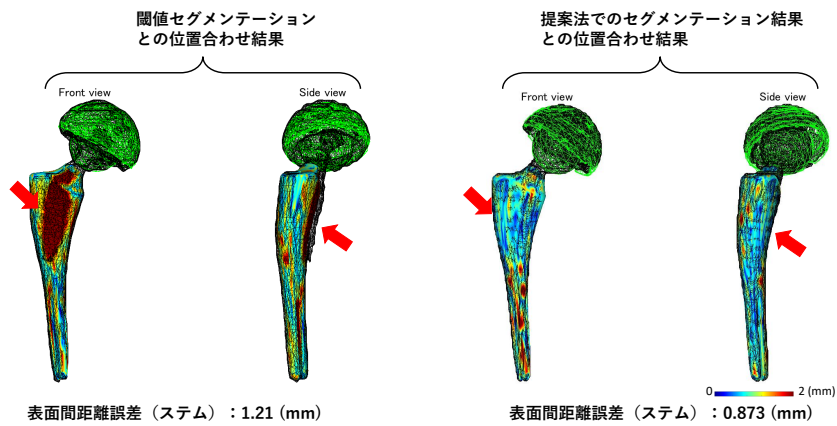


図 3.9 セグメンテーションされたインプラントと CAD モデルとの位置合わせ結果の比較. 左: 閾値セグメンテーションとの位置合わせ結果 右: 提案法でのセグメンテーション結果との位置合わせ結果.

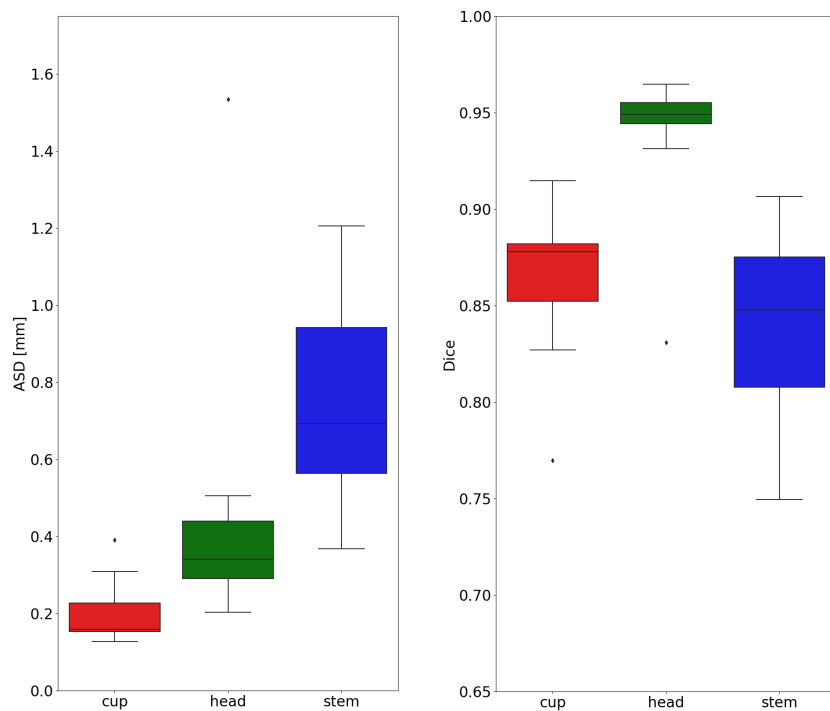


図 3.10 ステム機種が Accolade である症例のセグメンテーション精度. 左: Dice 係数, 右: ASD.

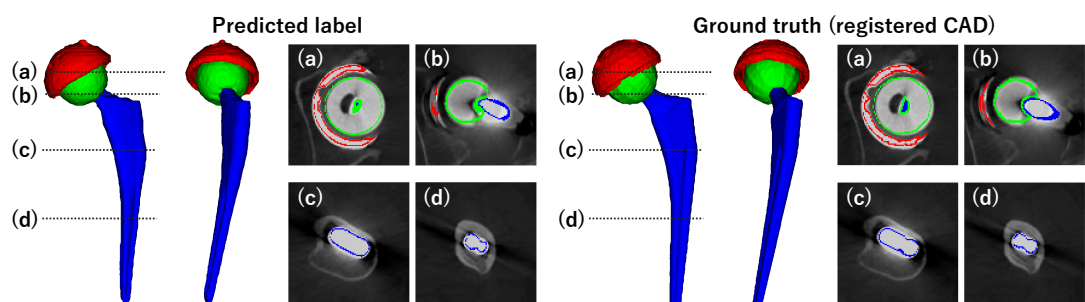


図 3.11 ステム（Accolade）の Dice 係数が最高値であった症例（カップ: 0.915, ヘッド: 0.965, ステム: 0.907) [WW,WL]=[1750, 4500]HU

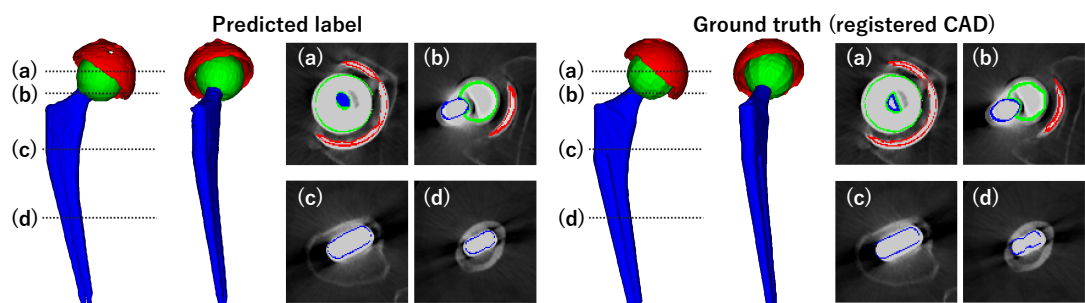


図 3.12 ステム（Accolade）の Dice 係数が中央値であった症例（カップ: 0.850, ヘッド: 0.960, ステム: 0.848) [WW,WL]=[1750, 4500]HU

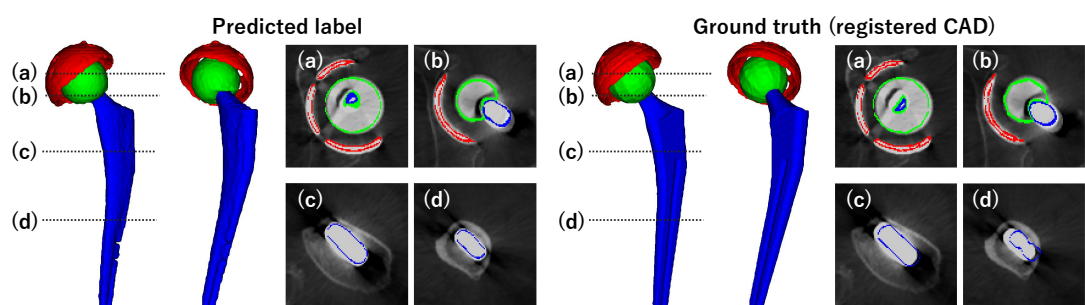


図 3.13 ステム（Accolade）の Dice 係数が最低値であった症例（カップ: 0.901, ヘッド: 0.949, ステム: 0.750) [WW,WL]=[1750, 4500]HU

3.5. 考察

3.5.1. セグメンテーション精度

実験 1 において，ステムのセグメンテーションでは Dice 係数で 0.930 ± 0.0290 mm, ASD で 0.359 ± 0.169 mm と高い精度が得られた．また CAD モデルとの位置合わせ（図 3.8）においても，閾値を用いた場合よりも誤差が統計的に有意に低下した．このことから，U-net は学習データに含まれる機種（Centpillar）をセグメンテーションできていると考えられる（図 3.9）．

一方，Accolade を用いた実験 2 では Centpillar でのテスト精度と比較して精度は低下し，Dice 係数で 0.839 ± 0.0479 mm, ASD で 0.749 ± 0.250 mm であった．Centpillar と Accolade の形状を横断面で比較すると，どちらも楕円に近い形状であるが，特に短軸方向の膨らみに違いが見られる（図 3.5，図 3.11）．Centpillar で学習した U-net は Accolade の症例においても，短軸方向にやや膨らみをもつような形状を認識しており（図 3.11），形状の違いがセグメンテーション精度に影響を及ぼしていると考えられる．ただし，セグメンテーションされたラベルと CAD モデルとの位置合わせ（図 3.8）では Centpillar と同様に，閾値を用いた場合よりも誤差が統計的に有意に低下しており，U-net は学習データに含まれない機種であっても従来法と比較すると正確な形状を認識することが可能であると考えられる．

3.5.2. 今後の課題及び展望

本研究の限界点として，まず，データセットのバリエーションが不十分である点が挙げられる．学習データに含まれない機種への汎用性についてのより詳細な評価を行うためには，データセット内のインプラント形状のバリエーションをさらに増強する必要がある．特にステムには本研究で対象とした single wedge 型と anatomical 型以外に 4 種類の形状が存在し [44]，それらの形状での評価が課題である．日本国内では多岐に渡る形状・機種のステムが使用されており [50]，様々な形状での評価は重要であると考えられる．定性評価にはなるが，CAD モデルの情報が紐づいていない大規模データベースに適用し，提案法の有効性を検証することも課題として挙げられる．

次に、正解データの正確性が課題である。3.2.1 節で述べた通り、本研究ではインプラントのサイズを最適化するパラメータに含めず事前に得られた値を使用しているが、サイズは術中に医師の判断によって変更される場合がある [41]。術中のサイズ変更を考慮し、姿勢パラメータだけでなくサイズを含めた最適化を行う事でより正確な学習データセットの構築を図る。また、提案法の結果を使用すると CAD モデルとの位置合わせ精度が向上した結果を踏まえ、その位置合わせで得られた結果を正解データとして再度使用し、データセットを増強する方法も考えられる。

今後の展望としては、以下の 2 点が挙げられる。

1. セグメンテーション精度の向上
2. 金属アーチファクト低減への応用

精度向上について、本研究では CT 画像の横断面を用いた 2 次元でのセグメンテーションを行ったが、3 次元でのセグメンテーション手法 [42] に拡張することで精度向上が期待できる。しかしながら 3 次元でのセグメンテーションでは、2 次元での学習に比べて学習データに含まれる機種種の 3 次元形状、未知の機種種に対する精度が低下する事が考えられる。

次に、提案法によって得られた 3 次元形状は金属アーチファクト低減に応用可能である。金属アーチファクト低減手法 (2.2.1 節) に示す中でも、サイノグラムの補正に分類される手法では、サイノグラム上での金属部分のセグメンテーション精度が非常に重要である [25, 43]。[43] ではサイノグラム上で金属の手動セグメンテーションを行い、金属のセグメンテーション精度が金属アーチファクト低減に及ぼす影響を検証している。従来法の多くは式 (2.3.1) に示すような閾値でのセグメンテーションを行っているが、提案法によって得られたインプラントの 3 次元形状を金属アーチファクト低減手法に用いる事で、金属アーチファクト低減精度を向上させることが可能であると考えられる。

3.6. 小括

本章では、人工股関節全置換術の術後 CT 画像から、カップ・ヘッド・ステムの 3 部位の 3 次元形状を自動でセグメンテーションする手法を提案し、その有効性

を定量的に評価した．提案法によって，推論時には CAD モデルなしに，カップ・ヘッド・ステムの 3 次元形状を取得する事が可能である．実験では学習データに含まれない機種を使用し，未知の機種に対する提案法の適用可能性について初期的検討を行った．現時点での限界点としてデータセット内のバリエーションの少なさや正解データの正確性について述べ，また，精度向上および金属アーチファクト低減への応用の観点から今後の展望を示した．

第 4 章 総括

本研究では，人工股関節全置換術の術後 CT 画像セグメンテーションとして (1) 筋骨格セグメンテーション (2) インプラントセグメンテーション の 2 つを提案した．

第 2 章では金属アーチファクト低減と筋骨格セグメンテーションを組み合わせた手法を提案し，シミュレーション実験と実画像実験から提案法の有効性を定量的に評価した．さらに，シミュレーション実験と実画像実験の両方でアーチファクト低減によって不確実性が低下することを示し，今後の展望としてアーチファクト低減の定量評価への筋骨格セグメンテーションの不確実性の応用について述べた．

第 3 章では，CAD モデルを正解データとして学習した U-net による，金属アーチファクトを含む術後 CT 画像からのインプラントセグメンテーションについて述べた．実験では特に，学習データに含まれないインプラント機種を使用し，未知のインプラントへの提案法の汎用性について初期的検討を行った．展望として，提案法で得られた 3 次元形状を用いた金属アーチファクト低減について述べた．

このように 2 章・3 章で述べた提案法によって，術後 CT 画像から，個々の筋骨格・個々のインプラント（カップ・ヘッド・ステム）の 3 次元形状の認識が可能となった．また，得られた結果を用いることで臨床的意義のある解剖学的パラメータの解析が可能になる．例えば著者らの研究グループではこれまでに，提案法を用いて術前計画でのインプラント設置位置と術後 CT 画像から計測されたインプラント設置位置の違いを検証し，報告している [45]．

今後はさらに，同一患者の術前術後で対応したデータセットを使用して解析を進め，過去症例の術後 CT 画像の解析結果を活用した手術計画の自動立案が期待される．従来，人工股関節全置換術の手術計画自動立案 [3] では，過去症例の術前情報とそれに基づいて医師が立案した手術計画のみを学習データとしていた．これに対し，提案法を基に得られた術後情報を学習データに加える事で，過去症例の術後評価も考慮した手術計画の立案が可能であると考ええる．

謝辞

本研究を進めるにあたり，日頃より有益な助言と丁寧な指導を賜りました奈良先端科学技術大学院大学 先端科学技術研究科 情報科学領域 生体医用画像研究室の佐藤嘉伸教授に深く御礼申し上げます。また，学内発表の際には貴重なご助言・ご意見を頂きました，情報科学領域 光メディアインタフェース研究室 向川康博 教授，バイオサイエンス領域 遺伝子発現制御研究室 別所康全 教授に深く感謝申し上げます。そして日頃からの密な打ち合わせや議論の中で多くの有益なご助言・丁寧なご指導をいただきました生体医用画像研究室 大竹義人 准教授，Mazen Soufi 助教，日朝祐太 助教，上村圭亮 助教に深く感謝申し上げます。

また，大変貴重なデータのご提供ならび手動トレース作成へのご協力，さらに医師の観点から大変貴重なご意見をいただきました，菅野伸彦 教授，高尾正樹 講師，田村和則 先生，佐藤龍一 先生をはじめとする大阪大学大学院 医学研究科 器官制御外科学の皆様に深く感謝申し上げます。

大学院生活では，入学当初より丁寧な指導を頂きました，鈴木裕紀 氏（大阪大学大学院 医学系研究科 人工知能画像診断学共同研究講座 特任助教），福田紀生 氏（脳情報通信融合研究センター 研究員），同じチームで研究に取り組み，また資料作成などでは多大な協力をして頂いた槇野大樹 君，そして出張の手配や日頃の心配事での相談など大学院生活全般を支えて頂きました田川郁代 秘書をはじめとする生体医用画像研究室の皆様に感謝申し上げます。

最後に，修士課程修了まで多大な金銭的・精神的支援を頂いた家族に深く，厚く感謝し，ここに謝辞とさせていただきます。

参考文献

- [1] 川原 康弘. 股関節・骨盤の画像診断. p22 メディカル・サイエンス・インターナショナル, 2017
- [2] 大谷 卓也, 飯田 寛和. 人工股関節置換術 [THA・BHA] 人工関節の基礎知識, 基本テクニック, 合併症対策. p38 メディカ出版, 2016
- [3] I. Otomaru, M. Nakamoto, Y. Kagiya, M. Takao, N. Sugano, N. Tomiyama, Y. Tada, and Y. Sato, “Automated preoperative planning of femoral stem in total hip arthroplasty from 3D CT data: Atlas-based approach and comparative study,” *Medical Image Analysis*, vol. 16, no. 2, pp. 415–426, 2012.
- [4] F. Yokota, Y. Otake, M. Takao, T. Ogawa, T. Okada, N. Sugano, and Y. Sato, “Automated muscle segmentation from CT images of the hip and thigh using a hierarchical multi-atlas method,” *International Journal of Computer Assisted Radiology and Surgery*, vol. 13, no. 7, pp. 977–986, 2018.
- [5] Y. Hiasa, Y. Otake, M. Takao, T. Ogawa, N. Sugano, and Y. Sato, “Automated Muscle Segmentation from Clinical CT using Bayesian U-Net for Personalized Musculoskeletal Modeling,” *IEEE Transactions on Medical Imaging*, no. I, pp. 1–1, 2019.
- [6] H. Lee, F. M. Troschel, S. Tajmir, G. Fuchs, J. Mario, F. J. Fintelmann, and S. Do, “Pixel-Level Deep Segmentation: Artificial Intelligence Quantifies Muscle on Computed Tomography for Body Morphometric Analysis,” *Journal of Digital Imaging*, vol. 30, no. 4, pp. 487–498, 2017.
- [7] R. Hemke, C. G. Buckless, A. Tsao, B. Wang, and M. Torriani, “Deep learning for automated segmentation of pelvic muscles, fat, and bone from CT studies for body composition assessment,” *Skeletal Radiology*, 2019.
- [8] A. D. Weston, P. Korfiatis, T. L. Kline, K. A. Philbrick, P. Kostandy, T. Sakinis, M. Sugimoto, N. Takahashi, and B. J. Erickson, “Automated Abdominal Segmentation of CT Scans for Body Composition Analysis Us-

- ing Deep Learning,” *Radiology*, vol. 290, no. 3, pp. 669–679, 2019.
- [9] D. Iwana, N. Nakamura, H. Miki, M. Kitada, T. Hananouchi, and N. Sugano, “Accuracy of angle and position of the cup using computed tomography-based navigation systems in total hip arthroplasty,” *Computer Aided Surgery*, vol. 18, no. 5-6, pp. 187–194, 2013.
 - [10] T. Shoji, T. Yamasaki, S. Izumi, H. Murakami, K. Mifuji, M. Sawa, Y. Yasunaga, N. Adachi, and M. Ochi, “Factors affecting the potential for posterior bony impingement after total hip arthroplasty,” *Bone and Joint Journal*, vol. 99B, no. 9, pp. 1140–1146, 2017.
 - [11] K. Uemura, M. Takao, T. Sakai, T. Nishii, and N. Sugano, “Volume Increases of the Gluteus Maximus, Gluteus Medius, and Thigh Muscles After Hip Arthroplasty,” *Journal of Arthroplasty*, vol. 31, pp. 906–912.e1, apr 2016.
 - [12] T. Ogawa, M. Takao, Y. Otake, F. Yokota, H. Hamada, T. Sakai, Y. Sato, and N. Sugano, “Validation study of the CT-based cross-sectional evaluation of muscular atrophy and fatty degeneration around the pelvis and the femur,” *Journal of Orthopaedic Science*, no. xxxx, 2019.
 - [13] M. Sakamoto, Y. Hiasa, Y. Otake, M. Takao, Y. Suzuki, N. Sugano, and Y. Sato, “Bayesian segmentation of hip and thigh muscles in metal artifact-contaminated ct using convolutional neural network-enhanced normalized metal artifact reduction,” *Journal of Signal Processing Systems*, pp. 1–10, 2020.
 - [14] L. Gjestebj, B. De Man, Y. Jin, H. Paganetti, J. Verburg, D. Giantsoudi, and G. Wang, “Metal Artifact Reduction in CT: Where Are We After Four Decades?,” *IEEE Access*, vol. 4, pp. 5826–5849, 2016.
 - [15] E. Meyer, R. Raupach, M. Lell, B. Schmidt, and M. Kachelrieß, “Normalized metal artifact reduction (NMAR) in computed tomography,” *Medical Physics*, vol. 37, no. 10, pp. 5482–5493, 2010.
 - [16] L. Gjestebj, H. Shan, Q. Yang, Y. Xi, B. Claus, Y. Jin, B. De Man, and G. Wang, “Deep Neural Network for CT Metal Artifact Reduction with

- a Perceptual Loss Function,” in *Proceedings of The Fifth International Conference on Image Formation in X-ray Computed Tomography*, vol. 1, pp. 439–443, 2018.
- [17] M. B. Ranzini, J. Henckel, M. Ebner, M. J. Cardoso, A. Isaac, T. Vercauteren, S. Ourselin, A. Hart, and M. Modat, “Automated postoperative muscle assessment of hip arthroplasty patients using multimodal imaging joint segmentation,” *Computer Methods and Programs in Biomedicine*, vol. 183, 2020.
 - [18] O. Ronneberger, P. Fischer, and T. Brox, “U-net: Convolutional networks for biomedical image segmentation,” *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 9351, pp. 234–241, 2015.
 - [19] J. Wang, S. Wang, Y. Chen, J. Wu, J. L. Coatrieux, and L. Luo, “Metal artifact reduction in CT using fusion based prior image,” *Medical Physics*, vol. 40, no. 8, pp. 1–8, 2013.
 - [20] W. A. Kalender, R. Hebel, and J. Ebersberger, “Reduction of CT artifacts caused by metallic implants,” *Radiology*, vol. 164, no. 2, pp. 576–577, 1987.
 - [21] E. Meyer, R. Raupach, B. Schmidt, A. H. Mahnken, and M. Kachelriess, “Adaptive normalized metal artifact reduction (ANMAR) in computed tomography,” *IEEE Nuclear Science Symposium Conference Record*, pp. 2560–2565, 2012.
 - [22] P. Isola, J. Y. Zhu, T. Zhou, and A. A. Efros, “Image-to-image translation with conditional adversarial networks,” *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, vol. 2017-January, pp. 5967–5976, 2017.
 - [23] Y. Gal and Z. Ghahramani, “Dropout as a Bayesian approximation: Representing model uncertainty in deep learning,” *33rd International Conference on Machine Learning, ICML 2016*, vol. 3, pp. 1651–1660, 2016.
 - [24] A. Kendall and Y. Gal, “What uncertainties do we need in Bayesian deep learning for computer vision?,” *Advances in Neural Information Processing*

- Systems*, vol. 2017-December, no. Nips, pp. 5575–5585, 2017.
- [25] Y. Zhang and H. Yu, “Convolutional Neural Network Based Metal Artifact Reduction in X-Ray Computed Tomography,” *IEEE Transactions on Medical Imaging*, vol. 37, no. 6, pp. 1370–1381, 2018.
 - [26] G. T. Herman, “Correction for beam hardening in computed tomography,” *Physics in Medicine and Biology*, vol. 24, pp. 81–106, jan 1979.
 - [27] Y. Kyriakou, E. Meyer, D. Prell, and M. Kachelrieß, “Empirical beam hardening correction (EBHC) for CT,” *Medical Physics*, vol. 37, no. 10, pp. 5179–5187, 2010.
 - [28] K. He, X. Zhang, S. Ren, and J. Sun, “Delving deep into rectifiers: Surpassing human-level performance on imagenet classification,” in *Proceedings of the IEEE International Conference on Computer Vision*, vol. 2015 Inter, pp. 1026–1034, 2015.
 - [29] D. P. Kingma and J. L. Ba, “Adam: A method for stochastic optimization,” in *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, pp. 1–15, 2015.
 - [30] S. Ioffe and C. Szegedy, “Batch normalization: Accelerating deep network training by reducing internal covariate shift,” *CoRR*, vol. abs/1502.03167, 2015.
 - [31] Z. Zhong, L. Zheng, G. Kang, S. Li, and Y. Yang, “Random Erasing Data Augmentation,” 2017.
 - [32] L. R. Dice, “Measures of the amount of ecologic association between species,” *Ecology*, vol. 26, no. 3, pp. 297–302, 1945.
 - [33] M. Styner, J. Lee, B. Chin, M. Chin, O. Commowick, H. Tran, S. Markovic-Plese, V. Jewells, and S. Warfield, “3d segmentation in the clinic: A grand challenge ii: Ms lesion segmentation,” *Midas Journal*, vol. 2008, pp. 1–6, 2008.
 - [34] J. Wang, J. H. Noble, and B. M. Dawant, “Metal artifact reduction for the segmentation of the intra cochlear anatomy in CT images of the ear with 3D-conditional GANs,” *Medical Image Analysis*, vol. 58, 2019.

- [35] D. Wu, K. S. Kim, B. Dong, and Q. Li, “End-to-end abnormality detection in medical imaging,” *CoRR*, vol. abs/1711.02074, 2017.
- [36] A. Shrivastava, T. Pfister, O. Tuzel, J. Susskind, W. Wang, and R. Webb, “Learning from simulated and unsupervised images through adversarial training,” in *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, vol. 2017-Janua, pp. 2242–2251, 2017.
- [37] Z. Chang, D. H. Ye, S. Srivastava, J. B. Thibault, K. Sauer, and C. Bouman, “Prior-guided metal artifact reduction for iterative X-ray computed tomography,” *IEEE Transactions on Medical Imaging*, vol. 38, pp. 1532–1542, jun 2019.
- [38] H. Miki, N. Sugano, K. Yonenobu, K. Tsuda, M. Hattori, and N. Suzuki, “Detecting cause of dislocation after total hip arthroplasty by patient-specific four-dimensional motion analysis,” *Clinical Biomechanics*, vol. 28, no. 2, pp. 182–186, 2013.
- [39] M. A. Hegazy, M. H. Cho, M. H. Cho, and S. Y. Lee, “U-net based metal segmentation on projection domain for metal artifact reduction in dental CT,” *Biomedical Engineering Letters*, vol. 9, no. 3, pp. 375–385, 2019.
- [40] N. Hansen, “The CMA Evolution Strategy: A Comparing Review,” *Towards a New Evolutionary Computation*, vol. 102, no. 2006, pp. 75–102, 2007.
- [41] T. Ogawa, M. Takao, T. Sakai, and N. Sugano, “Factors related to disagreement in implant size between preoperative CT-based planning and the actual implants used intraoperatively for total hip arthroplasty,” *International Journal of Computer Assisted Radiology and Surgery*, vol. 13, no. 4, pp. 551–562, 2018.
- [42] Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger, “3D U-net: Learning dense volumetric segmentation from sparse annotation,” *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 9901 LNCS, pp. 424–432, 2016.

- [43] M. Stille, B. Kratz, J. Müller, N. Maass, I. Schasiepen, M. Elter, I. Weyers, and T. M. Buzug, “Influence of metal segmentation on the quality of metal artifact reduction methods,” in *Medical Imaging 2013: Physics of Medical Imaging*, vol. 8668, p. 86683C, 2013.
- [44] H. S. Khanuja, J. J. Vakil, M. S. Goddard, and M. A. Mont, “Cementless femoral fixation in total hip arthroplasty,” *Journal of Bone and Joint Surgery - Series A*, vol. 93, no. 5, pp. 500–509, 2011.
- [45] 榎野 大樹, 阪本 充輝, 大竹 義人, 高尾 正樹, 日朝 祐太, 中西 裕紀, 鍵山 善之, 菅野 伸彦, 佐藤 嘉伸. 人工股関節全置換術術後 CT の定量解析によるカップ・ステムの設置統計モデルの構築, 第 28 回 日本コンピュータ外科学会, 2019
- [46] 阪本 充輝, 日朝 祐太, 大竹 義人, 高尾 正樹, 菅野 伸彦, 佐藤 嘉伸. GAN を用いた人工股関節全置換術術後 CT 画像の金属アーチファクト低減, 第 38 回 日本医用画像工学会, 2019
- [47] 松下 任彦, 橋本 伸朗, 福元 哲也, 前田 智, 中馬 東彦, 平井 奉博, 浦田 泰弘. 人工股関節手術における内閉鎖筋の温存と CT 画像評価. 整形外科と災害外科, 65(1), pp. 5-9, 2016.
- [48] 坂井 孝司, 関節外科における材料のニーズの現在と未来, まてりあ, 53 巻, 4 号, p.139-142, 2014
- [49] CMA-ES Source Code. http://cma.gforge.inria.fr/cmaes_sourcecode_page.html (2020/01/27 アクセス)
- [50] 日本人工関節学会 (2018), THA 人工関節登録調査集計 2006 年 2 月～2018 年 3 月, <https://jsra.info/jar-report.html> (2020/03/11 アクセス)