

Cone-beam CT 画像補完のための擬似劣化画像を用いた 自己教師あり一貫性モデル

中嶋 惇[†] 御手洗 彰[†] 中尾 恵[†]

[†] 京都大学大学院医学研究科 人間健康科学系専攻 〒606-8507 京都市左京区聖護院川原町 53

E-mail: †nakajima.jun.32x@st.kyoto-u.ac.jp

あらまし 本研究では, Cone-beam CT (CBCT) 画像の多様なアーチファクト補完を目的とし, CBCT の擬似劣化画像を用いた自己教師あり一貫性モデルを提案する. CT 画像から異なるシミュレーション条件の擬似劣化画像を作成し, 一貫性モデルの学習に適用した. 実験では, 頭頸部, 胸部, 腹部の CT-CBCT 画像ペアを用いて補完性能を比較し, CBCT 画像よりも擬似劣化画像の学習が, 高い補完性能を実現することを確認したので報告する.

キーワード 潜在一貫性モデル, CBCT 画像補完, 画像生成

1. はじめに

Cone-beam Computed Tomography (CBCT) は円錐状に X 線を照射する小型かつ可動式の 3 次元計測装置である. この特性から, 放射線治療や外科手術の際に, 治療時の体位で 3 次元画像を取得できるという臨床的利点がある. 一方で, X 線の照射方法や画像再構成法が CT と異なり, CBCT 画像は CT 画像と比べ画質が劣り, CT 値の精度が低下しやすい. 従来は画像処理や統計的な性質を活用した最適化などにより改善が図られてきたが, 複数の物理的要因から生じる複雑なアーチファクトの改善には限界がある. この課題に対し, 機械学習を用いて CBCT 画像を補完することによって, CT 画像に近い品質の画像を得ることが期待されている.

機械学習に基づく医用画像の画質改善は多く研究されており [1], 近年の研究では, 条件付き拡散モデル [2] が広く利用されている. 逆拡散過程において, 補完対象の画像を条件として与えることで画像変換を行うモデルで, 頭頸部 CBCT 画像の補完 [3] では, 高精度な合成 CT 画像を生成した. しかし, 解剖学的構造保持や, 画像生成時間の長さに課題があり, 放射線治療時などのリアルタイムな利用には制約がある.

本研究では, 潜在一貫性モデルの自己教師あり学習による CBCT 画像の画像補完を目的とする. 潜在一貫性モデル [4] を用いることで学習時の記憶容量と計算量を低減し, 高速な生成を可能とする. 近年, 自己教師あり学習による CBCT 画像の画像補完 [5] が報告されており, 人工的なアーチファクトのシミュレートによる擬似劣化画像の作成により, 構造を維持した補完が達成されている. 一方, 上記研究の擬似劣化画像の作成方法は経験的な設定に基づき, その指針は明らかでない. この課題に対し, 本研究では擬似劣化画像の作成に撮像時の物

理特性を考慮した劣化関数を導入し, 作成した擬似劣化画像を擬似 CBCT 画像として用いる. さらに, 頭頸部, 胸部, 腹部の 3 領域に対応できる汎用性を備えたモデルを目指し, 異なる劣化条件で擬似 CBCT 画像を作成し, 画像補完タスクの自己教師あり学習を行う. 位置合わせ済みの CT-CBCT ペア画像を用いて, 各条件の違いが CBCT 画像の補完に与える影響を調査する.

2. 提案手法

本研究では, 診断や治療時において CBCT 画像が得られる状況を想定する. 提案モデルは, 得られた CBCT 画像を拡散モデルのガイドとし, 出力としてモデルの予測結果である生成画像 (synthetic CT, sCT) を得る枠組みとして設計する.

2.1 撮像時の物理特性を考慮した擬似 CBCT 画像

実験に用いる擬似 CBCT 画像の作成方法を説明する. CBCT 画像は CT 画像と比較して, 画質劣化やアーチファクトの影響を受けやすいことが知られており, そのような劣化は様々な要因によって引き起こされる [6]. 例えば, 物理的要因から発生するものとして Scatter や Beam hardening, 幾何学・装置的要因のものに Ring artifact, Motion artifact, Saturation, 再構成におけるアルゴリズム上の要因として Aliasing などが知られている.

本研究では, CBCT 画像に生じ得るアーチファクトのうち, 再構成画像の画質に対する影響が大きい 6 つの要素をシミュレートした劣化関数を導入する. 提案する劣化関数は式 (1) のように定義される.

$$I_d = D(I; \theta, m, b, s_c, p, s_a), \quad (1)$$

ここで, I は 3D-CT 画像であり, D は投影計算と逆投影計算

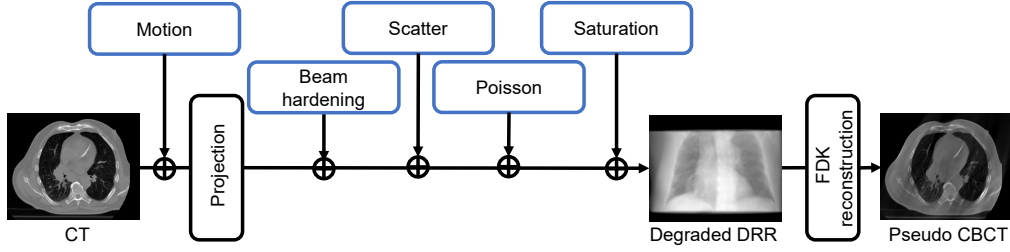


図 1 投影計算における劣化関数の導入による擬似 CBCT 画像作成の流れ

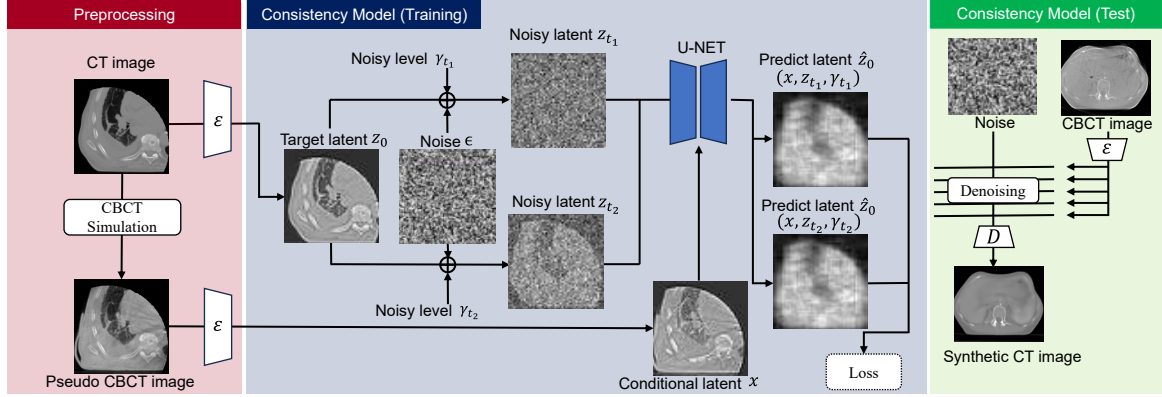


図 2 条件付き潜在一貫性モデルの構成とデータ処理の流れ

を含む劣化関数, I_d はシミュレート対象の擬似 CBCT 画像である. 劣化関数 D は次のパラメータを属性値として持ち, これら値の調整によって, 3D-CT 画像に対して多様かつ異なる強度のアーチファクトを合成する.

- Geometry parameter θ : 撮像時の投影方向
- Motion m : 被写体の動きや臓器の運動
- Beam hardening b : スペクトル依存の吸収
- Scatter s_c : 散乱線の影響による平滑化
- Poisson p : 検出される光子数の統計的なゆらぎ
- Saturation s_a : 検出素子における飽和

擬似 CBCT 画像の作成の流れについて図 1 に示す. まず, デジタル再構成 X 線画像 (Digitally Reconstructed Radiograph, DRR) を生成する際の投影方向 θ をランダムに決定する. この時, 撮像対象の動きや臓器の運動をシミュレートするため, 3D-CT 画像に対し, パラメータ m に基づいて, 画素空間内でボクセルを微小に摂動させる. 次に, 投影計算の過程において, 劣化関数の各パラメータ b, s_c, p, s_a に基づき, 画素値の非線形変換, 平滑化, ショットノイズの付与, 輝度値のクリップ等の演算を行い, 劣化 DRR 画像を得る. 最後に, 生成された劣化 DRR 画像を Feldkamp-Davis-Kress (FDK) 再構成によって画像空間に逆投影を行い, 擬似 CBCT 画像を得る.

2.2 潜在一貫性モデル

擬似 CBCT 画像をガイドとした画像生成に用いる条件付き潜在一貫性モデルを説明する. 図 2 に条件付き潜在一貫性モデルの処理の流れを示す. CT 画像及び作成した擬似 CBCT 画像から学習済み VQVAE [7] を用いて潜在表現 z_0 及び x を得る. 学習時には, z_0 に対してランダムな 2 つの時刻 $t = t_n, t_m$ ($1 \leq t_n, t_m \leq T-1$, ($t_n \neq t_m$)) に対応するノイズを付与する.

ここで, T は画像が完全にノイズになる時刻のことであり, 本研究では先行研究 [5] を参考に $T = 1000$ を用いた. 異なる時刻のノイズが付与された潜在表現に対し, 条件 x の下で z_0 を直接予測するようにモデルをトレーニングする. この時, 異なるノイズ時刻から一貫した潜在表現を復元するように損失を導入することで, 少ない推論ステップでの安定した画像生成を可能とする. 推論時には最初に CBCT 画像から学習済み VQVAE によって潜在表現を取得する. 次に, 内部で生成したガウシアンノイズに対して, CBCT 画像の潜在表現をガイドとして 1 ステップでノイズ除去を行う. 最後に, 得られた潜在表現を学習済み VQVAE によってデコードし, 最終的な生成画像 sCT を生成する. 以上の推論処理はスライス単位でなされるが, 3 次元的に連続的な画像補完となるようすべてのスライスにおいて同一の初期ノイズを用いる.

3. 実験

提案モデルの性能評価を目的として, 本研究では Syn-thRAD2025 Grand Challenge [8] において配布された CT-CBCT ペア画像群を用いて実験を行った.

3.1 対象データ

実験で用いたデータベースは, 頭頸部, 胸部, 腹部の 3 つの部位を含んだ計 774 症例から構成される. 実験では, 各部位のデータ数が概ね均等になるようにトレーニングデータ 714 症例, バリデーションデータ 15 症例, テストデータ 45 症例に分割した. 画像サイズは 512×512 pixel に統一し, 各画素値は $[-1000, 1000]$ にクリップ後, $[-1, 1]$ に正規化した.

3.2 実験条件

提案モデルの補完性能の評価には, 平均絶対誤差 (Mean

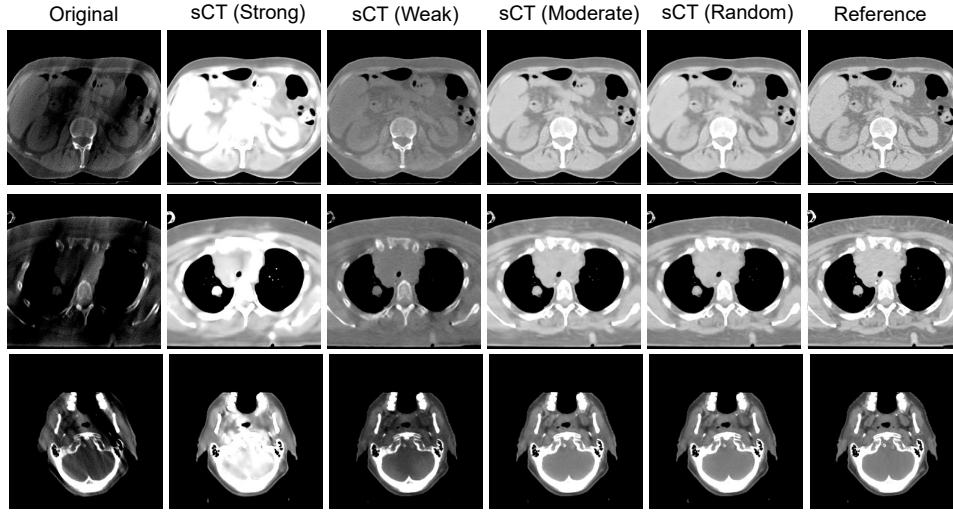


図 3 擬似 CBCT 画像の補完によって得られた生成画像と CT 画像の比較. Original: 擬似 CBCT 画像, sCT (Strong), sCT (Weak), sCT (Moderate), sCT (Random): 異なる 4 条件による学習済みモデルによって得られた生成画像, Reference: CT 画像

表 1 部位ごとの擬似 CBCT 画像補完性能

Region	MAE [HU] ↓					SSIM ↑				
	Original	Strong	Weak	Moderate	Random	Original	Strong	Weak	Moderate	Random
Thorax	64.7 ± 22.6	19.5 ± 0.4	16.4 ± 0.5	5.1 ± 0.2	5.6 ± 0.2	0.672 ± 0.058	0.975 ± 0.001	0.979 ± 0.001	0.988 ± 0.001	0.988 ± 0.001
Abdomen	75.2 ± 25.5	22.8 ± 4.2	17.8 ± 3.1	4.9 ± 0.6	5.6 ± 0.8	0.665 ± 0.05	0.974 ± 0.007	0.979 ± 0.005	0.988 ± 0.003	0.988 ± 0.003
Head-neck	38.0 ± 5.1	22.3 ± 2.1	17.7 ± 1.4	5.2 ± 0.2	5.7 ± 0.3	0.730 ± 0.031	0.972 ± 0.003	0.978 ± 0.002	0.987 ± 0.001	0.987 ± 0.001

Absolute Error, MAE) と構造的類似度指数 (Structural Similarity Index Measure, SSIM) [9] を用いた. 学習は計算時間の関係上, バッチサイズを 16, epoch 数 200 とし, バリデーションデータに対する MAE が最良となった epoch のモデルをテストに用いた. 学習で用いた擬似 CBCT 画像は, 2.1 節で示した劣化関数を用いて作成した. 各劣化パラメータは, CBCT 画像に見られるアーチファクト強度を参考に設定した. 具体的には, θ : $[-105, 75]$, m : $[0, 2]$, b : $[3e - 4, 3e - 3]$, s_c : $[1e - 3, 1e - 2]$, p : $[100, 10000]$, s_a : $[1e - 60, 1e - 80]$ の範囲とした. CBCT 画像の補完を目的とした自己教師あり学習において, 適切な擬似 CBCT 画像の生成条件は明確ではない. そこで, 上記パラメータ範囲に対して以下の 4 条件で擬似 CBCT 画像群を作成し, 学習データとして用いた.

- Weak: 最低値に固定
- Moderate: 中間値に固定
- Strong: 最高値に固定
- Random: 最低値から最高値の間でランダムに設定

3.3 擬似 CBCT 画像に対する性能評価

異なる条件で作成した自己教師データが擬似 CBCT 画像補完に与える影響を調査した. 3.2 節で示した各条件で作成した擬似 CBCT 画像を用いてモデルの訓練を行い, 4 つの学習済みモデルの補完性能をそれぞれ評価した. テストには Moderate 条件によって生成した擬似 CBCT 画像を用いた.

図 3 に擬似 CBCT 画像の補完によって得られた生成画像を示す. 左からテストに用いた擬似 CBCT 画像 (Original), 各学習済みモデルによる生成画像 sCT, 及びターゲットとなる CT 画像 (Reference) である. sCT (Strong) は, 画像中央部の

輝度補正が過剰であり, かつ, 全体的なコントラストや微細構造の再現性が最も低い結果となった. sCT (Weak) では, 画像中央部の輝度補正が不十分であるが, コントラストや微細構造の再現性は高かった. sCT (Moderate) と sCT (Random) はともに参照画像に近い外観を有したが, 微細構造の再現については sCT (Moderate) が若干優れる傾向が見られた.

生成された画像の MAE および SSIM を算出した. 表 1 に MAE および SSIM の平均値を各部位ごとに示す. 表 1 より, いずれの部位においても Moderate が最も MAE が小さくなった. 一方, SSIM は Moderate と Random で同等となった.

3.4 CBCT 画像に対する性能調査

3.3 節において良い性能が示された Moderate, Random 条件のモデルに加え, 位置合わせ済みの CT-CBCT 画像ペアを用いて学習したモデル (Real) の 3 つのモデルに対して CBCT 画像の補完性能を調査した. Real 条件における CT-CBCT 画像ペアは剛体位置合わせがなされているが, 細部構造や生体内部の空気の含有量, 位置などに差異があり, すべてのボクセルを対象に CT 画像を真値とした評価は適切とは言えない. そこで本実験では, 生成画像から典型例として数例を選出し, 関心領域での MAE 及び SSIM を算出した.

図 4 に Real, Moderate, Random の学習条件によって得られた学習済みモデルに CBCT 画像 (Original) を適用して得られた生成画像 (sCT) を示す. これら画像から, sCT (Real) はアーチファクトの補完が十分になされていたが, コントラストや微細構造の再現性は低くなった. sCT (Moderate) では, 腹部及び胸部ではアーチファクトの補正が不十分な場合が見られ, 頭頸部においては過剰な輝度補正が見られた. sCT (Random)

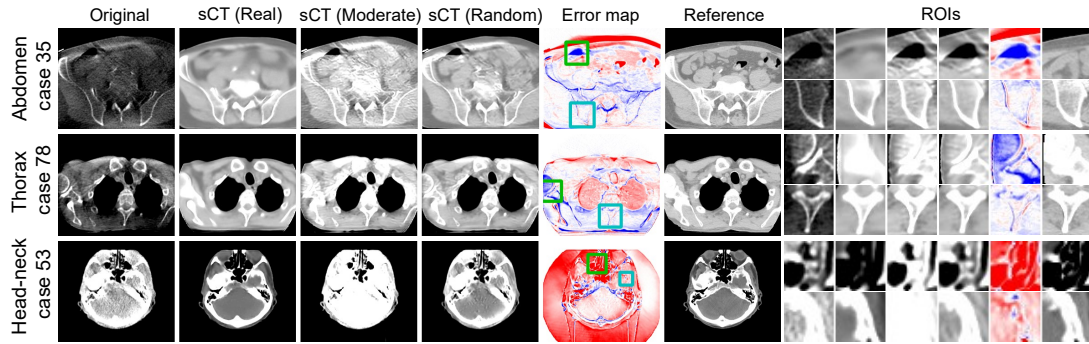


図 4 CBCT 画像の補完によって得られた生成画像と CT 画像の比較. Original: CBCT 画像, sCT (Real), sCT(Moderate), sCT(Random): 異なる 3 条件による学習済みモデルによって得られた生成画像, Error map: sCT (Random) と CT 画像の差分画像, Reference: CT 画像, ROIs: 各画像の Error map で示した関心領域

表 2 関心領域における CBCT 画像補完性能

Metric	Case	Original	Real	Moderate	Random
MAE [HU] ↓	Abdomen 35	275.4	45.3	<u>40.8</u>	36.4
	Thorax 78	271.4	51.8	<u>47.8</u>	36.7
	Head-neck 53	125.9	89.3	383.4	<u>97.7</u>
SSIM ↑	Abdomen 35	-0.365	0.615	<u>0.799</u>	0.817
	Thorax 78	-0.314	0.650	<u>0.690</u>	0.750
	Head-neck 53	0.669	<u>0.675</u>	0.450	0.753

では一部にアーチファクトの残存が見られたものの、骨格と軟部組織間の境界のコントラストが表現され、参照画像に近い外観となった。

Error map は sCT(Random) と CT 画像 (Reference) 間の差分マップであり、Reference と比較して赤は高値、青は低値を表す。Error map 内の二つの関心領域 (ROI) を拡大した画像をそれぞれ右側に示す。sCT(Real) では元画像から構造変化が見られる箇所がある一方、擬似画像を用いて学習した sCT(Moderate) 及び sCT(Random) では解剖学的構造や含有するガスの特徴が維持されていた。

典型例に対する定量評価を目的に、図 4 右側に並べた二つの関心領域のうち、下側の領域を対象に MAE 及び SSIM を算出した結果を表 2 に示す。本領域は、元画像と参照画像間の構造差が小さい骨格周辺を選択している。表 2 より、腹部、胸部領域では、sCT(Random) が MAE, SSIM とともに最も良い値を示した。一方、頭頸部領域の MAE は sCT(Real) の方が良くなった。頭頸部の sCT(Random) を観察すると、Random モデルでは同部位に特徴的な高輝度アーチファクトが十分に補正されておらず、MAE 悪化の要因となっていることが確認された。以上より、劣化関数のパラメータを可変として生成した擬似 CBCT 画像で学習したモデルは汎化性能を獲得し、多様かつ広い劣化強度分布を持つ CBCT 画像に対し安定した補完性能を示すことが判明した。一方、CBCT-CT 画像を直接学習に用いたモデルでは、画像間の構造の不一致を誤って学習し、解剖学的構造の変化や過度な平滑化に伴う微細構造の再現性の低下が生じる可能性が示唆された。

4. おわりに

本研究では、CBCT 画像に見られる多様なアーチファクト

の補完を目的として、自己教師あり一貫性モデルを提案した。異なる条件の劣化関数によって擬似 CBCT 画像を作成し、自己教師あり学習に適したシミュレーション条件を探索した。頭頸部、胸部、腹部領域の CT-CBCT 画像ペアを用いて補完性能の比較を行い、CBCT 画像が有する劣化の強度分布を包含する可変パラメータを用いて生成された擬似 CBCT 画像を学習したモデルが、CBCT 画像に含まれる解剖学的構造を維持しつつ汎化性能を有する画像補完を実現することを確認した。

文 献

- [1] X. Yi, E. Walia, and P. Babyn. Generative adversarial network in medical imaging: A review. *Medical Imaging Analysis*, Vol. 58, p. 101552, 2019.
- [2] C. Saharia, W. Chan, H. Chang, C. A. Lee, J. Ho, T. Salimans, D. J. Fleet, and M. Norouzi. Palette: image-to-image diffusion models. *ACM Transactions on Graphics*, Vol. 42, No. 2, pp. 15:1–15:4, 2023.
- [3] J. Peng, R. L. J. Qiu, J. F. Wynne, C. Chang, S. Pan, T. Wan, J. Roper, T. Liu, P. R. Patel, D. S. Yu, and X. Yang. CBCT-based synthetic CT image generation using conditional denoising diffusion probabilistic model. *arXiv preprint, Art no.:2303.02649*, 2023.
- [4] S. Luo, Y. Tan, L. Huang, and J. Li. Latent consistency models: Synthesizing high-resolution images with few-step inference. *arXiv preprint, Art no.:2310.04378*, 2023.
- [5] N. Murahashi, M. Nakamura, and M. Nakao. Cone-beam ct image quality enhancement using a latent diffusion model trained with simulated cbct artifacts. *arXiv preprint, Art no.:2603.26014*, 2026.
- [6] R. Schulze, U. Heil, D. Groß, DD. Bruellmann, E. Dranischnikow, U. Schwanecke, and E. Schoemer. Artefacts in CBCT: a review. *Dentomaxillofacial Radiology*, Vol. 40, No. 5, pp. 265–273, 2011.
- [7] H. Isogai, N. Murahashi, M. Nakamura, and M. Nakao. Limited-angle ct reconstruction using multi-volume latent consistency model. *arXiv preprint, Art no.:2603.05183*, 2026.
- [8] A. Thummerer, E. van der Bijl, A. J. Galapon Jr, F. Kamp, and M. Maspero. SynthRAD2025 grand challenge dataset: Training, 2025. <https://zenodo.org/records/15373853/> [アクセス日: Jun. 2025].
- [9] W. Zhou, A. C. Bovik, H. R. Sheikh, and E. P. Simoucellli. Image quality assessment: from error visibility to structural similarity. *IEEE Trans Image Processing*, Vol. 13, No. 4, pp. 600–612, 2004.