

Noise2Noise を用いた低線量 CT 画像のデノイジング法

A Method for Denoising from Low-dose CT Images Based on Noise2Noise

澤田 修志*, 神谷 亨*, 村上 誠一**, 李 光旭***

*九州工業大学, **純真学園大学, ***天津工業大学

Shuji Sawada*, Tohru Kamiya*, Seiichi Murakami**, Li Guangxu***

*Kyushu Institute of Technology, ** Junshin Gakuen University, ***Tiangong University

E-mail address: sawada.shuji710@mail.kyutech.jp

1. はじめに

CT 検査は、がんや腫瘍の有無や広がり調べ、治療の効果を判定する精密検査の一方法である。日本は人口 100 万人あたりの CT 装置の設置台数が先進国の中で最も多く、年間あたりの医療被ばく線量は世界平均の約 6 倍以上となっている[1]。放射線による治療が多大な恩恵をもたらしている一方で、検査時の医療被ばくが問題視されている。医療分野での被ばくは、健康被害と直接的な関係は証明されていないものの、年月が経過してがんや白血病を引き起こすリスクが増加する確率的影響があると示唆されている[2]。これらの問題解決のため、CT 撮影時の放射線の線量を下げることが解決策の一つとして提案されている。しかし、低線量 CT 画像は通常線量と比較して画像の鮮鋭度が低下し、ノイズが増加する。そのため、読影医師による判断が困難になり、病変部の未検出が生じる可能性がある。図 1 に低線量 CT 画像の例を示す。同図より、線量の違いによる画質の変化が大きいことが確認できる。

そこで、必要な情報を保持したまま低線量 CT 画像のノイズ除去を行う画像処理手法の開発が求められている。画像を対象とした畳み込みニューラルネットワーク (CNN: Convolutional Neural Network) は、従来の画像処理手法をはるかに凌ぐ性能を持つことが確認されており、多くの手法が提案されている[3]。そこで、本稿では学習時にノイズなしの画像を利用しない Noise2Noise の学習法を用いた独自の CNN モデルを提案し、オープンデータである子豚の低線量 CT 画像[4]を用いた結果について述べる。

2. 手法

2.1 Noise2Noise

本稿では、ノイズ除去の深層学習モデルとして、ノイズを含む画像を 2 つ用意し、それらをペアとして学習させる Noise2Noise[5]を用いる。Noise2Noise では、想定するノイズが所定の条件下であれば、ターゲットとするデータにノイズが含まれていても学習が可能であるとされている。特に、CT や MRI などといった元々ノイズがないデータの取得が難しい条件下では、有効性が高いとして注目されている。一般に、ノイズ除去タスクにおける

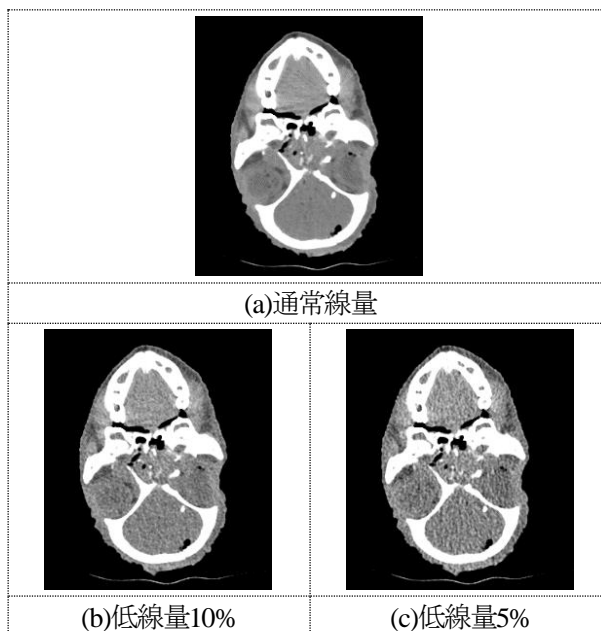


図1 低線量CT画像の例

学習は、式(1)のように \hat{x}_i で表されるノイズ画像と、 y_i で表される正解画像をペアとして学習し、損失を最小化させるようなフィルタ f_θ を求めることにより学習を行う。一方で、Noise2Noise はノイズ画像 \hat{x}_i と別のノイズ画像 \hat{y}_i のペアで学習を行う。Noise2Noise の学習式を以下の式(2)に示す。

$$\operatorname{argmin} \sum_i L(f_\theta(\hat{x}_i), y_i) \quad (1)$$

$$\operatorname{argmin} \sum_i L(f_\theta(\hat{x}_i), \hat{y}_i) \quad (2)$$

一般に、低解像度画像 x に対応する高解像度画像 y は一意に定まらず、高解像度画像 y は複数存在する。そのため、フィルタ f_θ は y の平均値を出力するような学習が進むため、ぼかし処理を生成する。ここで、 y の平均値を出力するフィルタ f_θ が得られるとわかっているなら、その関数が得られる条件下で y を \hat{y} に置き換えることが可能となる。この方法により、クリーンなデータの分布や、その分布におけるノイズの分布を必要としないため、ノイズデータのみを収集することによりノイズ低減処理が可能となる。

2.2 ネットワークアーキテクチャ

具体的なネットワークアーキテクチャとしては、医療画像分野におけるセマンティックセグメンテーションで有効性が示された U-Net[6]をベースモデルとして用いる。U-Netはエンコーダデコーダがスキップ接続されており、入力画像に近い次元数の特徴マップの連結により、画像復元の際の位置情報の補完を可能とする。また、Noise2Noiseの学習モデルでは、セグメンテーションタスクにおける U-Net とはチャンネル数や層数が異なる形状をとる。本稿では、ベースモデルを参考に、Noise2Noiseで用いられるデノイズタスク用 U-Net に対して改良を加える。実験ではセグメンテーション用 U-Net アーキテクチャと比較することにより、提案手法の優位性を示す。

2.3 Attention 機構

Attention とは、入力されたデータのどこに注目すべきかを動的に特定する仕組みで、自然言語処理や Transformer など高度なモデルで用いられ、最高精度を更新し続けている機構の一つである。本稿ではチャンネルごとに重みづけを行い、強調すべき特徴を抽出する Channel Attention[7]と、二次元画像中のどこに注目するかを決定する Spatial Attention[8]の二つの Attention 機構からなる CBAM(Convolutional Block Attention Module)[8]の導入により、特徴マップの補完を図る。CBAMの構造を図2に示す。

2.3.1 Channel Attention

Channel Attention の構造を図3に示す。Channel Attention は、特徴チャンネル間の依存関係を使用することとなるため、周波数成分フィルタリングのような作用を発揮し、周波数の違いが重要であるデノイズタスクにおいては有用であるとされている。Channel Attention のアテンションマップ M_c は以下の式(3)で計算される。

$$M_c = \sigma \left(\text{MLP}(\text{AvgPool}(F)) + \text{MLP}(\text{MaxPool}(F)) \right) \quad (3)$$

ここで、 F は特徴マップ情報、 σ はシグモイド活性化関数、AvgPool は平均値プーリング、MaxPool は最大値プーリングを表す。平均値プーリングならびに最大値プーリングを施した特徴マップに対し、1つの隠れ層を持つ多層パーセプトロン(MLP)を適用する。その後、2つの特徴ベクトルを統合し、シグモイド関数により活性化を行う。

2.3.2 Spatial Attention

Spatial Attention の構造を図4に示す。Spatial Attention は、画像中にながどこにあるかを強調するようなマスクを作成するため、エッジを含むノイズ以外の要素を強調することを目的として導入する。Spatial Attention のアテンションマップ M_s は以下の式(4)により計算される。

$$M_s = \sigma \left(f^{7 \times 7} \left((\text{AvgPool}(F)); (\text{MaxPool}(F)) \right) \right) \quad (4)$$

ここで、 $f^{7 \times 7}$ はフィルタサイズ 7×7 の畳み込み演算を表す。チャンネル軸に沿って最大値、平均値プーリングを適用し、獲得した特徴マップを連結することにより空間方向に特化した特徴の抽出を可能とする。

2.4 提案モデル

本稿で提案するモデルを以下の図5に示す。デノイズタスク用に改良された U-Net に対し、エンコーダ、デコーダ側ともに畳み込み層に Attention 機構の CBAM を導入することにより、保持したいエッジ情報の抽出を図った。また、損失関数は MSE、最適化アルゴリズムには Adam を使用し、学習率 0.001、 $\beta_1 = 0.9$ 、 $\beta_2 = 0.999$ とし、エポック数 1000 で学習を行った。

3. 実験と結果

本稿では、子豚の低線量 CT 画像(通常線量に対して線量が 10%、5%の CT 画像)906 枚に対し、作成した改良型 U-Net を適用し、Noise2Noise の学習法にならった教師なし学習を行った。学習データについては、通常線量 906 枚に、標準偏差 $\sigma = 40$ の正規分布に従ったガウシアンノイズを付与したものを使用した。学習に使用したノイズ画像の例を図6に示す。

また、評価指標として画質の客観的評価指標である PSNR(Peak Signal to Noise Ratio)を使用した。PSNR の定義式は以下の式(5)に示す。ここで MSE(Mean Squared Error) は平均二乗誤差を表し、2枚の画像の画素ごとの差分の二乗平均を表す。また、MAX は画像がとりうる最大の画素値を表す。

$$PSNR = 10 \log_{10} \frac{MAX^2}{MSE} \quad (5)$$

3.1 実験結果

実験では、テストデータとして低線量 CT 画像 906 枚を使用し、各モデルを適用した出力画像と平均 PSNR の値より、比較検討を行った。比較するモデルについては、(1)ベースモデルである U-Net、(2)デノイズタスク用 U-Net、本稿の提案手法である(3) デノイズタスク用 U-Net + CBAM である。通常線量と低線量 CT 画像とを比較した平均 PSNR、ならびに各モデルを適用した平均 PSNR を表1に示す。また、通常線量画像ならびに低線量画像、各モデルを適用した画像を図7,図8にそれぞれ示す。

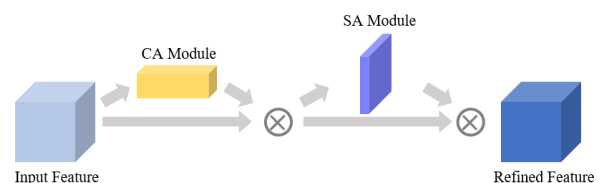


図2 CBAM の構造

るため、モデルの改良や異常組織部位に注目したノイズ除去を検討する必要がある、それらは今後の課題である。

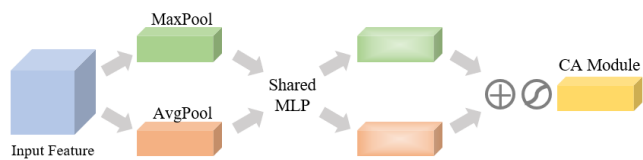


図3 Channel Attention Module

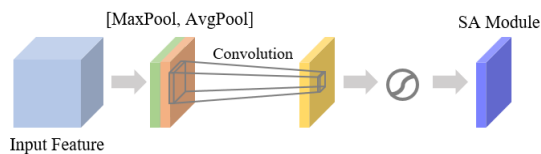


図4 Spatial Attention Module

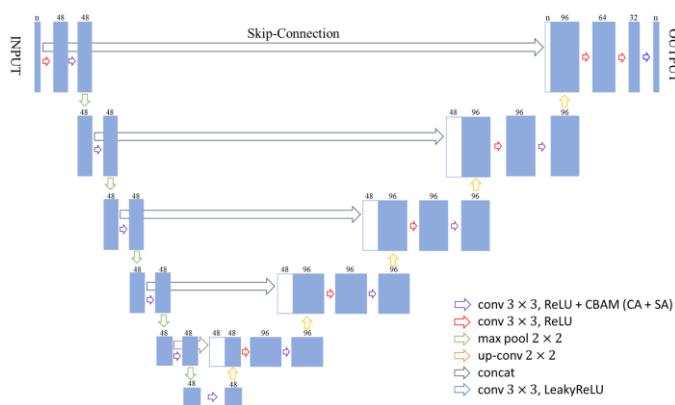


図5 提案ネットワーク

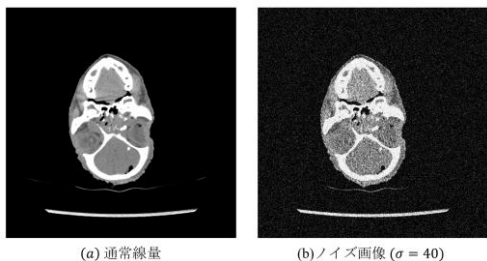


図6 ノイズ付与CT画像の例

本実験より、U-Netの改良ならびにCBAMの導入により他モデルと比較して高いPSNRを獲得した。これは、デノイズタスクではセグメンテーションタスクほどの特徴量が必要ないため、U-Netの長所を生かしつつ、エッジを保持したノイズ除去が可能になったと考える。また、CBAMの導入により、なにがどこにあるかの情報が保持され、低線量により失われたエッジ情報を復元することができたことが性能向上の一因であると考えられる。

4. おわりに

本論文では、Noise2Noiseの学習法を用いたAttention機構を搭載する改良型U-Netによるノイズ除去を行った。子豚の低線量CT画像5%、10%に適応した結果、それぞれPSNRが29.1154、31.2809となり、提案モデルによる画質改善が実現できた。今後は、さらなる精度向上を図

表1 実験結果

	平均 PSNR	
	低線量 5%	低線量 10%
通常線量 (100%)	26.1749	29.0530
Model(1)	27.8482	30.6732
Model(2)	27.8942	30.4973
提案手法	29.1154	31.2809

参考文献

- [1] 環境省, 保健化学物質対策, 放射線による健康影響等に関する統一的な基礎資料, https://www.env.go.jp/chemi/rhm/kisoshiryo/pdf_h29/2017tk1s02.pdf (access:2023/09/10).
- [2] 鳥取市立病院放射線部, 医療被ばくの基礎知識, <https://hospital.tottori.tottori.jp/files/20191216165341.pdf> (access:2023/09/11).
- [3] Hu Chen et al., "Low-dose CT via convolutional neural network", Biomed. Opt. Express 8, pp.679-694, 2017.
- [4] X. Yi, Recent publication., <https://onedrive.live.com/?authkey=%21APVzpTOocrEDIdk&cid=FD4F832E5E42083E&id=FD4F832E5E42083E%21582&parId=root&action=locate> (access: 2023/01/10).

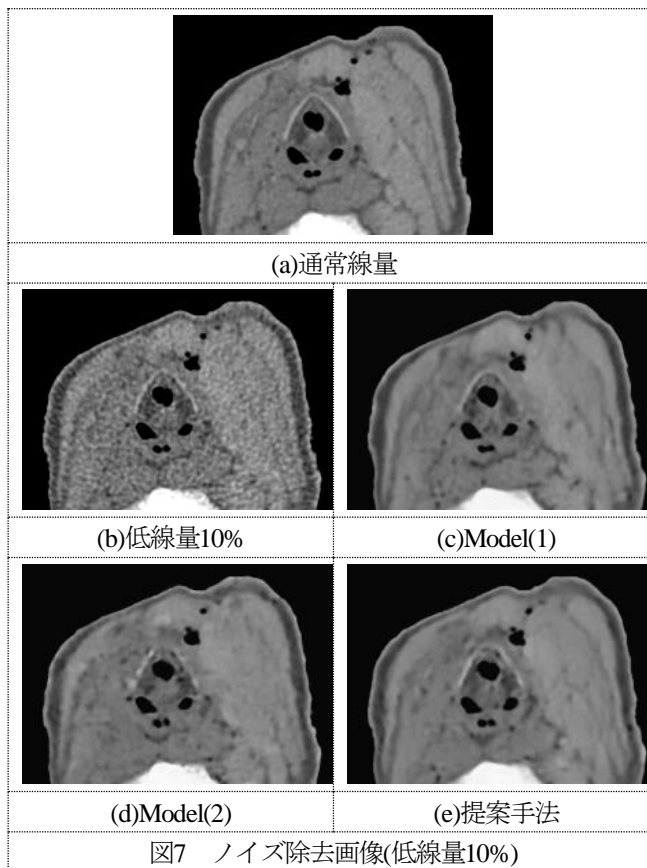
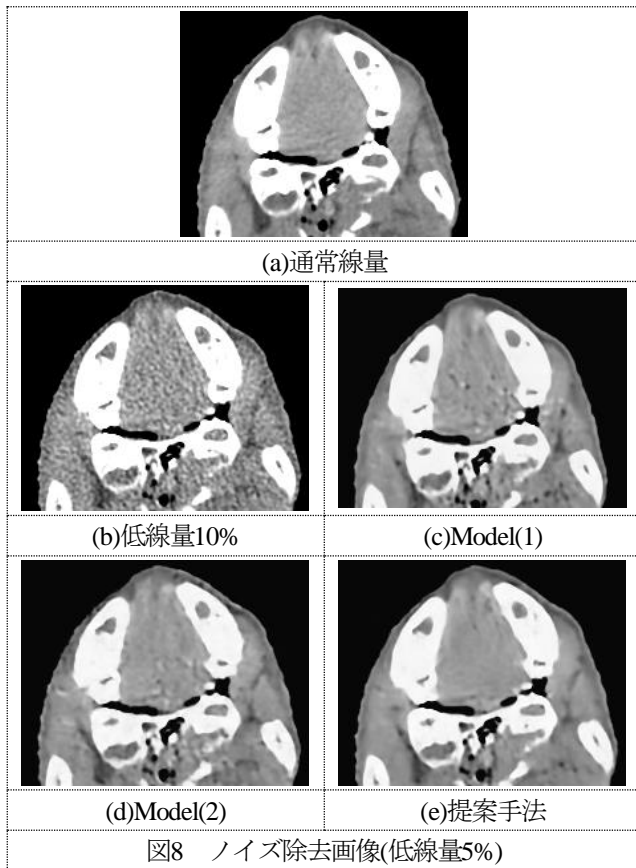


図7 ノイズ除去画像(低線量10%)



- [5] Lehtinen Jaakko et al., “Noise2Noise: Learning image restoration without clean data”, arXiv:1803.04189v3, 2019.
- [6] O. Ronneberger et al., “U-Net: Convolutional networks for biomedical image segmentation”, Medical Image Computing and Computer-Assisted Intervention, pp.234-241, 2015
- [7] Yulun Zhang et al., “Image super-resolution using very deep residual channel attention networks”, European Conference on Computer Vision, pp.286-301, 2018.
- [8] Sanghyun Woo et al., “CBAM: Convolutional block attention module”, European Conference on Computer Vision, pp.3-19, 2018.