

# Deep CNN を用いた低線量 CT 画像のデノイジング処理

貞松 勇太<sup>†</sup> 村上 誠一<sup>‡</sup> 神谷 亨<sup>†</sup> 李 光旭<sup>§</sup>

<sup>†</sup>九州工業大学 〒804-8550 福岡県北九州市戸畑区仙水町 1-1

<sup>‡</sup>純真学園大学 〒815-8510 福岡県福岡市南区筑紫丘 1-1-1

<sup>§</sup>天津工業大学 天津市西青区賓水西道 399 号

E-mail: kamiya@cntl.kyutech.ac.jp

## あらまし

CT 検査は日本で広く用いられ、国民の健康に多大な恩恵をもたらしている。一方で、放射線被ばくリスクも存在する。この問題を解決するため、撮影する際の線量を減らす試みが行われている。しかし、放射線量を減らすと画像にノイズが発生し、画質低下を招くという問題が発生する。そこで本論文では、Deep CNN の活性化関数に Mish 関数を用いることにより、画像の必要情報を失うことなく、効果的なノイズ除去を行うための画像解析法を提案する。豚の全身スライス CT 画像を用いた実験では、通常線量 CT 画像との PSNR(Peak Signal to Noise Ratio)による画質評価を行い、提案手法の有用性を検証した。

キーワード Low-dose CT Image, Convolutional Neural Network, Deep CNN, Mish Activation

## 1. 序論

CT(Computed Tomography)とは、X 線を発生させる X 線管と、それを検出する検出器が体の周りを回転し、体を透過した X 線を検出し体内を検査する技術である。骨や臓器によって放射線の吸収率が異なるため、その違いをコンピュータ処理し体の断面図の画像が生成される。CT 検査は治療前に腫瘍の有無や位置、他の臓器への転移、治療の効果を非侵襲的かつ正確に検査することができるため広く用いられ、健康寿命促進に多大な恩恵をもたらしている。

特に日本では CT 検査装置の導入が進んでいる。経済協力開発機構(OECD: Organization for Economic Cooperation and Development)の調査によると、2018 年時点での人口 100 万人当たりの CT 台数は日本が OECD 加盟国の中で最多であり、日本が世界的に見ても群を抜いて CT の普及が進んでいることがわかる[1]。このように、日本では多くの病院で最先端の放射線検査を受けることができ、世界一の長寿国になる要因の一つとなっている。

一方で、CT 検査は X 線を使用するため、放射線被ばくリスクが懸念される。人工放射線による被ばく線量は個人差があるものの、日本の被ばく量平均は、世界の平均被ばく量を大きく上回っている[2]。これは、CT 検査が容易に受けることが可能な環境に起因している。日本ではこの過剰な医療被ばくが問題になっており、諸説はあるが他の先進 15 か国と比較すると、X 線検査に伴う 75 歳までの発がん率は最も高い数値である[3]。

この問題を解決するため、CT を撮影する際の線量を少なくする取り組みが行われている。しかし、撮影する際の線量を少なくすると図 1 に示すように、得られる画像にノイズが発生し、解像度の低下を招き、その画像を用いた診断の信頼度が低下するという問題が発生する。そのため、古くから画像中のノイズ除去に関する研究が行われている。その一つにフィルタを用いた処理がある。しかし、単一フィルタを用いた処理では、ノイズを除去すると画像の鮮鋭度が低下し、鮮鋭度を上げようとするとノイズが増加する問題がある。そのため、近年では畳み込みニューラルネットワーク(CNN: Convolution Neural Network)を用いた手法が目まぐるしく注目されている。この手法を用いることにより画像の鮮鋭度を落とすことなく、ノイズを効果的に除去できる[4]。そこで、本論文では低線量 CT 画像のデノイズを目的とした CNN モデルの提案を行う。なお、実験では、学習の精度向上のため活性化関数の改善を行った後、画質の客観的評価により、従来手法との比較を行う。

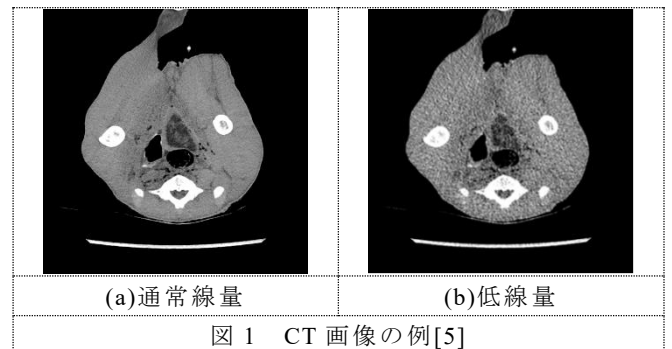


図 1 CT 画像の例[5]

## 2. 手法

### 2.1. ネットワーク構造

本論文では Deep CNN[6]をベースとした CNN モデルを採用する。Deep CNN は、元画像からノイズを除去するだけでなく、低解像度の画像を鮮明化することが可能な、高速で効果的な CNN の一種である。これは、最初の層の“DilatedConvolution+ReLU”，中間層である5つの“Dilated Convolution+Batch Normalization+ReLU”，最後の層である“Dilated Convolution”の7層で構成されている。第一層から最終層まで(3×3)畳み込み、Dilated Convolution の拡張係数はそれぞれ 1, 2, 3, 4, 3, 2, 1 としており、各中間層における特徴マップは 64 としている。ネットワークの構造を図 2 に示す。

この Deep CNN に Ghilzadeh-Ansari らの研究[4]を参考に、エッジ検出層、スキップコネクションと中間層を追加したネットワークを最終的なベースとなる構造とする。

そのため本論文では、Ghilzadeh-Ansari らの Deep CNN 構造中の ReLU 関数に代わって Mish 関数を採用することにより効率的に学習を行い、精度向上を目指した。実験では、Parametric ReLU 関数と比較することにより提案手法の優位性を示す。提案手法のネットワーク構造を図 3 に示す。

### 2.2. エッジ検出層

画像処理において、画像中の物体の境界を検出する技術をエッジ検出という。本論文ではネットワークの性能を向上させるため、単純なエッジ検出技術を採用する。Sobel edge detection operator[7]は画像の二次元勾配を計算し、画像を 3×3 フィルタで畳み込むことにより、高い空間周波数成分を持つ領域を強調するものである。本論文のエッジ検出層は、4 つのソーベルカーネルを非学習型フィルタとして持つ畳み込み層である。この層での出力画像は入力画像と連結されネットワークに与えられる。

### 2.3. 活性化関数

Mish 関数[8]は、2019 年に発表された Mish 活性化関数である。多くのベンチマークテストで ReLU 関数より良い結果を残している。Mish 関数の概形は図 4 に示され、式は以下のように定義される。

$$f(x) = x \tanh(\text{softpuls}(x)) = \tanh(\ln(1 + e^x)) \quad (1)$$

Mish 関数は次の 4 つの特徴がある。

(1) 上限が存在しない。

Mish 関数は ReLU 関数と同様、入力値が正の場合、

出力値はそのままの値を用いる。そのため、勾配消失問題が解消され、効率的な学習が可能となる。

(2) 下限が存在する。

下限が存在し、強い正則化効果がもたらされる。

(3) 負の値を持つ。

ReLU 関数では負の値は 0 として出力されるため、学習時に重みの更新が行われない可能性が存在していた。Parametric ReLU のように負の値を出力することによりその問題を解消する。

(4) 微分しても連続である。

ReLU 関数では、0 のときに不連続であり微分可能でないため、勾配に基づく最適化において好ましくない問題が引き起こされる。しかし、図 4 に示すように、Mish 関数では微分しても関数が連続であるため、この問題を解決することができる。

本論文では、CNN モデルの活性化関数に注目する。Thakur らの研究[9]により、Parametric ReLU 関数を画像のノイズ除去 CNN に用いることにより、ReLU 関数を用いたモデルよりも画像のノイズ除去性能が向上する。Mish 関数は多くの画像認識テストにおいて Parametric ReLU 関数を上回る精度を記録している。

## 3. 実験と結果

### 3.1. データセットの詳細

本論文で使用したデータセットは豚の頭部から骨盤までの全身 CT 画像[5]であり、通常線量と低線量がそれぞれ 906 枚である。撮影条件は表 1 に示す。

### 3.2. 入力画像の生成

本論文では、画像サイズ 512×512 の DICOM 形式を BMP 形式に変換して用いる。教師画像に通常線量 CT 画像、対応する訓練画像を通常線量の 5% の低線量 CT 画像とする。また、学習時は元画像を 64×64 パッチに分割し、学習を行う。小さなパッチに分割することにより、CNN がより多くの境界情報を参照でき、境界のアーティファクトを避けることができる。また、小さなサイズの学習サンプルを用いるため、メモリリソースを節約することができる。

### 3.3. 評価指標

本論文では、5 分割交差検証による手法の性能評価を行う。各モデルに対し、各データセットの平均 PSNR(Peak Signal to Noise Ratio)と全データセットの平均 PSNR により評価を行なう。

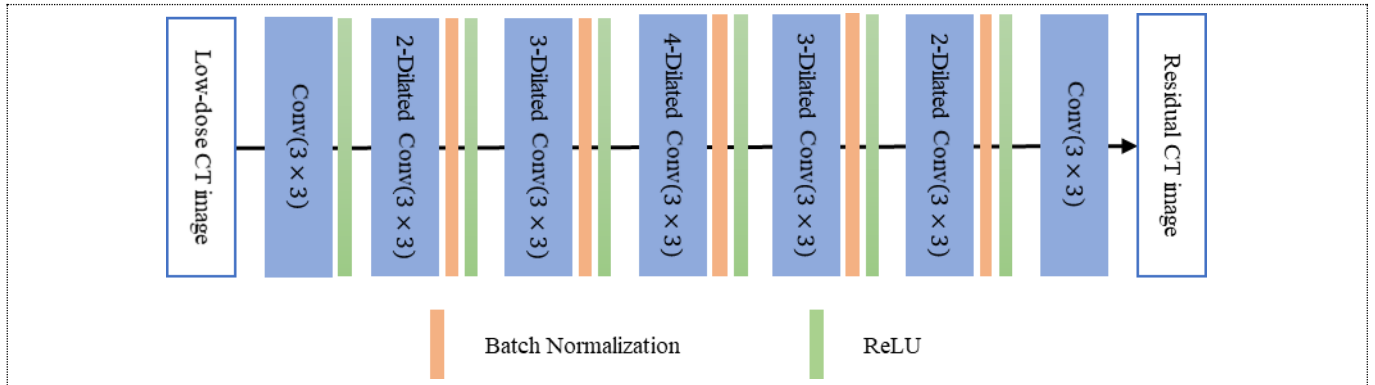


図2 基本 Deep CNN 構造

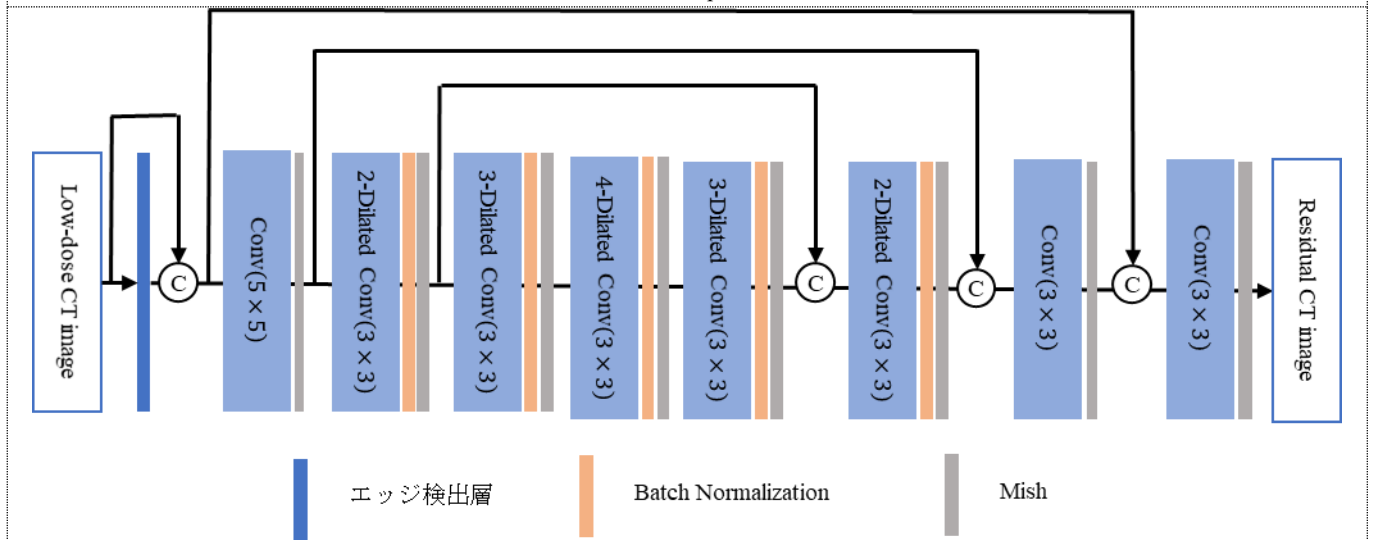


図3 提案手法のアーキテクチャ

表1 CT画像の撮影条件

	通常線量	低線量
ガントリ回転速度(s)	1	1
スライス厚(mm)	0.625	0.625
管電圧(kV)	100	100
管電流(mAs)	300	15
CTDI <sub>vol</sub> (mGy)	30.83	1.54
DLP(mGy-cm)	943.24	47016
実効線量(mSv)	14.14	0.71

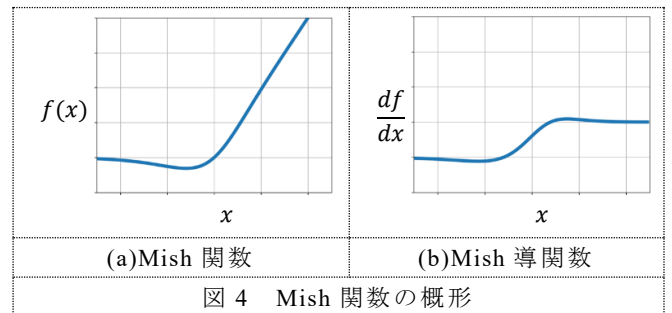


図4 Mish関数の概形

表2 実験結果

	Dataset1	Dataset2	Dataset3	Dataset4	Dataset5	Average
低線量	29.2964	28.8387	24.5737	24.6464	23.7427	26.2168
(1)	34.1369	32.6565	30.8397	29.3309	26.2228	30.6374
(2)	34.3928	32.7425	31.272	30.0846	28.936	31.4528
(3)	34.3291	32.8247	31.3137	30.0213	29.0026	31.4995
(4)	34.5262	32.4766	31.356	30.3084	28.9905	31.5134

### 3.4.実験結果

実験では、低線量 CT 画像、(1)Deep CNN + ReLU, (2) Ghilizadeh-Ansari らの Deep CNN + ReLU, (3) Ghilizadeh-Ansari らの Deep CNN + Parametric ReLU, (4)本論文の提案手法である Ghilizadeh-Ansari らの Deep CNN + Mish, これらの5つの出力画像と通常線量の CT 画像との PSNR を比較した。また、Adam のパラメータは、学習率 0.001,  $\rho_1=0.9$ ,  $\rho_2=0.999$  とし、エポック数 20 で学習を行う。各データセットでは、ネットワークの train 用 60%, validation 用 20%, test 用 20%に分割している。得られた結果を表 2 に示す。また、出力画像例を図 5 に示す。同図よりほかのモデルに比べ、提案手法である model(4)のデノイジング効果が視認できる。

### 4.考察と結論

本論文では、Deep CNN をベースとし、4 種類の CNN モデルを構築し、豚の全身 CT 画像に対し、デノイジング処理を行った。実験結果より、活性化関数の変更は、ノイズ除去精度を向上させることが確認された。また、提案手法である model(4)における平均 PSNR が、他手法と比べ最も良い値を得ることができ、本提案手法の有効性が確認された。

今後の展望として、同じ部位でのデータセットでの検証が挙げられる。本論文で用いたデータセットは、豚の頭部から骨盤にかけての全身 CT 画像での検証であるため、同一部位のみのデータセットを用いた場合における、精度向上への影響を検証する必要がある。しかし、問題点として、データセットの不足が挙げられる。本論文での CNN は、通常線量と低線量のペア画像が必要となるが、それらのデータセットを集めるのは困難である。そのため、通常線量 CT 画像にポアソン分布に基づいた疑似ノイズを付与することで、データセットを作成し、検証を行いたい。

### 文 献

- [1] OECD, Health at a Glance 2019: OECD indicators, pp.193, OECD publishing, Paris, 2019.
- [2] 環境省, 放射線による健康影響等に関する統一的な基礎資料 平成 30 年度版, pp.64-66, 2019.
- [3] A. Berrington de González et al., “Risk of cancer from diagnostic X-rays: estimates for the UK and 14 other countries.” The Lancet, vol. 363, pp.345-351, 2004.
- [4] M. Gholizadeh-Ansari et al., “Deep learning for low-dose CT denoising using perceptual loss and edge detection layer” Journal of digital imaging vol. 33(2), pp.504-515, 2020.
- [5] X Yi, Recent publication., [https://docs.google.com/forms/d/e/1FAIpQLSdnIxHhX\\_fDC4FSqnTe\\_BbvF3MR4N1hkIMT0T5SHwcmqP0Q/viewform](https://docs.google.com/forms/d/e/1FAIpQLSdnIxHhX_fDC4FSqnTe_BbvF3MR4N1hkIMT0T5SHwcmqP0Q/viewform) (access: 2022.2.8)
- [6] K. Zhang et al., “Learning deep CNN denoiser prior for image restoration”, IEEE Conference on Computer Vision and Pattern Recognition, vol. 2, pp.3929-3938, 2017.
- [7] I. Sobel, “An isotropic 3×3 image gradient operator.” Machine vision for three-dimensional scenes, pp. 376–379, 1990.
- [8] M. Diganta, “Mish: A self regularized non-monotonic neural activation function” arXiv preprint arXiv:1908.08681, 2019
- [9] R. S. Thakur et al., “PReLU and edge-aware filter-base image denoiser using convolutional neural network”, IET Image Processing, 14(15), pp.3869-3879, 2020.

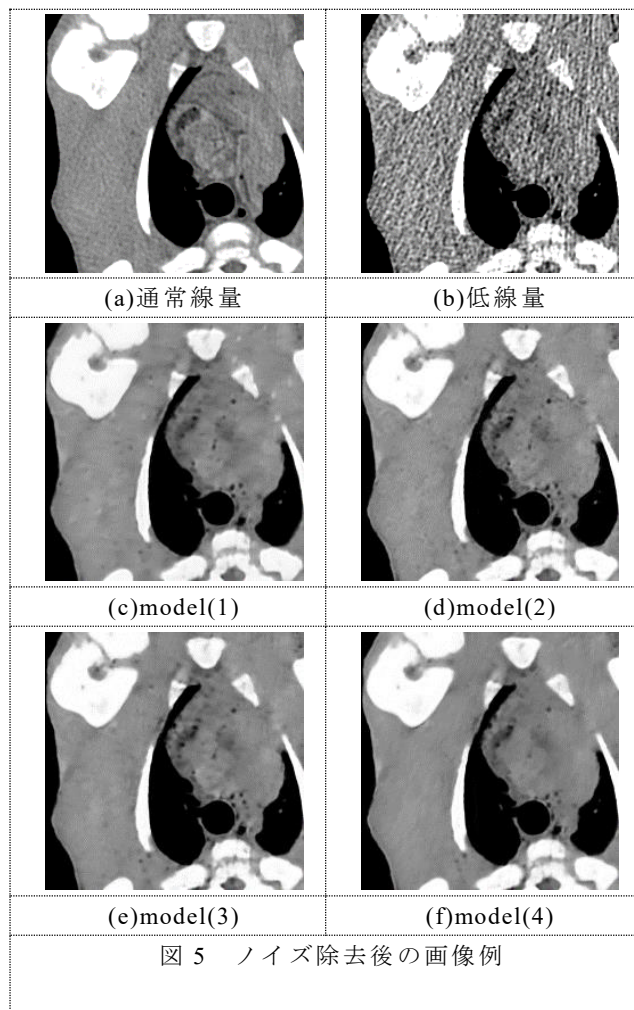


図 5 ノイズ除去後の画像例