

# Data Augmentation for Medical Image Segmentation

Sohei YAMAKAWA<sup>\*</sup>, Ryoya KOGACHI<sup>\*</sup>, Keiko ONO<sup>\*\*</sup>, Erina MAKIHARA<sup>\*\*</sup>, Akitaka YAGUCHI<sup>\*</sup> and  
Daisuke TAWARA<sup>\*\*\*</sup>

(Received October 14, 2021)

Semantic segmentation of medical images aims to associate a pixel with a label in a medical image without human interaction. The success of semantic segmentation using deep learning techniques is contingent on the availability of a large amount of imaging data with corresponding labels provided by experts. In contrast, a large amount of labeled medical image data is not available in many cases. In this study, we investigate an efficient data augmentation method using various filters. The experimental results showed that the combination between Gaussian and Median filters is adequate for semantic segmentation for bone images.

**Key words :** Semantic Segmentation, U-Net, Diagnostic Imaging, Data Augmentation

**キーワード :** セマンティックセグメンテーション, U-Net, 画像診断, データ水増し

## 医用画像セグメンテーションのための画像水増し法の検証

山川 蒼平 · 古我知 亮弥 · 小野 景子 · 槇原 絵里奈 · 矢口 瑛貴 · 田原 大輔

### 1. はじめに

近年、医療現場では人不足が問題となっており、機械学習の活用に注目されている。医療現場で機械学習を用いることで、人手不足の解消や人件費の削減、医療サービスの生産性向上などが見込める。医療現場における機械学習の活用例として医療用画像診断が挙げられ、CT スキャンやMRIなどの画像診断に用いられる。また、日本は人口当たりのCT 機器台数が世界一であるが<sup>1)</sup>、人口当たりの放射線科医の数は世界と比較して少ない。8 か国についてCT の検査数の合計を放射線診断医の数と比較した場合に、放射線診断医 1

人当たりの検査数は日本が最も多い<sup>2,3)</sup>。これに伴い、1 検査の画像診断に費やすことができる時間が短くなり、見落としや誤診の危険性が増すほか<sup>4)</sup>、肺癌 CT 検査では受診者 1 人に対して 100 枚以上の写真を確認するなど身体的、精神的にも負担が大きいことが問題視されている。さらに、過疎地など医師が少ない地域では診断ができないという問題も指摘されており、機械学習によって画像診断を行うことで医師の作業負担軽減へ向けた取り組みが進んでいる。このような問題を解決するために U-Net<sup>5)</sup> を用いたセマンティックセグメンテーションが注目されている。しかし、大量

<sup>\*</sup> Graduate School of Science and Engineering, Doshisha University, Kyoto  
Telephone:+81-90-4214-9852,E-mail:yamakawa.sohei@mikilab.doshisha.ac.jp

<sup>\*\*</sup> Faculty of Science and Engineering, Doshisha University, Kyoto  
Telephone:+81-774-65-6930,E-mail:kono@mail.doshisha.ac.jp

<sup>\*\*\*</sup> Faculty of Advanced Science and Technology, Ryukoku University, Kyoto  
E-mail:datawara@rins.ryukoku.ac.jp

のピクセル毎にラベル付けされた画像が必要であるため、コストがかかるという課題が挙げられる。そこで本研究では医療画像に有効な画像フィルタの組み合わせを明らかにする。性能検証では、先行研究<sup>5)</sup>に用いられている、Accuracy や Dice 係数に加え、本研究では IoU (Intersection over Union) を追加し、複数の評価指標での検証を行うことで、画像の水増し手法の選定による学習データ削減の実現可能性を検証する。

## 2. セマンティックセグメンテーション

セマンティックセグメンテーションは、画像の各画素にラベルやカテゴリを関連付ける手法である<sup>6)</sup>。特徴的なカテゴリを形成する画素の集まりを認識するために使用され、自動運転や医療用画像処理、工業用検査などに利用される。自動運転に用いられるセマンティックセグメンテーションを Fig.1 に示す。図より、道路と自動車が分離されており、正しく領域分類されていることがわかる。畳み込みニューラルネットワーク (CNN) を用いたセマンティックセグメンテーションには、主に全層畳み込みネットワーク (FCN)<sup>8)</sup> や SegNet<sup>9)</sup>、U-Net などの手法が用いられている。セマンティックセグメンテーション学習プロセスを以下に示す。

1. 画素にラベルが付けられた画像の集まりを解析。
2. セマンティックセグメンテーションのネットワークを作成。
3. 画像を画素カテゴリに分類するためネットワークの学習を行う。
4. ネットワークの精度を評価

## 3. U-Net

### 3.1 概要

U-Net とは、セマンティックセグメンテーションの一種であり、生物医学画像のセグメンテーションを行うために提案された手法である。特徴として、学習時間が非常に短いことが挙げられる。細胞の識別精度を競う大会において最も良い成績を残しており<sup>5)</sup>、領



Fig. 1. Semantic segmentation for self-driving.<sup>7)</sup>

域抽出において優れた学習モデルであることがわかる。通常の CNN (Convolutional Neural Network) とは異なり、物体の局所的特徴と位置情報を統合して作られた学習モデルである。通常の CNN によって行われる画像認識では、畳み込み層が物体の局所的な特徴を抽出する役割を担い、プーリング層で局所の特徴は残しつつ画像を縮小している。縮小する理由は、次の層で画像を扱いやすくするためである。しかし、位置情報が失われる問題がある。位置情報の有無が FCN と CNN との違いである。局所の特徴と位置情報を統合する方法は、CNN で用いられていた、全結合層を畳み込み層に置き換える。畳み込み処理を加えていくと、局所の特徴は抽出できるが、位置情報についての特徴は失われる。しかし、畳み込み後、特徴マップを保持しておき、後で逆畳み込みをする画像に足し合わせることで位置情報に着目した結果が得られる。

### 3.2 U-Net の構造

U-Net の構造を Fig.2 に示す。畳み込み層、プーリング層、cropping 層で構成され、U 字型のネットワークになっていることから U-Net という名がつけられた。U 字型ネットワークの左側は収束経路といい、畳み込みとプーリングにより深い層ほど特徴が局所的で位置情報が曖昧に、浅い層ほど特徴は曖昧で位置情報が正確になる。U 字型ネットワークの右側は拡散経路といい、逆畳み込みにより、特徴を保持したまま画像を復元することができる。両方の層において、画像サイズが同じものを深い層から段階的に統合することによって、局所の特徴を保持したまま全体的位置情報の復元を行うことが可能である。

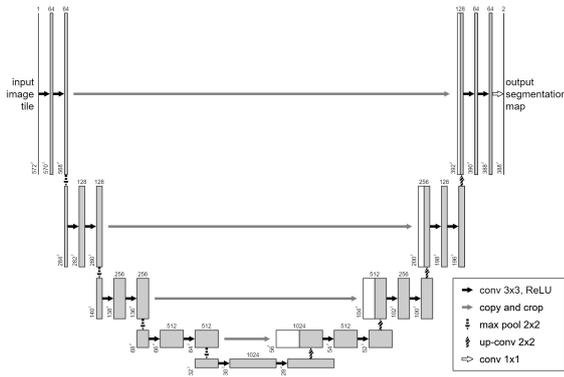


Fig. 2. Structure of U-Net.

3.3 収束経路

収束経路とは、Fig.2 左半分に位置しており、特徴マップを縮小させる役割をもつ<sup>10)</sup>。特徴マップとは、畳み込みによって得られた多次元配列のことである。カーネルサイズ 3 × 3 の畳み込み層、カーネルサイズ 2 × 2 の Max Pooling 層で構成されており、特徴抽出を行う。また、Max Pooling 層によるダウンサンプリング処理で特徴マップの数が 2 倍になる。

3.4 拡散経路

拡散経路とは Fig.2 の右半分に位置しており、特徴マップの解像度を上げる。カーネルサイズ 2 × 2 の up-convolution 層、対応する収束経路の特徴マップをコピーし外挿する cropping 層、カーネルサイズ 3 × 3 の畳み込み層で構成されており、入力画像におけるセグメンテーションを決定する。また up-convolution 層によるアップサンプリング処理で特徴マップの数が半分になる<sup>11)</sup>。また、位置情報保持のために cropping 層で収束経路の特徴マップをコピーし拡散経路の特徴マップと共に畳み込み処理を行うことで、位置情報を保持、また特徴抽出の両立が可能となる。

3.5 Overlap-tile 法

Overlap-tile 法とは、訓練データの拡張のために画像のミラーリングを行う手法である。ピクセル単の画像認識では全ての画素を個別に入力として扱うが、この方法では 1 枚の画像に対する処理全体を通して膨大な計算を要する。そこで、U-Net はミニバッチと呼ばれる複数の入力の集合をセットとして扱う入力形式を用いており、これを Overlap-tile 法によって実装して

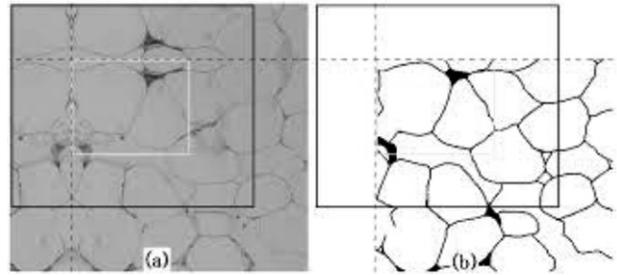


Fig. 3. Example of before(a) and after(b) segmentation pro-cessing by Overlap-tile method.<sup>11)</sup>

いる<sup>11)</sup>。Overlap-tile 法では入力として与えられた画素のセットに対して、周辺画素を用いて Fig.2 に示すレイヤー処理を行い、セグメンテーション結果を出力する。処理の例を Fig.3 に示す。Fig.3(a) の黄枠内の領域を入力して Fig.3(b) の黄枠内のセグメンテーション結果を得るために、U-Net では黒枠内部の部分の情報を利用する。入力の不足している領域は、Fig.3(a) の破線の外側部分のようにミラーリング処理を施して補う。

4. 水増し用フィルタの検証

4.1 検証概要

検証概要として、まず 8 種類の水増し手法を 1 つずつ削除し検証を行う。この検証を行うことでどの水増し手法が精度に影響を与えるか明らかにすることができる。次に精度に最も影響を及ぼす水増し手法とその他の水増し手法の組み合わせによる検証を行う。検証条件は、水増し手法として LUT 変換、メディアンフィルタ、ヒストグラム均一化、ノイズ付加を用いる。また本検証では LUT 変換にはハイコントラスト、ローコントラスト、ノイズ付加にガウシアンノイズ、ゴマ塩ノイズを使用している。Fig.4 に水増し後の画像を示す。また、各水増し手法の特徴を以下に示す。

**LUT 変換** 入力の数値 (RGB 値) に対応して出力すべき別の数値を記載した表を用い画像変換を行う。

**メディアンフィルタ** 画像をぼかし、画素値の変化をなめらかにする。

**ヒストグラム均一化** コントラストが悪く、明るさが偏っている画像の全体的なバランスを調整する。

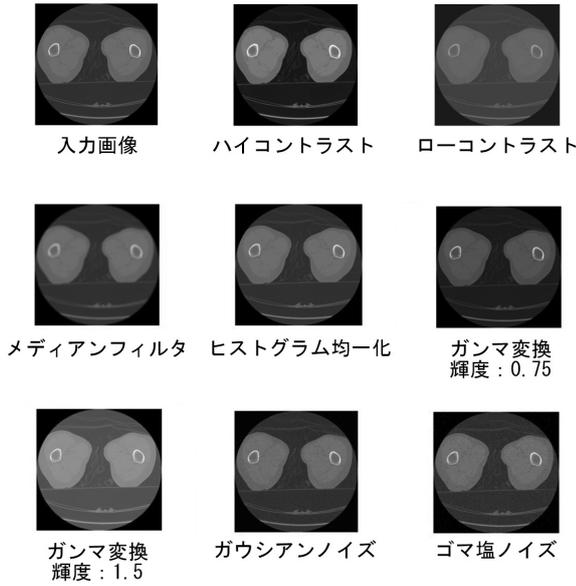


Fig. 4. Comparison of data augmentation.

**ガウシアンノイズ** 各ピクセルごとに現在の輝度を中心に適当な標準偏差で正規分布乱数をふり輝度を変更する.

**ゴマ塩ノイズ** 位置に独立でランダムにある頻度をもって現れるノイズ. あるピクセルでノイズが載るとそのピクセルデータはゼロになるか最大値になる.

#### 4.2 評価指標

既存の手法では背景部分における精度を考慮していないため, 本研究では輪郭の漏れや若干のはみ出しを厳密に評価できる IoU を採用した. はじめにその妥当性を検証する. 検証条件として, 画像サイズ  $128 \times 128$  の CT 画像を 2000 枚用意し, 水増し数を 18000 枚, Epoch 数を 200, バッチ数を 32 とした. 評価指標は, Accuracy と IoU を用いる. 概式をそれぞれ式 (1), 式 (2) に示す.

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \quad (1)$$

$$IoU = \frac{TP}{TP + FP + FN} \quad (2)$$

TP: True Positive    TN: True Negative

FP: False Positive    FN: False Negative

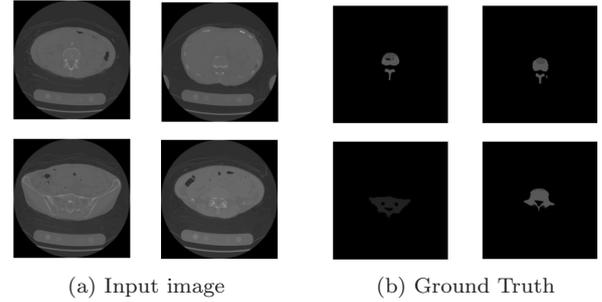


Fig. 5. Input image and Ground Truth.

Table 1. Impact of data augmentation on evaluation indicators.

除外処理	Accuracy	Loss	IoU
ハイコントラスト	0.995	0.011	0.924
ローコントラスト	0.994	0.009	0.938
ガンマ変換 (輝度：0.75)	0.993	0.011	0.921
ガンマ変換 (輝度：1.5)	0.995	0.010	0.919
メディアンフィルタ	0.994	0.012	0.929
ヒストグラム均一化	0.995	0.010	0.920
ガウシアンノイズ	0.991	0.014	0.898
ゴマ塩ノイズ	0.994	0.011	0.919

IoU 算出ではまず, Ground Truth, 出力画像をピクセルごとに比較するためにラベルを設定する. Ground Truth, 出力画像のピクセルごとの変換が終わった後, TP, FP, FN の値をカウントする. TP の場合, Ground Truth, 出力画像のピクセルが一致した場合 TP をカウントする. ただし, IoU の算出では背景部分を考慮しないため, 背景部分を正しく認識した場合カウントしない. FP の場合, Ground Truth, 出力画像のピクセルが不一致のとき, 背景部分を物体として認識した場合 FP をカウントする. FN の場合, 物体がある箇所を背景部分として認識した場合 FN をカウントする.

#### 4.3 入力画像および Ground Truth

本研究で扱う入力画像および Ground Truth を Fig.5 に示す. 入力画像として背骨の CT 画像, Ground Truth としてそれを部位ごとに赤, 緑, 黄色に色分けされた画像を用いる.

#### 4.4 水増しの組み合わせによる精度結果

水増し処理を 1 つずつ削除した場合における Accuracy, Loss, IoU を Table 5.1 に示す. Table 1 では, 8 種類の水増し手法から 1 種類手法を削除した場合の



Fig. 6. Comparison between without padding removal and with gaussian noise removed.

各評価指標のテスト時の精度を示している. Table 1 より, ガウシアンノイズを排除した場合, Accuracy, Loss, IoU ともに精度が最も低く算出された. 結果からガウシアンノイズが水増し処理として有効であることが言える. 次に水増し処理を削除していないとき, およびガウシアンノイズを削除したときの出力画像を Fig.6 に示す. 水増し処理を削除しない場合の出力画像とガウシアンノイズを削除した場合の出力画像を比較すると, 違いが明確に表れガウシアンノイズが精度に影響を与えるということがいえる.

#### 4.5 ガウシアンノイズとその他水増し処理の組み合わせによる精度検証

Table 1, Table 2 においてガウシアンノイズの有無が精度に影響を与えることを示すことができた. 本節ではガウシアンノイズとその他の水増し処理を加えたデータセットを用いて学習することで精度にどのような影響があるのかを検証する. ガウシアンノイズと他の水増し手法 1 種類を用いた場合の精度を Table 2 に示す.

Table 2 より, ガウシアンノイズとメディアンフィルタを施したデータセットを用いた場合, 精度が最も向上した. このことからガウシアンノイズとメディアンフィルタを施したデータセットの組み合わせが学習データにおいて有効である可能性を示すことができた.

次に, ガウシアンノイズとメディアンフィルタの組み合わせによる実際の出力画像を Fig.7 に示す.

Table 2. Verification of accuracy in combination with gaussian noise.

選択処理	Accuracy	Loss	IoU
ハイコントラスト	0.995	0.015	0.935
ローコントラスト	0.994	0.013	0.920
ガンマ変換 (輝度: 0.75)	0.994	0.019	0.900
ガンマ変換 (輝度: 1.5)	0.994	0.013	0.921
メディアンフィルタ	0.996	0.010	0.948
ヒストグラム均一化	0.994	0.012	0.921
ゴマ塩ノイズ	0.994	0.020	0.884

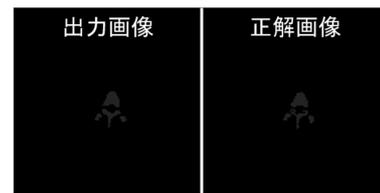


Fig. 7. Output result of the combination of gaussian noise and median filter.

Fig.7 より, 正しく学習できていることがわかる. また, 水増し処理を削除していない場合と精度が変わらないため, 学習データの削減が可能である.

## 5. 結論

本研究では評価指標である Accuracy に加え IoU を追加し, 複数の評価指標での検証, 画像の水増し手法の選定による学習データ削減の実現可能性の検証を行った. その結果, 以下の結論を得た.

1. IoU は Accuracy より厳しい値を取るが, epoch=100 付近で値が安定する.
2. ガウシアンノイズの有無が Accuracy, Loss, IoU に最も影響を与える.
3. ガウシアンノイズとメディアンフィルタを施したデータセットを用いた場合, 最も精度が高い.

以上の結論から, 水増し手法を選ぶことで高い精度を保ちつつ学習データの削減の可能性を示すことができた. 今後の研究方針として, 本研究では画像の水増しのみ観点からコスト削減について検証したが, アルゴリズムの変更や U-Net の構造を変更するなどの観点からコスト削減の実現可能性について検証する.

本研究は JSPS 科研費 19H02032 の助成を受けて実

施した。ここに謝意を表す。

### 参考文献

- 1) Indicators, “O. E. C. D.”, *Health at a Glance* (2019).
- 2) Y. Nakajima, K. Yamada, K. Imamura, and K. Kobayashi, “Radiologist Supply and Workload: International Comparison”, *Radiation medicine*, **26**[8], 455-465 (2008).
- 3) K. K. Kumamaru, A. Machitori, R. Koba, S. Ijichi, Y. Nakajima, and S. Aoki, “Global and Japanese Regional Variations in Radiologist Potential Workload for Computed Tomography and Magnetic Resonance Imaging Examinations”, *Japanese journal of radiology*, **36**[4], 273-281 (2018).
- 4) 名古屋大学医学部附属病院, “腎癌術後フォロー中, 原発性肺癌進行の発見が遅れた事例報道公表用資料”, 2015年12月21日.
- 5) O. Ronneberger, P. Fischer, and T. Brox, “U-net: Convolutional Networks for Biomedical Image Segmentation”, *International Conference on Medical image computing and computer-assisted intervention*, **Springer, Cham**, 234-241 (2015).
- 6) 藤田広志, 原武史, 医療 AI とディープラーニング 医用画像のためのディープラーニング-実践編-, (オーム社, 東京都, 2019), p. 81.
- 7) X. Huang, P. Wang, X. Cheng, D. Zhou, Q. Geng, and R. Yang, “The Apolloscape Open Dataset for Autonomous Driving and its Application”, *IEEE transactions on pattern analysis and machine intelligence*, **42**[10], 2702-2719 (2019).
- 8) J. Long, E. Shelhamer, and T. Darrell, “Fully Convolutional Networks for Semantic Segmentation”, *Proceedings of the IEEE conference on computer vision and pattern recognition*, 3431-3440 (2015).
- 9) V. Badrinarayanan, A. Kendall, and R. Cipolla, “Segnet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation”, *IEEE transactions on pattern analysis and machine intelligence*, **39**[12], 2481-2495 (2017).
- 10) 小松里奈, ゴンサルベス タッド, “U-Net による手書き文字画像内のノイズ除去” 人工知能学会全国大会論文集第 32 回全国大会, 4M101-4M101 (2018).
- 11) 水野雄太, 瀬尾茂人, 渡邊誓旅, 竹中要一, 平松拓郎, 後藤剛, 河田照雄, 松田秀雄, “Deep learning を用いた脂肪組織画像における細胞の認識”, 研究報告数理モデル化と問題解決 (MPS), **10**, 1-6 (2016).