

Deep Residual 3D U-Net と 3DCNN による CT 画像セグメンテーションの改善

Improvement of CT image segmentation by Deep Residual 3D U-Nets and 3D-CNNs

二宮 啓太 ^{*1}
Keita Ninomiya

古山 良延 ^{*2}
Yoshinobu Furuyama

太田 丞二 ^{*2}
Joji Ota

須鎗 弘樹 ^{*3}
Hiroki Suyari

^{*1}千葉大学 工学部 情報画像学科

Department of Informatics and Imaging Systems, Faculty of Engineering, Chiba University

^{*2}千葉大学医学部附属病院 放射線部
Department of Radiology, Chiba University Hospital

^{*3}千葉大学 大学院工学研究院
Graduate School of Engineering, Chiba University

Segmentation of medical images with high precision and speed is an important task in many medical scenes. One such method for this task is GraphCut based on energy minimization problem. However, in GraphCut, it is difficult to perform segmentation completely and automatically if adjacent pixel values are similar. There are many methods for this problem, but most of them are not suitable in speed. In deep learning methods, automatic segmentation is possible because of its capability of capturing complicated features. In this research, we propose a model incorporating 3D U-Net extended with Residual Unit and 3DCNN for correcting segmentation results.

1. はじめに

医療現場において高精度で高速な医療画像セグメンテーションは重要な課題である。特に一刻を争う救急医療においてその需要が高い。現在はグラフカットなどの人の手を借りた作業が一般的であり、これらの手法では多くの時間を必要とするのが現状である。時間短縮のために、最近では深層学習の Fully Convolutional Networks[Jonathan 15] や U-Net[Olaf 15] による手法が多く研究されている。しかし、これらのモデルでは臓器の種類によっては、セグメンテーションが困難なことがある。その原因として、臓器が高い可変性を有していることが挙げられる。特に本研究で対象とした脾臓は、臓器のセグメンテーションの中でも困難な課題として知られている。本研究では脾臓のセグメンテーションの精度向上のために、Residual Unit を組み込んだ 3D U-Net と、セグメンテーション結果の訂正を行う 3DCNN を用いたモデルを提案する。

2. 従来モデル

2.1 3D U-Net

[O. Cicek 16] では U-Net を三次元に拡張した 3D U-Net が提案されている。3D U-Net は U-Net の畳み込み層を三次元の畳み込み層に置き換えたもので、三次元の特徴量を捉えることができるネットワークである。二次元の U-Net で三次元画像を処理するには、二次元スライス画像にしなければならない。この際に、スライスした方向の連続性を情報量として失っており、U-Net が三方向の全ての軸の連続性を利用するには、U-Net を三つ用意する等の工夫が必要である。それに対して、3D U-Net は入力に三次元画像が使用可能で、複数のネットワークを準備する必要がない。この点で U-Net より優れており、セグメンテーションの精度も二次元の U-Net と比較して向上したことが報告されている [O. Cicek 16]。

2.2 Deep Residual U-Net

[Zhengxin Zhang 17] では U-Net に Residual Unit を組み込んだ Deep Residual U-Net を提案している。Residual Unit とは [Kaiming He 15] に用いられた手法で、ネットワークを深くした際に起こる学習が進まない現象を緩和する働きを持つ。その構造は単純で図 1 のように出力に対して入力を加算するだけである。[Zhengxin Zhang 17] では、他の手法と比較して高い精度を出したことが報告されている。

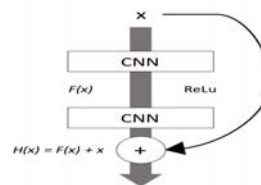


図 1: Residual Unit の構造の一例

3. 提案手法

3.1 提案モデル

本論文では 2.1 節に Residual Unit を組み込んだモデルと、3DCNN の二つを組み合わせたモデルを提案する。図 2 は提案ネットワークの概略図である。Deep Residual 3D U-Net と 3DCNN の詳細は 3.2 節と 3.3 節にて説明する。

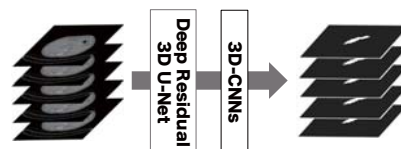


図 2: 提案モデル

3.2 Deep Residual 3D U-Net

図 3 に Residual Unit を組み込み、より層を深くした 3D U-Net を示す。Encoder 層及び Decoder 層の詳細は図 4 と図 6 で

連絡先: 二宮啓太, 千葉大学 工学部 情報画像学科, 263-8522 千葉市稲毛区弥生町 1-33, email: nino-keita1727@chiba-u.jp

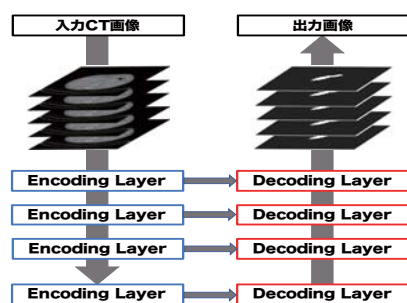


図 3: 提案する Deep Residual 3D U-Net

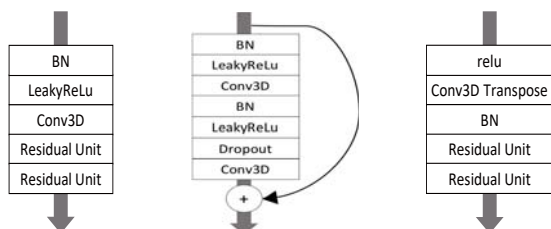


図 4: Encoder Layer 図 5: Residual Unit 図 6: Decoder Layer

ある。二次元の Deep Residual U-Net は 2.2 節で紹介したが、提案する Deep Residual 3D U-Net は更に Residual Unit を加えることで層を深くした。本モデルでは畳み込み層のフィルタサイズを $3 \times 3 \times 3$, Conv3DTranspose 層は $2 \times 2 \times 2$ とした。

3.3 3DCNN

3DCNN は三次元の畳み込み層を三つ重ねたものとした。フィルタサイズは初めの二層を $4 \times 4 \times 4$, 三層目を $1 \times 1 \times 1$ とした。3DCNN では入力に学習済みモデルによるセグメンテーション結果と CT 画像を入力として与える。本モデルを提案した目的は学習済みの手法に適応することでもとのセグメンテーション結果を高精度化する役割を期待している。

4. 評価実験

4.1 データ

データセットは [Holger R. Roth 16] の Pancreas-CT を使用した。このデータセットには 82 人分の CT 画像があり、学習毎に学習データとして 50 人分、学習中の検証データとして 10 人分、評価用のテストデータとして 20 人分をランダムに選択する。この CT 画像は体の横断面を写した二次元画像であるが、それを層のように積み重ねることによって、三次元画像として捉えることができる。本実験では、膵臓が写った二次元 CT 画像を層のように重ねることで三次元画像とし、この三次元画像より膵臓を含む直方体を抜き出し、縦と横を 64 にリサイズを行なったものをデータとして用いた。また、CT 値は水を 0, 空気は -1000 とした値であり、組織ごとにおおよそその値が決まっている。本実験では 0~300 を膵臓のセグメンテーションを行うのに重要なデータと考え、範囲外のは値を 0 とした。その後、-1~1 の範囲に正規化を行い入力画像とした。CT 画像の一枚を図 7 に示す。

4.2 損失関数

損失関数にはセグメンテーション問題において、一般的に利用されるダイス係数を 1 から減算したものを用いた。ダイス係数を式 (1) に、損失関数を式 (2) に示す。ダイス係数は画像の類似度を計算できる関数であり、0~1 の範囲を持ち、1 に近

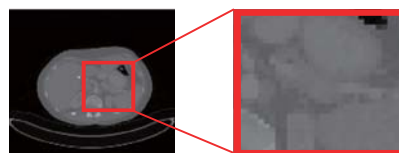


図 7: 入力三次元 CT 画像のうち、横断面 CT 画像の一枚



図 8: データセットの例

いほど正解に近いことを示す。A は教師データ、B はモデルによる予測結果とする。

$$\text{ダイス係数} = \frac{2 \times |A \cdot B|}{|A| + |B|} \quad (1)$$

$$\text{損失関数} = 1 - \text{ダイス係数} \quad (2)$$

4.3 実験環境

今回の実験で筆者が使用した環境を表 1 に示す。

表 1: 実験環境

OS	CentOS7
メモリ	60GB
プロセッサ	Intel(R) core(TM) i7-5820K CPU @ 3.30GHz
GPU	NVIDIA GeForce GTX TITAN X 12GB

4.4 評価方法

提案モデルによる出力結果は入力と同じサイズのボリュームデータであり、各ボクセルには膵臓であるかの確率が格納されている。今回は 2 クラス分類問題であるため、50%以上である場合は膵臓、50%未満である場合は膵臓でないとする。予測結果の評価には、式 (1) に示すダイス係数を用いた。

4.5 実験結果・考察

実験結果を表 2 に示す。表 2 は各手法 5 回ずつ実験を行なった結果である。これより、ダイス係数の平均、最小、最大の全てにおいて、提案手法である Deep Residual 3D U-Net + 3DCNN が最も高い値を記録していることがわかる。図 9 は各手法の平均をグラフ化したものである。このグラフを確認すると、図 4 と図 6 のような Encoding Layer, Decoding Layer を実装した図 3 の Deep Residual 3D U-Net は、平均で 5.2% の精度向上を果たしている。更に図 2 のように 3DCNN を通すことで、1.0% の精度向上が見られた。

次に 3D U-Net と Deep Residual 3D U-Net + 3DCNN の精度のばらつきに関して確認する。図 10 は二つの手法をそれぞれ 5 回ずつの実験を行い、その結果を箱ひげ図にしたものである。比較を行うと、Deep Residual 3D U-Net + 3DCNN は精度の向上だけでなく、ばらつきが小さくなっていることがわかる。

最後に実行速度について確認する。救急医療の現場で用いられること等を想定する以上、実用的な実行時間でなければな

表 2: 3D U-Net と提案手法の比較

ダイス係数	3D U-Net	Deep Residual 3D U-Net(提案)	Deep Residual 3D U-Net + 3DCNN(提案)
平均	0.688	0.740	0.750
最小	0.336	0.325	0.361
最大	0.855	0.879	0.882

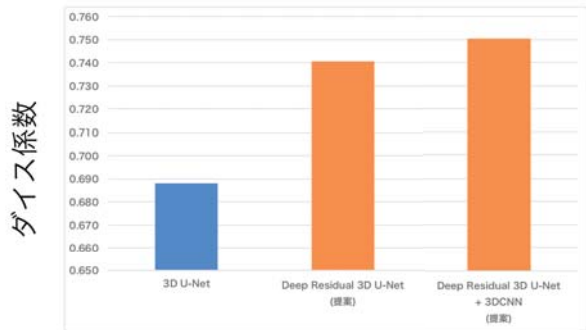


図 9: 各手法の平均ダイス係数の比較

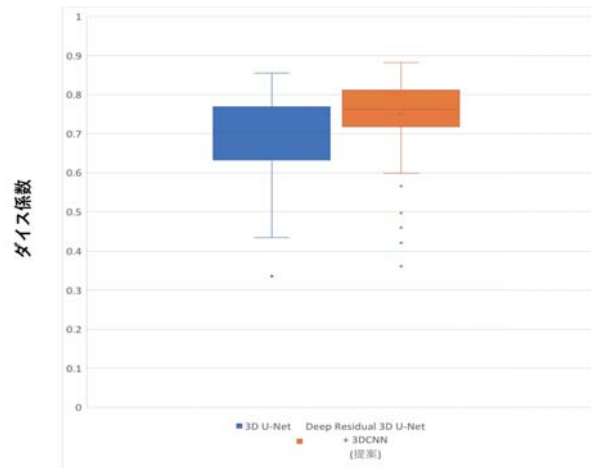


図 10: U-Net と Deep Residual 3D U-Net + 3DCNN の箱ひげ図

らない。従来手法及び、提案手法の実行時間を表 3 に示す。この時の実行環境は表 1 と同じである。比較を行うと、従来手法より実行時間は長くなっているが、0.38s と実用上問題のない実行時間となった。

表 3: 実行時間

	3D U-Net	Deep Residual 3D U-Net + 3DCNN(提案)
1 人あたり (s)	0.18	0.38

5. おわりに

本研究では 3D U-Net に Residual Unit を組み込んだ Deep Residual 3D U-Net と、セグメンテーション結果をより改善するように学習する 3DCNN の二つを組み込んだモデルを提案した。どちらも膵臓のセグメンテーションの精度向上に繋がった。しかし、欠点が存在する。例えば、今回は膵臓が写った二次元 CT 画像のみを利用して、三次元に再構築した三次元画像をデータセットとして利用したが、これは GPU の主記憶メモリが不足していたためである。より広い範囲からのセグメン

テーションを行うには、この問題を解決するためにアルゴリズムの改良を検討しなければならない。

また、本実験では対象を膵臓に限定したが、実際の医療現場では他の臓器も対象である。更に対象が一つとは限らず、同時にいくつかの臓器のセグメンテーションを要求されることがある。今回提案したネットワークがこれらのことに対応できるかを今後検証する必要がある。

参考文献

[Jonathan 15] Jonathan Long, Evan Shelhamer, and Trevor Darrell: Fully Convolutional Networks for Semantic Segmentation, CVPR 2015 pp 3431-3440

[Olaf 15] Olaf Ronneberger, Philipp Fischer, and Thomas Brox: U-Net: Convolutional Networks for Biomedical Image Segmentation, Medical Image Computing and Computer-Assisted Intervention - MICCAI 2015 pp 234-241

[O. Cicek 16] O. Cicek , A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger. 3d u-net: Learning dense volumetric segmentation from sparse annotation, arXiv preprint arXiv:1606.06650, 2016

[Zhengxin Zhang 17] Zhengxin Zhang, Qingjie Liu, et la.: Road Extraction by Deep Residual U-Net, arXiv:1711.10684v1 [cs.CV] 29 Nov 2017

[Kaiming He 15] Xiangyu Xhang, Shaoqing Ren, Jian Sun: Deep Residual Learning for Image Recognition arXiv:1521.03385 10 Dec 2015

[Holger R. Roth 16] Holger R. Roth, Amal Farag, Evrim B. Turkbey, Le Lu, Jiamin Liu, and Ronald M. Summers. (2016): Data From Pancreas-CT. The Cancer Imaging Archive. <http://doi.org/10.7937/K9/TCIA.2016.tNB1kqBU>

[Roth HR 15] Roth HR, Lu L, Farag A, Shin H-C, Liu J, Turkbey EB, Summers RM. DeepOrgan: Multi-level Deep Convolutional Networks for Automated Pancreas Segmentation. N. Navab et al. (Eds.): MICCAI 2015, Part I, LNCS 9349, pp. 556 - 564, 2015.