

## 心臓核医学における心筋抽出のための U-Net の拡張に関する一検討 A Study on Extensions of U-Net for Myocardial Extraction in Nuclear Cardiology

和田 直史<sup>†</sup> 菊池 明泰<sup>‡</sup> 鈴木 昭弘<sup>†</sup> 川上 敬<sup>†</sup>  
Naofumi Wada Akihiro Kikuchi Akihiro Suzuki Takashi Kawakami

### 1. はじめに

本研究は、心臓核医学検査における核医学画像から心筋領域を自動で抽出し、診療放射線技師の読影支援および医師の診断支援を行うことを目的としている。医用画像における臓器領域の抽出については既に多くの研究がなされており、近年は深層学習を用いて画素単位の領域抽出が可能となっている。しかしながら、これまでの研究の多くは CT や MRI を対象としたものであり、核医学を対象とした研究は未だ多くない。核医学画像は、体内に投与した放射性同位元素を含む薬剤が臓器に集積した様子を、体外からガンマカメラを使って撮影することで得られる。このことから、解剖学的構造を画像化する CT や MRI とは異なり、機能的情報を視覚化したものであると言える。また、CT や MRI に比べて空間解像度が低く、ノイズを多く含むことが特徴である。

本稿では、文献 [1] と同様、Semantic Segmentation の一手法である U-Net を用いて核医学画像から心筋領域を抽出する。概念図を図 1 に示す。さらに、U-Net にさまざまな拡張手法を導入した場合の抽出精度および計算コストを実験により評価し、有効性を検証する。

### 2. U-Net の拡張

2015 年に Ronneberger らによって提案された U-Net [3] は、FCN [2] と同様に畳込み層で構成され、Skip connection を用いた Encoder-Decoder 構造を持つ。シンプルな構造であるにもかかわらず、画素単位の高精度な Segmentation が可能であることから、主に医用画像の分野で注目を集めた。また、CT や MRI では 3 次元ボリュームデータが得られることから、3 次元畳込みを用いて U-Net を拡張した V-Net [4] や 3D U-Net [5] が提案された。3 次元化は高い精度が得られる一方で、膨大な計算コストが問題となっている。2 次元の U-Net においても様々な拡張モデルが提案されている。例として、残差ネットワーク構造を導入した Residual U-Net (以下、RU-Net) [7] や、Recurrent 構造を導入した R2U-Net [8]、Attention 構造を導入した Attention U-Net (以下、AU-Net) [9] などがある。近年では、深さ (レイヤー数) の異なる U-Net をアンサンブルさせた構造を持つ U-Net++ [10, 11] が提案され、種々の医用画像で精度の向上を実現している。これら U-Net を拡張したモデルは、いずれもネットワーク構造を複雑化させ、表現能力を向上させることで高精度化を図っている。

一方で、モデルを複雑化させずに高精度化する手法も提案されている。文献 [6, 12] では、畳込みを 3 次元化するのではなく、図 2(a) のように、Axial, Sagittal, Coronal の 3 つの断面それぞれで 2 次元画像として処理した結果を多数決で統合する “2.5 次元化” を行っている。また、複

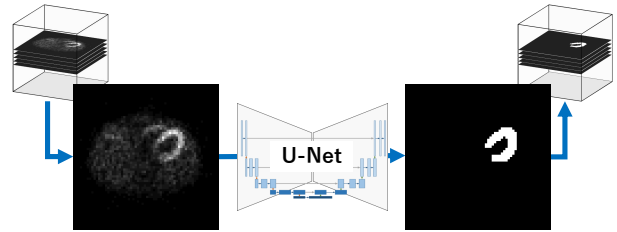


図 1 U-Net を用いた核医学画像における心筋抽出

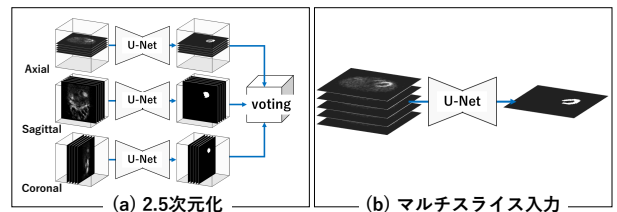


図 2 2.5 次元化とマルチスライス入力

数スライスを複数 channel としてネットワークに入力する “マルチスライス入力” も提案されている [12, 13]。マルチスライス入力は、スライス間の関係を考慮しつつ、畳込みは 2 次元で行うため、計算量の増加はほとんどない。文献 [12, 13] では、複数スライスを入力して複数スライスを推定しているが、本研究では図 2(b) のように複数スライスを入力して 1 スライスを推定する方式を用いる。

### 3. Semantic Segmentation の損失関数

Semantic Segmentation で使用する損失関数には様々なものがある [14]、ここでは、Binary Cross Entropy (BCE) 損失  $\mathcal{L}_{bce}$ 、Dice 損失  $\mathcal{L}_{dice}$ 、BCE-Dice 損失  $\mathcal{L}_{bdice}$ 、Tversky 損失  $\mathcal{L}_{tv}$ 、Focal Tversky 損失  $\mathcal{L}_{ftv}$  の 5 種類を用いる。各損失関数は次式 (1)-(5) で表される。

$$\mathcal{L}_{bce} = \sum \{-y \log \hat{y} - (1-y) \log (1-\hat{y})\} \quad (1)$$

$$\mathcal{L}_{dice} = 1 - \frac{2 \sum y \hat{y}}{\sum y + \sum \hat{y}} \quad (2)$$

$$\mathcal{L}_{bdice} = 0.5 \mathcal{L}_{bce} + \mathcal{L}_{dice} \quad (3)$$

$$\mathcal{L}_{tv} = 1 - \frac{\sum y \hat{y}}{\sum y \hat{y} + \alpha \sum (1-y) \hat{y} + (1-\alpha) \sum y (1-\hat{y})} \quad (4)$$

$$\mathcal{L}_{ftv} = \mathcal{L}_{tv}^{\frac{1}{\gamma}} \quad (5)$$

ここで、 $\Sigma$  は画素の総和、 $y$  は正解値、 $\hat{y}$  は推定値である。また、 $\alpha, \gamma$  は係数であり、 $\alpha = 0.3, \gamma = 0.75$  とした。

### 4. 評価実験

#### 4.1. 実験条件

実験には実際に病院で撮影された臨床データを使用し、1 患者あたり 2 つの状態 で撮影した 367 名分のデータか

<sup>†</sup>北海道科学大学 工学部 情報工学科

<sup>‡</sup>北海道科学大学 保健医療学部 診療放射線学科

ら、無効なデータを除いた694個体を使用した。1つの個体は64×64画素のAxial断面画像64枚で構成されたDICOMデータである。学習、検証、テストのデータは、それぞれ553個体(35,392枚)、71個体(4,544枚)、70個体(4,480枚)とし、テストデータに学習時と同じ患者が含まれないよう分割した。教師データとなるラベル画像は、診療放射線技師の資格を持つ専門家1名が44,416枚(694個体×64枚)の画像全てを目視で判断し、手作業で心臓領域をマークした。評価に使用するモデルは、U-Net [3], RU-Net [7], R2U-Net [8], AU-Net [9], U-Net++ [10]とし、U-Netはレイヤー数1~4( $L_1 \sim L_4$ )の4パターンを用いた。U-Net以外のモデルは、レイヤー数を4とした。学習時のパラメータは、エポック数100、バッチサイズ32とし、オプティマイザーにはAdamを使用した。また、学習率の初期値を $1.0 \times 10^{-3}$ とし、validation lossが減少しない場合は学習率を減衰させていくことで学習を安定化させた。実験に使用したPCのスペックは、CPU: Intel Core i9-9820X, メモリ: 64GB, GPU: RTX2080Ti × 2枚である。

#### 4.2. 評価値

本実験では抽出精度を測る評価値としてDice係数を用いる。正しく心筋を検出した画素数をTP(True Positive)、誤検出の画素数をFP(False Positive)、未検出の画素数をFN(False Negative)とすると、Dice係数は次式で表される。

$$Dice = \frac{TP}{TP + \frac{1}{2}(FP + FN)} \quad (6)$$

本実験では、テストデータ70個体に対する1個体あたりの平均Dice係数を評価値として用いる。また、スライスごとのDice係数、Precision, Recallを算出し、それらの値が閾値0.7を下回る場合を、それぞれ誤って判定したスライス(False slice), 誤検出スライス(FP slice), 未検出スライス(FN slice)と定義する。

#### 4.3. 実験結果

図3は、損失関数、マルチスライス入力、2.5次元化についてそれぞれDice係数を比較したグラフである。モデルは全て4層のU-Net( $L_4$ )を使用した。図3(a)(b)より、損失関数はBCE-Dice、入力スライス数は7の場合に、それぞれ最も高い値を示した。図3(c)では2.5次元化により精度向上が見られたが、3断面それぞれに対してモデルが必要となる上、計算量が3倍になるため、コスト面が問題となる。表1は、1個体あたりの平均Dice係数、1個体あたりのFalse slice・FP slice・FN sliceの平均枚数、学習済みモデルのファイルサイズ(Size)、1個体あたりの平均推定時間(Time)をモデル別に示している。図3の結果に基づき、損失関数はBCE-Dice、入力スライスは7とし、2.5次元化はなしとした。表1より、総合的にはU-Net++が最も精度が高いと考えられるが、1層のU-Net( $L_1$ )でも十分高いDice係数が得られており、多層化やモデルの複雑化による精度向上はほとんど見られなかった。これは、本実験で対象とした核医学画像が64×64画素の小さい画像であり、且つ心臓の位置や形状に大きな個人差がないため、必ずしも表現能力の高い複雑なモデルは必要なく、マルチスライス入力により計算コストを抑えながら精度を向上させることが可能であることを示している。

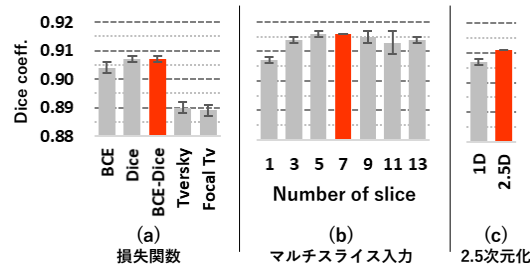


図3 モデルの複雑化がない手法の比較結果 (Model: U-Net ( $L_4$ ))

表1 性能比較 (損失関数: BCE-Dice, 入力スライス数: 7, 2.5次元化: なし)

Model	Dice coeff.	False slice	FP slice	FN slice	Size (MB)	Time (ms)
U-Net ( $L_1$ )	<b>0.917</b> ±0.021	1.277	0.920	0.651	<b>4.8</b>	<b>174.8</b>
U-Net ( $L_2$ )	0.915 ±0.020	1.020	0.726	0.611	21.8	208.2
U-Net ( $L_3$ )	0.912 ±0.022	0.989	0.497	0.751	89.4	241.3
U-Net ( $L_4$ )	0.916 ±0.020	0.886	0.551	0.611	359.7	357.0
RU-Net	<b>0.917</b> ±0.021	0.857	0.540	0.666	596.0	498.6
R2U-Net	0.914 ±0.022	0.969	0.663	0.603	523.7	528.3
AU-Net	0.916 ±0.021	0.969	0.637	<b>0.574</b>	365.9	419.5
U-Net++	0.916 ±0.019	<b>0.846</b>	<b>0.454</b>	0.711	405.5	469.5

#### 5. おわりに

本研究では、核医学画像からの心筋抽出において、U-Netをベースとした複数のモデルおよび拡張技術について比較評価を行った。今後の課題として、心筋抽出後の軸設定の自動化や心外集積除去などへの活用を検討する。

#### 謝辞

実験データをご提供頂いた、金沢大学大学院先進予防医学研究科中嶋憲一先生および金沢大学附属病院米山寛人先生に感謝の意を表す。また、本研究の一部は、JSPS科研費JP21K12629の助成を受けたものである。

#### 参考文献

- [1]岡田 他: “U-Netを用いた心臓核医学画像における領域検出に関する研究,” 計測自動制御学会SI部門講演会, 2019.
- [2]J. Long et al.: “Fully Convolutional Networks for Semantic Segmentation,” In *CVPR*, 2015.
- [3]O. Ronneberger et al.: “U-Net: Convolutional Networks for Biomedical Image Segmentation,” In *MICCAI*, 2015.
- [4]F. Milletari et al.: “V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation,” In *3DV*, 2016.
- [5]Ö. Çiçek et al.: “3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation,” In *MICCAI*, 2016.
- [6]X. Zhou et al.: “Deep Learning of the Sectional Appearances of 3D CT Images for Anatomical Structure Segmentation based on An FCN Voting Method,” *Medical Physics*, 44, **10**, 2017.
- [7]Z. Zhang et al.: “Road Extraction by Deep Residual U-Net,” *IEEE Geoscience and Remote Sensing Letters*, 15, **5**, pp.749-753, 2018.
- [8]M. Z. Alom et al.: “Recurrent Residual Convolutional Neural Network based on U-Net (R2U-Net) for Medical Image Segmentation,” arXiv:1802.06955, 2018.
- [9]O. Oktay et al.: “Attention U-Net: Learning Where to Look for the Pancreas,” In *MIDL*, 2018.
- [10]Z. Zhou et al.: “Unet++: A Nested U-net Architecture for Medical Image Segmentation,” In *DLMI, ML-CDS*, 2018.
- [11]Z. Zhou et al.: “UNet++: Redesigning Skip Connections to Exploit Multiscale Features in Image Segmentation,” *IEEE Medical Imaging*, 39, **12**, pp. 1856-1867, 2020.
- [12]Q. Yu et al.: “Thickened 2D Networks for Efficient 3D Medical Image Segmentation,” arXiv:1904.01150, 2019.
- [13]後藤 他: “CT画像における複数スライスを用いた小疾病領域検出,” *精密工学会誌*, 86, **12**, pp.1020-1025, 2020.
- [14]S. Jodon: “A Survey of Loss Functions for Semantic Segmentation,” In *CIBCB*, 2020.