

修士論文

深層学習を用いた 3 次元 PET 画像からの
腎臓および大動脈のセグメンテーション

里川 航亮

奈良先端科学技術大学院大学

先端科学技術研究科

情報理工学プログラム

主指導教員：金谷 重彦 教授

計算システムズ生物学 研究室（情報科学領域）

2023 年 3 月 7 日提出

本論文は奈良先端科学技術大学院大学先端科学技術研究科に
修士(工学)授与の要件として提出した修士論文である。

里川 航亮

審査委員：

主査	金谷 重彦	(情報科学領域 教授)
	松本 健一	(情報科学領域 教授)
	小野 直亮	(情報科学領域 准教授)
	MD.Altaf-Ul-Amin	(情報科学領域 准教授)
	黄 銘	(情報科学領域 助教授)

深層学習を用いた 3 次元 PET 画像からの 腎臓および大動脈のセグメンテーション*

里川 航亮

内容梗概

機械学習の発展に伴い、医師不足の問題を解決するためにこれまで多くの自動診断モデルが提案されてきた。医療画像の一種である PET 画像は、主のがん細胞の有無や血流の可視化のために用いられるが、CT や MRI などの医療画像と比較して使用率の低さなどから研究対象として積極的に扱われず、PET 画像に関する先行研究の多くは脳や心筋が対象となっており、腹部臓器を扱う際はその複雑性から CT 画像と同時に扱われることが多い。しかし、世界中に多くの患者がいるとされる糖尿病や腎不全の検査では PET 画像を用いて腎臓を流れる血流を可視化し、より精密な診断を行う場合がある。その際、医師たちは 3 次元時系列 PET 画像から腎臓部分を手動で抜き出し血流に関する評価を行うため、手動での作業は彼らにとって時間のかかる負担の大きい作業となっているといった問題が挙げられる。この問題に対処するために、本研究では畳み込みニューラルネットワークモデルを用いて特定領域の抽出を試みる。具体的には、深層学習モデルである U²-Net と U-Net の 2 種類の畳み込みモデルを用いて 3 次元 PET 画像から腎臓と大動脈のセグメンテーションを行った。その後、それぞれのモデルの有効性を検証した結果、PET 画像のみからでも血流推定に有効な腎臓および大動脈の自動抽出が可能であることが確認された。

キーワード

腎臓, 畳み込みニューラルネットワーク, セグメンテーション, データ拡張, PET 画像

*奈良先端科学技術大学院大学 先端科学技術研究科 修士論文, 2023 年 3 月 7 日.

Kidneys and aorta segmentation from 3D PET images using deep learning*

Kosuke Rikawa

Abstract

Due to development of deep learning, many automated diagnosis models have been proposed in order to resolve the problem of lacking doctors. PET image, one of medical images, is mainly used for detecting cancer cells or visualizing blood flow, but it is not treated actively as a research target because of low usage rate compared with medical images such as CT and MRI. Also most of previous studies related with PET images are targeting on the brain and myocardium, furthermore when dealing with abdominal organs, they are often used with CT because of their complexity. However, PET images may be used to visualize blood flows in kidneys for detail diagnosis of diabetes and kidney failure, which are known as a worldwide disease. In the process, physicians manually extract the kidney area from 3D time-series PET images and evaluate the blood flow, hence it is time consuming and burden task for them. To address this problem, we attempt to extract specific regions using a convolutional neural network in this study. In detail, we used two types of deep learning models, U²-Net and U-Net, to segment kidneys and aorta from 3D PET images. The validity of each model was verified, and then it was confirmed that automatic extraction of kidneys and aorta for effective blood flow estimation was possible even from PET images only.

Keywords:

Kidney, CNN, Segmentation, Data Augmentation, PET image

*Master's Thesis, Graduate School of Science and Technology, Nara Institute of Science and Technology, March 7, 2023.

目次

図目次	v
表目次	vi
第 1 章 序論	1
1.1 研究の背景	1
1.2 深層学習と医療	1
1.3 PET 画像	2
1.4 血流推定への PET 画像の利用	2
1.5 腎臓と PET 画像	4
1.5.1 腎臓のはたらき	4
1.5.2 糖尿病と腎不全	4
1.5.3 糖尿病や腎不全の診断方法	4
1.6 PET 画像に関する現状と課題	5
1.7 本研究の目的	6
1.8 本研究の取り組み	6
第 2 章 関連研究	7
第 3 章 データセット	9
3.1 概要	9
3.2 正解画像の作成	10
第 4 章 PET 画像からのセグメンテーション	11
4.1 提案手法	11
4.1.1 U ² -Net	11
4.1.2 U-Net	14
4.2 実験	15
4.2.1 Data Augmentation	16
4.2.2 学習の詳細条件	17

第 5 章	結果	19
5.1	評価指標	19
5.2	セグメンテーション結果	20
第 6 章	考察	23
6.1	U ² -Net と U-Net の精度比較	23
6.2	連続値マスク適用の検討	24
6.3	血流推定へのセグメンテーション領域の適用	28
第 7 章	結論	30
	謝辞	31
	参考文献	32

図目次

1.1	PET 画像の例	3
1.2	時間と血流の関係	5
3.1	時系列データからのデータセットの作成	10
3.2	本研究で使用するデータセットの一部	10
4.1	ReSidual U-Block の詳細 (a) と ReSidual U-Block 接続 (b) およ び Residual 接続 (c)	12
4.2	本研究で使用する U ² -Net の構成	13
4.3	本研究で使用する U-Net の構成	15
4.4	実験の概要図	16
4.5	本研究で適用した Data Augmentation 手法	17
5.1	Dice の説明図	20
5.2	1 人目のセグメンテーション結果	21
5.3	2 人目のセグメンテーション結果	22
6.1	予測画像の 3 次元立体表示	24
6.2	正確に大動脈のセグメンテーションを行えなかった例	25
6.3	<補正前> 大動脈の半径 r のスライス毎の変化	26
6.4	<補正後> 大動脈の半径 R のスライス毎の変化	27
6.5	U ² -Net と U-Net から得られた領域を用いた血流推定	28

表目次

3.1	各データの詳細情報	9
4.1	各 RSU ブロックの詳細	14
4.2	ネットワークの学習条件	18
5.1	実験結果	20

第 1 章 序論

1.1 研究の背景

現在、日本をはじめとする多くの先進国の間で少子高齢化が進んでおり、医療・ヘルスケアに対する需要は、今後ますます大きくなると考えられる。EU 諸国を中心とした 38 の先進国で構成された OECD(経済協力開発機構)の統計によると、人口 1000 人あたりの医師数は日本では 2.6 人と少なく [1]、医師の不足が大きな課題となっている。また、X 線画像や CT などの医療画像を用いた診断には専門的な知見が必要とされるが、専門知識を有した放射線診断医の不足や高齢化なども問題視されており [2]、今後この動きはさらに加速すると考えられている。

これらの問題を解決するために、近年では深層学習を画像解析に用いることで、診断の自動化の動きが高まっている。このような技術の 1 つに、AI を用いて医師の医療画像診断を支援する自動診断技術があり、乳がん、肺結節、動脈瘤など様々な疾患の診断に活用されている [3]。特に、CT や MRI などの医療画像撮影装置の発展・普及により、高解像度な医療画像の取得が容易になったため、画像内の特定領域を抽出する研究が進展しており [4, 5]、2012 年に Alex らの研究チームが開発したモデルが Image Net で優勝して以降 [6]、Deep Neural Network を用いた手法の中でも、畳み込みニューラルネットワーク (CNN : Convolutional Neural Network) が注目を集めるようになり、医療画像解析の分野でも広く活用されている。

1.2 深層学習と医療

深層学習の発展に伴い、医療分野でもこれらの技術の導入が進められ、遺伝子情報の解析、新薬の開発、AI ロボットによる手術支援などその活躍は多岐にわたる。特に、画像診断への応用は活発であり、これまで多くの医療画像や疾患が研究の対象となってきた。Abhir らは深層学習を X 線画像に適用し、肺炎画像と正常画像の高い精度での自動分類に成功した [7]。また、Naser らは脳の MRI 画像から腫瘍を自動抽出するモデルの提案を行っている [8]。このように現在は深層学習によって、より正確な診断の実現や医師不足の問題解決といったことが期待されている。

1.3 PET 画像

医療画像の 1 つに PET 画像がある (図 1.1)。PET とは、Positron Emission Tomography(陽電子放出断層撮影) の略称であり、陽電子を放出する放射性薬剤を被検体に注入し、体内の電子と対消滅した時に生じる γ 線を計測することで、体内の分子や細胞の機能を空間的・時間的に画像化したものである [9]。放射性薬剤などの放射性物質が体内に注入されてから半分の量になるまでに要する時間を半減期といい、投与する薬剤の種類によってその長さは異なる。半減期は短い方が被検体の放射性被ばくは少なくなるが、一方であまりにも短すぎると検査を行うことが困難になる。近年の PET 検査とは主に FDG-PET 検査を指し、がん細胞の有無の検査に用いられる。FDG(^{18}F -fluorodeoxyglucose) とは放射性フッ素を付加したブドウ糖のことで、半減期は 109 分である [10]。FDG はブドウ糖と同様に体組織によって摂取されるが、すぐには代謝されないので組織内に長く停滞する。この性質を利用し、がん細胞発見のための PET 検査では、まず FDG の注射が行われる。一定の時間が経過すると、がん細胞が正常な細胞よりも多量の FDG を吸収する。そして、がん細胞に集積した FDG が発する γ 線を PET 検査で撮影することによって、体内のがん細胞に取り込まれたブドウ糖の分布を可視化することができる。

また、PET 画像は CT と比較される場合が多い。PET 画像は代謝などの機能に関する観点から疾患を診断するのに対し、CT は臓器や骨格の形状に関する観点から疾患を診断する。また、CT 検査では胸部・腹部など部位を限定して撮影することが多いが、PET 検査では全身を一度に撮影することが可能である。使用する放射性薬剤の種類に依存するが、1 回の撮影による放射線被ばく量は、CT 検査よりも PET 検査の方が少ないことが報告されている [11]。現在では、PET 検査と CT 検査を組み合わせた PET-CT 検査が一般的であり、一度の検査で両方の画像を重ね合わせて表示することで、診断精度の向上が行われている。

1.4 血流推定への PET 画像の利用

先述したように、PET 画像は主のがん細胞の有無を診断するために用いられることが多いが、中には生体内を流れる血流の測定や可視化に用いられる場合がある。 H_2^{15}O は主に血流の測定に用いられる放射性薬剤で、この薬剤が血管中を流れる際

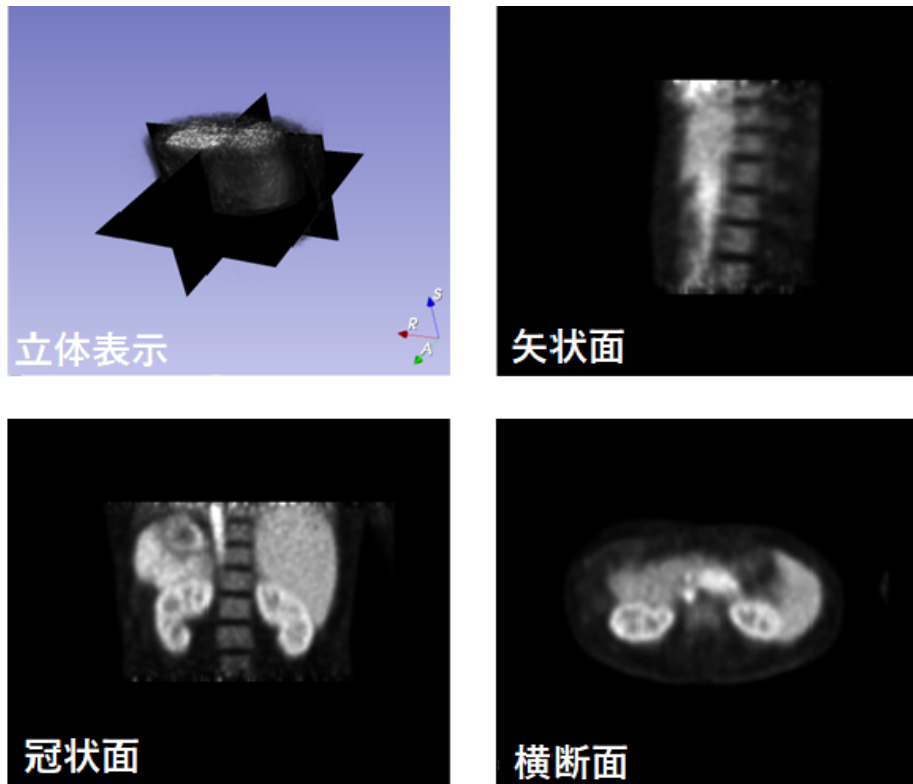


図 1.1 PET 画像の例

に発する陽電子を PET 装置で撮影し、 $H_2^{15}O$ による輝度の時間変化を解析することによって、血流を定量的に評価することができる。Bruehlmeier らは、低酸素状態が悪性腫瘍の治療にどのような影響を与えるかを調べるために、放射性薬剤である ^{18}F MISO と $H_2^{15}O$ を用いて、脳腫瘍患者の脳内における血流の測定・可視化を行い、PET 画像で脳腫瘍の低酸素状態を空間的に表示できること、そして腫瘍の大きさと低酸素状態の間に相関が見られないことを示した [12]。また、PET 画像を用いて心筋血流の定量的な測定を行った研究がある [13]。この研究では、 $^{13}NH_3$ と $H_2^{15}O$ の 2 つの放射性薬剤を用いて非侵襲的な血流の測定が可能であることが示されている。Peretti らは、PET 画像の一種である PIB スキャンを用いて脳灌流とグルコース消費量に着目し、患者パターンや放射線科医の知識に依存しないアルツハイマー病の診断方法を提案している [14]。PIB では ^{11}C でラベルした放射性薬剤

を用いて撮影を行うが、この研究では PIB 局所脳血流画像を用いた診断結果と実際に使用される FDG 画像を用いた診断結果を比較して、PIB スキャンによる代用の有効性が示されている。

1.5 腎臓と PET 画像

1.5.1 腎臓のはたらき

腎臓は胸部の背面に位置する臓器で、左右に1つずつ位置しており、形はそら豆に例えられることが多い。腎臓は構造上さらに皮質と髄質に分けることができ、外側の皮質では血液を濾過して原尿が生成され、内側の髄質では原尿の約99%が再吸収され最終的な尿が生成される [15]。腎臓のはたらきには主に老廃物の排泄、水分・電解質のバランス維持、血圧の調整、ホルモンの生成などがあり、体内を常に最適な環境に保つ機能が備わっている [16]。

1.5.2 糖尿病と腎不全

糖尿病はインスリンの働きが不十分なために血液中を流れる糖が増加する疾患である。世界中に多くの糖尿病患者が存在し、糖尿病を発症することによって脳卒中や網膜症をはじめとした様々な合併症を引き起こす可能性が高くなることが知られている。また、糖尿病は腎臓の働きとも密接な関係があり、高血糖状態が続くと血管の狭窄や動脈硬化により高血圧が生じ、腎臓の皮質内にある毛細血管が密集した糸球体が障害を受け、腎臓の濾過機能が低下する [15]。次第に老廃物が体内を循環するようになり、糸球体がさらに障害を受け、やがて正常な尿を生成する機能が失われる腎不全を発症する可能性が高くなる。

1.5.3 糖尿病や腎不全の診断方法

現在、糖尿病や腎不全の診断方法については、尿検査や血液検査を用いて診断を確定させるのが一般的となっている [17]。しかし、中には PET 画像を用いて腎臓周辺の血流を調べる場合がある。図 1.2 は経過時間と Aorta(大動脈) および

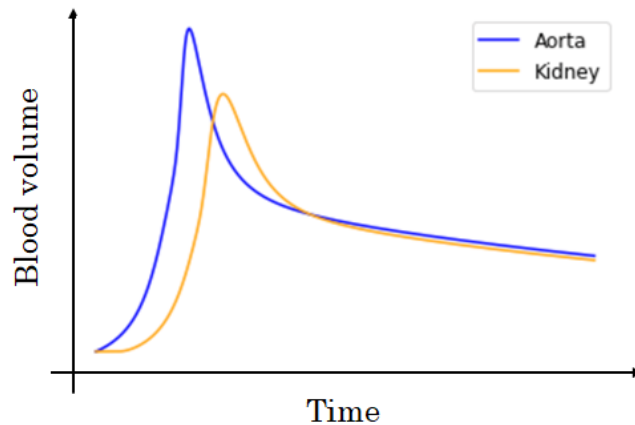


図 1.2 時間と血流の関係

Kidney(腎臓)に流れ込む血液との関係を表した簡略図である。これらの器官が正常に活動をしている場合、図 1.2 に示すようにまず大動脈に血液が流れ込み、その後少し遅れて腎臓に血液が流れ込むため、2つの山の形が並んだような概形となる。しかし腎臓に異常が診られる場合、腎臓に適切な量の血液が流れ込まなくなるため、黄色で示された波形に変形が起こるようになる。このように、糖尿病や腎不全の精密検査では血流を知ることが重要であり、PET 画像に映る腎臓とその周辺器官の輝度の時間的変化を解析することによって血流を把握して診断に用いる場合がある。

1.6 PET 画像に関する現状と課題

先述したように、糖尿病や腎不全の検査では、PET 画像を用いて腎臓を出入りする血液量の変化を把握し、より精密な検査をする場合がある。PET 画像から腎臓などの関心領域をセグメンテーションして血流に関する評価を行うが、その際医師たちは立体的な 3 次元 PET 画像から関心領域を手動で抜き出すため、彼らにとって時間のかかる負担の大きな作業となっている。これまで自動化によって医師の診断を支援するための研究は盛んに行われてきたが、それらの多くは X 線画像、CT、MRI などの利用頻度の高い医療画像が中心に行われており、PET 画像を積極的に

扱っていない傾向がある。いくつかの研究では PET 画像からの血流推定に関する研究が行われているが、それらの多くは脳あるいは心筋を対象としており、複雑な構造や組織同士の癒着が多く見られる腹部臓器を対象としているものは少ない。いくつかの研究の中には腹部臓器を対象としているものもあるが、それらの多くは PET と CT を併用している研究が多い。したがって、腹部臓器である腎臓の血流を定量的かつ正確に推定するために、他の医療画像よりも解像度が低い PET 画像のみから高い精度で関心領域をセグメンテーションできる機械学習モデルの構築が望まれる。

1.7 本研究の目的

本研究では、深層学習の手法を用いて 3 次元 PET 画像から腎臓および腹部大動脈部分を抽出するモデルを提案する。対象部位を自動抽出できるセグメンテーションモデルを構築することで、医療現場での診断の自動化を促進させ、医師の負担を軽減することを目的とする。

1.8 本研究の取り組み

本研究では、医療現場における診断の自動化に貢献するために、医療画像の 1 つである 3 次元 PET 画像から腎臓および大動脈部分をセグメンテーションするモデルを提案する。具体的には、畳み込みニューラルネットワークモデルの一種である U²-Net と U-Net の 2 つの深層学習モデルを用いて、3 次元腹部 PET 画像とそれに対応した腎臓・大動脈のアノテーション付き正解画像を使った教師あり学習を行った。学習後は Accuracy と Dice の 2 つの評価手法を用いてそれぞれのモデルの有効性を検証した。得られた結果から U²-Net と U-Net のそれぞれの特徴や差異について考察し、さらには数種類の Data Augmentation(データ拡張)手法を適用することで、セグメンテーション精度の向上を試みた。

第 2 章 関連研究

機械学習の発展に伴い、これまでに多くの深層学習モデルの研究や開発が進められてきた。セマンティックセグメンテーションとは、画像認識技術の 1 つで、画像内にあるすべてのピクセル (画素) に対して正解ラベルを割り当てる手法のことを指す [18]。医療画像の分野でもセマンティックセグメンテーションに関する研究がされており、これまでに従来手法の改良や新規モデルの開発などが行われてきた。

畳み込みモデルをベースとした先行研究の 1 つに、3 次元 CT 画像からの腎臓腫瘍のセグメンテーションを行った研究がある [19]。この研究では、輪郭強調やデータセットの効率的な分割が前処理段階で行われた後、計算量削減のために、スライスされた 3 次元立体画像を 3 スライスずつ z 軸方向に 1 スライドさせながら畳み込みモデルに入力されている。これにより、複数の評価指標で低い計算コストながら 3 次元畳み込みモデルに匹敵する精度が報告されている。PET 画像を扱った先行研究に、肺周辺の 2 次元 PET 画像からの肺結節のセグメンテーションがある [20]。この研究では畳み込みモデルである U-Net を用いてセグメンテーションが行われているが、畳み込み層の数を変更しながら精度の検証がされている。結果として、原著の U-Net よりも少ない 14 層の畳み込み層で最も良い精度を収めているが、このようになった理由として筆者らは限られたデータセットで大量のパラメータを最適化することの困難さや、関心領域が画像全体に対して非常に小さいことに起因する畳み込み過程での結節部位の消失を挙げている。

Transformer は元々自然言語処理の分野で広く活用されていた深層学習モデルであるが、Vision Transformer[21] の開発により、画像認識の分野でも Transformer が使用されるようになった。医療画像に Transformer をベースとしたモデルを適用している先行研究に、3 次元 CT 画像から複数の臓器をセグメンテーションしたものが [22]。この研究では、畳み込み層と画像内にある物体同士の類似度の計算を行う Self-Attention 層が交互に配置されたモデルを使用しており、畳み込みのみで構成されたモデルよりも低い計算コストで学習できることが述べられている。8 つの臓器についてセグメンテーションが行われたが、臓器ごとに精度にばらつきが見られるといった問題点が挙げられている。Swin Transformer[23] を医療画像のセグメンテーションに適用した先行研究が存在する [24]。この研究の提案

モデルでは、エンコーダーとデコーダーの両方に Self-Attention 計算を行う Swin Transformer ブロックを採用し、大腸のポリープ写真のセグメンテーションで他の最先端手法よりも上回る結果を残している。

医療画像以外のセマンティックセグメンテーションの例として、U²-Net を用いた物体検出がある [25]。U²-Net は Residual U-block と呼ばれるの U-Net のような構造を複数個つなげた形となっており、事前学習をせずにスクラッチを用いた学習のみで高精度なセグメンテーションを実現している。また、U²-Net よりも画像から抽出する特徴量を減らし、計算量を大幅に削減した U²-Net[†] も提案されており、オリジナルモデルに匹敵する精度を達成している。

Transformer を用いたモデルでは、モデルの汎用性を高めるために大規模なデータセットを用いた学習が必要になる [21]。しかし、本研究ではデータセットの数に限りがあるため、畳み込みモデルである U²-Net と U-Net を用いて 3 次元 PET 画像からの腎臓・大動脈のセグメンテーションを行う。そして、得られた結果に対して比較や考察を行うことで、より高精度なセグメンテーション方法の提案を試みる。

第 3 章 データセット

3.1 概要

本研究では、データセットにフィンランドにある Turku 大学から提供していただいた 3 次元 PET 画像を用いた。データ数は 10 例であり、いずれの患者も腎臓に目立った疾患を持たない健康体である。各データの詳細情報については表 3.1 に示す通りである。

表 3.1 各データの詳細情報

subject	1	2	3	4	5	6	7	8	9	10
Sex (M/F)	M	M	M	M	M	M	M	M	M	M
Age	23	24	27	22	22	22	26	23	22	20
Hight (cm)	181	165	167	171	178	177	171	177	163	172
Weight (kg)	74	64	55	60	60	66	68	64	50	65
BMI	22.6	23.5	19.7	20.5	18.9	21.1	23.3	20.4	18.8	22.0

提供された 10 例の PET 画像の放射性薬剤には $H_2^{15}O$ が使用されている。複数ある放射性薬剤の中でも $H_2^{15}O$ は主に血流の可視化に用いられる薬剤で、PET 画像に映る放射性薬剤による輝度を解析することによって血液量を推定することができる。つまり、腎臓内のある一部分が明るくなっているほど、その部分に血液が流れ込んでいると解釈することができる。また、今回提供された画像データは時系列データとなっており、1 人の患者につき 420 秒間に 33 回のスキャン (撮影) が行われている。しかし、スキャンされたそれぞれの 3 次元画像は非常にぼやけて映っている場合や全体が上手く映っていない場合があったため、より鮮明な画像を得るために 420 秒間に撮影された 33 個の 3 次元画像を時間軸方向に足し合わせたものを 1 つのデータとしてセグメンテーションを行った (図 3.1)。

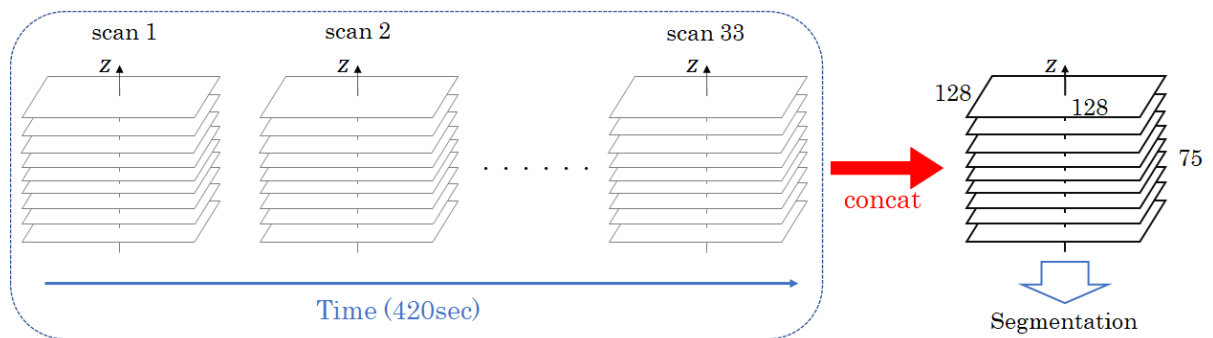


図 3.1 時系列データからのデータセットの作成

3.2 正解画像の作成

提供されたデータセットには 3 次元 PET 画像の原画像のみが含まれており、腎臓と大動脈のアノテーションは含まれていなかった。そこで本研究では、腎臓と大動脈の正解画像を取得するために、セグメンテーション用のライブラリの 1 つである ITK-SNAP を用いて腎臓及び大動脈領域の抽出を行った。具体的には ITK-SNAP に実装されているセグメンテーションツールを使って、指定した閾値の範囲にある領域の探索を行い、その後手動で補正をして正解画像を作成した。実際に使用するデータセットの例は図 3.2 に示す通りである。左から順に (a) 原画像、(b) 作成した正解画像、(c) 重ね合わせた図であり、(b) 中の緑色の部分が大動脈、黄色の部分が腎臓である。

完成した 10 例すべてのデータセットの解像度は、 $(H, W) = (128, 128)$ であり、スライス数は 75 である。

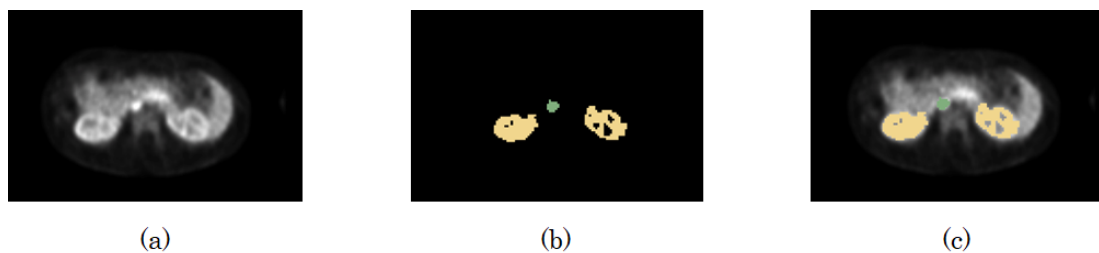


図 3.2 本研究で使用するデータセットの一部

第 4 章 PET 画像からのセグメンテーション

4.1 提案手法

本研究では、 U^2 -Net と U-Net の 2 種類の畳み込みニューラルネットワークモデルを用いて 3 次元 PET 画像からの腎臓および大動脈のセグメンテーションを行った。本研究におけるセグメンテーションとは、画像内にあるすべてのピクセルに対してラベルを割り当てるセマンティックセグメンテーションのことであり、ここでは背景、腎臓、大動脈の 3 種類のラベルを割り当てる。次からは本研究で使用した U^2 -Net と U-Net について概説する。

4.1.1 U^2 -Net

U^2 -Net は物体の切り抜きのために開発されたセグメンテーションモデルの 1 つで、画像内にある最も顕著な物体を検出し、背景と分離させることができる [25]。医療画像の分野に限らず、これまで多くのセグメンテーションモデルが開発されてきたが、[25] の筆者らは分類タスクのために開発された機械学習モデルがセグメンテーションタスクに応用されていることや、低い計算コストで高解像度の特徴マップを適切に扱えないかといったことに疑問を持ちこのモデルを提案している。この問題に対処するために U^2 -Net ではスクラッチのみで訓練が可能になっており、さらには次項で述べる U-Net のような構造に設計することで計算コストを大幅に抑えることを可能にした。

U^2 -Net は Residual U-Block(RSU) と呼ばれる構造が基本単位となっており、それが複数個組み合わされてモデルが構築されている。図 4.1(a) は RSU の構造である。 H, W は画像サイズ、 C_{in}, C_{out} はそれぞれ入力画像と出力画像のチャンネル数、 M は中間層でのチャンネル数を表している。RSU は Encoder-Decoder 構造をもつ左右対称の設計がされており、Encoder 側でダウンサンプリングされながら異なる解像度で効率的に計算されるので、大域的な情報と局所的な情報の両方を損なわずに特徴マップが生成されることが期待できる。また、Encoder 側から Decoder 側へは横方向に接続がされているため、画像内での物体の位置情報も保持されると

考えられる。また、 L は RSU における Encoder-Decoder 構造の深さを表しており、 L の値が大きくなるほど RSU の構造は深くなり、畳み込み操作やプーリング操作が増えるため、より局所的な特徴を捉えることができるようになる。

最初に畳み込み層に入力された画像 $x(H \times W \times C_{in})$ は、1 つ目の畳み込み層でチャンネル数が C_{out} の特徴マップ $F_1(x)$ に変換される。得られた特徴マップ $F_1(x)$ は、図 4.1(a) に示された U-Net のような構造を持った RSU に入力されることによって、マルチスケールな情報量を含んだ特徴マップ $U(F_1(x))$ に変換される。そして、図 4.1(b) に示すように、 $U(F_1(x))$ と $F_1(x)$ を Residual な接続で足し合わせることによって、局所的な情報とマルチスケールな情報の両方を持った $U(F_1(x)) + F_1(x)$ を得ることができる。

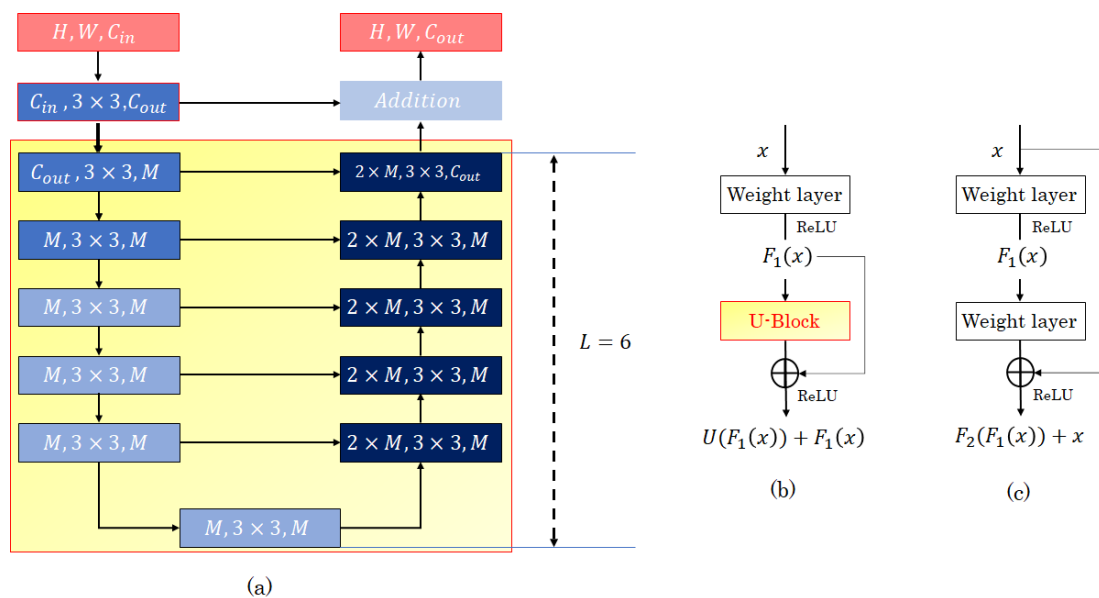


図 4.1 ReSidual U-Block の詳細 (a) と ReSidual U-Block 接続 (b) および Residual 接続 (c)

図 4.1(c) は ResNet に用いられている Residual 接続である [26]。ResNet は畳み込みニューラルネットワークモデルの 1 つで、畳み込み層を過度に増加させると精度が落ちるといった問題に対処するために、Shortcut connection を用いて残差関数を学習させることを試みたモデルである。入力 x を最初の畳み込み層に入力し

特徴マップ $F_1(x)$ を得た後、それを次の畳み込み層に入力にすることで特徴マップ $F_2(F_1(x))$ に変換される。さらに最初の入力 x を Shortcut connection によって足し合わせることで局所的な情報と元々の情報を併せ持った $F_2(F_1(x)) + x$ が得られる。しかし RSU を用いると、局所的な情報とマルチスケールな情報の両方を持った $U(F_1(x)) + F_1(x)$ に変換されるので、よりグローバルな情報が反映されることが期待できる。

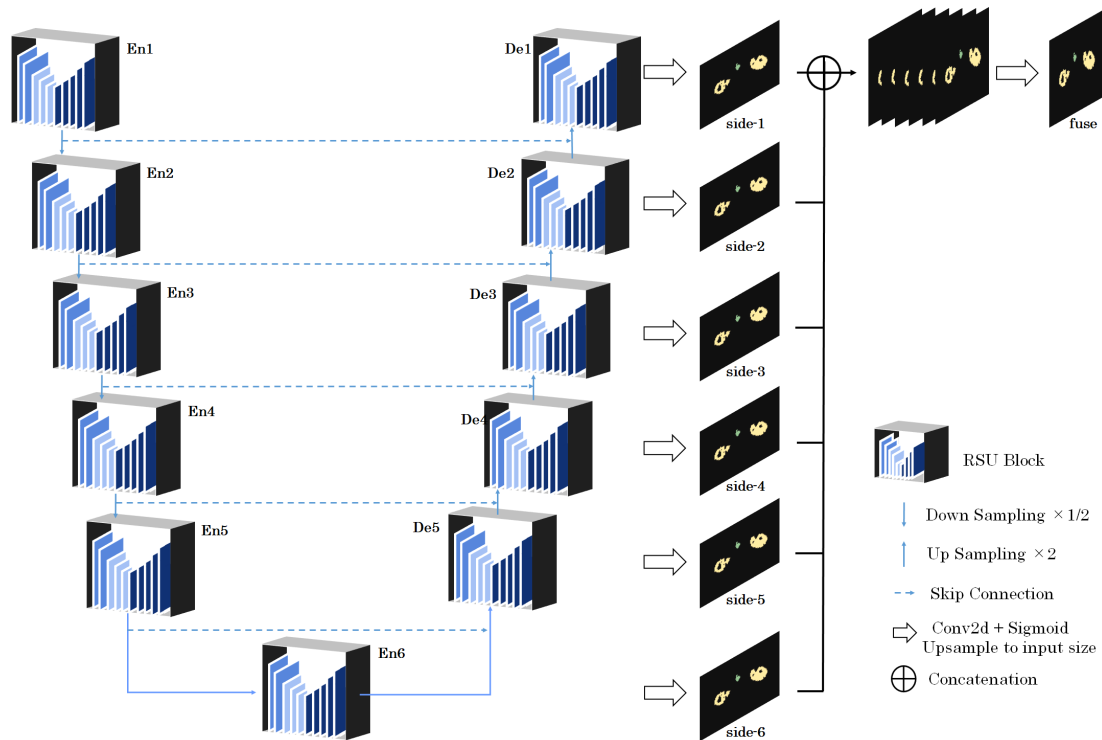


図 4.2 本研究で使用する U²-Net の構成

U²-Net 全体の構造は、図 4.2 に示す通りである。11 個の RSU ブロックが U-Net のように左右対称に並べられており、6 つの Encoder、5 つの Decoder、そして各ステージで出力された特徴マップを足し合わせるためのモジュールで構成されている。各ステージでの RSU ブロックの深さ L 、Input channel 数 C_{in} 、中間層での channel 数 M 、Output channel 数 C_{out} は表 4.1 にある通りである。En6、De5、De4、De3、De2、De1 の 6 つのステージで出力された特徴マップは、入力画像と同じ大きさに Upsampling された後、6 枚すべてが足し合わされ、畳み込み処理と

Sigmoid 関数によって処理が行われて最終的な予測マップが出力される。

表 4.1 各 RSU ブロックの詳細

stage	En1	En2	En3	En4	En5	En6	De5	De4	De3	De2	De1
L	7	6	5	4	4	4	4	4	5	6	7
C_{in}	3	64	128	256	512	512	1024	1024	512	256	128
M	32	32	64	128	256	256	256	128	64	32	16
C_{out}	64	128	256	512	512	512	512	256	128	64	64

4.1.2 U-Net

本研究でセグメンテーションに使用する 2 つ目のモデルは U-Net である [27]。U-Net は医療画像のセマンティックセグメンテーションのために提案された CNN であり、全結合層を持つ代わりに畳み込み層のみで構成された Fully Convolutional Network(FCN)[28] の一種である。U-Net では FCN の構造をもとに、さらには転置畳み込み (逆畳み込み) やスキップ接続が導入されている。転置畳み込みでは、畳み込み演算の逆の計算を行うことで圧縮した特徴を目的のサンプルサイズに拡大することができ、スキップ接続では前半の Encoder で計算した特徴マップを後半の Decoder に受け渡すことで、各層での特徴や位置情報を失うことなく演算することが可能となっている。U-Net 全体としては、Encoder-Decoder 形式の U 字形をした対称的な構造をしており、Encoder で入力画像を数回にわたり畳み込みながら特徴を抽出し、Decoder では Encoder から受け取った特徴マップを逆畳み込みで大きくしながら入力画像と同じサイズの確率マップを出力する。

本研究で使用する U-Net の構造は図 4.3 に示す通りである。主に 2 次元畳み込み層、Batch Normalization、ReLU 関数から構成されており、Down Sampling と Up Sampling には Max Pooling を使用した。

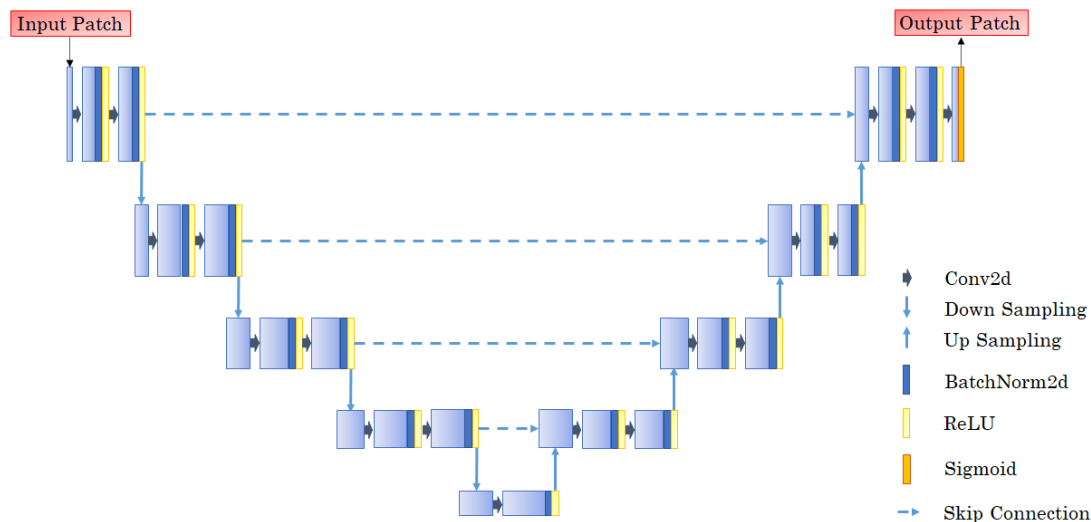


図 4.3 本研究で使用する U-Net の構成

4.2 実験

前章で述べた 3 次元 PET 画像のデータセットを用いて、U²-Net と U-Net の 2 種類の畳み込みニューラルネットワークモデルによる腎臓および大動脈のセグメンテーションを行った。本研究における実験の概要図は図 4.4 に示す通りである。本研究ではデータ数が 10 例と非常に限られているため、第一段階では Data Augmentation によってデータ数の増加を試みる。今回適用した Data Augmentation 手法は、立体画像を傾ける 3 次元 Rotation と画像内の輪郭を強調するラプラシアンフィルタである。また、Data Augmentation の有効性を検討するために、データの増しをせずにオリジナルのデータのみでも実験し、結果の比較を行った。次に U²-Net もしくは U-Net のセグメンテーションモデルに画像の入力を行うが、計算コスト削減のために、3 次元 PET 画像を一度に入力するのではなく、z 軸に垂直な面で切ったスライスを 3 枚ずつ、そして z 軸方向に 1 枚ずつスライドさせながら (縦, 横, 高さ) = (128, 128, 3) のサイズでモデルへの入力を行った。最後に 2 つのモデルからの出力結果に対して比較・考察をし、結果を立体図形で表示することで視覚による定性的な評価も行った。以下では Data Augmentation ならびに学習方法についての詳細を述べる。

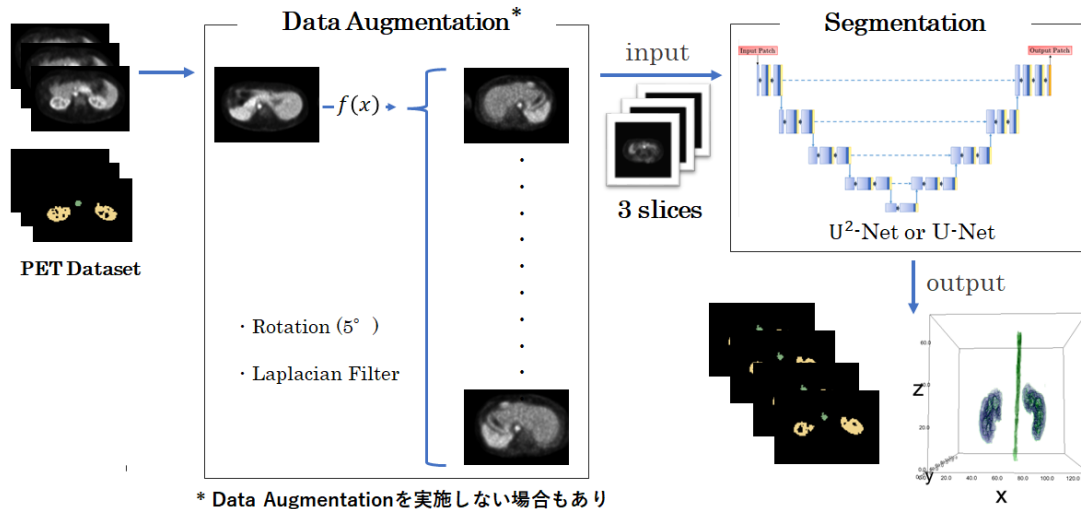


図 4.4 実験の概要図

4.2.1 Data Augmentation

本研究ではデータ数が非常に限られているため、Data Augmentation でデータ数を増加することを試みた。適用した手法は 3 次元 Rotation とラプラシアンフィルタ [29] である。3 次元 Rotation では、図 4.5 のように 3 次元 PET 画像を x 軸方向に ± 5 度、 y 軸方向に ± 5 度のそれぞれ 4 方向に立体画像を傾けた。

ラプラシアンフィルタは画像内の物体の輪郭を強調する際に用いられるフィルタで、画像内のエッジを検出することに用いられる。PET 画像は、CT や MRI などの他の医療画像と比較して、解像度が低くぼやけて見えることが知られている。したがって、輪郭を強調することでシャープな画像を生成できるラプラシアンフィルタは腎臓および大動脈の抽出に有効であると考えた。本研究では画像処理ライブラリの 1 つである OpenCV[30] を用いてラプラシアンフィルタの実装を行った。ラプラシアンフィルタでは 2 次微分で画像内のエッジを検出するが、微分を 2 回行うことで 1 次微分フィルタよりもより勾配変化の大きい箇所を強調することができる。勾配の変化 $\Delta grad$ は以下の式で計算される。

$$\Delta grad = \frac{\partial^2 grad}{\partial x^2} + \frac{\partial^2 grad}{\partial y^2}$$

本研究では 3×3 のサイズの 4 近傍のカーネルを使用したので、上記の式を用いて計算するとカーネルの値は図 4.5 の通りとなる。また、モデルに入力する際ラプラシアンフィルタに通さない場合はオリジナルの PET 画像を 3 スライスずつモデルに入力するが、ラプラシアンフィルタに通す場合、オリジナルの PET 画像 3 スライスとそれぞれのスライスをラプラシアンフィルタに通した 3 スライスの計 6 スライス、つまり (縦, 横, 高さ) = (128, 128, 6) のサイズで入力する。

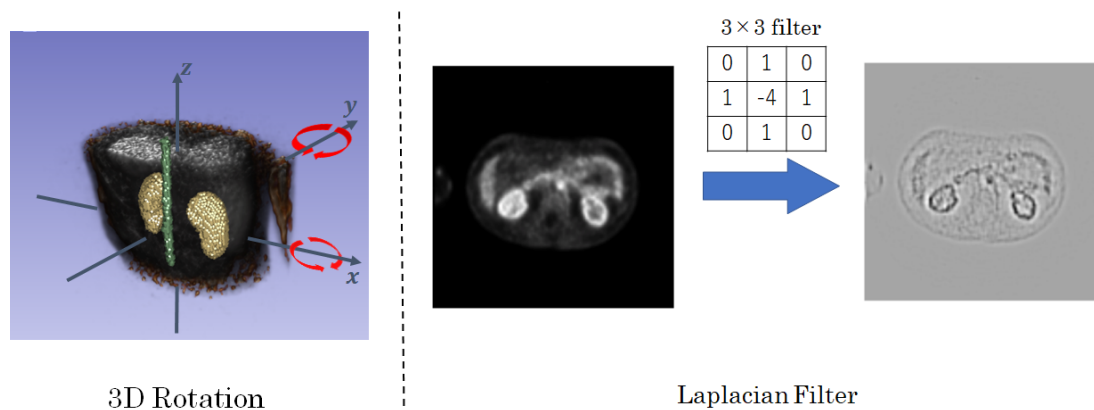


図 4.5 本研究で適用した Data Augmentation 手法

4.2.2 学習の詳細条件

ここではセグメンテーションモデルの学習段階に関する条件について述べる。使用する 3 次元 PET 画像のデータ数は 10 例と限られているため、本研究では学習用に 8 例、テスト用に 2 例のデータセットを使用した。また、モデルに画像を入力する際、本研究では計算コスト削減のために 3 スライスずつモデルへの入力を行うが、データの偏りをなくして過学習を防ぐために、データセット (バッチ) を作成する際にはシャッフル機能を利用し、ランダムな順序でモデルへ入力した。ネットワークの学習パラメータの更新には Adam[31] を使用し、学習率は 1.0×10^{-4} に設定した。バッチ数は 32 に設定し、200epoch の学習を行った (表 4.2)。また、深層学習ネットワークの構築には、Pytorch[32] を使用している。

次に損失関数について述べる。U²-Net で使用した損失関数は、以下の式で定義

表 4.2 ネットワークの学習条件

	損失関数	Batch Size	Epoch	Optimizer	学習率
U ² -Net	BCE	32	200	Adam	1.0×10^{-4}
U-Net	CE + Dice Loss	32	200	Adam	1.0×10^{-4}

される [21]。

$$\mathcal{L} = \sum_{n=1}^N w_{side-n} \ell_{side-n} + w_{fuse} \ell_{fuse}$$

ここで、N は図 4.2 中に示す side-1 から side-6 で出力される画像の枚数であり、今回は $N = 6$ である。 w_{side-n} は side-1 から side-6 までのそれぞれのステージの損失にかける重みを表しており、 w_{fuse} も同様に最終的に出力する損失にかける重みを表している。実験では、すべての w を 1.0 に統一して学習を行った。 ℓ_{side-n} と ℓ_{fuse} は Binary Cross Entropy(BCE) を表しており、以下の式で定義される。

$$\ell = - \sum_{(r,c)}^{(H,W)} [P_{G(r,c)} \log P_{S(r,c)} + (1 - P_{G(r,c)}) \log(1 - P_{S(r,c)})]$$

ここで (r, c) は画像内の座標、 (H, W) は画像のサイズ、 $P_{G(r,c)}$ と $P_{S(r,c)}$ はそれぞれ点 (r, c) における正解値と予測マップの出力値を表している。

U-Net の損失関数については、Cross Entropy(CE) と Dice Loss の和を用いており、以下の式で定義される。

$$\mathcal{L} = \sum_{(r,c)}^{(H,W)} \left\{ -P_{G(r,c)} \log \left(\frac{\exp\{P_{S(r,c)}\}}{\sum_{label} \exp\{P_{S(r,c)}\}} \right) + \left(1 - \frac{2P_{G(r,c)}P_{S(r,c)}}{P_{G(r,c)} + P_{S(r,c)}} \right) \right\}$$

Dice loss はクラス不均衡なデータに対しても学習が進むように設計された損失関数である [33]。Dice loss では、予測領域と正解領域がどのくらい重なっているのかを表す Dice 係数と呼ばれる指標に基づいている。学習時の詳細な条件については表 4.2 に示す通りである。

第 5 章 結果

5.1 評価指標

U²Net と U-Net によるセグメンテーション結果を評価するために、本研究ではまず評価指標の一種である Accuracy を用いて精度を確認した。Accuracy の定義は以下の通りである。

$$Accuracy = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N 1(y_{ij} = \hat{y}_{ij})$$

本研究では、入力画像のすべてのピクセルに対して背景、大動脈、腎臓の 3 種類のラベルのうちいずれかを割り当てる。セマンティックセグメンテーションにおいて、Accuracy は予測画像内のそれぞれのピクセルが正解画像のものとのくらい一致しているのかを示し、 y_{ij} は予測画像内の点 (i, j) における予測値、 \hat{y}_{ij} は正解画像内の点 (i, j) における正解値を表している。

しかし、本研究で扱うデータセットは画像内の大半が背景であり、大動脈や腎臓の領域が限られている。このように偏りのあるデータセットを用いる場合、仮に大動脈や腎臓部分が上手く抽出されずに、ほとんどが背景の予測画像が出力されても Accuracy は高い値となってしまふ。より実用的なモデルを開発するために、ここでは評価指標である Dice も使用して精度の評価を行う。Dice の定義は以下の通りである。

$$Dice = \frac{2 * TP}{2 * TP + FP + FN}$$

簡潔に述べると、Dice は予測領域が正解領域とどれだけ重なっているかを表した指標であり、図 5.1 のように表現することができる。また、ここでは大動脈や腎臓の領域に正しいラベルを割り当てることができた例を TP、真のラベルは大動脈もしくは腎臓であるピクセルに背景のラベルを割り当てた例を FN、真のラベルは背景であるピクセルに大動脈や腎臓のラベルを割り当てた例を FP としている。

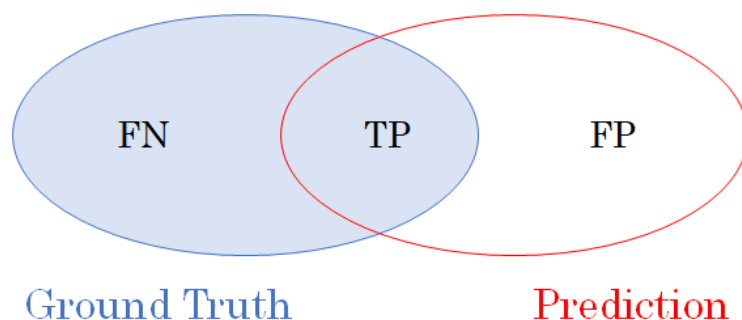


図 5.1 Dice の説明図

5.2 セグメンテーション結果

ここでは、U²-Net と U-Net の 2 つのモデルによるセグメンテーション結果を示す。それぞれのモデルについて、「Data Augmentation を行わない」、「3 次元 Rotation を行う」、「Laplacian Filter に通す」の 3 種類の方法で学習を行った。テスト用に使用したデータ数は患者 2 人分の 2 例であり、表 5.1 には 2 例のテスト用データの結果を平均して記した。

表 5.1 実験結果

model	U ² -Net			U-Net		
	Augmentation	None	Rotation	Laplacian	None	Rotation
Accuracy (%)	80.9	77.8	74.8	85.3	80.1	80.4
Dice (%)	82.2	81.0	74.5	84.9	86.4	82.1

図 5.2 および図 5.3 は「Data Augmentation を行わない」と「3 次元 Rotation を行う」場合の U²-Net と U-Net のセグメンテーション結果を抜粋したものである。青色の部分は大動脈、水色の部分は腎臓を表しており、1 列目の GT(Ground Truth) は正解画像である。また、 z は z 座標を表しており、今回は 10 スライスおきにそれぞれの画像を表示した。

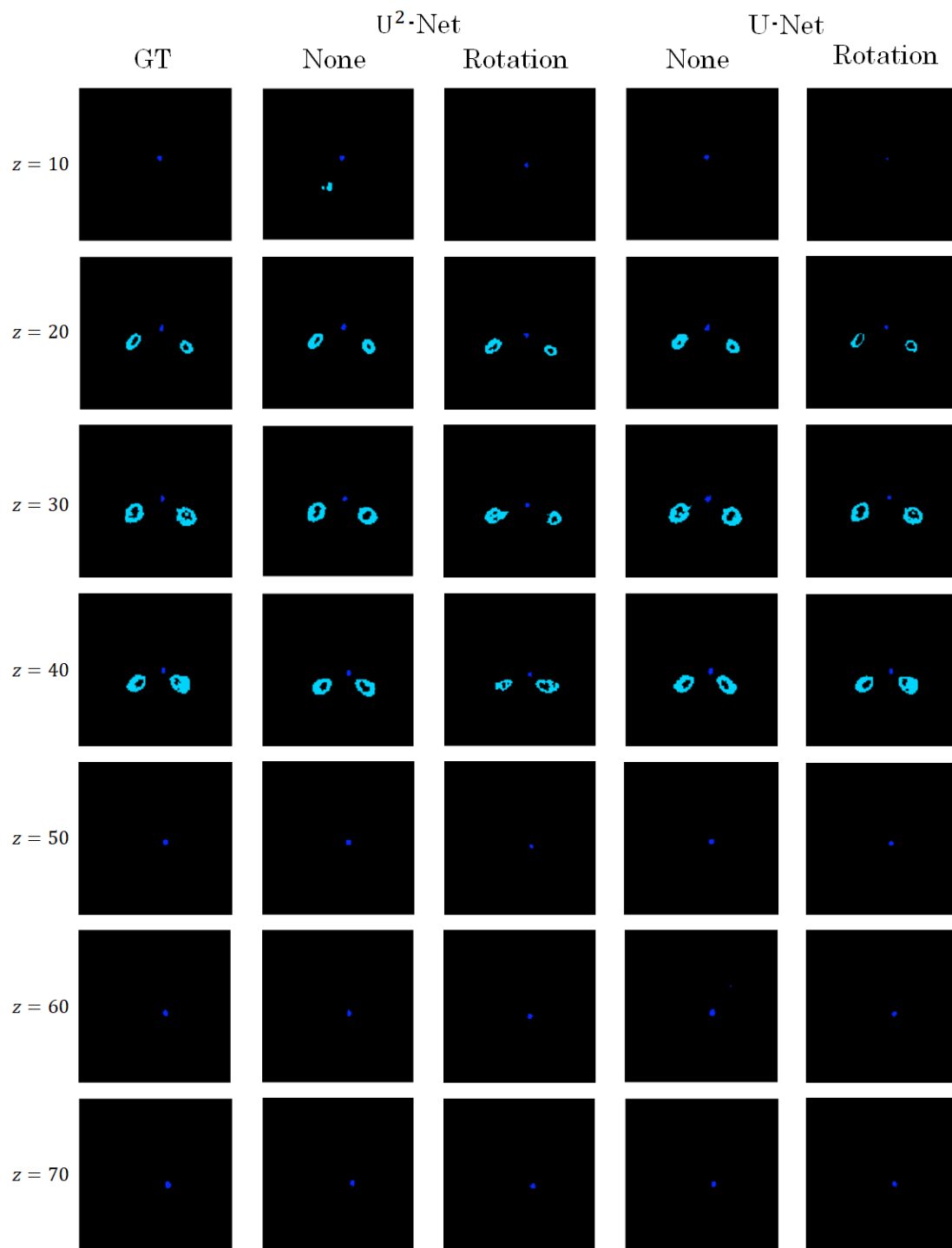


図 5.2 1 人目のセグメンテーション結果

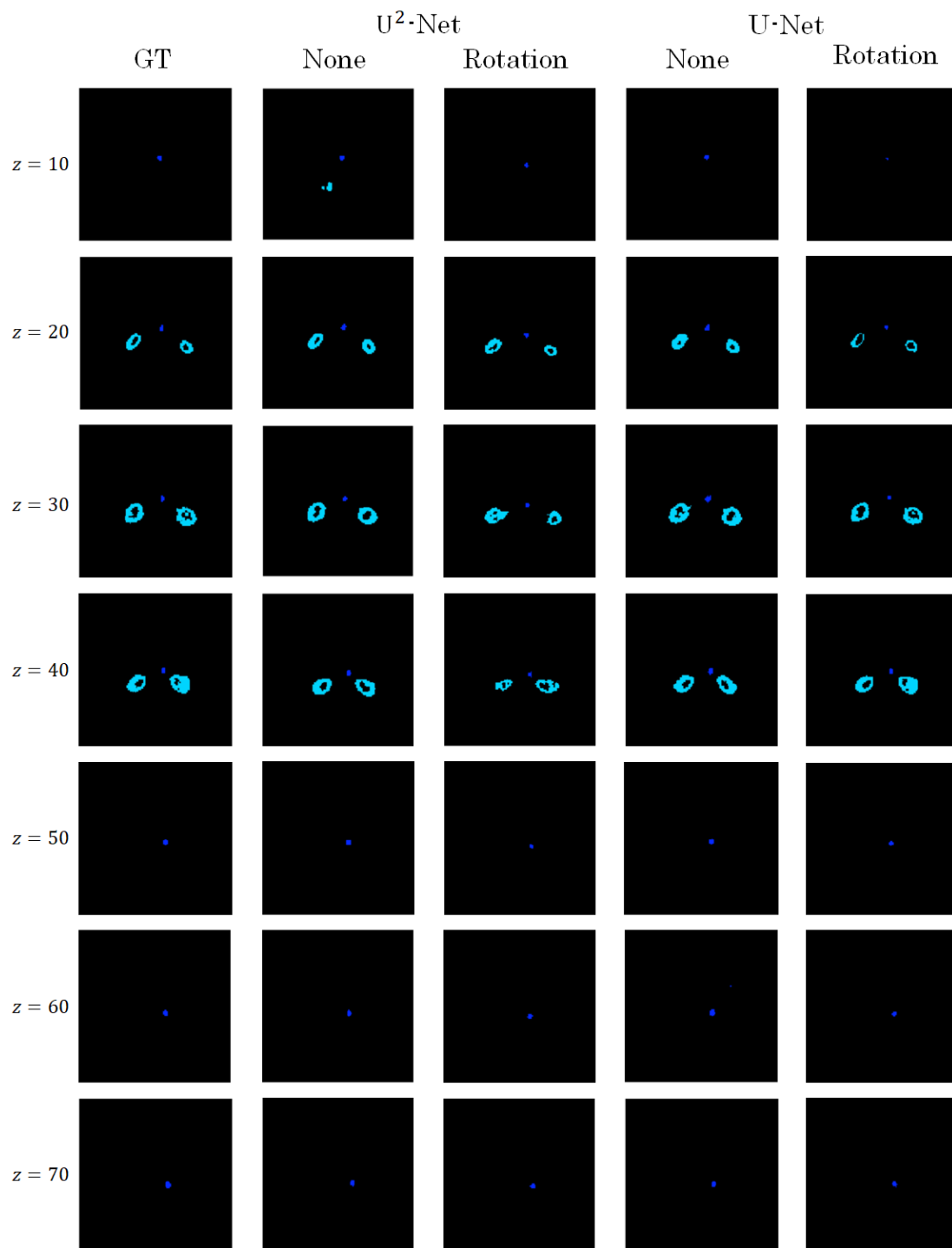


図 5.3 2 人目のセグメンテーション結果

第 6 章 考察

6.1 U²-Net と U-Net の精度比較

表 5.1 より、Data Augmentation を行わない場合、表 5.1 より U²-Net と U-Net の両方のモデルにおいて、2つの指標で 80% を上回る結果となったが、U-Net の方が精度が優れていることがわかる。精度に差が生まれた理由として、モデルの構造の違いを考えることができる。図 4.2 に示すように、U²-Net は複数の解像度の画像を用いて損失関数の計算を行い、出力の際も複数の解像度の出力を足し合わせて最終的な予測を行っている。本研究では、解像度が (縦, 横) = (128, 128) の PET 画像を用いており、図 4.2 中の En6 では (縦, 横) = (4, 4) まで圧縮される。[25] で提案された U²-Net は高解像度な画像用に開発されたモデルであるため、本研究のような低解像度の画像には適していない可能性が考えられる。また、図 6.1 に示した立体画像は、Data Augmentation を行っていない際の予測画像の一例であり、左から順に入力画像、正解画像、U²-Net の出力結果、U-Net の出力結果となっている。精度が下がった原因として、図 6.1 中の赤色の点線で囲まれたように余分な腎臓領域が出力されていることが考えられ、それらは腎臓下部に集中している。このことは他のテスト用データでも確認されたため、腎臓が映っている部分から映らなくなる部分にかけての特徴表現に乏しかったと思われる。

Data Augmentation 適用後の結果については、適用前よりも全体的に精度が下がった印象を受ける。PET 画像は低解像度で輪郭がぼやけるといった欠点を補うために、Laplacian Filter で輪郭を強調することを試みたが、2つのモデルで Accuracy と Dice の両方の指標が下がっていることから、PET 画像内の有益な特徴が Laplacian Filter によって失われていると考えられる。したがって、今後は PET 画像の輝度の分布を保ちながら輪郭の強調を行うことが必要である。3次元 Rotation については、データセットの少なさを補うために行ったが、精度を大きく改善することはできなかった。医療画像のセグメンテーションでは、抽出する臓器の位置を大幅に変更することは領域の特定に悪影響を及ぼすことが知られているため、本研究では回転角の大きさを 5 度に設定したが、精度が下がっていることが多いため、抽出したい領域の位置が変わらないような Data Augmentation が有効で

あると考える。また、3次元 Rotation を適用した結果 U-Net の Dice のみ精度が上がっていることが確認されたが、その理由については今後の研究で説明する必要がある。

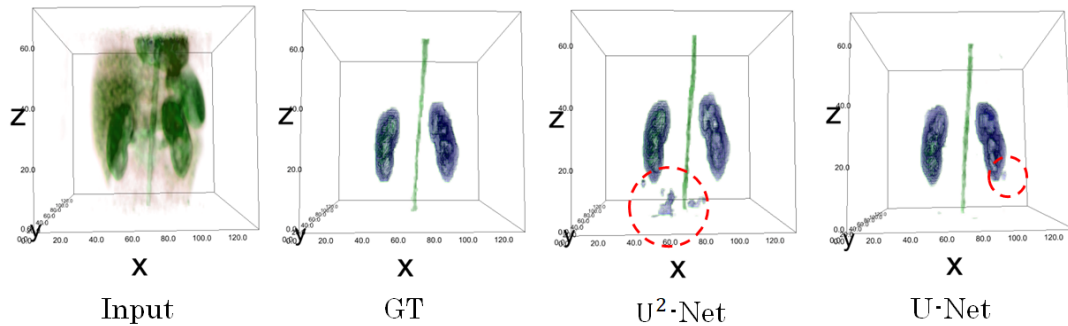


図 6.1 予測画像の 3 次元立体表示

6.2 連続値マスク適用の検討

精度が下がった原因として、図 6.2 に示すように青色で示された大動脈部分を適切に出力できなかったことも挙げられる。この現象は U^2 -Net および U-Net のいずれのモデルでも確認されたが、将来的に抽出した領域を血流推定の研究に適用する場合、大動脈が分かれて出力される予測や何も出力されない予測は医療現場での使用の際に効果的でないと考える。このように大動脈部分が出力されない原因として、正解ラベルが離散値で表現されていることが考えられる。本研究では{0:背景, 1: 大動脈, 2: 腎臓}の 3 つのラベルのうちいずれかがすべてのピクセルに割り当てられているが、{0:背景}と{1:大動脈}の境界部分を連続値を用いてなめらかに変化させることで、予測と正解データの間の損失の計算が連続的に変化するようになり、より適切に学習ができると考えた。以下では、大動脈部分のマスクを 0 から 1 の連続値に変更した連続値マスクの作成方法について説明する。

まずは、 z 軸に垂直な面で切断された 2 次元の正解画像内にある{1:大動脈}が割り当てられているピクセルの座標をすべて取得し、これらの座標の重心 $G(\bar{x}, \bar{y})$ を求める。{1:大動脈}が割り当てられているピクセルの座標を

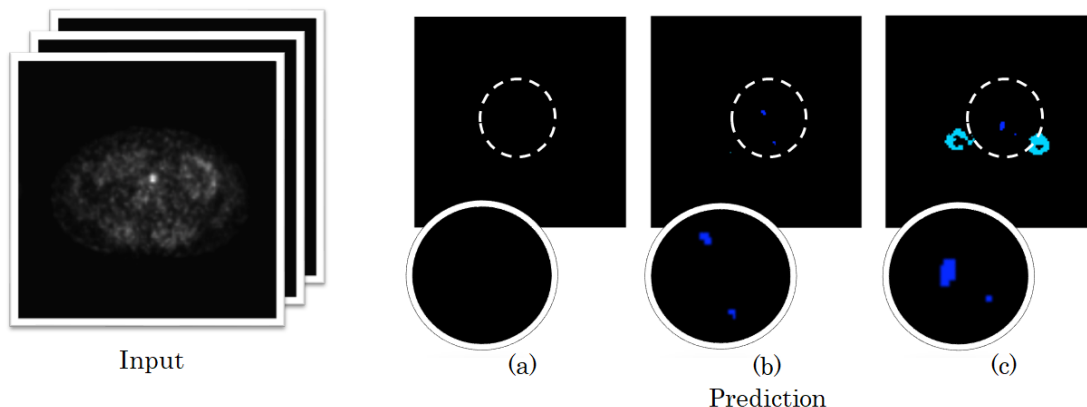


図 6.2 正確に大動脈のセグメンテーションを行えなかった例

$(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)$ とすると、 $G(\bar{x}, \bar{y})$ は以下の計算で求められる。

$$G(\bar{x}, \bar{y}) = \frac{1}{N} \sum_{i=1}^N (x_i, y_i)$$

次に $G(\bar{x}, \bar{y})$ と各点 $(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)$ の距離から、大動脈の半径 r を求める。 r は以下の計算で求められる。

$$r = \frac{1}{N} \sum_{i=1}^N \sqrt{(x_i - \bar{x})^2 + (y_i - \bar{y})^2}$$

$G(\bar{x}, \bar{y})$ と r をスライスごとに計算するが、スライスごとに半径 r の大きさが異なる。図 6.3 はある患者 1 人の 3 次元 PET 画像の各 z 座標でのスライスに映る大動脈の半径 r の大きさを表したグラフである。図 6.3 からわかるように、半径 r のスライスごとの変化に振動が確認されるため、大動脈の 3 次元立体画像に再構成した際に太さにばらつきがある大動脈が形成されると予想される。そこで本研究では、なめらかで連続的な大動脈を形成するために、ガウス過程 (Gaussian Process) による回帰 [34, 35] を用いて半径 r の補正を行った。

ガウス過程回帰とは、ある入力 x とその出力 y の関数関係 $y = f(x)$ を考えたときに、入力 $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N$ (N :自然数) の出力 $\mathbf{f}=(f(x_1), f(x_2), \dots, f(x_N))$ が平均 μ 、共分散行列 \mathbf{K} を分散とするガウス分布 $\mathcal{N}(\mu, \mathbf{K})$ に従うことを利用して、未

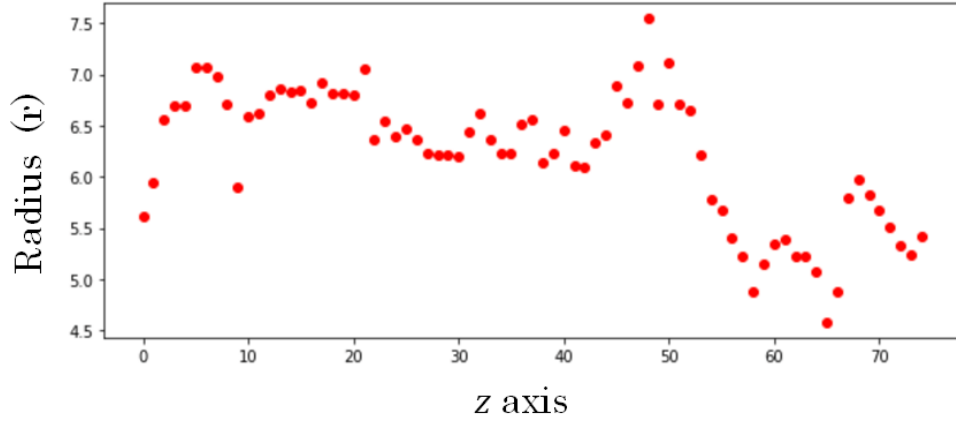


図 6.3 <補正前> 大動脈の半径 r のスライス毎の変化

知の入力 x_{new} に対応する y_{new} を予測する問題のことで、データのばらつきやノイズを考慮しながら非線形な関数を用いて回帰をすることができる [36]。

共分散行列 \mathbf{K} の各成分の値 k は、入力 \mathbf{x} の 2 点 x, x' の距離から算出され、本研究ではガウスクーネル関数を用いて計算した。式は以下の通りである。

$$k(\mathbf{x}, \mathbf{x}') = \theta_1 \exp\left\{-\frac{(\mathbf{x} - \mathbf{x}')^2}{2\theta_2^2}\right\} + \lambda\delta(\mathbf{x}, \mathbf{x}')$$

ここで、 $\theta_1, \theta_2, \lambda$ はガウスクーネルの特性を決定するパラメータであり、本研究では、 $\theta_1 = 1.0$ 、 $\theta_2 = 100$ 、 $\lambda = 0.1$ に設定した。 $\delta(x, x')$ はガウスクーネルにノイズを加える項で、 $x = x'$ ならば $\delta(x, x') = 1$ 、それ以外は $\delta(x, x') = 0$ となる。

既知の入力 \mathbf{x} とそれに対応する出力 \mathbf{y} から、新たな入力 \mathbf{x}_{new} に対応する y_{new} の予測分布は以下の式で求めることができる。

$$p(y_{new} | \mathbf{x}_{new}, \mathcal{D}) = \mathcal{N}(k_*^T \mathbf{K}^{-1} \mathbf{y}, k_{**} \mathbf{K}^{-1} k_*)$$

ここで、 $\mathbf{k}_* = (k(\mathbf{x}_{new}, \mathbf{x}_1), k(\mathbf{x}_{new}, \mathbf{x}_2), \dots, k(\mathbf{x}_{new}, \mathbf{x}_N))^T$ 、 $k_{**} = k(\mathbf{x}_{new}, \mathbf{x}_{new})$ である。これらの式を用いて大動脈の関数の補正を行った結果が図 6.4 に示す。

青色の線で引かれたグラフが補正後の大動脈の半径 R で、元々の赤色の点で示された補正前の半径 r を結ぶよりもなめらかな曲線であることがわかる。また、薄青色で塗られたエリアは誤差の範囲を表しており、今回は平均値 \pm 標準偏差を誤差の

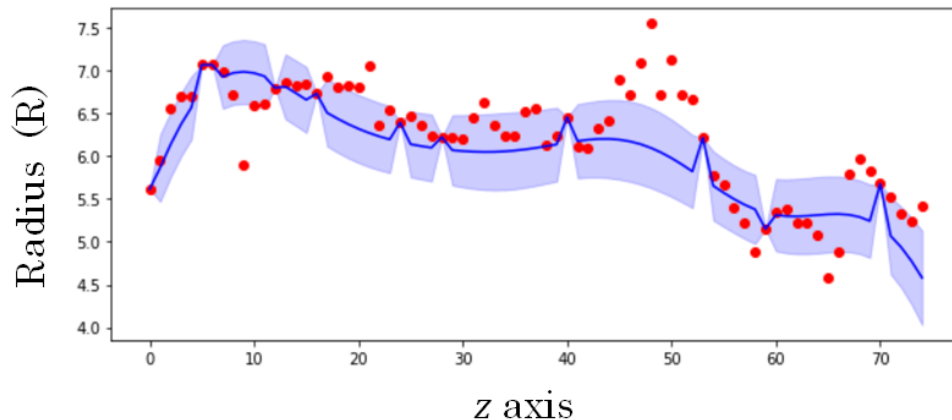


図 6.4 <補正後> 大動脈の半径 R のスライス毎の変化

範囲とした。図 6.3 および図 6.4 は患者 1 人分のグラフを示しているが、10 人分のすべてのデータについてこの操作を行った。

次に、補正した半径 R の大動脈内にあるピクセルに数値を割り当てる方法について説明する。半径 R の大動脈にある重心 $G(\bar{x}, \bar{y})$ のラベルを 1 とし、円の外側になるにつれて 0 に近づくようにラベルを付与する。この時、与えるラベルの数値はガウス分布 $\mathcal{N}(\mu, \sigma)$ に従うように設定し、ここでは $\mu = 0$ 、 $\sigma = R/2$ とした。

以上の手順で作成した大動脈の連続値ラベルを用いて、Data Augmentation を適用せずにセグメンテーションを行った結果、図 6.2(a) に示すような大動脈が何も出力されない失敗例は完全に無くことができたが、(b) や (c) に示すような大動脈が 2 つ以上に分かれて出力される点については改善することができなかった。したがって今後の展望として、本研究で使用したデータセットの大動脈のように、画像全体に対して抽出したい領域が小さい場合でも効率的に学習が行えるような損失関数の設計やデータセットの前処理が必要であると考ええる。また、本研究では大動脈の概形を円と仮定して連続値マスクの作成を行ったが、腎臓のように単純な図形で置き換えることができないような領域に対しても連続値マスクが作成できる方法を考案することは、将来的に有効であると考ええる。

6.3 血流推定へのセグメンテーション領域の適用

ここでは、得られたセグメンテーション結果を用いた血流推定の方法について述べる。U²-Net と U-Net の 2 つのモデルによるセグメンテーション結果を用いて比較を行うが、今回は表 5.1 中の Dice が最も高いものを使用した。つまり、U²-Net の場合は Data Augmentation を行わずに学習したモデル、U-Net の場合は 3 次元 Rotation を行って学習したモデルのセグメンテーション領域を用いて血流推定を行った。

今回得られた大動脈および腎臓の領域は、図 3.1 に示すように 420 秒間に撮影された 33 個の立体画像を時間軸方向に足し合わせた画像から得られたセグメンテーション結果である。したがって、まずは得られた大動脈と腎臓の領域を時間軸方向に足し合わせる前の画像である scan1, scan2, ..., scan33 に重ね合わせた。次に、大動脈と腎臓のそれぞれの領域内にあるボクセル値を足し合わせ、その後ボクセルの個数で割ることによって各 scan の大動脈と腎臓のボクセル値の平均を求めた。横軸に経過時間、縦軸に計算したボクセル値をとり、グラフにして表したものが図 6.5 である。

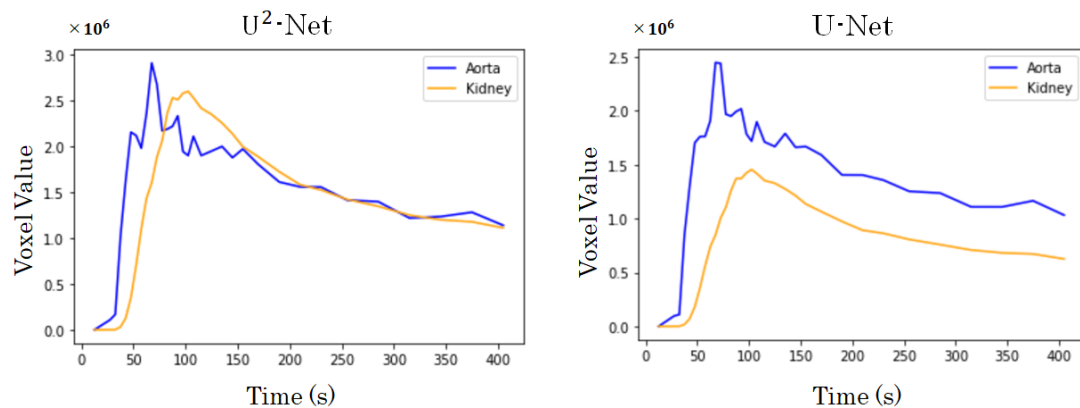


図 6.5 U²-Net と U-Net から得られた領域を用いた血流推定

図 6.5 はテスト用データの患者 1 人分の結果であり、横軸は撮影を開始してからの経過時間 (秒)、縦軸はボクセル値 (輝度) の大きさを表している。青色で表されたグラフは大動脈 (Aorta) 内を流れる血液量の時間変化であり、橙色で表されたグ

ラフは腎臓 (Kidney) 内を流れる血液量の時間変化である。U²-Net と U-Net の両方で図 1.2 で表されるような理論値から求められるグラフに似た概形を描けていることがわかる。また、本研究で使用しているデータは腎臓や大動脈に目立った疾患を持たない健康体であるが、得られたグラフから図 1.2 のような健常者の血流カーブに近いグラフを描けていることも確認できた。しかし、青色で表された大動脈のグラフの波形が U²-Net と U-Net の両方で滑らかに描けていないことから、今後はより正確な大動脈領域の抽出が必要である。さらに、本研究のようにセグメンテーション結果から予測された血流のグラフが理論値から求められたグラフや医師が作成したグラフなどの正解データとどのくらい適合するのかを正確に測れる評価指標や手法の考案も必要である。

第7章 結論

本研究では、医師たちが糖尿病や腎不全の検査で PET 画像を用いて血流の可視化をする際に大動脈や腎臓を手動で抽出していることを課題に捉え、彼らの負担を軽減することを目的に、畳み込みニューラルネットワークモデルによるセグメンテーションモデルの提案を行った。PET 画像に関する従来研究では、脳や心筋を対象としたものが多く、腹部臓器が対象となっている研究では、組織の複雑性などから CT と併用して扱われるものが多いが、本研究では U²-Net と U-Net の 2 つの深層学習を用いて PET 画像のみからの血流推定に有効な腎臓および大動脈のセグメンテーションを試みた。実験では、3次元 Rotation の Data Augmentation を適用してモデルを学習した後、U-Net でセグメンテーションを行った結果、評価指標の 1 つである Dice で 86.4% の精度を確認することができた。さらに、得られた腎臓と大動脈の領域を用いて 3次元 PET 画像を解析することで、血流の時間変化を表したグラフが描けることを示した。このことから、本研究で抽出した腎臓と大動脈の領域を用いて血流の推定を行うことは十分可能であると考えられる。今後の展望として、より適切な血流グラフを描くために、より正確な大動脈および腎臓の領域抽出が必要となる。本研究では Data Augmentation や連続値マスクの導入によって精度の向上を試みたが、今後は PET 画像に最適な Data Augmentation 手法の発見や提案が求められる。また、本研究で示したような血流の予測グラフが医師が作成した血流グラフや理論的に求められるグラフなどの正解データとどのくらい適合するかを測れる評価指標や手法の提案も必要であると考えられる。

謝辞

本研究は、多くの方々からの御協力のおかげにより進めることができました。はじめに、本研究を進めるにあたり、日頃より有益な助言と丁寧な御指導を賜りました、奈良先端科学技術大学院大学 先端科学技術研究科 情報科学領域 計算システムズ生物学研究室 金谷重彦教授に深く感謝申し上げます。

研究発表や論文執筆の際に、副指導教員として貴重な御意見を賜りました、奈良先端科学技術大学院大学 先端科学技術研究科 情報科学領域 ソフトウェア工学研究室 松本健一教授に深く感謝申し上げます。

本研究を進めるにあたり、私が疑問に感じた部分やわからない部分を懇切丁寧に対応していただき、研究テーマの決定から修士論文の作成まで様々なご享受を賜りました、奈良先端科学技術大学院大学 先端科学技術研究科 情報科学領域 計算システムズ生物学研究室 小野直明准教授に深く感謝申し上げます。

研究室ゼミや授業、研究室での生活を通して多くのことを学ばせていただいた奈良先端科学技術大学院大学 先端科学技術研究科 情報科学領域 計算システムズ生物学研究室 Md.Altaf-UI-Amin 准教授ならびに黄銘助教授のお2人に深く感謝申し上げます。

日頃の研究室での生活や私生活を含め多くの面から支えてくださった計算システムズ生物学研究室内のスタッフ、メンバーの皆様に深く感謝申し上げます。皆様のおかげで大変充実した大学院生活を送らせていただくことができました。

最後になりますが、大学院修了まで多大な支援を頂いた家族に深く感謝し、ここに謝辞とさせていただきます。

参考文献

- [1] OECD. “Health at Glance 2021”. 2021
- [2] 竹中大祐 “「肺癌検診 State of the art」 肺癌検診の基本から最新動向まで” 肺癌 2020; 60: 895-897
- [3] K.Doi. “Computer-aided diagnosis in medical imaging: historical review, current status and future potential”, Computerized medical imaging and graphics : the official journal of the Computerized Medical Imaging Society, 31(4-5):pp.198–211, 2007.
- [4] N.Khalili, P.Moeskops, N.H.P.Classens, et.al, “Automatic segmentation of the intracranial volume in fetal MR images”, Fetal, Infant and Ophthalmic Medial Image Analysis, pp.42-45, 2017
- [5] T Okada, M G Linguraru, M Hori, et.al, “Abnominal multi-organ segmentation from CT images using conditional shape-location and unsupervised intensity priors”, Medical Image Analysis, vol.26, issue 1, pp.1-18, 2015
- [6] Alex Krizhevsky, Ilya Sutskever, Geoffrey E. Hinton, “Image Net Classification with Deep Convolutional Neural Networks”, Advances in Neural Information Processing Sysyem 25(NIPS 2012)
- [7] Bhandary Abhir, Prabhu G. Ananth, Rajinikanth Venkatesan, Thanaraj K. Palani, et. al, “Deep-learning framework to detect lung abnormality–A study with chest X-Ray and lung CT scan images”, Pattern Recognition Letters, volume 129, pp. 271-278, 2020
- [8] Mohamed A. Naser, M. Jamal Deen, “Brain tumor segmentation and grading of lower-grade glioma using deep learning in MRI images”, volume 121, June 2020, 103758
- [9] 北村圭司 “PET. 日本医用画像工学ハンドブック編集委員会辺” 医用画像工学ハンドブック, 医用画像工学会, 東京, 2012; pp.293-320
- [10] 川野剛, 井上登美夫 “FDG-PET を理解するための基礎知識” 日本消化器病学会, 2004, 101 巻, 12 号, pp.1288-1291
- [11] Gunnar Brix, Ursula Lechel, Gerhard Glatting, Sibylle I. Ziegler, et.al, “Ra-

- diation Exposure of Patients Undergoing Whole-Body Dual-Modality ^{18}F -FDG PET/CT Examinations”, *The Journal of Nuclear Medicine*, vol46(4), pp.608-613, April 2005
- [12] Matthias Bruehlmeier, Ulrich Roelcke, Pius A. Schubiger, Simon Mensah Ametamey, “Assessment of Hypoxia and Perfusion in Human Brain Tumors Using PET with ^{18}F -Fluoromisonidazole and ^{15}O -H $_2\text{O}$ ”, *The Journal of Nuclear Medicine*, vol45(11), pp.1851-1859, November 2004
- [13] Philipp A. Kaufmann, Paolo G. Camici, “Myocardial Blood Flow Measurement by PET: Technical Aspects and Clinical Applications”, *The Journal of Nuclear Medicine*, vol46(1), pp.75-88, January 2005
- [14] Debora E. Peretti, David Vallez Garcia, Fransje E. Reesink, Janine Doorduin, et.al, “Diagnostic performance of regional cerebral blood flow images derived from dynamic PIB scans in Alzheimer’ s disease”, *EJNMMI Research*, Number 59, July, 2019
- [15] 坂井建雄 “初心者のための腎臓の構造”, *日腎会誌*, 2001, 43(7), pp.572-579
- [16] Christian Kurts, Ulf Panzer, Hans-Joachim Anders, Andrew J. Rees, “The immune system and kidney disease: basic concepts and clinical implications”, *Nature Reviews Immunology*13, 16 September 2013, pp.738-753
- [17] 小坂樹徳 “” 糖尿病の診断”, *日本内科学会雑誌*, 1982, 73(6), pp.759-763
- [18] Yanming Guo, Yu Liu, Theodoros Georgiou, Michael S. Lew, “A review of semantic segmentation using deep neural networks”, *International Journal of Multimedia Information Retrieval*, 7, 2018, pp.87-93
- [19] Luana Batista da Cruz, Domingos Alves Dias Junior, Joao Otavio Bandeira Diniz, et.al, “Kidney tumor segmentation from computed tomography images using DeepLabv3+ 2.5D model”, *Expert Systems with Applications*, Volume 192, 15 April 2022, 116270
- [20] Yongzhou Lu, Jinqiu Lin, Sheng Chen, Hui He, Yuantao cai, “Automatic Tumor Segmentation by Means of Deep Convolutional U-Net With Pre-Trained Encoder in PET Images”, *IEEE Access*, Volume:8, 17 June 2020, pp.113636-113648

- [21] Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, et.al, “An Image is Worth 16×16 Words: Transformers for Image Recognition at Scale”, Computer Vision and Pattern Recognition, 22 October 2020, arXiv preprint arXiv:2010.11929
- [22] Hong-Yu Zhou, Jiansen Guo, Yinghao Zhang, Xiaoguang Han, Lequan Yu, Liansheng Wang, Yizhou Yu, “nnFormer: Volumetric Medical Image Segmentation Via a 3D Transformer”, Computer Vision and Pattern Recognition, 7 September, 2021, arXiv preprint arXiv:2109.03201
- [23] Ze Liu, Yutong Lin, Yue Cao, Han Hu, Yixuan Wei, Zheng Zhang, Stephen Lin, Baining Guo, “Swin Transformer : Hierarchical Vision Transformer using Shifted Windows”, Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), 2021, pp.10012-10022
- [24] Ailiang Lin, Bingzhi Chen, Jiayu Xu, Zheng Zhang, Guangming Lu, “DS-TransUNet: Dual Swin Transformer U-Net for Medical Image Segmentation”, IEEE Transactions on Instrumentation and Measurement, Volume: 71, 30 May 2022, 4005615
- [25] Xuebin Qin, Zichen Zhang, Chenyang Huang, Masood Dehghan, Osmar R. Zaiane, Martin Jagersand, “U²-Net: Going Deeper with Nested U-Structure for Salient Object Detection”, Pattern Recognition, Volume 106, October 2020, 107404
- [26] Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun, “Deep residual learning for image recognition”, Proceedings of the IEEE conference on computer vision and pattern recognition, pp.770-778, 2016
- [27] Olaf Ronneberger, Philipp Fischer, Thomas Brox, “Unet: Convolutional Networks for Biomedical Image Segmentation”, Medical Image Computing and Computer-Assisted Intervention - MICCAI 2015, pp.234-241
- [28] Jonathan Long, Evan Shelhamer, Trevor Darrell, “Fully Convolutional Networks for Semantic Segmentation”, Computer Vision and Pattern Recognition, 2015, pp.3431-3440
- [29] Hui Kong, Hatice Cinar Akakin, Sanjay E. Sarma, “A Generalized Lapla-

- cian of Gaussian Filter for Blob Detection and Its Application”, IEEE Transactions on Cybernetics, Volume:43, Issue:6, pp.1719-1733, 09 January, 2013
- [30] Bradski Gary, “The openCV library.”, Dr. Dobb’s Journal: Software Tools for the Professional Programmer, vol 25, pp.120-123, 2000
- [31] Diederik P. Kingma, Jimmy Ba, “Adam: A Method for Stochastic Optimization.”, ICLR(Poster), 2015
- [32] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer et.al, “Pytorch: An Imperative Style, High-Performance Deep Learning Library”, NeurIPS 2019, pp.8024-8035, 2019
- [33] Fausto Milletari, Nassir Navab, Seyed-Ahmad Ahmadi, “V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation”, 2016 Fourth International Conference on 3D Vision, pp.565-571, 2016
- [34] Christopher K. I. Williams, Carl Edward Rasmussen, “Gaussian Process for Regression”, Advances in neural information processing systems, 1995, vol8
- [35] Rasmussen, Carl Edward, “Gaussian processes in machine learning”, Summer school on machine learning, pp.63-71, 2004
- [36] 持橋大地 , 大羽成征, 『ガウス過程と機械学習』 , 講談社, 2019