

卒業論文
Bachelor Thesis

カプセルネットワークを用いた
CT画像からの脂肪領域抽出
Segmentation of Fat Region on CT Images
using Capsule Network

指導教官 杉本千佳准教授
Advisory Professor Chika SUGIMOTO

平成31年3月13日
February 27th, 2019

横浜国立大学 理工学部 数物電子情報系学科
Department of Mathematics, Physics & Computer Engineering
College of Engineering Science, Yokohama National University

1564301

竹下 智章
Tomoaki TAKESHITA

あらまし

医療現場において、近年の人工知能技術、取り分け深層学習の発達をもとに、ニューラルネットワークをはじめとする深層学習を用いた画像処理技術により医用画像診断を支援する研究が多く為されている。医用画像の深層学習では、特定の臓器や器官などに対応する領域を画像上で色分けするセグメンテーション (抽出) の機械学習手法が用いられる。医用画像のセグメンテーションは、医師の読影を補助する、あるいは画像内から特定領域を抽出し、別の診断アーキテクチャが診断を行うために前処理と画像のマスキングを行うなどの利点があるため、様々な手法が研究されている。

医用画像の診断でしばしば課題になるのが、組織の形状や位置などの医用画像間での特徴差である。身体の断層面の画像である CT や MR の医用画像において、部位によっては、撮影する身体上の位置がわずかに移動するだけで画像内の組織の特徴が大きく変化する。つまり、組織同士の形状および位置関係に差異がある状態となる。よって、このような部位におけるニューラルネットワークを用いた画像診断では、特徴差を持った医用画像という条件下で全ての画像に対し同一の高い精度で抽出ないしは検出を行うことが求められる。

さらに深層学習では、分類や抽出などのタスクの特徴を学習する際に大量のデータが必要となる。データ量の確保が重要となるが、症例数の少なさや、医師等の専門家による注釈 (アノテーション) 付きのデータを大量に得るコストなど観点から必ずしも多くのデータを得られるとは限らないのが実情である。

そこで本研究では、撮影位置の移動で組織の形状が大きく変化し、かつ組織の位置関係に個人差がみられる部位の医用画像に対して、高精度のセグメンテーションが可能なモデルを導出することを目的とする。また、そのモデルが、学習サンプルのデータ量が少量の段階においても特徴を学習可能であり、未学習データに対して高精度に抽出が可能であることを示す。

本論文では、画面内の物体の構造や位置関係の認識と学習に優れた性質を持つカプセルネットワークを用いたモデルを、画像同士特徴の差があり、データ量の少ない医用画像のセグメンテーションに適用する。学習するデータに関するこれらの条件下において、カプセルネットワークの特徴たるカプセルレイヤーによる空間的な特徴の認識が、ピクセルレベルでの従来型の学習手法と比較して高い精度での領域抽出を可能とすることを示す。

目次

第1章	序論	1
1.1	研究背景と目的	1
1.2	本稿の内容と構成	3
第2章	カプセルネットワークを用いた抽出手法	4
2.1	カプセルネットワーク	4
2.2	本研究における領域抽出モデル	6
第3章	評価実験	8
3.1	データセット	8
3.2	評価方法	8
3.3	tiramisu との比較実験	9
3.4	実験結果	9
3.5	考察	11
第4章	結論	14
4.1	まとめ	14
4.2	今後の課題	14

目 次

1.1	本稿の流れ	3
2.1	カプセルネットワークの構造	4
2.2	カプセルネットワークにおける画像の再構成	5
2.3	SegCaps の構造	6
3.1	データセットのサンプル	8
3.2	tiramisu の構造	10
3.3	セグメンテーションのサンプル	12
3.4	モデルの学習状況	13

表 目 次

3.1	評価指標	9
3.2	各モデルのパラメータ数 (入力サイズ 512×512 の場合)	9
3.3	実験結果	11

第1章 序論

1.1 研究背景と目的

医療現場において、CT (コンピュータ断層撮影法) や MRI (磁気共鳴映像法) のように特性の異なる複数の撮影法 (モダリティ) を使い分けることや、両者を総合的に加味することでの診断を行うことは、症状の分析タスクにおいて重要となる。これに対し、近年の人工知能技術、取り分け深層学習の発達をもとに、画像処理技術により医用画像診断を支援する研究が多く為されている [1]。その内容は専門家の判断を自動化するもの [2] から、専門家でも客観的判断が困難な症例を定量的に評価するもの [3] まで多岐にわたる。近年では、人工知能を用いた画像処理分野において、ニューラルネットワークの一種である CNN (畳み込みニューラルネットワーク) を派生させた研究が盛んであり、これにより医用画像の解析分野においても CNN を応用した研究が数多く為されている [4]。

医用画像に対する深層学習を用いた診断法において、特定の臓器や器官などに対応する領域を画像上で色分けする目的でセグメンテーション (抽出) の機械学習手法が用いられている。医用画像のセグメンテーションは、関心領域提示による医師への診断補助や、自動鑑別における特定領域の自動抽出においても必要であるため、様々な手法が研究されている。近年の CNN の派生技術を用いた研究では、位置合わせを行った複数のモダリティの医用画像の特徴を CNN で抽出して統合を行う研究 [7] や、GAN (敵対性生成ネットワーク) を利用することで元の医用画像からセグメンテーションマスク画像を生成するという研究 [8] が医用画像セグメンテーションの分野で為されている。

さて、医用画像の診断をニューラルネットワーク等の深層学習によって行う上でしばしば問題になるのが、組織の形状や位置などの医用画像間での特徴差である。身体断面の画像である CT や MR の医用画像において、部位によっては、撮影する身体上の位置が移動することで組織の特徴が大きく変化する。特徴差が少なくなるように医用画像を得られるのであればそれが望ましいが、断面の位置変動による特徴変化の大きな部位の医用画像の撮影の際に、特定の形状となるような特定の場所に撮影断面を合わせることは困難である。また、このように位置による特徴差が大きな部位においては、各組織の構造や位置には個人差があるため、異なる患者の医用画像間で位置を対応づけることも難しい。つまり、組織同士の形状および位置関係に差異がある状態となる。よって、このような部位における深層学習を用いた画像診断では、特徴差を持った医用画像という条件下で全ての画像に対し

同一の高い精度で抽出ないしは検出を行うことが求められる。

さらに深層学習では、分類や抽出などのタスクの特徴を学習する際に大量のデータが必要となる。そこで、一定量まで人力で用意したデータセットを複製し、特定のデータの変形や差し換えを行うことでデータを水増しするデータ拡張を行うことが一般的である。画像処理の場合では、画像の拡大・縮小および回転などの変形を行うことでデータ拡張が為される。しかし、データ拡張はデータ量を増やすことができるが、特徴が元のデータと似てしまう傾向がある [9]。これにより、拡張を行ったデータで学習を行った際に、元データと異なる性質を持つ未知のデータに対して精度が下がってしまう (過学習を起こす) 可能性がある。これは個人差や断層面位置によって身体中の組織の構造や位置関係が異なる医用画像においては特に顕著になると考えられる。つまり、データ拡張による画像データの増しが学習データに無い特徴を持った患者の医用画像の診断精度を下げてしまう可能性が考えられる。よって、医用画像の深層学習による診断の分野ではデータ量の確保が重要となるが、症例数の少なさや、医師等の専門家による注釈 (アノテーション) 付きのデータを大量に得るコストなど観点から必ずしも多くのデータを得られるとは限らないのが実情である [9]。

上記を踏まえて、医用画像のセグメンテーションにおいて深層学習による判断を行う際、その学習モデルに求められる性質は、画像間で組織の形状やそれらの位置に差異があるデータに対して空間的な特徴を学習し、一様に高い精度で領域の抽出を行うことができる性質、および学習に用いるデータの少なさに対しても耐性を持ち、データ量が少ない段階においても高い精度の抽出を行うことができる性質である。

そこで本研究では、わずかな撮影位置の移動で組織の形状が大きく変化し、かつ組織の位置関係の個人差が大きな部位の医用画像に対して、高精度のセグメンテーションが可能なモデルを導出することを目的とする。また、そのモデルが、学習サンプルのデータ量が少量の段階においても特徴を学習可能であり、未学習データに対して高精度に抽出が可能であることを示す。

従来型の CNN およびその派生系のアーキテクチャを用いたモデルの場合、画素同士の強度と互いの位置関係といった画素特徴 (ピクセルレベル特徴) を抽出する傾向にある [2]。これらのモデルを組織の形状や位置の特徴に差がある画像のセグメンテーションに適用した場合、ピクセルレベルの学習では単純な画素の位置関係を学習してしまい、精度が低くなってしまう。このようなモデルでは、データ量を多く確保することや、ネットワーク中に含む層の数や深さを増やす (パラメータ数を多くする) ことで高い精度を実現する [11] ことが可能である。すなわち、抽出する特徴の数を増やすことで精度を高めようとする試みである。しかし、データ数が少量に限られるため単純なデータ量の増加は今回の研究の目的にそぐわない。また、パラメータ数を増加させる場合、膨大な数のパラメータを調整する必要があり、処理を行うコンピュータの性能要求が非常に高くなる。つまり、この研究に用いるべきモデルは単純な画素の配置ではなく、それよりも上位の概念たる組織の特徴

や, さらに上位となるそれらの位置関係や大きさの比率などの空間的特徴を認識および学習することのできるモデルである. そこで, 今回の研究では, 画面内の物体の構造や位置関係の認識と学習に優れた性質を持つカプセルネットワーク [10] を用いたモデルを, 画像同士特徴の差があり, データ量の少ない医用画像のセグメンテーションに適用する. 学習するデータに関するこれらの条件下において, 本モデルの特徴たるカプセルレイヤーによる空間的な特徴の認識および学習が, ピクセルレベルでの従来型の学習手法と比較して高い精度での領域抽出を可能とすることを本研究では示す.

1.2 本稿の内容と構成

本稿の内容は図 1.1 のようになる。

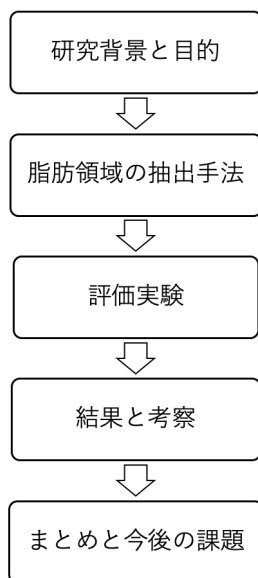


図 1.1: 本稿の流れ

第2章 カプセルネットワークを用いた抽出手法

本研究では、医療用画像撮影装置によって撮影された 2D の骨盤領域画像から、脂肪領域を抽出する手法を構築する。スライス位置によって画像上の形状が変化する脂肪領域について、カプセルネットワークによるセグメンテーションアーキテクチャに学習させることによって抽出を行う。本項では、元となるカプセルネットワークおよび、脂肪領域抽出用に用いるセグメンテーションモデルについて説明する。

2.1 カプセルネットワーク

カプセルネットワークとは、2017 年に Hinton らによって発表された論文 [10] で登場したニューラルネットワークの新技术である。通常のニューラルネットワークのニューロンをカプセルという新しい概念に置き換えることで、(1) 空間的特徴の抽出に優れる、(2) データのクラス不均衡に頑健、(3) データ量の少なさに耐性がある、などの性質を持つとされる。

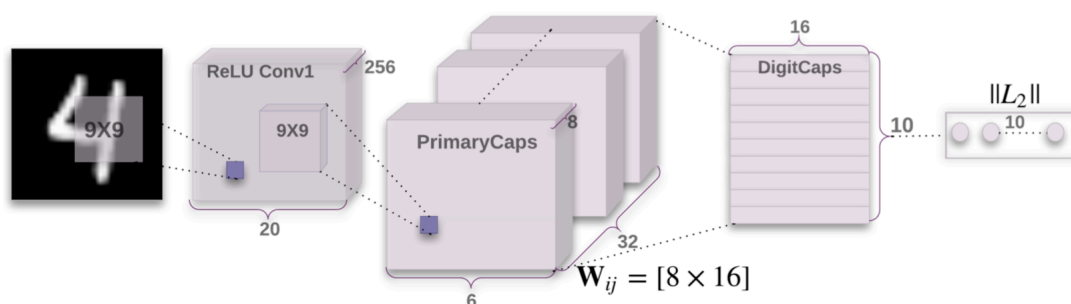


図 2.1: カプセルネットワークの構造

このネットワークでは、通常の CNN において一般的に用いられる畳み込み層、プーリング層、完全結合層に対し、プーリング層を廃止し、完全結合層の概念の拡張が行われている。図 2.1 は手書き数字のデータセットである MNIST に対し、該当する数字に分類を行うモデルである。この例を用いて説明する。最初に入力画像

に対して 2D 畳み込みを行うまでは通常の CNN と同じであるが、一般的には畳み込み層の次で挿入されるプーリング層がカプセルネットワークの場合は存在しない。プーリング層は通常、畳み込み層で得られた画素の位置に対して過学習を防ぐ目的と、画面サイズを小さくしてパラメータ数を少なくする目的で用いられる。しかし、プーリング層は画素の画面内での位置への依存を緩和し過学習を防止すると同時に、位置関係の情報を失ってしまうとされる。通常の CNN アーキテクチャをもつモデルでは、このトレードオフを畳み込み層とプーリング層の組み合わせで同時に解決することが困難であるため、複雑な位置関係の学習を行う場合はこれらの層を組を増やしたり、ネットワークの分岐と合流を行うなどの試みで対応が為されている [11]。カプセルネットワークは後述するカプセルによってプーリング層に関するトレードオフを同時に満足する性質を持つため、プーリング層を廃することが可能となる。

次に、第 2 の畳み込みと同時に、テンソルは多次元のベクトルを持つ形に変形され、最初のカプセル層 (図中の PrimaryCaps 層) に配される。カプセル層ではこの多次元ベクトルが単位となり、通常のニューラルネットワークにおけるニューロンに代わってカプセルという概念として導入される。この層は従来型のネットワークにおける完全結合層に相当するが、通常の完全結合層におけるニューロンはスカラーである。これが通常の CNN との大きな違いである。カプセル内のベクトル成分は、画像中に存在する特定のエンティティの様々な特性を表す。これらには、位置、サイズ、方向、色相、テクスチャなど、さまざまな種類のパラメータが含まれる (パラメータが何の特徴を表しているかは具体的にはわからない)。カプセル層は通常の完全結合層のように層を深くすることが可能で、入力側に近いほど物体の大きさや形状などの下位の概念を表し、出力側に近いほどそれらの位置関係や比率などの空間の全体的な特徴を表すようになる。このカプセル層によってピクセルレベルの学習よりも上位の概念となる空間的特徴の認識と、単純な画素の画面内位置への依存の防止を同時に満足することが可能となる。カプセル層の最終層 (図中では DigitCaps 層) では 1 つのカプセルが確率ベクトルを表し、最終的な出力はこの最終カプセル層をスカラーに変換することで得られる。

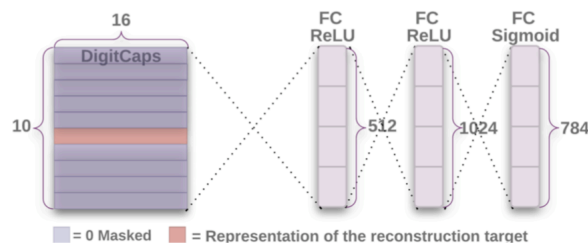


図 2.2: カプセルネットワークにおける画像の再構成

カプセルネットワークの特徴として、図 2.2 のように、最終カプセル層のカプセル

ルより元画像の復元を行い、その復元性も学習の評価基準とするとともに、復元に用いる従来型の完全結合ネットワークも同時に学習を行う点がある。つまり、入力層から最終カプセルをエンコーダとみなした場合におけるデコーダも同時に学習を行う。

2.2 本研究における領域抽出モデル

上記のようなカプセルネットワークの構造を持ち、その空間的特徴の認識が可能となる性質を持ったセグメンテーションモデルを本研究では用いる。SegCaps は、カプセルネットワークをもとに [12] で考案されたモデルである。

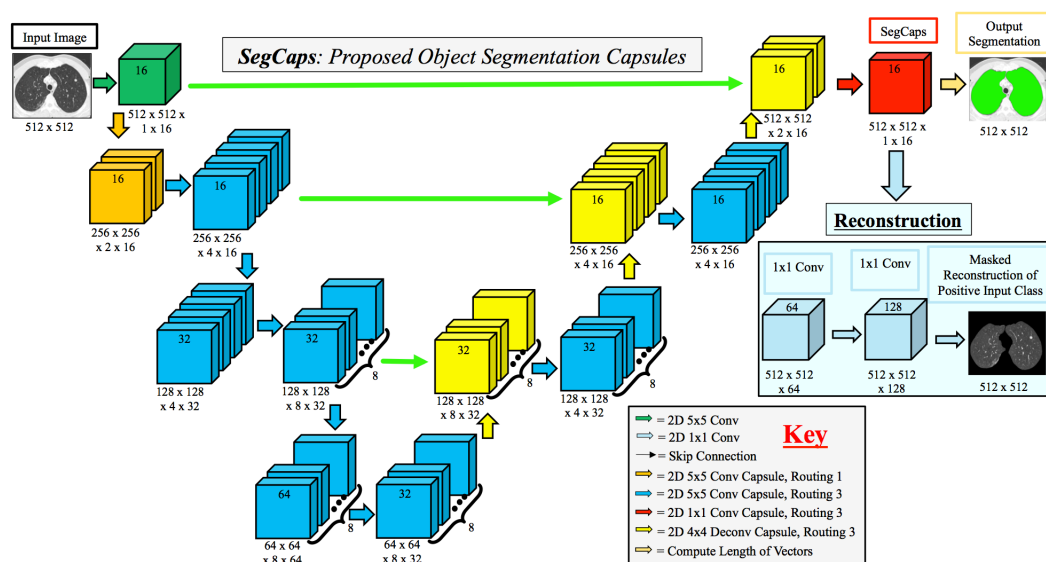


図 2.3

SegCaps は、[10] のカプセル層がカプセル 1 つを単位とする全結合層であったのに対し、CNN の畳み込み層がニューロンを 2 次元に配置するように、カプセルを 2 次元に配置する。ある層の中のカプセルは、前の層で近接するカプセルのみと結合されて畳み込まれる。すなわち、ニューロンをカプセルに置き換えた畳み込み層である。カプセルによる完全結合層を畳み込み層 (ConvCapsule) に置き換えることで、大幅にパラメータ数を少なくしながらも空間的特徴の抽出を可能としている。SegCaps は、この畳み込みカプセル層をセグメンテーションモデルである U-Net[13] の畳み込み層に対応する位置に配置した構造になっている。U-Net のように畳み込み層とアップサンプリング層で構成され、入力層側から出力層側の同じ段階の層にスキップコネクションが行われている。これにより、U-Net のように入力画像中の対応する位置を維持しながら、空間的特徴を学習することが可能なカプセルネットワークの性質を持ったセグメンテーションモデルが実現される。最終

カプセル (図中で SegCaps 層) は入力画像と同じサイズにカプセルが並べられ, 多次元確率ベクトルであるカプセル 1 つは, 入力画像にに対応する位置でのマスクの存在確率を表している. この SegCaps 層での確率ベクトルを sigmoid 関数でスカラー量の確率に変換される. この確率を一定の閾値でマスクが存在するか否かのバイナリ情報として表したマップが入力画像に対する予測マスクとなる. SegCaps は畳み込みカプセル層を採用することにより, 512×512 の入力の場合, 同一の深度を持つ U-Net よりもパラメータ数が $1/30$ ほどに抑えられる. SegCaps においても最終カプセル層から入力画像を復元するが, [10] では完全結合層で入力画像への再構成が行われていたのに対し, 畳み込み層によって復元される構造になっている. これによって, パラメータ数が大幅に抑えられている.

ここで, 少量のデータを学習に用いた際に生じる問題がある. ニューラルネットワークの学習では, 学習に用いるトレーニングデータ, 精度の評価に用いるテストデータの他に, 学習の状況を検証するためのバリデーションデータが用いられる. バリデーションとは, 学習を開始する前にランダムで選んだ学習に用いないデータ (テストデータよりも少量) を用いて, 学習の 1 エポックごとに精度を検証することである. これにより, モデルが学習データに対して最適化されすぎてしまう過学習をしていないかを, 未学習のデータで逐一評価しながら学習の進行を監視することができる. トレーニングデータでの精度に対してバリデーションデータでの精度が低くなる場合はモデルが過学習しているということになる. 今回の研究のように, 少量の画像データで学習をする場合, バリデーションデータの枚数は数枚とかなり少数になる. このバリデーションデータを用いて学習状況の検証を行う際, バリデーションデータの選び方によっては画像間の特徴が似たようなものに偏りやすく, 場合によってはトレーニングデータ以上に適合してしまう可能性があり, 指標として正しく機能しない. そこで本研究では, 通常, バリデーションデータには行わないデータ拡張を施すことで擬似的にバリデーションデータに多様性を持たせる. これにより, 未学習データへの精度が過大評価されることを抑制し, 過学習防止の指標として正しく機能することを狙った. 具体的には, バリデーションステップごとにランダムで選出されたバリデーション画像に, さらにランダムで回転や平行移動, 拡大等のアフィン変換を施し, バリデーションに用いた.

第3章 評価実験

3.1 データセット

評価実験には, 組織の形状や位置に差のある医用画像として, 骨盤領域を撮影した CT 画像 5 人分 (48 枚) に対し, 脂肪領域を塗り分けたマスク画像のセットを用いた. 各画像の身体上における撮影間隔 (スライス間隔) は 2.5mm で, 1 人分は最少で 8 枚, 最多で 11 枚の画像で構成される.



[1] CT 画像



[2] マスク画像

図 3.1: データセットのサンプル

3.2 評価方法

入力された CT 画像からモデルによって予測されたマスク画像と, 正解のマスク画像の Dice 係数と Jaccard 係数を用いて評価を行う. Dice 係数と Jaccard 係数はどちらも集合の類似性の評価関数である. それぞれの式は, 類似性を比較する集合 (この場合では抽出された領域と正解の領域) を X, Y としたときに表 3.1 に表され, その範囲は $0 \sim 1$ となる. Dice 係数は形状の類似性におけるヒトの判断感覚に近いとされる. Jaccard 係数は差集合 (片方の領域にしか含まれない領域) が影響しやすく, 無関係の領域を抽出した場合に数値が小さくなる.

表 3.1: 評価指標

Coefficient	
Dice	$\frac{2 X \cap Y }{ X + Y }$
Jaccard	$\frac{ X \cap Y }{ X \cup Y }$

3.3 tiramisu との比較実験

同一のデータセットを用いて既存のニューラルネットワークを用いたセグメンテーションモデルである, tiramisu[14] との比較実験を行った. tiramisu は, 図 3.2 のように U-Net を派生させたモデルで, Fully Connected Network(FCN) 型のモデルである. FCN とは, CNN の完全結合層を 1×1 の畳み込み層で置き換えたモデルで, tiramisu はこれにより U-Net よりも大幅にパラメータ数を抑えながら高精度な抽出が行えるように改良したものである. 512×512 の入力画像に対し, エンコーダとデコーダで構成される層を SegCaps と同じ深さ 3 層になるモデルを構築した際, パラメータ数が SegCaps の 1.5 倍程度である. よって, パラメータ数が同程度で, U-Net ベースの類似の抽出手法として精度の比較対象モデルとした.

表 3.2: 各モデルのパラメータ数 (入力サイズ 512×512 の場合)

Model	Pram.
SegCaps	1,425,649
tiramisu	2,314,208

モデルの学習には 3 人分のデータ (約 30 セット), バリデーションにはそこから 3 セットをランダムに取り出して用いた. また, テストデータは残る 2 人分 (約 20 セット) とした. 本評価実験においてはデータセットの量が少ないため, それらの組み合わせを 3 通り (トレインデータ: テストデータのセット数がおおよそ 3:2 になるように調整した) で変えた Split による交叉検証を行った.

3.4 実験結果

表 3.3 に各モデルによる各 Split のデータでの抽出領域と正解領域の形状類似性評価係数と, 全 Split の平均数値を示す. また, 各モデルによる抽出領域の比較サン

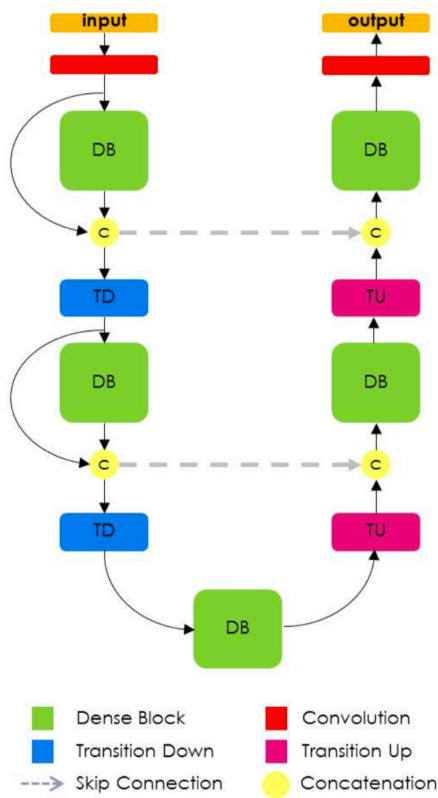


図 3.2: tiramisu の構造 : 2つのトランジションダウンのあるダウンサンプリング行程と2つのトランジションアップのあるアップサンプリング行程で構成される. ダウンサンプリング経路では, 密ブロックへの入力とその出力と連結される.

プル画像を図 3.3 に示す.

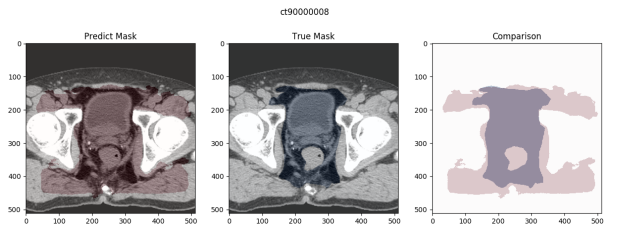
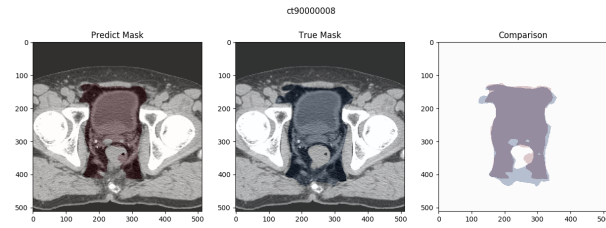
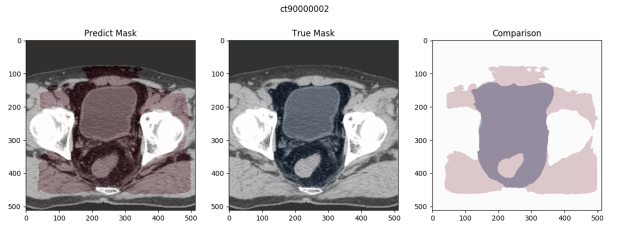
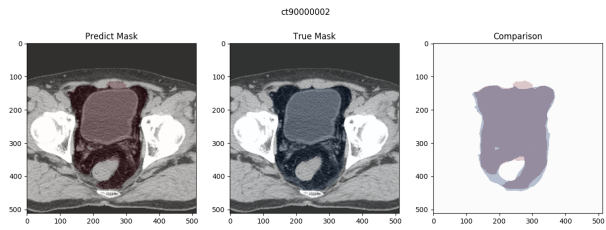
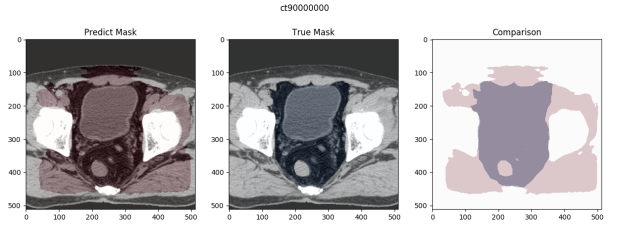
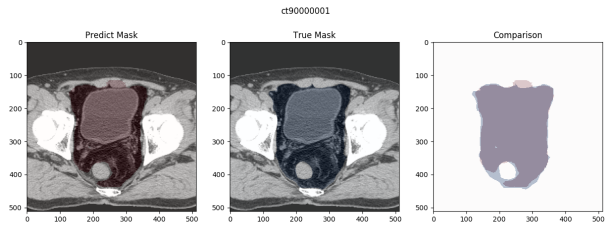
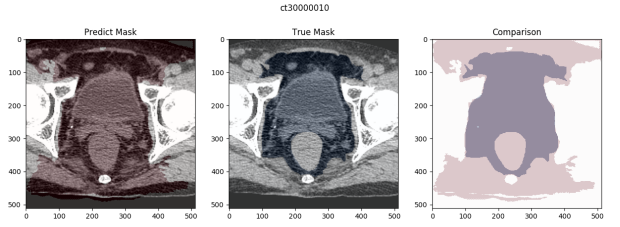
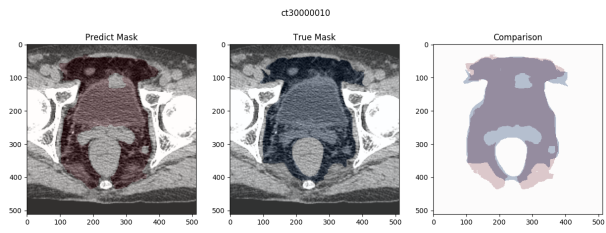
表 3.3: 実験結果

	Split	Dice (%)	Jaccard(%)
SegCaps	Split-1	92.10	85.42
	Split-2	86.52	76.48
	Split-3	90.63	83.56
	平均	89.75	81.82
tiramisu	Split-1	71.15	56.27
	Split-2	75.17	60.94
	Split-3	74.59	59.48
	平均	73.63	58.89

3.5 考察

いずれの split でも Dice 係数, Jaccard 係数ともに SegCaps を用いた場合の精度が高いという結果となった. これはセグメンテーション結果のサンプル画像を見ても明らかで, SegCaps による抽出が概ね対象の領域を抽出できているのに対し, tiramisu による抽出では無関係の領域を抽出している. Jaccard 係数は対象の領域を抽出しきれない場合よりも, 無関係の領域を多く抽出してしまった場合に低い値となるため, 2 モデル間では tiramisu が Dice 係数よりも Jaccard 係数において大きく劣る結果と合致している. ここで, 各モデルの学習状況 (図 3.4) を見ると, tiramisu ではトレーニングの精度が 2 エポック程度で高精度に収束するが, バリデーションの精度が低く, 安定しないことから過学習を起こしていることが分かる. 一方で, SegCaps はトレーニング精度とバリデーション精度ともに同程度に収束することから, 少量のデータに対しても過学習を起こすことなく, 高精度に学習することができていることが分かる. これらの学習状況からも, SegCaps が少量の学習データでも未学習データへの対応性を維持しながらも高精度の学習が可能であることが示されている.

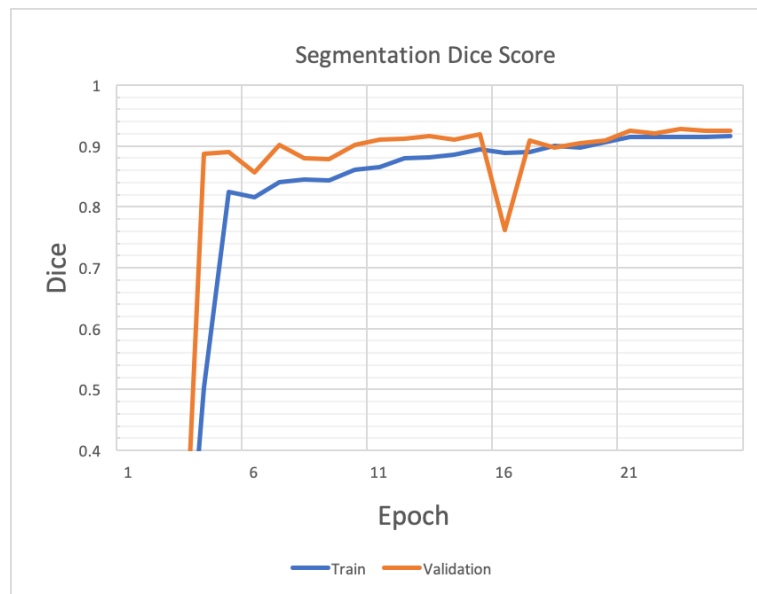
本実験で抽出対象となる脂肪領域は, 黒く見える領域と灰色に見える領域を含んでいるが, 画像下部の灰色の領域が除外すべき領域として対象領域に含まれていない. つまり, 画素強度的には同様の領域が対象領域と対象外領域に含まれていることになる. 従来型の手法である tiramisu は, 少量のデータによる学習では, 除外すべき領域と含めるべき領域の区別できず, 画素特徴を学習していることが疑われる. 一方で, SegCaps による学習では, 同程度の画素強度の領域でも区別ができおり, 周辺の組織の位置関係等の空間的特徴を学習していると考えられる.



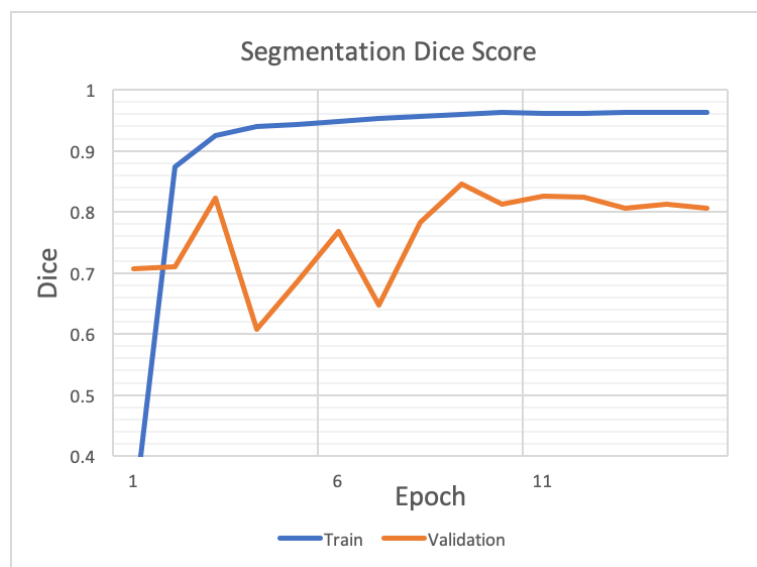
[1] SegCaps

[2] tiramisu

図 3.3: セグメンテーションのサンプル：医用画像中に赤の領域で表される領域抽出結果と、青の領域で表される正解の領域比較である。



[1] SegCaps の学習状況



[2] tiramisu の学習状況

図 3.4: モデルの学習状況：モデルの学習時における各エポック終了時のトレーニング精度 (青) とバリデーション精度 (橙) の推移

第4章 結論

4.1 まとめ

カプセルネットワークを用いたモデルにより、画像間で組織の形状や位置に差異があり、かつデータ量が少量の場合において、ピクセルレベルの学習モデルと比較して高い精度での脂肪領域抽出が可能であることが示された。

4.2 今後の課題

今回の研究においては、少ないデータ条件でカプセルネットワークによって脂肪領域という形状が変化する領域単体が高精度で抽出されることが示された。しかし、モデルが空間的特徴を捉えることが少量のデータで高精度に抽出できたことに寄与しているかは疑問の残る結果であった。SegCaps は現状、該当する領域であるか否かの2クラス分類しか行うことができない。これをマルチクラスの分類ができるように改良を加えることができれば、互いの位置関係を学習しながら複数種類の領域を同時に抽出することが可能となり、位置関係の理解が少量データの学習に寄与していることを示すことができると考えられる。

謝辞

本研究を進めるにあたり, 年間を通じてご指導いただきました杉本千佳准教授に深く御礼申し上げます. また, 論文の執筆に際して助言をくださった研究室の皆様, そして進学のお機会を与えてくださった両親へ感謝いたします.

参考文献

- [1] R Aarthi, S Harini :“A Survey of Deep Convolutional Neural Network Applications in Image Processing”, International Journal of Pure and Applied Mathematics, Vol. 118, No. 7, pp. 185-190, 2018
- [2] Ozan Oktay et al. :“Anatomically Constrained Neural Networks (ACNNs): Application to Cardiac Image Enhancement and Segmentation”, IEEE Transactions on Medical Imaging, Vol. 37, No. 2, pp. 384-395, 2018.
- [3] Qi Dou et al. :“Automatic Detection of Cerebral Microbleeds from MR Images via 3D Convolutional Neural Networks”, IEEE Transactions on Medical Imaging, Vol.35 , Issue 5, 2016.
- [4] Geert Litjens et al. : “A survey on deep learning in medical image analysis”, Medical Image Analysis, Vol. 42, pp. 60-88, 2017
- [5] Guotai Wang et al. :“Interactive Medical Image Segmentation Using Deep Learning With Image-Specific Fine Tuning”, IEEE Transactions on Medical Imaging , Vol. 37, Issue 7, pp. 1562-1573, 2018.
- [6] Fausto Milletari et al. :“V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation”, 2016 Fourth International Conference on 3D Vision (3DV), 2016.
- [7] Zhe Guo et al. :“Medical Image Segmentation Based on Multi-Modal Convolutional Neural Network:Study on Image Fusion Schemes”, IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), pp. 903-907, 2018.
- [8] Yuan Xue et al. :“SegAN: Adversarial Network with Multi-scale L1 Loss for Medical Image Segmentation”, Neuroinformatics, Vol. 16, Issue 3-4, pp. 383-392, 2018.
- [9] Blaine Rister et al. :“CT organ segmentation using GPU data augmentation, unsupervised labels and IOU loss”, 2018.
- [10] Geoffrey E. Hinton et al. :“Dynamic Routing Between Capsules”, Advances in Neural Information Processing Systems 30 (NIPS 2017), 2017.

- [11] Christian Szegedy et al. : “Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning”, Thirty-First AAAI Conference on Artificial Intelligence, 2017.
- [12] Rodney LaLonde, Ulas Bagci : “Capsules for Object Segmentation”, 1st Conference on Medical Imaging with Deep Learning (MIDL 2018), 2018.
- [13] Olaf Ronneberger et al. : “U-Net: Convolutional Networks for Biomedical Image Segmentation”, 2015.
- [14] Simon Jegou et al. : “The One Hundred Layers Tiramisu: Fully Convolutional DenseNets for Semantic Segmentation”, The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, pp. 11-19, 2017