

神経細胞セグメンテーションにおける 深層畳み込みアーキテクチャの適用方法に関する検討

高屋 英知^{1,a)} 竹市 裕介² 尾崎 まみこ² 栗原 聡^{1,b)}

概要: コネクティクスと呼ばれる分野では、電子顕微鏡連続切片画像の3次元再構築法により、神経系のミクロな構造の観察がなされている。しかし、再構築のためのラベリングを人手により行うことは多大なコストを要するため、近年では深層学習を用いたセグメンテーション手法が注目を浴びている。しかし、深層学習の適用においては、他のタスクと同様に、ハイパーパラメータの探索が課題となる。本研究では、画像セグメンテーションを行うために設計された Deep Contextual Network と呼ばれる深層畳み込みアーキテクチャの構造に注目し、特徴マップの可視化に基づいたハイパーパラメータの調整手法を検討する。具体的には、学習済み DCN の特徴マップを可視化し、ネットワークにおいて適切に働いていない箇所を目視により特定し、対応する逆畳み込みネットワークのフィルタサイズを拡大あるいは縮小して再学習を行う。2種類のデータセットを用いた実験の結果、フィルタサイズを変更することによる特徴マップの変化の様子が明らかになったとともに、提案手法によりセグメンテーションの精度が向上することが示された。

A Study of Method to Apply Deep Convolutional Architecture to Neuronal Structure Segmentation

TAKAYA EICHI^{1,a)} TAKEICHI YUSUKE² OZAKI MAMIKO² KURIHARA SATOSHI^{1,b)}

1. はじめに

生物の脳や感覚器における神経系の構造や接続状態を詳細に調べることは、それらの機能的意味を理解することにつながる。このような研究分野はコネクティクスと呼ばれ、連続切片画像の取得をはじめとした電子顕微鏡の技術とともに発展してきた。たとえば図1は、クロオオアリの単仲間識別感覚子の微細構造を3次元空間に描画したものである。多数の神経が束になっており、ところどころに瘤状構造が観察できる。

実際にこのような観察を行うためには、Serial block-face

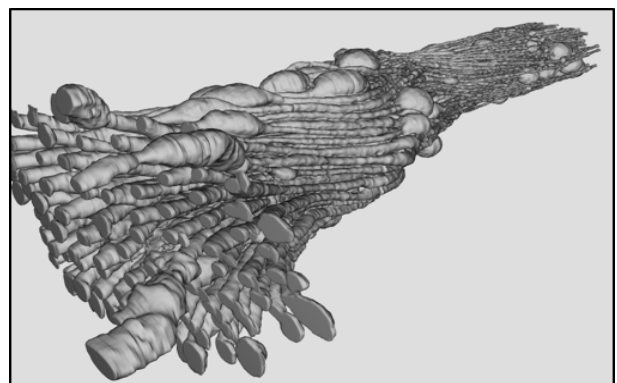


図1: クロオオアリ単仲間識別感覚子

¹ 電気通信大学 大学院情報理工学研究科
Graduate School of Informatics and Engineering - The University of Electro-Communications, Chofu, Tokyo 182-8585, Japan

² 神戸大学 大学院理学研究科
Graduate School of Science - Kobe University, Kobe, Hyogo 657-0013, Japan

a) etakaya@ni.is.uec.ac.jp

b) kuri@acm.org

scanning electron microscopy (SBF-SEM) などの電子顕微鏡によって高解像度の連続切片画像を取得し、それらを3次元に再構築する必要がある。しかし、それには数百枚にわたる画像に対して図2のように神経細胞のラベリングを行う必要があり、全てを人手によって行うことは膨大な時間と労力を要する。Kaynigら [1] は、これを自動化するた

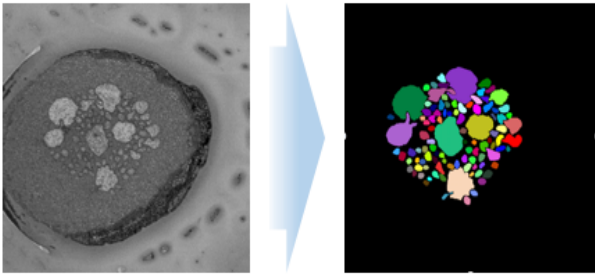


図 2: 神経細胞のラベリング

めの画像処理パイプラインを提案しており、そこでは神経細胞のセグメンテーションと、3次元構造における同一領域のグルーピングが段階的に行われる。このパイプライン方式では、特に最初のステージの精度が、後のステージに大きな影響を及ぼすため、より重要度が高いと考えられる。したがって、本研究では最初のステージである神経細胞セグメンテーションに焦点を当てる。

近年では、神経細胞セグメンテーションの精度を向上させるために、深層学習に基づく手法が数多く提案されている [2][3][4]。しかし、このタスクを解くために提案された深層学習アーキテクチャは、特定のデータセットで高精度なセグメンテーションを行うためのチューニングがなされている。そのため、他のデータセットに適用する際には、学習によって獲得されるパラメータとは別に、事前に決めておくべきハイパーパラメータを適切に設定する必要がある。一般に、ハイパーパラメータを決定するためには、ランダムサーチやグリッドサーチ、あるいは強化学習などを用いて適切な組み合わせを探索するが、深層学習では1度の計算にかかる時間が長いので、何度も繰り返し計算し直すのは多大な時間的コストがかかる。そこで本稿では、神経細胞セグメンテーションのタスクにおいて、ネットワークのアーキテクチャの特性に基づき、特定のハイパーパラメータを効率的に決定するための方法を検討する。

以下では、第2節にて深層学習を用いた神経細胞セグメンテーションについて概観し、第3節では特に関連タスクで精度が高いとされる Deep Contextual Network について述べる。第4節では、Deep Contextual Network の構造に基づいたハイパーパラメータの決定方法を検討し、第5節にて実験結果を報告する。第6節では、まとめと今後の展望について述べる。

2. 深層学習による神経細胞セグメンテーション

神経細胞セグメンテーションの問題に対し、深層学習に基づく手法が初めて提案されたのは、2012年であった。Ciresanら [2] は、畳み込みニューラルネットワークを用いて画像のピクセル毎に分類問題を解く方法により、従来手法を大きく上回った。こうして、深層学習による精度の高

さが浮き彫りになって以来、新たに提案される手法のほとんどは深層学習に基づいたものとなっている。そして、現在は Fully Convolutional Network [5] の登場により、それを基にした U-Net [4] や Deep Contextual Network [3] など、さらに精度の高いアーキテクチャも提案されている。本研究では、特に高い精度が報告されている Deep Contextual Network に焦点を当て、その精度をさらに引き出すための方法を検討する。

3. Deep Contextual Network

Deep Contextual Network (以下 DCN) は Fully Convolutional Network (以下 FCN) を拡張した畳み込みニューラルネットワークである。物体認識等で用いられる深層学習のアーキテクチャは通常、複数の畳み込み層と全結合層によって構成されているが、FCN は「全畳み込みネットワーク」とも訳されるように、すべての層が畳み込み層で構成されている。これにより、出力の次元を入力と等しくし、ピクセル毎の回帰および分類を行うことが可能になった。また、Nohら [6] は FCN でのプーリングによって小さくなった特徴マップのサイズを元に戻す際、逆畳み込み層と呼ばれるモジュールを導入している。

このような背景を受けて、Chenら [3] は図3に示す DCN を提案した。このアーキテクチャは主に2つのモジュールで構成されている。一方は畳み込み層とプーリング層によるダウンサンプリング部で、もう一方は畳み込み層と逆畳み込み層によるアップサンプリング部である。前者には抽象的な情報にもとづいて物体領域を認識する働きがあり、後者には境界線のような細かい部分を再構成する働きがある。

DCN で特に工夫がなされているのは、段階的に得られた特徴マップを最終的に足し合わせている点である。Chenらによれば、深い層で得られた大域的な情報を分類に用い、浅い層で得られた局所的な情報を位置特定に用いることを意図している [3] ののだという。たとえば後述するデータセット (図5) にみられる神経細胞の中には核のような物体が存在するが、色だけを見ると境界線と同様であるため、局所的な情報のみを用いると物体に属するピクセルは境界線として認識されてしまう。すなわち、このような物体を無視しつつ境界線の位置を正確に捉えることが、DCN の構造設計における Chen らの狙いである。DCN にみられるように、ダウンサンプリング部で階層的に得られた特徴マップをより深い層で直接利用する構造は、スキップコネクションと呼ばれる。この構造については、特に神経細胞セグメンテーションのタスクでの重要性が指摘されており [7]、先述した U-Net などのアーキテクチャでも採用されている。

DCN の具体的な構成要素としては、16の畳み込み層 (フィルタのサイズは 3×3 または 1×1) と3つのプーリン

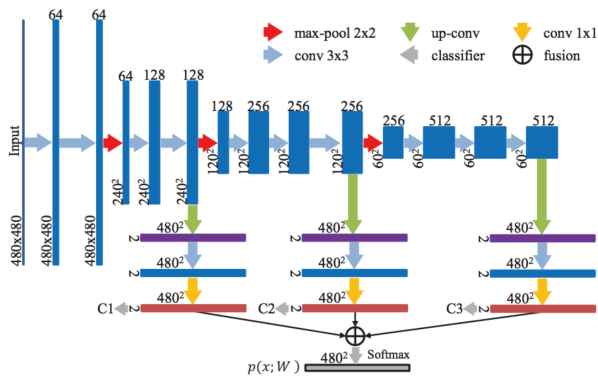


図 3: DCN のアーキテクチャ (Chen ら [3] より引用)

層 (ウィンドウのサイズは 2×2 , スライドは 2), 3 つの逆畳み込み層 (フィルタのサイズは浅い層から順に 4×4 , 8×8 , 16×16) がある. 活性化関数には ReLU (rectified linear unit) を用いる. また, 勾配消失を防ぐために, 補助的な分類器 C1, C2, C3 を導入している. これらの分類器と区別するため, 最終的に $p(x; W)$ を出力する分類機をメイン分類器と呼ぶことにする. ただし, $p(x; W)$ は, F1, F2, F3 の和に対するソフトマックス関数の出力である. また, 以下では C1, C2, C3 におけるソフトマックス関数に入力される特徴マップをそれぞれ F1, F2, F3 と表記する.

DCN の学習を行う際の誤差関数は次のように表せる.

$$L(X; \theta) = \frac{\lambda}{2} \left(\sum_c \|W_c\|_2^2 + \|W\|_2^2 \right) - \sum_c \sum_{x \in X} w_c \psi_c(x, l(x)) - \sum_{x \in X} \psi(x, l(x)) \quad (1)$$

第 1 項は正則化項で, W_c と W はそれぞれ補助的な分類器とメイン分類器のパラメータである. λ は正則化パラメータである. $\psi(x, l(x))$ はパラメータ W におけるメイン分類器の出力に対する交差エントロピー誤差である. $l(x)$ は画像 X に属するピクセル x に対するラベルである. $\psi_c(x, l(x))$ についても同様に, パラメータ W_c における C1, C2, C3 の出力に対する交差エントロピー誤差である. そして, w_c は補助的な分類器の誤差に対する割引率である. この割引率を学習の進みとともに減少させることで, C1, C2, C3 での勾配消失を避けつつ, それぞれが出力する特徴マップを利用してメイン分類器の最適化がなされることが期待される.

後述する実験では, 学習済みモデルにおいて, テストデータに対する各々の特徴マップを可視化することで, 階層的な情報が利用されていることを確認する.

4. 提案手法

本節では, 先述した DCN の構造をふまえ, その能力を最大限に引き出し, 高い精度での神経細胞セグメンテーションを行うためのパラメータチューニング方法を提案する. 提案手法の模式図を図 1 に示す. まず, Chen ら [3] が提

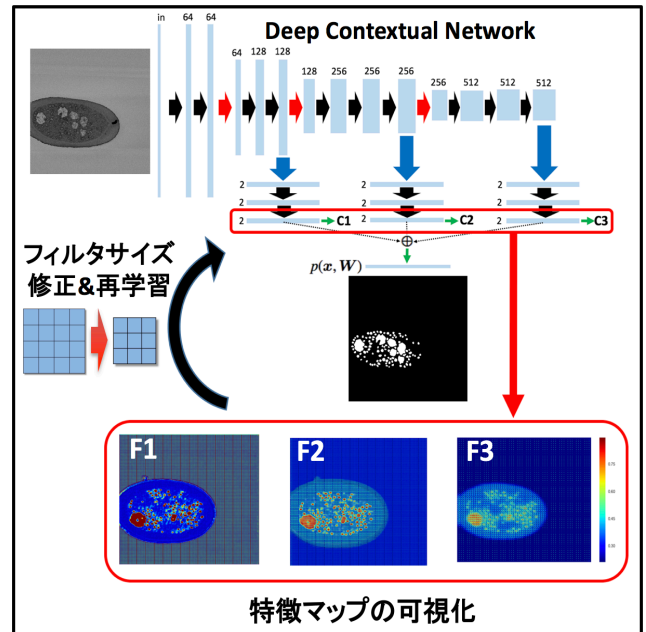


図 4: 提案手法の概要

案したパラメータを用いて DCN による学習を行う. そして, 検証用データを用いた推論の際に得られる 3 種類の特徴マップを目視によって観察し, 十分に働いていない特徴マップを特定する. さらに, 特定された特徴マップを出力するスキップコネクションにおける, 逆畳み込み層のフィルタサイズを変更し, 再度学習を行う.

DCN においてはスキップコネクションが重要な要素であり, それらの働きの中心となるのは逆畳み込み層であることから, 本研究では逆畳み込み層のハイパーパラメータであるフィルタサイズに注目している.

5. 実験

5.1 実験手順

まず, 次項で述べるデータセット A を用いて, フィルタサイズの変更が, 特徴マップの見た目やセグメンテーション精度にどのような影響を与えるかを検証した. 次に, データセット B を用い, 提案手法によるフィルタサイズの変更を行い, セグメンテーション精度の変化を検証した. 以下では, 各々を実験 1, 実験 2 と呼ぶこととする.

5.2 データセット及び評価指標

実験には, 「ショウジョウバエ 1 齢幼虫腹部連続切片画像 (図 5)」と, 「クロオオアリ単仲間識別感覚子連続切片画像 (図 2)」を用いた. 以下ではこれらをデータセット A, データセット B と呼ぶこととする. データセット A は, ISBI と呼ばれる国際会議のコンペで使用されたもので, DCN を含め, これまでに提案された深層学習アーキテクチャにおいて高い精度が保証されている. そのため, 今回はフィルタサイズ変更の影響を検証するために利用す

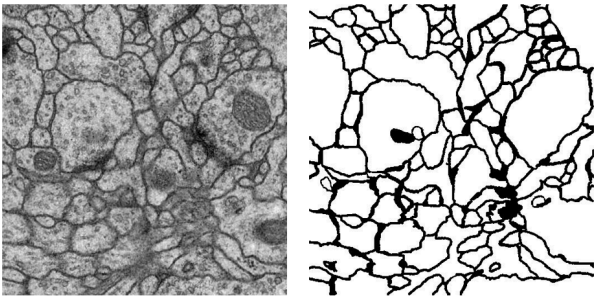


図 5: ISBI2012 データセット

るに留めた。訓練データ 30 枚のうち 3 枚を検証用データとし、汎化誤差の確認や特徴マップの可視化に用いた。画像サイズは 512×512 ピクセルである。

データセット B は、神戸大学理学研究科の尾崎研究室から提供された連続切片画像データで、377 枚の電子顕微鏡画像に対し、専門家によるラベル付けがなされている (図 2 右)。ここでは、ラベル画像で色の付いている部分のみが正解領域となるようにグレースケール変換したものを教師データとした。画像のサイズは 2048×2048 ピクセルであるが、実験では興味ある領域のみを 512×512 にクリッピングしたものを使用した。図 2 で示されているのは、クリッピング済みの画像である。なお、今回は 377 枚の画像のうち、先頭から 100 枚を訓練データとし、続く 10 枚をテストデータとした。ただし、訓練データのうちの最後の 10 枚は検証データとした。

いずれのデータセットにおいても、その量を補うためにデータ拡張を適用している。具体的には、ミニバッチを作成する際に各画像に対し、(1)90 度回転、(2)180 度回転、(3)270 度回転、(4)左右反転、(5)上下反転、(6)何もしない、のいずれかの処理をランダムで施した。

提案手法の評価には、画像セグメンテーションのタスクでしばしば利用される Intersection over Union (以下 IoU) を用いた。

5.3 実験設定および実装

実験 1 では、特徴マップ F1, F2, F3 について、それぞれに対応する逆畳み込み層におけるフィルタサイズを変更して学習を行った。以下では便宜的に、F1 に関わるフィルタサイズを「F1_filter」のように表記する。変更パターンは表 1 に示す 6 種類とし、デフォルトの設定である $k = 4$, $k = 8$, $k = 16$ の場合と比較した。ただし、 k はフィルタの 1 辺の長さを表している。

実験 2 では、実験 1 の結果をふまえたうえで、特徴マップの可視化に基づきフィルタサイズの変更を行い、精度が向上するかどうかを検証した。今回は、提案手法によって選択されたものとは異なる 12 種類の変更パターンを用意し、それらとの比較によって提案手法の有効性を検証した。

モデルの最適化には Adam を用い、その際のハイパーパ

表 1: フィルタサイズの変更パターン (k はフィルタの 1 辺の長さ)

	F1_filter	F2_filter	F3_filter
変更前	$k = 4$	$k = 8$	$k = 16$
F1_大	$k = 6$	$k = 8$	$k = 16$
F1_小	$k = 2$	$k = 8$	$k = 16$
F2_大	$k = 4$	$k = 10$	$k = 16$
F2_小	$k = 4$	$k = 6$	$k = 16$
F3_大	$k = 4$	$k = 8$	$k = 20$
F3_小	$k = 4$	$k = 8$	$k = 10$

ラメータは $\alpha = 0.001$, $\beta_1 = 0.9$, $\beta_2 = 0.999$ とした。正則化パラメータは、 $\lambda = 0.0005$ とした。バッチサイズは 10 とし、エポック数は 100 とした。ただし、特徴マップの可視化の際には、100 エポックのうち汎化誤差が最小となったタイミングで学習を止めたモデルを用いた。補助的な分類器の誤差に対する割引率 w_c は、1 を初期値とし、学習が進むにつれて減少させた (10 エポック毎に 0.1 倍)。以上の設定は全ての実験において共通である。

特徴マップの可視化にあたっては、学習済みモデルに対して検証データによる順伝播計算を行い、F1, F2, F3 が足し合わされる直前に記録しておき、それをヒートマップとして出力した。

実装は主に深層学習フレームワーク Chainer によって行い、モデルの学習には NVIDIA Tesla M40 GPU を用いた。

5.4 実験結果

まず、実験 1 において、フィルタサイズを変更しなかった場合の結果を図 6 に示す。

特徴マップの可視化結果から、DCN の設計時に意図されたように、F1 から F3 にかけてだんだん粗い検出がなされているのがわかる。次に、フィルタサイズを変更した場合の一例として、F2_filter を拡大して再学習を行った場合の特徴マップの変化を図 7 に示す。拡大前と比べると、誤検出がなくなっている部分と、検出漏れが現れている部分が観察できる。また、全体的にコントラストが弱くなっていることも分かる。これは、フィルタサイズを拡大したことで、より広い範囲の特徴が考慮されるようになったためだと考えられる。同様に、表 1 に示した 6 つの変更パターンにおける特徴マップの変化を分析したところ、フィルタサイズを拡大した場合はこれと同様の傾向が見られ、また、縮小した場合は逆に誤検出が増えて検出漏れが減り、コントラストが強くなる傾向があることが読み取れた。

次に、実験 1 の結果をふまえて行った、実験 2 の結果について述べる。まず、提案手法のフローに従い、デフォルトのフィルタサイズを用いてデータセット B の学習を行った。その出力と特徴マップの可視化結果を図 8 に示す。実験 1 の場合と同様、階層的な特徴マップが得られているが、

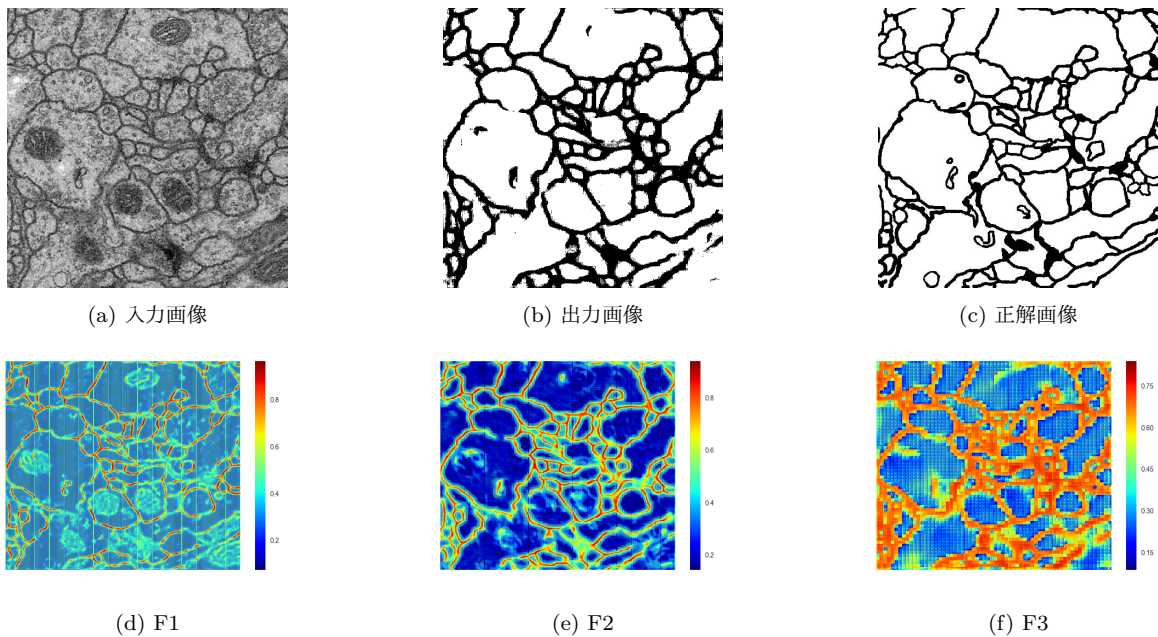


図 6: データセット A における学習結果 (フィルタサイズ変更なし)

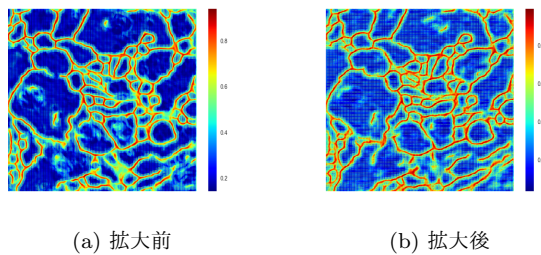


図 7: F2_filter の拡大による F2 の変化 (データセット B)

F3 における検出部位がぼやけてしまっているのが分かる。また、出力画像においては、正解画像と比べて細胞が小さめに検出されてしまっており、穴が空いてしまっている部分も見受けられる。ここで、出力画像において目立った誤検出も認められないこと、実験 1 の結果をふまえると、F3_filter を小さくすることが効果的だと考えられる。これにより、コントラストが強調されて一つ一つの細胞がはっきりと分割されることが期待される。この場合、F3_filter のサイズを $k = 14, 12, 10$ とするパターンが候補として挙げられる。^{*1} 図 9 および図 10 にて、 $k = 10$ とした場合の特徴マップと出力画像の変化を示す。

図 9 からは、フィルタサイズ変更時に意図したとおり、一つ一つの細胞がはっきりと分割されるようになっていることが分かる。また、図 10 からは、特徴マップの改善に伴い、出力画像において細胞内に穴が空いてしまっていた部分 (矢印で示されている) が改善されていることが分かる。

次に、IoU を用いて定量的な評価を行い、提案手法によって選択されたものとは異なる 12 種類の変更パターンと比

^{*1} DCN の設計上、F3_filter のサイズは 10 が最小となる

較した。その結果を表 2 に示す。

各パターンの IoU の値をフィルタサイズ変更前と比較すると、提案手法によって選択した $k = 4, 8, 10$ を含めた 15 パターンのうち、13 パターンにおいて精度が改善しており、特に $k = 4, 8, 10$, $k = 6, 8, 16$, $k = 2, 8, 20$ の場合に 0.05 以上の改善が見られる。この結果から、データセット B においては DCN のフィルタサイズを変更する余地が十分にあったと考えられる。そして、提案手法で $k = 4, 8, 10$ を候補として選択できたことは、特徴マップの可視化に基いてフィルタサイズを変更し、再学習を行うことの有効性を示しているといえる。

6. おわりに

神経細胞セグメンテーションをはじめ、電子顕微鏡連続切片画像を対象とした画像処理のタスクにおいては、DCN のようにスキップ接続を有する深層畳み込みアーキテクチャが有効である。しかし、通常はデータセットに応じてハイパーパラメータの調整をする必要があり、それには多大な計算コストがかかる。本研究では、DCN の持つ特徴マップの視認性に注目し、その可視化結果に基いて逆畳み込み層のカーネルサイズを変更する方法を提案した。2 種類のデータセットを用いた実験では、フィルタサイズの変更による特徴マップの変化を分析した上で、提案手法の評価実験を行った。その結果、提案手法によりセグメンテーション精度が向上することが示された。

今後の課題としては、グリッドサーチやランダムサーチなどのハイパーパラメータ探索手法との比較や、提案手法の自動化メカニズムの構築が挙げられる。特に周波数解析や強化学習などを用いることで、可視化に基づく判断を定

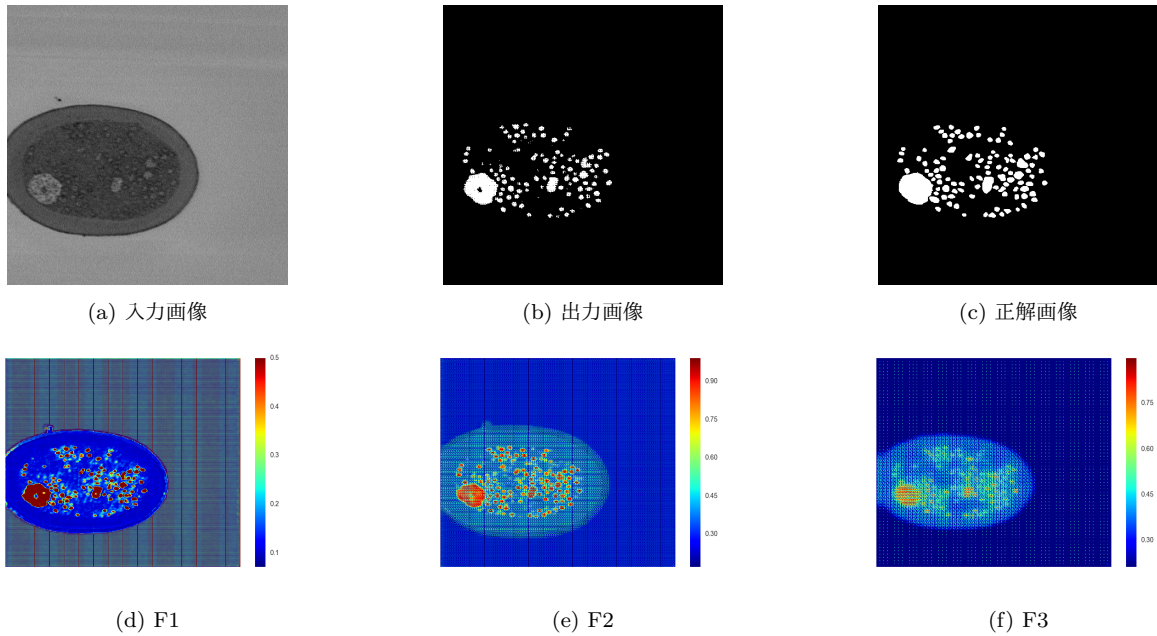


図 8: データセット B における学習結果 (フィルタサイズ変更なし)

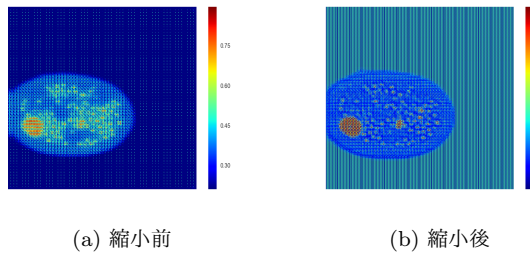


図 9: F3_filter の拡大による F3 の変化 (データセット B)

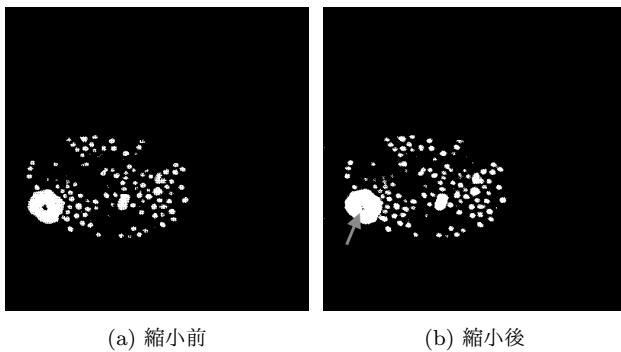


図 10: F3_filter の拡大による出力画像の変化 (データセット B)

量的かつ自動的に行えると考えられる。また、提案手法はスキップ接続を有するアーキテクチャとの親和性が高いと考えられるため、U-Netをはじめ、DCN 以外のアーキテクチャへの転用可能性も探していきたい。

謝辞 本研究では、NVIDIA 合同会社様のご厚意により、NVIDIA Tesla M40 GPU を使用させていただきました。心より感謝致します。

表 2: テストデータに対する精度評価 (データセット B)

フィルタサイズ	IoU	変更前との差
$k = 4, 8, 16$ (変更前)	0.6536	
$k = 4, 8, 10$ (提案手法)	0.7041	+0.0505
$k = 4, 8, 12$ (提案手法)	0.6859	+0.0323
$k = 4, 8, 14$ (提案手法)	0.6625	+0.0089
$k = 2, 8, 16$	0.6871	+0.0335
$k = 6, 8, 16$	0.7097	+0.0561
$k = 4, 6, 16$	0.5198	-0.1338
$k = 4, 10, 16$	0.6934	+0.0398
$k = 4, 8, 18$	0.6163	-0.0373
$k = 4, 8, 20$	0.6801	+0.0265
$k = 2, 4, 10$	0.6693	+0.0157
$k = 2, 4, 20$	0.6573	+0.0037
$k = 2, 8, 20$	0.7063	+0.0527
$k = 4, 10, 18$	0.6611	+0.0075
$k = 2, 6, 16$	0.6419	-0.0117
$k = 2, 8, 18$	0.6736	+0.02

参考文献

- [1] Kaynig, V., Vazquez-Reina, A., Knowles-Barley, S., Roberts, M., Jones, T. R., Kasthuri, N., Miller, E., Lichtman, J. and Pfister, H.: Large-scale automatic reconstruction of neuronal processes from electron microscopy images, *Medical image analysis*, Vol. 22, No. 1, pp. 77–88 (2015).
- [2] Ciresan, D., Giusti, A., Gambardella, L. M. and Schmidhuber, J.: Deep neural networks segment neuronal membranes in electron microscopy images, *Advances in neural information processing systems*, pp. 2843–2851 (2012).
- [3] Chen, H., Qi, X., Cheng, J.-Z., Heng, P.-A. et al.: Deep Contextual Networks for Neuronal Structure Segmentation., *AAAI*, pp. 1167–1173 (2016).
- [4] Ronneberger, O., Fischer, P. and Brox, T.: U-net: Convolutional networks for biomedical image segmentation,

International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, pp. 234–241 (2015).

- [5] Long, J., Shelhamer, E. and Darrell, T.: Fully convolutional networks for semantic segmentation, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3431–3440 (2015).
- [6] Noh, H., Hong, S. and Han, B.: Learning deconvolution network for semantic segmentation, *Proceedings of the IEEE International Conference on Computer Vision*, pp. 1520–1528 (2015).
- [7] Drozdal, M., Vorontsov, E., Chartrand, G., Kadoury, S. and Pal, C.: The importance of skip connections in biomedical image segmentation, *International Workshop on Large-Scale Annotation of Biomedical Data and Expert Label Synthesis*, Springer, pp. 179–187 (2016).