

電子顕微鏡連続切片画像セグメンテーションにおける 深層学習の利用に関する研究

Segmentation in Stacks of Electron Microscopy Images Using Deep Learning

高屋 英知^{*1}
Eichi Takaya

竹市 裕介^{*2}
Yusuke Takeichi

尾崎 まみこ^{*2}
Mamiko Ozaki

栗原 聡^{*1}
Satoshi Kurihara

^{*1}慶應義塾大学 大学院理工学研究科
Graduate School of Science and Technology, Keio University

^{*2}神戸大学 大学院理学研究科
Graduate School of Science, Kobe University

In the research field of connectomics, serial section electron microscopy images are reconstructed in three dimensions for the purpose of observing neural structure. Automatic segmentation of neuronal cells is important to reconstruct the images, and in recent years a method using deep learning has got a lot of attention. However, in the application of deep learning, there are many optional methods to improve segmentation accuracy. In this paper, we focus on the tuning method of hyperparameters. Deep Contextual Network (DCN), which is expected to be processed with particularly high precision among existing methods, has some feature maps with different spatial frequencies. We have an eye to this structure and propose hyperparameter tuning method based on visualization of feature maps. Experiments using two kinds of datasets showed that DCN accuracy can be improved by changing the kernel size of the deconvolution layer by the proposed method.

1. はじめに

コネクトミクスと呼ばれる分野では、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、その機能的な意味を理解することが目指されている [1]。神経構造を調べるための方法はスケールによって様々だが、特にミクロな構造を観察するためには、電子顕微鏡によって撮影された連続切片画像を取得し、それらを3次元に再構築する方法が用いられる。しかし、切片画像の再構築には、数百枚にわたる画像に対して図1のようなラベリングを行う必要があり、これを人手により行うことは多大なコストを要する。Kaynigら [2] は、3次元再構築を自動化するための画像処理パイプラインを採用しており、こ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を段階的に行うことで、ラベリングの自動化を実現している。そして近年では、このパイプラインにおけるそれぞれのステージでの処理精度を向上させるために、機械学習に基づいた様々な手法が提案されている。本稿では、パイプライン方式において全体の精度に与える影響が大きいとされる、神経細胞セグメンテーションのタスクに注目する。

神経細胞セグメンテーションのタスクについては、関連したコンペティション [3] をきっかけに、様々な手法が提案されてきた。特にCiresanら [4] によりDeep neural networkを用いた手法が提案され、その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くものとなっている。さらに、近年成果を挙げている例としては、Chenら [5] によるDeep contextual network (以下、DCN と呼ぶ) やRonnenbergerら [6] によるU-Netなどが挙げられる。いずれもネットワークの構築方法に独自性がみられるが、Longらの提案したFully convolutional network [7] (以下、FCN と呼ぶ) を基礎とし、中間層で得られた特徴マップを明示的に利用している点が共通している。しかし、あるデータセットにおいて高精度なセグメンテーションを行うためには、ただ上記のようなネットワークを適用するだけでは不十分である。例えば、

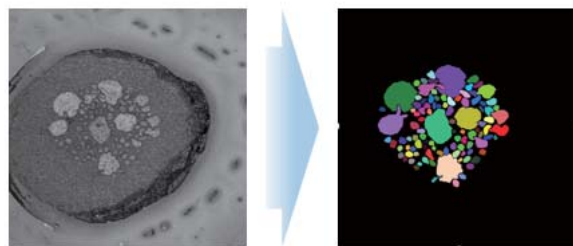


図 1: 連続切片画像のラベリング

ハイパーパラメータの適切な設定や、少ないデータを有効活用するための工夫、時系列性の考慮など、精度向上に寄与するメソッドが数多く存在する。本稿では、その中でもハイパーパラメータの設定に焦点を当て、既存のネットワークアーキテクチャのうち特に精度が高いとされるDCNの特性に基づいたハイパーパラメータ調整手法を検討する。具体的には、DCNの特徴マップの可視化に基づき、逆畳み込み層のフィルタサイズを変更するものである。以下ではDCNのアーキテクチャについて概観し、続いて提案手法について述べる。

2. Deep Contextual Network

DCNは、すべての層が畳み込み層で構成されるFCNを拡張した畳み込みニューラルネットワークである(図2)。物体認識等で用いられる畳み込みネットワーク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。そして、畳み込み層とプーリング層による特徴抽出を行うダウンサンプリング部と、抽出された特徴が逆畳み込み層によって入力画像と同サイズに復元されるアップサンプリング部を有することで、ピクセル毎の回復および分類を行えるようになっている。一方、DCNでは、ダウンサンプリング部で得られた段階的な特徴マップから複数のアップサンプリング部(以下ではスキップコネクションと呼ぶ)を構成し、最終的に足し合わせているの

連絡先: 高屋英知, 慶應義塾大学大学院理工学研究科,
〒223-8522 神奈川県横浜市港北区日吉3-14-1,
Email: takaya.wisdom0616@gmail.com

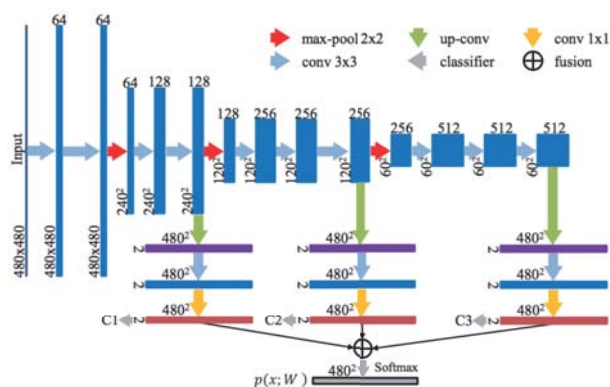


図 2: DCN のアーキテクチャ

が特徴的である。ここでは、深い層で得られた大域的な情報を分類に使い、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が意図されている [5]。DCN の有するスキップコネクションについては、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 [8]、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。

本研究では、予備実験として Chen ら [5] と同様のデータセットにて DCN の学習を行い、精度の検証および特徴マップの可視化を試みた。その結果、正則化手法として定評のあるバッチ正規化を導入することにより、特徴マップの可視化が明瞭に行えることが明らかになった。それに付随して、最終的に足し合わされる 3 種類の特徴マップは、各々が異なる粒度での特徴を保持しており、協調的に出力画像を生成していることが示唆された。さらに、スキップコネクションを欠落させることで、精度が低下することも明らかになった。

3. DCN 適用時におけるハイパーパラメータの決定

本研究では、DCN 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 [5] が提案したパラメータを用いて DCN による学習を行う。そして、チューニング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支流における、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。なお、以下では 3 種類の特徴マップを浅い層から順に F1, F2, F3 と呼ぶこととする。

4. 実験

4.1 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。

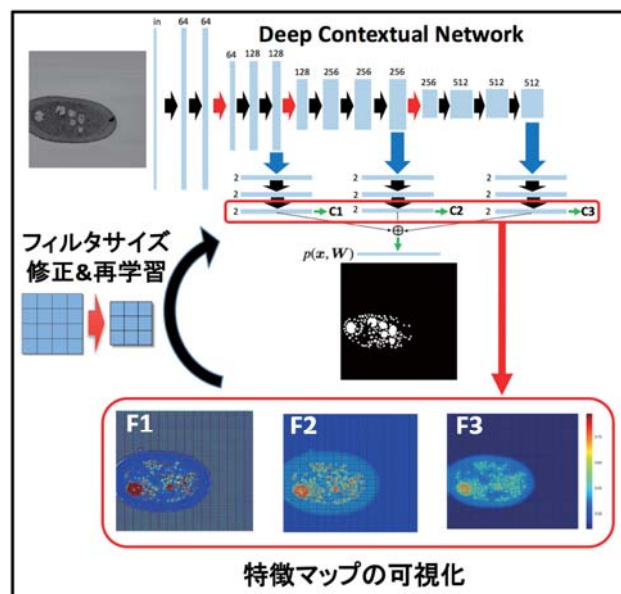


図 3: 提案手法の概要

4.2 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」と、「クロオオアリ巣仲間識別感覚子連続切片画像」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 1 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。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。

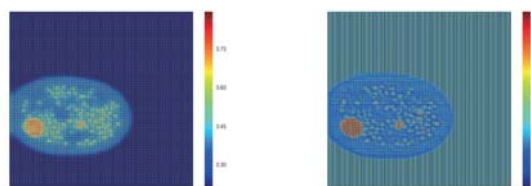
いずれのデータセットにおいても、その量を補うためにデータ拡張を適用している。具体的には、ミニバッチを作成する際に各画像に対し、(1)90 度回転、(2)180 度回転、(3)270 度回転、(4)左右反転、(5)上下反転、(6)何もしない、のいずれかの処理をランダムに施した。

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。

4.3 実験結果

データセット A を用いた実験では、フィルタサイズを縮小した場合に、対応する特徴マップのコントラストが強くなり、拡大した場合には弱くなる傾向があることが明らかになった。

データセット B を用いた実験では、特徴マップの可視化に



(a) 縮小前

(b) 縮小後

図 4: フィルタサイズの縮小による特徴マップの変化



(a) 縮小前

(b) 縮小後

図 5: フィルタサイズの縮小による出力画像の変化

より, F3 にフィルタサイズ縮小による改善の余地があると判断された. そして, フィルタサイズを縮小して再度学習を行うことで, 対応する特徴マップ F3 の変化がみられ (図 3), 出力画像の改善もみられた (図 4). また, その際の IoU の値も 0.6536 から 0.7041 へと改善しており, 擬似的にランダムな変更を施した場合と比較しても同等以上の精度であった (表 1).

5. まとめ及び今後の予定

本稿では, 電子顕微鏡画像セグメンテーションのタスクにおける深層学習の適用方法の 1 つとして, 特定のハイパーパラメータを決定する手法を検討した. DCN の特性をふまえ, 特徴マップの可視化結果に基づいてフィルタサイズを変更し, 再学習を行うことで, セグメンテーション精度の向上がみられた. 今後の課題として, 現状では目視で行うに留まっている, 可視化結果からフィルタサイズを決定する部分の自動化や, 各種ハイパーパラメータ探索手法との比較が挙げられる.

また, 本稿では深層学習適用時のハイパーパラメータ決定に着目したが, 今後は連続切片画像の時系列性を考慮したモデルや, 少量のデータを有効に活用するための方法についても検討する. インタラクティブセッションでは, それらの実験結果や議論についても併せて報告する予定である.

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.

参考文献

- [1] Olaf Sporns, Giulio Tononi, and Rolf Kötter. The human connectome: a structural description of the human brain. *PLoS computational biology*, Vol. 1, No. 4, p. e42, 2005.
- [2] Verena Kaynig, Amelio Vazquez-Reina, Seymour Knowles-Barley, Mike Roberts, Thouis R Jones,

表 1: テストデータに対する精度評価 (データセット B). k の値は左から順に, F1, F2, F3 に対応する逆畳み込み層のフィルタサイズを意味する.

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

Narayanan Kasthuri, Eric Miller, Jeff Lichtman, and Hanspeter Pfister. Large-scale automatic reconstruction of neuronal processes from electron microscopy images. *Medical image analysis*, Vol. 22, No. 1, pp. 77–88, 2015.

- [3] I Arganda-Carreras, HS Seung, A Cardona, and J Schindelin. Segmentation of neuronal structures in em stacks challenge-isbi 2012, 2012.
- [4] Dan Cireșan, Alessandro Giusti, Luca M Gambardella, and Jürgen Schmidhuber. Deep neural networks segment neuronal membranes in electron microscopy images. In *Advances in neural information processing systems*, pp. 2843–2851, 2012.
- [5] Hao Chen, Xiaojuan Qi, Jie-Zhi Cheng, Pheng-Ann Heng, et al. Deep contextual networks for neuronal structure segmentation. In *AAAI*, pp. 1167–1173, 2016.
- [6] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 234–241. Springer, 2015.
- [7] Jonathan Long, Evan Shelhamer, and Trevor Darrell. Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3431–3440, 2015.
- [8] Michal Drozdal, Eugene Vorontsov, Gabriel Chartrand, Samuel Kadoury, and Chris Pal. The importance of skip connections in biomedical image segmentation. In *International Workshop on Large-Scale Annotation of Biomedical Data and Expert Label Synthesis*, pp. 179–187. Springer, 2016.