

弱教師学習を用いた顕微鏡画像における細胞領域認識

西村 和也^{1,a)} Dai Fei Elmer Ker^{2,b)} 備瀬 竜馬^{1,c)}

概要

バイオ研究において、細胞形状は重要な指標である一方、機械学習による領域分割のための学習データの作成は、細胞種や顕微鏡の種類ごとに、細胞領域を個々にアノテーションするといった多大な労力を必要とするという課題がある。そこで、本研究では、個々の細胞の中心付近の座標といった簡易なアノテーション（弱教師）を用いた学習から、個々の細胞領域を認識することが可能な手法を提案する。提案手法では、まず CNN に細胞中心付近を推定する回帰問題を学習させる事で個々の細胞を検出する。学習した CNN は細胞形状を元に細胞検出を行っているという仮定の元、ネットワークの出力から入力方向へ重みをたどることで、個々の細胞の検出に貢献した画素を抽出し、個々の細胞領域に活用する。複数の非侵襲顕微鏡によって撮影された細胞画像に対して実験を行ない、従来手法より高精度な領域分割の実現を確認し、提案手法の有用性を確認した。また、学習データとして細胞核を染色した画像を使用することでアノテーションコストなしで細胞領域分割が実現できることを示した。

1. はじめに

バイオ分野の研究において、生体内に近い環境での研究は重要であり、非染色・非侵襲での観察を可能とする位相差顕微鏡が広く用いられる。また、細胞形状は、細胞の内部状態に関係が深く、様々な研究で細胞領域認識（セグメンテーション）が必要とされている。

しかし、位相差顕微鏡は、図 1 に示すように、細胞周辺の輝度が高くなる halo や細胞中心の輝度が低くなる shade-off 等のアーティファクトが存在し、セグメンテーションが難しい。一方、CNN を用いた手法では、学習データとして個々の細胞境界を示したアノテーションが大量に必要である。細胞形状は複雑なため、個々の細胞形状にアノテーションをつけることは非常に高コストで手間がかかる。そ

こで、より簡易なアノテーションから細胞形状を認識する方法が求められる。

本研究は、個々の細胞の中心付近へのアノテーションといった簡易なアノテーション（弱教師）を用いた学習から、個々の細胞領域を認識することを目的とする。本研究の鍵となる以下の主要な仮説を元に手法開発を行った。

- (1) 多数の複雑な形状を持つ細胞が混在している状態から、直接、個々の「細胞領域」を認識するのは難しいが、個々の細胞の「おおよその位置」を知ることができれば、細胞領域分割を容易にする。また、核染色蛍光像をこの正解データとして利用することができる。
- (2) 深層学習は、「個々細胞の中心付近の検出」という課題を学習する過程で、そこが中心であるという認識に「領域」を利用して学習している。

以上の仮説を元に、本研究では、細胞位置という弱教師を用いて細胞検出 CNN を学習し、個々の細胞ごとに検出に貢献した画素を解析することで、個々の細胞領域を認識することが可能な手法を提案する。本論文の主なコントリビューションは以下の三点である。

- 非剛体で不境界な境界を持つ細胞が、画像中に高密度に分布する状態から、個々の細胞領域を認識するというチャレンジングな課題に対して、細胞「位置情報」という「領域情報」が与えられない弱教師のみで、細胞領域を分割可能なフレームワークを提案。
- CNN のクラス分類課題において、どの画素が認識に貢献したかを可視化するのに利用される Guided Back-propagation [6] を、「位置検出」の画素貢献分析課題に適用し、空間分布において、細胞ごとに特定の位置ごとに解析可能な手法を提案。
- 複数の細胞種・顕微鏡画像を用いて検証し、従来手法と比べ、高精度で認識できることを示した。さらに、核染色蛍光像を利用した「アノテーション完全なし」での実験を行い、可能性を示した。

2. 関連研究

位相差顕微鏡等の非侵襲顕微鏡画像を対象とした細胞領域分割手法が提案されている。輝度特徴量に基づいた画像処理手法 [2]、Graph-cut 最適化による手法 [1]、位相差顕

¹ 九州大学

² 香港中文大学

a) kazuya.nishimura@human.ait.kyushu-u.ac.jp

b) elmerker@cuhk.edu.hk

c) bise@ait.kyushu-u.ac.jp

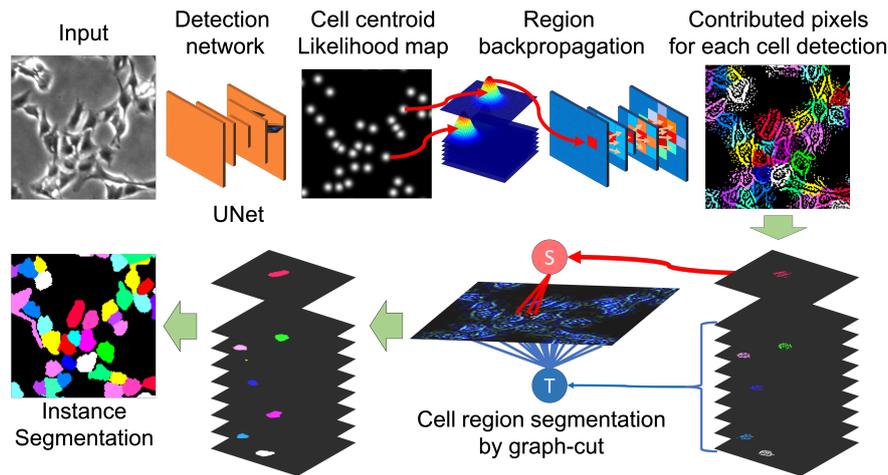


図 1 提案手法の概要

微鏡の光学モデルに基づいた手法 [7] 等、多数の手法が提案されている。これらの手法は、撮像環境の違いや細胞種（形状・厚み）の違いによる輝度値への影響によって、機能しないことが多い。また、細胞領域の学習情報を用いて、U-net ベースの領域分割ネットワークを学習する手法が多数提案されている [5] が、これらの手法は、細胞領域ごと境界をアノテーションした学習データが必要であり、細胞種ごとに作成するにはコストが高い。

一方、一般物体認識の分野では、クラスラベルや Bounding Box といった弱教師から、個々の物体領域を認識する手法が提案されつつある。Zhou ら [8] は、画像に対するクラスラベルから、クラス分類への貢献度マップにおけるピークを抽出し、個々の物体領域を識別する手法を提案している。しかしながら、一画像に同一物体が多数存在する場合、物体ごとにピークが出ない場合が多く、機能しない。Li ら [4] は、Bounding Box 学習データを用いて物体検出を行う ssd の学習を行い、CRF で個々の物体領域を分離している。この手法は、隣接する物体の Bounding Box が重複する領域を除いた Loss を用いることを特徴としているが、細胞は非剛体で複雑な形状で密接しているため、Bounding Box の重複領域が非常に多いという課題がある。これらの手法と異なり、提案手法はネットワークから個々の細胞の領域に関する情報を効果的に抽出し、複数の対象が密接に存在する場合でも分離可能なフレームワークを提案する。

3. 提案手法

提案手法の概要を図 1 に示す。提案手法は、手法の目標である「細胞領域」と比べ弱い教師である「細胞の中心付近の位置座標」を用いて、個々の細胞中心付近の尤度が高くなるような尤度マップ（図 2b）を出力するように U-net を学習する。出力された尤度マップにより、細胞ごとの中心座標を取得し、その中心領域から Backpropagation を行うことで、個々の細胞ごとに、検出に貢献した領域を抽出

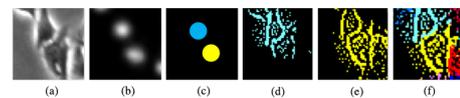


図 2 (a) 位相差像, (b) 尤度マップ, (c) 細胞毎の領域, (d),(e) 各細胞領域からの Backpropagation によって得られる貢献領域, (g)max を計算し、貢献領域を分離した結果

する。抽出した個々の領域はチャンネル C_u として表現され、個々の細胞 u とその他領域の 2 値分割を Group-cut を用いて行う。Graph-cut では、対応するチャンネル C_u を用いて、前景シードを作成し、その他のチャンネルを背景シードとして与えることで、領域分割を行う。

3.1 細胞検出

物体検出の学習データとしては、Bounding Box を用いることが一般的だが、上述したように、細胞は非剛体で複雑な形状で他の細胞と近接しているため、一つの Bounding Box に他の細胞を多数含む可能性がある。そこで、本研究では、個々の細胞の中心付近の座標を Groundtruth として利用する。これは、Bounding Box と比べ、より簡易なアノテーション法であるという利点もある。細胞の形状は不定形であるため、細胞の「中心」は曖昧であり、画像や細胞により多少の揺らぎがある。そこで、正解として与えられた中心を尤度のピークとして、中心から離れるに従って、ガウス分布的に次第に尤度が減衰するような細胞位置尤度マップ（図 2b）を正解データとして作成する。元画像を入力するとこの尤度マップを出力するように、Loss を尤度マップと出力の二乗誤差（MSE）として与え、回帰問題として U-net を学習する。ここで、この U-net の出力を y とする。

3.2 細胞検出に貢献した画素の抽出

テスト画像に対して、前節の細胞検出 U-net を用いて、細胞尤度マップを推定する。細胞中心付近は尤度が高くな

るように学習されていることから、尤度マップの値が閾値より高い領域(細胞中心領域)を抽出し、独立する前景領域を一つの細胞領域 $s_u \in \mathbf{S}$, $\mathbf{S} = \{s_{bg}, s_1, \dots, s_k\}$ とし, s_{bg} を背景領域とする. U-net による推定において, 細胞 u の中心領域の外側の形状特徴も, 細胞中心領域 s_u の尤度を高く出力していると考えられる. そこで, 提案手法では, 通常クラス分類の出力に貢献した領域を可視化する Guided Backpropagation を応用し, 注目細胞に貢献した領域を抽出する方法を提案する. 通常のクラス分類への適用と異なり, 細胞ごとに中心領域からの Backpropagation を行う. 細胞 u の領域抽出において, 出力のうち領域 s_u の領域に属さない画素を 0 とする. Backpropagation によって得られる逆伝搬を g とすると g は以下のように定式化できる.

$$g_{s_u}^{Out}(i) = \begin{cases} y(i) & \text{if } i \in s_u \\ 0 & \text{if } i \notin s_u. \end{cases} \quad (1)$$

i は座標を示す. 逆伝搬時に Relu を通過する際は, 順伝搬時かつ逆伝搬時に正となった伝搬のみを入力側の層へ伝搬する. 順伝搬または逆伝搬時に負の伝搬を伝えた場合は次の層へ値を伝搬しない. この逆伝搬は以下のように定式化できる.

$$g_{s_u}^l(i) = I(f^l(i)) \cdot I(g_{s_u}^l(i)) \cdot g_{s_u}^{l+1}(i) \quad (2)$$

$$I(x) = \begin{cases} 1 & \text{if } (x > 0) \\ 0 & \text{if } (x \leq 0). \end{cases} \quad (3)$$

f^l は l 層目の順伝搬時の l 層目の値, g^l は l 層目の逆伝搬の値を示す. I は指示関数である. 畳み込み層の逆伝搬は, ネットワークの学習時の Backpropagation と同様に計算する. この逆伝搬によって得られる貢献マップは図 2(d),(e),(f) のように近接する細胞領域も含む.

しかし, 各貢献マップの強度には明確な違いがあると考えられる. そこで, 細胞 u に対する貢献マップのうち他の細胞より強い貢献を示した領域を細胞 u の貢献マップと考える. 各領域から得た貢献マップを $g_{s_u}^0$ とすると最終的な個々の細胞毎の領域は C_u は以下の式により得る.

$$C_u(i) = \phi(\arg \max_k g_{s_k}^0(i)). \quad (4)$$

$$\phi(x) = \begin{cases} 1 & \text{if } (x = u) \\ 0 & \text{if } (x \neq u). \end{cases} \quad (5)$$

図 2g は貢献度合いを可視化した図である. 貢献度合いを考慮することで, 細胞毎の細胞形状が抽出できていることが確認できた.

3.3 Graph-cut

前節で得られた, 貢献マップを seed として使用し, 細胞毎に graphcut を行う. 細胞 u に対し, Graph-cut する場

表 1 F 値による各データセットの検出性能比較

Method	C2C12 [3]	GBM	B23P17	アノテーション無し
Zhou [8]	0.513	-	-	-
Bensch [1]	0.407	0.645	0.209	0.669
Chalfoun [2]	0.707	0.818	0.679	0.836
Ours	0.948	0.911	0.887	0.951

表 2 mean Intersection over Union (mIoU) による各データセットのインスタンスセグメンテーション精度評価

Method	C2C12 [3]	GBM	B23P17	アノテーション無し
Zhou [8]	0.157	-	-	-
Bensch [1]	0.268	0.117	0.265	0.143
Chalfoun [2]	0.405	0.265	0.225	0.364
Ours	0.514	0.472	0.464	0.478

合, 前景 seed には C_u を与え, 背景 seed にはその他の細胞の貢献領域をを与える.

4. Evaluation

4.1 dataset

位相差顕微鏡で撮影した線維芽系の幹細胞 C2C12, 位相差顕微鏡で撮影したウシ上皮細胞 B23P17, 微分干涉顕微鏡によって撮影された神経系の幹細胞 NSC といった複数の異なる顕微鏡及び異なる細胞種のデータセットを用いて評価を行った. これらのデータには, 弱教師データ及び評価用データとして, 細胞中心付近をアノテーションした情報が付与されている. 学習データとして, ネットワーク入力サイズ 320×320 に切り取った 24000, 1560, 360 枚を用いた. 評価用データは, それぞれ 144, 10, 10 枚の画像で細胞数としては計 796, 360, 633 個の細胞領域の正解データを用いて評価を行った.

さらに, 核染色蛍光像により, 完全にアノテーションなしでの領域認識の可能性を示すため, 同一の細胞培養画像を位相差顕微鏡と蛍光顕微鏡で交互に取ることで, 位置があった 2 枚のペア画像を取得し, 蛍光画像を細胞位置アノテーションの代わりに用いた. 学習データとして, 86 枚の画像を用いて, 評価用データとして 1 枚計 1306 個の細胞領域のアノテーション画像を用いた.

4.2 比較手法

精度評価のため Bensch [1], Chalfoun [2], Zhou [8]3 種類の手法と比較を行なった. Bensch はテスト画像の 1 つの細胞にピクセルレベルのマスク画像を与えた. Chalfoun [2] はデータセットの一枚の画像に対し細かいパラメータチューニングを行った. Zhou [8] はクラス分類の学習を行い, インスタンスセグメンテーションを行うため, 学習データとして細胞が写っていない画像が必要である. B23P17, GBM は共に密に細胞が存在しており, 学習データが作成不可能

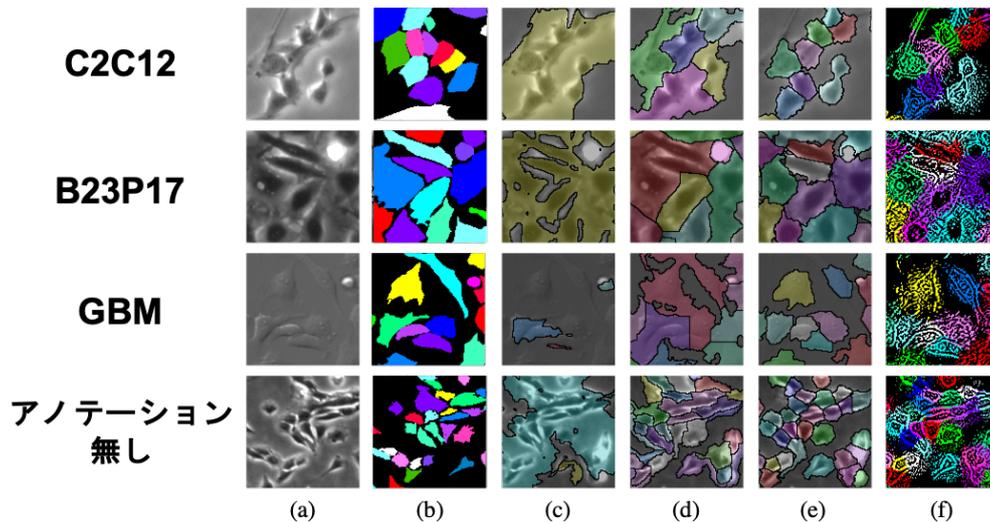


図 3 (a) 原画像, (b) groundtruth, (c) Bensch, (d) Chalfoun, (e) Ours, (f) Backpropagation

であったため, C2C12 のみに対し実験を行なった.

4.3 検出性能の比較

評価指標は F 値を使用する. 比較した結果を表 1 に示す. 表 1 はネットワークに回帰問題を与えたことで, 不定形である細胞への対応が可能となり, 高精度な検出が可能となったと考えられる. また, 表 1 より, 核染色蛍光像からアノテーション無しで学習した場合でも, F 値が 0.95 を示しており, アノテーション無しにおいても同様の細胞検出性能を持つ可能性が示された.

4.4 instance segmentation の評価

評価指標は Mean Intersection over Union を使用する. IoU は $IoU = \frac{tp}{tp+fn+fp}$, と定義する. 正解の細胞と, 推定結果の細胞のうち重複領域が大きい細胞との IoU の平均が mIoU である. 細胞毎の IoU の平均を示す. 図 1 に各手法によるインスタンスセグメンテーション結果および, 検出細胞領域毎の貢献領域を抽出した結果の例を示す. 図 1f から, 細胞検出ネットワークの出力から細胞毎の貢献領域を抽出することで細胞形状に近い領域が得られていることが確認できる. 従来手法の多くは, 細胞が密に存在する場合に対応できない. 一方, 提案手法では, 対象が密になると検出からの貢献画素が信頼度を増し, より高精度な領域推定が可能となることが示された. また, 図 1 からアノテーション無しの場合でも高精度なインスタンスセグメンテーションが実現可能性が示された. 表 2 より, 詳細なパラメータチューニングを行なった Chalfoun より高精度な結果が得られていることが分かる. 異なる顕微鏡, 細胞に対しても, 高い性能な領域推定能力を持つことが確認できる.

5. まとめ

細胞検出を行い, そのネットワークから個々の細胞領域

の成分のみを, インスタンスセグメンテーションを実現する手法を提案した. 実験により, 3 種類の手法と比較し複数の細胞が密に混在する状況下でも頑健推定を可能とすることが確認できた. また, 核染色蛍光像を利用することでアノテーション無しからインスタンスセグメンテーションが可能であることを実証した.

参考文献

- [1] Bensch, R. and Ronneberger, O.: Cell segmentation and tracking in phase contrast images using graph cut with asymmetric boundary costs, *2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI)*, IEEE, pp. 1220–1223 (2015).
- [2] Chalfoun, J., Majurski, M., Dima, A. and *et al.*: FogBank: a single cell segmentation across multiple cell lines and image modalities, *Bmc Bioinformatics*, Vol. 15, No. 1, p. 431 (2014).
- [3] Ker, D. F. E., Eom, S., Sanami, S. and *et al.*: Phase contrast time-lapse microscopy datasets with automated and manual cell tracking annotations, *Scientific data* (2018).
- [4] Li, Q., Arnab, A. and Torr, P. H.: Weakly- and semi-supervised panoptic segmentation, *Proceedings of the European Conference on Computer Vision (ECCV)*, pp. 102–118 (2018).
- [5] Ronneberger, O., Fischer, P. and Brox, T.: U-net: Convolutional networks for biomedical image segmentation, *Lecture Notes in Computer Science*, Vol. 9351, pp. 234–241 (2015).
- [6] Selvaraju, R. R., Cogswell, M. and *et al.*: Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization, *Proceedings of the IEEE International Conference on Computer Vision*, Vol. 2017-October, pp. 618–626 (2017).
- [7] Yin, Z., Kanade, T. and Chen, M.: Understanding the phase contrast optics to restore artifact-free microscopy images for segmentation, *Medical Image Analysis* (2012).
- [8] Zhou, Y., Zhu, Y., Ye, Q., Qiu, Q. and Jiao, J.: Weakly Supervised Instance Segmentation using Class Peak Response (2018).