

# クラス別敵対的ネットワークの統合による細胞画像のセグメンテーション

## Cell Image Segmentation by Integrating Generative Adversarial Network for Each Class

津田 大輝\*<sup>1</sup>  
Hiroki Tsuda

堀田 一弘\*<sup>1</sup>  
Kazuhiro Hotta

\*<sup>1</sup> 名城大学  
Meijo University

Human experts segment cell images manually now, and the criterion for segmentation varies on each expert. As a result, subjective results are obtained. If we develop an automatic segmentation method, we can obtain objective results by the same criteria. This paper proposes a cell image segmentation method using Generative Adversarial Network (GAN) with multiple different roles. The proposed method improved the segmentation accuracy in comparison to conventional pix2pix.

### 1. はじめに

近年, Generative Adversarial Network (GAN) [Goodfellow 14]を用いることにより, 本物らしい画像が生成できることが注目を集めている. さらに, GAN を拡張した pix2pix [Isola 17]や Cycle GAN [Zhu 17], Star GAN [Choi 18]等は Image-to-Image の変換法として幅広いタスクで汎用性を持つことが知られている. 特に pix2pix はセマンティックセグメンテーションに対しても有効であり, 医療分野や細胞生物分野に応用されている [Havaei 17], [Ji 15]. 従来, 細胞画像のセグメンテーションは専門家の手により行われてきたため, 主観的な結果になりがちであった. そのため, 深層学習技術を用いることにより同一の基準による客観的な結果を得たいという要望が多い. そこで, 本論文では細胞画像を細胞膜, 細胞核, 背景の 3 クラスにセグメンテーションする.

従来の pix2pix は生成器と識別器が一個ずつで構成されていた. しかし, pix2pix を複数クラスのセグメンテーション問題に適用する場合, 1 つの生成器だけにより複数クラスを同時にセグメンテーションすることは困難である. 細胞膜だけをセグメンテーションする生成器と細胞核だけをセグメンテーションする生成器などのように各生成器に異なる役割を与えれば, 効率的に学習ができ, 精度が向上すると考えた. そこで本論文では, 複数の生成器と識別器に異なる役割を与える.

ここでは 2 つのセグメンテーション法を提案する. 1 つ目は Dual Different Roles GAN (DDR-GAN)という方法であり, 2 つの生成器を用いて細胞膜と細胞核をセグメンテーションする. 2 つ目の手法は Triple Different Roles GAN (TDR-GAN)という方法であり, 3 つの生成器で細胞膜, 細胞核, 背景をそれぞれセグメンテーションする.

実験では, 細胞画像から細胞膜, 細胞核, 背景の 3 クラスをセグメンテーションする問題で評価する. 評価指標として IoU (Intersection over Union)を用いる. 従来手法である 1 個ずつの生成器と識別器を用いた pix2pix の結果と比較し, 提案手法の有効性を示した. さらに, 提案手法からいくつかの識別器を削除した方法の評価も行い, 提案手法の有効性を示した.

本論文の構成は以下の通りである. まず, 2 節では関連研究について述べる. 3 節では提案手法の詳細について説明する. 4 節では細胞画像のセグメンテーションにおいて提案手法を評価する. 最後に, 5 節でまとめと今後の課題を述べる.

### 2. 関連研究

近年, 様々なセマンティックセグメンテーション法が提案されている [Ronneberger 15], [Long 15], [Zhao 17]. pix2pix は生成器にエンコーダ・デコーダ構造の方法を用いて入力画像から学習画像に似た変換画像を生成する. そして, 生成された画像が本物か否かを判断する識別器と生成器を互いに競わせながら学習する敵対的ネットワークである. GAN の一種である pix2pix はセグメンテーションに使用することができ, 非敵対的ネットワークよりもセグメンテーションに関して優れた結果を示している [Luc 16].

複数の GAN を用いる手法もいくつか提案されている. Mixture Generative Adversarial Nets (MGAN) [Hoang 18]は複数の生成器を用いる.  $K$  個の生成器を用いて分布を生成し, それらを混合した分布により GAN の弱点である mode collapse 問題を回避した.他にも, Dual Discriminator Generative Adversarial nets (D2GAN) [Nguyen 17]は 2 つの識別器を用いる. この 2 つの識別器には異なる役割が与えられている. 1 つは学習画像を高得点に生成画像を低得点にするように学習を行う. もう一方の識別器は生成画像を高得点に学習画像を低得点にするよう学習し, 生成器はこれらの 2 つの識別器を騙すように学習する. これにより mode collapse 問題を回避できた. Generative Multi-Adversarial Networks (GMAN) [Durugkar 17]は複数の同じ識別器を用いる. これは最大得点を持つ識別器を選択し, それに対して学習することにより良好な解への収束が期待できる手法である.

複数の生成器または識別器を用いた従来法は存在するが, これらの手法は各々に特定の役割を与えなかった. 役割が異なる複数の GAN を用いることにより, 精度が向上すると考えられる.

### 3. 提案手法

本節では提案手法の詳細について述べる. 3.1 節では 2 つの提案手法に共通するネットワークの詳細について説明する. また, 3.2 節では DDR-GAN について述べ, 3.3 節では TDR-GAN について述べる.

#### 3.1 ネットワークの詳細

従来の pix2pix は 1 つずつの生成器と識別器によりセグメンテーションを行うため, 複数のクラスのセグメンテーション問題では精度が低下してしまう. そこで, クラス毎に生成器を用意して

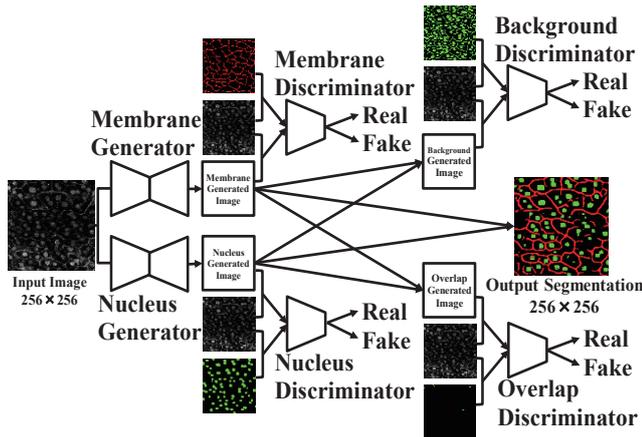


図 1. 提案手法 (a) DDR-GAN

学習を行い、その結果を統合することにより質の高いセグメンテーションを行う。また、クラス毎に識別器を用意することにより、様々な専門的観点から生成と本物か偽物か判断することができる。つまり、生成器により各クラスの確率マップを独立して生成し、それぞれの確率マップは対応する識別器に入力され、本物か偽物か判断される。また、それぞれ特徴マップからさらに判断したい部分だけを抽出することにより、該当部分だけを学習する識別器の作成も可能である。これは後述のそれぞれの提案手法にて説明する。

提案手法における生成器と識別器は従来手法とすべて同じ構造を使用する。生成器には U-Net [Ronneberger 15] を使用し、識別器として畳み込み層が 5 層の CNN を使用する。また、すべての生成器と識別器に対してフルスクラッチ学習を行う。

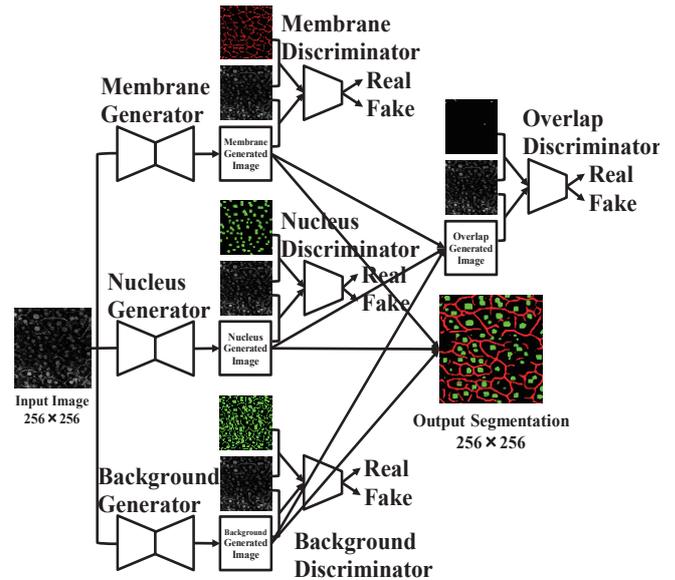
### 3.2 DDR-GAN

図 1(a) に DDR-GAN を示す。2 つの生成器により細胞膜と細胞核の確率マップを生成し、各クラスの結果はそれぞれの対応した識別器に入力される。また、細胞膜と細胞核のセグメンテーション結果から、細胞膜でも細胞核でもない部分を背景クラスとみなすことにより、背景のセグメンテーション結果を得ることが出来る。そして、識別器に背景クラスのセグメンテーション結果を入力する。学習に用いる正解画像は専門家が細胞膜と細胞核をそれぞれ独立して割り当てを行ったため、重複している箇所がある。この重複部分については、細胞膜を優先して学習を行うように正解画像の重複部分を細胞膜と定義し、学習画像とした。そして、それぞれの生成器により生成された画像からも同様に重複部分を抽出し、識別器により判断させる。つまり、1 つの生成器に対して 3 つの識別器から loss 関数の更新を行う。

細胞膜の生成器に対して loss 関数は以下のように学習した。

$$L_{CGAN}(G, D) = E_{x, y \sim P_{data}(x, y)} [\log D(x, y)] + E_{x, y \sim P_{data}(x), z \sim P_z(z)} [\log(1 - D(x, G(x, z)))] \quad (1)$$

$$L_{L1}(G) = E_{x, y \sim P_{data}(x, y), z \sim P_z(z)} [\|y - G(x, z)\|_1] \quad (2)$$



(b) TDR-GAN

$$G_{mem}^* = \arg \min_{G_{mem}} \max_{D_{mem}} L_{CGAN}(G_{mem}, D_{mem}) + \arg \min_{G_{back}} \max_{D_{back}} L_{CGAN}(G_{back}, D_{back}) + \arg \min_{G_{overlap}} \max_{D_{overlap}} L_{CGAN}(G_{overlap}, D_{overlap}) + \lambda \mathcal{L}_{L1}(G_{mem}) \quad (3)$$

式中の入力は 3 つあり、 $x$  は入力画像、 $y$  は正解画像、 $z$  は学習ラベルを表す。 $G(\cdot)$  は生成器が出力するセグメンテーション結果、 $D(\cdot)$  は識別器が出力する本物か偽物か判断した結果を表す。 $G_{mem}$  は細胞膜の生成器、 $D_{mem}$  は細胞膜、 $D_{back}$  は背景、 $D_{overlap}$  は重複部分の識別器を表す。 $G_{back}$  と  $G_{overlap}$  は細胞膜と細胞核のセグメンテーション結果から得られる背景と重複部分のセグメンテーション結果を表す。

細胞膜、背景、重複部分の識別器の式 (1) を 3 つと L1 正則化の式 (2) を組み合わせた式 (3) の loss 関数を用いて生成器の学習を行った。また、細胞核の生成器も同様の loss 関数を用いた。

### 3.3 TDR-GAN

図 1(b) に TDR-GAN を示す。3 つの生成器により、細胞膜、細胞核、背景の 3 つの確率マップを生成する。3 つの生成器により得られた結果はそれぞれの対応する識別器に入力される。また、DDR-GAN と同様に重複部分を抽出して識別器に入力し、判断をする。TDR-GAN では生成器 1 つに対して 2 つの識別器から loss 関数の更新を行う。

細胞膜の生成器に対して TDR-GAN では以下のように学習した。

$$G_{mem}^* = \arg \min_{G_{mem}} \max_{D_{mem}} L_{CGAN}(G_{mem}, D_{mem}) + \arg \min_{G_{overlap}} \max_{D_{overlap}} L_{CGAN}(G_{overlap}, D_{overlap}) + \lambda \mathcal{L}_{L1}(G_{mem}) \quad (4)$$

細胞膜と重複部分の識別器毎の式(1)が 2 つになった以外は DDR-GAN と loss 関数は同様である、また、細胞核、背景の生成器も同様の loss 関数を用いた。

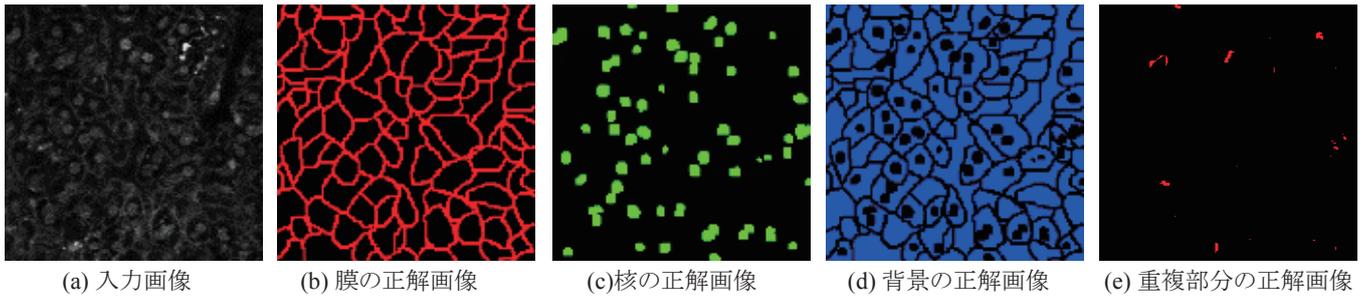


図 2. 細胞膜および細胞核に蛍光マーカーを発現するトランスジェニックマウスの肝臓の蛍光画像の一例

表 1. 提案手法の精度結果

method	Membrane[%]	Nucleus[%]	Background[%]	mIoU[%]
1. Comparison of DDR-GAN				
w/o Background & Overlap	36.61	58.74	70.54	55.30
w/o Background	37.13	59.62	70.43	55.73
w/o Overlap	36.28	61.18	70.38	55.95
Proposed method (Full model)	37.92	60.12	70.58	56.21
2. Comparison of TDR-GAN				
w/o Overlap	38.60	59.26	69.53	55.80
Proposed method (Full model)	38.83	58.77	70.06	55.89

表 2. 従来手法との精度比較

Method	Membrane[%]	Nucleus[%]	Background[%]	mIoU[%]
pix2pix	36.67	57.99	67.98	54.21
DDR-GAN	37.92	<b>60.12</b>	<b>70.58</b>	<b>56.21</b>
TDR-GAN	<b>38.83</b>	58.77	70.06	55.89

## 4. 評価実験

本節では提案手法の評価実験結果を示す. 4.1 節では今回の実験で使用したデータセットの詳細を述べ, 評価実験の結果を 4.2 節で述べる.

### 4.1 データセット

本論文では, 細胞膜および核に蛍光マーカーを発現するトランスジェニックマウスの肝臓の蛍光画像 50 枚を使用した. 画像の大きさは  $256 \times 256$  画素であり, 細胞膜, 細胞核, 背景の 3 クラスからなる. ここでは 30 枚を学習用画像, 10 枚を検証用画像, 10 枚を評価用画像とした [Imanishi 18].

図 2 に細胞画像の例を示す, 図 2(a) が細胞画像であり, 図 2(b) は専門家により細胞膜のラベルが付けられた正解画像である. 図 2(c) と (d) は同様に細胞核と背景の正解画像である. 図 2(e) は細胞膜と核の重複する部分を示す.

評価指標には Intersection over Union (IoU) および Mean IoU を採用した. これらは次式のように算出される.

$$IoU = \frac{TP}{TP+FP+FN} \quad (5)$$

$$mIoU = \frac{1}{c} \sum_c \frac{TP}{TP+FP+FN} \quad (6)$$

ここで, TP は正解した画素数, FP は別クラスの画素に対して誤識別した画素数, そして FN は別クラスとして誤識別した画素数を示し, c はクラス数を示す.

### 4.2 実験結果

役割の異なる pix2pix の有効性を示すために, 識別器の一部用いずに精度を評価する. 実験では全ての方法で独立に 5 回データセットの画像を変更し, 学習を行う. その平均値を精度評価に用いた. 表 1 にそれぞれの精度を示す. 表 1 から分かるように, 全識別器を用いた DDR-GAN の mIoU が最も高かった.

また, 背景と重複部分の識別器を取り除いた場合が最も精度が低い. それぞれの識別器の有効性を示している. また, TDR-GAN でも同様に重複部分の識別器がない場合, 精度が低下することが分かる. これらの結果から役割の異なる複数の pix2pix の有効性を示している.

表 2 は提案手法と従来手法である pix2pix の結果を示している. 2 つの提案手法は従来手法と比べてどちらも精度が向上した. また, DDR-GAN は TDR-GAN より良い結果である. これは, TDR-GAN が 1 つの生成器に対して 2 つの識別器を使用したのに対して, DDR-GAN では 1 つの生成器に対して 3 つの識別器が使用されるためである. 複数の識別器が各生成器を学習させるのに効果的であった.

従来手法の pix2pix と表 1 の背景と重複部分の識別器を取り除いた DDR-GAN を比較すると, 元の pix2pix より優れている. これはクラス毎の生成器と識別器の有効性を示している.

表 2 より, DDR-GAN は従来の pix2pix と比較して, 細胞膜で 1.25%, 細胞核で 2.13%, 背景で 2.6%, mIoU で 2.00% 向上していることが確認できる.

図 3 はセグメンテーション結果を示す. 左から順に入力画像, 正解画像, pix2pix の結果, DDR-GAN の結果, TDR-GAN の結果を示す. 下段の画像は上段の拡大図である. 拡大画像に注目した場合, 提案手法により細胞膜のセグメンテーションがうまく識別出来ている. また, 細胞核も正確に識別出来ていることが分かる. これらの結果から異なる役割を持つ複数の pix2pix を使用する提案手法の有効性を示した.

## 5. まとめ

本論文では, 役割が異なる複数の pix2pix を用いたセグメンテーション法を提案した. 細胞画像の実験では, 提案手法は従来の pix2pix と比較して 2.0% 改善された.

しかし, 提案手法は細胞画像に特化しており, 多くのクラスを含む他のデータセットでは提案手法をそのまま適用することは困難である. 適用するにはクラスを生成器毎に自動で分類する必要がある. ここには Gating ネットワーク [Hiramatsu 18] を用いることができると考えられる. これは今後の課題である.

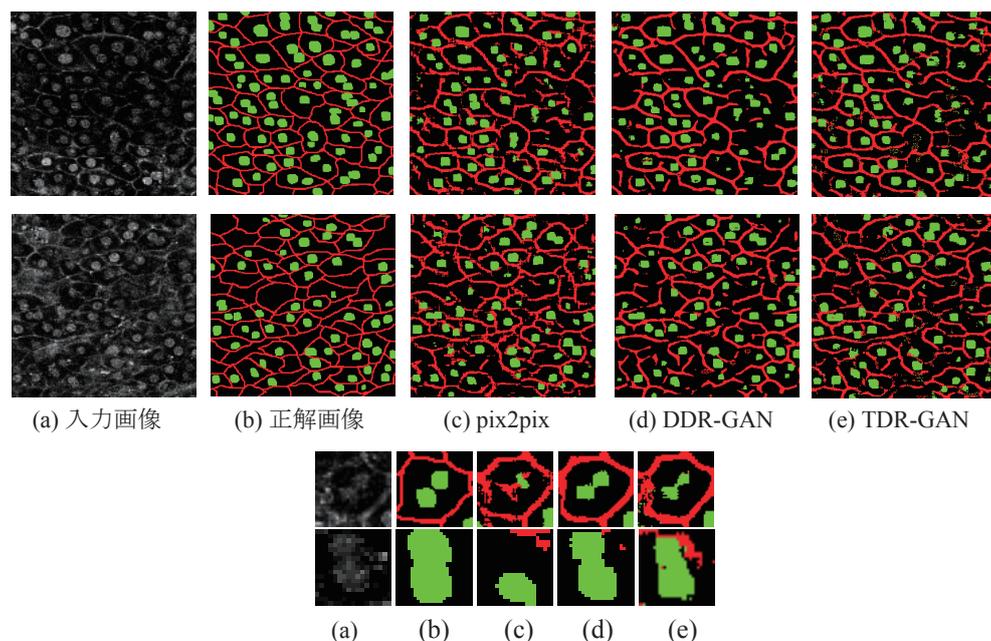


図 3. セグメンテーション結果 (a)入力画像 (b) 正解画像 (c) pix2pix [2] (d) DDR-GAN (e) TDR-GAN

### 謝辞

本研究の 1 部は、科学研究費新学術領域研究“レゾナンスバイオ”18H04746 によるものである。

### 参考文献

- [Goodfellow 14] I.J. Goodfellow, J.P. Abadie, M.Mirza, B. Xu, D.W. Farley, S. Ozair, A. Courville, and Y. Bengio.s, “Generative Adversarial Nets,” *Advances in Neural Information Processing Systems*, pp. 2672-2680, Dec, 2014.
- [Isola 17] P. Isola, JY. Zhu, T. Zhou, and A.A. Efros, “Image-to-Image Translation with Conditional Adversarial Networks,” *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 5967-5976, Jun, 2017.
- [Zhu 17] J. Zhu, T. Park, P. Isola, and A.A. Efros, “Unpaired image-to-image translation using cycle-consistent adversarial networks,” *IEEE International Conference on Computer Vision*, pp. 2223-2232, Jun, 2017.
- [Choi 18] Y. Choi, M. Choi, M. Kim, J.W. Ha, S. Kim, and J. Choo, “StarGAN: Unified Generative Adversarial Networks for Multi-Domain Image-to-Image Translation,” *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 8789-8797, Jun, 2018.
- [Havaei 17] M. Havaei, A. Davy, D.W. Farley, A. Biard, A. Courville, Y.Begio, C. Pal, P.M. Jodoin, and H. Larochelle, “Brain tumor segmentation with Deep Neural Networks,” *Medical Image Analysis*, vol. 35 pp. 18-31, Jan, 2017.
- [Ji 15] O.X. Ji, Y. Li, J. Cheng, Y. Yu, and M. Wang, “Cell image segmentation based on an improved watershed algorithm,” *International Congress on Image and Signal Processing*, pp. 433-437, Oct, 2015.
- [Luc 16] P. Luc, C. Couprie, S. Chintala, and J. Verbeek, “Semantic segmentation using adversarial networks,” *Advances in Neural Information Processing Systems Workshops*, Dec, 2016.
- [Ronneberger 15] O. Ronneberger, P. Fischer, and T. Brox, “U-net: Convolutional networks for biomedical image segmentation,” *Medical Image Computing and Computer-Assisted Intervention*, pp. 234-241, Oct, 2015.
- [Hoang 18] Q. Hoang, T.D. Nguyen, T. Le, and D. Phung, “MGAN: Training Generative Adversarial Nets with Multiple Generators,” *International Conference on Learning Representations*, Apr, 2018.
- [Nguyen 17] T. Nguyen, T. Le, H. Vu, and D. Phung, “Dual Discriminator Generative Adversarial Nets,” *Advances in Neural Information Processing Systems*, pp. 2667-2677, Dec, 2017.
- [Durugkar 17] I. Durugkar, I. Gemp, and S. Mahadevan, “Generative Multi-Adversarial Networks,” *International Conference on Learning Representations*, Apr, 2017.
- [Long 15] J. Long, E. Shelhamer, and T. Darrell, “Fully Convolutional Networks for Semantic Segmentation,” *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 8789-8797, Jun, 2015.
- [Zhao 17] H. Zhao, X. Qi, X. Wang, and J. Jia, “Pyramid Scene Parsing Network,” *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 8789-8797, Jun, 2017.
- [Hiramatsu 18] Y. Hiramatsu, K. Hotta, A. Imanishi, M. Matsuda, and K. Terai, “Cell Image Segmentation by Integrating Multiple CNNs,” *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2318-2324, Jun, 2018.
- [Imanishi 18] A. Imanishi, T. Murata, M. Sato, K. Hotta, I. Imayoshi, M. Matsuda, and K. Terai, “A Novel Morphological Marker for the Analysis of Molecular Activities at the Single-cell Level,” *Cell Structure and Function*, Vol43, No.2, pp.129-140, 2018