

生体蛍光観察動画像の深度を考慮した 深層学習による細胞追跡精度の改善

嶋田 彩人^{1,a)} 瀬尾 茂人^{1,b)} 繁田 浩功¹ 間下 以大^{1,2} 内田 穰³ 石井 優^{3,4} 松田 秀雄¹

概要: 二光子励起顕微鏡や蛍光タンパク質等の技術発展により、生体内の細胞をリアルタイムで観察可能となっている。細胞動態の解析を行うためには、まず細胞の動きを自動で追跡することが基本的なタスクである。しかし、観察する対象や用いる技術に依存して画像の見た目や細胞の特徴や性質が大きく変化するため、追跡のためのアルゴリズムや特徴量を個別に設計する必要があった。この問題に対して、追跡対象の物体の特徴を随時学習しながら追跡を行う教師あり学習による手法が提案されており、Action-Decision Network (ADNet) は畳み込みニューラルネットワークと強化学習を組み合わせる物体追跡を行う方法の1つである。

本研究では、転移学習により ADNet を蛍光生体イメージングで得られた動画像での細胞追跡に応用する方法を提案する。また、3次元の細胞動画像の深度の情報を疑似的な深度マップとして入力に用いることで、細胞の追跡精度が向上することを示す。

キーワード: 細胞画像処理, 物体追跡, 深層学習, 蛍光生体イメージング, 畳み込みニューラルネットワーク, 強化学習, 最大値投影法, 深度マップ

1. はじめに

近年の二光子励起顕微鏡を始めとした光学観察技術の発展により、多種多様かつ大量の細胞動画像が日々産出されている。例えば、二光子励起顕微鏡で生きたマウスの骨組織内を観察することにより、骨代謝における破骨細胞と骨芽細胞の分布と活動のバランスの変化の様子が [1]、また皮膚を観察することにより、顆粒球等の免疫細胞の動きが急性炎症やアレルギーといった条件によって変化する様を観察することができる [2]。このような動画像中から細胞の動きを抽出した結果を解析することで、薬効の定量的な評価や病態の解明に貢献することが期待される。一方で、バイオイメージングのハイスループット化が進む現在、手動による細胞の追跡は負担が大きく、また観測者の主観によるバイアスも危惧されるため、画像処理による自動追跡

技術の開発は喫緊の課題となっている。

二光子励起顕微鏡は、高い組織透過性と低い組織侵襲性により、生体内部の観察に適した方法である。蛍光の励起は焦点面でのみ起こるため Z 軸方向に高い空間解像度を持ち、焦点を Z 方向に移動しながら断層像の繰り返し撮影 (スタック) することで、対象を 3 次元的に観察することができる。加えて一定の時間間隔をおいて撮影を続けることにより 3 次元の動画像を得ることができ、これをタイムラプスという。ただし、3 次元でのタイムラプスの場合は、時間的な制約が存在し、Z 方向へのスタックの枚数とタイムラプスの時間間隔は反比例することになる。特に免疫細胞のような移動する細胞の動態を観察するような場合は、細胞が大きく動いてしまう前に次フレームの撮影をはじめなければいけないため、Z スタックの枚数は必然的に少なくなることになる。よってこのような場合には、そもそも Z 方向に広く観察することは行わず、比較的狭い領域の Z スタックを撮影し、最大輝度投影法 (Maximum Intensity Projection, MIP) 処理を行って 2 次元に投影した動画を解析対象とすることが行われる。MIP によって Z 方向の情報は失われるものの、画像としては S/N 比が向上し、細胞の検出が行いやすいという利点がある。

このような蛍光生体イメージングにおいても、観察する対象は多岐にわたるため、個別に適したアルゴリズムを設

¹ 大阪大学大学院情報科学研究科
Graduate School of Information Science and Technology, Osaka University

² 大阪大学サイバーメディアセンター
Cybermedia Center, Osaka University

³ 大阪大学大学院医学系研究科
Graduate School of Medicine, Osaka University

⁴ 大阪大学大学院生命機能研究科
Graduate School of Frontier Biosciences, Osaka University

a) a-simada@ist.osaka-u.ac.jp

b) senoo@ist.osaka-u.ac.jp

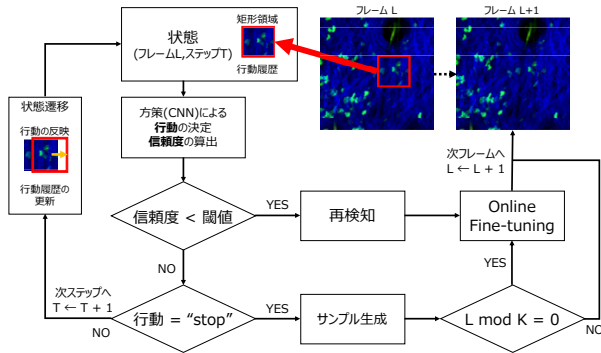


図 1 ADNet による追跡の流れ

Fig. 1 Pipeline of the tracking using ADNet.

計・調整するのは困難である。そこで本研究では、一般物体追跡のために開発された深層強化学習に基づく方法である ADNet[3] を、蛍光生体イメージング画像における細胞追跡へと応用する方法を提案する。これは初期値として与えた対象の特徴を随時更新しながら追跡を行う方法であり、細胞の特徴を個別に設計する必要がないのがメリットである。本研究では、一般物体追跡のデータセットによってニューラルネットワークを事前学習し、これを皮膚内の細胞の遊走を観察したデータに転移学習することで、細胞の追跡が可能であることを示す。また、蛍光生体イメージングにおいては、いずれの細胞も単色の輝点の集合として観察されることから、それぞれの細胞を見た目で区別することが難しいという問題もある。このアピランス情報の乏しさを補うために、本研究では MIP の際の深度の情報を疑似的な深度マップとして入力に加えることで、細胞の追跡精度の向上を試みる。

2. ADNet を用いた一般物体追跡

2.1 問題設定

ADNet を用いた際の物体追跡問題は、マルコフ決定過程に従う。マルコフ決定過程は、ステップ t における行動 $a_t \in \mathcal{A}$ 、状態 $s_t \in \mathcal{S}$ 、状態遷移関数 $s_{t+1} = f(s_t, a_t)$ 、報酬関数 $r(s_t)$ の 4 要素によって定義される。

図 1 は ADNet の追跡および、追跡中の処理の流れを表したものである。ADNet では、物体の位置を矩形の操作によって追跡する。状態 (矩形の位置/行動履歴) が方策に入力され、行動 (矩形の変形/移動) が決定される。決定された行動をもとに矩形が移動され、行動履歴が更新されることによって次の状態へと遷移する。この一連の流れをステップと呼び、1 フレーム中に数回のステップが実行される。行動を決定するための方策は教師あり学習と強化学習によって訓練される。

方策の学習については 2.2 節で説明を行い、以下では ADNet を用いた物体追跡問題における各要素の定義について説明を行う。

2.1.1 行動

行動空間 \mathcal{A} は以下の 11 種類の行動によって構成される。

- 矩形を上下左右に移動させる (4 種類)
 - 矩形を上下左右に 2 回移動させる (4 種類)
 - 矩形を拡大/縮小する (2 種類)
 - 現在フレームにおける矩形の移動を終了する (1 種類)
- それぞれの行動は 11 次元の one-hot 表現の形で扱われる。

2.1.2 状態

状態 s_t は (p_t, d_t) の組で定義される。ここで、 p_t はステップ t における画像中の矩形領域を表しており、 d_t はステップ t における過去 k 回の行動を表すベクトルである。矩形 p_t は、4 次元ベクトル $[x_t, y_t, w_t, h_t]$ で表現される。ここでは画像左上を原点とした座標系を用いており、 (x_t, y_t) はステップ t における矩形左上の座標を、 (w_t, h_t) はステップ t における矩形の幅、高さをそれぞれ表している。行動履歴 d_t における、それぞれの行動 a_{t-i} ($i = 1, 2, \dots, k$) は one-hot 表現として保存される。

2.1.3 状態遷移

ステップ t において、状態 s_t のもとで行動 a_t が選択された場合、状態遷移関数 $f(s_t, a_t)$ に従いステップ $t+1$ における状態 s_{t+1} が決定される。矩形の移動に関する状態遷移 $p_{t+1} = f_p(p_t, a_t)$ はそれぞれ次式で表される。

$$x_{t+1} = x_t \pm \alpha w_t \quad (1)$$

$$y_{t+1} = y_t \pm \alpha h_t \quad (2)$$

ここで、 α は移動倍率を表すパラメータである。

また、矩形 p_t の拡大縮小が選択された時の状態遷移は次式で表される。

$$x_{t+1} = x_t \pm \frac{\beta - 1}{2} w_t \quad (3)$$

$$y_{t+1} = y_t \pm \frac{\beta - 1}{2} h_t \quad (4)$$

$$w_{t+1} = \beta w_t \quad (5)$$

$$h_{t+1} = \beta h_t \quad (6)$$

β は拡大倍率を表しており、拡大が選択された場合は加算を、縮小が選択された場合は減算を行う。「終了」の行動が選択された場合は、現在フレームの状態遷移は終了し、報酬を受け取り次フレームの初期状態へと遷移する。また、状態遷移において、矩形の移動と同時に行動履歴 d_t の更新が $d_{t+1} = f_d(d_t, a_t)$ に従って行われる。行動履歴は k ステップの行動を保存するので、最も古い行動である a_{t-k} が破棄され、行動 a_t が新たに保存される。

2.1.4 報酬

あるフレーム l における真値 g_l が存在する場合、報酬関数は以下のように定義される。 T_l フレーム l における最終ステップ、つまり「終了」の行動が選択されたステップ $T_l = (t | a_{t,l} = \text{“終了”})$ とした時、

$$r(s_{t,l}) = \begin{cases} 0 & (t \neq T_l) \\ 1 & (IoU(p_{T_l}, g_l) > 0.7) \\ -1 & (IoU(p_{T_l}, g_l) \leq 0.7) \end{cases} \quad (7)$$

IoU は Intersection-over-Union(IoU) を表しており、 $IoU(p_a, p_b) = \frac{|p_a \cap p_b|}{|p_a \cup p_b|}$ で定義される。 $|p_a \cap p_b|$ 、 $|p_a \cup p_b|$ はそれぞれ2つの矩形領域 p_a 、 p_b の積集合部分のピクセル数、和集合部分のピクセル数を表している。ここで、ADNetを用いた物体追跡問題における報酬関数は、行動 a_t や遷移後の状態 s_{t+1} には依存しない。また、真値 g_l が存在しない場合は、「終了」の行動が選択された場合においても報酬 0 が与えられる。

2.2 ネットワークの学習

2.2.1 教師あり学習のためのラベル生成

教師あり学習において、ADNet は出力を2つ持つマルチタスク学習を行う。与えられた矩形領域に対して、一方では最適な「行動」を選択するようなタスクを学習し、もう一方では追跡対象である確率を表す「信頼度」を出力するタスクを学習する。そのためには、追跡対象物体を含む矩形（正のクラス）と含まない矩形（負のクラス）を多数生成し、クラスラベルをつける必要がある。また、それらの矩形に対して適切な行動ラベルをつける必要がある。

物体追跡問題における真値は矩形領域の座標として与えられ、それらを基に行動ラベルおよびクラスラベルを生成する。真値の矩形 g から座標やスケールを変化させた N 個の矩形 $\{p^{(i)}\}_{i=1}^N$ を生成する。生成された矩形に対して、2.1.1 節で定義した全ての行動を用いて矩形の移動を行う。その中で最も IoU が高かった行動を、矩形と対応する行動ラベル $a^{(i)}$ として扱う。

$$a^{(i)} = \arg \max_{a \in \mathcal{A}} IoU(f_p(p^{(i)}, a), g) \quad (8)$$

ここで $f_p(p, a)$ は矩形の移動に関する状態遷移関数を表している。さらに、生成された矩形 $p^{(i)}$ に対して、真値である矩形 p と一定より高い IoU を持つ場合は正のクラス、そうでない場合は負のクラスとして扱う。

$$y^{(i)} = \begin{cases} 1 & (IoU(p^{(i)}, g) > 0.7) \\ 0 & (otherwise) \end{cases} \quad (9)$$

$y^{(i)}$ は矩形領域 $p^{(i)}$ に対応するラベルを表している。これらによって、真値となる矩形領域から、矩形領域 $p^{(i)}$ 、行動ラベル $a^{(i)}$ 、クラスラベル $y^{(i)}$ を組として持つサンプル $(p^{(i)}, a^{(i)}, y^{(i)})$ が生成できる。

2.2.2 教師あり学習によるパラメータ更新

ニューラルネットワークは、2.2.1 節において説明した2種類のタスクについて、マルチタスク学習を行う。訓練に使用されるミニバッチは毎エポック、それぞれの動画像からサンプル (p_i, a_i, y_i) を N 個ランダムに抽出し作成する。

また、損失関数は以下のように定義されており、確率的勾配降下法を用いて最小化を行う。

$$L = \frac{1}{N} \sum_{i=1}^N L_1(a_i, \hat{a}_i) + \frac{1}{N} \sum_{i=1}^N L_2(y_i, \hat{y}_i) \quad (10)$$

ここで、 L_1, L_2 はクロスエントロピー関数を、 a_i, y_i はそれぞれ行動/クラスのラベルを、 \hat{a}_i, \hat{y}_i はそれぞれニューラルネットワークによって予測された行動/クラスを表している。

2.2.3 強化学習によるパラメータ更新

2.2.2 節で説明した方法によって教師あり学習によるニューラルネットワークの訓練を行った後、ネットワークをパラメータ化された方策として扱い、このパラメータを強化学習により更新していくことで方策を改善する。訓練済みのネットワークのパラメータを初期値に方策ネットワークとして強化学習を行う。強化学習のための事前学習として教師あり学習を行っておくことで、訓練が効率的に進行する可能性がある事が確認されている [4]。ここでは現在の方策を用いて追跡を行い、その結果によって報酬を受け取りパラメータを更新する。強化学習の目的は方策ネットワークの改善であるため、教師あり学習において訓練を行った信頼度を出力するタスクについては、強化学習において訓練は行わない。

強化学習におけるパラメータ更新および追跡シミュレーションの手順について、ここでは詳細な記述を行う。まず、シミュレーションに用いる動画像について、 L フレームのシーケンス $\{V_l\}_{l=1}^L$ およびそれらの真値 $\{G_l\}_{l=1}^L$ をランダムに抽出する。この時、シーケンス中の全フレームに真値を持っている必要は無く、最終フレームにのみ最低限真値を持っていれば良い。抽出したシーケンスに対して現在の方策ネットワークを用いて、追跡を行う。シーケンスにおける l フレーム目のステップ t における状態を $s_{t,l}$ とした時、これを方策となるニューラルネットワークに入力することで、行動 $a_{t,l}$ を獲得する。この時、以下の式に従って行動が選択される。

$$a_{t,l} = \arg \max_{a \in \mathcal{A}} p(a|s_{t,l}; W) \quad (11)$$

ここで、 $p(a|s)$ はニューラルネットワークの出力となる確率分布を、 W はニューラルネットワークにおける現在のパラメータを表している。状態 $s_{t,l}$ と行動 $a_{t,l}$ が決定した後、状態遷移関数によってステップ $t+1$ の状態 $s_{t+1,l}$ が決定される。また、各ステップの行動後には報酬が与えられるが2.1.4 節に従い、「終了」の行動が選択された場合は1もしくは-1の報酬が与えられ、それ以外の場合は0が与えられる。

1つのシーケンスの追跡が終了後、与えられた報酬を用いて方策となるニューラルネットワークのパラメータ W の更新を行う。パラメータ更新は確率的勾配上昇法 [5] に

よって行われ、勾配は以下の式で計算される。

$$\Delta W = \frac{1}{L} \sum_{l=1}^L \frac{1}{T} \sum_{t=1}^T \frac{\partial \log p(a_{t,l} | s_{t,l}; W)}{\partial W} \rho_{t,l} \quad (12)$$

ここで、 T はステップ数を表しており、 $\rho_{t,l}$ は以下の式で定義される。これは、フレーム l の最終ステップ T における報酬を表している。

$$\rho_{t,l} = r(s_{T,l}) \quad (13)$$

3. 生体蛍光観察動画の深度を考慮した細胞追跡精度の改善

3.1 概要

本手法では、一般物体データセットによって訓練されたADNetを拡張する。教師あり学習および強化学習を用いて細胞動画データセットの転移学習をする事で、3次元蛍光動画中の細胞追跡を行う。ネットワークの訓練および物体追跡の流れは、2章で示した手順に従う。3次元蛍光動画をMIPによって2次元画像に投影して解析するという特性上、Z軸の情報が失われてしまい、細胞が密集しているような状況下では追跡は困難になる。この問題を解決するために、MIPを行う際にZ軸方向の情報を擬似的な深度として保持し、ネットワークの入力に用いる。これによってMIP処理された蛍光動画を扱いながら、本来であれば失われるZ軸方向の深度情報を考慮した訓練を行うことができ、より正確な細胞追跡が可能になる。

3.2 深度成分の入力

MIPによって、Z軸方向の情報は失われてしまう。そこで、MIPにおいて投影される最大輝度値を持つz座標を保存しておき、それを物体追跡のための入力として加えることで、Z軸方向の情報を考慮した追跡を行う。ここでは、入力として扱うz座標の情報を疑似深度と呼ぶ。

疑似深度情報は、追跡対象となる細胞を表す蛍光のチャンネルから計算する。疑似深度画像は以下の式によって求められる。

$$D(x, y) = \arg \max_z I(x, y, z = k) \quad (k = 1, 2, \dots, Z) \quad (14)$$

ここで、 $D(x, y)$ は疑似深度画像の座標 (x, y) における値を、 $I(x, y, z)$ は対象とするチャンネルの3次元画像における座標 (x, y, z) の輝度値を表している。また、本手法は2次元動画中の物体追跡のためにも3次元動画画像同様に用いることができる。その場合は疑似深度を画像中の全画素について0として扱う。

3.3 ネットワーク構造

図2は、本手法で使用したCNNの構造を示している。こ

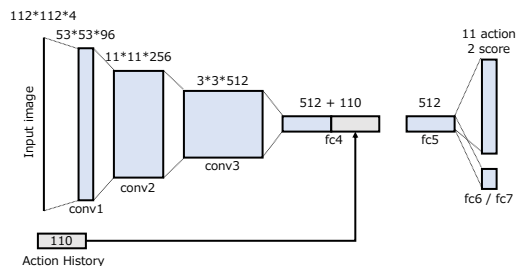


図2 ネットワーク構造

Fig. 2 Network Structure.

れは、図1の方策に該当する。ネットワーク構造はVGG-Mモデル[6]を用いている。全体で3層の畳み込み層(conv)と4層の全結合層(fc)からなり、fc4層およびfc5層の直後にはDropout層が使用されている。活性化関数には出力層ではソフトマックス関数を、それ以外の層ではReLUを用いている。畳み込み層の各パラメータは、ImageNet[7]によって訓練されたパラメータによって初期化されている。また、画像の入力部分は疑似深度を含むため、本手法では4チャンネルとなっている。ネットワークに入力される画像は矩形を112×112の大きさにリサイズしたものである。行動履歴の入力部分は、11行動の10ステップからなる110次元のベクトルで表現されており、入力画像の畳み込みで得られた特徴量とfc4層で結合されている。最終層であるfc6層では行動を表す確率分布を、fc7層では信頼度をそれぞれ出力する。

4. 実験

本節では提案手法の有効性を確認するための実験についての記述、および実験結果についての考察を行う。一般物体データセットおよび細胞動画データセットを用いて、深度を考慮することによる有効性を確認するための実験を行った。

4.1 実験データ

4.1.1 一般物体画像のデータセット

一般物体動画を用了ネットワークの訓練のために、VoT2013[8]、VoT2014[9]、VoT2015[10]、ALOV300[11]と呼ばれる4種類のデータセットを用いた。それぞれのデータセットには重複している動画が含まれているが、訓練時には重複を取り除き、合計371種類の一般物体動画を訓練に利用した。

4.1.2 細胞画像のデータセット

提案手法の学習と追跡精度の評価を行うため、皮膚の炎症を誘導したマウスの耳介を二光子励起顕微鏡によって観察/撮影した動画をデータセットとして用いた。EGFP(緑色蛍光タンパク質)によって緑色に標識されているのは主に顆粒球であり、この細胞の遊走を追跡するのが本実験の目的である。また、青色は第二次高調波発生(Sec-

表 1 実験に使用したパラメータ

Table 1 Parameters used for experiments.

パラメータ	値
教師あり学習エポック数 (一般物体)	30
強化学習エポック数 (一般物体)	30
教師あり学習エポック数 (細胞)	50
強化学習エポック数 (細胞)	10
矩形の移動倍率	0.03
学習率 (畳み込み層)	0.0001
学習率 (全結合層)	0.001
正クラスの矩形生成数 (Train)	150
正クラスの矩形生成数 (Test, フレーム $l=1$)	200
正クラスの矩形生成数 (Test, フレーム $l>1$)	30
負クラスの矩形生成数 (Train)	50
負クラスの矩形生成数 (Test, フレーム $l=1$)	50
負クラスの矩形生成数 (Test, フレーム $l>1$)	15
再検知とみなされる IoU の閾値	0.5
Fine-tuning が行われるフレーム間隔 (一般物体)	20
Fine-tuning が行われるフレーム間隔 (細胞)	5

ond Harmonic Generation, SHG) を捉えており、皮膚ではコラーゲンの線維が特異的に可視化されている。赤色は EGFP からのシグナルの漏れ込みや自家蛍光などである。動画の解像度は XY 方向に 512×512 ピクセルであり、1 ピクセルは $0.5 \mu\text{m}$ に相当する。また、Z 方向には $3 \mu\text{m}$ 間隔で約 15 枚撮影したスタックである。1 本の動画は、1 分間隔の 20 フレームからなる。動画は 15 本あり、各動画ごとにそれぞれ 5 個程度の細胞についてアノテーションされており、訓練と評価に用いるシーケンスの数は 81 系列となっている。

4.2 実験設定

一般物体動画画像データセットを用いてネットワークを訓練した後、その重みを初期値として細胞動画画像データセットを用いて転移学習を行う。細胞動画画像の訓練および評価のために、データセットを動画画像単位で 2 分割して交差検証を行った。表 1 に実験に使用したパラメータを示す。本手法で提案している追跡器は Python および深層学習フレームワークである PyTorch を用いて実装し、全ての実験は i7-6900K 3.20GHz CPU, 128GB RAM, Quadro GP100 GPU を搭載した計算機上で行われている。

4.3 評価手法

追跡器の性能を測るために、One-Pass Evaluation (OPE) [12] と呼ばれる手法を用いて Success rate と、Precision の 2 種類の指標に基づく評価を行う。Success rate は、あるシーケンスにおいて真値と予測結果との IoU が、与えられた閾値より高いフレームの割合を表す。Precision は、あるシーケンスにおいて真値と予測結果の中心間のユークリッド距離が、与えられた閾値よりも小さいフレームの割合を

表す。OPE では、評価に用いる全シーケンスについてこれらの値を、様々な閾値で計算しプロットする。また、定量的にプロット同士を比較するため、Success rate においては Area Under Curve(AUC) を、Precision においては閾値を 20 ピクセルにした時の評価値を用いている。

4.4 実験結果と考察

図 3 は、細胞動画画像データセットを用いた 2-fold 交差検証による性能評価である。“ADNet+Depth” は本論文で提案する疑似深度を入力した ADNet を、“ADNet” は従来の ADNet を表している。また、“SL” は教師あり学習のみ、“SL+RL” は教師あり学習と強化学習を用いて訓練したネットワークによる追跡結果を表している。疑似深度を用いた “ADNet+Depth” は学習の方法に関わらず、疑似深度を用いていない “ADnet” よりも Success rate, Precision とともに性能が改善していることが分かる。一方で、“ADNet+Depth” を用いた追跡において、教師あり学習と強化学習による訓練を行った場合は、教師あり学習のみの場合と比較して追跡性能が落ちている事が確認できる。これは、疑似深度の利用が、画像中の細胞の位置と行動の組み合わせの学習、つまり教師あり学習の精度向上に貢献している。しかし、強化学習によってシーケンスを用いた学習を行う際に、細胞の不規則な移動方法に対しての対応が出来ておらず、訓練によって方策が劣化してしまっている為であると考えられる。

図 4 は、追跡結果の例である。画像は最終フレームであり、1 フレーム目から最終フレームまでの矩形の追跡経路を画像上に出力している。疑似深度を用いることによって、図 4 上段のように周囲の細胞を誤認識せずに追跡が成功している。一方で、図 4 下段のように、追跡が失敗している例も挙げられる。誤追跡の一因として、本手法では追跡を行う矩形のアスペクト比が固定されている点が考えられる。細胞は一般物体とは異なり変形が頻繁に生じる。それに合わせて真値となる矩形も大きく変形することが多い。毎フレーム大きく変形する細胞に対して、本手法では初期フレームのアスペクト比を変更するような矩形の操作はできないため、追跡が困難な場合が起こりうると考えられる。

5. おわりに

本論文では、蛍光動画画像中の細胞追跡のために深度を考慮した深層学習手法を提案した。これは矩形の操作によって物体追跡を行う ADNet を拡張した方法である。細胞の追跡において、細胞動画画像のみからなる訓練に十分なデータセットを準備することは困難であるが、一般物体追跡のタスクのために公開されているデータセットを用いて訓練したモデルから転移学習を行うことで、比較的少数の細胞動画画像から細胞の追跡が可能になった。また、MIP によって 3 次元動画画像を 2 次元平面に投影する際に、本来であ

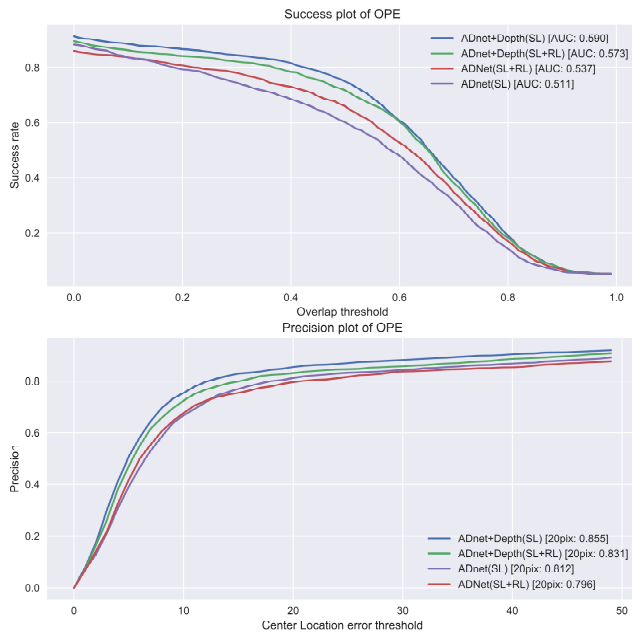


図 3 細胞動画画像データセットにおける Success plots と Precision による追跡手法の比較

Fig. 3 The success plots and precision of OPE on cell dataset.

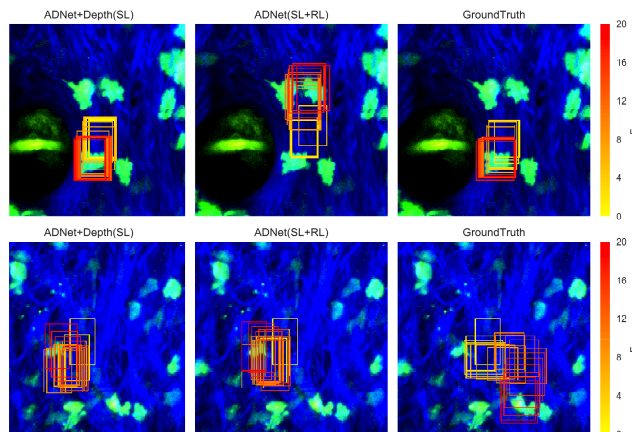


図 4 追跡結果の例

Fig. 4 Examples of tracking result and ground truth.

ば Z 軸方向の情報は失われてしまい、細胞が高密度に動画画像内に写っているような場合の対応が難しいが、最大値の Z 座標を疑似的な深度としてニューラルネットワークに入力することで、Z 軸方向の情報を考慮した追跡を行い、精度が向上することを確認した。実験では一般物体動画画像、細胞動画画像のデータセットを用いた訓練をしたネットワークを用いて、細胞動画画像データセットによる交差検証を行う事で、細胞追跡における疑似深度の有効性を確認した。

今後の課題として、細胞の不規則な挙動に対してアスペクト比を追跡中に変形できるような操作を訓練する事が挙げられる。また、細胞は観察中に分裂や消滅するといったような現象が確認できる場合がある。こういった現象を合わせて捉えられる方策の構築が挙げられる。

参考文献

- [1] Furuya, M., Kikuta, J., Fujimori, S., Seno, S., Maeda, H., Shirazaki, M., Uenaka, M., Mizuno, H., Iwamoto, Y., Morimoto, A. et al.: Direct cell-cell contact between mature osteoblasts and osteoclasts dynamically controls their functions in vivo, *Nature communications*, Vol. 9, No. 1, p. 300 (2018).
- [2] Kreisel, D., Nava, R. G., Li, W., Zinselmeyer, B. H., Wang, B., Lai, J., Pless, R., Gelman, A. E., Krupnick, A. S. and Miller, M. J.: In vivo two-photon imaging reveals monocyte-dependent neutrophil extravasation during pulmonary inflammation, *Proceedings of the National Academy of Sciences*, Vol. 107, No. 42, pp. 18073–18078 (2010).
- [3] Yun, S., Choi, J., Yoo, Y., Yun, K. and Young Choi, J.: Action-Decision Networks for Visual Tracking with Deep Reinforcement Learning, *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2017).
- [4] Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., Van Den Driessche, G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M. et al.: Mastering the game of Go with deep neural networks and tree search, *nature*, Vol. 529, No. 7587, p. 484 (2016).
- [5] Williams, R. J.: Simple statistical gradient-following algorithms for connectionist reinforcement learning, *Machine learning*, Vol. 8, No. 3-4, pp. 229–256 (1992).
- [6] Chatfield, K., Simonyan, K., Vedaldi, A. and Zisserman, A.: Return of the devil in the details: Delving deep into convolutional nets, *arXiv preprint arXiv:1405.3531* (2014).
- [7] Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A. C. and Fei-Fei, L.: ImageNet Large Scale Visual Recognition Challenge, *International Journal of Computer Vision (IJCV)*, Vol. 115, No. 3, pp. 211–252 (2015).
- [8] Kristan, M., Pflugfelder, R., Leonardis, A., Matas, J., Porikli, F., Cehovin, L., Nebhay, G., Gustavo, F. and Vojir, T.: The Visual Object Tracking VOT2013 challenge results, *Workshop on the VOT2013 Visual Object Tracking Challenge*, pp. 98–111 (2013).
- [9] Kristan, M., Pflugfelder, R., Leonardis, A., Matas, J., Cehovin, L., Nebhay, G., Vojir, T. and Fernandez, G.: The Visual Object Tracking VOT2014 challenge results, *Workshop on the VOT2014 Visual Object Tracking Challenge*, pp. 191–217 (2014).
- [10] Kristan, M., Matas, J., Leonardis, A., Felsberg, M., Cehovin, L., Fernandez, G., Vojir, T., Hager, G., Nebhay, G. and Pflugfelder, R.: The visual object tracking vot2015 challenge results, *Proceedings of the IEEE international conference on computer vision workshops*, pp. 1–23 (2015).
- [11] Smeulders, A. W., Chu, D. M., Cucchiara, R., Calderara, S., Dehghan, A. and Shah, M.: Visual tracking: An experimental survey, *IEEE transactions on pattern analysis and machine intelligence*, Vol. 36, No. 7, pp. 1442–1468 (2014).
- [12] Wu, Y., Lim, J. and Yang, M.-H.: Online object tracking: A benchmark, *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2411–2418 (2013).