

平成 30 年 6 月 13 日現在

機関番号：12602

研究種目：挑戦的萌芽研究

研究期間：2016～2017

課題番号：16K12530

研究課題名(和文)非裁量的な蛍光顕微鏡画像定量解析法の開発

研究課題名(英文)Methodological study of non-arbitrary quantification for fluorescent microscopy

研究代表者

二宮 洋一郎(NINOMIYA, Youichirou)

東京医科歯科大学・歯学部・非常勤講師

研究者番号：90237777

交付決定額(研究期間全体):(直接経費) 1,600,000円

研究成果の概要(和文):生命現象を研究する上で、実験や観察の結果を定量化することは不可欠である。従来、実験結果の定量化には操作する人、つまり術者の裁量が入るため、全く同じ結果であっても異なる術者が定量化すると得られる数値が違ってくる。そこで、機械学習を利用した定量化法を開発し、そのプロトタイプを蛍光顕微鏡画像の定量化に適用した。

画像を一定の大きさの矩形タイルに分割し、それぞれのタイルの輝度値の分布を中央値と四分位範囲の2つの要約統計量で表現した。この統計量を用いて機械学習を行い、タイルをパターン毎にクラス分類した。観察目的に合致するクラスのタイルのみを対象に数値データを取得し、術者の裁量に拠らない定量化を実現した。

研究成果の概要(英文):Non-arbitrary, non-biased measurement is an essential property for quantification of fluorescent microscopy images. Conventional methods of the image quantification has been heavily relied on human ability of cognition, and that means there are a plenty of room to errors and inconsistencies.

To eliminate these inconsistencies and to achieve the non-arbitrary quantification, I developed a non-supervised machine learning (ML) based method that utilized median and IQR (interquartile range) intensities of the fixed-size tiles. These image tiles were successfully classified by ML to few clusters. Selection of appropriate cluster can lead an extraction of particular group of tiles, which corresponds suitable set of biologically meaningful regions. The study indicated that ML-based classification combined with descriptive statistics values of tiling images, such as median and IQR, should be a non-arbitrary, non-biased quantification of the biological and medical images.

研究分野：発生生物学

キーワード：バイオイメージインフォマティクス 蛍光強度 要約統計量 クラス分類 機械学習 ベイズモデル

### 1. 研究開始当初の背景

蛍光顕微鏡画像の定量解析は現代の生物学の研究に不可欠な作業である。通法では細胞核や細胞質の染色像を利用して個々の細胞像をセグメントとして同定 (cell segmentation) し、目的分子の蛍光強度をそれぞれのセグメントについて計測する。この方法では、個々の細胞同定の成否がその後の定量解析の質を決定するため、cell segmentation に細心の注意を払い、様々なパラメータを術者が設定してノイズやアーチファクトの排除を試みなければならない。このことは、画像解析の最初の過程で術者のバイアスや裁量の入り込む余地が多く存在し、解析結果の一定性や再現性を損ねる要因となっている。さらに、cell segmentation のアルゴリズムやパラメータセットは対象となる細胞や組織の種類によって最適解が異なり、処理対象が異なるたびに術者による処理の最適化が必要となるため、画像処理のユニバーサル性の確保やハイスループットワークフローへの組み込みを難しくしている (Meijering 2012)。

### 2. 研究の目的

そこで、cell segmentation を行うことなく蛍光顕微鏡画像を定量計測できるアルゴリズムの開発を目的とした。まず、画像を一定の大きさのタイル(以下 grid と呼ぶ)の集合体に分割し、その個々の grid を構成する画素の蛍光強度を計測するアルゴリズムを編み出した。このアルゴリズムを発生途上のマウス胚の間葉組織に発現する遺伝子産物を定量解析する研究に適用し、野生型と変異型の表現型の差を定量化することに成功した (Zhao et al. 2015)。本研究課題では、このアルゴリズムを更に汎用性のある方法へ拡張し、設定するパラメータの数を最小にした方法を探索・開発する。また、その方法の妥当性を通法と比較することによって検証する。

さらに、計測した値を元にその蛍光顕微鏡画像の持つ特徴量を自動的に抽出し、解析の段階も非裁量的に行うことの可能なアルゴリズムを開発する。画像の特徴量を自動的に抽出するには深層学習が良く用いられる。しかし、これらの方法は対象に応じてハイパーパラメータの設定や最適化が必要となるなど、一般的な実験生物学者が日常的に使いこなすには困難な作業が残っている。そこで、本研究課題では既存の統計的分類手法を採用したアルゴリズムを構築し、上記の画像計測アルゴリズムと組み合わせることで、誰が術者であっても同じ画像からは常に同じ解析結果が最小限のパラメータ設定で得られることを目指した。

### 3. 研究の方法

様々な種類の蛍光顕微鏡画像を取得し、本研究課題で開発した画像計測・解析アルゴリズム群を適用する。同時に、cell segmentation が可能な画像については同法による計測・解析を行い、grid を用いた本法との比較検討を行う。細胞密度が高く、cell segmentation を適用し難い画像、例えば未分化 ES 細胞のコロニーや組織切片の顕微鏡画像において、本法がどの程度有効なのか検討する。

(1) grid をベースにした蛍光顕微鏡画像の計測・解析ソフトウェアパッケージを開発し、研究者コミュニティに提供する。計測・解析とともに統計処理を頻用することから、プラットフォームにはオープンソースの統計処理言語 R を採用し、ソフトウェアの提供も R のパッケージとして行う。

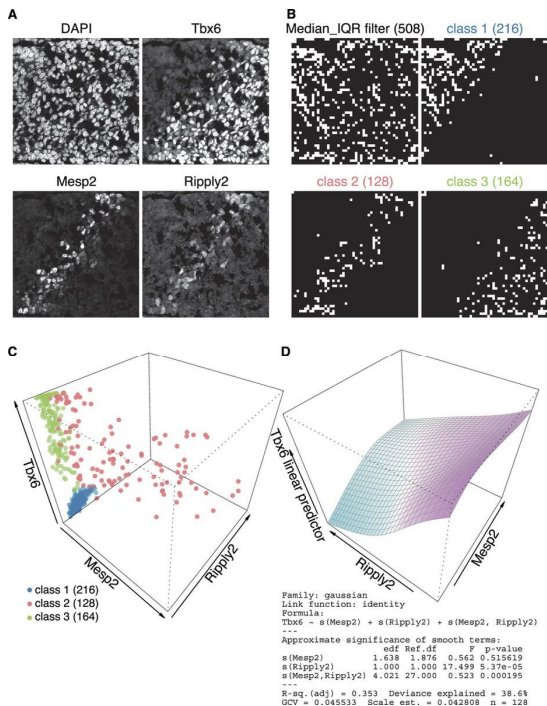
(2) パッケージの改訂・改良を継続する。特に組織切片の蛍光顕微鏡画像の計測・解析を効率よく行えるような改良を重点的に検討し、cell segmentation ではほぼ不可能であった組織切片のハイスループットデータ処理を行えるパッケージを開発する。

(3) 本パッケージの理論的詳細を学術雑誌に投稿し、リファレンスとなる文献を発表する。

### 4. 研究成果

#### (1) 研究の主な成果

Grid をベースとした新しい画像定量化アルゴリズム、GBIQ (grid based image quantification) を開発し、研究者コミュニティにオープンソースとして公開した。このアルゴリズムでは、細胞の特定の画分、たとえば画像の中で細胞核だけを分画する際に grid の要約統計量を元にして特徴量を抽出する。教師なし機械学習を行って特徴量を元にクラス分類を行い、細胞核を含む grid のみを計数することができた。さらに、この細胞核を含む grid のみを解析対象として、細胞核に局在する複数種類の蛋白質の発現量を定量化することができた。この定量化は、従来よく用いられてきた培養細胞に対するものではなく、生体組織の薄切試料に対して行ったものである。Cell segmentation を利用した従来法では、組織切片のように細胞密度が高い試料に適用するには数々のハードルがあり、数多くのパラメータを厳密に追いつく作業が必須であった。GBIQ では設定するパラメータは1つだけであり、術者が誰でも、同じ画像からは常に同じ定量結果を得ることができる。



### 付図説明

A. 細胞核を染色する DAPI の蛍光顕微鏡画像に対して GBIQ を適用し、まず細胞核を含む grid のみを得た。その細胞核をマスクとして、3 つの核蛋白質(Tbx6, Mesp2, Ripply2)の発現強度を定量化した。B. A で得た 3 つの核蛋白質の発現量を説明変数として、は発現パターンの違いを教師なし機械学習で 3 つにクラス分類した。クラス 2 は Mesp2, と Ripply2 が共有している。C. それぞれの核蛋白質発現量を XYZ 軸にとり、クラス分類を色分けでプロットした。D. クラス 2 の grid に関して、統計モデルを設定して 3 つの核蛋白質の相関関係をプロットした。ここでは、Tbx6 と Ripply2 は逆相関関係にあり、Mesp2 と Tbx6 には相関関係はほとんど観察されない。このことは、Ripply2 が Tbx6 の分解に強く関わっており、Mesp2 と Ripply2 はほぼ同じ発現パターンを示すものの、その機能は異なっていることを示唆している。実際、生化学的実験により、Tbx6 の分解には Ripply2 が直接関わり、Mesp2 は作用しないことが確認されている。

さらに、機械学習でクラス分類した grid についてのみ解析対象とすることで、統計モデルの構築検証に必要なデータからノイズを除去し、精度を高めることにも成功した(付図参照)。ここでは、画像の中から発現パターンの異なる細胞核を 3 つのクラスに分類し、そのうちの一つのクラスのみを解析対象とした。この細胞核分画と解析対象の絞り込みの処理に術者の裁量は全くなく、蛍光顕微鏡画像の読み込みから定量化まで術者が指示したパラメータは grid の大きさだけである。この解析対象から得た定量結果を元に統計モデルを作成し、核蛋白質間の相関関係を推定した。その結果、一見同じように発現する

核蛋白質でも、異なるネットワークの構成要因であり、機能は異なることを示すことができた。この画像から推定した機能の相違は、生化学的実験によって正しいことが証明され、GBIQ は画像から表現型の特徴量を抽出し、その相違を描くバイオイメージインフォマティクスのツールとして利用可能であることを示した。

以上の結果をまとめて国際学術雑誌に報告し、あわせてソースコードを公開(<https://github.com/yo-ninomy/DemoScripts>)することで幅広い研究者コミュニティが GBIQ を利用できる仕組みを作った。

### (2) 得られた成果の国内外における位置づけとインパクト

画像を grid に区切って要約統計量を取得し、教師なし機械学習でクラス分類することで、生物・医療画像を非裁量的に定量化することを示した。このアルゴリズムは世界に類例を見ない方法であり、国外においても評価されている。GBIQ は深層学習のようなブラックボックス化した学習過程を用いていないため、どの特徴量をどの程度取り入れてモデルを作ったのか、後解析が簡単にできる利点がある。このことは、付図で例にあげた蛋白質ネットワークの推定にも貢献しており、数枚の蛍光顕微鏡画像から蛋白質の作用を明らかにすることができる強力なツールになり得ることを示している。

### (3) 今後の展望

GBIQ のもつ非裁量性と簡便性を活かして、研究室から日々生成される生物・医療画像を定量化し、これまでバイオインフォマティクスに対して遅れていたバイオイメージインフォマティクスを推進することができる。特に、これまでは難しかった表現型の定量化を大規模に行い、遺伝型に次ぐ表現型のビッグデータ整備に資する研究としたい。また、ベイズモデルによって実験条件の違いをモデル化した統計モデルを構築し、複数の実験結果から生物学的に意味のある定量データを効率良く、かつ簡便に取得できる環境を作りたい。

### <引用文献>

1 Meijering, E. Cell segmentation: 50 years down the road. *IEEE Signal Processing Magazine* 29(5)2012 140-145

2 Zhao, W., Ajima, R., Ninomiya, Y. and Saga, Y. Segmental border is defined by Ripply2-mediated Tbx6 repression independent of Mesp2. *Developmental Biology* 400(1)2015 105-117

### 5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

〔雑誌論文〕(計 1 件)

1 Ninomiya, Y., Zhao, W. and Saga, Y.:  
GBIQ: a non-arbitrary, non-biased method  
for quantification of fluorescent images.  
Scientific Reports 6, 2016, Article  
number: 26454, 査読有  
DOI: 10.1038/srep26454

〔図書〕(計 1 件)

1 独立行政法人情報処理推進機構 AI 白書編  
集委員会編(中島 秀之、浅田 稔、川上 量生、  
北野 宏明、喜連川 優、辻井 潤一、松尾 豊、  
浅川 伸一、麻生 英樹、安宅 和人、和泉 潔、  
石田 亨、石塚 満、井之上直也、上野 達弘、  
尾形 哲也、小田 悠介、金広 文男、河原 達  
也、清水 亮、庄野 逸、武田 英明、田所 諭、  
谷口 忠大、富山 和彦、中原 啓貴、二宮洋  
二郎、野田五十樹、原田 達也、比戸 将平、  
平田 圭二、松井 俊浩、松岡 聡、松原 仁、  
宮尾 祐介、山川 宏、大淵 栄作、大山 洋介、  
國吉 康夫、佐藤 育郎、野村 哲弘、滝澤 豪、  
田中 伸彦、松本真太郎、中 智晴、中村 敬  
子、関根 久、金山 恒二、松田 成正、富田 達  
夫、川浦 立志、米田 健三、和田 恭)、角川  
アスキー総合研究所、AI 白書 2017、2017、  
360

〔その他〕

ホームページ等

蛍光顕微鏡画像の簡便かつ非裁量的な定量  
解析法を開発

[https://www.nig.ac.jp/nig/ja/2016/05/research-highlights\\_ja/20160524.html](https://www.nig.ac.jp/nig/ja/2016/05/research-highlights_ja/20160524.html)

6 . 研究組織

(1)研究代表者

二宮 洋一郎 (NINOMIYA, Youichirou)

東京医科歯科大学・歯学部・非常勤講師

研究者番号：90237777