

## 最適化および機械学習とバイオイメージインフォマティクス応用

内田 誠一

九州大学大学院システム情報科学研究所 〒819-0395 福岡市西区元岡 744  
(2017 年 1 月 28 日受理)

### Bioimage-informatics with Mathematical Optimization and Machine Learning Techniques

Seiichi UCHIDA

Faculty of Information Science and Electrical Engineering, Kyushu University  
Motooka 744, Nishi, Fukuoka, 819-0395, Japan  
(Received on January 28, 2017)

**Abstract :** This paper introduces bioimage-informatics, which is a rather new interdisciplinary research field by collaborations between biology and image informatics. Recent biology often needs to prove their biological findings through appropriate quantitative and objective analysis of bioimages, such as microscopic images or videos. Typical image analysis techniques are not sufficient for the bioimage analysis because bioimages has many difficulties ; for example, less spatio-temporal resolution, more noise, less appearance information, and more ambiguous object boundaries. We, therefore, need to develop new image analysis techniques with enough robustness to deal with those difficulties. An important approach to realize robust analysis is to introduce optimization methods and machine learning methods into image analysis techniques. This paper explains basic idea of those methods and their usefulness for bioimage-informatics.

**Keywords :** Bioimage-informatics, pattern recognition, image analysis, optimization, machine learning

#### 1. はじめに～バイオイメージインフォマティクスとは何か？

バイオイメージインフォマティクスとは、生物学と画像情報学の協働研究分野である。端的に言えば、生命現象を捉えた画像について、画像情報学技術を用いて解析することを目的とする。すなわち主観的な目視観察に留まらずに、客観的で再現性のある定量解析を計算機上で実現する。

生物学における画像観察の利用は、生命現象を直接観察するという意味で、極めて本質的である。特に対象の空間的分布（個数、面積・体積、位置関係、粗密状況など）、アピランス（形状、テクスチャ、色など）、そしてそれらの時間的変化（増減、移動、変形）を把握・解析するには、画像を利用するより他はない。医学における本学会の役目と同様、生物学においてもバイオイメージインフォマティクスの必要性はますます高まってくるであろう。

バイオイメージインフォマティクスが対象とする画像には様々なものがある。その代表は、顕微鏡による生体内部の画像である。本誌の読者には言うまでもないことであろうが、具体的には原子や分子レベルの画像から、オルガネラ、細胞、そしてその集合体としての組織レベルまで、実に様々である。日々高度に進歩するイメージング技術により、ミクロの世界の状況が静止画像や動画として観察されている。

さらにより一般的なカメラで撮影された個体全体の画像も対象となる。例えば線虫やゼブラフィッシュ、マウスなどの挙動観察がその例である。なおカメラは一般的でも、それら生物個体の周囲状況は様々である。線虫ならばシャーレ、ゼブラフィッシュならば水槽、マウスならばオープンフィールド、と言った具合である。こうした周囲状況の違いはその後の画像解析技術にも影響を及ぼす。また、単一個体の場合もあれば、複数個体の場合もあり、こ

れらの差異も用いるべき画像解析に影響する。

言うまでもなく、医用画像情報学とバイオイメージインフォマティクスは非常に類似した分野である。利用される画像情報処理技術も相当共通している。これは医学と生物学が共に生命現象を扱うことから当然とも言えよう。意味がないことかもしれないが、敢えて区別するとすれば、以下のような違いがある。すなわちバイオイメージインフォマティクスでは、線虫やゼブラフィッシュさらには植物や人工生物を扱うなど、より広い生物種を対象とする。その結果、扱う画像試料も多様なようである。さらには、分子細胞生物学レベルからより個体群が織りなすマクロ行動なレベルまでもが解析対象となる。一方で、病理観察や内視鏡観察はあまり一般的ではない。

本稿では、バイオイメージインフォマティクスについて、いくつかの側面から概観する。扱う課題の多様性からそれらの詳細を網羅することは不可能なため、各論には入らず一般論を主とする。これから参画を志す（非情報系の）学生向けとも言える内容なので、物足りない点はご容赦いただきたい。2章および3章では、本章に引き続いて、バイオイメージインフォマティクスという分野を特徴づけるべく、その独特の難しさおよび面白さについて述べる。4章および5章では、少しだけ技術的側面に触れる。具体的には、それぞれ最適化および機械学習に関する。いずれも扱う解析課題の困難性に打ち勝つべく、半ば必然的にバイオイメージインフォマティクスに導入されるものである。ご存知の読者も多いと思うが、それぞれ基本的な考え方と、導入しうる技術の例について述べる。

#### 2. バイオイメージインフォマティクスの難しさ

一般的なカメラで撮影された一般的な情景や物体の画像を扱うのに慣れていると、バイオイメージインフォマティ

クスで扱うバイオ画像の解析には独特の技術的困難性があることに気づく。第一は、分子レベルや細胞内レベルでのミクロな現象を顕微鏡で拡大観察しているため、時空間解像度やコントラストを十分に確保できない点である。これらに対しては、超解像技術や高輝度な蛍光物質の開発、光シート顕微鏡などなども活発に行われている。ただし、こうした技術進歩があると、よりミクロな現象の観察に利用されうる。すなわち、進歩があっても上記の問題は存在しうる。

第二はアピランス情報の圧倒的な不足である。その最たる例は、蛍光観察による分子レベルでの観察である。個々の分子は単なる輝点として観察されるために、(当然ではあるが)それらをアピランスで区別することはできない。輝点数の定量化だけであれば問題ないが、各分子をトラッキングしようとするとなると問題になる[2]。分子ほどではないにしろ、細胞群のトラッキングでも類似の問題は起こる。すなわち同一組織内で隣接する細胞は互いに似通っているため、それらを個々に区別するのは困難である。さらには個体レベルでもアピランス情報は不足する。例えばメダカ群の泳動を捉えて行動習性を解析する場合、個々のメダカを区別する必要がある、そしてそれはやはり容易ではない。ある意味で、雑踏の人流を把握するよりも困難である。メダカは洋服を着てくれないのである<sup>1</sup>。

第三はデータ量の少なさである。より正確には撮影試料の少なさとも言える。解析対象の生物試料を準備し、かつ撮影するには、それ相当のコストがかかる。従って、利用し得る事例の数としては、例えば、数万オーダーを期待するのは無理である。実際には、わずか数例で結論を導き出さなくてはならない課題も多い。なお生物学においても「データ量が多い」という話はよく出てくる。ただし、事例数が多いのではなく、1事例あたりのデータ量が多いことも多い。特に画像のサイズが非常に大きいことが多く、いわゆる4D画像となると、1事例あたり数ギガバイトのファイルであることも珍しくない。この状況をデータ解析の立場から見れば、データが極めて高次元であることを意味しており、事例数の多寡とは無関係である。むしろ高次元になると、後述する機械学習を利用する場合、その次元数に応じてより多くの事例が無くては意味のある結果が出せないため、注意が必要である。

第四は解析課題の特異性である。生物学者は、それぞれ独自の課題を独自の生命体を対象として実施している。従って解析課題は非常に多様である。さらに類似課題であっても撮影機器や撮影状況が異なることもある。このため、「少数の手法を構築しておけば、それらを汎用的に利用できる」というシナリオにはなりにくい。換言すれば、往々にして、個々の課題にテラーメイドでソフトウェア開発をすべきことになる。

第五は正解情報(Ground-truth)の不足である。一般に画像情報学では、正解情報のある画像試料を利用して手法を設計・評価し、その後正解情報のない画像試料においてその手法を実利用する。従って正解情報のある画像が多ければ多いほど、よりよい手法を開発できる。しかしながら、上述の通り、バイオイメージインフォマティクスの課題は多様で一般性に乏しく、大量の正解情報を付与するコストがかかけられない場合が多い。さらに「何が正解か」という点が常に問題になる。医用画像診断であれば、病変していると思われる箇所について組織を採取し、検査することで、そこが真の病変箇所か否かの正解を付けることは可能である。生物学においても同様の診断課題であれば、同様のアプローチが採れる可能性はある。しかし、例えば対

象数のカウントや形状定量化、そしてトラッキングと言ったバイオイメージインフォマティクスに類出する課題においては、対象の状況が曖昧といった理由により、人間ですら一貫した正解を付けられない場合も非常に多い。

### 3. バイオイメージインフォマティクスの面白さ

バイオイメージインフォマティクスには、前節で述べた困難性があったとしても、それらに立ち向かうに足る面白さも多く含まれる。そしてそれらはある意味困難性の双対にもなっている。具体的には以下のとおりである。

第一の面白さは、課題の困難性は、それだけ画像情報学にとって検討課題が多いということを意味している。一般的な画像に比べて、解像度はコントラストが低く、さらにアピランス情報が使えないということは、それだけロバストな手法が望まれるということ意味する。この点は、後述するように「機械学習」や「最適化」といった技法を利活用するモチベーションとなる。なお、既存手法の組合せで済まない場合には、新規な手法を考案する必要があり、画像情報学的にも新たなアイデアをもたらす可能性もある。

第二に、少数事例しか入手できないという状況下で、役立つ手法を開発する必要がある。昨今の深層学習の隆盛については、後述のとおり、単に学習理論の発達だけではなく、データ(事例)が大量にあることが大きい。換言すれば、事例が十分に活用できない場合、十分な事例を前提とした深層学習は利用できないことを意味する。従来は、統計的検定により、少数な事例でもある程度の結論を見出してきた。しかし分布の正規性などいくつかの仮定が暗に含まれている。少数事例の扱いというのは、今後の知能研究・データ解析の中でもますます重要になってくるであろう。

第三に、サイエンスとエンジニアリングの接点としての面白さもある。筆者のようなエンジニアは、常々「効率性」や「精度」、そして「何にどれぐらい役に立つか」が求められている。誤解を恐れずに言えば、生物学者が取り組んでいる研究課題は、そうした工学的価値観では扱えない価値を持っている。もちろんどちらが良いとか悪いという話ではなく、それらが接する場を少しずつ拓いていくという学際協働研究ならではの面白さを強調したい。

第四に、あくまで情報系大学人としての私見であるが、こうした場を学生諸君に経験させることには、大きな教育的効果があると考ええる。情報系の学生諸君の多くは、システムインテグレーター(SIer)として就職する。彼らは、顧客のニーズを把握し、それに対するソリューション技術を提案・実装する必要がある。バイオイメージインフォマティクスのような学際協働研究も同様で、用語の壁を越えながら、解析課題を把握し、真に必要なかつ実現可能な技術を提供する必要がある。こうしたプロセスの経験には、コミュニケーションのセンスを磨く効果が期待できる。

## 4. 数理最適化とバイオイメージインフォマティクス

### 4.1 数理最適化とは

数理最適化とは、目的関数・制御変数・制約条件により数理的に表現される問題(最適化問題)を扱う。具体的には、目的関数(基準)を最大化もしくは最小にするように、制御変数を決定する問題である。その際、変数の取り得る値について、何かしらの制約条件を与えることも多い。ご存知の読者も多いと思われるが、本節ではこれを概説する。

数理最適化は、我々の日常にもあふれている。以下に三

例を挙げる。

- 昼食のメニューを選ぶのも一種の最適化問題である。この場合、「なるべく腹いっぱい」が目的関数の一例であり、「選ぶメニュー」が制御変数であり、「所持金」が制約条件となる。なお、メニューを複数選べる場合、状況は少し複雑になるだろう。
- ある市内に学校を1校開校するときに、どこにその学校を配置するか、という問題もある[3]。この場合、「学校の位置」が制御変数となる。「学校までの距離を全家庭について総和したもの」が最小化すべき目的関数の一例である。このまま無制約かもしれないし、例えば、「消防署からは100メートル以内」といった（不等式）制約条件が入る場合もあるだろう。
- A地点からB地点までの経路を見つける問題（最適経路問題）も典型的な最適化問題である。この場合、「なるべく短時間で」が目的関数の一例であり、「経路」が制御変数（正確には変関数）、そして「経路の平滑性（経路が急に変わったりしない）」が制約条件となる。

最適化問題の解（すなわち目的関数の最大値もしくは最小値を与える制御変数値）は、あらゆる制御変数値の中で、最も望ましいもの、もしくは最も望ましいものに近いものである。従って、「何となく問題を解く」よりは、「最適化問題として最適解を求める」過程を経たほうが、よりよい解が求まる。先の昼食の例で言えば、何となく解くよりは、より腹いっぱいになるメニューを選べるだろう。

上記の三例は、互いに大きく異なった最適化問題であり、それぞれ全く異なった解法で解かれる。メニューの例は「離散最適化（組合せ最適化）問題」、学校は「連続最適化問題」、そして経路は「変分問題」である。こうした違いは制御変数の違いに因る。すなわちそれぞれ、整数値（メニューは有限離散個なので1, 2, のように整数値で考えられる）、連続値（学校は二次元平面上で任意の位置に配置できる）、そして関数（経路は関数で表現される）である。厄介なことに、これらに加え、制約の有無や、問題の凸性、微分可能性の有無などによっても解法がさらに異なってくる。

さらに、大局的最適化と局所的最適化のどちらを目指すかによっても解法は異なる。大局的最適化とは、考え得る全ての解の中で最も良いことが保証された解を与える方法である。一方で、局所的最適化は近似解法とも言え、大局的最適化と同じ解を保証しないものの、なるべくそれに近い解を与える方法である。大局的最適化では計算量が非常にかかる場合でも、局所的最適化ならば高速に準最適解が求まる場合が多い。例えばある組合せ最適化問題について、その大局的最適化には（NP-completeのような）爆発的な計算量が必要とされる場合、局所最適化を用いるほうが現実的である。一方で、大局的最適解のほうがより高精度すなわち基準を最大化もしくは最小化できるため、定式化を工夫することで、多項式時間で大局的最適解が求まるようにすることもある。

## 4.2 数理最適化と画像情報学

さて、画像情報学においても、最適化に基づく方法が様々に利用されてきた。そのメリットを直感的に言えば、最適化を用いれば、可能なすべての画像処理結果の中で、ある基準で最も良い結果を導出してくれる。そもそも一般に画像処理では「こういう処理結果になってほしい」という要請に対してパラメータを制御するのであるから、要請＝目的関数＋制約条件で、パラメータ＝制御変数と考えれば、最適化問題に帰着することになる。さらに、次章で述

べる機械学習も、その基盤は最適化であるから、機械学習に基づく画像処理はすべて最適化に関係する。

最適化応用のわかりやすい例は二値化である。有名な大津の二値化は、あらゆる二値化しきい値の中で、輝度ヒストグラムの分離度を最大化するものを求める。一方、グラフカットによる二値化では、「元々の画素値となるべく変わらない」かつ「隣接画素の白黒がなるべく一致する」の二つの基準を目的関数として、二値化結果画像を求める[4]。このように二つ基準がある場合、両者の荷重和で一つの基準にすることが多い。その際の重みは結果に大きく影響する。例えば、グラフカット二値化の場合、隣接画素の一致性の重みを非常に大きくすれば、画像全体が白もしくは黒一色になる。

他にも、最適化に基づく画像情報処理は、枚挙に暇がない。対象追跡（トラッキング）は、そもそも一種の変分問題である（追跡経路＝変関数）。画像認識も、識別関数の最適化問題である。領域分割も、二値化がその最も単純な場合であることから理解できるように、最適化問題と深い関係がある。例えば、Markov Random Field (MRF) に基づく領域分割法ならば、各画素について、領域に関連した確率変数を定義し、その確率変数値の最適化を図ることで、領域分割結果が求まる。複数画像の位置合わせ問題についても、Low Rank Optimization という比較的新しい最適化問題の枠組みで解かれた例もある[5]。

## 4.3 最適化とバイオイメージインフォマティクス

バイオイメージインフォマティクスにおいては、最適化に基づく画像処理を積極的に利用すべきと思われる。手動でのパラメータ設定や試行錯誤的な処理のほうが手軽に思えるかも知れないが、以下に述べる三つの理由により、最適化を用いたほうが、最初は面倒そうでも結局は近道であろう。

第一に、先述の通りバイオ画像については、低解像度やノイズ、そしてアピランス情報の少なさ等のために通常の画像処理以上に解析が困難であり、従ってそこで用いる画像処理にはロバスト性が求められる。ところで、最適化の目的関数に処理結果に対する「理想」を組み込んでおけば、それになるべく合致した結果が求まる。この理想はpriorと呼ばれる。前節で述べたグラフカット二値化において「隣接画素の白黒がなるべく一致する」という項がpriorに相当する。このpriorを含めて最適二値化することで、ノイズの多い画像であっても、白もしくは黒の領域がごま塩状に点在せず、まとまった領域になる。様々なバイオ解析課題においても、理想的な処理結果の性質をpriorとして定式化して最適化に組み込めば、悪条件に対するロバスト性を確保できる。

第二に、パラメータを自動に最適化することで、再現性や客観性を担保できる点である。先に述べたように、大津の二値化を用いれば、しきい値が最適化の枠組みで自動決定される。この最適しきい値による二値化結果が多くの場合適切であることに加え、同じ画像であれば「誰が何度やっても」同じしきい値そして同じ二値化結果が求まるという点が重要である。すなわち再現性の担保である。さらにしきい値決定の背後には数式で明確に定義された目的関数があり、感覚的なものは一切ない。すなわち客観性の担保である。もちろん最適化の前段階である目的関数の設計の部分は人間が行うのであって、そこに主観が入る余地はある。大津の二値化も、同氏が定めた基準下での最適しきい値である。しかし絶対的に定まる最適基準は一般に存在しないこと<sup>2</sup>を考えれば、現状の枠組みでも十分に客観的

と言える。

第三に、関係が複雑で、試行錯誤的な方法では手に負えないような問題も、最適化を使うことでシステマティックに解ける場合がある。先の二値化の問題にも複雑な関係が隠れている。すなわち、「隣接画素の白黒がなるべく一致する」という要請があったとすると、Fig.1に示すように、ある画素  $p_{ij}$  の白黒の決定には、上に隣接する画素  $p_{i-1,j}$  が影響する。そして連鎖的に、その左に隣接する画素  $p_{i-1,j-1}$  が影響し、さらにはその下  $p_{i,j-1}$ 、そしてその右  $p_{i,j+1}$ 、すなわち自分自身が影響することになる。要するに白黒決定時に考慮すべき隣接関係がループ構造になっており、謂わば卵と鶏の関係になっている。こうした構造が全画素において存在するため、試行錯誤的な方法では全く手に負えないことになる。これに対し、グラフカットを用いた二値化ならば、すべての関係を考慮しながらも一括して全画素の白黒を最適に決定できる。

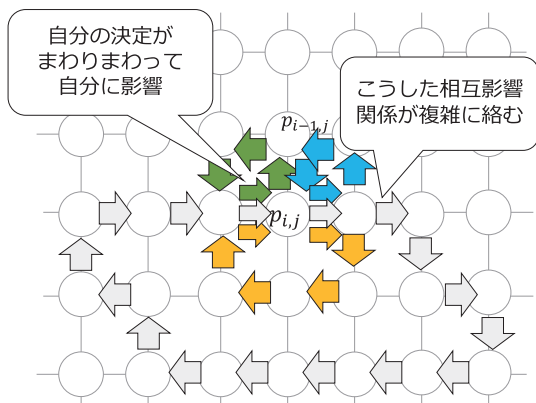


Fig.1 試行錯誤的な二値化を困難にする画素間の依存関係

#### 4.4 最適化の応用例

複雑な関係があっても最適化の枠組みでシステマティックに解ける場合について、筆者らの試みを一つ紹介したい。これは一種の多物体同時追跡問題である。メダカ  $K$  匹が水槽を泳動する動画が与えられ、各個体の移動軌跡を求める問題を想像いただければよい。いま、各時刻（フレーム）で、 $K$  匹それぞれが画像中のどこにいるかを検出できたとして、あとは隣接フレーム間でそれらを繋いでゆけばよい。個体数  $K=3$  の場合の様子を Fig.2 に示す。

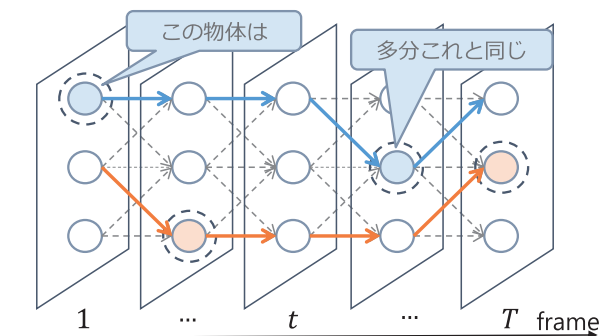


Fig.2 多物体同時追跡問題における非マルコフ的制約

この繋いでゆく処理は、想像以上に困難である。各メダカのアピランスが明確に違っていれば、それを手掛かりにすればよい。しかし2章でも述べたように、互いに似通っており区別が付きにくい。特に泳動中は動きボケが生じ、ますます区別がつかなくなる。ある個体は隣接フレームで近くにいる（すなわち1フレームでの移動量は大きくない）という基本的仮定は極めて重要であるが、交差や接

触を起こすので、それだけでは不十分である。

他に考え得るのは、静止時のアピランスである。メダカは時々泳動をやめているので、その際には動きボケがなく、体の模様が比較的良好に見える。これを手掛かりにする。すなわち、「時刻  $t$  であそこに静止していた個体は、（途中でどう移動したかはわからないが）時刻  $t' (> t)$  においてここに静止している個体と似ている」という手掛かりである。動画全体を通して、こうした手掛かりが幾つか見つければ、それらを用いて全個体の経路を決定すればよい。

このメダカ追跡問題は、動画全体での大域的最適化問題として捉えるべきである。なぜなら、隣接フレーム毎に  $K$  対  $K$  の対応付けを繰り返す方法では、上記の手掛かりを利用できないためである。さらに、1 個体ずつ順に経路を確定していく方法も、後の個体になるほど不自然な経路を選ばざるを得なくなる。やはり動画全体の情報を用いて、時刻 1 から  $T$  までの全  $K$  経路の一括最適化すなわち大域的最適化が望ましい。

結果的にこの問題は、非マルコフ的多物体同時追跡とも言える、特殊な（そして数理的に極めて興味深い）問題になる。マルコフ的とは隣接フレーム間のみを見ておけばよい状況であり、これに対し非マルコフとは上述の手掛かりのように時間的に離れたフレームの関係をも見る必要がある状況である。隣接フレームでは近くにいるという基本的仮定だけならば、マルコフ的な問題である。

マルコフ的で済むならば、文献[6]にあるように、ネットワークフロー最適化問題の一種に帰着し、多項式時間で（すなわち膨大な計算量を要せずに）大域的最適解が求まる。直感的に言えば、Fig.2のネットワークの各エッジを水1リットルまで流せるパイプ考え、左から水3リットルを流して右から排出させるのである。ここで各エッジにはコストが付与される。大きな移動を意味するようなエッジには大きなコスト、小さな移動なら小さなコストを付与し、そこに水が流れればそれだけのコストがかかるとする。全体として最もコストが小さくなるように3リットルの水の流れ（フロー）を求めると、その流れが大域的最適解となる<sup>3</sup>。文献[7]ではネットワークフローに基づいた細胞群トラッキング法が提案されている。

一方、手掛かりを利用するとなれば非マルコフ的になり、文献[6]の最適化をそのまま利用できなくなる。しかしながら、多少複雑にはなるものの、同文献を発展させた枠組み（二段階の最適化）で準最適解を求めることは可能である。詳細は本記事の範囲を超えるので割愛する。興味ある読者は文献[8]を参照されたい。

Fig.3は、この手法を用いた結果例である。メダカの代わりに4個の白いボールが盤上を転がる。うち2個にはそれぞれ緑と赤の小さなシールが貼られている。時々それが見えることで、そのボールがそこに存在することがわかり、非マルコフ的な手掛かりとなる。中央付近でボールは時々隠れ（オクルージョン）を起こすので、こうした手掛かりを利用できなければ追跡は全く容易ではない。一方、それが適切に利用できると、シールが見えていない部分も含め、全てのボールについて安定した追跡が実現されている。

## 5. 機械学習とバイオイメージインフォマティクス

### 5.1 機械学習とは

機械学習を最も一般に定義すれば、関数  $y=f(x)$  を複数の事例から推定する問題となるだろう。ここで事例とは一般に、入力  $x_n$  と、それに対する理想的な出力値  $y_n$  の組



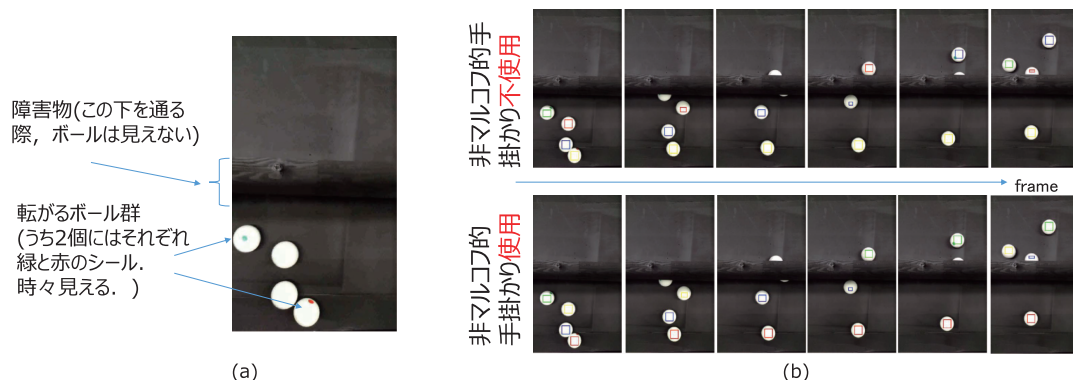


Fig.3 非マルコフ的多物体同時追跡の実験結果例。(a) 実験条件。(b) 結果例。

$(x_n, y_n)$  ( $n=1, \dots, N$ ) である。こうした事例は学習データとも呼ばれる。推定すべき関数  $f$  は、少なくとも  $y_n \sim f(x_n)$  を満たす必要がある。そして、学習パターン  $(x_n, y_n)$  だけでなく、未知の  $x$  についても、対応すべき  $y$  を正しく予測できる必要がある。未知データに対して正しい結果  $y$  を与える能力は、汎化能力と呼ばれる。学習データについては、すでに正解を知っているの、この未知データに対する汎化能力こそが機械学習結果の価値となる。

機械学習の典型例は回帰問題やクラス分類問題である。回帰問題は、説明変数  $x \in \mathbb{R}^d$  と目的変数  $y \in \mathbb{R}$  の組  $(x_n, y_n)$  が与えられたとき、それらの関係を最適記述する関数  $f$  を求める問題である。クラス分類問題は入力パターン  $x \in \mathbb{R}^d$  についてそのクラス  $y \in \mathcal{C}$  ( $\mathcal{C}$  はクラスラベルの集合) を与える関数  $f$  を求める問題である。他にも、異常検出は入力パターン  $x \in \mathbb{R}^d$  についてその異常度  $y \in \mathbb{R}$  を与える関数  $f$  を求める問題である。さらに、ノイズ除去のための最適フィルタリングでは、原画像  $x \in \mathbb{R}^{M \times N}$  について処理後画像  $y \in \mathbb{R}^{M \times N}$  を与えるフィルタ  $f$  を求める問題になる。

機械学習では  $x$  と  $y$  の間に潜む関係性  $y=f(x)$  を推定するために、前節で述べた数理最適化が利用される。具体的には、少なくとも学習データ  $(x_n, y_n)$  については  $y_n \sim f(x_n)$  となるべきなので、例えば二乗誤差  $\sum \|y_n - f(x_n)\|^2$  を最小にするように  $f$  を定める最適化問題となる<sup>4</sup>。ただし汎化能力を持つ必要がある。すなわち未知のデータ  $x$  についても正解  $y$  を与えるように関数  $f$  を最適化しなくてはならない。この一見無理な要求に応えるためには、 $f$  に関する何らかの事前知識が有効である。例えば変数  $x$  と  $y$  の間に比例関係がありそうならば関数を  $f(x)=ax+b$  のように“モデル化”してやればよい。そして  $\sum \|y_n - f(x_n)\|^2$  を最小化する  $a$  と  $b$  を求めればよい。

関数  $f$  に汎化能力を持たせるためには、上記のモデル化に加え、大量の学習データの利用も重要である。要するに、事例が多ければ多いほど「全くの未知データ」に出会う可能性が低くなるわけである。一方で大量の学習データの準備にはコストがかかることが多い。例えば分類問題ならば、まず  $\{x_n\}$  を十分に集め、次にそれぞれについて、目視など人間による観察により正解ラベル  $y_n$  を与える必要がある。「顔」や「机」などを対象とした一般的な画像認識課題では、いわゆるクラウドソーシング (crowd sourcing) [10] が良く利用される。対象画像を提示する web インタフェースを準備し、任意の人々にその画像が机なのか顔なのかを回答させることで、大量データに対する正解ラベル収集を行っている。その際、回答に対して適切なインセンティブを与えることで、回答精度を高めることができる。

## 5.2 深層学習の隆盛

昨今、深層学習が大きな話題を呼んでいる。その実体は Deep Neural Networks (DNN) と呼ばれる機械学習技術である。DNNの中で画像解析に頻繁に利用されるものが (Deep) Convolutional Neural Networks (CNN) である。DNNやCNNの歴史は古い。特にCNNの元祖は1979年に福島によって発表されたネオコグニトロン[11]であることが知られている。その後の改良は多々あるものの40年近く前の方法のリバイバルには次の2点が大きい。第一は、GPUなど安価で高性能な計算機リソースの普及である。第二の、そしてより本質的な理由は、大量の学習データが入手可能になった点である。DNNやCNNの詳細に関する記述は割愛する。興味ある読者は、充実した解説書[12]や本学会誌の記事[13]を参照されたい。また分類問題におけるCNNの利用については、筆者による平易な解説記事[14]もある。

DNNやCNNも関数  $y=f(x)$  であることに変わりはないが、非常にパラメータが多い大規模なモデルとなっている点で他手法と大きく異なっている。要するに、調節可能な“つまみ”が膨大についているので、それらを適切に設定してやれば、相当複雑な関数  $f$  でも実現できる。逆に言えば、これらのつまみを何とかして適切に設定する必要がある。従って、大量の学習データが必要になるのである。学習データはある意味で関数  $f$  の最適化の際に制約として効いてくるため、膨大な自由度を持ったDNN/CNNに意味のある最適解を求めるためには、自由度に応じた数の学習データが必要になる。

学習データが大量にある場合について、深層学習がクラス分類問題において高い精度を発揮することの実例を紹介したい。文献[15]では、手書き数字画像 (1枚の画像に手書きされた数字パターンが1つ写っているもの) の分類 (認識) を深層学習で行った結果がレポートされている。全部で82万画像があり、すべてに目視で正解ラベル (“0” から “9” までの全10クラス) が付与されている。これらのうち90% (74万) 画像を学習パターンとしてCNN (LeNet と呼ばれる最もシンプルなCNN) を学習し、残り10%で学習後のCNNの分類性能を評価したところ、99.88%の精度を得た。同じ74万画像をそのまま用いて最近傍法<sup>5</sup>で分類したところ98.94%であった。誤分類率でみれば、1.06%が0.12%まで減少したわけで、CNNの利用によりおよそ10倍の精度向上があったことを意味する。Fig.4は、関連して行った英大文字画像 (26クラス) の分類結果例である。約7千種類の異なるフォントを準備し、90%で学習、10%で評価したところ、同図のような人間にも分類困難な文字でも読めている。以上のように1クラスあたり数千~数万以上の学習パターンが準備できれば、CNNは極めて高い分類精度を示す。



Fig.4 CNNで正しく分類できた英大文字画像の例. 左上から右下へ“A”から“Z”まで各文字種4画像例を提示.

データが少ない場合は、CNNを用いても高い精度は出ない. 文献[15]では学習データ数を変えて精度評価を行っている. そこでは、学習データを74万画像の1/10, 1/100, 1/1000にした場合、上記数字画像の分類精度は99.63, 98.36, 93.18%と落ちていくことが明らかになっている. もちろん、こうした劣化はCNN以外の分類手法でも起こる. ここで言いたいのは、CNNの優位性を享受したいならば、大規模な学習データが不可欠という事実である.

### 5.3 バイオイメージインフォマティクスにおける機械学習応用

先述の通り機械学習も数理最適化に基礎をおいているために、バイオイメージインフォマティクスに適した頑健性が期待される. 実際、バイオイメージインフォマティクスにおける分類問題における特徴量や識別器の決定には、以前から機械学習が用いられてきた (例えば文献[16-19]など). 最近では、前節で述べた深層学習の導入も進みつつある[20, 21]. 今後も、分類問題だけでなく、セグメンテーションやトラッキング、位置合わせ、ノイズ除去、超解像など、様々な分野での利用が進むであろう.

ただし、バイオイメージインフォマティクス課題に機械学習を利用するにあたり、幾つかのハードルがある. 第一に学習データが少ない場合が多い点である. 特に独特の画像を扱っている課題においては、利用可能な正解ラベルが付いた大規模なデータセットの存在はあまり期待できない. この場合、大規模な学習データが前提となる深層学習は利用できないことになる. クラウドソーシング[10]も、エキスパートの眼が必要なバイオイメージインフォマティクス課題において利用するのは困難であろう. 転移学習 (本来の課題と類似したデータセットを用いて学習しておき、その後実際の課題の小規模データセットを用いて追加学習する技術) の利用や、データ拡張 (少数のラベル付きデータを意図的に変形させて疑似的にデータを増やす技術) などの工夫が必要となるだろう. 分類問題ならば、少数の学習データでも比較的良好に機能するSupport Vector Machine (SVM) の利用が考えられる[14].

第二は、真の正解ラベルを付与しづらいケースが多い点である. 例えば、領域分割をしようとした際、バイオ画像では領域が人間にとっても曖昧な場合があり、境界周辺に

ついて正確にラベリングされた学習データの準備は不可能であろう. 対象検出の場合も、それが本当に目的としている対象なのかどうか、曖昧な場合も多い. これらの場合、正解ラベルに揺らぎが生じやすく、期待した性能が達成されない場合がある. 一方、「遺伝子型がわかっている個体群についてそれらを表現型で区別できるか」という課題のように、正解ラベルが明確にわかっている場合もある. また、上記の曖昧な場合であっても、境界付近は曖昧でも、境界から遠い部分は明確に区別できる場合もあるので、それら明確な部分からのみ学習データを準備するという方法もあるだろう.

第三は、問題がアンバランスになりやすいという点である. 例えば2クラス分類問題では、正常データに比べて異常データが極端に少ない場合がある. 正常・異常が同程度の状況と対照して、この状況をアンバランスと呼ぶ. この極端にアンバランスな分類問題では、常に正常と分類したほうが、分類成功率が高くなる可能性がある. すなわち、一切何も学習せずに単に「正常」と言い続ける分類器のほうが、努力して「正常・異常」を分離しようとする分類器よりも性能が高いと評価されてしまうリスクがある. 領域分割や対象検出においても、ターゲットとしている領域が背景部分に比べて非常に小さい場合がある. この場合も、すべてを背景と回答するほうが、トータルの性能としては高くなる場合もあり、注意が必要である.

## 6. まとめ

本稿では、画像情報学と生物学との協働研究分野であるバイオイメージインフォマティクスについて概観した. この比較的新しい学際分野で扱われる画像は、肉眼観察できないミクロな世界をイメージング技術により拡大し可視化したようなものも多い. その結果、計算機による画像処理や解析が困難である. 解析のロバスト性を向上させるために、最適化や機械学習が用いられている. 最適化は、考えられる画像処理結果の中で最も合目的なものを自動で探してくれる. 機械学習は事前に準備しておいた事例を効率的に利用した画像解析を実現できる.

画像情報学は、最適化や機械学習の進展とともにこれからも進展していくであろう. そしてその進展は、ダイレク

トにバイオイメージインフォマティクスを介して、生物学に恩恵を施すであろう。その一方で、生物学の進展により新たな解析課題が発生し、その結果、バイオイメージインフォマティクスを介して画像情報学に課題が提示されるかもしれない。このように、バイオイメージインフォマティクスは、生物学と画像情報学の両輪を結ぶ車軸として、両分野の win-win な関係を構築しながら、その重要性を益々高めていくことになるであろう。

## 註

- 1 Idtracker というメダカのアピアランスを使ったトラッキング手法も提案されている[1]。原理的にはメダカ体表のテクスチャを利用し、個体を区別する。筆者らも利用してみたが、撮影条件に対する制約がかなり厳しいのか、期待する精度が出なかった。特にメダカは高速に泳動するため、ぼけてしまい、テクスチャをシャープに捉えられない点も問題になった。
- 2 「どのような人間がベストか」について絶対的な基準が存在しないことと同様である。
- 3 最適解において、各パイプに流れる流量は自動的に整数（この場合 1 リットル）になることが保証されている。従って流量 0.5 といった中途半端な状況はないので、最終的には 3 本の経路だけを確定できる。「自動的」というのは、問題を解く際に流量の整数性を仮定しなくても、結果的に整数（0 もしくは 1）になるという意味である。整数性が必要になると計算量が爆発的に増加するので、この性質は極めて好都合である。興味ある読者は完全単模性（total unimodularity）について参照されたい。文献[6]でも触れられている。
- 4 このように最小二乗法も、機械学習の立派な一手法なのである[9]。
- 5 最近傍法とは、分類対象と全ての学習データを比べ、最も類似しているもののクラスを分類結果とする方法である[14]。本実験の場合、1 画像の分類のために 74 万画像と比べる必要があるため、相当の計算量が必要となる。

## 参考文献

- [1] Pérez-Escudero A, Vicente-Page J, Hinz RC, et al.: idTracker: Tracking individuals in a group by automatic identification of unmarked animals, *Nature Methods*, 743-748, 2014.
- [2] 藤崎顕彰, フォンヤオカイ, 内田誠一他: ネットワークフロー最適化手法に基づく細胞内粒子群の追跡, 電子情報通信学会技術研究報告, PRMU2013-64, 2013.
- [3] 金森敬文, 鈴木大慈, 竹内一郎他: 機械学習のための連続最適化, 講談社, 東京, 2016.
- [4] 石川博: グラフカット, 情報処理学会研究報告 2007-CVIM-158 - (26), 2007.
- [5] Bise R, Zheng Y, Sato I, et al.: Vascular registration in

- photoacoustic imaging by low-rank alignment via foreground, background and complement decomposition, *Int. Conf. Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, 326-334, 2016.
- [6] Berclaz J, Fleuret F, Türetken E, et al.: Multiple object tracking using k-shortest paths optimization, *IEEE Trans., PAMI*-33(9), 1806-1819, 2011.
- [7] Bise R, Yin Z, and Kanade T: Reliable cell tracking by global data association, *Proc. IEEE Int. Symp. Biomed. Imag.*, 1004-1010, 2011.
- [8] 徳永誠, 内田誠一: 任意の非マルコフ的制約下での多物体追跡, 電子情報通信学会技術研究報告, PRMU 2016-107, 2016.
- [9] 杉山将: イラストで学ぶ機械学習 — 最小二乗法による識別モデル学習を中心に, 講談社, 東京, 2013.
- [10] 鹿島久嗣, 小山聡, 馬場雪乃: ヒューマンコンピュータとクラウドソーシング, 講談社, 東京, 2016.
- [11] 福島邦彦: 位置ずれに影響されないパターン認識機構の神経回路のモデル — ネオコグニトロン —, 電子通信学会論文誌 J62-A(10), 658-665, 1979.
- [12] 岡谷貴之: 深層学習, 講談社, 東京, 2015.
- [13] 庄野逸, 鈴木聡志, 木戸尚治: ディープラーニングの医用画像への応用, *医用画像情報学会雑誌*, 33(4), 75-80, 2016.
- [14] 内田誠一: パターン認識入門 ~ 人工知能の基盤技術, プラズマ・核融合学会誌, 92(11), 839-849, 2016.
- [15] Uchida S, Ide S, Iwana B, et al.: A further step to perfect accuracy by training CNN with larger data, *Proc. 15th Int. Conf. Frontiers in Handwriting Recognition (ICFHR)*, 405-410, 2016.
- [16] Orlov N, Shamir L, Macura T, et al.: WND-CHARM: multi-purpose image classification using compound image transforms, *Pattern Recognition Letters*, 29, 1684-1693, 2008.
- [17] Shamir L, Delaney JD, Orlov N, et al.: Pattern recognition software and techniques for biological image analysis, *PLoS Computational Biology*, 6(11), 1-10, 2010.
- [18] Sommer C and Gerlich DW: Machine learning in cell biology – Teaching computers to recognize phenotypes, *J. Cell Science*, 126, 5529-5539, 2013.
- [19] Dao D, Fraser AN, Hung J, et al.: CellProfiler Analyst: Interactive data exploration, analysis and classification of large biological image sets, *Bioinformatics*, 32(20), 3210-3212, 2016.
- [20] Angermueller C, Pärnamaa T, Parts L: Deep learning for computational biology, *Molecular Systems Biology* 12, 878, 2016.
- [21] Teikari P, Santos M, Poon C, et al.: Deep learning convolutional networks for multiphoton microscopy vasculature segmentation, *arXiv*, 2016.