

生物学×人工知能

寺澤 武

最近、人工知能 (AI) という言葉を至るところで聞くようになってきた。「人類はAIに支配される」や「AIが人間の仕事を奪う」などの、やや物騒な表現もされることがあるAIであるが、本稿では「やっぱり便利な道具であるAIのすすめ」という立場で、初学者の方を対象に「生物学に対してAIって何ができる?」という内容を紹介したい。

昨今のAIブームは2006年ころに登場した「深層学習 (ディープラーニング)」と呼ばれる技術が牽引役になっていると言われる。人工のニューラルネットワークを多層に及ぶ階層構造で構築したネットワークを使用して機械に学習させるので「深層学習」と呼ばれる。詳細な説明は割愛するが、この技術の利点は「特徴量を自動的に学習」できることである。特徴量とは、「特徴が数値化されたもの」であり、この特徴量をもとに機械は物体を認識したり属性を判別したりする。「イヌ」を例にあげると、土佐闘犬のような強固な顔つき、大柄でたれ耳の犬もいれば、ポメラニアンのような小柄で愛くるしい顔つきのフワフワしたような毛をもつ犬もいる。犬種分類を機械にさせる場合、従来の画像認識では、たとえば「身体の大きさ」や「毛の色」「毛の長さ」「口の口角の上がり具合」のような特徴量となりえるパラメータを検討し、同一の犬種でもさまざまな個体がいるため、それらの個体差に依らないパラメータ閾値を探索する。認識率が悪ければ特徴量を追加や削除し、再度パラメータ閾値を探索、という開発プロセスを繰り返す。ヒトには何となく分かるのに、それをコンピュータにやらせるのはとても難しいということはよくあり、画像認識アルゴリズムの開発作業はある種の職人芸のようなものである。これに対し、深層学習では「イヌの画像」と「犬種」という情報を与え機械に学習させることで、自動的に特徴量と閾値とが決定される。そのため専門のエンジニアでなくてもある程度のプ

ログラムの素養があれば高精度な画像認識が可能になった。さらに、プログラミングも不要でデータだけ与えることで学習機能を利用できるツールも登場¹⁾するなど今後よりいっそう、深層学習の活用が進むと予想される。

深層学習の「生物学」への応用先は、時系列データ解析など多岐にわたるが、ここでは先にあげた画像認識を例にあげ、その活用法について触れてみたい。生物の状態を観察し記録した画像を「良い状態」と「悪い状態」に分類したいとする。その場合、関連する一連の画像データを深層学習させると、新しく取得した画像が「良い状態」か「悪い状態」かを自動的に判断させることができる。これは画像認識の「分類」タスクと呼ばれる。たとえば、正常細胞 (良) だけの画像か、がん細胞 (悪) が含まれる画像か分類できるし、がん細胞の種類も情報として与えると、これらの種別を判別・分類することもできる²⁾。画像内の特定の種類のものを画像中から探す「物体検出」もでき、その個数を数えたりすることもできる。それらの画像中の位置が大事であれば、その情報も取得できる。さらに、画像中の各画素単位でその画素が何のクラスか分類したい場合は、「セマンティックセグメンテーション」を行えば良い。たとえば、図1はその例であるが、画素ごとに「細胞」か「背景」かを分類することで、図1右図に示すように細胞 (白抜き) 画素だけ判別できる。画像中に占める細胞の画素数を数えることで細胞の占める面積を定量化できる³⁾。さらに、たとえば同じ細胞種であっても、別の個体として画素ごとに区別したい場合には、「インスタンスセグメンテーション」と呼ばれる手法により可能となる。深層学習の欠点として、一般的に学習に用いる多量の画像データが必要なことに加え、それぞれの画像に正しい分類を表す正解ラベルを付け加える作業が必要であり、これらの準備が非常に骨の折れる点である。しかしながら一度学習させパフォーマンスが得られれば、画像の定量的かつ詳細な解析が行うことができ、生物学のさらなる発展に貢献するであろう。興味を持たれた方はぜひトライしていただきたい。

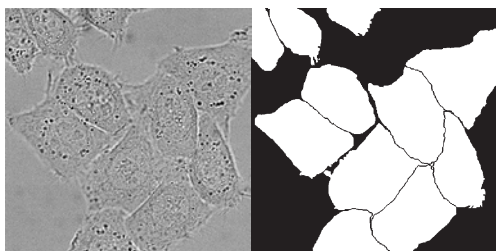


図1. 細胞画像のセマンティックセグメンテーションの例³⁾. 黒-背景, 白-細胞としてpix単位で分類.

- 1) Cloud AutoML: <https://cloud.google.com/automl?hl=ja> (2021/4/13).
- 2) Toratani, M. et al.: *Cancer Res.*, **78**, 6703 (2018).
- 3) Ronneberger, O. et al.: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015* (Navab, N. et al. Ed.) p. 234, Springer (2015).