

研究報告

注目対象の動きと状態を計測する汎用アルゴリズムの開発 ～定量科学研究に深層学習を用いる方法の検討報告～

新谷正嶺

中部大学生命健康科学部 生命医科学科

要 旨

本報告文は、深層学習がなぜ高い表現力を持つのかの説明をした後、畳み込み層 1 相のシンプルな深層学習アルゴリズムを用いて、深層学習はどのように基礎研究に応用できるのかの具体例を示す。その例とは、動画データの中で、注目したい観察対象の位置推定を行い、その推定位置の近傍の画像を用いて観察対象の状態分類の学習を行えば、高い精度の状態識別が行なえるというものである。具体的に、収縮ユニットであるサルコメアが振動のパターンを作り出している筋原線維に対して、注目するサルコメアの位置を捉えつつ、その注目しているサルコメアが収縮しているか伸展しているかを識別する解析と、手ぶれで揺れる動画の中で電線にとまっている鳥の挙動を識別する解析を、同じ解析アルゴリズムで行なった。どちらの解析でも、数十枚以下の画像による 2,30 回の学習(CPU ノート PC で数秒で完了)で識別率 100%の状態認識に至った。すなわち、位置推定などの精密解析を行う際に、その推定位置近傍の微小画像を状態識別に用いることで、手動だと大変でかつ再現性を疑われる、対象の状態に関する情報を容易に取得できることが分かった。

1. はじめに

現在、深層学習による画像認識の技術が注目を集めている。チェスや将棋、囲碁のトッププロに人工知能が勝利したり、画像集の分類問題で人工知能アルゴリズムが人よりも高い正答率を吐き出せるようになった(2015 年 2 月、人のエラー率 5%に対して 3.5%のエラー率)。これまで人間が行なわなければならないと考えられている多くの仕事を人工知能に任せられるようになってきている。なぜそのような技術革新が進むようになったのか？その仕組みの肝は、ニューラルネットワークと呼ばれる技術である[1]。ニューラルネットワークとは、複数の入力に重みをつけて合算し、その合計が閾値を越えたら高い値(1)を出力し、超えなかったら低い値(0)を出力する素子、ニューロンを多数繋げたネットワークである。このニューロン 1 つは、入力の合計が閾値以上か否かの判別しか出来ないが、その集合体は、非常に高い表現力を持つ。端的に言えば、コンピューターと同じ論理演算を行なえる。コンピューターは論理回路の集合体で出来ている。論理回路とは複数の場合もある 0 または 1 の入力に対し

て 1 つの 0 または 1 の出力を返す演算を行なう回路のことである。この論理回路の基本回路は入力のすべてが 1 のときは 1 を出力し、そうでなければ 0 を出力する AND 回路、入力のすべてが 0 のときは 0 を出力し、そうでなければ 1 を出力する OR 回路、1 つの入力に対して反対の出力をする NOT 回路の 3 つであるが、そのすべてを 1 つのニューロンで再現できる。簡単のために入力が 2 つの場合を考えると、ニューロン発火の閾値が 1 以上であれば、入力の 2 つともが 1 でないと 1 の出力をしない AND 回路になり、ニューロン発火の閾値が 1 以下であれば、2 つの入力のうち 1 つでも 1 であれば、1 の出力をする OR 回路になる。図 1 の左下図に示すように、1 つのニューロンは 1 本の判定直線と捉えることができる。この図 1 左下図の判定直線を傾きそのままに切片を大きくすると AND 回路になる。論理回路は基本回路を組み合わせることで様々な演算を実現する。例えば、基本回路に準じる回路として XOR 回路がある。入力が 2 つの場合では、入力が 1 と 0 であれば 1 の出力を行なうが、入力のどちらも 1 または 0 であれば、0 を出力する回路である。この

XOR 回路は 1 つのニューロンで表現することが出来ない。しかし、2 本の判定直線を用いれば表現できる(図 1 右下図)。もう少し丁寧に言うと、2 つのニューロンによる 2 本の判定直線と、その 2 本の判定の組み合わせに対して判定を行なう 1 つのニューロンによる 2 層ニューラルネットワーク(ニューロン 3 つ)で XOR 回路を表現できる。

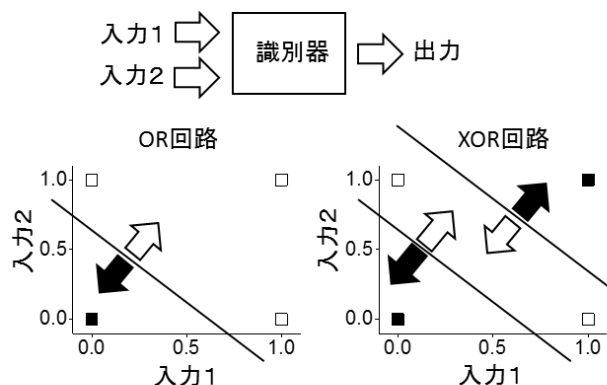


図1. OR回路とXOR回路に相当する識別器

この(多層の)ニューラルネットワークに対して、学習による各ニューロン(判定直線)のチューニングを行なうことで、様々なアルゴリズムを実現する技術、これがニューラルネットワークである。ネットワークの層を深くすることで表現力を高めること、学習によって適切なアルゴリズムを作っていくことが重要な特徴であるため、深層学習、あるいは深層ニューラルネットワークとも呼ばれている。

このニューラルネットワークという技術は今のところ、画像認識の分野で特に盛んに活用されている。前述の囲碁でトッププロに勝った人工知能(AlphaGo)は、Deep Q-Network と呼ばれるアルゴリズムを使用していたが、これは碁の盤面を画像として捉える強化学習アルゴリズムである[2]。画像分類は言わずもがな、画像認識の問題そのものである。

生命科学の分野においても画像データは重要な情報源である。医学系分野でも医療画像の診断に関する研究が盛んに進められている。皮膚がんか否かを識別させるためのデータセットの分類問題で、皮膚科医に匹敵するレベルの正答率を達成したアルゴリズムも報告されている[3]。しかし、これらの研究には多数の学習用データセットと計算リソースが必要になるため、研究の敷居が高く、また、学習の結果出来上がったアルゴリズムの解釈が困難であるため、基礎研究との相性が悪い。

生命科学の基礎研究に、ニューラルネットワークの技

術を活用することが出来ないか、具体的なアルゴリズムを模索し、作成してみた。その結果、それなりに汎用性のある、活用できそうなアルゴリズムを作成できたので、今回、その報告を行なう。具体的には、撮像領域を動き回る観察対象の動きと状態を解析できるアルゴリズムを作成した。対象の時々刻々の位置は、テンプレートマッチングと呼ばれる、テンプレート(型)と類似する領域を探索するアルゴリズムで推定した[4,5]。そして、その推定位置近傍の画像をニューラルネットワークで識別することで、注目対象の状態を判別した。同じアルゴリズムで、骨格筋の収縮ユニットであるサルコメアの位置と収縮・弛緩の状態変化、電線にとまる鳥の位置と振る舞いを識別することに成功した。

2. 方法

2.1 筋原線維の実験データの準備

本研究で用いた筋原線維のデータは、早稲田大学先進理工学部物理学科の石渡研究室にて、以前、筆者が取得したデータを用いた。この筋原線維は、同大学の生物実験管理委員会の承認を得た実験プロセスで調製した、ウサギの腸腰筋由来のものである[6]。筋原線維は、収縮力が発揮できる溶液環境において、溶液成分の濃度変化がなくても、自ら収縮と弛緩を繰り返す振動状態になる性質を備えている [6-9]。今回用いたデータは、ADP-SPOC (ADP Spontaneous Oscillatory Contraction) と呼ばれる条件で振動している筋原線維のデータを使用した[6,7]。

2.2 電線にとまる鳥の観察データの準備

2017 年 12 月、中部大学の正門近くで、デジタルカメラで撮影したデータを用いた。

2.3 アルゴリズム作成

本研究で使用したアルゴリズムは全て python 言語によって作成した。また、注目対象の位置推定アルゴリズムの作成にはオープンソースライブラリーである numpy, OpenCV を使い、注目対象近傍画像を用いた状態分類アルゴリズムの作成には、numpy, pandas, chainer を用いた[10,11]。

3. 結果&議論

3.1 自発振動する筋原線維のサルコメア位置と状態

筋原線維は、サルコメアと呼ばれる収縮ユニットが直列に連なった集合体である。図 2 に示すように、位相差像で筋原線維を観察すると、暗い領域と明るい領域が交互に現れる縞状の構造が観察される。暗い領域を A 帯、明るい領域を I 帯と呼ぶ。この I 帯の中央に細い暗線があり、これがサルコメアを仕切る Z 線と呼ばれる構造体である。サルコメアの収縮は A 帯の隙間に I 帯の一部が滑り込むことによって実現する。ゆえに、A 帯の構造は位相差像の光学顕微観察においては比較的不変である。よって、この A 帯の輝度パターンを型として、画像中の A 帯の位置のテンプレートマッチングを行なった(図 2)。このとき、型と類似する領域の推定に、前フレームの推定位置の近傍のみを用いることで、同一の A 帯の位置を追跡続けることに成功した(図 3)。この推定位置を中心に隣接する I 帯が含まれる程度の範囲の画像を、画像中をダイナミックに動くサルコメアの状態判別のための画像とした(図 3)。識別する状態は収縮状態か弛緩状態かの 2 択とした。ADP-SPOC 条件という、一定濃度の溶液条件下で振動している筋原線維では、急激なカルシウム濃度変化などの、個々のサルコメアの振動のタイミングを強制的に揃える因子が存在していない。ゆえに、同じ時間に収縮しているサルコメアも弛緩しているサルコメアも共在する。その為、画像全体を用いた識別でサルコメアの状態を推定することは出来ない。

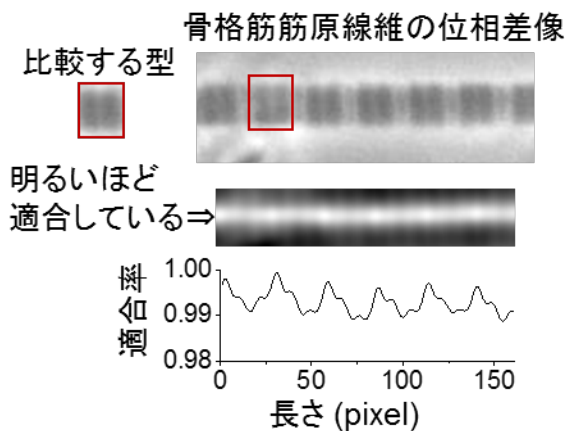


図2. A帯を型にした筋原線維のテンプレートマッチング

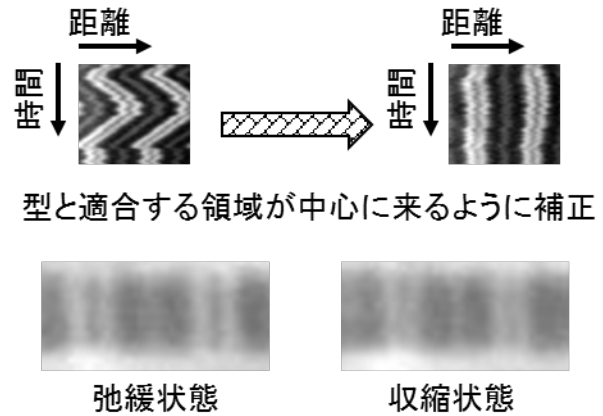


図3. テンプレートマッチングの結果に基づく判定用画像作成

状態識別のための学習器は、図 4 に示すネットワークを用いた。畳み込み層では 3×3 のフィルター4 種を用い、続くプーリング層では、 3×3 の最大値取得フィルターを用いた。活性化関数には ReLu 関数を用い、出力の手前の全結合層は 10 個のノード、そこから 2 個の出力に至るネットワークを作成した。

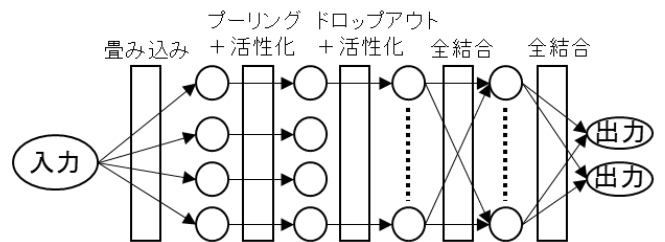


図4. 状態識別のための学習器ネットワーク

このネットワークは最大限にシンプルな畳み込みニューラルネットワークであるが、図 3 下方に示すような収縮状態と弛緩状態それぞれ 15 枚ずつのサルコメアの画像セットで半分を訓練データ、半分を学習後のテストデータとする学習を行なったところ、100%の状態識別率を誇った(図 5)。学習は 90 回繰り返したが、20 回学習を回した時点で訓練データ、テストデータ共に正答率は 100%に達した。なお、筆者のノート PC(CPU)で 90 回の学習は 5 秒で終了した。これほど短い時間で学習を終了できたのは、学習ネットワークがシンプルであることと、位置推定アルゴリズムを組み合わせ、学習する画像のサイズを小さく出来たことが理由であると考えられる。A 帯ベースの位置推定の結果と合わせると、サルコメアの位置と収縮・弛緩の状態を計測するアルゴリズムが出来上がった。

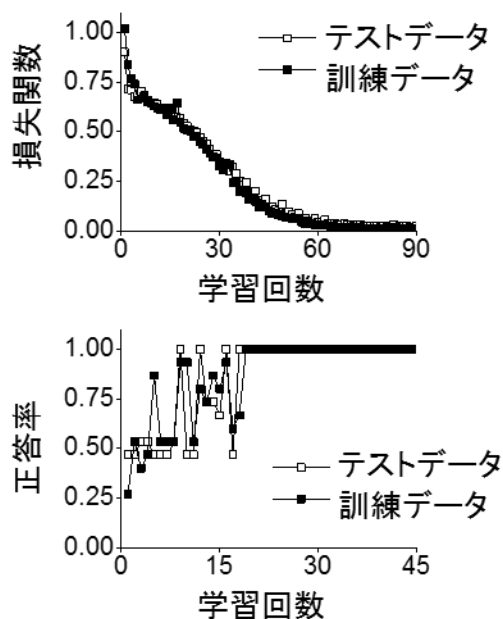
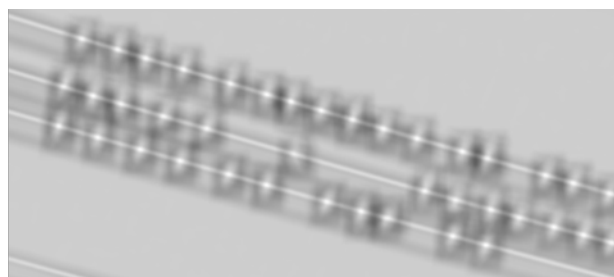


図5. サルコメアの状態識別の学習結果

3.2 電線にとまる鳥の状態識別

この型推定と深層学習を組み合わせた位置と状態の識別アルゴリズムは、推定に用いる型と状態識別の学習対象を変更すれば、別の実験系にも適応できることが期待される。そこで、昨年の12月にデジタルカメラで撮影した、電線に止まる鳥の動画に対して、このアルゴリズムを適応した。着目対象は電線にとまっている鳥とした。用いたデジタルカメラには手振れ補正の機能がないこともあり、動画中、電線にとまって動いていないはずの鳥が、カメラのゆれに伴って揺れていたが、テンプレートマッチングにより1羽の鳥の位置を追跡し、その検出位置が定位置になるよう画像の方をシフトさせることで、手振れのないカメラで撮影したような、ゆれを大きく低減させた動画像に変換することに成功した。テンプレートマッチングによる追跡可能な鳥たちは図6に示すように認識できる。この検出した鳥のしぐさに対して、状態判定を行なった。具体的には、鳥は羽づくろいを行なっているか否かを判定対象とした(図6)。判定画像はサルコメアの実験系と同様に追跡対象の近傍画像である。追跡した鳥の推定位置を中心とした局所画像90枚のセット(訓練・テスト用に半数分割)で学習したところ、学習回数30回ほどで、100%の正答率で識別することに成功した(図7)。



明るいほど適合している↑



比較する型



静止



羽づくろい

図6. 電線にとまる鳥の位置と状態

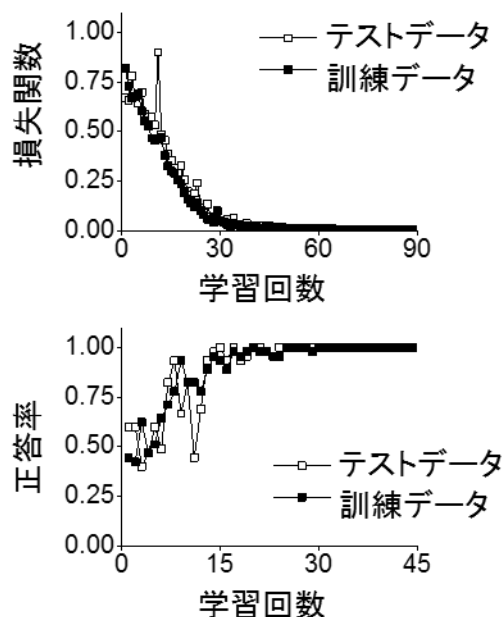


図7. 電線にとまる鳥の状態識別の学習結果

4. 結論

注目対象の推定位置近傍の局所画像を用いて識別判定を行なうことで、注目対象の位置と状態を計測できるアルゴリズムを開発できた。注目対象を中心に据えた小さい画像を識別判定およびそのための学習に用いることで、識別器のネットワークを簡素なものにしつつ高い識別率を持たせられた。判定画像の小ささとネットワークのシンプルさは判定に必要な処理時間を短くし、学習済みの識別器の判断則の分析(事後解釈)を容易にする。大きく異なる2種類の実験系に適応できたことから、この

アルゴリズムは汎用性が期待できる。例えば、明滅するホタルの位置を追跡しつつ、発光しているか否かの状態も判別するなどという処理も実現できるかもしれない。

5. 謝辞

本研究は特別研究員奨励費(S.A.S.,15J07373)および若手研究(B)(S.A.S., 17K15102)の助成を受けている。

引用文献

- [1] Johnson J, CS231n; Convolutional Neural Networks for Visual Recognition (<http://cs231n.github.io/>)
- [2] Mnih V et al., Human-level control through deep reinforcement learning. *Nature*. 2015 518, 529-33.
- [3] Andre E et al., Dermatologist-level classification of skin cancer with deep neural networks., *Nature*. 2017 542, 115-118.
- [4] Brunelli, R Template Matching Techniques in Computer Vision: Theory and Practice, Wiley, ISBN 978-0-470-51706-2, 2009 ([1] TM book)
- [5] Qingzong T et al., A new micropatterning method of soft substrates reveals that different tumorigenic signals can promote or reduce cell contraction levels. *Lab on Chip*. 2011 11, 2231-2240.
- [6] Shimamoto Y, Suzuki M, and Ishiwata S. Length-dependent activation and auto-oscillation in skeletal myofibrils at partial activation by Ca^{2+} . *Biochem Biophys Res Commun*. 2008 366, 233-238.
- [7] Ishiwata S, Shimamoto Y, Fukuda N. Contractile system of muscle as an auto-oscillator. *Prog Biophys Mol Biol*. 2011 105, 187-198.
- [8] Shintani SA et al., Sarcomere length nanometry in rat neonatal cardiomyocytes expressed with α -actinin-AcGFP in Z-discs. *J Gen Physiol*. 2014 143, 513-524.
- [9] Shintani SA, Oyama K, Fukuda N, Ishiwata S. High-frequency sarcomeric auto-oscillations induced by heating in living neonatal cardiomyocytes of the rat. *Biochem Biophys Res Commun*. 2015 457, 165-170.

[10] OpenCV (Open Source Computer Vision Library) (<https://opencv.org>)

[11] Tokui S et al., Chainer: a Next-Generation Open Source Framework for Deep Learning, NIPS, 2015.

Title: Development of a general-purpose algorithm to measure motion and state of attention object ~ Study report on method of using deep learning for quantitative science research ~

Author(s): Seine A. Shintani

Address(es): Department of Biomedical Sciences, College of Life and Health Sciences, Chubu University, 1200 Matsumoto-cho, Kasugai, Aichi 487-8501, Japan

Keywords: Deep learning, Object tracking, SPOC, Convolution neural network