BRAIN and NERVE 2019 年 7 月号特集「神経学と人工知能」

人工知能と脳科学の現在とこれから
Artificial Intelligence and Brain Science: the Present and the Future

銅谷賢治[1]・松尾豊[2]
Kenji Doya[1] & Yutaka Matsuo[2]

1: 沖縄科学技術大学院大学
Okinawa Institute of Science and Technology Graduate University
2: 東京大学
The University of Tokyo

人工知能と脳科学は、「知能の工学的実現は生物の脳のしくみにとらわれるべきではない」という考えと、「現存する高度な知能の実現例から学ぶべきだ」という考えの間で、接近と乖離を繰り返しながら互いに進化して来た。本稿では、まず今日の人工知能の到達点、脳科学と生命科学へのインパクトについて概観したのち、脳科学の進歩が次世代の人工知能にいかに貢献し得るかについて議論する。

Artificial intelligence and brain science have kept a swinging relationship with opposing views: “Artificial realization of intelligence should be free from biological constraints.” and “We should reverse-engineer the best existing implementation of intelligence.” In this article, we first review today’s achievements of artificial intelligence and its impacts on brain science and life science. We then discuss how the progresses in brain science can contribute to the future developments in artificial intelligence.

キーワード: 人工知能、脳科学、深層学習、強化学習、内部モデル
Keywords: artificial intelligence, brain science, deep learning, reinforcement learning, internal model

1. はじめに

「脳と人工知能の関係ってどうなってるんですか？」というのはよく聞かれる質問である。人工知能の研究開発では、「電子回路で知能を実現するために、生物の脳のしくみになどにとらわれる必要はない」という考えは根強くある。一方で、「人間の脳のような高度な知能の実現例があるのだから、それをリバースエンジニアリングしない手はない」という考えも当然ある。歴史的に見ても、人工知能と脳科学は接近と乖離を繰り返しながら互いに進化して来た。

前世紀までの人工知能は、エキスパートの知識や技能を抽出してプログラム化するものであり、知識表現やその探索が主要な技術であり、必ずしも脳の構造や動作に習う必要はなかった。一方今世紀インターネットの時代になると、大量のデータからの機械学習が人工知能の主要な技術となり、そこでは「深層学習(deep learning)」[1 2 3 4]と呼ばれる、脳の回路構造にならった手法が他を遥かに凌駕する高い性能を示したことから、脳型情報処理への関心が再び高まって来ている。さらに脳科学の側でも、イメージング、シーケンシングなどの実験

計測技術の進歩により得られるようになった膨大なデータから、科学的、医学的な知見を得る上で、機械学習を中心とした人工知能技術の活用は必須のものとなってきた。

本稿では、まず今日の人工知能の到達点、脳科学と生命科学へのインパクトについて概観したのち、脳科学の進歩が次世代の人工知能にいかに貢献し得るかについて議論する。

2. 人工知能の到達点

2.1 深層学習の進展

人工知能は、1956 年のダートマス会議 [1] [2] からスタートして以来、60 年以上の歴史をもつ分野である。近年特に注目されている理由は、複数の要因に分けられる。大別すると、i) 社会における IT 技術の重要性が改めて認識されてきたこと、ii) スマホやウェブ、センサ（あるいは IoT）などの技術により大量のデータを活用できる基盤が整ってきたこと、iii) 深層学習の技術が急速に進展していること、に分けられる。このうち i)や ii)は、技術的な要因というよりは、人々の認知が変わってきたこと、およびデータを取得・処理する環境が整ってきたという社会的な要因による変化である。一方で、iii)は純粋に技術的な進展であり、長い間重要と思われつつも実現することができなかった「深い」構造を持ったモデルを使って学習をすることが可能になってきた。

深層学習は 2012 年の NIPS 会議で、Hinton らのグループによって画像認識での圧倒的な性能が示されて以来 [5]、画像認識、音声認識、言語処理などのさまざまな領域において、目を見張る成果を挙げ続けている。画像認識では、2015 年ごろから特定のデータセットにおいて人間の水準を超え始め [6]、産業的な活用が一気に進んでいる。その一例が、医療における画像診断であり、レントゲンや CT、MRI などの画像からの診断、眼底検査、内視鏡による検査、皮膚病の診断、病理診断など、さまざまな活用が世界中で行われている [7]。

また、顔認証も急速に活用が進む技術のひとつである。羽田空港や成田空港での出国審査で顔認証のシステムが活用され始めている。これは、1 対 1 認証と呼ばれる比較的容易な技術である。一方で、従業員の入退室管理を顔認証で行う、決済を顔認証で行うという場合には、1 対 n 認証と呼ばれ、より難しい技術になる。こうした技術は、深層学習によって一気に実用段階に進み [8]、中国のベンチャー企業である SenseTime や Megvii（Fact++という技術を開発している）を始め、世界中で活用が進んでいる。

一方、自然言語処理の分野でも 2018 年には大きな性能の向上が相次いだ。従来は Word2Vec という1層のニューラルネットワークにより、ある語の周辺の語を推測するという問題を解くことで、ひとつの語をベクトルに置き換えるという手法が有名であったが [9]、ELMo と呼ばれる手法では、回帰結合神経回路(RNN)の一種である LSTM [10] を、双方向に接続したものを多層に積重ねて連結する手法が提案され、大きな性能の向上があった [11]。さらに、BERT という手法では、RNN ではなく、transformer という手法を双方向に多層に用いることで、入力文の含意判定や質問応答、固有表現抽出などさまざまなタスクにおいて、それまでの最高精度を大きく上回った [12]。

transformer[13] は、最近大きく注目されている技術であり、複数の自己アテンションをもった仕組みを用い、辞書を引くように key と value をセットにして、query に近い key の value を取ってくるような働きを行うことができる。自己アテンションは、ニューラルネットワークの隠れ層の情報自身が、どこの情報を用いるかを指定することになるため、中でどのような処理が行われているのかを理解することは非常に難解であるが、RNN よりも長距離の依存関係を取り出すことができる。Hinton らの提案したカプセルネットワーク [14] と原理的に近い。

ほかにも深層学習の活用としてアルファ碁が有名であり [15] [16]、また、さまざまなゲーム環境で動作を学習する技術が深層強化学習の研究として行われている [17] [18] [19]。

2.2 未だ到達していない点

こうした深層学習の進展は華々しいものではあるが、現段階の深層学習が達成していることは、人間の知能の多面的で複合的な働きからみると限定的である。深層学習は、これまでの人工知能の進展のボトルネックとなっていたパターン認識の部分で大きなブレークスルーをもたらしたものではあるが、そのブレークスルーの先に、これまでの人工知能領域でのさまざまな研究との融合が控えていることを理解すべきである。

まず第一に、いまの深層学習は画像認識分野で大きな飛躍を見せているものの、環境の状態の潜在構造を見つけ出すものにはなっていない。例えば、深層強化学習では、最初に状態空間が定義され、その空間での状態遷移モデルを使うモデルベース、モデルは使わず状態遷移の系列から直接学習するモデルフリー、それぞれの手法が提案されている。しかし、そもそものこの状態空間をどう定義するかという問題は、「状態表現学習」と呼ばれるが、未だに有効な解法がない難問である。

このような問題意識のもと、2018 年には、David Ha らが world models という論文を発表した [20]。変分オートエンコーダ（VAE）[21] を用いて、2次元のゲームを行う際の状態空間の定義とその遷移をモデル化しようというものである。また、DeepMind の研究者らは、生成クエリーネットワークという手法を提案した [22]。これは3次元空間内の視点とそこからの情景の見え方の組みから空間の内部表現を獲得し、クエリーとなる任意の視点からの見え方を答えられるようにするものであり、これも変分オートエンコーダがベースになっている。こうした研究の先に、状態表現、およびその遷移をデータから獲得すること、それによって、環境をよりコンパクトな表現で表し、強化学習や模倣学習等の手法が有効に働くようになることが期待される。

また、前節では、言語処理において大きな進展が続いていることを紹介したが、これも見方を変えると「人間ではとても考えられないほどの大量の言語データを使って、やっと人間並みの精度がでている」状態であり、人間の精度を大きく超えている画像処理とは状況は根本的に異なる。この理由は、古くからの人工知能分野の言い方を借りれば、「シンボルグラウンディング」を解いていないということになる。

これをどう解決するのかのヒントは見え始めているように思う。言語的な表現をベースに、画像・映像を生成する深層生成モデルを使ってシーンを「想像」することである。こうした技術は、部分的ではあるが実現されつつある。DeepMind の Hassabis らは、知能において想像（imagination）が予測やプランニング、意味理解において基盤的な役割を果たすと指摘している [23]。深層生成モデルの研究として、下絵に色をつけるとか、ピカソ風の絵を描くといった分かりやすい例が多いが、現実世界をモデル化し、次に何が起こるかを想像することを可能にすること、生物で言えば「脳内シミュレーション」[24] [25] を行う能力こそが、知能において決定的に重要な機能と考えられる。

またその先には、言語の学習が現実世界のモデルの学習にどう役立っているのか、演繹的な推論などの枠組みを機械学習の観点からどのように解釈することができるのか、コミュニティを形成し知識を蓄積することは機械学習の言葉でどう記述できるのか、意識や意志などがどのようにモデル化され得るのかなど、さまざまな興味深い課題がある。こうした課題が少しずつ目の前に迫ってきていることは、人工知能の 60 年以上の研究の歴史の中でも、極めてエキサイティングな瞬間であろう。そして、人工知能の研究の進展が脳科学にもたらす知見も多いであろうし、逆に、脳科学からヒントをもらうことが必要になる部分もたくさん出てくる

のではないかと考えられる。

3. 人工知能の脳科学、生命科学へのインパクト

今日、脳科学を含む生命科学では、連続切片電子顕微鏡、2光子顕微鏡や超解像度顕微鏡、次世代シーケンサー、常時装着型生体センサーなどから得られる膨大なデータを、いかに活用して新たな科学的発見や医療につなげるかが問われている。そこでは深層学習を含む人工知能技術の活用が不可欠なものとなりつつある。以下2つの例を見てみよう。

3.1 コネクトミクス

コネクトミクスは、脳の全体または一部の神経回路の結合を網羅的に同定しようという試みである [26]。これは一昔前までは夢物語でしかなかったが、米国 BRAIN イニシアチブが脳の全ニューロンの活動を計測するというゴールを設定するなどがぜん現実味を帯びてきた。例えばミクロレベルのコネクトミクスでは、脳の 1 ミリ角程度のサンプルを、連続切片電子顕微鏡で撮像し3次元再構成することにより、その全て細胞の形態とそれらの間のシナプス結合を可視化、定量化する。そこでは画像処理の複数の問題を解決しなければならないが、最大のものは連続切片間のどの領域が同じ細胞に属しているかを塗り分ける問題である。Seung らはそこに、畳み込み神経回路モデル [27] を応用することを提案し、高精度、高スループットでの処理を実現した [28]。その学習には教師データを揃えることと、塗り分けた結果の確認作業が必要であるが、Seung らはインターネット経由でゲーム感覚でこの作業に参加することを広く呼びかけ、「クラウドソーシング」による脳研究の雛形を作った [29]。

生体画像からの組織抽出の問題はCTやMRIからの画像診断でも基本となる処理であり、近年多階層の生成モデルであるU-Net[30]など、高精度の手法が提案され実用化されつつある。

コネクトミクスではニューロンの隣接関係を知ることができるが、それだけでは各シナプスがどれだけの結合強度を持ち、回路がどういう動作をするかを理解することは難しい。近年、2光子顕微鏡とカルシウム感受性蛍光タンパクの改良が進み、脳の0.5ミリから数ミリ四方の領域内の全ニューロンの活動を網羅的に記録することが可能になってきた。そこではまず、画像内で個々のニューロンが占める画素を同定し、ニューロンごとの蛍光変化の時系列を得ることが課題になる。その解決に向けて、PCA, ICA, NMF などの統計的機械学習手法が適用され、実用的なツールが提供されている [31-33]。

さらに、計測された多数のニューロンの活動パターンから、その背後にあるシナプス結合を推定するという問題も重要である [34]。そこでは相関係数に基づく従来の「機能結合」だけでなく、グランジャー因果性やそれを非線形系に拡張した転移エントロピー法が用いられる。またさらに、神経回路の確率的生成モデルを仮定し、一定の事前知識や仮定のもとでベイズ推定の原理により結合パラメタを推定するという手法も多く用いられるようになった [34, 35]。

これらは脳／生体画像処理への統計的機械学習手法の応用のごく限られた例にすぎず、今後、脳／生命科学の実験研究者にとって、さまざまな機械学習手法の原理とそれらの仮定や問題点の理解は必須のものとなることが予想され、そのためのチュートリアルコースなども活況を呈している。

3.2 計算精神医学

うつ病、統合失調症、自閉症などの精神疾患が心臓病など他の病気に比べて困難な点のひとつは、レントゲンや血液検査などによる明確なバイオマーカーが存在せず、症状に関する問診が唯一の診断基準となっている点である。その限界を越えるべく近年隆盛しつつあるのが「計算精神医学」のアプローチである[36]。そのひとつの戦略は、ヒトの認知機能を強化学習やベイズ推定などの数理モデルで捉えた上で、そのアルゴリズムや回路の誤動作として精神疾患を理解し、診断と処方につなげようという計算論的なものである[37]。

また一方で、患者や対照群の MRI、遺伝子多型、メタボロームなどの大量のデータに統計的機械学習アルゴリズムを適用することにより、診断や治療効果の予測、疾患サブタイプの同定などを行おうというデータ駆動のアプローチも活発化している。筆者らのグループは、広島大学の精神科医のグループとの共同で、構造 MRI、機能 MRI、血液バイオマーカーを含む多次元のデータに Group Lasso という教師あり学習アルゴリズムを適応することによって、うつ病の診断を行う可能性を示した[38]。また、Multiple co-clustering という教師なし学習アリゴリズムにより、抗うつ薬の有効性に関わる疾患サブタイプを同定している[39]。

また、ATR、東大、京大、広島大などをつなぐ共同研究では、安静時 fMRI データにスパースロジスティック回帰というアルゴリズムを適用することにより、自閉症に特有の機能的結合を同定するとともに[40]、自閉症、統合失調症、うつ病に関連する機能的結合の類似性と相違が明らかになった[41]。

さまざまな精神疾患の患者を含む脳構造、脳活動、遺伝子多型などの大規模なデータベースを構築しデータマイニングを可能にしようという試みは世界的に進行しつつあり、今後そこに関心を持った人工知能研究者が参入することにより、精神医学に新たな展開がもたらされることが期待されている。

4. 脳科学／人間科学から次世代の人工知能へ

Hubel & Wiesel[42] により 1950 年代に発見された視覚系の構造と機能をもとにした深層学習[27] は、近年の人工知能の急速な発展の起源となるものであるが、これからの人工知能のさらなる進化に役立つような新たな脳科学の知見は何かあるのだろうか[23]？

4.1 エネルギー効率

深層学習の計算には GPU(画像処理ユニット)が多用されるが、一般に GPU は消費電力が高くその冷却のためにも電力がかかる。例えば AlphaGo[15] には GPU176 台が使われ、その消費電力は 250 キロワット程度と言われる。一方ヒトの脳の消費エネルギーは 20 ワット程度とされており、桁違いに少ない。この違いはどこから来るのだろうか。シリコン半導体と神経細胞という物質レベルの違い、2進数デジタル演算と電位やスパイクタイミングによるアナログ演算との違い、誤差逆伝播法と自己組織化という学習方式の違い、と様々なレベルでの違いがある。しかしシリコン素子でも局所的にはアナログ演算を行い通信にはスパイクを使ったり、それに伴う素子のばらつきやノイズにロバストな処理方式を探る試みが neuromorphic computing というキーワードのもと、盛んに進められつつある[2]。

4.2 データ効率

深層学習ではまた、大量の学習データが必要とされる。AlphaGo の成功の背後には、生身の人間では一生かかっても経験できないほど膨大な数の対戦経験がある。インターネットや

シミュレーターから膨大なデータを高速に取り込めることは人工知能の強みではあるが、実ロボットの制御や変化する環境での学習など、使えるデータ量が限られたケースでは、いかにデータ効率の高い学習を行うかは大きな課題である。

人間、特に子供たちが視覚や音声の認識、歩行や手指の運動制御、さらに言語の活用を日常経験の中からすみやかに学習していく能力は驚くべきものがある。そのしくみは発達心理学や認知科学で盛んに研究されつつ未だ解明の途上であるが、Lake らは以下の3つの要素を特に重要なものとして指摘している [43]:1) 世界の因果的なモデルを獲得し説明や理解を行うこと。2) 物理や心理の学習を助ける直感的な理論生得的に持つこと。3)経験の構造化と再構成(compositionality)とメタ学習(learning to learn)により新たな状況に対応すること。これらの要素を人工知能に取り込もうという試みはすでに進みつつあり、前述の world models[20]、生成クエリーネットワーク [22] などはその例である。

4.3 自律性

今日の人工知能システムは、物体認識、ゲーム、自動運転など、ある特定の目的を設計者が定め、その学習に必要なデータやシミュレーターを用意し、深層学習回路の構造とパラメタを設定し走らせることで作成される。これは自律的、自発的に学習し成長し進化していく生命体を担う脳との決定的な違いである。特定の入出力関係を与えられる教師あり学習に比べれば、目標を達成するための出力を探索的に学習する強化学習はより自律性が高いと言えるが、そこでも報酬関数は設計者により決められたものである。また報酬関数を設計する側でも、目標の達成とコストの削減、危険の回避などの要素をいかにバランス良く取り込むかが難点である。

人間や動物の報酬系は進化の歴史の中で形成されたものであるが、飲食や痛みなど自己保存に関わるもの、交配や養育など自己複製に関わるものがその根幹にある。さらに人間や高等動物の場合、好奇心、美の欲求、理解欲、達成欲など、生存や繁殖には必ずしもつながらず、時に矛盾するような内発的な報酬を持ち、その結果として文化、技術や知識の探索が進められ、今日の科学や産業、社会が形成され進化し続けている。

このような知の探索と活用のサイクルに人工知能が加わっていくためには、自分自身の目標ないし報酬を進化させる機能が必要であり、それをめざした研究も人工知能と人工生命との間で進められている [44, 45]。また、好奇心などの内発的報酬をいかに数理的に定義し実装するかという研究も盛んに展開している [46, 47]。

4.4 社会性

人工知能エージェントがそれ自体の目標を自律的に発見し追及することが可能になれば、人工知能により新たな科学や技術が発見され、新たなサービスや産業が創生されることも期待できる一方、その暴走への懸念も高まる。人工知能エージェントが人類を敵とすべき必然性は考えにくいが、その目標追求の副作用として人間に危害や不利益を与えたり、さらには悪意や敵意を持った人間に操作される形で脅威となる可能性は十分にあり得る。

今後進化する人工知能エージェントの暴走や悪用の危険性は正しく予測評価すべきものだが、実はすでに世にはびこる人間たちも相当に危険な生き物である。人間の利己主義や悪意、敵意の暴走を抑えるために、脳には痛みへの共感や不平等への忌避を促す報酬機構が生得的に備わっており [48, 49]、またその生み出す社会には伝統や宗教、ルールや制度が形成されてきた。例えば特定の個人やグループに無制限な権力の集中が進むと必ずや腐

敗や抑圧が起こるという歴史的教訓から、任期付きで改選される首長制、三権分立、地方自治、独占禁止法、労働争議権、情報公開制度など、今日の民主主義のルールが形成されてきた。

これら脳と社会の進化のなかで獲得されてきた知恵は、人工知能エージェントの暴走や悪用を防ぐ上でも貴重な手段となり得る。単一の人工知能システムプが社会の全てを制御するのではなく、多様なプログラムが相互監視し合いながら協調連携する仕組みを作っていくことも重要な課題である。

5. おわりに

以上、人工知能の今日、その脳科学への応用、さらに脳科学がこれからの人工知能研究に果たし得る役割について概観を試みた。筆者らは科研費新学術領域「人工知能と脳科学の対照と融合」(http://www.brain-ai.jp)を 2016 年度から推進しており、その主催／共催する合同ワークショップや国際シンポジウム等での議論がもととなっている。もちろん本稿では捉えきれていない重要な進展や可能性もあるであろうし、読者のみなさんからもぜひ議論を寄せていただけると有難い。

このような可能性と課題を踏まえた上で何より重要なのは、情報科学と脳科学の両方に知識と関心を持つ若手研究者の育成である。深層学習の分野での研究開発は世界的に急速に進んでおり、それにキャッチアップするだけでも大変であるが、そこでさらにオリジナルな理論や応用を提案するには、確かな知識技量とフレッシュな視点が強く求められる。筆者らの新学術領域でもそのための若手サマースクールの企画などを行なっているが、学部、大学院でのコースや研究機関などでの組織的な育成の必要性はますます高まっている。

脳と人工知能の境界領域の国際会議、NeurIPS や COSYNE などで、日本からの発表は残念ながら少なく、中国勢などの勢いに押され気味である。日本の研究者も世界的な研究ネットワークの中に積極的に身を投じ、連携し切磋していくことが求められる。その中で、神経回路の理論研究を先導してきた [27, 50] 日本のコミュニティの強みが生かせることを願っている。

文献

1. 岩澤有祐, 鈴木雅大, 中山浩太郎, 松尾豊監訳: 深層学習 (Goodfellow I, Bengio Y, Courville A 著: Deep Learning). 角川書店, 2018.
2. 銅谷賢治監訳: ディープラーニング革命 (Terrence J. Sejnowski 著: The Deep Learning Revolution). ニュートンプレス, 2019.
3. 松尾豊: 人工知能は人間を超えるか：ディープラーニングの先にあるもの. 角川書店, 2015.
4. 松尾豊, NHK「人間ってナンだ?超 AI 入門」制作班: 超 AI 入門—ディープラーニングはどこまで進化するのか. NHK 出版, 2019.
5. Krizhevsky, A, Sutskever, I, Hinton, GE: ImageNet classification with deep convolutional neural networks. Advances in Neural Information Processing Systems, 25, 1090–1098. 2012.
6. He, KZ, X.; Ren, S.; and Sun, J. : Deep residual learning for image recognition. arXiv, 1512.03385, 2015.
7. Litjens, G, et al.: A survey on deep learning in medical image analysis. Medical Image Analysis, 42:60-88, 2017.
8. Wang, M, Deng, W: Deep face recognition: A survey. arXiv, 1804.06655, 2018.

9. Mikolov, T, Sutskever, I, Chen, K, Corrado, GS, Dean, J: Distributed representations of words and phrases and their compositionality. NIPS 2013, 2013.
10. Hochreiter, S, Schmidhuber, J: Long short-term memory. Neural Comput, 9:1735-1780, 1997.
11. Peters, ME, et al.: Deep contextualized word representations. NAACL 2018, 2018.
12. Devlin, J, Chang, M-W, Lee, K, Toutanova, K: BERT: Pre-training of deep bidirectional transformers for language understanding. arXiv, 1810.04805, 2018.
13. Vaswani, A, et al.: Attention is all you need. NIPS 2017, 2017.
14. Sabour, S, Frosst, N, Hinton, GE: Dynamic routing between capsules. NIPS 2017, 2017.
15. Silver, D, et al.: Mastering the game of Go with deep neural networks and tree search. Nature, 529:484-489, 2016.
16. Silver, D, et al.: The predictron: End-to-end learning and planning. ICML, 1612.08810, 2017.
17. Silver, D, et al.: A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play. Science, 362:1140-1144, 2018.
18. Moravcik, M, et al.: DeepStack: Expert-level artificial intelligence in heads-up no-limit poker. Science, 356:508-513, 2017.
19. Vinyals, O, et al.: AlphaStar: Mastering the Real-Time Strategy Game StarCraft II. 2019.
20. Ha, D, Schmidhuber, J: World models. arXiv, 1803.10122, 2018.
21. Kingma, DP, Welling, M: Auto-encoding variational Bayes. International Conference on Learning Representations (ICLR), 2014.
22. Eslami, SMA, et al.: Neural scene representation and rendering. Science, 360:1204-1210, 2018.
23. Hassabis, D, Kumaran, D, Summerfield, C, Botvinick, M: Neuroscience-inspired artificial intelligence. Neuron, 95:245-258, 2017.
24. Fermin, AS, et al.: Model-based action planning involves cortico-cerebellar and basal ganglia networks. Sci Rep, 6:31378, 2016.
25. Funamizu, A, Kuhn, B, Doya, K: Neural substrate of dynamic Bayesian inference in the cerebral cortex. Nat Neurosci, 19:1682-1689, 2016.
26. Seung, S: Connectome: How the Brain's Wiring Makes Us Who We Are. Mariner Books, 2013.
27. Fukushima, K: Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. Biol Cybern, 36:193-202, 1980.
28. Helmstaedter, M, et al.: Connectomic reconstruction of the inner plexiform layer in the mouse retina. Nature, 500:168-174, 2013.
29. Arganda-Carreras, I, et al.: Crowdsourcing the creation of image segmentation algorithms for connectomics. Front Neuroanat, 9:142, 2015.
30. Ronneberger, O, Fischer, P, Brox, T: U-Net: convolutional networks for biomedical image segmentation (2015). arXiv preprint arXiv:1505.04597.
31. Maruyama, R, et al.: Detecting cells using non-negative matrix factorization on calcium imaging data. Neural Networks, 55:11-19, 2014.
32. Friedrich, J, Zhou, P, Paninski, L: Fast online deconvolution of calcium imaging data. PLoS Comput Biol, 13:e1005423, 2017.
33. Pachitariu, M, Stringer, C, Harris, KD: Robustness of Spike Deconvolution for Neuronal Calcium Imaging. J Neurosci, 38:7976-7985, 2018.
34. Magrans de Abril, I, Yoshimoto, J, Doya, K: Connectivity inference from neural recording data: Challenges, mathematical bases and research directions. Neural Netw, 102:120-137, 2018.
35. Pillow, JW, et al.: Spatio-temporal correlations and visual signalling in a

complete neuronal population. Nature, 454:995-999, 2008.
36. Montague, PR, Dolan, RJ, Friston, KJ, Dayan, P: Computational psychiatry. Trends Cogn Sci, 16:72-80, 2012.
37. 国里愛彦, 片平健太郎, 沖村宰, 山下祐一: 計算論的精神医学: 情報処理過程から読み解く精神障害. 勁草書房, 2019.
38. Shimizu, Y, et al.: Toward probabilistic diagnosis and understanding of depression based on functional mri data analysis with logistic group LASSO. PLoS One, 10:e0123524, 2015.
39. Tokuda, T, et al.: Identification of depression subtypes and relevant brain regions using a data-driven approach. Sci Rep, 8:14082, 2018.
40. Yahata, N, et al.: A small number of abnormal brain connections predicts adult autism spectrum disorder. Nat Commun, 7:11254, 2016.
41. Yahata, N, Kasai, K, Kawato, M: Computational neuroscience approach to biomarkers and treatments for mental disorders. Psychiatry Clin Neurosci, 71:215-237, 2017.
42. Hubel, DH, Wiesel, T. N.: Receptive fields of single neurones in the cat's striate cortex. Journal of Physiology, 148:574-591, 1959.
43. Lake, BM, Ullman, TD, Tenenbaum, JB, Gershman, SJ: Building machines that learn and think like people. Behav Brain Sci, 40:e253, 2017.
44. Elfwing, S, Uchibe, E, Doya, K, Christensen, HI: Darwinian embodied evolution of the learning ability for survival. Adaptive Behavior, 19:101-120, 2011.
45. Uchibe, E, Doya, K: Finding intrinsic rewards by embodied evolution and constrained reinforcement learning. Neural Networks, 21:1447-1455, 2008.
46. Jung, T, Polani, D, Stone, P: Empowerment for continuous agent--environment systems. Adaptive Behavior, 19:16-39, 2011.
47. Baldassarre, G, et al.: Intrinsic motivations and open-ended development in animals, humans, and robots: an overview. FiCS, 2014.
48. Haruno, M, Frith, CD: Activity in the amygdala elicited by unfair divisions predicts social value orientation. Nat Neurosci, 13:160-161, 2010.
49. Fermin, AS, et al.: Representation of economic preferences in the structure and function of the amygdala and prefrontal cortex. Sci Rep, 6:20982, 2016.
50. Amari, S: A thoery of adaptive pattern classifiers. IEEE Transactions on Electronic Computers, EC-16:299-307, 1967.

謝辞

本研究は JSPS 科研費 JP16H06561, JP16H06562, JP16H06563, JP16K21738 の助成を受けたものです。