

人工知能と生物知能——深層学習から神経科学への寄与

小口峰樹 (Mineki Oguchi)

玉川大学脳科学研究所

人工知能研究は、現在、コンピューターの大幅な性能向上とビッグデータの活用に裏打ちされ、深層学習の数々の成功例に見られるように隆盛期を迎えている。今日の機械学習によって実現される人工知能の基本原理は、もともと、生物の神経回路における情報処理についての洞察を通じて得られたものである。しかし、近年では、人工知能研究と神経科学は「別々の言語を喋っている」(Marblestone et al., 2016, *Front Comput Neurosci*) とされる状態であり、人工知能の黎明期における密な連携と比較すると、両者の間にはわずかな交流しか成立していない。神経科学は、神経活動に関する新たな計測技術や操作技術の開発に後押しされ、ここ 20~30 年、様々な生物学的機能に関してそれらを実現する神経回路の働きを闡明するような成果を生み出してきた。しかしながら、今日の人工知能の発展を主に指導してきたのは、そうした成果を通じて得られた知見ではなく、モデルのパラメータ設定における効率的な最適化を実現するための数学的な洞察である。

本発表では、こうした没交渉の時代を超えて、今日における人工知能と神経科学がいかにして協働しうるかを考察する。特に、深層学習をはじめとする人工知能の側から神経科学に対してどのような寄与が成立しうるかを中心に検討する。

近年の神経科学の研究では、機能的核磁気共鳴画像法 (fMRI) によって得られたボクセルデータや、多点電極を用いて得られた活動電位データ、硬膜下皮質表面電位 (ECoG) を用いて得られた局所場電位データなどの大規模データに対して、機械学習によるデコーディング技術を用い、そこにどのような情報が潜んでいるかを解析するという手法がしばしば利用されている。こうした研究においては、人工知能は神経科学をサポートする解析の道具としての寄与に留まっている。こうした水準に留まらず、人工知能の計算機構から生物の脳に関する機能解明に結びつくような洞察を得ることは可能だろうか。たとえば、DeepMind の開発した AlphaGo を考えよう。われわれは、AlphaGo の計算機構を考察することで、卓抜した囲碁の一手を可能とする脳の機能に迫ることはできるだろうか。

こうしたアプローチからの寄与を実現する上では以下の二つの問題が存在している。第一に、高度な深層学習においては、その中間層においてどのような内部表象が存在するのかがブラックボックス化されているという問題である。第二に、深層学習を通じてある課題が解決されたとしても、その解決を導いた計算機構が生物の脳が実装している計算機構とは異なる可能性があるという問題である。これらの問題は、人工知能研究を通じて生物の脳機能の解明を目指すことを困難にすると考えられる。では、これらの問題はどのように解決されうるのだろうか。

これらの問題は、少なくとも階層性の明確な感覚皮質での情報処理においては回避可能であると考えられる。たとえば、Yamins らは、視覚認識課題を用いて多層ニューラルネットワークモデルを人間のパフォーマンスにマッチするように学習し、その出力層

と中間層がそれぞれ腹側経路の最終段階と途中段階にあたる下側頭皮質と第4次視覚野の神経活動を予測できることを確認した (Yamins et al., 2014, *PNAS*)。こうした方向性においては、モデルの階層構造と感覚皮質の階層構造との対応性をより高いレベルで実現し、各階層に対応する脳部位からの神経活動データとのフィッティングを通じてモデルの取捨選択を行うことで、高い生物学的妥当性を有する人工知能モデルを得ることができる。そして、こうしたモデルについて、各層でどのようなコスト関数を用いて最適化問題が解決されているのかを見ることで、対応する脳部位での計算原理に関して洞察を得ることができると期待される。

しかし、このような解決は、生物の知能の実現に関わるより高次の脳領域での情報処理には適用が困難であると予想される。たとえば、推論や思考に関わる前頭前野は異なる複数の脳領域に区分可能であるが、感覚皮質のような明確な階層構造は存在しておらず、また、そもそも何を入力データとして用いるべきかも判然としない。

本発表では、以上で述べてきた論点を解説するとともに、生物知能を実現する高次脳機能の解明に人工知能がどのように寄与するかという最後に述べた課題について、現在主流であるパターン認識ではなく「モデル構築」を行う人工知能の実現を通じて解決を目指すというアプローチから検討を行う。