

I 医療における AI 研究開発の最新動向と将来展望

4. 基盤モデルの現状と将来展望

小田 昌宏 名古屋大学情報基盤センターデータサイエンス研究部門

「ChatGPT」などの文章生成 AI や、CT 画像を与えると所見文の案を生成するマルチモーダル AI が登場している。このような生成 AI を構築する上では、汎用的な判断能力を持つ基盤モデル (foundation model) の活用が必須である。現在、基盤モデルの研究開発が活発に進められており、医療用の基盤モデルが多数登場している。基盤モデルを活用することで、診断や治療支援だけでなく、医療関連業務の効率化に貢献する AI が今後充実すると見込まれる。本稿では、AI 開発の重要な要素である基盤モデルの解説、医療用基盤モデルの研究動向および今後の展望について述べる。

医療支援 AI の状況

高度化が進むコンピュータの判断能力を活用し、診断や治療支援を行う仕組みがこれまで数多く提案されている。特にコンピュータ支援診断 (computer-aided diagnosis : CAD) において応用が進み、CT 画像、X 線画像、内視鏡画像、眼底画像などを対象とした商用 CAD システムが多数登場している。これらの CAD システムの多くは、画像を入力すると判断結果を出力する画像識別 AI を使用していた。

近年の生成 AI 関連技術の発展に伴い、自然な文章を生成できる文章生成 AI、文章からの画像生成、または画像からの文章生成ができるマルチモーダル AI が登場している。文章生成 AI として、OpenAI 社の ChatGPT や Google 社の「Gemini」などがあり、文章を入力すると、自然な返答を文章で返すことができる。マルチモーダル AI では、医用画像から文章生成を行う AI が次々と開発されている。名古屋大学と国立情報学研究所 (NII) は、2つの CT 画像を比較・経時解析し、日本語の所見文を生成する AI を開発した¹⁾。また、東京大学が開発した Asagi は、さまざまな一般画像から日本語の説明文を生成することが可能である²⁾。海外では、マルチモーダル AI を使用した商用の医療支援サービスが登場している。フランスの Therapixel 社の「MammoScreen」は、マンモグラフィ画像から腫瘍の自動検出

および読影レポート文章案の自動生成が可能である。アメリカの RADPAIR 社は読影レポート文章案の自動生成サービスを提供している。

文章生成 AI やマルチモーダル AI といった新たな形の AI を開発するための重要な要素に、基盤モデルがある。文章生成 AI では言語基盤モデル (language foundation model) が使用され、その多くは大規模言語モデル (large language model : LLM) である。画像処理 AI では画像基盤モデル (vision foundation model) が使われる。そして、マルチモーダル AI では、画像基盤モデルと言語基盤モデルを組み合わせた vision language model (VLM) を使用するケースが多い。これら基盤モデルを使用することの利点や基盤モデルの例について、この後に述べる。

基盤モデル

従来型の AI 開発では、教師あり学習 (supervised learning) を使用して特定のタスク向けの深層学習モデルを構築していた。例えば、腫瘍有無のアノテーションが付与された X 線画像を用いて、X 線画像からの腫瘍判別を行う深層学習モデルを構築していた。この方法の問題点は以下の2点である。

① タスク特化型モデル：特定のタスク向けに構築された深層学習モデルは、ほかのタスクへの応用性が限定されている。例えば、特定の腫瘍と画像モダリティ用に構築されたモデルを、