



働き方改革から
2040年問題まで
医療の課題に
立ち向かう

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)を使用する場合が多い。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、