

IV 乳腺画像診断におけるAI技術の進歩

# 1. 乳腺領域における生成AIによるAI-CADの新たな展開への期待

藤田 広志 岐阜大学工学部

生成AI(ジェネレーティブAI: generative AI)により、いまや第四次AIブームがやってきたとも言われており、「生成AIは新たな産業革命であり、未来を拓く」(東京大学・松尾 豊)<sup>1)</sup>との新たな期待がある。医療AIもこのブームに乗り、乳腺画像診断支援の領域でも、またさらなる大きな変化が起きようとしている。本稿では、そのような最近の状況を概観する。

## AI-CADの高度化・進化

現状のAI-CADの「目的の多様化」「高度化・進化」、そして「利用形態の多様化」を図示すると、図1のようにまとめられる<sup>2)</sup>。乳腺領域における米国食品医薬品局(FDA)承認の医療AI製

品(SaMD)の総数は、2017年7月の乳房MRにおけるCADxを皮切りに、2023年5月のデジタル乳房トモシンセシス(digital breast tomosynthesis: DBT)におけるCADE/x(CADE + CADxの両機能あり)までに13品目であり、決して多くはない<sup>3)</sup>。また、内訳はCADE/xが7、CADxが2、CADtが4である。このように、学術論文は千のオーダーで出版されているが、商用化への道は決して容易ではないということである。

RSNA 2022でも教育講演などで指摘されていたが、商用化製品についても、多くは市場導入後の実臨床世界におけるプロスペクティブな臨床評価が必要とされる。臨床導入前のAI-CAD製品の評価の方法、あるいは導入後のAI-CAD製品のQA/QCのプロトコールの

検討なども、まだ課題として残る。AI単体での性能よりも、AI+医師でどれだけ性能が出せるかについても、まだ十分なデータが出そろっていない。スクリーニングにおけるプライマリ診断(正常症例と異常症例を選別する)としての役割についても、プロスペクティブな評価が必要である。RSNA 2022のGilbertのKeynote講演では、少なくとも3つの商用AIアルゴリズムで、50%の正常マンモグラムを1%以下のがん見逃しの精度でルールアウト(除外診断: rule out)できると報告しており、これは放射線科医以上の性能という。いずれにしても、過去の従来型CADの失敗例を繰り返さないように、今度こそしっかりした臨床評価を行う必要がある。

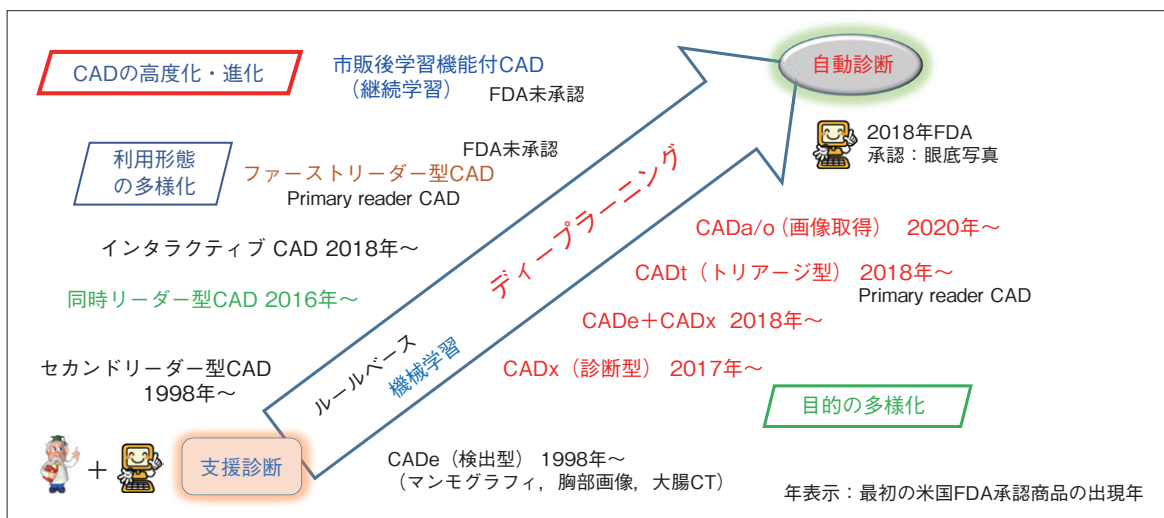


図1 AI-CADの「目的の多様化」「高度化・進化」、そして「利用形態の多様化」<sup>2)</sup>  
 ファーストリーダー型CADはルールアウトとも呼称される。生成AIの導入により、対話型AI-CAD, 説明型AI-CAD, 総合情報型AI-CADなどにさらなる高度化・進化が起きつつある。



## 生成AI

現状のCADで利用されているAI技術は、CNNによるディープラーニングモデルであり、「認識AI/識別AI (discriminative AI)」と区分化され、周知のように文字認識や画像認識で大成功が収められている。

これに対して、「生成AI」は、静止画、動画、オーディオ、文章、プログラミングコードを生成、もしくはこれらを組み合わせて生成するAIである。その技術は、データ/コンテンツから学習するディープラーニング (特にTransformerが特異的な役割を果たしている) より構築された大規模言語モデル (large language model: LLM) であり、いま話題のテキスト生成のOpenAI社の「ChatGPT」などがある。また、画像生成では、主にテキスト入力に基づく「Stable Diffusion」が代表的である。加えて、OpenAI社の言語モデルの最新版 (GPT-4) には画像認識機能が追加され、これからおもしろい展開が期待できそうである。

図1の各種CADも、例えば、ChatGPTとの融合により、対話型でこれまでよりずっと使い勝手の良いシステム (図2) に変貌を遂げていくと予想される<sup>7)~10)</sup>。また、モデルも、基盤モデル (foundation model) \*<sup>1, 11)</sup> でAIシステムの作り込みができるようになるであろうから、現状のAI-CADはさらに脱皮していくことになるであろう。

## 「乳がん予防とスクリーニング」に関するChatGPTへの質問

ChatGPTに対して、一般市民が健康上のアドバイスを求めるケースが急増しているという。これを受けて、この先進的な技術をどう適切に利用すべきか、専門家はその回答能力の検証を進めている。例えば、米国メリーランド大学医学部の研究チームでは、「乳がんの予防とスクリーニングに関して、ChatGPTが

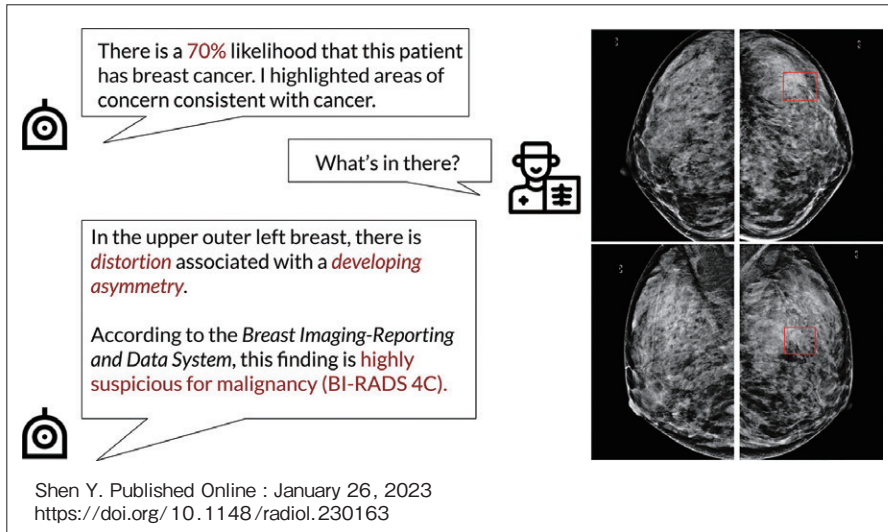


図2 スクリーニング用マンモグラムの診断に関して臨床医 (右) にその推論を説明する対話型人工知能 (AI) ベースのAI-CADシステム (左)

(Shen, Y., Heacock, L., Elias, J, et al. : ChatGPT and other large language models are double-edged swords. Radiology, 307 (2) : e230163, 2023. より許可を得て引用転載)

## 三次元マンモグラフィ (DBT)

まず、乳腺領域の画像AIの動向の一つとして、DBTについて少し触れておこう。

6月に発表された最新統計によると、DBTの占める割合はデジタルマンモグラフィ (DM) の46%に達している、とのことであり、2~3年以内にはDMの数を追い越すと予想されている<sup>4)</sup>。また、マンモグラフィセンターでは、87%の施設に少なくとも1台のDBTが設置されているという。しかしながら、読影の負担が課題となっており、AI-CADに期待がかかっている。

DMおよびDBTの読影用AI単体での性能評価を、系統的レビューとメタ分析 (過去に行われた複数の研究結果を統合し、より信頼性の高い結果を求める統計解析手法) で行ったYoonらの最近のRadiology論文では<sup>5)</sup>、次のように結論づけている。「DMでスクリーニングするための単体AIは、放射線科医と同等かそれ以上の性能が発揮されているが、それに比較して、DBTによるスクリーニング検査の読影におけるAIシステムの

性能を評価する研究の数はまだ不足している」と。そして、今後の取り組みとして、「レトロスペクティブな結果が、将来のがん検出の改善とプロスペクティブな設定でのスクリーニングプログラムの最適化の両方に確実につながるように、さまざまな実施戦略と継続的な品質管理に焦点を当てる必要がある」と結んでいる。

韓国のLunit社から2023年5月に公開されたRadiology-AIの論文では<sup>6)</sup>、従来型のCNN (convolutional neural network) ベースのディープラーニングモデル (スライス単位の2Dモデルと3Dモデルの両方) と、自然言語処理で成功したTransformerモデルの画像版であるVision Transformer (VT) モデルとを、それぞれ5000症例を超えるDBT症例で学習させて、米国内の9つの施設症例で乳がん分類性能の比較・検討を行っている。その結果、VTは2D-CNNよりも性能が良く (スライス間の情報を有効に活用するため)、3D-CNNに比べると、同一性能でもより効率的 (計算量を75%削減) であることを示しており、VTモデルを取り入れたDBTのためのAIシステムのさらなる展開に期待したい。

\* 1 基盤モデルとは？

これからの新しいAIモデルであり、ラベルの付いていないデータによってまず事前学習されたモデルのこと。これを用いて、タスクごとにわずかなファインチューニング作業をし、利用目的に合った最終的なモデルが作られる<sup>11)</sup>。2021年にスタンフォード大学の「Human-Centered AI Institute (HAI)」(人間中心のAI研究所)によって広められた用語。OpenAI社のGPT-3やStable Diffusionにも使われているCLIPなどがその代表例である。

適切に回答できるのか？」という目的の調査研究を行い、その内容を *Radiology* 誌に報告している<sup>12)~14)</sup>。

同論文では、「乳がん予防とスクリーニング」に関する25の質問を作成し、各質問をChatGPTに3回入力し、その回答を乳腺放射線科医が評価している。その評価として、①適切、②不適切、③信頼できない(矛盾した回答)、の3分類とし、その結果、25の質問のうち、88%の22の回答が「適切」と評価された。不適切とされた1問は、「COVID-19 ワクチン接種の時期にマンモグラフィ検査を受けるべきか」という質問に対するものであった。その回答は、古い情報に基づき「COVID-19 ワクチン接種でマンモグラフィ検査を4~6週間待機する」という回答で、同項目は2022年2月に、すでに「待機を推奨しない」とガイドラインで改定済みの内容であった。また、信頼できないとされた2問は、乳がんの個人リスクに関する質問と、マンモグラフィをどこで受けられるかという質問で、質問のたびに回答が大きく変化し、一貫性が見られなかったという。この論文では、ChatGPTが88%の質問に正しく回答でき、一般ユーザーが簡単に理解できるように、情報を消化しやすい形にまとめるという利点があるという。一方で、よく知られているように、ChatGPTが間違えることもある。そこで、ChatGPTのユーザーは、この技術が新しいもので検証が十分ではないことを認識し、現時点ではまだ医療者のアドバイスに頼るべきであると著者は語っている。

## Google社の最新のモデル：Med-PaLM2

Google社は医療上の質問に対する独自の大規模言語モデルとして、Med-PaLM<sup>15)</sup>を開発している。これは、2022年末に、米国医師免許試験(USMLE)形式の問題で最初に合格点を越えたモデルだという。また、その最新モデルであるMed-PaLM2では、USMLEの問題で85.4%

の精度に達したという。現在、X線やマンモグラムなどの情報を統合するマルチモーダル機能が追加されており、最終的にはこのモデルが医用画像から得られる情報を解釈できるようになり、放射線科医が読影とレポートの両方を行う支援装置としてこのモデルを使用できるようになると期待される<sup>16)</sup>。Med-PaLM2は、安全で役立つユースケースを特定するために、2023年の夏後半にクラウド顧客の小グループにフィードバックを提供する予定という。

### ◎

Microsoft社がOpenAI社へ100億ドル(約1.3兆円)を投資するという驚きのニュースが流れたのはまだ記憶も新しい(2023年1月23日)。その後、WordやExcelなどにもChatGPTが導入され、ついにはWindowsにも追加されるという。これからは仕事のやり方も研究の進め方も何もかもが大きく変わってしまうであろう、と予想するのは難くない。Google社は6月14日、画像検索ツール「Google レンズ」の新機能として、肌や爪、髪などの写真でそれらの状態を検索できるようにした、と発表している。これにより、「皮膚の状態を画像で検索」できるが、「検索結果は情報提供のみを目的としており、診断ではありません。アドバイスについては、かかりつけの医療機関に相談してください」という旨のただし書きがある。乳腺画像診断にも、このようなツールがすぐにでもできそうである。ただし、医療AIの進歩が急激に起きているが、利点の裏にあるリスク面も見逃せず、今後のさらなる議論が欠かせない。

本稿の内容も、数か月先には完全に陳腐なものとなってしまおうであろう。当面、これらのニュースから目が離せないし、これからの新しい時代が急速度でどのように変わっていくのかを見るのは、楽しみであり、また興味深くもある。なお、生成AIについては、本誌7月号の特集記事も参照されたい<sup>17)</sup>。

### ●参考文献

- 1) 人工知能学会 生成AI“活用へ環境整備を”“リスクや懸念も”。NHK, 2023.  
<https://www3.nhk.or.jp/news/html/20230607/k10014092021000.html>
- 2) 藤田広志：AIで進化する医用超音波技術 in 2022. 乳房AI総論. 超音波TECHNO, 35 (1): 33-36, 2023.
- 3) ACR Data Science Institute AI Central.  
<https://aicentral.acrdsi.org/>
- 4) Fornell, D. : 87% of mammography centers now have 3D breast imaging systems. Health Imaging, 2023.  
[https://healthimaging.com/topics/medical-imaging/womens-imaging/breast-imaging/87-mammography-centers-now-have-3d-breast-imaging-systems?utm\\_source=newsletter&utm\\_medium=hi\\_news](https://healthimaging.com/topics/medical-imaging/womens-imaging/breast-imaging/87-mammography-centers-now-have-3d-breast-imaging-systems?utm_source=newsletter&utm_medium=hi_news)
- 5) Yoon, J.H., Strand, F., Baltzer, P.A.T., et al. : Standalone AI for breast cancer detection at screening digital mammography and digital breast tomosynthesis : A systematic review and meta-analysis. *Radiology*, 307 (5) : e222639, 2023.
- 6) Lee, W., Lee, H., Lee, H., et al. : Transformer-based deep neural network for breast cancer classification on digital breast tomosynthesis images. *Radiol. Artif. Intell.*, 5 (3) : e220159, 2023.
- 7) Shen, Y., Heacock, L., Elias, J., et al. : ChatGPT and other large language models are double-edged swords. *Radiology*, 307 (2) : e230163, 2023.
- 8) Wang, S., Zhao, Z., Ouyang, X., et al. : ChatCAD : Interactive computer-aided diagnosis on medical image using large language models. 2023.  
<https://arxiv.org/abs/2302.07257>
- 9) 山下裕毅：ChatGPTがレントゲン画像を分かりやすく説明 中国の研究者ら「ChatCAD」開発. ITmedia, 2023.  
<https://www.itmedia.co.jp/news/articles/2303/01/news070.html>
- 10) Zhao, Z., Wang, S., Gu, J., et al. : ChatCAD+ : Towards a Universal and Reliable Interactive CAD using LLMs. 2023.  
<https://arxiv.org/abs/2305.15964>
- 11) 基盤モデルとは？ IBM, 2022.  
<https://www.ibm.com/blogs/think/jp-ja/what-are-foundation-models/>
- 12) Murphy, H. : ChatGPT offers “pretty amazing” recommendations on breast cancer screening, but oversight remains critical. Health Imaging, 2023.  
[https://healthimaging.com/topics/artificial-intelligence/chatgpt-offers-recommendations-breast-cancer-screening?utm\\_source=newsletter&utm\\_medium=hi\\_monthly](https://healthimaging.com/topics/artificial-intelligence/chatgpt-offers-recommendations-breast-cancer-screening?utm_source=newsletter&utm_medium=hi_monthly)
- 13) Haver, H.L., Ambinder, Bahl, M., et al.: Appropriateness of breast cancer prevention and screening recommendations provided by ChatGPT. *Radiology*, 307 (4) : e230424, 2023.
- 14) ChatGPTが「乳がん関連の健康アドバイス」で有効性を示す. The Medical AI Times, 2021.  
<https://aitimes.media/2023/04/12/13160/?5392>
- 15) Med-PaLM : A large language model from Google Research, designed for the medical domain.  
<https://sites.research.google/med-palm/>
- 16) Gkramani, Z. : Introducing PaLM2. Google, 2023.  
<https://blog.google/technology/ai/google-palm-2-ai-large-language-model/>
- 17) 特集：医療AIの普及拡大とさらなる展開 医療からヘルスケアサービスまで発展に向けた現状と展望. *INNERVISION*, 38 (7), 2023.