

第30回学術総会/パネルディスカッション 2

乳がん検診で AI をどのように使えばよいのか

乳がん検診領域における AI-CAD の開発と実用化の現状を探る

岐阜大学工学部

藤田 広志

要旨：従来型 CAD (computer-aided diagnosis) は、マンモグラフィを中心に商用化も実現され一定の成果が上がってはいるが、性能面など普及を阻む問題点も少なからず存在した。そのハードルを乗り越える新たな技術としてディープラーニングがあり、CAD はいま AI-CAD としてさらなる進化/多様化を遂げつつある。最近の著名な論文誌には、AI-CAD が医師と同等かそれ以上の性能を示すもの、複数の AI-CAD の性能を比較し有効性を検証したもの、AI-CAD により従来 CAD より偽陽性数減少や recall 数を減らせたもの、医師のワークロードを大幅に軽減すべく検診における正常画像と異常の可能性ある画像を AI-CAD により自動分類しその有効性を示したものなどがある。オランダの ScreenPoint Medical 社の CEO は、現在、商用化されているマンモグラフィ AI-CAD の性能は放射線科医の平均読影レベルに達しており、2025年頃にはハイレベルの放射線科医のレベルかそれ以上に到達するだろうと予測している。ただ、それでも自動診断ではなく、あくまで「AI-CAD+医師」の協働により、「AI-CAD」単独または「医師」単独よりもさらに性能が向上するというのが現状であろう。本稿では、基調講演で講演した最新の AI-CAD の進捗状況、解決されるべき問題点、乳がん検診における新しい AI-CAD の今後の利活用について、内容の一部を更新して解説する。

索引用語：人工知能 (AI)、ディープラーニング、AI-CAD、乳がん検診

はじめに

米国で最も早く Food and Drug Administration (食品医薬品局, FDA) の認可を得て商用化に成功したコンピュータ支援診断 (computer-aided diagnosis, 以下 CAD) システムは、レーザーデジタイザでフィルムをデジタル化して用いる検診マンモグラフィのためのセカンドリーダー型の CAD システムで、それは1998年のことであった^{1, 2)}。すなわち CAD システムの商用化の歴史は、20年程度のものである。その後、いくつかの画像診断領域で、複数の撮像モダリティについて CAD システムの商用化が続いた。しかしながら、つい最近までは、CAD に対する進捗・普及は停滞気味の感が否めない状況であった。

従来型のマンモグラフィ CAD について特徴をまとめると、以下ようになる。

- ・世界初の FDA 承認を得て商用化が始まり、20年余の歴史がある。
- ・医師がまず CAD なしで読影し、続いて CAD の検

出結果を参照し、その後、総合的に医師が最終診断を行う、と定義される「セカンドリーダー型」で、すべて検出型 CAD (CADe と表記する) である。CAD 単体での性能は医師のレベルよりも劣っていても、セカンドリーダー型として利用することにより、医師の最終診断の結果向上がみられる。

- ・米国では2001年にマンモグラフィ CAD の使用に対して保険会社の医療報酬補助が付いて、急速な普及が始まった。
- ・米国では約92%の検診マンモグラフィで CAD が利用される (2016年時点)³⁾。
- ・偽陽性候補数がまだ多く、要精検数が増加するなど性能面で問題が残る。CAD システムの開発コストが大きい。臨床現場でのワークフローが悪い。
- ・その後の大規模なマルチ施設での実臨床研究で、医師の診断性能が改善されないという結果を示す研究が複数報告され⁴⁻⁶⁾、有効性が大きく問われるようになった。

このようななか、昨今の第三次人工知能 (AI) ブームになり、その原動力ともいえる機械学習の一手法である「ディープラーニング (深層学習と和訳される)」技術により、CAD はその性能が大きくパワーアップするようになっている^{1, 2, 7-9)}。AI とディープラーニング

別冊請求先：〒501-1194 岐阜市柳戸 1-1 (医学部棟内)

岐阜大学工学部 藤田広志

e-mail address: hiroshi.fujita.gifu@gmail.com

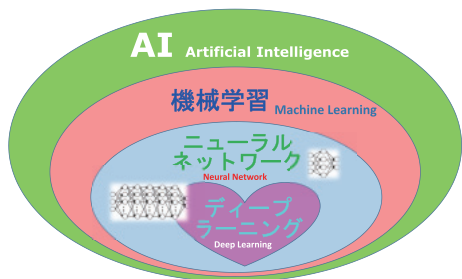


図1. AI からディープラーニングまで。AIはコンピュータに人間のような知能をもたせる研究分野である。機械学習(machine learning)は、経験に基づいた学習をさせることで機械(コンピュータ)を賢くさせるAIにおける重要な技術の一つである。ニューラルネットワーク(neural network)は、脳の情報処理の働きをモデルにした機械学習の技術。ディープラーニング(deep learning)は、深層学習とも呼ばれ、従来のニューラルネットワークを多層化することや、新しい学習技術、あるいはコンピュータの性能向上などにより進化した機械学習の手法で、昨今のAIブームを牽引している。

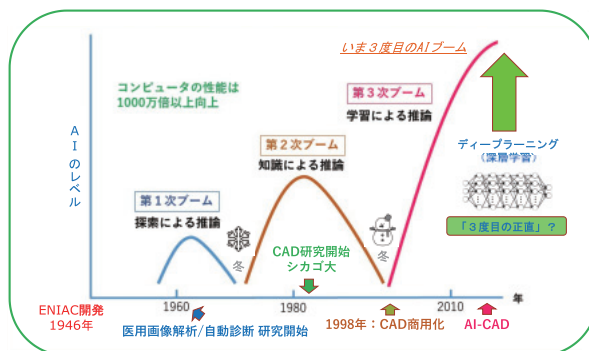


図2. 3回のAIブームとCAD研究の歴史的推移

の関係を図1に、AIとCAD研究の概略的な歴史を図2に示す。詳細は文献2などを参照されたい。なお、実は従来型CADにおいても広くAIの技術が使われていたが、最近のディープラーニング型では特にAI-CAD、ときには単にAIと呼称されることが多い。

第30回日本乳癌検診学会学術総会におけるパネルディスカッション「乳がん検診でAIをどのように使えばよいのか」(2020年11月23日)における筆者の基調講演の主な内容が、その後、「乳房画像におけるAI応用の現状」と題して、学術誌『乳癌の臨床』(36(1): 7-17, 2021年2月)⁹⁾に掲載されたが、本稿はこれに加筆修正して転載されたものであることをご了解いただきたい。

1. AI-CADの原動力：ディープラーニング

ディープラーニング技術がCADシステムのエンジン部分に活用されることにより、「従来型(伝統的)CAD」は、いま「新生CAD」(ここでは“AI-CAD”と呼ぶ)として、高度化が始まった。

従来型のCAD開発においては、画像の中の認識対象の特徴量(例えば、がん領域の形状や濃淡情報)を、設計者(人間)が考案し、それらを駆使して判断過程をルール化するソフトウェア(アルゴリズム)を人がすべて構築した(ルールベース法)。近年では、それらのデータ(特徴量)からルールを自動構築して識別を行う種々の機械学習の技術が使われている(ランダムフォレストやサポートベクターマシンなど)。

これに対して、ディープラーニングの利点は、この

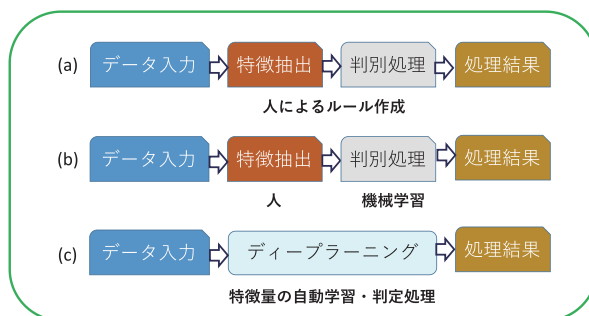


図3. CADの進化。(a)ルールベース法によるエキスパートシステム型CAD、(b)それに従来の機械学習を取り入れたもの、(c)ディープラーニング型AI-CAD

ような特徴量を自ら作り出し(ときには人間が気づかないような特徴量も含めて)、識別処理もすべて学習処理で行うことができる点にある。すなわち、従来型CADは図3の(a)や(b)のような開発過程に従ったのに対して、AI-CADは同図(c)に示す開発過程を経るため、性能が向上しかつ開発の手間が相当省けるようになった。従来では開発に何年もの期間を要したものが、ディープラーニングを使えば数か月でも開発が可能になった。

さて、このようなディープラーニングとは何であろうか？ ディープラーニングは人間の脳の神経回路網を、人工的にコンピュータで実現した人工ニューラルネットワーク(neural network: NN)の進化版である。NNの進化版と称した理由は、過去の2回のAIブームとも呼応して、すでに過去に2回のNNブームがあったためである。いま画像認識で話題のディープラーニングは、図4に示すような構造のもので、畳み込みニューラルネットワーク(convolutional neural network: CNN)と呼ばれる。畳み込み層、プーリング層、全結合層の3種類の層構造が多重に積み重なった構造であり、たくさんの種類のもの(モデル)が次々と考案されている。例えば、マンモグラフィの関心領域を同図の入力として、出力に「正常/良性/悪性」の分類(鑑別)が可能である。また、マンモグラフィ全体

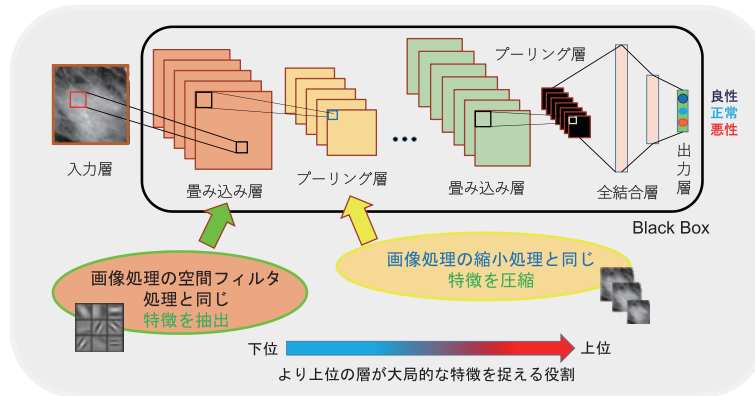


図 4. 畳み込みニューラルネットワークによる良悪性鑑別の例

を入力として、出力では異常部位を矩形の枠で囲む(検出)、あるいは病変部位を画素単位で決定する(領域抽出、セグメンテーション)、さらには解剖学的な構造(例えば、乳腺領域や胸筋領域など)を指摘する(認識)ことも可能である。また、5年後にがんが発症するかどうかの推定(予測)のようなことも不可能ではない。画質改善(超解像、ノイズ減少等)にも応用されている。

ディープラーニングでは、大量の学習データによってハイパーパラメータを調整することにより、システムができていく。そこで、ディープラーニング型の AI-CAD はデータ駆動型(data driven)といわれ、性能を決める勝敗は画像データの収集のパワーに依存する。単にデータの量が多いのみならず、データの質も重要であり、また、学習時に正解データも必要で、その作業はアノテーション(注釈付け)と呼ばれ、例えば、腫瘍の辺縁領域とかその有無や種類などの情報を付与する。ディープラーニングに必要な道具は、ハードウェアとしてのコンピュータさえあれば、ソフトウェアはオープンソースで簡単に入手できるため、今や誰でもディープラーニングを試せる時代(AI の民主化と呼ばれる)になっている。基本的にプログラミングなしで、単にブロック図を組み立てる感覚で簡単にディープラーニングを構築できるツールも公開されている¹⁰⁾。

2. AI によるマンモグラフィ CAD の性能向上の例

乳房画像診断領域におけるディープラーニング型の AI-CAD により、いかに従来の伝統的な CAD よりも性能が向上したのかを示す研究事例が、凄い勢いで各種の著名な論文誌に報告されている。以下では、そのようないくつかの研究例を紹介する。

1) 偽陽性候補の削除

偽陽性(FP)候補数に関して、従来型 CAD と AI-

CAD とを比較検討した報告がある¹¹⁾。実験では、従来型 CAD として FDA 承認済みの ImageChecker™ (Version 10.0, Hologic 社製)を利用し、AI-CAD には FDA 未承認のプロトタイプ of cmAssist™ (CureMetric 社製)を使用している。250症例の FFDM (Full Field Digital Mammography)マンモグラフィを用い、平均 69%の FP が削減でき(石灰化83%、腫瘍56%)、症例当たりの読影時間が17%減少されている。これにより、従来型 CAD では83%の症例で FP 候補が存在したものが、AI-CAD では52%の症例までの存在と大きく FP 数が削減されている。

2) 良性の精密検査症例の削減

マンモグラフィ検診での不要な精密検査数を減らせるかどうかを研究した報告がある¹²⁾。実験では、14,860枚(3,715患者)の検診マンモグラフィを用いて、ディープラーニングで悪性、良性、精密検査(recall)に回った良性の3ケースの自動分類を行っている。技術的には段階的な転移学習法^{脚注1)}を利用しており、自然画像のデータセット (ImageNet)で学習済みディープラーニングモデルを DDSM (Digital Database for Screening Mammography)データセット(フィルムをデジタル化して作成)でパラメータをチューニングして、さらに FFDM データで詳細なチューニングを行っている。このようなモデルチューニング法で性能が向上することを示すと同時に、リコールされた良性画像を悪性および陰性の症例から区別できることを示した。

脚注1 転移学習 (transfer learning) : ある領域で学習したこと(学習済みモデルという)を別の領域に役立たせ、効率的に学習させる方法。転移学習を用いたディープラーニングモデルは、限られたデータから高精度なモデルが作成できたり、ゼロから学習させるよりも短時間で学習ができるメリットがある。

3) 専門医と同等かそれ以上の性能

AI 単独だけの性能で、少なくとも医師と同等、あるいは医師以上の読影の精度を出す事例が少なからず報告されている。

例えば、Google の研究者らからの Nature 誌への報告として¹³⁾、イギリス(UK)のマンモグラフィデータベースとして2施設から計25,856枚、米国(USA)のデータベースとして1施設から計3,097枚を収集して実験を行った大規模なものがある。いくつかの項目の実験が行われ、AI-CAD は、医師と同等、あるいは医師6名の平均より性能がよかったこと、偽陽性数が減少したこと(USA データで5.7%、UK データで1.2%)、偽陰性が減少したこと(USA データで9.4%、UK データで2.7%)、UK データで学習した AI-CAD モデルを USA データでテストしても性能は落ちないという汎化性を示せたこと、UK の医師の二重読影においてセカンドリーダーの負担を88%も減少できたこと(第一読影医と AI-CAD との意見が異なる12%のケースについてのみ、第二の読影医が意見を出すという利用法で)を示している。

また、ニューヨーク大学の研究者らも、20万症例の100万枚を超えるマンモグラムでディープラーニングモデルを開発・評価し、経験豊富な放射線科医と同等のレベルであったことを、工学系の著名な論文誌に示している¹⁴⁾。

Nature Medicine 誌に掲載された最近の論文¹⁵⁾では、米国、英国、中国の5つの大規模なデータセットで AI-CAD を学習し、医師の性能を上回る結果や早期の乳がん指摘(45.8%)を示すとともに、異なるマンモグラフィ装置や異なるベンダーの装置、あるいはトモシンセシスへ応用してもうまくいくことまでも含めて報告している。なお、中間期乳がんを AI-CAD が医師よりも検出することができることを示した2つの研究成果が、北米放射線学会(RSNA)2020で報告されている。

このように AI 研究用に大規模のマンモグラフィデータが使われるようになり、素晴らしい研究報告が行われるようになってきている。スウェーデンのカロリンスカ研究所からも、ディープラーニング学習用に利用可能という200万枚という非常に大規模なマンモグラフィデータベース(cohort of screen-aged women: CSAW)が開発されており、文献16に記載の研究者に申し込めば利用可能である。

4) 乳がん発症の予測モデル

文献17の論文では、マサチューセッツ総合病院の患者4万人分からなる9万枚のマンモグラム、乳腺濃度

スコア、患者のがん発症に関する臨床データを使用して、データから一見してそれとはっきりわからない、そして医師でも識別できないパターンを特定し、人種に関係なく乳がんを発症の5年前に予測できるディープラーニングベースの予測モデルの開発に成功している。問題点として、実験は単一施設のデータであり、かつ単一ベンダーの装置によるマンモグラフィであること、ディープラーニングのブラックボックス性を挙げている。これらの指摘は、決して本研究に限られるものではない。

また、文献18の研究では、質問票とマンモグラフィ乳腺濃度を基本とした乳がんリスク予測モデルとディープラーニングをベースにしたモデルの比較を、カロリンスカ大学病院の2,283人のデータ(後に278人ががんと診断されている)を用いて実験を行っている。その結果、ディープラーニングによるモデルの予測はより正確性があり、より進行性のがん(aggressive cancer)の偽陰性率は低くなることを示している。

3. CAD の進化・多様化

以上で説明したように、ディープラーニングにより CAD の目的の多様化、利用形態の多様化、さらには高度化・進化が起きている(図5)。本章では、これらについて、説明する。詳細は文献などを参照いただきたい^{1, 2, 7-9)}。

1) 目的の多様化

長年にわたり利用されているのは検出を支援する『CADE』のみであった。2017年に診断型と呼ばれる悪性度を数値で定量的に示すことにより鑑別診断を支援する『CADx』が、まずは乳房 MRI を対象として FDA の初の承認を得て商用化されている。最近では、『CADE+CADx』と分類される両方の機能をもつマンモグラフィあるいはトモシンセシス用の AI-CAD も商用化されている。

CAD 技術を応用・拡張し、放射線科医が読影する前の撮影直後の画像を分析して、対処の緊急性の有無

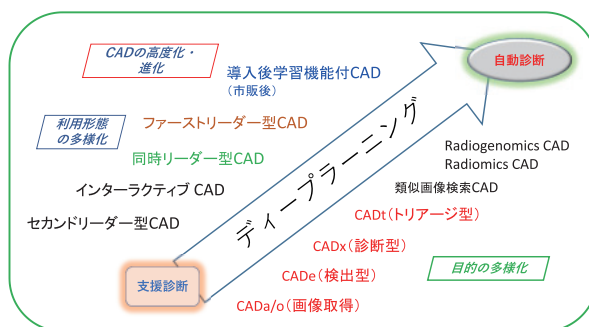


図5. ディープラーニングによる CAD の新潮流

を専門医に提示・警告する(ワークリストの優先順位付けの最適化; prioritization)システムがある。これは“トリアージ型 CAD”と呼ばれ(AI as triage), 最近の FDA の分類では『CADt』と表示される。脳卒中をトリアージするものが FDA 承認第 1 号(2018年)であり, 2020年に米国メディケアによる保険補助にも, AI 医療機器として初めて成功している。最近ではマンモグラフィ用の CADt も商用化に至っている。例えば, 現在までに少なくとも 6 つの画像診断領域でトリアージ CADt の商用化に米国で成功しているイスラエルの Zebra Medical Vision 社のマンモグラフィ CADt がある。同社のホームページでは¹⁹⁾, 「トリアージ製品は, スクリーニングスタディの先読みとして機能し, 放射線科医が最初に疑わしいものに対処できるようにする。あるいは, 症例を経験豊富な放射線科医と経験の浅い放射線科医に分けることができる。追加の使用例は, 放射線科医によって読み取られた後に結果を提供することである。これにより, スクリーニングプログラムの 2 番目のリーダーおよびセーフティネットとして機能する」と説明している。2021年 4 月には, DeepHealth 社の 2D マンモグラフィと 3D マンモグラフィのトリアージ製品も, FDA の承認を得ている。AI によるスクリーニングワークリストの優先順位付けツール(AI-driven screening worklist prioritization tool)としている。

このような CAD におけるトリアージでは, ①がん候補症例をトリアージして画像読影の優先順位付けのみを目的とし, 医師はすべての画像を最終的に読影するものと, ②トリアージして正常例を選別し, 医師は異常例のみを読影してワークロードを改善するものに分けられるようである。後者については, 次項のファーストリーダー型 CAD のところでも触れる。

『CADa/o』は2020年 2 月に新しい種類の CAD 機器として FDA 承認されたもので, コンピュータ支援取得/最適化と呼ばれ, 心臓超音波検査に対する AI ガイド画像システム“Caption Guidance”という商品があり, COVID-19による心疾患の検査支援に有益と期待される。

2) 利用形態の多様化

セカンドリーダー型 CAD(図 6a)に続いて, 医師が画像上の気になる箇所をクリックしたときにのみマンモグラフィ CAD の結果が表示されるものとして「インタラクティブ(interactive)CAD」(図 6b)があるが, オランダベースの ScreenPoint Medical 社の transparaTM が2018年11月に初めて FDA で承認され商用化された。CADe+CADx の機能を有している。同

社のホームページには, 100万枚以上の画像でモデルを学習していると謳われている。同製品の有効性を示す論文も出されており, インタラクティブであるため読影時間は増えないという²⁰⁾。

また, CAD の結果を最初から参照する「同時リーダー型(concurrent reader)CAD」(図 6c)として, 2016年 FDA 承認の QView Medical 社の 3D 乳房超音波 AI-CAD(QVCADTM)や, あるいは iCAD 社の 3D 乳房トモシンセシス(ProFound AITM)で商用化されている。最近では, 2020年 7 月にフランスの Therapixel 社が, CADe+CADx 型のマンモグラフィ AI-CAD で FDA 承認を得ている。また, 同 AI-CAD の利用により, 読影時間はほぼ同じで, 特異度に影響なく医師の検出感度が増加したとの論文が出ている²¹⁾。

さらに高度なものに「ファーストリーダー型(first reader)CAD」(図 6d)があり, これは最初に CAD が単独で解析処理を行い, 医師がチェックすべき画像と明らかに正常でその必要がないものを選定するタイプの CAD であり(AI triages negative cases), 正常と判定したものに対しては“半自動診断”に相当し, 機能によってはレポート作成まで行う。これにより, 医師は異常と判定された症例のみに集中できる。ほとんどが正常症例である検診での利用が強く望まれる。このとき, 異常候補に対してこのように単に選別処理をするのみのトリアージのタイプ(CADt)と, CAD ががん候補の検出(CADe)や鑑別解析(CADx)もしてその結果を示すタイプとが考えられる。後者では, 医師は異常ケースと不確定なケースのみを読影する。2020年に欧州の薬事承認といえる CE 規格認証を得たものでは(ドイツの Vara 社の製品), マンモグラフィ乳がん検査で97%もの正常例を除外するという。

このような正常ケースを事前に選別して, 異常のある可能性が高い症例の読影に医師が集中できるようにして, 読影の負荷を減らすような利用法の可能性を検討した研究論文を紹介する²²⁻²⁶⁾。例えば, カロリンスカの Dembrower らは²³⁾, 乳がん症例547, 正常症例 6,817を用いて, 韓国の Lunit 社の AI-CAD(ver. 5.5.0)でがん存在の予測スコア機能を AI トリアージツールとして利用し, 放射線科医のマンモグラフィの作業負荷を50%以上削減することを示している。また, 米国 UCLA の Kyono らは²⁴⁾, 7,000人以上のデータセットを使用し, 正常マンモグラフィの識別率を自律型放射線科医アシスタント(AURA)というモデルを作成・検討している。その結果, 陰性的中率=99%に固定して, がん有病率が15%のときに正常識別率が34%, がん有病率が1%のときに同91%であり, これにより医師は困難な症例に集中するために時間をより

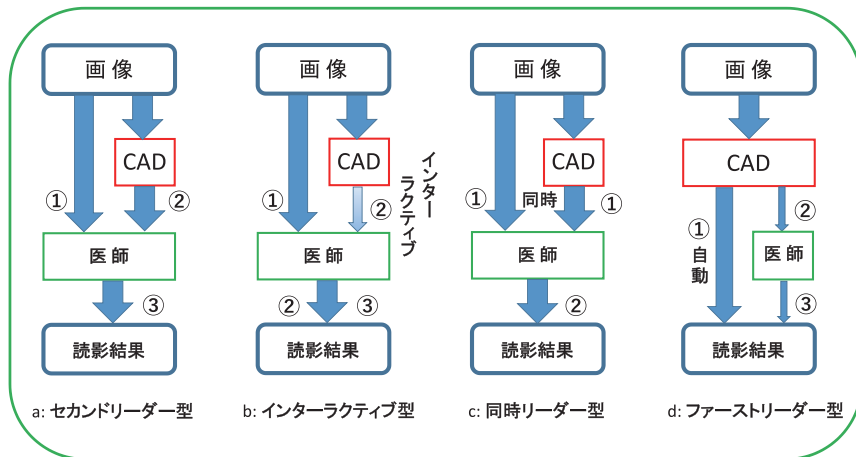


図 6. CAD の利用形態の多様化

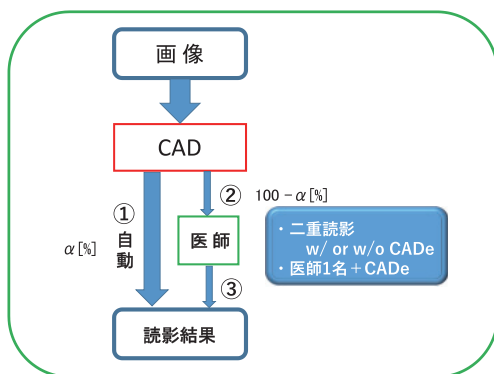


図 7. 検出 CADe のがん候補存在確率提示機能または専用のトリージ CADt を利用するファーストリーダー型マンモグラフィ検診 CAD 戦略

有効に使えるようになり、誤診を減らせる可能性がある」と報告している。米国 MIT の Yala らの論文では²⁵⁾、22万余の症例でトリージ実験を行い、ワークロードを減少させる可能性を見出し、検出感度を変ることなく特異度を増すことできたとの結果を示している。これらはすべて 2D マンモグラフィのみを対象とした研究であったが、2021年掲載の文献26では、ScreenPoint Medical 社の AI-CAD におけるがん存在スコアをトリージ機能として利用して、二重読影下のデジタルマンモグラフィ (DM) またはデジタル乳房トモシンセシス (DBT) を使用した乳がん検診で、がんの検出率を 5% 以上減らすことなく作業負担を減らすことができるかどうかを検証するために、「自律型 AI トリージスクリーニング戦略 (autonomous AI triaging screening strategy)」という研究を行っている。その結果、AI 利用により、2D でも 3D スクリーニングでも作業負担を最大 70% 減少できる可能性があり、また DM スクリーニングから DBT スクリーニングに移行すると 30% のワークロード削減、25% 感度向上、



図 8. 世界初で米国 FDA 認可された米国 IDx 社が開発した自律型 AI 診断システム。糖尿病性網膜症を検出する眼科用 AI 医療機器「IDx-DR」(<https://dxs.ai/>より)

27% リコール率の低下があったとの結果を得ている。今後の新しい検診形態として、検出型 CADe の機能を活用あるいはトリージ CADt を利用して、正常症例の α [%] をまず取り除き、その後の残りの $(1 - \alpha)$ [%] の乳がんの可能性のより高い症例を、従来型の二重読影(このとき CADe なし、あるいは CADe ありが考えられる)もしくは CADe 利用による医師 1 名の読影という形態のマンモグラフィ検診 CAD 戦略が考えられるであろう(図 7)。

3) 高度化・進化

CAD の高度化・進化で今後期待される CAD として、まず、「導入後(市販後)学習機能付 CAD」が挙げられる。これは学習機能をもったディープラーニングの出現により、臨床現場に AI-CAD の導入後(市販後)でも、施設固有の新しい画像データの追加・再学習により、プログラムが自動でアップデートされ、ど

んどん賢くなる AI-CAD である。ただし、性能が改善される可能性もあるため、それを防止する対応策など、解決されるべき技術的な問題点が多くあり、本邦の薬事承認に関するガイドラインでも触れられている²⁷⁾。商用化までにはまだ時間がかかりそうである。

2018年4月、「IDx-DR」という名称の眼底画像を対象とした AI 検査機器(DRは糖尿病性網膜症の diabetic retinopathy の略、図 8)が、FDA の最終販売承認を取得した。この装置により、「DRを検出：専門医の受診を勧める」あるいは「DRは未検出：12か月以内の再検査を勧める」という出力結果が得られる。本機器は、写真もしくは解析結果の診断が眼科医でなくてもプライマリケアドクターが利用できるもので、新しいジャンルの AI 医用機器である。AI ドクターとか自律診断(autonomous diagnosis)と呼ばれるようになっている。ただ、注意が必要なのは、米国の2つの医学放射線学会からも共同の声明が出されているように²⁸⁾、このような自律的な AI の臨床への導入では、さらなる安全性なども十分検証されなければならないことである。

4. 医師との協業が重要

マンモグラフィにおける AI-CAD については、多くの成果が論文で出されている昨今であり、AI 単独でもかなりの高い性能を示す例が多い。しかしながら、現時点でもっとも有効的な使い方は、やはり診断支援として AI-CAD を医師がパートナーとして利用すべきであるという点は、従来からの CAD の基本的な考え方と何ら変わりない。いくつかの研究論文報告があるが^{14, 29-31)}、ここでは以下に3点について紹介しよう。

カロリンスカ研究所からの論文では²⁹⁾、商用化されている3社の AI-CAD を利用して、検診マンモグラフィの乳がん739症例、正常8,066症例を使い、各種の比較実験を行っている。そして、最良の商用 AI-CAD は今後の前向きな臨床研究として独立した読影者として十分な性能を示し、また、AI-CAD と医師の組み合わせは、医師2名の二重読影よりの性能がよかったとの結果を報告している。

44か国から126のチームが参加して2016年末から開始された大規模マンモグラムを利用した“デジタルマンモグラフィ Dream Challenge”の結果を纏めた論文では³⁰⁾、AI 単独では医師を超える性能は示されなかったが、“協調”によりトータルの精度が向上することを示している。

韓国 Lunit 社の AI-CAD(CE マーク取得済)を使用した研究では³¹⁾、3か国で5施設から収集された約17

万枚のマンモグラムを使用し(内訳は、韓国：14万以上、米国：1.5万、英国：6.5千)、生検で確認された独立した癌陽性例36,000件以上のデータセットが含まれ、マンモグラフィ関連の AI 研究のなかでは最大規模のがんデータとなっている。そして、AI を使用する前と使用した後で、放射線科医の性能が大幅に向上したことが示された。AI だけでは乳がんの検出感度が88.8%だったのに対し、放射線科医だけでは75.3%であったが、放射線科医が AI の補助を受けた場合には、精度は9.5%上昇して84.8%になったという。

5. 他のモダリティにおける AI 応用

乳房超音波像へのディープラーニング応用については、上述の QView Medical 社のように、3D 超音波像の検出型 AI-CAD や Koios Medical 社の Koios DS for breast という悪性度を自動推定する AI-CAD が FDA の承認を得た製品が出現している。わが国でも研究が盛んであり、例えば Fujioka らは³²⁻³⁴⁾、良・悪性の鑑別においてディープラーニングが放射線科医より感度でも特異度でも高い性能を示している。また、超音波の動画像に対してリアルタイム処理が可能な AI 応用システムの開発なども企業との共同研究体制が進められている³⁵⁾。ほかにも多くの開発事例が国内外でみられるが、それについては他の解説論文等³³⁻³⁷⁾を参照されたい。ただ、検診というような枠組みでの乳房超音波 AI の利活用に関する研文は、まだ行われていないようである。

MRI についても上記したように、すでに診断型の CADx として商用化されたものがあるが、学術論文については、解説文献を参照されたい^{36, 38)}。T1 強調(造影剤有と無)の MRI 画像情報(radiomics)に加え70の遺伝子情報(genomics)をディープラーニングに同時に入力し、4分類(pathological stage, ER status, PR status, HER2 status)を予測する CAD(radiogenomics CAD)³⁹⁾なども報告されている。

研究における障壁

1. 学習データ不足
 - ⇒ 各種領域での公共の画像データベースのさらなる開発
 - ⇒ 少数の教師データ、あるいは教師なし学習法の開発
2. 実験室研究・検討から実臨床研究・検証へ

AIの実装における課題

3. AIの効果的な利用法(支援診断, 半自動診断, 自動診断)
4. 既存のシステムへの組み込みとシームレス化
5. 薬事承認と保険償還

倫理的側面

6. AI診断に対する責任の所在(現在は医師に)
 - ⇒ 医師? 病院? ベンダー?
7. ディープラーニングのブラックボックス性の解決
 - ⇒ 説明可能なAI(explainable AI)研究へ
8. 医療関係者から患者までのAIへの理解・教育の推進

図 9. 医療画像支援 AI における諸課題

おわりに

医療画像支援AIでは図9に示すような諸課題がある。

最後に、以下は本稿のまとめである。

- ・ディープラーニング型のAI-CADの開発、実用化により、伝統的CADに比べて性能が格段に向上している。
- ・AI単体で医師の性能を凌ぐ研究事例が増えているが、医師との協調によりさらに性能が向上するため、そのような利用方法の有効性・妥当性を示す研究も少なくない。
- ・乳がん診断のためのCADの種類もCADe(検出), CADx(診断), CADe+CADx, CADt(トリアージ)と多様化し、またCADの利用方法もセカンダリー、インターラクティブ、同時リーダー、ファーストリーダー型と増えてきた。
- ・検診において、正常と異常症例を最初にふるい分けることにより、医師の読影負担を減らす手段にAI-CADは利用できるようなると期待される。ただし、利用環境に応じた実証実験が必要である。
- ・二重読影における片方の医師の代わりに利用できる可能性もある。
- ・AI-CADの実臨床におけるプロスペクティブな実証による検証研究が望まれる。また、ディープラーニングにおけるブラックボックス性が少しでも解決され、「説明可能なAI」として利用できることが望まれる。
- ・最後に、AI-CADの利用に際し注意すべきは、診断・治療を行う主体は医師であり、医師はその最終的な判断の責任を負うこととなる(医政医発1219第1号、平成30年12月19日:「人工知能(AI)を用いた診断、治療等の支援を行うプログラムも利用と医師法第17条の規定との関係について」)。

【文献】

- 1) 藤田広志: 乳房領域へのAI応用の歴史とこれから。臨床画像, 35(10): 1129-1138, 2019
- 2) 藤田広志監・編: 医療AIとディープラーニングシリーズ〈2020-2021年版〉はじめての医用画像ディープラーニング——基礎・応用・事例, オーム社, 東京, 2020
- 3) Keen JD, Keen JM, Keen JE: Utilization of computer-aided detection for digital screening mammography in the United States, 2008-2016. J Am Coll Radiol, 15 (1 Pt A): 44-48, 2018
- 4) Fenton JJ, Taplin SH, Carney PA, et al: Influence of computer-aided detection on performance of screen-

ing mammography. N Engl J Med, 356 (14): 1399-1409, 2007

- 5) Fenton JJ, Abraham L, Taplin SH, et al: Effectiveness of computer-aided detection in community mammography practice. J Natl Cancer Inst, 103: 1152-1161, 2011
- 6) Lehman CD, Wellman RD, Buist DSM, et al: Diagnostic accuracy of digital screening mammography with and without computer-aided detection. JAMA Internal Med, 175 (11): 1828-1837, 2015
- 7) 藤田広志監: 学ぶ! 究める! 医療AI——ディープラーニングの基礎から研究最前線まで, iv-MOOK vol.1, インナービジョン, 東京, 2020
- 8) 井川房夫, 藤田広志編著: これだけでわかる! 医療AI, 中外医学社, 東京, 2021
- 9) 藤田広志: 乳房画像におけるAI応用の現状. 乳癌の臨床, 36(1): 7-17, 2021
- 10) 藤田広志(監修), 福岡大輔(編著): 医療AIとディープラーニングシリーズ〈2020-2021年版〉標準 医用画像のためのディープラーニング入門編, オーム社, 東京, 2020
- 11) Mayo RC, Kent D, Sen LC, et al: Reduction of false-positive markings on mammograms: A retrospective comparison study using an artificial intelligence-based CAD. J Digit Imaging, 32 (4): 618-624, 2019
- 12) Aboutalib SS, Mohamed AA, Berg WA, et al: Deep learning to distinguish recalled but benign mammography images in breast cancer screening. Clin Cancer Res, 24 (23): 5902-5909, 2018
- 13) McKinney SM, Sieniek M, Godbole V, et al: International evaluation of an AI system for breast cancer screening. Nature, 577: 89-94, 2020
- 14) Wu N, Phang J, Park J, et al: Deep neural networks improve radiologists' performance in breast cancer screening. IEEE Trans Med Imaging, 39 (4): 1184-1194, 2020
- 15) Lotter W, Diab AR, Haslam B, et al: Robust breast cancer detection in mammography and digital breast tomosynthesis using an annotation-efficient deep learning approach. Nat Med, 27: 244-249, 2021
- 16) Dembrower K, Lindholm P, Strand F: A multi-million mammography image dataset and population-based screening cohort for the training and evaluation of deep neural networks—the cohort of screen-aged women (CSAW). J Digit Imaging, 33 (2): 408-413, 2020
- 17) Yala A, Lehman C, Schuster T, et al: A deep learning mammography-based model for improved breast cancer risk prediction. Radiology, 292 (1): 60-66, 2019
- 18) Dembrower K, Liu Y, Azizpour H, et al: Comparison of a deep learning risk score and standard mammographic density score for breast cancer risk prediction. Radiology, 294 (2): 265-272, 2020
- 19) Zebra Medical Vision, Inc: AI Solutions. MAMMO

- SOLUTION. <https://www.zebra-med.com/mammo-solution> (Accessed Jan. 18, 2021)
- 20) Rodríguez-Ruiz A, Krupinski E, Mordang JJ, et al: Detection of breast cancer with mammography: Effect of an artificial intelligence support system. *Radiology*, 290 (2): 305-314, 2019
 - 21) Pacilè S, Lopez J, Chone P, et al: Improving breast cancer detection accuracy of mammography with the concurrent use of an artificial intelligence tool. *Radiol Artif Intell*, 2 (6): e190208, 2020
 - 22) Rodríguez-Ruiz A, Lång K, Gubern-Merida A, et al: Can we reduce the workload of mammographic screening by automatic identification of normal exams with artificial intelligence? A feasibility study. *Eur Radiol*, 29 (9): 4825-4832, 2019
 - 23) Dembrower K, Wählin E, Liu Y, et al: Effect of artificial intelligence-based triaging of breast cancer screening mammograms on cancer detection and radiologist workload: A retrospective simulation study. *Lancet Digit Health*, 2 (9): E468-E474, 2020
 - 24) Kyono T, Gilbert FJ, van der Schaar M, et al: Improving workflow efficiency for mammography using machine learning. *J Am Coll Radiol*, 17 (1 Pt A): 56-63, 2020
 - 25) Yala A, Schuster T, Miles R, Regina t al: A deep learning model to triage screening mammograms: A simulation study. *Radiology*, 293 (1): 39-46, 2019
 - 26) Raya-Povedano JL, Romero-Martín S, Elías-Cabot E, et al: AI-based strategies to reduce workload in breast cancer screening with mammography and tomosynthesis: A retrospective evaluation. *Radiology*, 300 (1): 57-65, 2021
 - 27) 次世代医療機器・再生医療等製品 評価指標作成事業 (平成29年度 人工知能分野審査 WG 報告書, 平成30年 3月). http://dmd.nihs.go.jp/jisedai/Imaging_AI_for_public/H29_AI_report_v2.pdf (Accessed May 24, 2021)
 - 28) “Public Workshop – Evolving Role of Artificial Intelligence in Radiological Imaging;” Comments of the American College of Radiology (June 30, 2020). https://www.acr.org/-/media/ACR/NOINDEX/Advocacy/acr_rsna_comments_fda-ai-evolvingrole-ws_6-30-2020.pdf (Accessed Aug 2, 2021)
 - 29) Salim M, Wählin E, Dembrower K, et al: External evaluation of 3 commercial artificial intelligence algorithms for independent assessment of screening mammograms. *JAMA Oncol*, 6 (10): 1581-1588, 2020
 - 30) Schaffer T, Buist DSM, Lee CI, et al: Evaluation of combined artificial intelligence and radiologist assessment to interpret screening mammograms. *JAMA Netw Open*, 3 (3): e200265, 2020
 - 31) Kim HE, Kim HH, Han BK, et al: Changes in cancer detection and false-positive recall in mammography using artificial intelligence: A retrospective, multi-reader study. *Lancet Digit Health*, 2 (3): E138-E148, 2020
 - 32) Fujioka T, Kubota K, Mori M, et al: Distinction between benign and malignant breast masses at breast ultrasound using deep learning method with convolutional neural network. *Jap J Radiol*, 37 (6): 466-472, 2019
 - 33) 藤岡友之, 森美央, 山鹿絵美, 他: 乳がん画像診断における AI の研究・開発の動向 2) 超音波. *INNERVISION*, 35(8): 66-67, 2020
 - 34) Fujioka T, Mori M, Kubota K, et al: The utility of deep learning in breast ultrasonic Imaging: A Review. *Diagnostics*, 10 (12): 1055, 2020
 - 35) 林田哲, 北川雄光: AI の乳房超音波検査への応用に向けて解決すべき諸課題. *INNERVISION*, 35(7): 41-43, 2020
 - 36) 村松千左子, 藤田広志: 乳腺領域におけるディープラーニング研究の最新動向. *INNERVISION*, 34(7): 46-48, 2019
 - 37) Mahmood T, Li J, Pei Y, et al: A brief survey on breast cancer diagnostic with deep learning schemes using multi-image modalities. *IEEE Access*, 8: 165779-165809, 2020
 - 38) Sheth D, Giger ML: Artificial intelligence in the interpretation of breast cancer on MRI. *J Magn Reson Imaging*, 51 (5): 1310-1324, 2020
 - 39) Yoon HJ, Ramanathan A, Alamudun F, et al: Deep radiogenomics for predicting clinical phenotypes in invasive breast cancer. *Proc SPIE (14th Int Workshop on Breast Imaging, IWBI 2018)*, 10718: 107181H-1-107181H-6, 2018

Exploring the current status of the development and practical application of AI-CAD in the field of breast cancer screening

Hiroshi Fujita

Faculty of Engineering, Gifu University

Conventional CAD (computer-aided detection / diagnosis) has been commercialized mainly in mammography and has achieved some results, but there are not a few problems that hinder its widespread use such as performance. Deep learning is a new technology that overcomes this hurdle, and CAD is now undergoing further evolution / diversification as AI-CAD. Recent prominent journals include those that show that AI-CAD performs as well as or better than physicians, those that compare the performance of multiple AI-CADs and verify their effectiveness, those that show that AI-CAD can reduce the number of false positives and the number of recalls compared to conventional CAD, and those that show the effectiveness of AI-CAD in automatically classifying normal and possibly abnormal images in screening (triage type) to significantly reduce the workload of physicians. The CEO of ScreenPoint Medical in the Netherlands predict that the performance of mammography AI-CAD currently on the market will reach the average radiologist's interpretation level, and will reach the level of high-level radiologists or higher around 2025. However, even so, the current situation is that the performance will be further improved compared to "AI-CAD" alone or "doctor" alone, not by automatic diagnosis, but by the collaboration of "AI-CAD + doctor". This paper describes the progress of the latest AI-CAD, the problems to be solved, and the future utilization of the new AI-CAD in breast cancer screening, based on the contents of the author's keynote lecture, with some of the contents updated.

Key words: artificial intelligence(AI), deep learning, AI-CAD, breast cancer screening
