

[AI 画像診断が医療現場を変える]

1 AI 画像診断の全体像と将来の展望

— 医師を助ける “第三の目” —

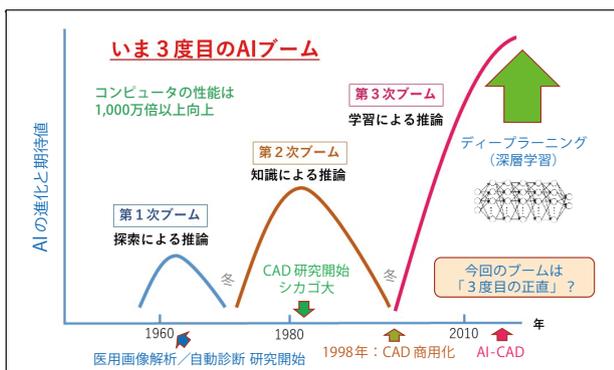
基
般

藤田広志 | 岐阜大学工学部

画像診断システムの開発史

医用画像診断において「誤診」はどうしても避けられない課題である。特に集団検診のように大量に画像が発生するところでは、誤診が起りやすく、そのため医師のコンピュータによる診断支援への要望は昔から強く、昨今のように「画像の洪水」が起きている時代には、ますますその期待が大きい。また、本邦は、対人口比のCT・MRI保有台数で世界1位を誇っているが、読影（画像診断）をする放射線科の医師は大幅に不足で、したがって医師1人あたりの読影件数も世界1位という苦い実情がある。

コンピュータによる画像診断への取り組みには、「支援診断」（Computer-Aided Diagnosis, 以下CAD）あるいは最終的な「自動診断」（Automated Diagnosis, Autonomous Diagnosis）を目指すとして、すでに1960年頃からの長い研究開発の歴史がある（図-1参照）。



■ 図-1 AIの進化と期待およびCADの開発の推移

最初に実用化に成功したのは、乳がん検出を目的としたコンピュータ支援検出（Computer-Aided Detection：以下CADe）システムである。これは、シカゴ大学で基礎研究されたシステムをベースに、米国のベンチャー企業R2 Technologies, Inc.（現Hologic）が、乳がんX線検査に使用される画像（以下、マンモグラフィ）を対象に開発した商品で、1998年に米国の食品医薬品局（Food and Drug Administration：以下FDA）の審査・承認を得て、商用化に至っている。当時はまだアナログ（連続情報）と称せられるフィルムが使われていた時代であり、レーザーディジタイザによってフィルムをスキャンし、コンピュータで取り扱えるデジタルデータ（離散情報）にまず処理する必要がある。この1998年はしばしば“CAD元年”と呼ばれる。

マンモグラフィCADに続いて、肺がんX線検査で利用される胸部単純X線写真や胸部CT画像の肺がん（結節陰影）検出のCADや、大腸CT検査におけるポリープ検出のCADなどが米国で次々と商用化された。

しかしながら、実用化（＝商用化）されつつ本格的に普及に成功した例は、マンモグラフィCADに限定された。その最大の成功要因は、マンモグラフィCADの利用に対して、2000年初頭から医療保険の補助金支援が可能になったことである。2016年には乳がん検診マンモグラフィの画像読影の92%で

CADが使われるようになったという^{☆1}。

このようなCADの利用には、使用法に厳密な定義があった。それはコンピュータの結果を「第2の意見」として医師が利用するもので、『セカンドリーダー型 (Second Reader) CAD』といわれ、

- ① 医師はまずCADの結果なしで画像を単独読影し、その後、
- ② CADの結果（“第2の意見”）を参照して、医師が最終診断を下す、

というものである（図-2）。これには、最初からCADの指摘箇所のみを参照してほかの箇所を観（診）なくなってしまうことを防ぎ（CADも見落としがあるため）、あるいは拾いすぎ（偽陽性）候補を医師が精査する機会を与えるという効果がある。しかし、一方で、読影時間が増えてしまう傾向もあった。

CAD商用化歴の約20年の間に、いろいろな問題点が明らかになってきた。特に、① 開発コストが大きい、② 検出性能（偽陽性候補の数も問題）がまだ不十分（マンモグラフィ検診では要精検率が增加するなど）、③ 商用化に至ったのはすべて検出支援型のCADeであり、検出部位ががんかどうかを解析・識別する（鑑別）診断支援型のCADx（後述）は、FDAの承認例が皆無、④ 実臨床で検証したところ有効性がでなかった、とする複数の論文の出現^{☆2}、⑤ 特定の病変検出にしか対応できていない、⑥ 臨床現場での使い方が面倒！（ワークフローが

^{☆1} Keen et al., J Am Coll Radiol. 2018;15:44-48

^{☆2} Fenton et al., N Engl J Med. 2007;356:1399-1409



■ 図-2 コンピュータ支援診断 (CAD) の概念図

悪い)、の理由が挙げられる。

AI-CADとして新時代に

トロント大学のG. Hintonが、2016年秋にトロントで開催された機械学習に関するある国際会議で、「5年以内にディープラーニングは専門医（放射線科医）のレベルに達するだろう」という主旨の衝撃的なスピーチをし、大きな話題となった。

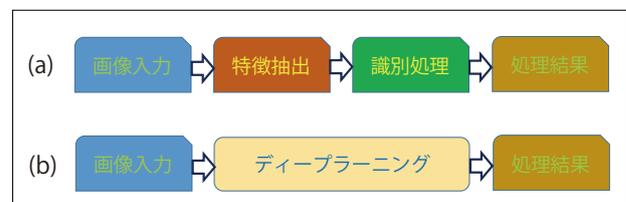
この予測を実証するかのように、2017年1月に、スタンフォード大学のA. Estevaらが、皮膚がんの診断にディープラーニングを活用した結果をNature誌に報告している^{☆3}。彼らは約13万枚の皮膚疾患の画像を収集し、「メラノーマ（悪性黒色腫）」や「良性腫瘍」をAIで学習させた結果、皮膚科医と同等の精度で皮膚がん診断できたという。さらに、その翌年、ハイデルベルグ大学のH. A. Haenssleらの「人間対機械」との主題の論文では、10万枚以上の画像で開発されたディープラーニングと皮膚科医がメラノーマを鑑別したところ、その精度は皮膚科医 vs AIが87% vs 95%となり、AIに軍配が上がっている^{☆4}。

いまや伝統的 (Traditional) とも呼ばれる「従来型CAD」では、画像の中の認識対象の特徴量を、設計者（人間）が苦勞して考案・作成したのに対して（図-3 (a)）、ディープラーニング（図-4）の利点は自ら特徴量を作り出す（学習する）ことができる点にある（図-3 (b)）。

ルールベース法に基づくものや、ディープラーニ

^{☆3} Esteva et al., Nature. 2017;542:115-118

^{☆4} Haenssle et al., Ann Oncol. 2018;29:1836-1842



■ 図-3 (a) 従来型CADと (b) ディープラーニング型AI-CADの開発過程の相違

特集
Special Feature

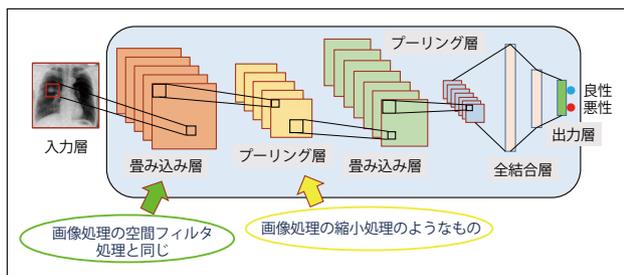
ング以前の従来の機械学習を用いるもの（分類・識別処理にニューラルネットワークやサポートベクターマシンを使う部類）までを「従来型CAD」と呼び、ディープラーニング型のCADを「AI-CAD」と呼んで区別する。

進化・多様化

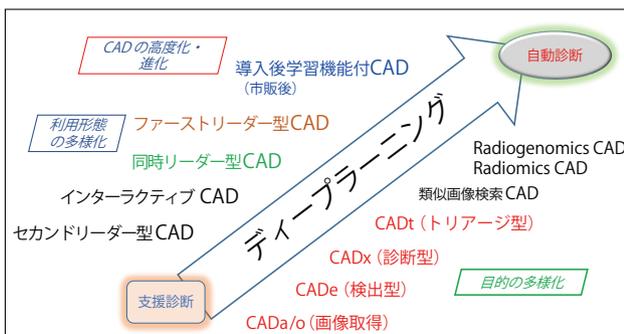
ディープラーニングの出現により、AI-CADとして、伝統的CADはいま大きく進化しかつ多様化している（図-5参照）。

利用形態の多様化

1998年にFDA承認のマンモグラフィCADが出現して以降、伝統的CADはすべて「セカンダリー型CAD」であったが、ディープラーニング時代になり、2016年以降、医師がCADをどの順番でどのように読影に利用するかによって、4種類のAI-CADが開発され、順次商用化されつつある。



■ 図-4 畳み込みニューラルネットワーク (CNN) による悪性鑑別への応用



■ 図-5 ディープラーニングによるCAD (AI-CAD) の進化・多様化

インタラクティブCAD

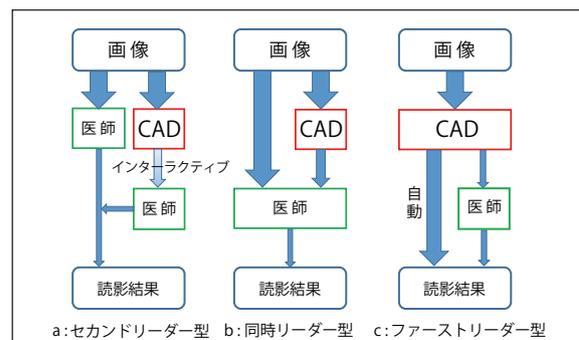
図-6 aは前述のセカンダリー型であるが、医師が画像上の気になる個所をクリックしたときにのみCADの結果が表示されるものとして、「インタラクティブ (Interactive) CAD」が派生し、マンモグラフィで商用化されている。これにより、セカンダリー型CADでも読影時間の増加なしに乳がんの検出率が良くなるという。

同時リーダー型CAD

「同時リーダー型 (Concurrent Reader) CAD」は、CADの結果（たとえば、検出マーカ）を読影の最初から見て参考にするものであり（図-6 b）、ワンステップであるため、セカンダリー型よりも読影時間の短縮が期待される。これは、1症例あたり大量の画像データの読影が必要な画像診断領域で、かつ検診での利用に期待される。少なくとも3つの同時リーダー型CADe (検出型) がFDAの認可を得ており、胸部CT用 (3次元画像) のシステムが筆頭である (2016年9月)。これは胸部画像中の血管陰影を抑制する部分と、種類の異なる結節陰影を検出する部分とで構成される。ほかには、3D乳房超音波画像のための同時CADシステムや、乳房トモシス (連続断層画像) を対象とした同時CADシステムの商用化例がある。

ファーストリーダー型CAD

「ファーストリーダー型 (First Reader) CAD」は、最初にCADが単独で読影処理を行い、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



■ 図-6 CADの代表的な3つの利用形態

いものを選定し、その画像の解析結果も提示するタイプのCADである(図-6 c)。これは半自動診断に相当する。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。現時点では、まだ、そのようなFDA承認の商品は市場に誕生してはいない。ただ、2020年に欧州の薬事承認と言えるCE規格認証を得たものがあり、これはマンモグラフィ乳がん検査で97%の正常例を除外するという。

高度化：導入後学習機能

学習機能を持ったディープラーニングの出現により、いよいよ期待が高まっているより高度化されたAI-CADに、臨床現場に導入後(市販後、すなわち商用展開後)にも、新しいデータの追加・再学習により、プログラムがアップデートされ、どんどん賢くなるCAD、すなわち「導入後(市販後)学習機能付CAD」がある。

米国ベンチャー企業から心臓MR画像を対象とした世界初のクラウドベースの心疾患診断支援AIシステムが、2017年1月にFDA承認を得ている。導入後、クラウド上に蓄えられた新規データによって再学習される仕組みにより、事後学習されたシステムのプログラム更新が、年に5回程度実施されているという。さらに高度化した導入後学習機能付の本格的なAI-CADシステムが、遠からず主流になるであろう。これにより、各医療施設の状況に応じた“独自のCAD”も夢ではなくなる。

利用目的の多様化

図-5に示すように、CADの目的も従来の検出型のCADeから多様化が起きており、FDA承認を取得し商用化に成功を収めているものが少なからず出現している。がんの良悪性鑑別を行い、悪性を提示する診断型のCADx、トリアージ(優先付)を目的としたトリアージ型のCADt、類似症例を検索して医師に提示することにより、CADの解析結果への確信度を高めるCAD、治療の予後予

測を画像解析により支援するラジオミクスCADやさらに遺伝子情報をも取り入れたラジオジェノミクスCADなどもある。FDAのCOVID-19対策として25日間という超短期間のレビューで認証されたという、心臓超音波検査に対するAIガイド画像システム“Caption Guidance”という商品があるが、COVID-19による心疾患の検査支援に有益と期待されている。これは2020年2月に新しい種類のCAD機器として承認されたもので、コンピュータ支援取得/最適化：CADa/oと呼ばれる。

CADe から CADx へ

1998年に初のCADがFDA承認されて以来、最近までのすべてのFDA承認のCADは、異常部位の検出を目的としたCADeに限定されていた。

2017年7月になり、ついにこの壁が破られた。これは米国ベンチャー企業が開発した商品で、乳房のMRIを対象とする診断支援型のCADxである。この乳房MRI用のCADxは、候補病変部位に対して「QI Score」という悪性を表すインデックスを提示するもので、類似画像検索・提示機能をも有している。

トリアージ型 CAD (CADt)

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を専門医に提示・警告するシステムがある。これは、従来のCADeあるいはCADxではなく異なる分類となるため、“トリアージ型CAD”と呼び、最近のFDAの分類ではCADtと表示される。

FDA初承認を得た商用化第一号は、救急(ER)患者のCT画像における主幹動脈閉塞(Large Vessel Occlusion: LVO)を特定するLVO脳卒中プラットフォームである(2018年2月)。脳卒中の疑いが確認されたことを脳卒中専門医(脳神経科医)に警告するもので、画像をこの医師のスマートフォンに直接送信する機能がある。95%を超えるケースで、専門医への自動通知で通知時間が平均52分短縮され、これにより治療時間までの

短縮が可能になったという。

さらに重要なのは、同プラットフォームは、AIソフトウェアの新技術追加支払いとして最初の保険 (Medicare) 適用が2020年9月に認められ、脳卒中が疑われる患者にアルゴリズムを使用するごとに最高1,040ドルが支払われるという。

その後、2019年以降、各社から乳がん、頭蓋内出血、肺塞栓症、頸椎骨折などに関するCADtが商用化されている。

眼底写真 AI-CAD：自律型

眼底写真は、緑内障のような目の病気のみならず、高血圧性網膜症（高血圧症、動脈硬化の程度判定）や糖尿病網膜症など、目の病気以外を由来とする病変の早期発見を目的として、検診でも多く用いられている。Googleの研究者らは、2018年3月に、心疾患予測について約28.4万人の患者から収集した眼底画像データでディープラーニングを用いて学習したモデル（テストには2つの独立したグループの計約1.3万人の患者の画像を使用）を発表している^{☆5}。この論文は非常に興味深く、予測できた因子に、喫煙の習慣の有無、性別、年齢、血圧値を含んでいる。このような眼底写真のみからは、これまで人間（＝眼科医）が決して予測できなかった診断情報まで予測でき、大きな注目を浴びている。

眼底写真の画像診断に対して、実用化という点では一歩先んじた興味ある以下のような事例が出てきた。

2018年4月、“IDx-DR”という名称のAI検査機器（DRは糖尿病網膜症のDiabetic Retinopathyの略）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カ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ドクターとか自動診断（Autonomous Diagnosis）と呼ばれるようになっていく。

ただ、注意が必要なのは、2020年6月、米国の2つの医学放射線学会が、このような自律型システムや導入後学習機能付CADの承認は、安全性の面から時期尚早として、連邦政府に承認のブレーキをかけるように要請書を提出している。

医用画像データについて

データが重要

すでに多くの事例が示しているように、ディープラーニングのCADへの影響力は計りしれないが、ディープラーニング型のAI-CADはData Drivenであるため、勝負は画像データの収集パワーに依存する要素が大である（図-7）。「量」が重要ではあるが、「Garbage in, Garbage out」と言われるように、その「質」も重要である。

医用画像の世界は、データ数は希薄と言われているが、最近の学会や企業からの報告事例を見ると、それでもまず千症例クラスでシステムを構築し、一万症例規模でシステム性能を向上させ、次に10万症例規模で実用化を目指す、本格的な世界展開の商用化には100万症例規模で、という大雑把な印象がある。

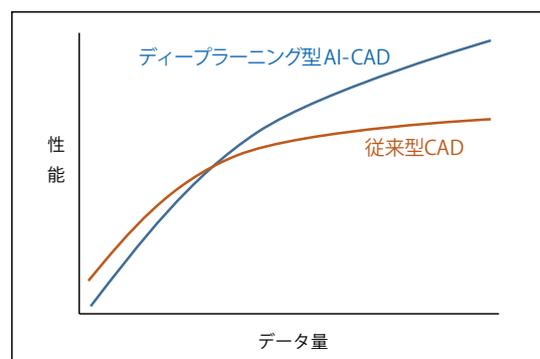


図-7 新旧CADにおける性能 vs データ量の変化

☆5 Poplin et al., Nat Biomed Eng. 2018;2:158-164

転移学習や GAN

自然画像（2次元画像）で学習させたディープラーニングモデルを、医用画像に応用するオーソドックスな「転移学習」（Transfer Learning）の技術は、医用画像へのディープラーニング応用では欠かせない技術である。また、医用画像は3次元画像であるケースも多く、各種の3次元医用画像で学習された専用のモデルも開発・公開され始め、その転移学習とさらなるファインチューニングによる手法に期待される。

また、「機械学習において、この10年間で最もおもしろいアイデア」（Facebook社のAI研究所LeCun所長）といわれる「敵対的生成ネットワーク」（Generative Adversarial Network: GAN）は、模擬病変作成などにも応用されており、データ不足の解消にも寄与すると結果も少なからず報告されている。

連合学習

倫理的な制約でデータを他施設から集約できないという悩みはつきものであるが、データ不足解消の手段の1つでもあり、プライバシー保護の観点からもそれを解決する手段として連合学習（Federated Learning）と呼ばれる手法が注目されている。これは、ディープラーニングのモデルのみを共有するものであり、匿名性・安全性が高く、患者データを共有しないので、プライバシーを保護できるという特徴が生まれる。

公共の画像データベース

公共の大規模な医用画像データベースも増えている。たとえば、米国国立衛生研究所（NIH）からは、約11万枚の胸部X線画像や1万症規模の体幹部CTデータベース（DeepLesion）、スタンフォード大学からは22万枚を超える胸部X線画像（CheXpert）の大規模データベースがすでに公開されている。また、多くのコンテストも開催され、共通の画像デー

タでソフトウェアの比較も可能である。

国内では、日本医学放射線学会、日本消化器内視鏡学会、日本病理学会、日本眼科学会、日本皮膚科学会、日本超音波医学会が、日本医療研究開発機構（AMED）の支援の下、AI診断支援システム開発も目指した画像などのデータの収集、データベースの構築が現在も行われており、そのデータ収集のクラウド基盤は、国立情報学研究所（NII）に設置された「医療ビッグデータ研究センター」が担っている。早期の公開が望まれる。

さらに、医療分野におけるデータ利活用に向けた国の施策として、匿名加工医療情報の円滑かつ公正な利活用の仕組みが整備され、いわゆる「次世代医療基盤法」が2018年5月に施行された。そして、2019年12月になってようやく初の事業者の認定がなされている。

COVID-19 画像診断への応用

昨今のコロナ禍において、COVID-19による肺炎画像診断に対して、すでに多くのAIによる研究論文がある。以下、いくつかの例を紹介する。

2020年3月、COVID-19のAI研究を念頭においた胸部X線画像データベースがウォータールー大学の研究者とカナダのDarwinAI社によって公開され、COVID-NetというDLモデルもGitHubで公開されている。ロシアの研究者らは千を超えるCT画像DBを作成している^{☆6}。

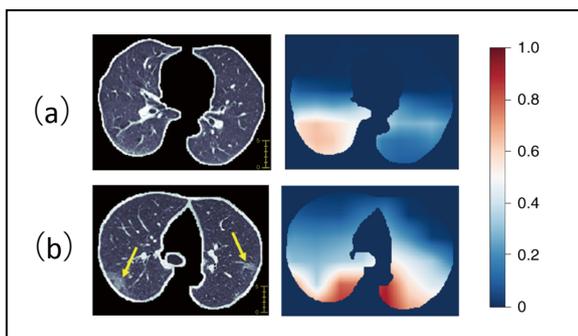
米国マウントサイナイ医科大学の研究者らは、肺のCT画像鑑別などを基にAI（従来の機械学習やCNN）による診断支援の可能性を報告している^{☆7}。図-8に一例を示す。2つのCNNモデル（CTのスライス同定モデルと診断モデル）から成るCOVID-19予測モデルは、人間の医師（経験10年の放射線科医）より検出精度が高い結果を示

☆6 <https://arxiv.org/abs/2005.06465>

☆7 Mei et al., Nat Med. 2020;6:1224-1228

した (83.6% vs 74.6%)。また、CT 画像データに加えて、患者の年齢・性別、症状の有無、血液検査結果をも機械学習させた総合的な予測モデルは、さらに精度が向上した。注目すべきは、放射線科医が CT 検査で陰性と診断したが、PCR 検査では陽性であった症例に対して、その 68% について AI モデルがすでに陽性と判定した点であり、早期の隔離措置を行うべき観点から重要であるという。

2020 年 6 月には、肺 CT 画像をから COVID-19 による肺炎を検出する国内 2 社からのプログラム医療機器に対して、本邦の薬機承認が行われている。COVID-19 疑い例に対して積極的に CT 検査を行い、その結果、CT 所見を契機に COVID-19 の診断に至った症例も少なくないと言われ、このような製品が臨床現場で有効に利用され、早期の正確な診断に寄与されることを期待したい。なお、厚生省 Web ページには、「新型コロナウイルス感染症又は関係する症状を対象とした医療機器の承認・認証状況」というサイトがあり、最新の情報が発信されている。



■ 図-8 ディープラーニング(CNN)へ入力されたCT画像(左)とその出力結果のヒートマップ表示(右:赤いほど、感染確率が高い)。(a)発熱と感染歴のある51歳女性で、CNNモデルでは右下葉に異常(右肺下端の白色部分)が認められたが、2名の放射線科医はこれをCT陰性としたケース。(b)感染歴があり発熱と湿性咳嗽(しっせいがいそう:痰のからんだ咳、ゴボゴボといった咳)を呈した52歳女性で、両方の肺の辺縁部のスリガラス状の不透明度(矢印)があり、医師もCNNも正しく検出した例。転載許諾を得て改変・引用[Mei et al., Nat Med. 2020;6:1227]

問題点

最近では、データの量のみではなく、異なる施設の画像データも含めないと、AIを取り扱った研究論文として認められないようになっている学会もある。画像データの量と質、撮像装置の機種や撮像条件の相違、性別、人種などなど、ディープラーニングに多くの依存パラメータがあり、これらの制約を取り除かないと、実臨床現場では使えない。管理された研究環境よりも実臨床はもっと複雑で多様であるからである。従来型CADのときでも同様であったが、実験室レベルでのいわゆる後ろ向き(Retrospective)評価のみならず、多様な状況下での実臨床現場でのAI-CADの有効性評価が重要である。米国の北米放射線学会(RSNA)からは、実臨床でのAI評価に役立つ9つの重要な考慮リストを提示している^{☆8}。

AI-CADの有効性を示す論文が多数増産されている昨今であるが、最近Nature誌に掲載された「AIの透明性と再現性」と題する論文では、そのような研究成果において、他の医療機関でも実証できるように、コンピュータコード、予測モデル、テスト設定、データなどを共有できるように要求している^{☆9}。

ディープラーニング問題点の1つに、導き出された結果の理由を説明できない、というブラックボックス性がある。そこで、何とか理由を導き出しホワイトボックス化しようとの試みがなされている。たとえば、Grad-CAM(Gradient-weighted Class Activation Mapping)と呼ばれる手法で、ディープラーニングが画像のどの部分に反応しているかを視覚的にヒートマップ(図-8参照)で表示する方法があり、CADが見当違いの箇所を異常の候補として指摘しているようなケースでは、一目瞭然に医師が判定でき、これは役立ちそうである。それでも、これだけではその根拠の十分な説明とは言いがたい。

また、学習データ不足は医療分野では常に付きも

☆8 Bluembe et al., Radiology. 2019;294:487-489

☆9 Haibe-Kains et al., Nature. 2020;586:E14-E16

のであり、深層学習の次なるイノベーションとして期待される正解データのないデータから特徴を抽出する「教師なし学習」や、あるいは学習データから正解を自ら作り出す「自己教師あり学習」などの開発が待たれる。

最終判断は医師

ディープラーニング技術のCADへの導入により、いまこのようにAI-CADとして、従来の技術的な限界を突破して、新しい画像支援診断の時代に突入している。AI-CADにより、従来のCADでは達成できなかったより診断の精度向上がなされ、多種疾病、多種モダリティにも対応でき、常時ONでシームレスに利用できるようになりつつある。また、アプリ形式に必要なAI-CADをダウンロードして(あるいはクラウド上で)利用する枠組みもできつつある。導入後もより賢く進化するAIが出現するのも時間の問題であろう。

AI-CADの高度化・多様化の時代には、「AIを理解して正しく使わないドクターは、AIを賢く利用するドクターに駆逐される」、と言われるような昨今であり、画像診断領域でのAIの正しい理解と積極的な活用が望まれる。ある研究者が述べている。「診断放射線医学におけるAIサポートへの移行は、自動運転車の採用のように進める必要がある。ゆっくりと慎重に、信頼を構築し、安全に重点をおいてシステムを改善する」と。

最後に、現状は『最終判断は医師によるものでなければならぬ』。2018年12月、厚生労働省からの通達「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の利用と医師法第17条の規程との関係について」では、「……診断、治療等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……**」と注意を喚起している。10年後にはもしかすると状況は変わっているかもしれないが、現在はこれが鉄則である。

本稿では、画像診断支援領域におけるAI導入の歴史から現状と課題、将来展望などについて、主に開発技術よりはむしろ臨床実利面から概説したが、診断に限らず治療や手術支援などAI応用領域は幅広い。より詳細な内容に興味ある読者は、参考文献1)～6)などを参照されたい。本研究の一部は、科研費・基盤研究(B)(19H03599)による。

参考文献

- 1) 藤田広志(監修, 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・事例一, オーム社(2020).
- 2) 藤田広志(監修), 福岡大輔(編): 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一, オーム社(2020).
- 3) 藤田広志(監修), 原 武史(編): 医療AIとディープラーニングシリーズ 2021-2022年版 標準 医用画像のためのディープラーニング実践編一, オーム社(2020).
- 4) 藤田広志(監修), 上杉正人, 平原大助, 齋藤静司(編): 医療AIとディープラーニングシリーズ Pythonによる医用画像処理入門, オーム社(2020).
- 5) 藤田広志(監修), 有村秀孝, 諸岡健一(編): 医療AIとディープラーニングシリーズ 放射線治療AIと外科治療AI, オーム社(2020).
- 6) 藤田広志(企画, 監修): 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(2020).

(2020年11月2日受付)

■藤田広志 hiroschi.fujita.gifu@gmail.com

岐阜大学工学部特任教授/名誉教授、藤田医科大学および中国・鄭州大学客員教授、電子情報通信学会・フェロー、医用画像情報学会・名誉会長、専門: 医用画像の計算機支援画像診断学。