

AI と人間が共創する医学・医療の未来

川上 英良

千葉大学大学院医学研究院・人工知能 (AI) 医学
理化学研究所・情報統合本部・先端データサイエンスプロジェクト・医療データ数理推論チーム

The Future of Medicine and Medical Care co-created with AI and Human

Eiryō Kawakami

Department of Artificial Intelligence Medicine, Graduate School of Medicine,

Chiba University

Medical Data Mathematical Reasoning Team, Advanced Data Science Project (ADSP),

RIKEN Information R&D and Strategy Headquarters

Key words: 人工知能, 機械学習, 層別化, 非侵襲モニタリング, human-in-the-loop

近年, 深層学習を始めとする人工知能 (AI) 技術が飛躍的な進歩を遂げ, 医療・ヘルスケア分野においても様々な応用例が報告されている。タスクによっては専門医を超える診断精度を示すことがある一方, AI 医学研究が進むにつれ, 実装に際して様々な課題も顕在化してきている。そこで本稿では医療・ヘルスケア分野での AI の導入状況や現状での課題について取り上げ, 今後の AI 医学研究の発展の方向性について議論する。

人工知能研究の歴史

現在私たちの社会の中で広く活用が進んでいる人工知能だが, その研究の始まりは 1950 年代に遡り, これまで平坦ではない道のりをたどってきた。1950 年代後半~1960 年代, コンピューターによる推論や検索が可能になり, 特定の課題に対し解を示せるようになったことから, 第 1 次 AI ブームとなった。しかし, 迷路やパズルのような単純な課題は解けるものの, 現実の課題は解決できないことから, ブームは去り冬の時代が訪れた。

1980 年代には大量のデータを与えて人工知能に推論させるプログラムが多数開発され, 第 2 次 AI ブー

ムが到来した。が, 人の手でコンピューターが認識可能な形でデータを記述する必要があるため, 扱えるデータ量が少ない, さらに知識の寄せ集めでは現実社会の複雑な課題に対処できないという問題により, 再び冬の時代が到来した。

2010 年以降, 現在も続いている第 3 次 AI ブームでは, 機械学習・特徴表現学習が実用化された。機械学習ではビッグデータから直接パターン, ルールの学習が可能となり, 幅広い分野で AI の活用が進んだ。医療・ヘルスケア分野においても疾患の診断・予測を中心に様々な用途での応用が進んでおり, いくつかの事例を紹介したい。

医学研究における AI の活用事例

AI の画像解析への応用事例として, 機械学習の一種である深層学習を用いた皮膚がんの分類の研究について紹介する。皮膚がんの臨床スクリーニングでは視覚的に診断を行うが, 皮膚病変の外観が多様なため, 画像を利用した自動分類は本来非常に困難である。Esteva ら (2017)¹は AI に約 13 万枚の画像のピクセル情報と疾患ラベルを学習させることで, 皮膚科医と同等以上 (90~95%) の診断精度を達成した。

Correspondence to Eiryō Kawakami, Department of Artificial Intelligence Medicine, Graduate School of Medicine, Chiba University, 1-8-1 Inohana, Chuo-ku, Chiba-shi, Chiba 260-8670, Japan

E-mail: eiryō.kawakami@chiba-u.jp

Journal Website (<https://www.nms.ac.jp/sh/jmanms/>)

また AI を用いた疾患の発症予測では、急性腎障害の事前予測研究が有名である²。約 70 万人分の入院患者の電子カルテの時系列データを学習することで、急性腎障害を発症の最大 48 時間前に予測することに成功した。

これらの AI による学習様式は End-to-End と表現され、画像などの入力データと病名などの正解ラベルのみを AI に与えることで、入力と正解の間の関係性を自動で学習させる。この学習様式では、途中で人の手が介在しないため、データさえ大量に与えることで、専門家も認識していなかったような入力データの特徴を抽出し、高い精度を達成することができる。実際にタスクが限定されている状況下では、臨床医を超える精度を見せる事例も報告されている³。

このような顕著な成果を上げている一方で、現在の AI 医学研究において下記のような課題も生じている。

①モデルがブラックボックス化し、予測の根拠が不明確になりがちである

②診断や既知の病態の検知が中心で、新たな発見に繋がるのが少ない

③多くの疾患データは発症後に病院で取られるため、予防につながらない

これらの課題への対処について次節で触れていきたい。

AI のホワイトボックス化

End-to-End での学習では上述の通り効率的に高い精度を達成できる一方で、途中の情報処理過程が人間に理解できないものになってしまうことも多く、人間では考えられないようなエラーをしたり、データからのバイアスを受けたりする事例も報告されている。

AI が陥りやすいエラーやバイアスはいくつかに類型化されるが、そのうちの一つがデータセットシフトと呼ばれ、学習に用いたデータセットと異なる性質を持つデータに対して予測が上手くいかなくなるというものである。医療・ヘルスケアでは、データセットシフトの要因は施設間差、測定機器の違い、年齢層の違い、と数多く存在する。例えば、ある年代のデータで学習された急性腎障害を予測するモデルが、その後急性腎障害の発症頻度が減るにつれて偽陽性率が上がってしまったという事例が報告されている⁴。

また、人の目では判別できない程度の僅かなノイズを加えることで、AI による判定が大きく影響を受けてしまう、Adversarial attack と呼ばれる問題も提起

されており⁵⁶、こちらは人間と AI の認識の違いに起因するものである。

これらの問題は、①のモデルのブラックボックス化、すなわち AI の情報処理過程が複雑化し、人間に理解できないものになっていることが大きな要因である。特に医療・ヘルスケア領域において、AI が起こしたエラーが人間に理解、説明できないということは極めて好ましくない。

このような課題に対し、AI の情報処理過程の説明性を高め、予測精度と解釈性を両立するための試みが「説明可能な AI (XAI)」である。2017 年に、米国の国防高等研究計画局 (DARPA) が XAI プログラムをスタートさせて以降、世界的に AI の説明性を高めるための研究が進んでいる。

XAI の具体的な手法としては、特徴量の説明・可視化、解釈可能なモデル開発、モデルの近似・簡略化が挙げられる。予測に重要な特徴を説明・可視化するために、ゲーム理論に基づき、予測における変数の寄与を個別にスコア化する SHAP (SHapley Additive exPlanations)⁷ や画像中の予測に重要だった領域をヒートマップで可視化する Grad-CAM (Gradient-weighted Class Activation Mapping)⁸ に加えて、複雑過ぎるモデルを臨床応用に向けて単純な決定木などのモデルで近似する⁹ といった方向性が提案されている。

AI に基づく疾患の再定義

ここまで扱った事例はいずれも AI を用いた疾患の診断・検知を目的としたものであったが、次に著者らが行った AI に基づく疾患の再定義の事例について紹介する。この研究では、卵巣腫瘍の良・悪性、早期・進行期といった性質を術前に計測される 30 項目程度の血液検査データのみに基づいて予測することを試みた¹⁰。良・悪性については、ランダムフォレストという機械学習を用いると 90% を超える高い精度で予測することができ、統計的手法であるロジスティック回帰 (86.7%) よりも高い値が得られた。

一方、早期・進行期の判定を同じランダムフォレストを用いて行くと、その精度は 70% 程度に留まった。早期がんと進行がんは播種、転移の有り無しで区別されるが、小さな播種や転移は検出されずに見落とされることもある。すなわち「正解」として与えたラベルが正しくない可能性がある。そこで、著者らは機械学習に良性・早期・進行期といった正解を与えずにパターンを学習する教師なし学習を行った。教師なしの

ランダムフォレストで症例同士の類似度を計算し、似た症例を近くに、パターンが異なる症例を遠くに配置すると、ある程度良性・早期・進行期ごとに固まった配置となった。良性腫瘍、進行がんの配置にはほぼ重なりがない一方、早期がんは良性腫瘍と重なるものと、進行がんに重なるものに分かれる。興味深いことに、良性腫瘍と重なる良性腫瘍（集団1）は5年以外にほとんど再発せず、死亡は0例、逆に進行がんと重なる早期がん（集団2）は5年以内に1/4程度が再発か死亡することが分かった。これは従来知られている卵巣がんの分類とは異なる早期がんの分類で、術前血液検査に基づいて早期がんを層別化できたと考えられる。疾患分類は経験的に設定されてきたものも多く、医師によって診断が異なる疾患も多くある。このような場合にデータに基づいて疾患をとらえ直すには、教師なし機械学習による層別化解析は有効なツールとなる。

疾患の非侵襲モニタリング

がんや生活習慣病を始めとする慢性疾患に対しては、ある時点で高精度に診断することよりも、疾患発症や病態進行の長期的な時間経過を理解・予測し、未然に重篤な変化を予防することが重要となる。そのためには、疾患の長期的モニタリングデータを蓄積するとともに、時系列データを扱うことができるような手法、特に日常的に低・非侵襲に行える計測手段が必要となる。従来の医療は、患者が来院して診察や検査が行われるが、計測頻度が低く、疾患発症前のデータが基本得られないという問題がある。

そこで近年活用が期待されているウェアラブルデバイスは、心拍数や活動量、血糖値といったバイタルデータを連続的に測定することができ、病院では見ることができない患者の日常活動データに基づいた疾患の検知を可能にするかもしれない。ウェアラブルデバイスを用いた研究としては、Apple Watchの心拍変動時系列データに基づくAIによる疾患の高精度な検出を行ったApple Heart Studyが有名である^{11,12}。約6万人のApple Watch（心拍変動）のデータから、糖尿病を85%、高コレステロール血症を74%、高血圧を81%の精度で予測している。心拍変動時系列データに対してAIを適用することで、人間も認識できないような時系列パターンを抽出し、疾患診断に活用できた一例と言える。このようなウェアラブルデバイスからAIによって抽出できる時系列パターンはDigital Biomarkerと呼ばれ、様々な疾患の早期検知や日常

ヘルスケアへの活用が期待されている¹³。また生活習慣病のモニタリングに留まらず、Digital Biomarkerに基づいた生活習慣への介入提案まで行う事例も報告されており¹⁴、今後Digital Biomarkerは慢性疾患の進行の流れを遅らせる、時にはより健康な状態に引き戻すための介入の鍵となるかもしれない。

AIと人間の共創に向けて

AIは大量のデータに基づいて、タスクによっては臨床医と同等以上の精度を出すことができるくらいパワフルなツールとして成長している。人間の知識量、処理能力には限界があるので、AIを活用することで医療の水準を高めることも可能になっている。一方で、AIの情報処理過程が見えにくくなりがちなので、医療・ヘルスケアにおいて、診断や予測の根拠を明確にし、誤った診断を下してしまった場合に適切な説明ができるようなモデルを構築するといった対応が今後さらに重要となるだろう。AIモデルの出力を人間がチェックし、適切でないときはフィードバックを行うことで、学習データやAIモデルを適切に更新することができる。AIと人間が、共に医学というドメインに正しい知識を集積するループを回していくことが、医療・ヘルスケア分野でのAIと人間が共創への道につながるだろう。

Conflict of Interest：開示すべき利益相反はなし。

文 献

1. Esteva A, Kuprel B, Novoa R, et al: Dermatologist-level classification of skin cancer with deep neural networks. *Nature* 2017; 542: 115-118.
2. Tomašev N, Glorot X, Rae J, et al: A clinically applicable approach to continuous prediction of future acute kidney injury. *Nature* 2019; 572: 116-119.
3. Ardila D, Kiraly A, Bharadwaj S, et al: End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. *Nature Medicine* 2019; 25: 954-961.
4. Davis S, Lasko T, Chen G, Siew ED, Matheny ME: Calibration drift in regression and machine learning models for acute kidney injury. *J Am Med Inform Assoc* 2017; 24: 1052-1061.
5. Goodfellow I, Shlens J, Szegedy C: Explaining and Harnessing Adversarial Examples. arXiv 2015. doi: <https://doi.org/10.48550/arXiv.1412.6572>
6. Finlayson S, Chung H, Kohane I, Beam A: Adversarial Attacks Against Medical Deep Learning Systems. arXiv 2018. doi: <https://doi.org/10.48550/arXiv.1804.05296>
7. Lundberg S, Lee S: A unified approach to interpreting model predictions. *Advances in neural*

- information processing systems 2017; 30.
8. Selvaraju R, Cogswell M, Das A, et al.: Grad-cam: Visual explanations from deep networks via gradient-based localization. In Proceedings of the IEEE international conference on computer vision 2017; 618–626.
 9. Fujii J, Aoyama S, Tezuka T, Kobayashi N, Kawakami E, Inaba Y: Prediction of Change in Pelvic Tilt After Total Hip Arthroplasty Using Machine Learning. J Arthroplasty 2022; 38: 2009–2016.e3.
 10. Kawakami E, Tabata J, Yanaihara N, et al.: Application of Artificial Intelligence for Preoperative Diagnostic and Prognostic Prediction in Epithelial Ovarian Cancer Based on Blood Biomarkers. Clin Cancer Res 2019; 25: 3006–3015.
 11. Turakhia M, Desai M, Hedlin H, et al.: Rationale and design of a large-scale, app-based study to identify cardiac arrhythmias using a smartwatch: The Apple Heart Study. Am Heart J 2019; 207: 66–75.
 12. Ballinger B, Hsieh J, Singh A, et al.: DeepHeart: Semi-Supervised Sequence Learning for Cardiovascular Risk Prediction. Thirty-Second AAAI Conference on Artificial Intelligence 2018; 32.
 13. Coravos A, Khozin S, Mandl K: Developing and adopting safe and effective digital biomarkers to improve patient outcomes. npj Digital Medicine 2019; 2: Article number: 14.
 14. Zeevi D, Korem T, Zmora N, et al.: Personalized Nutrition by Prediction of Glycemic Responses. Cell 2015; 163: 1079–1094.

(受付：2023年9月30日)

(受理：2023年10月13日)

日本医科大学医学会雑誌は、本論文に対して、クリエイティブ・コモンズ表示 4.0 国際 (CC BY NC ND) ライセンス (<https://creativecommons.org/licenses/by-nc-nd/4.0/>) を採用した。ライセンス採用後も、すべての論文の著作権については、日本医科大学医学会が保持するものとする。ライセンスが付与された論文については、非営利目的の場合、元の論文のクレジットを表示することを条件に、すべての者が、ダウンロード、二次使用、複製、再印刷、頒布を行うことができる。