

PRESS RELEASE

2019年12月18日
理化学研究所
日本医科大学
日本医療研究開発機構

がんの未知なる特徴を AI が発見

—がんの画像から、再発に関わる新たな知識を自力で獲得—

理化学研究所（理研）革新知能統合研究センター病理情報学チームの山本陽一朗チームリーダー、日本医科大学泌尿器科の木村剛准教授らの共同研究グループ^{*}は、医師の診断情報が付いていない病理画像から、がんに関わる知識を AI が自力で獲得する技術を開発し、がんの再発の診断精度を上げる新たな特徴を見つけることに成功しました。

本研究成果は、手術後の高精度ながんの再発予測法として、個々に合った治療選択に生かされるとともに、画像から新たな知識を獲得するための自動解析手法として役立ちます。さらに、ブラックボックスといわれている AI の解析根拠をひも解く一歩として、医療において安心して使用できる AI の実現に貢献すると期待できます。

今回、共同研究グループは、1枚あたり100億画素以上の前立腺病理画像から、AI が画像上のがんの特徴を、人に教わることなく自動で取得し、それを人間が理解できる情報として出力する技術の開発に成功しました。AI が見つけた要素には、今日までに世界中で使われているがんの診断基準のほか、専門家も気づいていなかったがん領域以外の部位の特徴が含まれていました。これらの要素の再発予測性能を確かめるために、三つの大学病院の15,000枚以上の病理画像（AI 学習用の分割画像にすると約960億枚に相当）で検証したところ、現在の診断基準よりも高い精度で再発予測ができました。加えて、病理医の診断と合わせて使うことで、予測精度をさらに上げることができました。

本研究は、英国のオンライン科学雑誌『*Nature Communications*』（12月18日付）に掲載されました。

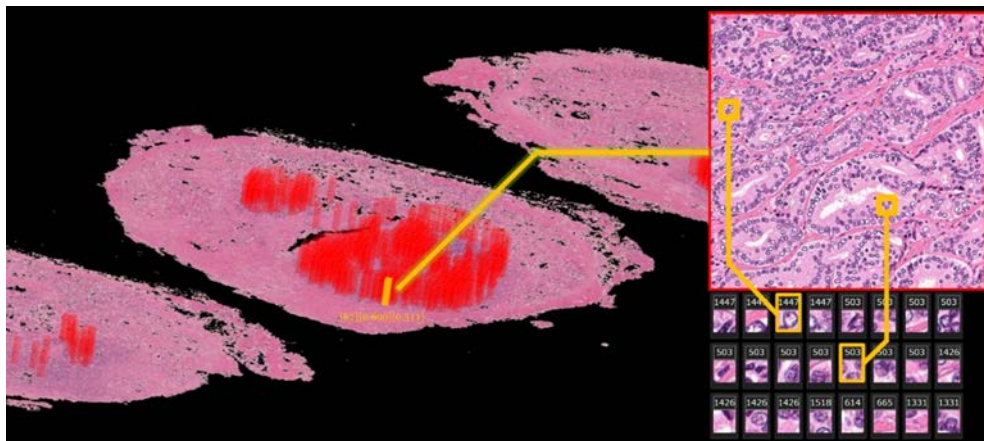


図 人に教えられることなく、がんの特徴を AI が自動で発見（3D 病理画像）

※共同研究グループ

理化学研究所 革新知能統合研究センター

目的指向基盤技術研究グループ

病理情報学チーム

チームリーダー

技師

山本 陽一朗 (やまもと よういちろう)

沼田 康志 (ぬまた やすし)

森川 啓 (もりかわ ひろむ)

堤 光太郎 (つつみ こうたろう)

客員主管研究員

福本 学 (ふくもと まなぶ)

客員研究員

赤塚 純 (あかつか じゅん)

(日本医科大学 泌尿器科)

遺伝統計学チーム

チームリーダー

田宮 元 (たみや げん)

(東北大学 東北メディカル・メガバンク機構)

研究員

植木 優夫 (うえき まさお)

副センター長

上田 修功 (うえだ なおのり)

日本医科大学

泌尿器科

教授

近藤 幸尋 (こんどう ゆきひろ)

准教授

木村 剛 (きむら ごう)

解析人体病理学

教授

清水 章 (しみず あきら)

愛知医科大学病院 病理診断科

教授

都築 豊徳 (つづき とよのり)

助教

高原 大志 (たかはら たいし)

助教

露木 琢司 (つゆき たくじ)

聖マリアンナ医科大学 腎泌尿器外科

講師

中澤 龍斗 (なかざわ りゅうと)

北里大学 北里研究所病院 病理診断科

教授

前田 一郎 (まえだ いちろう)

信州大学 医学部 病理組織学教室

教授

菅野 祐幸 (かんの ひろゆき)

社会医療法人栗山会 飯田病院

副病院長

土屋 眞一 (つちや しんいち)

※研究支援

本研究は、日本医療研究開発機構 (AMED) 「Medical Arts の創成に関する研究 (外科、がん、看護、リハビリ等) の新たな医療技術やソフトウェアの開発」の研究課題「機序の異なる人工知能の多重解析による癌コンパニオン診断システムの開発 (研究代表者: 山本陽一朗)」による支援を受けました。

1. 背景

人工知能（Artificial Intelligence, AI）の技術は、翻訳から自動運転に至るまでさまざまな分野で利用されています。そして近年では、AI の持つ高い可能性を医療に応用するための研究が盛んに行われています。

しかし、現在の AI 技術の主流であるディープラーニング（深層学習）^[1]では、学習にビッグデータを必要とするため、医師の診断情報が付いた大量の医療画像をどのように集めるかが、実用化に向けた課題となっていました。

また、AI における解析根拠はブラックボックスだといわれています。というのは、AI の解析根拠はニューラルネットワーク^[2]上の無数の重みとして保存されることから、人間は数学的なメカニズムは理解できても、AI による解析根拠を直接的に理解することが困難であるためです。医療への応用には、AI の解析根拠が重要視されており、現存する医学知識を上回る新知見の獲得のためにも、病理画像のように豊富な情報を含むデータから、機械学習^[3]を通して「人間が理解できる情報」を自動で取得する技術が求められていました。

2. 研究手法と成果

共同研究グループは、複数のディープラーニングと非階層型クラスタリング^[4]を用いることで、病理画像から人間が理解できる情報を自動で取得する新たな AI 技術の開発に成功しました。今までは、医師が教えた診断を AI が学習する「教師あり学習^[5]」と呼ばれる手法が医療分野では主に使用されてきましたが、教師以上の分類はできないという限界がありました。本研究では医師の診断を必要としない「教師なし学習^[6]」により獲得した特徴を、人が理解できるように変換し、再発期間のみを用いた最適な重み付けを AI に行わせることで、これまで不可能であったがんの未知なる情報の獲得を目指しました（図 1）。

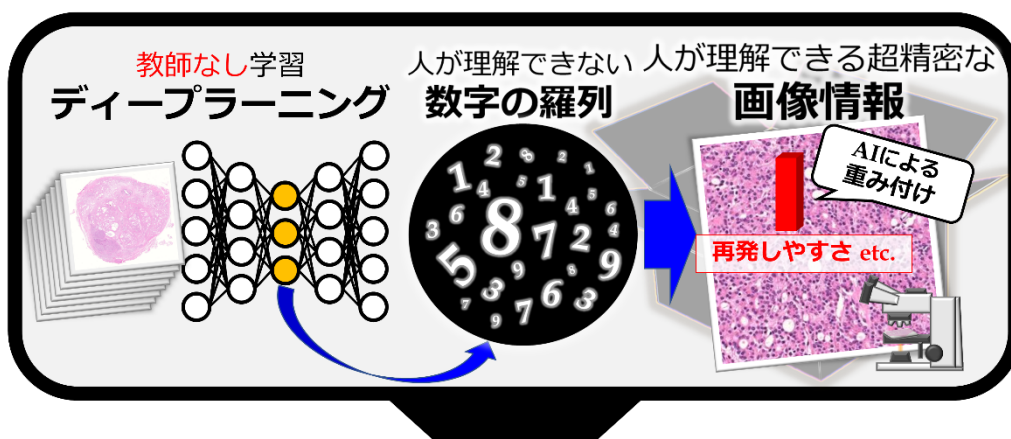


図 1 手法の基本原理

医師の診断情報が付いていない画像に教師なし学習のディープラーニングを適用し、背後にあるデータ構造を捉えた特徴を獲得する。AI が獲得した特徴は人が直接的に理解できない数字の羅列となっているため、数値の規則性を手がかりに、新しく開発した技術を用いて人間が理解できる高解像度の画像情報に変換する。その後、これらの画像に対して AI が最適な重み付け（意味付け）を行う。

この新しい技術を、医師の診断情報が付いていない 100 億画素を超える全埋埋・全割^[7]した前立腺の病理画像（AI 学習用の分割画像にすると、合計約 11 億枚に相当）に対して適用したところ、病理画像と予後情報のみから、詳細に分類されたがんの情報を AI に自動で抽出させることに成功しました（図 2）。この AI が作成した分類には、現在世界中で使用されているがん分類（グリソンスコア^[8]）が含まれており、さらに、これまで専門家も気づいていなかった「がん領域以外の間質^[9]の変化」も、がんの再発の診断精度を上げる特徴として読み取られました。

なお、AI により見つけられた病理学的特徴は論文の中で、AI が作成した初めての病理画像アトラス^[10]として閲覧できます。

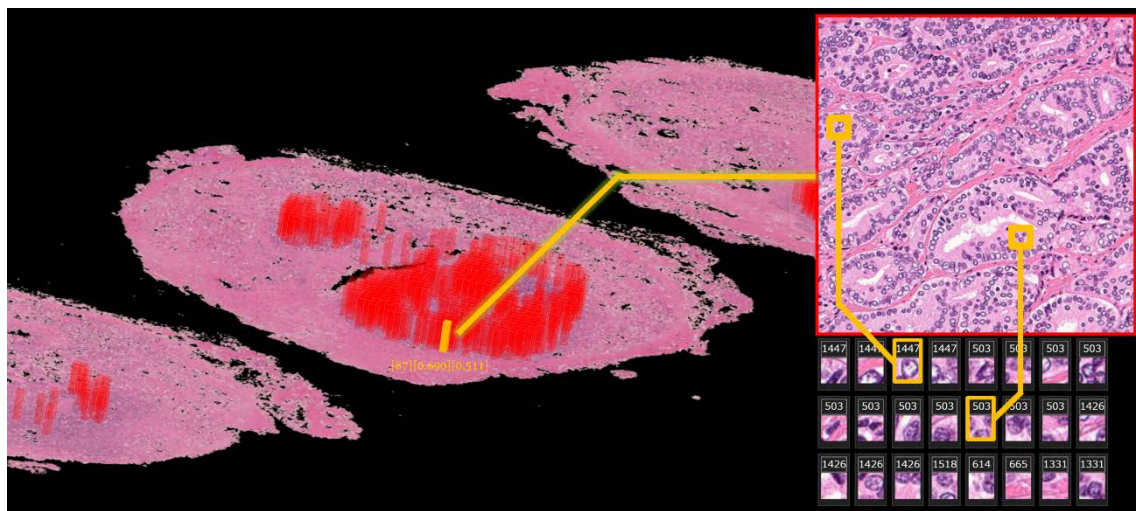


図 2 前立腺病理標本の連続切片に対する 3D 病理画像

左の 3D 病理画像上における赤い色の領域は、AI が人に教えられることなく、がんの特徴を自動で発見した部分。3D 病理画像上の黄色の領域に対応するがんの特徴が、AI によって弱拡大画像（右上）と強拡大画像（右下）として提示されている。

次に、AI が見つけたこれらのがんの特徴が再発予測に役立つかを確認するため、日本医科大学病院の 20 年間分の 13,188 枚の前立腺の病理画像（AI 学習用の分割画像にすると約 860 億枚に相当）を用いて、がんの予後予測の検証を行いました。その結果、現在世界中で使用されている前立腺がんの診断基準（ $AUC^{[11]} = 0.744$ ）よりも高い精度（ $AUC = 0.820$ ）で再発予測ができることが分かりました（図 3）。

さらに、日本医科大学病院の症例だけを用いて AI に学習させた結果が、聖マリアンナ医科大学病院と愛知医科大学病院においても利用できるかどうかを調べました。これら二つの大学病院の 2,276 枚の前立腺の病理画像（AI 学習用の分割画像にすると約 100 億枚に相当）に対して検証したところ、日本医科大学における予測精度とほぼ同等の再発予測ができることが分かりました（ $AUC = 0.845$ ）。これは、今回開発された技術により、AI が病院や大学といった施設や地域を越えて、一般化された情報を学習したことを示しています。

最後に、AIが見つけた特徴と病理医の診断を組み合わせることで再発予測をしたところ、それぞれが単独で予測するよりも、さらに予測精度を上げることができました（図3）。この結果は、AIと人間は病理画像の解析に対して得意とする点が異なり、お互いに補い合うことで精度を上げることができることを示しています。

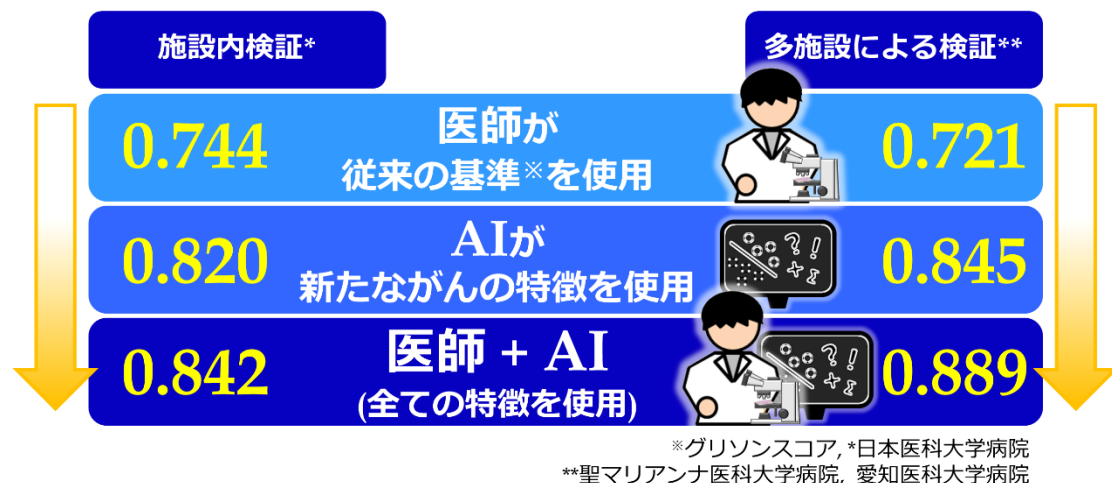


図3 前立腺がんの再発に係る予測精度の比較（AUC）

図中の数値は、前立腺がんにおける術後1年の生化学的な再発に対するAUCを示す。左が施設内検証（日本医科大学）、右が多施設による外的妥当性検証（聖マリアンナ医科大学病院と愛知医科大学病院）のデータ。

3. 今後の期待

医療においてAIを安心して使用するためには、医師が理解可能な根拠を示すことができる技術が不可欠です。さらに、情報量に富んだ画像から、人間が理解できる情報を引き出すことで、既存の基準を超えた新たな知識の獲得が可能になります。

本研究成果は、手術後の高精度ながんの再発予測法として、個々に合った治療選択に貢献するとともに、画像から新たな知識を獲得するための自動解析手法として役立ちます。さらに、ブラックボックスといわれているAIの解析根拠をひも解く一歩として、医療において安心して使用できるAIの実現に貢献すると期待できます。

4. 論文情報

<タイトル>

Automated acquisition of explainable knowledge from unannotated histopathology images

<著者名>

Yoichiro Yamamoto, Toyonori Tsuzuki, Jun Akatsuka, Masao Ueki, Hiromu Morikawa, Yasushi Numata, Taishi Takahara, Takuji Tsuyuki, Kotaro Tsutsumi, Ryuto Nakazawa, Akira Shimizu, Ichiro Maeda, Shinichi Tsuchiya, Hiroyuki Kanno, Yukihiro Kondo, Manabu Fukumoto, Gen Tamiya, Naonori Ueda, and Go Kimura

<雑誌>

Nature Communications

<DOI>

10.1038/s41467-019-13647-8

5. 補足説明**[1] ディープラーニング（深層学習）**

機械学習の手法の一つ。人工知能（AI）の急速な発展を支える技術であり、多層のニューラルネットワークによる機械学習の手法。

[2] ニューラルネットワーク

脳内の神経細胞（ニューロン）のネットワークと類似した構造を持つ数理モデルのこと。

[3] 機械学習

明示的にプログラムしなくても学習する能力をコンピュータに与える研究分野であり、データをもとにして予測や判断を行うモデルを抽出することができる。

[4] 非階層型クラスタリング

データをいくつかの階層構造を持たないグループに分割する機械学習の手法の一つ。

[5] 教師あり学習

入力データとペアとなる正解データを用いて学習し、予測モデルを構築する機械学習の手法の一つ。分類、回帰などが教師あり学習に含まれる。

[6] 教師なし学習

入力データのみが存在しペアとなる正解データが存在しない状況で、データの背後にある未知の構造を取り出す機械学習の手法の一つ。クラスタリングなどが教師なし学習に含まれる。

[7] 全包埋・全割

摘出された臓器全体をパラフィン等で固め、約 3～5mm 幅で全領域において断面の病理スライドを作成すること。通常の病理標本と比べ各段に多い情報量を持つ。

[8] グリソンスコア

前立腺組織を顕微鏡で検査し、がんの悪性度を評価するために用いられる指標。世界中の病院で使用されている。最も多い病変のタイプと 2 番目に多い病変のタイプを判定し、その数値の合計で 2～10 の 9 段階に分類する。グリソンスコアが高いほど、

がんの悪性度が高いと評価され、その後の治療方針に影響する重要な要素の一つとなる。原案は1966年に米国のドナルド・グリソンにより提唱された。

[9] 間質

臓器の機能の特徴づけている実質を支える部分。本研究では、前立腺がんや、その発生源地である前立腺上皮細胞以外の、線維芽細胞を含む結合組織を指している。

[10] 病理画像アトラス

病理画像の意味が提示された図表集のこと。論文中の Supplementary Figure 7-36 に、AIにより見いだされた計30枚の再発に関わる病理画像が掲載されている。

[11] AUC

検査などの性能を表す際に使用されるグラフの一つである ROC (Receiver Operating Characteristic) 曲線を作成したときの、グラフ下部の面積のこと。0から1までの値をとり、値が1に近いほど判別能が高い。AUCは Area Under the Curve の略。

6. 発表者・機関窓口

理化学研究所 革新知能統合研究センター

目的指向基盤技術研究グループ 病理情報学チーム

チームリーダー 山本 陽一朗 (やまもと よういちろう)

日本医科大学 泌尿器科

准教授 木村 剛 (きむら ごう)



山本 陽一朗



木村 剛



<機関窓口>

理化学研究所 広報室 報道担当
TEL : 048-467-9272 FAX : 048-462-4715

学校法人日本医科大学 総務部 広報課
TEL : 03-3822-2131 FAX : 03-3822-2822
E-mail : hp-kouhouka [at] nms.ac.jp

<本事業に関する問い合わせ先>

日本医療研究開発機構
臨床研究・治験基盤事業部 臨床研究課 ICT 基盤研究グループ
TEL : 03-6870-2229 FAX : 03-6870-2246
E-mail : rinsho-ict[at]amed.go.jp

※上記の[at]は@に置き換えてください。
