

技術動向レポート

肺がん検診における AI（人工知能）実用化に向けた 福島県立医科大学との共同研究

情報通信研究部		
首席コンサルタント	チーフコンサルタント	コンサルタント
永田 毅	佐野 碧	岩淵 耕平
社会政策コンサルティング部		
課長	調査役	
山崎 学	井高 貴之	

1. はじめに

肺がんは世界的規模で増加傾向にある疾患で⁽¹⁾死亡率も高く、この疾患の制御は、全世界の喫緊の課題であり、特に肺がんは早期の場合には自覚症状を伴わないために発見が遅れる事が課題となっており、我が国も同様の課題を抱えている(図表1)。早期発見に向け、現在では、CTやMRIの普及により非侵襲な手段による精密検査が可能であるものの、時間もコストもかかることから検診率が低いのが現状であり、広く普及している胸部X線撮影による肺がん検診は大きな意義を持つ。

我が国では健康増進法に基づいて集団健診、施設健診などが行政(自治体)主導で行われており、健康維持管理に役立っている。成人に対する健診における肺がん検診を目的とした胸部X線画像の読影においては、「肺がん検診手引き」(日本肺癌学会編)にて処理フローが規定されており、読影担当医師2名で1枚の胸部レントゲンフィルム(デジタル画像)をダブルチェック方式で読影する(二重読影、一次読影)。この段階で

異常と指摘された症例に対しては比較読影(二次読影)で最終評価が行われて、精度の高い判断がなされている。

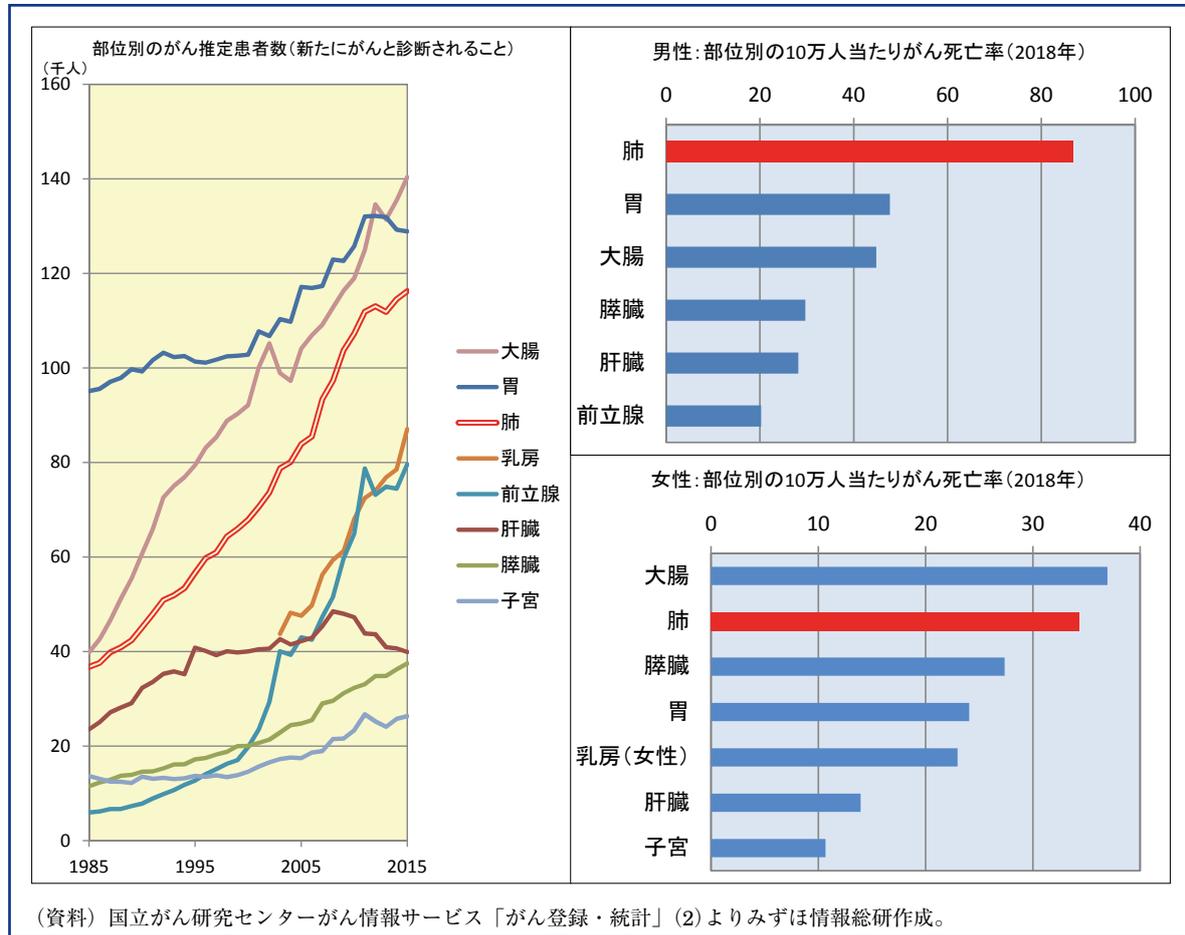
読影作業に従事する医師の負担は決して軽いものではない。最近では政府主導で働き方改革が議論されており、医師に関してもその働き方を改善するための議論がなされて久しいが、医療に対する国民のニーズは年々高まっており、医師の負担が一向に軽減していない問題が存在する。また、医師の長時間にわたる読影作業は、作業効率の低下を招き、陰影の見過ごしにつながる可能性があるため、医療安全管理上も課題とされる。

上記のような問題を鑑み、当社と公立大学法人福島県立医科大学は、2019年より集団健診の画像読影作業におけるAIによる異常検知の実用化を目的とした共同研究に取り組んでいる⁽³⁾(図表2)。

肺がんの最終的な診断は医師が行うため、我々は主に以下の目的で本共同研究を実施している。

- ① 一次読影において、AIの判断を参考に医師が診断を行うことで、診断の質を落とすこ

図表1 日本におけるがん推定罹患数と死亡率



となく、医師の負担軽減、医師の労働時間の短縮を図る。

- ② 医師の比重を、一次読影から二次読影業務にシフトすることで、二次読影の質の向上を実現する。
- ③ 肺がん検診を遠隔で支援することを検討し、専門性の高い読影医の偏在による地方格差の解消を実現する。

本共同研究は現在進行中であり、現在は、図表2における最初のステップである、過去の検診症例・臨床症例での精度評価を行っている。本稿ではまず、国内の肺がんを含むがん検診全般における AI の活用状況について概観した上で、本共同研究の現状について報告する。

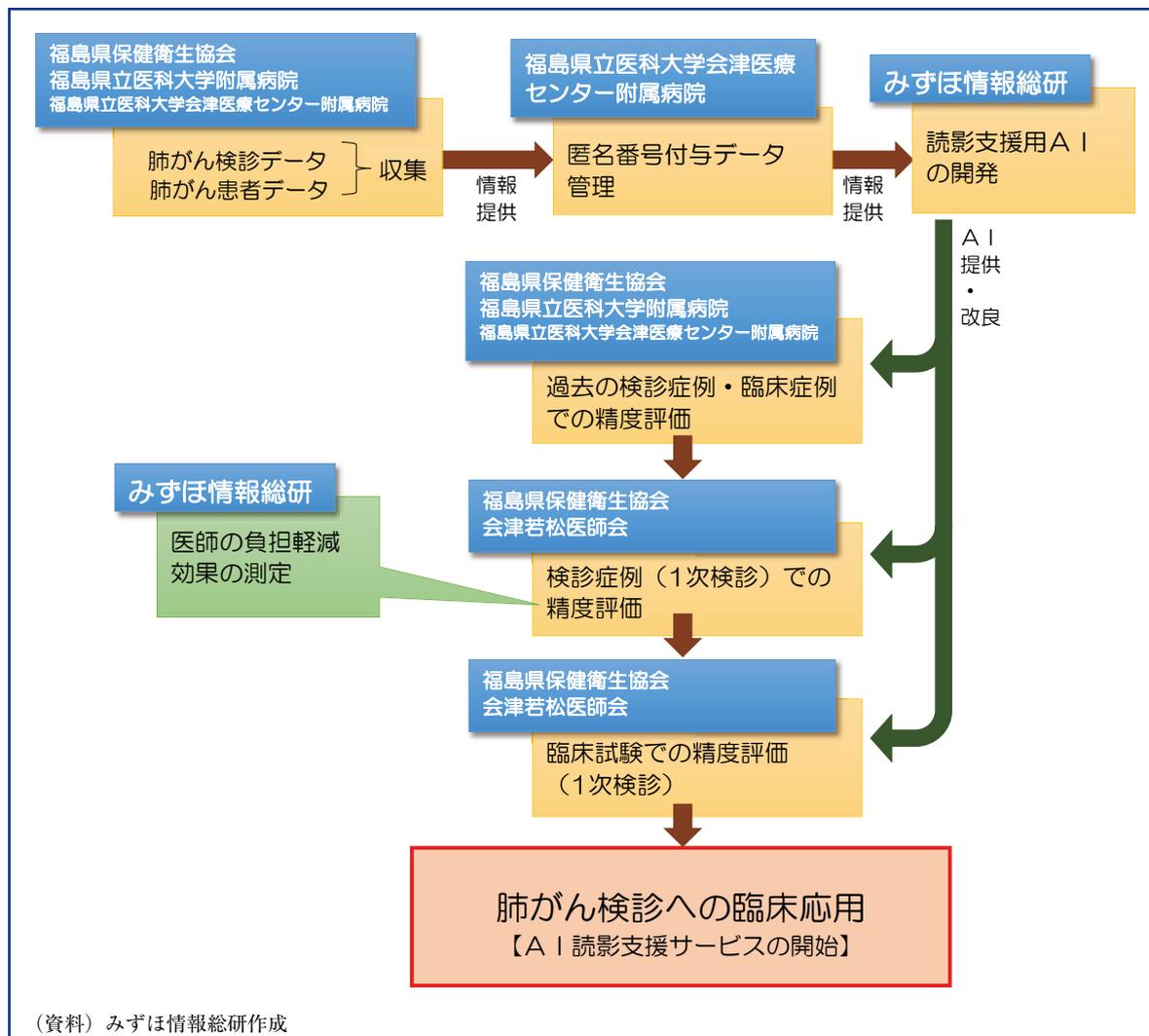
2. がん検診全般における AI の研究・活用状況

ここで、国内のがん検診における AI の活用状況について、現状を概観しておきたい。

理化学研究所と国立がん研究センターの共同研究グループは、ディープラーニングとデータ拡張技術による早期胃がんの高精度な自動検出法を確立した⁽⁴⁾。陽性的中率(AI が陽性と判断した画像中、実際に陽性であった割合)は93.4%、陰性的中率(AI が陰性と判断した画像中、実際に陰性であった割合)は83.6%を達成している。

オリンパス株式会社は、AI 搭載の内視鏡画像診断支援ソフトウェア「EndoBRAIN」の国内

図表2 共同研究の概念図



販売を2019年3月から開始した⁽⁵⁾。EndoBRAINは、製造販売承認を受けた内視鏡関連の国内初のAI搭載ソフトウェアであり、腫瘍性・非腫瘍性ポリープを自動的に判別し診断を支援する。機械学習にはSVM（サポートベクターマシン、機械学習手法の一種）を用いており、製造販売承認に当たって6万9,142件の教師データを使用し、その精度は、感度（陽性のうち検査で陽性とされた割合）が97%、正診率（データが正しく診断された割合）98%を達成している。国立研究開発法人日本医療研究開発機構（AMED）の研究事

業に採択され、昭和大学、名古屋大学、サイバネットシステム株式会社が研究開発を行い⁽⁶⁾、サイバネットシステム株式会社が2018年12月に薬機法の製造販売承認を取得した。

日本電気株式会社と国立がん研究センターの共同研究グループは、大腸潰瘍性ポリープを、ディープラーニングを基礎としたAIを用いて内視鏡検査時に即時に発見するシステムの開発に成功した⁽⁷⁾。隆起型の病変に関しては、感度（陽性のうち検査で陽性とされた割合）98%、特異度（陰性のうち検査で陰性とされた割合）99%

を達成している。

がん研究会有明病院と株式会社 AI メディカルサービスの共同研究グループは、AI により、内視鏡検査時のリアルタイム胃がん検出支援システムを構築することに成功し、全胃がんの 94.1% を検出することに成功した⁽⁸⁾。

株式会社日立製作所は肺がん検診における 3 次元 CT 像に対して、ルールベース手法とディープラーニングをハイブリッドさせた AI を適用し、充実性結節病変の検出率 93.4% を達成した⁽⁹⁾。

コニカミノルタ株式会社は、医用画像 AI ベンチャーのパイオニアである Enlitic, Inc.、丸紅株式会社と、共同開発契約を締結し、胸部単純 X 線における異常部位検知を行う AI の開発を進めると発表した⁽¹⁰⁾。

株式会社エヌ・ティ・ティ・データは 2019 年 4 月にインドで AI を用いた胸部単純 X 線画像の遠隔読影サービスを始めた⁽¹¹⁾。インドのスタートアップであるディープテック(プネー市)と連携し、複数の医療機関向けに、画像を AI で分析して各疾病リスクを確認、診断したりレポートを作成・提供している。

このように、三大がん(胃がん、肺がん、結腸がん)に対して、様々な機関が AI の活用を試みしており、その多くが読影医に迫る精度を報告していることから、今後もこうした流れは拡大していくと考えられる。

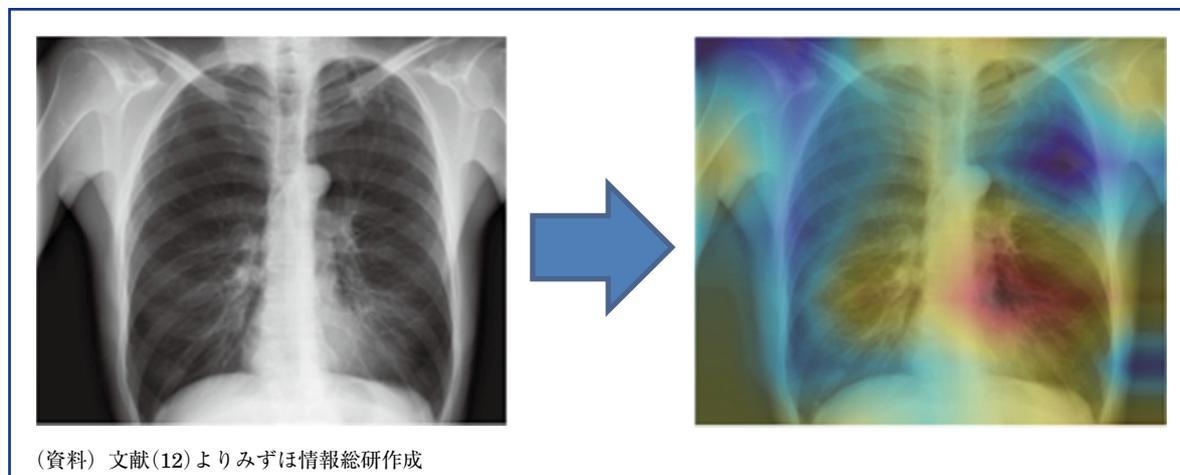
本共同研究が対象とする肺がんに対する AI の活用は、従来は三次元 CT 像を対象とした研究がほとんどであった。X 線画像は、二次元画像であるため、三次元 CT 像よりも情報が少なく、正常と異常の見分けがつきにくいいため、難易度が高かったためと思われる。しかし、近年の画像解析技術および AI の発展により、この問題を克服できると我々は考えている。

3. AI を活用した X 線画像からの異常検知 (過去症例を用いた後ろ向き精度評価)

本研究を始めるに当たり、まず先行研究についての論文調査を行い、精度の検証を行った。論文調査の結果、胸部 X 線画像を対象とし、AI で肺がんを検知する研究は見つけることができなかったが、肺疾患(結節、結核、肺炎、気胸等)を検知する研究がいくつか見つかった。そこで、X 線画像における見え方が比較的肺がんに近い結節を対象として、論文を調査した結果、複数の先行研究の中から、NIH (National Institutes of Health、アメリカ国立衛生研究所) の Chest X-ray データセット(10万枚)を教師データとして利用し、ディープラーニングの一種である DenseNet を改良した手法で学習した研究⁽¹²⁾が最も良い精度(AUC=0.78)を示していたため、当該論文⁽¹²⁾を精度検証における先行研究の対象として選定した。なお、AUC とは、ROC 曲線(横軸に偽陽性率(陰性データを陽性と予測した率)、縦軸に真陽性率(陽性データを陽性と予測した率)を取ったもの)の面積であり、0 から 1 の値を取り、1 に近いほど精度が高いことを示す。さらに、この手法は、図表 3 に示すように、異常の尤度を可視化することが可能であり、AI が異常と判断した箇所について目視で確認することが可能である。

我々はまず、文献⁽¹²⁾のネットワーク構造に相当するプログラムを開発し、結節を対象として NIH の Chest X-ray データセット(10万枚)に対して学習を行い、論文で示された精度が得られることを確認した。そして次に、図表 4 の 2 つの学習ケースについて、肺がん検知の精度比較を行った。検診データについては、両ケースとも ImageNet⁽¹³⁾ による training 結果を初期値とし、最終的な学習は福島保健衛生協会から提供を受けたデータ(正常 401 枚 + 肺がん疑い 452 枚

図表3 改良 DenseNet による異常検知例(肺炎)⁽¹²⁾



図表4 学習ケース

ケース	学習ステップ
A	初期値 ImageNet ⇒ NIH データ (10万枚)による pre-training ⇒ 福島保健衛生協会のデータ (853枚)による学習
B	初期値 ImageNet ⇒ 福島保健衛生協会のデータ (853枚)データによる学習

(資料) みずほ情報総研作成

図表5 交差検定による精度検証結果

ケース	感度 (SEN)	特異度 (SPE)	AUC
A	0.75	0.60	0.74
B	0.73	0.75	0.80

感度：陽性のうち検査で陽性とされた割合 特異度：陰性のうち検査で陰性とされた割合
 AUC：ROC 曲線における面積。ROC 曲線とは横軸に偽陽性率(陰性データを陽性と予測した率)、縦軸に真陽性率(陽性データを陽性と予測した率)を取ったもの。0から1の値を取り1に近いほど精度が高い。

(資料) みずほ情報総研作成

= 853枚)で行ったが、ケース A は間に NIH データ (10万枚)による pre-training を行った。なお、いずれのケースも、X 線画像に対しては、輝度分布の平均と分散を一致させる標準化を行い、画像間の輝度のばらつきを吸収する前処理を行った。

図表5に交差検定における精度検証の結果に

ついて示す。ケース B の方が高い精度が得られており、先行研究の結節の結果(AUC=0.78)に匹敵する精度が得られている。また NIH データによる pre-training は、今回のケースでは逆効果であることが分かる。一般的に、ディープラーニングで高い精度を得るためには、大量の教師データが必要であり、NIH の Chest X-ray

データセット(10万枚)は、本研究のデータ不足を補う目的で利用したものであるが、NIHのデータには、新旧様々な機器で撮影したデータが含まれており、撮影環境や受診者の姿勢が一樣でないため、同一機種・同一環境で撮影された福島保険衛生協会のデータと相性が良くなかった可能性がある。よって、今後実用化に向けて、機種・撮影環境・受診者の姿勢等のバラつきを前処理で吸収し、一定の精度を保証することが課題の1つになると考えている。

また、今後の精度向上に向け、質の高い国内の検診データをさらに収集し、教師データの母数を増やしていくことが有効と考えており、既に、福島県立医科大学付属病院、福島県立医科大学会津医療センター付属病院から提供を受けるなど、データ拡充を進めている。一方で、肺がん検診において肺がんと診断される比率は1%を大きく下回るため、陰性データと比較して陽性データは圧倒的に母数が少ないのが実情である。よって、今後は、陽性データ不足を補うデータ拡張技術も必要になると考えられる。

まとめると、今後、我々は本研究の実用化にむけて、

- ・より肺がん検知に適したディープラーニングのネットワーク構造の検討。
 - ・質の高い国内の検診データをさらに収集することによる、教師データの拡充。
 - ・機種・撮影環境・受診者の姿勢等の違いを前処理により吸収する技術の検討。
 - ・データを幾何学的に変換し教師データの母数を水増しするデータ拡張技術の検討
- 等の様々な工夫を組み合わせることで精度向上を果たしていく予定である。

4. おわりに

我々は、AIによる肺がん検知について、従来適用されていなかった胸部X線画像を用いてそ

の精度を検証し、他の肺疾患(結節)と同程度の精度が得られることを確認した。実施方法・算出方法によってかなりの差があるが、胸部X線の肺がん検診における読影医の感度(陽性のうち検査で陽性とされた割合)は0.63~0.88、特異度(陰性のうち検査で陰性とされた割合)は0.95~0.99とされている⁽¹⁴⁾。我々の精度(図表5)と比較すると、感度は読影医に近い精度が実現できていることが分かる。特異度は読影医と比較すると精度が低い、ケースBの結果(特異度0.75)は、正常データの75%は正常と判断できることを示しており、今後さらなる精度向上を達成できれば、医師の負担軽減に大きく貢献できる可能性があると考えている。

本研究は、肺がん検診において広く利用されている胸部X線検査においてAIを活用する点に大きな意義があり、本研究の実用化を果たすことで、肺がん検診の有効性の向上および受診率の向上につなげていきたいと考えている。さらに、肺がん検診におけるAIの有用性を福島県発の業績として世界へ発信し、他地域での活用も促すことで、我が国の肺がんの早期発見および治療に少しでも貢献していきたいと考えている。

謝辞

本共同研究では、福島県立医科大学 呼吸器外科の鈴木 弘行教授、樋口 光徳准教授に多大なご協力とご指導をいただきました。ここに感謝の意を表します。また、貴重な検診データをご提供いただいた福島保健衛生協会、福島県立医科大学付属病院、福島県立医科大学会津医療センター付属病院に心から感謝いたします。

注

(1) <http://www.ciesin.org/IC/who/MortalityDatabase.html>

(2) https://ganjoho.jp/reg_stat/statistics/stat/index.

- html
- (3) <https://www.mizuho-ir.co.jp/company/release/2019/medicheck0729.html>
 - (4) Sakai, Yoshimasa, et al. "Automatic detection of early gastric cancer in endoscopic images using a transferring convolutional neural network." 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE, 2018.
 - (5) <https://www.olympus.co.jp/news/2019/nr01157.html>
 - (6) Misawa, Masashi, et al. "Accuracy of computer-aided diagnosis based on narrow-band imaging endocytoscopy for diagnosing colorectal lesions: comparison with experts." International journal of computer assisted radiology and surgery 12.5 (2017): 757-766.
 - (7) 浜本隆二. "人工知能技術を用いたがんの本態解明及び診断・治療への応用." 日本分子腫瘍マーカー研究会誌 34 (2019) : 16-18.
 - (8) Ishioka, Mitsuaki, Toshiaki Hirasawa, and Tomohiro Tada. "Detecting gastric cancer from video images using convolutional neural networks." Digestive Endoscopy 31.2 (2019): e34-e35.
 - (9) 影山昌広, 外: 新しい畳み込みニューラルネットワーク設計法を用いた肺がん CT 検診向け結節検出 CAD システム, 第75回日本放射線技術学会総会学術大会(2019.4)
 - (10) <https://www.konicaminolta.com/jp-ja/newsroom/topics/2019/0410-02-01.html>
 - (11) <https://www.nikkan.co.jp/articles/view/00509541>
 - (12) Rajpurkar, Pranav, et al. "Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning." arXiv preprint arXiv: 1711.05225 (2017).
 - (13) <http://www.image-net.org/>
 - (14) https://ganjoho.jp/med_pro/pre_scr/screening/screening_lung.html