

5. AIの乳房超音波検査への応用に向けて解決すべき諸問題

林田 哲 / 北川 雄光 慶應義塾大学医学部一般・消化器外科

対策型・任意型検診において、現在の乳房一次スクリーニングとして頻用されているのは、マンモグラフィ検査および乳房超音波検査である。マンモグラフィ検査は診療放射線技師が撮影した個々の患者の画像を読影認定医がまとめて読影・判定を行うのに対して、乳房超音波検査は被検者の乳房を医師や診療放射線技師が直接調べ、リアルタイムに所見を記載する生理機能検査に分類される。乳房超音波検査の精度は機材の性能や環境、観察者によるプローブの操作法や経験の深さ、疾患に対する知識や所見の取り方などのほか、検査施行者の能力と主観にも大きく左右される。そのため、検査精度の均てん化を目的に、2003年に米国放射線専門医学会により、乳房超音波所見報告における用語の基準化のためにBI-RADS分類が作成され、所見用語の統一化や均質な診断が行われるようになってきた¹⁾。

近年、乳がん検診における乳房超音波検査の役割が増し、本邦においても対策型検診への導入の動きが活発化した結果、すでに多くの自治体で導入が進んでいるため、これを担う医師や診療放射線技師の育成、検査精度向上は急務であると考えられる。しかし、長い検査経験があるベテランであっても、見落としの危険や病変の解釈を巡る判断の誤りなどに対して、常に不安感を持ちながら検査を行っているのが実態であり、ましてや検査経験の浅い診療放射線技師・医師の診断能力は一朝一夕には向上しない。乳がん患者はこの30年間一貫して増加傾向にあり、いまや日本

人女性の11人に1人が乳がんの診断を受け、女性が罹患する最も頻度が高いがんとして認知されている²⁾。このような社会的・医学的な背景から、今後スクリーニング目的で行われる乳房超音波検査の需要が増加すると考えられており、検査技師および医師の育成や診断技術の均てん化が重要な課題である。この課題を解決するために、ディープラーニング技術を用いた人工知能(AI)の応用を検討するに至った。

当初よりこのシステムの開発に当たっては、乳房超音波検査における病変の有無とその良悪性の判定に、人間の能力を超えた客観性を持つことをめざしている。すなわち、人の目では判断しきれない超音波画像の特徴をAIが読み取り、BI-RADS判定による要精密検査の判断や、良悪性を瞬時に提示可能とすることを到達すべき目的とした。これを実現するために、AIを専門的に社会応用しているフィックスターズ社が提供する畳み込みニューラルネットワーク(convolutional neural network: CNN)をベースとしたAI技術に対して、慶應義塾大学が提供するアノテーション済み乳房超音波画像を教師データとしてシステム構築を行ってきた(図1)。現状では感度・特異度共に高い精度で腫瘍性病変の良悪性の判定を0.01秒以下で行い、リアルタイム診断にも対応可能なシステムが構築されている(図2)。本システムの用途・アプリケーションを現在検討中であるが、さまざまな問題点を検討する必要がある。本稿では、開発の途中で得た気づきや今後解決すべき問題点などを挙げていきたい。

これまで意識していなかった乳房超音波検査の特性

超音波検査の特性は、前述の「生理機能検査」に分類されることである。CT・MRIやマンモグラフィなどは、一度撮影してしまえば画像が残り、後から複数の読影者が客観的にこれを評価し、読影結果を記載することが可能である。超音波検査も画像として残すことは可能であり、後に評価の指針の一つとして検討が必須であるが、最も重要視されるべきは観察者によるリアルタイムな検討と、それに基づく主観的な印象・判断である。アノテーション作業を行った教師データによるAIの教育には、画像の数にもよるが数十時間必要である。一方で、判定に要する時間はある程度の速度を持つGPUを用いることで、前述のとおり0.01秒以下に抑えることができ、静止画の連なりである動画に対する判定が技術的に可能である。上記の生理機能検査の特性から、リアルタイムに検査中の動画を判定することにより、乳房腫瘍の良悪性を明示することで、検査技師の主観的判断をサポートするシステムが有用と考え開発に成功した。しかしながら、実際にこれを用いて検査を行った検査技師からの評価は散々であった。

なぜリアルタイム支援システムに対して検査者からの支持を得られなかったのかについて聞き取り調査を行った。その結果われわれが認識したことは、「乳房