

超音波検査のワークフロー改善を目的とした断面認識AIの開発

富士フイルム㈱ メディカルシステム開発センター

竹本 昂生



【はじめに】

高齢化社会の進行や医療現場における慢性的な人材不足を背景として、限られた医療資源の中で高品質な医療を提供するための業務効率化が求められている。超音波検査は非侵襲的かつ簡便な診断法として広く活用されており、検診から日常診療まで幅広い場面で利用されている。一方で、超音波検査は操作者の技量に依存する側面が大きく、検査品質の維持や標準化、さらには検査ワークフローの効率化が重要な課題となっている¹⁾。

近年、Deep Learningをはじめとする人工知能(AI)技術の発展により、医用画像認識の性能は大きく向上している。超音波画像に対しても、病変検出、臓器認識、画質改善などさまざまな応用が検討されており、検査支援技術としての活用が期待されている。特に、超音波診断装置が検査中の状況を自動的に認識し、その結果に基づいて装置操作の一部を自動化できれば、操作者の負担軽減および検査ワークフローの改善につながる可能性がある。

本稿では、超音波検査のワークフロー改善を目的として開発を進めている画像認識技術の一例として、超音波画像から検査中の断面を自動認識する断面認識AIを取り上げる。日本消化器がん検診学会、日本超音波医学会および日本人間ドック学会の3学会合同で策定された腹部超音波検診判定マニュアルにおける推奨記録25断面¹⁾を対象とし、畳み込みニューラルネットワーク(Convolutional Neural Network: CNN)を用いて超音波画像と対応する断面との関係を学習させるフレームワークを構築した。また、当社内で取得した健常者データを用いて性能評価を行い、その有効性について検討した。

【腹部超音波検査の課題】

腹部超音波検査は、肝臓、胆嚢、膵臓、脾臓、腎臓および腹部大動脈など複数の臓器を対象とする検査であり、検診や日常診療において広く実施されている。腹部超音波検診判定マニュアルでは、検査の質的向上と標準化を目的として対象臓器や記録方法、判定基準などが定められている。一方で、超音波検査は装置、操作者および被検者の状態によって診断精度が左右されることが知られており、検査品質の維持と標準化が重要な課題となっている¹⁾。

実際の腹部超音波検査では、被検者受け入れ、本人確認、被検者情報の入力、被検者ポジショニング、息止めの指示、画像描出、画像保存、PACSへの画像転送、被検者送り出しおよびレポート入力など、多くの工程が連続して実施される。特に画像描出の工程では、対象臓器をくまなく走査しながら病変の有無を観察するとともに、規定された断面を撮像し記録として保存する必要がある。また、異常所見が認められた場合には、病変を示す画像の保存に加え、病変径の計測や位置情報の記録などの追加作業も必要となる。

近年の腹部超音波検診では、検査の標準化と見落とし低減を目的として推奨記録25断面の取得が推奨されている。しかし、撮像断面数の増加は検査時間の延長や操作者の作業負荷増加につながる事が指摘されている。実際に、推奨記録25断面の撮像手順を見直すことで検査時間の短縮が可能であることが報告されており、記録画像取得に関するワークフロー改善の重要性が示されている²⁾。

さらに、超音波診断装置は腹部領域だけでなく、循環器、乳腺、産婦人科、血管など多様な診療領域に対応するため、多数の画像処理機能や計測機能を搭載している。その結果、トラックボール、タッチパネル、各種ボタンおよびつまみなど複数の入力デバイスを用いた複雑な操作が必要となる。操作者はプローブを操作しながら画像を観察し、同時に装置設定や記録操作を行う必要があり、高い集中力と熟練した操作技能が要求される。

また、日本超音波医学会は、超音波検査がプローブ把持、画像観察および装置操作を同時に行う作業であり、不自然な姿勢や長時間の反復作業によって操作者の身体的負担や眼精疲労が生じることを指摘している³⁾。そのため、検査品質を維持しながら定型的な装置操作や記録作業を効率化し、操作者の負担軽減を図ることが求められている。

以上のような背景から、当社は画像認識技術を活用し、超音波診断装置が現在の検査状況を自動的に把握するとともに、検査中に発生する定型的な装置操作の一部を自動化する技術の開発を行った。次の章では、その要素技術の一つとして、腹部超音波検診で用いられる推奨記録 25 断面を自動認識する断面認識 AI について述べる。

【腹部断面認識AIの開発】

当社は、学会が定める推奨記録25断面を対象として、同様の対象臓器を撮像したBモード画像84,716枚(426症例)を収集した。取得された各画像に対し、超音波検査経験を有する臨床検査技師3名が腹部超音波検診判定マニュアルに基づき1から25までの断面ラベルを付与し、教師データとして採用した。教師データ作成にあたっては、画像中に表示されるボディマークやアノテーションが断面分類とは無関係な特徴であり、作成するAIによる断面分類モデルがそれらに過度に依存する可能性があることから、画像編集ソフトウェアを用いて事前に除去した。学習には畳み込みニューラルネットワーク(Convolutional Neural Network: CNN)を用いた。対象臓器の個人差やプローブ操作による画像変動を模擬するため、回転、平行移動、拡大縮小などのデータ拡張(Augmentation)を適用し、学習時に多様な画像を生成することでモデルの汎化性能(未知データに対する予測性能)向上を図った。これらの教師データを用いて、臓器の形状や輝度分布などの画像特徴と対応する断面との関係を学習させ、断面分類モデルを構築した。

Deep Learningによる学習では、学習データに過度に適合することで未知データに対する性能が低下する過学習(Overfitting)が生じる場合がある。この過学習を抑制するため、検証データに対する認識性能が最良となった時点で学習を終了するEarly Stoppingを採用した。

推論時(AIによる画像判定時)には、Bモード画像を学習済みモデルへ入力し、画像解析を実施する。モデルは入力画像に対して25断面それぞれへの確率分布を出力する。最終的に最も高い確率を示した断面を認識結果として取り扱う(図1)。

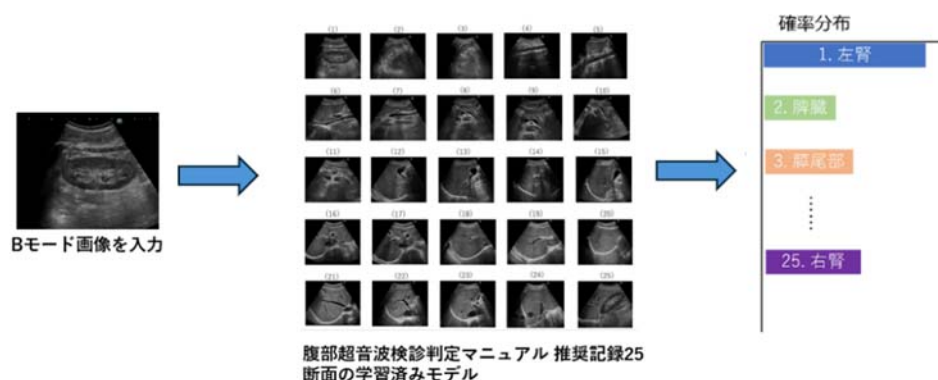


図1 Deep Learning による画像認識のフロー

モデルの性能検証には、426症例全体を用いた5分割交差検証 (Cross Validation) を実施した (図2)。各試行では 426症例を 5群に分割し、そのうち 4群を学習データ(Training)、残り 1 群を評価データ (Validation) として使用した。

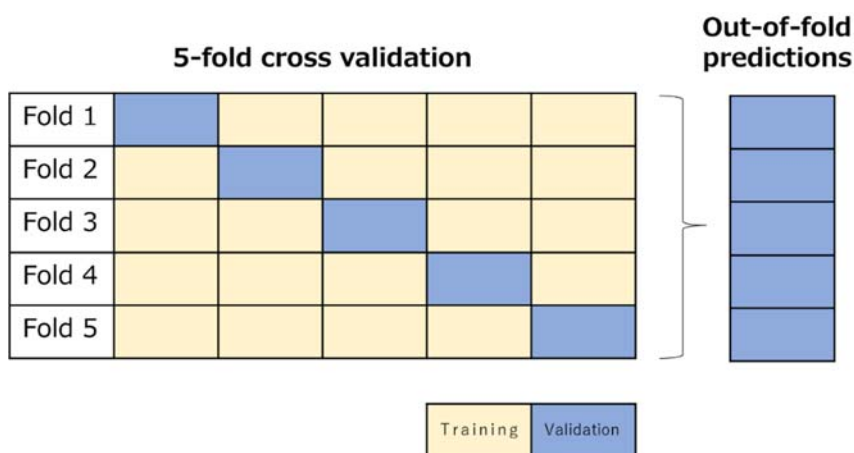


図2 5 分割交差検証

【モデルの性能評価】

未知のデータに対する分類性能を評価するために、社内ボランティア 4 例の健常者データを取得した。評価指標として、断面ごとの適合率 (Precision)、再現率 (Recall) を採用した (表1) (図3)。適合率 (Precision) はモデルが予測した断面のうち正しく分類できた割合を示し、再現率 (Recall) は各断面に属する画像のうちモデルが正しく予測できた割合を示す。

表1 評価指標の基本用語

	正解 (正)	正解 (負)
認識結果 (正)	True Positive (TP)	False Positive (FP)
認識結果 (負)	False Negative (FN)	True Negative (TN)

$$\text{Precision} = \frac{TP}{TP + FP}$$

$$\text{Recall} = \frac{TP}{TP + FN}$$

図3 評価指標の定義

表2 に断面ごとの予測精度の一部を示す。

表2 断面ごとの予測精度

断面名	適合率(%) (Precision)	再現率(%) (Recall)
1) 左肋間走査:左腎	100	100
2) 左肋間走査:脾臓	100	100
3) 左肋間走査:脾臓・脾尾部	100	100
4) 心窩部縦走査:腹部大動脈	100	100
5) 心窩部縦走査:肝左葉(肝縁)	100	100
6) 心窩部縦走査:下大静脈・肝左葉・尾状葉(S1)	100	100
7) 心窩部縦走査:脾頭部(脾鉤部)	25	50
8) 心窩部横走査:脾体部	100	50
9) 心窩部横走査:脾体部(拡大で主脾管径計測)	25	100
10) 左肋骨弓下斜走査:脾尾部	100	75
11) 心窩部斜走査:脾頭部	33	50
{ 12)～24) 省略	-	-
25) 右肋間走査:肝後下区域(S6)・右腎	100	100

順番と断面名は腹部超音波検診判定マニュアル 推奨記録 25 断面に準ずる。

さらに、図4に断面ごとの認識結果を混同行列形式で示す。縦軸を正解断面ラベル、横軸をAIの判定結果として二次元プロットしたものである。対角線上に集まり濃い色になるほど、正解とAIの判定が一致していることを意味する。本モデルでは大部分が対角線上に集中しており、AIが正解を高い精度で再現できていることがわかる。

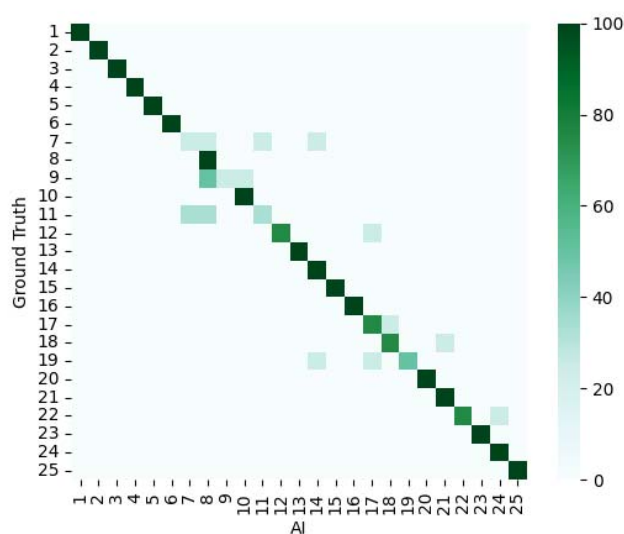


図4 断面毎の再現率(混同行列形式)

断面別に評価すると、腎臓(断面 1)、脾臓(2)および腹部大動脈(4)のように特徴的な形状や輝度パターンを有し、他断面との類似性が低い断面では高い認識精度が得られた(図5)。一方、脾臓(7)~11))では、消化管ガスや体格差の影響により描出状態のばらつきが大きく、肝区域断面では隣接区域間で類似した画像特徴を有することから、相対的な精度低下が認められた(図6)。また、混同行列からも誤分類の多くは解剖学的に近接した肝区域間で発生していることが確認された。



図5 AI は4)心窩部縦走査：腹部大動脈と正答



図6 正解は11)心窩部斜走査：脾頭部だが、
AI は8)心窩部横走査：脾体部と誤答

【検査のワークフロー向上のために】

前章で示した断面認識AIは、推奨記録25断面に対して高い認識性能を示した。この断面認識結果を活用することで、超音波診断装置は現在撮像されている断面を自動的に判別し、その結果に応じた各種操作の自動化が可能となる。

本技術は当社最新装置においてボディマーク／プローブマーク機能に応用されている*1。従来の超音波検査では、操作者が断面をフリーズした後、撮像断面に対応するボディマークを選択し、さらにプローブマークの回転および位置調整を手動で行う必要があった。このため、記録画像を保存するたびに複数回の装置操作が発生していた。一方、本技術では、断面認識AIが現在の撮像断面を自動認識することで、図7に示す①～④の操作を自動化している。これにより、断面認識結果に基づき、ボディマークおよびプローブマークが自動設定される。認識結果には誤りが生じる可能性があるため、操作者は設定内容を確認し、必要に応じて修正する。これらの操作は超音波検査において日常的かつ繰り返し実施される作業であり、操作回数の削減は検査時間短縮に寄与すると考えられる。



図7 ボディマーク/プローブマークの自動化

超音波画像の断面認識AIを用いて、装置操作を自動化する取り組みの一例を示した。今後はボディマーク設定だけでなく、撮像条件や画質パラメータの自動調整など、さらなるワークフロー改善への応用が期待される。

*1 AI技術のひとつであるDeep Learningを用いて開発・設計した。自動的に装置の性能・精度が変化することはない。自動化機能の結果は操作者の確認が必要である。必要に応じて結果を修正することが可能である。

【まとめ】

本稿では、超音波検査のワークフロー改善を目的として開発を進めている画像認識技術の一例として、腹部超音波検診における推奨記録25断面を対象とした断面認識AIについて述べた。

腹部超音波検査では、対象臓器の観察に加えて記録画像の保存や装置操作など多くの作業が必要であり、検査品質の維持と検査効率向上の両立が求められている。そこで、CNNを用いた断面認識AIを開発し、426症例、84,716枚のBモード画像を用いて学習および評価を行った。その結果、推奨記録25断面の認識タスクに対して十分に高い認識性能を有することを確認した。

さらに、本技術をボディマーク／プローブマーク機能へ応用することで、従来は操作者が手動で行っていた設定操作の自動化を実現した。これにより、超音波検査における反復操作を削減し、検査ワークフローの改善に寄与できる可能性を示した。

今後は、さらなる認識性能の向上に加え、画像認識技術を活用した装置操作支援や検査支援機能への展開を進めることで、超音波検査の効率化と検査品質の向上に寄与していきたい。

【参考文献】

- 1) 一般社団法人日本消化器がん検診学会：腹部超音波検診判定マニュアル改訂版(2021年):日本消化器がん検診学会雑誌, Vol.60(1) Jan.2021.
- 2) 山本敏樹, 小川真広, 金子真大, ほか:腹部超音波検診判定マニュアル基準断面の診断装置内蔵ソフトの撮影手法変更が workflow に与える影響. 日本消化器がん検診学会雑誌, 60(3):366-374, 2022.
- 3) 一般社団法人日本超音波医学会 機器及び安全に関する委員会:超音波検査者が安全・快適で健康的に働くための提言ー作業関連筋骨格系障害と眼の障害を予防するための機器と作業環境ー.