

Communication of the Imaging Group of the JSRT

画像通信

2026年4月

Vol.49 No.1(通巻 96)

☆ 教育講演

「非線形再構成時代の CT 画像評価: 普遍性の喪失にどう向き合うか？」

茨城県立医療大学大学院 瓜倉 厚志

☆ 第 98 回画像部会 『パラダイムシフト?! 臨床画像評価の未来』

1. 「臨床画像評価のための物理評価」

東北大学病院 小野寺 崇

2. 「ゼロから学ぶ ROC 解析 ～判断基準とカットオフ値で理解する「観察者実験」の仕組み～」

西南学院大学 馬場 祥吾

3. 「臨床画像の評価および読影における課題への取り組み」

福岡大学筑紫病院 田畑 成章

4. 「臨床画像評価の未来 ～AIを使った画質評価法～」

滋賀医科大学医学部附属病院 大橋 昂平

☆ 技術紹介

①「骨密度測定装置の技術解説」

GE ヘルスケア・ジャパン株式会社 矢ヶ部 真理

②「画像の再現性向上をアシストする光学カメラソリューションの最前線」

キヤノンメディカルシステムズ(株) 鈴木 省吾・菊地 航希

☆ 読者のページ

①「第 46 回医用画像処理プログラミングセミナーに参加して」

横浜栄共済病院 柿原 妃菜子

②「第 46 回医用画像処理プログラミングセミナーを受講して」

順天堂大学大学院 隈本 神太郎

③「第 15 回 ROC セミナー参加記 「わからない」から「少しわかる！」へ、ROC の見え方が変わった体験」

山口労災病院 皆川 圭太

④「第 15 回 ROC セミナーに参加して」

社会医療法人 北農会恵み野病院 千葉 峻介

⑤「第 12 回臨床画像評価セミナー ～マンモグラフィの画像評価の基礎～ に参加して」

愛知県がんセンター 小林 杏花

⑥「第 12 回臨床画像評価セミナー参加記 ～マンモグラフィの画像評価の基礎～ を受講して」

医療法人志成会 阪本胃腸・外科クリニック 平松 佐和子

☆ 専門部会講座(入門編): 「X 線画像の視覚評価と試料数の決定について」

和歌山県立医科大学附属病院 山崎 純

☆ 専門部会講座(入門編): 「Python 入門とパイプコーディングの可能性」

岡山大学学術研究院 福井 亮平

☆ 国際会議案内・報告: 「RSNA2025 の参加記」

金沢大学 川嶋 広貴

☆ 画像部会「研究情報サイトの紹介」

岐阜大学 教育学部技術教育講座 福岡 大輔

☆ 2025 年度事業報告・2026 年度事業計画

☆ 画像部会入会案内



JAPANESE
SOCIETY
OF
RADIOLOGICAL
TECHNOLOGY

公益社団法人 日本放射線技術学会
画像部会

第 99 回 画像部会予告

日 時：2026 年 10 月 23 日(金)～25 日(日) の第 54 回日本放射線技術学会秋季学術大会期間中

会 場：G メッセ群馬（高崎市）

1. 核医学部・会画像部会 合同企画

テーマ「画像の成り立ちと AI：物理的基礎から画像処理の実装まで」（仮題）

2. 画像部会・核医学部会 合同企画

テーマ「生成 AI が拓く放射線技術の新展開：臨床・研究・業務の革新」（仮題）

臨床画像評価セミナーの開催予定

第 13 回臨床画像評価セミナー :2026 年 未定

会場：未定

医用画像処理プログラミングセミナーの開催予定

第 47 回医用プログラミングセミナー :2026 年 7 月 26 日(日)

会場:オンライン実施 (予定)

画像部会委員 氏名・所属・電子メール

東出 了 (画像部会長)	鈴鹿医療科学大学	raryo@suzuka-u.ac.jp
小野寺 崇	東北大学病院診療技術部放射線部門	shu.onodera.e2@tohoku.ac.jp
川嶋 広貴	金沢大学医薬保健研究域保健学系	kawa3@med.kanazawa-u.ac.jp
佐保 辰典	小倉記念病院	saho-t@kokurakinen.or.jp
平原 大助	帝京平成大学	d.hirahara@thu.ac.jp
福井 亮平	岡山大学	rfukui@okayama-u.ac.jp
山崎 明日美	神戸常盤大学	a-yamazaki@kobe-tokiwa.ac.jp
山崎 純	和歌山県立医科大学附属病院	jun.ymsk.77ppp@gmail.com
由地 良太郎	東海大学医学部付属八王子病院	ryotaro.yuji@gmail.com

画像部会についてご意見やご希望等がありましたらご連絡ください。

画像部会に関する情報は、以下の web ページをご利用ください。

日本放射線技術学会：<http://www.jsrt.or.jp>

画像部会：<http://imgcom.jsrt.or.jp>

非線形再構成時代の CT 画像評価: 普遍性の喪失にどう向き合うか?

CT image evaluation in the era of nonlinear reconstruction: How should we respond to the loss of universality?

茨城県立医療大学大学院保健医療科学研究科放射線技術科学領域
国立がん研究センター中央病院放射線診断科

瓜倉 厚志

要旨

X 線 CT における画像再構成法が filtered back projection から逐次近似再構成や deep learning-based reconstruction へと移行する中で、画質評価のパラダイムは大きな転換期を迎えた。非線形システムでは、従来の線形システムで培われた評価の「普遍性」が成立せず、空間分解能や画像ノイズが線量や被写体コントラストに対してアダプティブに変化する。本稿では、非線形画像の登場に伴う評価手法の変遷を概説し、タスクベース画像評価の重要性について解説する。また、CT 領域で先行して経験された問題提起を通じて、最新の非線形画像処理の実装が進む digital radiography 領域の研究者に向けた、モダリティを越えた評価体系構築の必要性を示す。

1. はじめに

X 線 CT では、主に新しい画像再構成法への移行に伴い、非線形画像に対応する評価法が一般化してきた。画像評価における非線形システムへの対応は、線形システムで培った普遍性の喪失を意味し、画質評価のプロセスに多大な困難と複雑さをもたらしている。普遍性を持たない測定結果は、極めて限定された条件下、あるいは特定の臨床条件を想定した場合にのみ適用可能であり、それを他の条件へ外挿して反映することは容易ではない。すなわち、非線形画像評価では、評価の目的（タスク）を明確にすることが必須である。タスクといっても様々であり、画像再構成法、線量、被写体径、評価対象の径やコントラスト、その他の撮像パラメータ等、多くの組み合わせがあり極めて煩雑である。本稿では、X 線 CT における非線形画像評価について、その変遷といくつかの具体例を示し概説する。

2. 線形から非線形画像への変遷

X 線 CT は、1968 年に Hounsfield によって発明され(1)、1972 年に商用機が稼働（日本国内では 1975 年から導入）したが、1977 年には早くも米国医学物理学会 (American Association of Physicists in Medicine: AAPM) から AAPM report No. 1 が刊行された(2)。AAPM report は、医学物理学および関連分野における技術情報の作成と普及を目的としたものであるが、その第一弾として CT スキャナの性能を評価する

実践的方法を提供した。その後、日本国内においても X 線 CT の性能評価指標が提案され、CT 技術の発展とともに改訂が行われてきた。また、一般的な CT 画像に関する研究の画質評価では、しばしばこの性能評価項目に基づく物理的画質指標が用いられ、あるいはそれを応用することで空間分解能や画像ノイズの客観的評価が行われてきた。

CTに限ったことではないが、性能評価のガイドラインで用いられるような物理指標は、画像の線形性に基いている。CTでは、長い間 filtered back projection (FBP) が画像再構成法として用いられてきた。FBPで再構成された画像は、X線量、幾何学的条件に応じた画質変化を示すため、多くの場合において線量と画質の間には線形関係が成立する。また、診断部位に応じて複数のフィルタ関数が用いられるが、それらの空間周波数特性(空間分解能やノイズの周波数成分)は線量に依存せず一定であるため、周波数解析をするうえでも普遍的であった。

2000年代になり、画像再構成の過程に非線形処理を加えることで空間分解能を維持しつつ画像ノイズを低減する非線形フィルタが商用機で使用可能になった(3,4)。非線形フィルタの登場後、JSRTの学術大会でもその評価法について議論があったことを記憶している。また、非線形フィルタに対応した画質評価法について先駆的な論文が国内から発表された(5,6)。しかしながら、当時の非線形フィルタは一部の機種での稼働に限られていたことに加え、必ずしも積極的に使用されたわけではなかったため、多列化や高速化といった当時の主要なトピックに比べるとニッチな領域であった。

2007年、逐次近似再構成法(iterative reconstruction: IR)が臨床機に搭載され、大きな注目を集めた。IR画像は、上述の非線形フィルタと同様に“空間分解能を維持しつつ画像ノイズを低減する”というたい文句で登場したが、世界的な医療被ばくに対する関心の高まりも相まって多くの研究が行われ、爆発的に論文数が増加した。IRの臨床導入から数年は、IRの画像特性について極めて肯定的な結果を示す研究が多く、Willeminckらのレビューによると線量低減率(achieved dose reduction)は、対象部位にもよるが20-76%を達成しつつ画質を維持するといったものであった(7,8)。国内でも同様の発表が多く行われていたが、空間分解能、画像ノイズ、低コントラスト検出能などの物理的指標と視覚的印象の乖離については早期から議論の対象であった。一方、一部の放射線科医師からは、日本の診療放射線技師は画質変化に敏感過ぎるところがある(気にしすぎであり慣れば問題ない)と言われたこともあった。図1は、同一患者の腹部CT画像であり、通常線量で撮像されFBPで再構成された画像(左)と、1/8の線量で撮像されIRで再構成された画像(右)である。腹部大動脈に配置したregion of interest内のCT値の標準偏差はほぼ同等であるが、低線量IR画像は大まかな構造を除いて診断には不適であることが一目瞭然である。また、ノイズの質感が異なることから、“違和感”、“blotchy”、“pixilated image texture”などと表現され、従前から用いられてきた物理的画質指標をIRのような非線形画像に適用することの限界は明らかであった。

筆者らは、IR画像の空間分解能が線量やIRの処理強度に依存することを客観的に示した(9,10)。また、同時期にRichardらは、IRに対応した空間分解能測定法として、非線形システムのタスクベース評価のための modulation transfer function (MTF) 測定法として MTF_{Task} を提案し、IR画像の MTF_{Task} が線量や被写体コントラストに依存することを明らかにした(11)。その後、 MTF_{Task} は、task transfer function (TTF)と呼ばれて広く普及している。

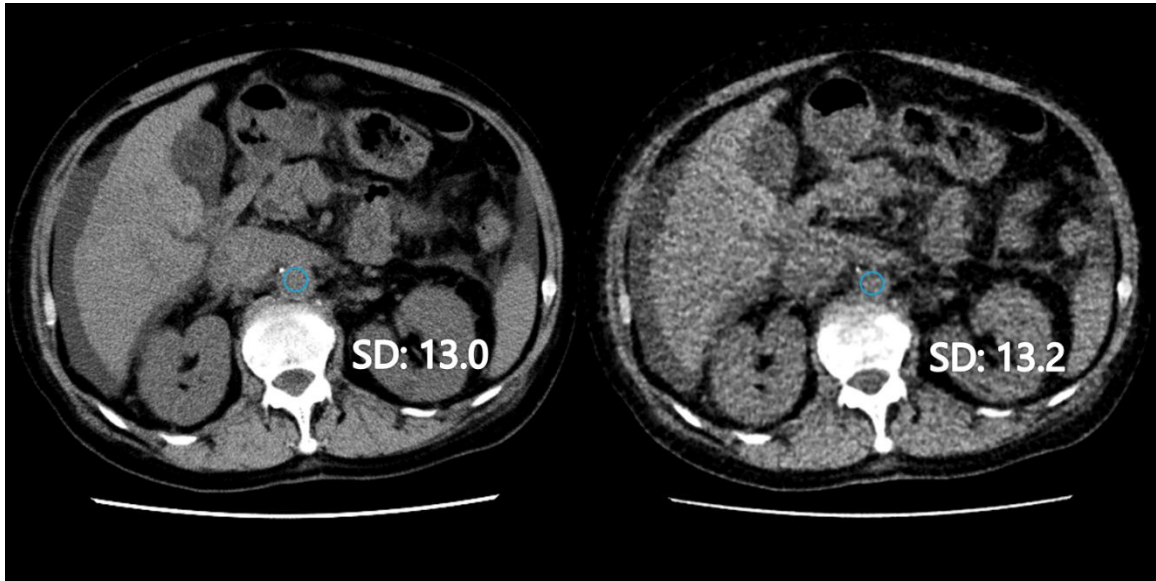


図1 通常線量のFBP(左)と、1/8線量のIR(右)画像. SD: 関心領域(青丸)内のCT値の標準偏差.

低コントラスト検出能についても、それまで発表された過激な線量低減から、読影試験に基づいた冷静な臨床研究も散見されるようになり、IRはcontrast-to-noise(CNR)のようなimage qualityを改善したが、低コントラスト検出能は維持できないといった報告が行われるようになった(12, 13). 非線形画像に対応した低コントラスト検出能の物理指標についてもいくつかの方法が提案され、線量低減による低コントラスト検出能の低下が明らかになった(14-16). 当時提案された空間分解能や低コントラスト検出能の評価手法は、線形システムではあまり問題にならなかった被写体コントラストや線量の変化に伴うアダプティブな画質変化を捉えたことが画期的であった。また、腫瘍などの信号径に対応した空間周波数領域のノイズ特性に注目することで、視覚的な検出能との乖離を解消した。その後も同様の研究は国内外から発信され、いくつかの手法が提案されている。

2018年には、artificial intelligence(AI)技術のひとつである深層学習に基づく再構成法(deep learning-based reconstruction: DLR)が臨床機で使用可能になり、IR以上のノイズ低減や周波数特性の維持が可能となりつつある。さらに、AAPMは、Report No. 233(TG233)において、非線形画像を前提にしたCTの画像評価法を体系立ててまとめた(17)。その後、TG233は日本国内でも多く引用され、近年のCT画像評価の標準化に貢献している。

3. 空間分解能

一般的に、空間分解能は、“ノイズのない環境において二つの物体を互いに区別する能力”と定義される。CTでは、スライス面に直交する細径の金属ワイヤをインパルス信号としてpoint spread functionを取得し、MTFを算出する方法が用いられ、日本産業規格(JIS)の性能評価項目にも掲載されている。金属ワイヤのCT値は1,000 Hounsfield unit(HU)を超えるため、標準的な線量におけるバックグラウンドノイズは無視できることから、ほぼ“noiseless field”となる。この測定手法を非線形画像に適用した場合、ワイヤから取得したMTFは、その高いCNRから低線量下でも低下することなく維持される(10)。ここで問題なのはワイヤ法の測定条件である。ワイヤ法では、極めて高いコントラスト、画像ノイズが

問題にならない線量，狭小の field of view で取得した画像を用いる．前項でも触れたとおり，IR や DLR のような非線形処理は，コントラスト，ノイズ，撮像パラメータに応じてアダプティブな画像特性を示すため，ワイヤ法で取得した MTF は，極端な条件下の解像特性を表しているに過ぎない．実際に，低線量条件の IR 画像でも皮質骨の境界などは比較的解像特性を維持する．

CT の撮像条件や対象コントラストに応じた IR 画像の空間分解能を反映するには，低線量（高ノイズ）かつ低コントラスト条件における TTF を測定することも想定されるが，“noiseless field” の定義から逸脱するため測定誤差要因となる．この問題を解決するために，筆者らは同じ位置で取得した大量の画像を加算平均化することで，低線量画像でもほぼ“noiseless field” を再現し，TTF 測定誤差を低減した(10)．加算平均化を用いることによって，TTF 測定の堅牢性が向上するため，異なる線量，コントラスト，被写体径等の様々なタスクに対応した TTF 測定が可能となった．また，TTF 測定では，タスクに応じたコントラストを再現しやすいエッジ法を用いることが多い．特に，ある程度の径 (> 30 mm) を有する円形ロッドに対して，ロッドの中心から放射状にエッジを取得してプロファイルを合成する circular edge 法が主流である．しかしながら，放射状に取得した edge spread function を精度よく合成することは容易ではない．こうした測定の煩雑さを解消するため，日本 CT 技術学会は，会員向けに CT 画像計測を効率化するソフトウェア (CTmeasure, <https://www.jsct-tech.org/members/software/>) をいち早く提供した(18)．測定法の提案とともに，こういった計測自動化ソフトウェアの普及は，非線形画像の空間分解能評価において大きな役割を果たしている．

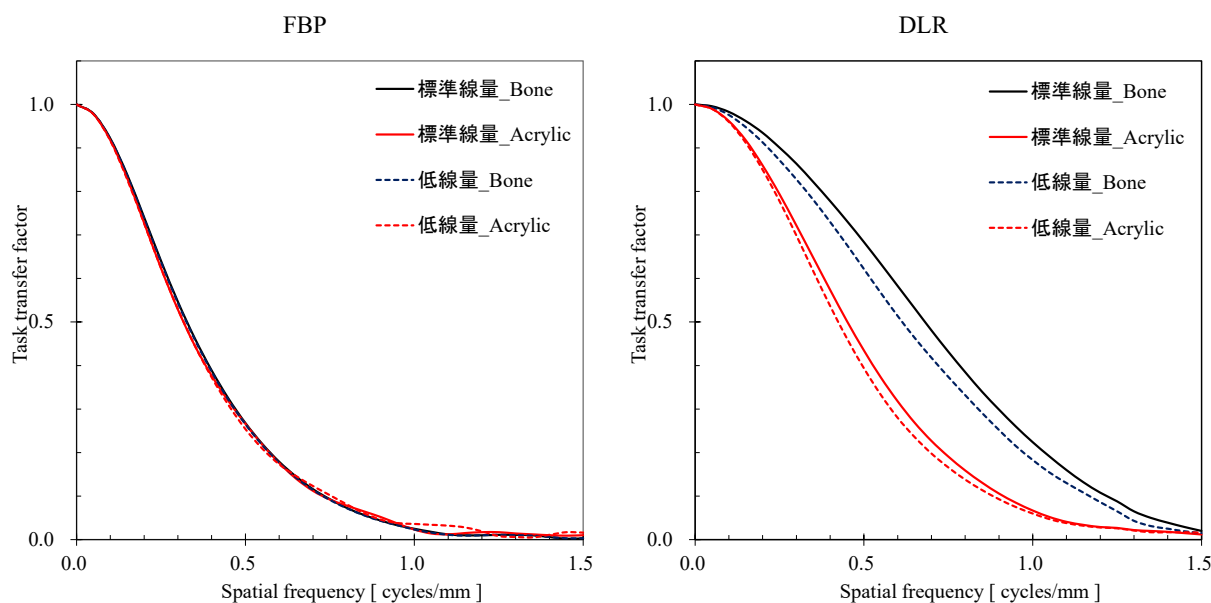


図 2 同一ファントム内の異なるコントラストから計測した TTF. 線形システムである FBP(左)と非線形システムである DLR(右)．

図 2 は，同一ファントム（464 型 ACR CT ファントム，Gammex）内に配置された 2 種類のロッド（Bone: 955 HU; Acrylic: 120 HU）の加算平均化処理画像を circular edge 法で TTF 解析した結果である．FBP 画像では，線量やコントラストに関わらず同等の TTF を示しているが，DLR は，低い線量で TTF が低下し，コントラストによって TTF が異なる．この TTF は非線形画像に特徴的であり，画像ノイズによって TTF が変わるだけでなく，ひとつの画像内にある構造のコントラストに依存して異なる解像

特性を示しているということであり、タスクベース解析の条件設定に極めて慎重な配慮が求められる（面倒くさい）ことを示している。

4. 画像ノイズ

JIS の性能評価や臨床の線量最適化における画像ノイズ指標として、CT 値の標準偏差（standard deviation: SD）がしばしば用いられる。SD は、検出器に到達した X 線光子数のゆらぎによって各画素の CT 値が変動する程度の指標として用いられ、計測が容易であること、線形システムでは線量の最適化指標としても有用であることなどから多用されてきた。しかしながら、前述したとおり非線形画像では CT 値の SD だけで画質の良し悪しを判断することが困難になり、ノイズの周波数特性を含めた評価が必要となった。Noise power spectrum (NPS) は、空間周波数特性を反映可能な指標であり、IR 画像などの画像特性変化を捉えることができるため、CT のノイズ評価でもよく利用される。NPS カーブを観察することで、どの空間周波数領域のノイズ量に変化が生じたか把握することが容易である。TG233 では、NPS の最大値における空間周波数 (f_p) や、NPS カーブの平均空間周波数 (f_A) がノイズテクスチャを表す指標として示されている。図 1 に示したような低線量 IR 画像では、 f_p が低空間周波数領域にシフト（peak shift と呼ばれる）するため、相対的に低周波ノイズが増加し、粒の大きいノイズが目立つ画像になっている。

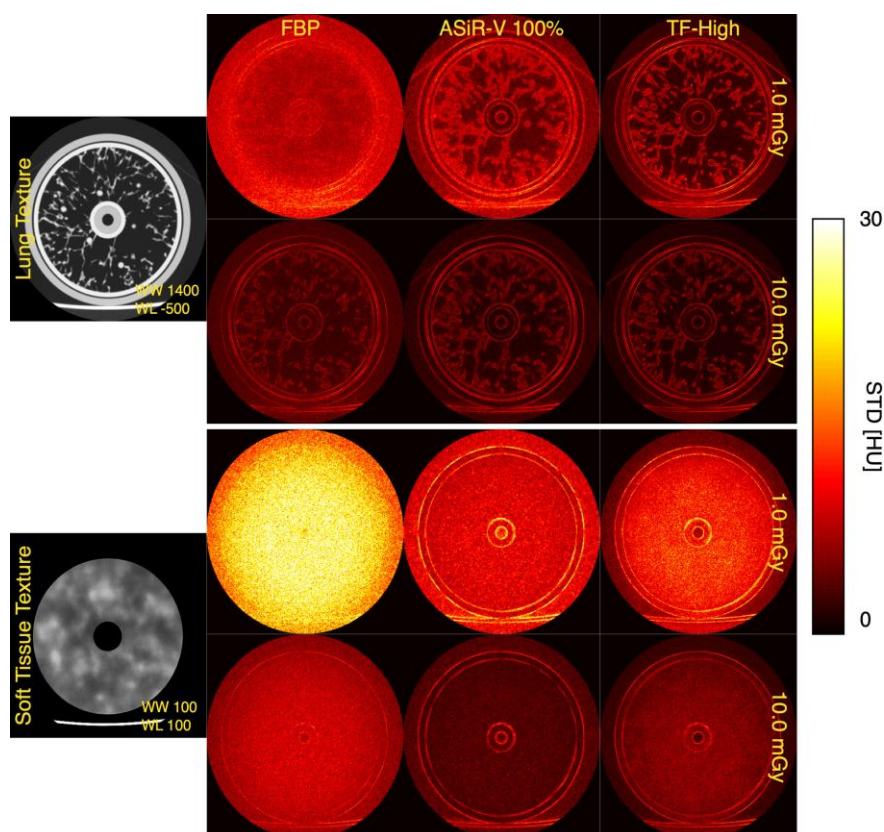


図 3 FBP と DLR アルゴリズムにおけるノイズ分布の比較. FBP ではノイズが空間的に均一であるのに対し、IR (ASiR-V) や DLR (TF-High) では被写体の構造に依存してノイズが不均一に分布している。(Solomon ら(19)より許諾を得て転載。©2020 American Association of Physicists in Medicine.)

さらに面倒なことをひとつ挙げると、非線形画像におけるノイズ低減の程度は、画像内でも均一ではない。なぜなら、画像内のどの程度の構造を信号とみなし、どこまでをノイズ低減の対象とするかという処理特性が、各 IR や DLR のアルゴリズムに強く依存するからである。一般的に、FBP のノイズ評価は均一物質の CT 値の SD を計測するが、IR や DLR ではノイズが画像内で一様であるとは限らず、淡い腫瘍や複雑な構造のある場所ではノイズ分布がまだらになることがある。TG233 では、このような画像ノイズの空間分布を可視化・定量化するアプローチとして noise map を示している。図 3 に、FBP と DLR における noise map の例を示す(19)。FBP ではノイズが空間的に均一であるのに対し、IR や DLR では被写体の構造に依存してノイズが不均一に分布している。すなわち、構造の境界などはノイズを積極的に低減せず、平坦部分はノイズが低い。Noise map は画像内のノイズ分布を可視化する方法であるが、同一個所を複数回 (>20 回) スキャンする必要があるため、臨床画像による評価は今のところ現実的ではない。

5. 低コントラスト検出能

筆者らは、腫瘍などの信号径の周波数に対応した noise power value をノイズ指標とした CNR (low-contrast object specific CNR: CNR_{Lo}) を提案し、観察者試験とよく一致することを報告した(14, 16)。また、Samei らは、MTF, NPS, およびタスク (検出対象のコントラスト・径) を統合した detectability index (d') を線量最適化のツールとして提案した(15)。Ichikawa らの研究グループは、システム固有の性能指標として system performance function (SP 関数)を提案した(20)。後者 2 つの指標は、 $MTF^2(u) / NPS(u)$ によって空間分解能の影響をキャンセルし、システム固有の性能を評価する pre-whitening 理論に基づいている。この考え方は、X 線画像の基本的物理指標である noise equivalent quanta (NEQ)や detective quantum efficiency (DQE)に通じるものである。NEQ や DQE は、主にシステムの入力に対する出力の効率を評価する関数であるのに対し、 d' や SP 関数は検出対象の情報 (タスク関数) を組み合わせることで、非線形システムにおいても適用可能な客観的評価指標に発展させたものである。

これらの指標の背景には、信号検出理論に基づき観察者を数理モデル化して検出能を評価する model observer の考え方がある。すなわち、特定のタスク (病変の径、形状、コントラストなど) を設定し、その信号が存在する場合と存在しない場合の画像から判別スコアを得て、検出能を定量化するという考え方である。Model observer test を用いると、IR や DLR のようにノイズ特性や空間分解能が単純な線形モデルではない場合でも、タスクに対する性能として比較が可能となる。

一方で、model observer の結果は、タスク設定やファントム、背景構造の有無、およびノイズの空間的非一様性などに依存するため、目的とする臨床タスクを適切に反映させることが重要である。したがって臨床条件に反映するには、model observer により客観指標によって比較・最適化を行い、必要に応じて代表条件で観察者試験により妥当性を確認する、という位置付けが現実的である。

6. おわりに

本稿では、X 線 CT における非線形画像の登場とその変遷を示し、画像評価の普遍性が喪失した現状においてどのようなアプローチで画像解析が行われているかについて可能な限りシンプルに解説した。

Digital radiography (DR)では、CTよりも前から非線形画像処理が臨床で使用されている。しかしながら、DRでは画像診断に提供されている非線形処理画像に対する画像評価が未だ整備されていない。この状況は国内に限ったことではないが、診断画質に基づく線量最適化を行うためには、CTと同様に非線形画像への対応が必須である。最近、筆者らの研究グループでもDRの非線形画像評価に取り組んでいるが、DRにはCTとは異なる測定上の問題点も経験している。いずれにしても、非線形画像の評価において普遍性は成立せず、その評価プロセスが極めて煩雑であることは変わらないであろう。DR画像にもdeep learningに基づく最新の画像処理が導入されている中で、「一見すると通常線量と同等以上の良好なノイズ特性に見えても、実際には淡い陰影が消失している」という事態が起こっているかもしれない。我々は画像を扱う専門家として、このような評価の煩雑さに直面しても、診断に十分な画質を有する画像を提供し続ける使命がある。CT領域で培われたタスクベース評価のパラダイムは、DRにおける評価体系構築の重要な足がかりとなるはずである。そのためにも、常に“タスク”が何かを明確にすることが極めて重要である。

参考文献

1. Hounsfield GN. Computerized transverse axial scanning (tomography). 1. Description of system. Br J Radiol. 1973;46(552):1016-22.
2. Judy P, Balter S, Bassano D, McCullough E, Payne J, Rothenberg L. Phantoms for performance evaluation and quality assurance of CT scanners. 1977.
3. 塚越伸介. 量子ノイズ低減化フィルタ(量子フィルタ)による低被ばく時の画質向上. ワークショップ-よりよい撮影技術を求めて(その73)-「腹部CT検査のエビデンスガイドライン」; 2004: 日本放射線技術学会 撮影分科会.
4. 杉原 直樹. 最新マルチスライス CT. Medical Imaging Technology. 2005;23(2):63.
5. Okumura M, Ota T, Tsukagoshi S, Katada K. New method of evaluating edge-preserving adaptive filters for computed tomography (CT): digital phantom method. Japanese Journal of Radiological Technology. 2006;62(7):971-8.
6. Mori I, Machida Y. Deriving the modulation transfer function of CT from extremely noisy edge profiles. Radiological physics and technology. 2009;2(1):22-32.
7. Willeminck MJ, de Jong PA, Leiner T, de Heer LM, Nievelstein RA, Budde RP, et al. Iterative reconstruction techniques for computed tomography Part 1: technical principles. Eur Radiol. 2013;23(6):1623-31.
8. Willeminck MJ, Leiner T, de Jong PA, de Heer LM, Nievelstein RA, Schilham AM, et al. Iterative reconstruction techniques for computed tomography part 2: initial results in dose reduction and image quality. Eur Radiol. 2013;23(6):1632-42.
9. Urikura A, Ichikawa K, Hara T, Nakaya Y, editors. A new method to evaluate spatial resolution of iterative reconstruction images in CT2012: European Congress of Radiology-ECR 2012.

10. Urikura A, Ichikawa K, Hara T, Nishimaru E, Nakaya Y. Spatial resolution measurement for iterative reconstruction by use of image-averaging techniques in computed tomography. *Radiol Phys Technol.* 2014;7(2):358-66.
11. Richard S, Husarik DB, Yadava G, Murphy SN, Samei E. Towards task-based assessment of CT performance: system and object MTF across different reconstruction algorithms. *Med Phys.* 2012;39(7):4115-22.
12. Schindera ST, Odedra D, Raza SA, Kim TK, Jang HJ, Szucs-Farkas Z, et al. Iterative reconstruction algorithm for CT: Can radiation dose be decreased while low-contrast detectability is preserved? *Radiology.* 2013;269(2):511-8.
13. Love A, Olsson ML, Siemund R, Stalhammar F, Bjorkman-Burtscher IM, Soderberg M. Six iterative reconstruction algorithms in brain CT: a phantom study on image quality at different radiation dose levels. *Br J Radiol.* 2013;86(1031):20130388.
14. 瓜倉厚志, 市川勝弘, 原孝則, 西丸英治, 中屋良宏. 逐次近似再構成画像の低コントラスト検出能評価. *JSCT; 2014: Japanese Society of Computed Tomography.*
15. Samei E, Richard S. Assessment of the dose reduction potential of a model-based iterative reconstruction algorithm using a task-based performance metrology. *Med Phys.* 2015;42(1):314-23.
16. Urikura A, Hara T, Ichikawa K, Nishimaru E, Hoshino T, Yoshida T, et al. Objective assessment of low-contrast computed tomography images with iterative reconstruction. *Phys Med.* 2016;32(8):992-8.
17. Samei E, Bakalyar D, Boedeker KL, Brady S, Fan J, Leng S, et al. Performance evaluation of computed tomography systems: summary of AAPM task group 233. *Medical physics.* 2019;46(11):e735-e56.
18. Inoue T, Ichikawa K, Hara T, Ohashi K, Sato K, Kawashima H. Validating computer applications for calculating spatial resolution and noise property in CT using simulated images with known properties. *Radiol Phys Technol.* 2024;17(1):238-47.
19. Solomon J, Lyu P, Marin D, Samei E. Noise and spatial resolution properties of a commercially available deep learning-based CT reconstruction algorithm. *Medical Physics.* 2020;47(9):3961-71.
20. Miura Y, Ichikawa K, Fujimura I, Hara T, Hoshino T, Niwa S, et al. Comparative evaluation of image quality among different detector configurations using area detector computed tomography. *Radiological Physics and Technology.* 2018;11(1):54-60.

第 98 回画像部会

「パラダイムシフト?! 臨床画像評価の未来」

臨床画像評価のための物理評価

東北大学病院 診療技術部放射線部門 **小野寺 崇**

1. はじめに

近年の X 線撮影では検出器として高性能の Flat Panel Detector (FPD) が用いられており、従来にくらべて高画質な画像を提供することが可能となっている。また、画像処理においても以前から用いられているエッジ強調・ノイズ低減に加え、人工知能技術を用いたノイズ低減処理なども提供されており、今後も X 線画像の高画質化は進行すると推測される。

医療の発展とともに X 線撮影の撮影目的も多様化している。病変の検出だけでなく、デバイスの位置確認や近年では動態撮影による機能評価も行われており、目的達成に必要なとされる線量・画像処理が異なることは明白である。臨床画像評価を行う前に、撮影目的の理解・画像処理特性の把握が非常に重要になると考える。

前述した技術革新により、当然ながら撮影線量低減は可能である。しかし、その根拠が曖昧なまま線量低減を目指しているケースが散見される。Computed radiography (CR) システム使用時の撮影条件を流用する、あるいは感覚的に線量低下を図っているなど撮影条件の最適化には依然として課題が残る。また、画像処理においても特性を十分に理解しないまま使用し、その実力を存分に発揮できていないということはないだろうか。

本稿では検出器特性把握のための物理評価、撮影ジオメトリ・画像処理を含む臨床画像評価のための物理評価に大別し、それぞれ概説していく。

2. 検出器特性把握のための物理評価

検出器の評価については画像部会主催のセミナーにおいても長い期間をかけて取り扱ってきた。X 線像は検出器の変遷とともに高画質化が進んできたといっても過言ではなく、検出器の評価は非常に重要である。過去の画像通信でも非常に多くの記述があるだろうが、今一度、検出器の物理特性について復習する。

2-1. 入出力特性

デジタルシステムの入出力特性は露光量と raw データのデジタル値の関係を示している (Fig.1)。

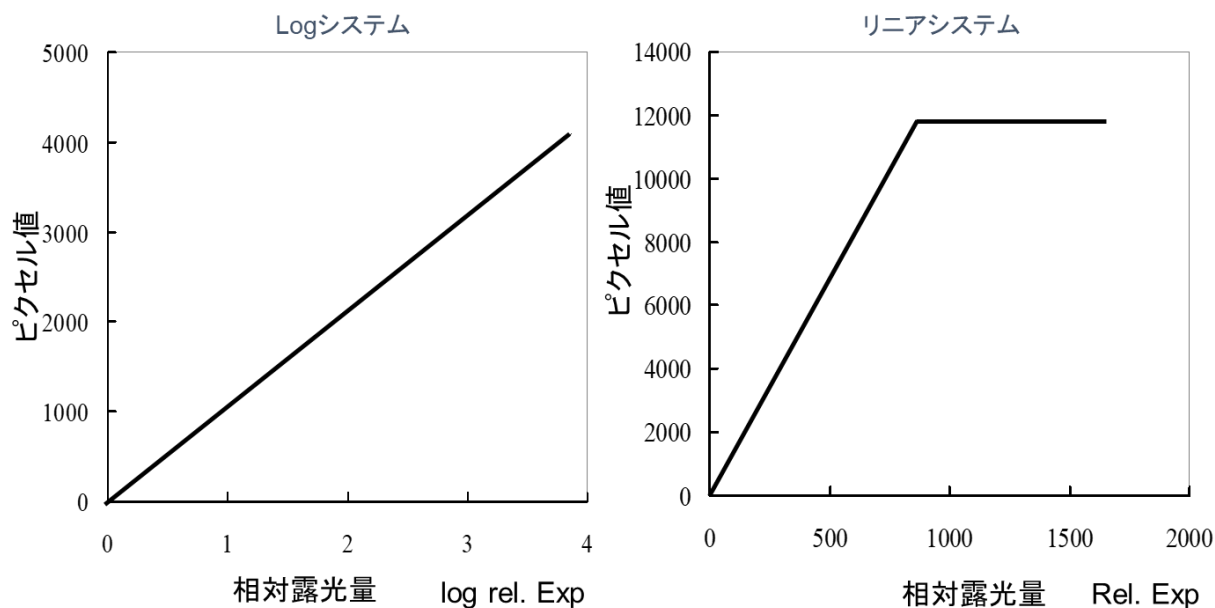


Fig.1 デジタルシステムの入出力特性

入出力特性からは①システムの直線性、②ダイナミックレンジ、③システムコントラストなどの情報が得られる。また、後述する presampled MTF、Normalised Noise Power Spectrum (NNPS)を算出する際の有効露光量変換のツールとしても使用するため、測定は必須である。

2-2. Presampled MTF

システムの解像特性は、出力画像の鮮鋭性を決定する因子である。現代の放射線画像における解像特性の評価法には空間周波数領域で評価を行うレスポンス関数である変調伝達関数 (Modulation Transfer Function: MTF)を用いることが一般的となっている。

デジタルシステム「検出器」の解像特性評価においては、位置不変性の問題をクリアするため presampled MTF (Fig.2)を用いる。Presampled MTF はサンプリング以前のアナログ成分を評価しており、デジタルシステムの解像特性を表す代表的な指標である。

Presampled MTF へ影響を及ぼす因子として、検出器の種類 (直接変換方式 FPD、間接変換方式 FPD、イメージングプレートなど)、サンプリング間隔などがある。また、デジタルシステムは様々な構成要素から成り立っており、画像処理部や画像表示部も含め総合的に評価する場合は、各部の MTF を測定するなどオーバーオール観点から検討しなくてはならない。

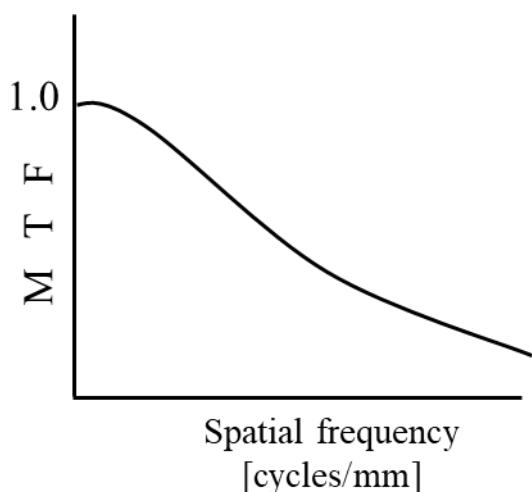


Fig.2 Presampled MTF

2-3. NNPS

X 線画像の粒状性の物理評価項目として、RMS 粒状度や NNPS (Fig.3) がある。RMS 粒状度は均一画像の標準偏差により信号値のばらつきを示した値であるが、いわゆるノイズの全体量を表すことしかできない。

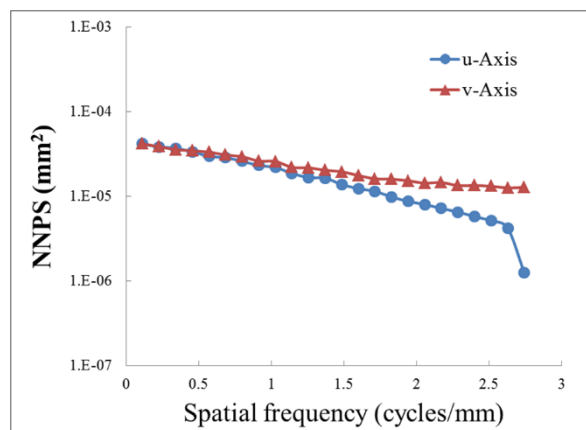


Fig.3 NNPS

一方、Fig.3 に示すように NNPS は presampled MTF と同様に横軸が空間周波数で表すことができる。つまり、NNPS は病変検出を妨害するノイズについて「量」と「質」の両面から評価することができる。

検出器の NNPS は臨床画像評価をするうえで非常に重要である⁽¹⁾。病変検出が主な撮影目的である場合を想定する。検出すべき病変は必ずある程度の大きさをもつため、信号スペクトルは必ず低周波数側に集中する。従って、NNPS の低周波数域が低コントラスト検出能を規定するといっても過言ではない。後述するタスクベースによる評価が臨床画像の NPS にも適用可能になるのであれば、視覚評価の結果との不一致問題も無くなるのではと期待している。

3. 臨床画像評価(物理評価)

デジタル X 線画像には画質向上を目的に様々な画像処理が加えられている。前述したとおり、検出器の評価には presampled MTF や NNPS が用いられるが、これらは画像の線形性が成立している前提で算出されるものであるため、非線形な画像処理が加えられた臨床画像の評価には適さない。近年、X 線 CT 分野において評価対象に類似した被写体を用い、実臨床の撮影条件・画像処理にて解析画像を得るタスクベースによる評価手法が提案され⁽²⁾、X 線画像分野においても多くの研究者が踏襲する形で臨床画像の解像特性評価を試みている。

タスクベースによる解像特性評価には circular edge 法が用いられる。被写体が円柱形であることから例えば胸部 X 線像の解像特性を評価する場合は、腫瘍の性状評価や肋骨抽出などが対象となるであろう。著者自身もこの手法を用いて胸部 X 線画像の各条件(被写体厚の違い・画像処理の有無)による解像特性を評価し、散乱性補正処理を加えた場合、通常画像処理(周波数処理、ダイナミックレンジ圧縮処理のみ)に対して低周波数領域で大きく解像特性が向上することを確認している(Fig.4)。

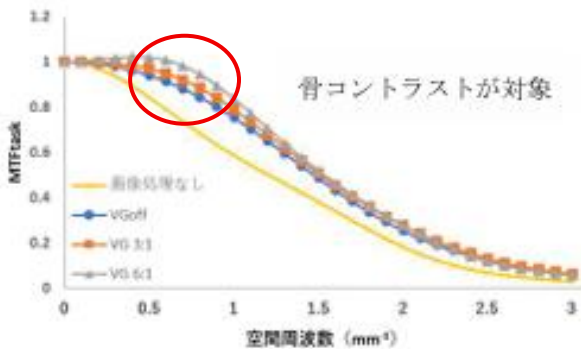


Fig.4 胸部画像の解像特性(骨コントラスト対象)

4. 臨床画像評価の実際

我々、診療放射線技師は画像を提供する立場であるが、多くの施設において臨床画像を活用・読影しているのは医師である。医師が知りたい情報が画像に含まれているか、あるいは認識できるかを評価することは非常に重要であり、本来の臨床画像評価とは、臨床における診断能や有用性を調査することなのだと思える。

ここまで物理評価のことを中心に記述してきたが、やはり臨床画像評価には視覚評価は欠かせない。物理評価は線量設定の根拠作り・画像の定量的評価に用い、臨床的有用性の評価には視覚評価が肝となる。具体的には Receiver Operating Characteristics (ROC) 解析と統計解析ということになるが、詳細は後述される先生方の卓越した良文を参照されたい。

5. さいごに

本稿では検出器の物理評価、臨床画像の物理評価について概説した。X 線画像における臨床画像評価においては診断能(診断の正確さ)もさることながら、撮影線量についても評価しなくてはならない。被ばくと画質の最適化を念頭に、最適化画像の診断能をより良くする発想が臨床画像を向上させるものだ結論づけたいと思う。

参考文献

- (1) 森 一生. 近年の X 線 CT 画像の非線形的特性と画質の物理評価について. 東北大医保健学科紀要 22(1):7-24. 2013
- (2) Richard S, Husarik DB, Yadava G, et al. Towards task-based assessment of CT performance: system and object MTF across different reconstruction algorithms. Med Phys 2012; 39(7): 4115-4122.

ゼロから学ぶ ROC 解析

～判断基準とカットオフ値で理解する「観察者実験」の仕組み～

西南学院大学 情報処理センター 馬場 祥吾

1. はじめに

医療画像は体内の生体情報を画像化し、その画質が診断精度に大きく影響する。そのため、鮮鋭度、粒状性といった物理評価に加え、視覚評価を併せて行う場合が多い。視覚評価は定量性と再現性に欠けるといった問題があったが、Receiver Operating Characteristic (ROC) 評価は心理的、主観的評価でありながら、客観性も兼ねた画質評価法として確立された。ROC 解析は、観察者の刺激に対する応答特性や Artificial Intelligence (AI) などのコンピュータシステムより得られた予測結果を ROC 曲線や ROC 曲線下面積である Area Under the Curve (AUC) で定量化する方法である。一般的に、ROC 曲線が左上 (FPF, TPF) = (0.0, 1.0) へ近づくほど、つまり AUC が 1.0 に近いほど、診断能が高いことは広く知られているが、その理論的背景にある信号と雑音の確率分布 (両正規分布) やカットオフ値に基づく判断基準を正しく理解しなければ、誤った解釈につながる可能性がある。

本稿は、2024 年に発行された「画像通信 Vol. 42 No. 1 (通巻 92)」に掲載された「ROC 解析の理論と AUC による統計的検定」の内容を改訂したものであり、今回は、両正規分布とカットオフ値に基づく ROC 曲線の形成理論について加筆を行った。本稿を通じて、ROC 評価の理論的背景について、より深く理解していただけることを期待する。

2. 信号検出理論と ROC 解析について

信号検出理論 (Signal Detection Theory) は、雑音の中から信号を検出する際の観察者の決定行動を分析するものであり、心理学などの様々な分野へ応用されてきた。ROC 解析は、レーダーシステムが飛来する敵の航空機を鳥などの群れと判別する際の性能を評価するために開発された信号検出理論であり、1960 年以降に放射線医学分野に導入され、その応用が始まった。放射線医学分野では、以下の目的に対して ROC 解析が用いられる。

- ・モダリティ間の診断能比較
- ・検査法または手技間の診断能比較
- ・観察者の診断能評価
- ・新しいモダリティ、検査法、手法の評価
- ・最適なカットオフ値の推定
- ・コンピュータ支援診断システム (Computer-Aided Diagnosis : CAD) の性能評価

特に ROC 解析の利点は、病変を信号と見なすことで、異なったモダリティ間の比較が可能となった点であり、高度医療が進み様々なモダリティが医療に利用される昨今においては ROC 解析への期待は大きい。

また近年では、従来の観察者実験の評価にとどまらず、ディープラーニングや機械学習といった AI や CAD の性能評価へと適用範囲が拡大しており、ROC 解析の重要性は以前にも増して高まっている。

2-1. ROC 解析の実験手順

ROC 解析の観察者実験手順は、以下の通りとなる。

- 1) 実験の目的を明確にする：ROC 実験に用いられる画像は、ROC 解析を用いて何を評価するかによって決定される。そのため、統計的に有用な画像枚数を確保するためにも、明確な実験目的を設定し、臨床画像もしくは擬似画像を準備しなければならない。
- 2) 実験の計画：観察試料の作成と選択に加えて観察者の種別と選択を行う。観察試料の枚数は統計的に多いほど有利となるが、観察者の目の疲労も考慮し、信号あり画像が 150～250 枚であれば、統計的誤差も 2～3%に抑えることができる。また、観察試料の難易度に偏りがあった場合、観察者の反応は正規分布からかけ離れたものとなってしまう、統計的な有意性を示すことが困難となる。したがって、仮に観察者の判断基準が 5 段階評価で正規分布に従うとすれば、Fig.1 のような観察試料の難易度を調整した選出、もしくは作成が観察者実験を成功に導くために必要となる。また、観察者の選択では観察者個々の能力を把握して、目的に応じた選択が理想的である。観察者の人数も多ければ、統計的精度が向上するため、少なくとも 5 名程度で 1 名の観察者に対して複数回の観察実験を行う必要がある。
- 3) 学習：観察者に対しては事前にどのような目的の実験かを理解するために、信号の大きさや難易度、加えて、観察者が二者択一的な回答となることを避け、強い確信度の回答が全体の 10%程度になるように説明する。また、学習用画像の選出では、観察者に無用なバイアスを与えないよう注意深く選出する。
- 4) 実験の実施：学習や実験条件、実験環境などが適正に整った下で実験を開始する。
- 5) 実験結果の解析：ROC 曲線より、カットオフ値や AUC の算出、また、AUC による統計的検定などの分析を行う。

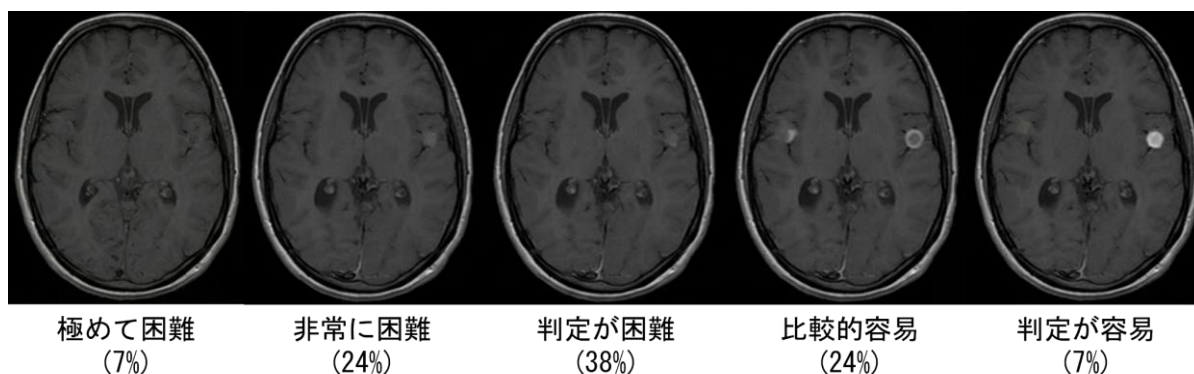


Fig.1 信号強度を調整した難易度別観察試料作成の例：頭部 MRI 画像は、ImageJ のサンプルデータである「MRI-Stack」から抽出した 1 スライスを使用した。この画像に対し、生成 AI を用いて信号強度を段階的に変化させた擬似腫瘍を合成したものである。なお、括弧内の数値は、観察試料の総枚数に対する各画像の含有率の目安を示している。

2-2. 刺激-反応行列

ある画像を観察したとき、観察者にいくらかの幅があったとしても「陽性」と「陰性」のいずれかの応答を示す。また画像が「疾病あり」と「疾病なし」のいずれかとすれば、Table 1 に示すような2×2の刺激-反応行列を示す。まず、診断結果が「疾病あり」で正しく「陽性」と判断した場合は真陽性 (True Positive : TP)、誤って「陰性」と判断した場合は偽陰性 (False Negative : FN) と呼ばれる。一方、診断結果が「疾病なし」で正しく「陰性」と判断した場合は真陰性 (True Negative : TN)、また、誤って「陽性」と判断した場合は偽陽性 (False Positive : FP) と呼ばれる。

Table 1 刺激-反応行列の対応表

刺激	反応	陽性	陰性
	信号あり (疾病あり)	真陽性 True Positive (TP)	偽陰性 False Negative (FN)
信号なし (疾病なし)	偽陽性 False Positive (FP)	真陰性 True Negative (TN)	

ここで、「陽性」と「陰性」を区別する読影や検査の信頼度は、一般的に感度 (Sensitivity) と特異度 (Specificity) で定義される。感度は「疾病あり」画像に対して診断結果が「陽性」である割合を指し、「疾病あり」画像の総枚数は TP と FN を足した数となるため、感度は式 (2.1) で算出できる。

$$\text{感度} = \frac{TP}{TP + FN} \quad (2.1)$$

対して、特異度は「疾病なし」画像に対して診断結果が「陰性」である割合を指し、「疾病なし」画像の総枚数は TN と FP を足した数となるため、特異度は式 (2.2) で算出できる。

$$\text{特異度} = \frac{TN}{FP + TN} \quad (2.2)$$

また、Table 1 の刺激-反応行列において、感度は真陽性率 (True Positive Fraction : TPF)、特異度は真陰性率 (True Negative Fraction : TNF) に対応する。したがって、 $1.0 - \text{感度} = \text{偽陰性率}$ (False Negative Fraction : FNF) であり、 $1.0 - \text{特異度} = \text{偽陽性率}$ (False Positive Fraction : FPF) になる。

しかし、感度と特異度は、「陽性」と「陰性」を識別する上で、観察者の固有の能力を表現しているが、それがどの程度の割合で正しく判断されるかは不明である。そこで、適中度は診断結果が実際に正しかった割合を示す指標である。診断結果が「陽性」で観察試料が「疾病あり」であった割合を陽性適中率 (Positive Predictive Value : PPV)、診断結果が「陰性」で観察試料が「疾病なし」であった割合を陰性適中率 (Negative Predictive Value : NPV) という。PPV は式 (2.3)、NPV は式 (2.4) で算出される。

$$\text{陽性適中率(PPV)} = \frac{TP}{TP + FP} \quad (2.3)$$

$$\text{陰性適中率(NPV)} = \frac{TN}{FN + TN} \quad (2.4)$$

2-3. ROC 曲線とカットオフ値

評定実験の方法には、観察者が各試料を判断基準ごとの数段階のカテゴリーに分類する評定確信度法と、観察者が各試料に対して判断基準に応じたスコアを付ける連続確信度法がある (Fig. 2).

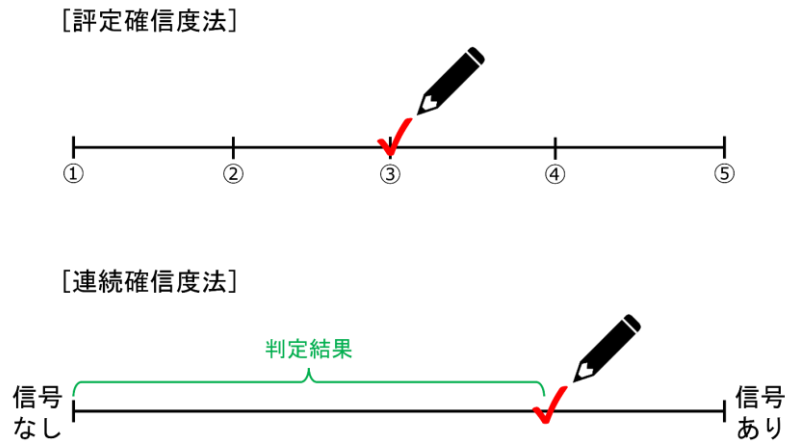


Fig. 2 評定確信度法と連続確信度法

評定確信度法は Fig. 2 に示すように数段階のカテゴリーを用いて評価する方法であり、「①信号 (疾病) は存在しない」、「②信号 (疾病) は存在しないだろう」、「③どちらともいえない」、「④信号 (疾病) は存在するかもしれない」、「⑤信号 (疾病) は存在する」のように 5 段階のカテゴリーを設ける場合が多い。それに対して、連続確信度法は、カテゴリーを設定せずに Fig. 2 に示すような連続スケール (左端: 「信号 (疾病) なし」、右端: 「信号 (疾病) あり」) を利用し、観察者には評定結果を自由にマークしてもらい、確信度はスケールの左端からマークまでの距離を求めることで算出する。

Table 2 は、ある検査診断能の調査目的で、確実に脳腫瘍と分かっている患者 100 人と確実に健常であることがわかっている 100 人に画像検査を実施し、1 人の観察者が「①腫瘍は絶対ない」、「②腫瘍は存在しないだろう」、「③どちらともいえない」、「④腫瘍は存在するかもしれない」、「⑤腫瘍は必ず存在する」の評定確信度法 (5 段階評価) を行った結果の例である。

Table 2 脳腫瘍の判定結果

	① 絶対ない	② 存在しない だろう	③ どちらとも いえない	④ 存在する かもしれない	⑤ 必ず存在する	計
脳腫瘍	3	0	20	37	40	100
健常	30	7	36	27	0	100

カットオフ値とは、定量データを区切るために使用される基準値である。Table 2 に対して、評価基準ごとにカットオフ値を設定した表を Table 3～6 に示す。{①}を「腫瘍なし」、{②③④⑤}を「腫瘍あり」のカットオフ値で設定すれば Table 3 のような表となり、{①②}を「腫瘍なし」、{③④⑤}を「腫瘍あり」のカットオフ値で設定すれば Table 4 のような表となる。Table 5 および 6 も同様にカットオフ値を設定して得られた表となる。また、Table 3 から感度と特異度を算出すると感度は $97/(97+3)=0.97$ 、特異度は $30/(70+30)=0.30$ となる。同様に、Table 4 の感度は $97/(97+3)=0.97$ 、特異度は $37/(63+37)=0.37$ 、Table 5 の感度は $77/(77+23)=0.77$ 、特異度は $73/(27+73)=0.73$ 、Table 6 の感度は $40/(40+60)=0.40$ 、特異度は $100/(0+100)=1.0$ となる。一般的に、感度と特異度はトレード・オフの関係にあり、感度を高くしようとすると特異度は低くなり、特異度を高くしようとすると感度は低くなる。

Table 3 {①}vs.{②③④⑤}

	陽性	陰性	計
脳腫瘍	97	3	100
健常	70	30	100

Table 4 {①②}vs.{③④⑤}

	陽性	陰性	計
脳腫瘍	97	3	100
健常	63	37	100

Table 5 {①②③}vs.{④⑤}

	陽性	陰性	計
脳腫瘍	77	23	100
健常	27	73	100

Table 6 {①②③④}vs.{⑤}

	陽性	陰性	計
脳腫瘍	40	60	100
健常	0	100	100

ROC 解析に用いられる ROC 曲線は、横軸に FPF (1.0-特異度)、縦軸に TPF で表され、「腫瘍あり」と「腫瘍なし」の境界を決定する判断基準を変化させた場合の感度と特異度をプロットしたグラフである。Fig. 3 は、Table 3～6 のカットオフ値ごとに算出された FPF と TPF をプロットした ROC 曲線である。つまり、Table 3 の感度と特異度の結果より得られる ROC 曲線上の座標点は、 $(FPF, TPF) = (0.97, 0.70)$ となり、同様に、Table 4 の ROC 曲線上の座標点は、 $(FPF, TPF) = (0.97, 0.63)$ 、Table 5 の ROC 曲線上の座標点は、 $(FPF, TPF) = (0.77, 0.27)$ 、Table 6 の ROC 曲線上の座標点は、 $(FPF, TPF) = (0.40, 0.0)$ となる。また、Fig. 3 には $(0.0, 0.0)$ と $(1.0, 1.0)$ が追加してあるが、これは $\{\phi\}$ vs $\{①②③④⑤\}$ と $\{①②③④⑤\}$ vs $\{\phi\}$ でカットオフ値を設定した座標であり、強い確信度のカットオフ値(「①腫瘍は絶対ない」という確信度は判定基準： $-\infty$ 、「⑤腫瘍は必ず存在する」という確信度は判定基準： $+\infty$)といえる。

観察者実験で得られる ROC 曲線の AUC は 0.5～1.0 の範囲をとり、面積が大きいほど精度が高いと評価できる。また、検査結果による AUC の評価判定の基準は 0.5～0.7 が低い精度、0.7～0.9 が中程度、0.9～1.0 が高い精度と評価されることが多い。しかし、診断目的や診断の難易度によって AUC に対する評価基準は変化するため、考察の際は十分な議論が必要である。

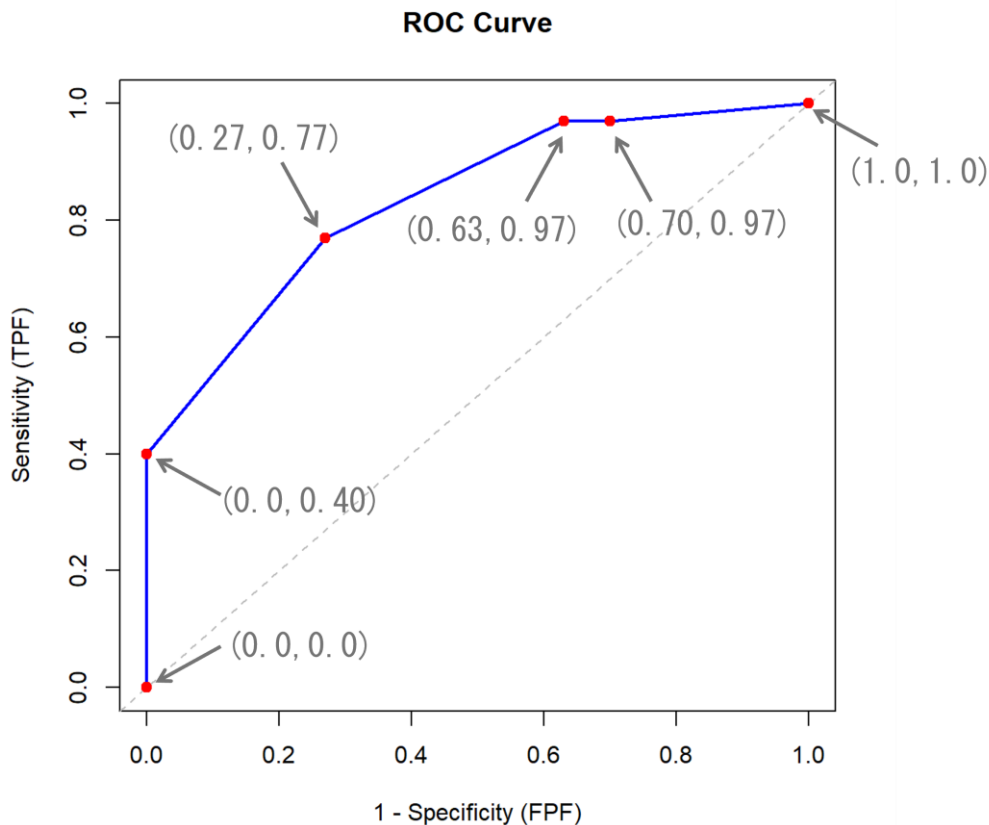


Fig. 3 Table 3~6の結果より作成したROC曲線

2-4. 両正規分布と判定基準の設定によるROC曲線

前述の通りROC曲線はFig. 3に示すような横軸にFPF、縦軸にTPFを示すグラフであり、ROC曲線が左上 $(FPF, TPF) = (0.0, 1.0)$ へ近づくほど、診断能が高いことは広く知られている。本項では、よりROC評価に関する理解を深めるために、観察者の心理的ゆらぎとROC曲線に関する理論的背景について解説する。ROC曲線は、観察者の「疾病あり(信号+雑音)」画像に対する心理的ゆらぎと、「疾病なし(雑音)」画像に対する心理的ゆらぎが、それぞれ独立した正規分布に従うという両正規分布モデルの仮定に基づいて構築されている。正規分布は、代表的な連続型の確率分布で、自然界や人間社会の中の数多くの現象に当てはまり、統計学においては最も重要な確率分布の1つと考えられている。正規分布を仮定する理由は、中心極限定理に基づいて独立確率変数の総和は、変数が無限に近づくとき正規分布に収束するという事実と、変数の処理操作が容易であるという理由である。したがって、観察者実験において、各画像に対する観察者の心理的ゆらぎが、正規分布に従うと仮定すると、Fig. 4に示すような2つの山を持つ分布となる。ここで、Fig. 4の横軸は、観察者の心理的判定基準を示しており、実際の実験では、この連続的な判定基準を評定確信度法である「①信号(疾病)は存在しない」、「②信号(疾病)は存在しないだろう」、「③どちらともいえない」、「④信号(疾病)は存在するかもしれない」、「⑤信号(疾病)は存在する」までの5段階のカテゴリーに区切って回答しており、これらが潜在的に正規分布上に配置されていると仮

定する.

[疾病あり (信号+雑音)] 画像に対する観察者の心理的ゆらぎに対する正規分布を $f_s(x)$, [疾病なし (雑音)] 画像に対する観察者の心理的ゆらぎに対する正規分布を $f_n(x)$ とすると, $f_s(x)$ は式 (2.5), $f_n(x)$ は式 (2.6) で算出される.

$$f_s(x) = \frac{1}{\sqrt{2\pi}\sigma_s} \exp\left(-\frac{(x - \mu_s)^2}{2\sigma_s^2}\right) \quad (2.5)$$

$$f_n(x) = \frac{1}{\sqrt{2\pi}\sigma_n} \exp\left(-\frac{(x - \mu_n)^2}{2\sigma_n^2}\right) \quad (2.6)$$

以上の2つの正規分布より ROC 曲線が通過する座標点を算出する. 2つの独立した正規分布より, ROC 曲線が通過する座標点 (FPF, TPF) について FPF は $f_n(x)$, TPF は $f_s(x)$ よりカットオフ値を用いて算出できる. まず, $f_n(x)$ は [疾病なし (雑音)] 画像に対する確信度の分布を表しており, $f_n(x)$ の面積は FPF+TNF=1.0 となる. カットオフ値で FPF と TNF を分割する場合, カットオフ値を超える確信度 (カットオフ値より右側) の高い領域の面積を [疾病なし (雑音)] 画像に対して「疾病 (信号+雑音) あり」と判断した FPF と考えることができる. 同様に, $f_s(x)$ の面積は TPF+FNF であるため, カットオフ値を超える確信度の高い領域の面積を [疾病あり (信号+雑音)] 画像に対して「疾病 (信号) あり」と判断する TPF と考えることができる. したがって, ROC 曲線はカットオフ値を変化させたときの (FPF, TPF) を座標としてプロットし, それらの点を結ぶことで描かれる軌跡であるということが理解できる.

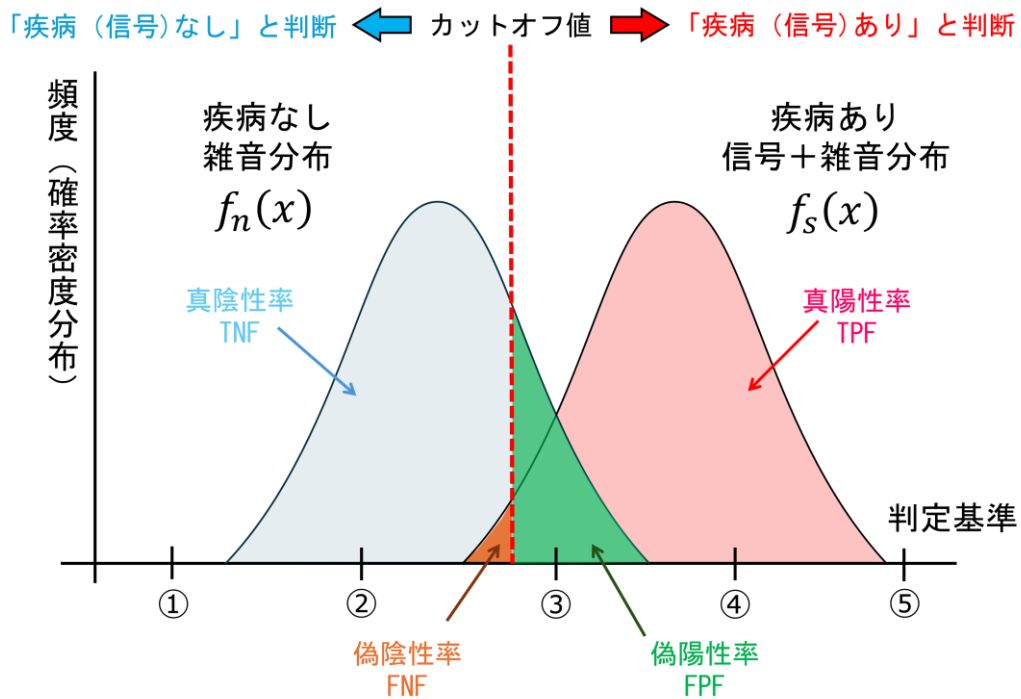


Fig. 4 両正規分布とカットオフ値の関係

2-5. 両正規分布間の距離と ROC 曲線の関係性について

ROC 曲線の軌跡は、両正規分布の近接度に依存し、両正規分布間の距離が近いほど AUC は低下する。Fig. 5 に両正規分布間の距離と ROC 曲線の関係を示す（赤線の正規分布：疾病あり [信号+雑音分布] 画像に対する観察者の心理的ゆらぎ、青線の正規分布：疾病なし [雑音分布] 画像に対する観察者の心理的ゆらぎ）。両正規分布の距離が遠いほど、AUC は 1.0 に近づき、逆に距離が近いほど 0.5 に収束する。両正規分布が完全に分離している場合、誤分類である FPF および FNF は発生しないため、カットオフ値の移動によって得られる (FPF, TPF) の座標点の軌跡は、FPF=0.0 のまま TPF=1.0 まで垂直に上昇した後、FPF=1.0 へと推移する理想的な矩形を描く。対照的に、両正規分布の重なりが増すと FPF および FNF の発生確率が増加するため、両正規分布の重複の割合に応じて (FPF, TPF) は変動し、AUC は減少する。両正規分布が完全に一致した場合、カットオフ値の FPF と TPF は常に等しい値をとるため、ROC 曲線は (0.0, 0.0) と (1.0, 1.0) を結ぶ対角線となる。これは、観察結果がランダムな判定と同義であることを意味している。

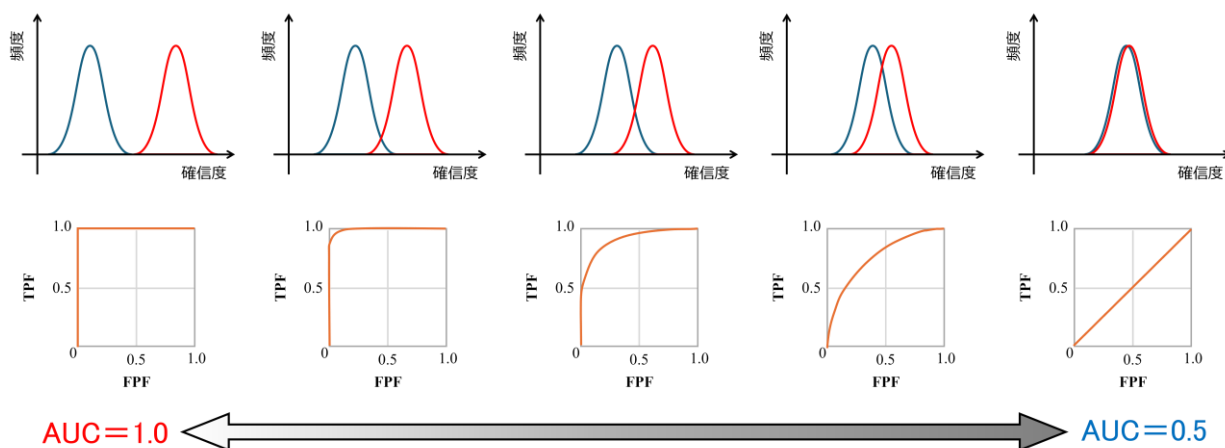


Fig. 5 正規分布の距離と ROC 曲線の変化

3. 視覚評価による統計的検定方法

ROC 解析は、観察者の読影に基づく評価法であるため、評価結果は必ずしも再現できるとは限らない。そのため、ROC 解析では統計的に解析する必要がある。ここでは、2つの診断能や検出能の差を評価する2群間の統計的検定方法について解説する。

3-1. 2群間のパラメトリック検定の手順

2群間のパラメトリック検定の手順について述べる。なお、検定方法の選択については後述する。

- 1) 有意水準の設定：2つのデータ間に差があるかどうかの判断基準に有意水準 (Significance level) を設定して判定する。有意水準 α は通常 5% または 1% に定められるが、医学統計では 5% が一般的に用いられる。
- 2) 仮説の設定：「2群間に差がある」という仮説を証明するためには、その逆の「2群間に差がない」

という仮説が矛盾することを証明する背理法を用いる。まず、「2 群間に差がない」という仮定 (H_0 : 帰無仮説) を設定し、本来の「2 群間に差がある」という仮説 (H_1 : 対立仮説) を設定する。

- 3) 統計量の計算: 2 群間の平均値を \bar{x}_1 , \bar{x}_2 とすると, 2 群間の平均の差 $\bar{x}_1 - \bar{x}_2$ を統計量とする。この統計量は 2 群間の差を表すものであればなんでも良く, 検定方法の違いは統計量の違いとなる。
- 4) 確率を求める: H_0 の下で統計量 $\bar{x}_1 - \bar{x}_2$ の生じる確率を求め, 差の有意性を判定する。p 値は, $\bar{x}_1 - \bar{x}_2$ の標準誤差で標準化すると, 自由度 $df = n_1 + n_2 - 2$ の t 分布に従うことを利用して求める。

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

ここで, s は両群の分散 s_1^2 , s_2^2 から求めた合成分散 s^2 の平方根として求まる。

$$s = \sqrt{\frac{s_1^2(n_1 - 1) + s_2^2(n_2 - 1)}{n_1 + n_2 - 2}}$$

また, $s \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$ が $\bar{x}_1 - \bar{x}_2$ の標準誤差にあたる。

- 5) 判定: t 分布表より, 自由度 df , 有意水準 α の t 値 (t_α) を調べ, 標本の t 値と比較する。
 $|t| \leq t_\alpha$ つまり $p \geq \alpha(0.05)$ のとき, 有意差があるとはいえない (H_0 を棄却できず, 判定保留)
 $|t| > t_\alpha$ つまり $p < \alpha(0.05)$ のとき, 有意差あり (H_0 を棄却し, H_1 を採用)

3-2. 検定方法の選択

検定方法には大きく分けて, パラメトリック検定とノンパラメトリック検定があり, 検定前に 2 群間の正規性と等分散性を確認する必要がある。統計学的に正規性を検定する方法はシャピロ・ウィルクの検定 (Shapiro-Wilk test), もしくは平均値と標準偏差が既知の場合には K-S 検定 (Kolmogorov-Smirnov test) が主に用いられ, 帰無仮説に「データが正規分布に従う」と仮定する。つまり, p 値が有意水準 α よりも大きい ($p \geq 0.05$) 場合, 帰無仮説を棄却できないことから, 「正規分布に従っている」と仮定する。また, 統計学的に等分散性を検定する方法は, F 検定やパートレット検定 (Bartlett test) などが用いられ, 帰無仮説に「データの分散が等しい (等分散)」と仮定する。正規性の検定と同様に p 値が有意水準 α より大きい ($p \geq 0.05$) の場合, 帰無仮説を棄却できないことから, 等分散性を仮定する。

対応なし 2 群間の検定とは, 異なった 2 群の対象から得られたデータ間の統計的検定方法である。2 群のデータの分布が正規分布に近似 (パラメトリック) し, 分散が等しい (等分散) 場合, Unpaired t 検定 (Student's t-test) を用いることができる。また, 正規性はあるが, 等分散性とみなせない場合, 自由度の値をコントロールして p 値と棄却域を調整する Welch の t 検定を用いることができる。データの正規性を確保できない場合, 母集団の分布に対して, 特別な仮定をおく必要のないノンパラメトリックな検定であるマン・ホイットニーの検定 (Mann-Whitney test) が用いられる。なお, マン・ホイットニーの検定はウィルコクソン順位和検定 (Wilcoxon rank-sum test) と統計学的に等価な検定である。

対応あり 2 群間の検定とは, 同じ対象から得られた 2 群のデータ間における統計的検定方法であり, 各群のデータの分布が正規分布に近似 (パラメトリック) し, 分散が等しい (等分散) 場合, paired t 検

定が用いられる。また、独立2群間の検定と同様、データの正規性を確保できない場合、母集団の分布に対して、特別な仮定をおく必要のないノンパラメトリックな検定であるウィルコクソン符号付き順位和検定 (Wilcoxon signed-rank test) が用いられる。検定方法の選択フローはFig.6に示す。

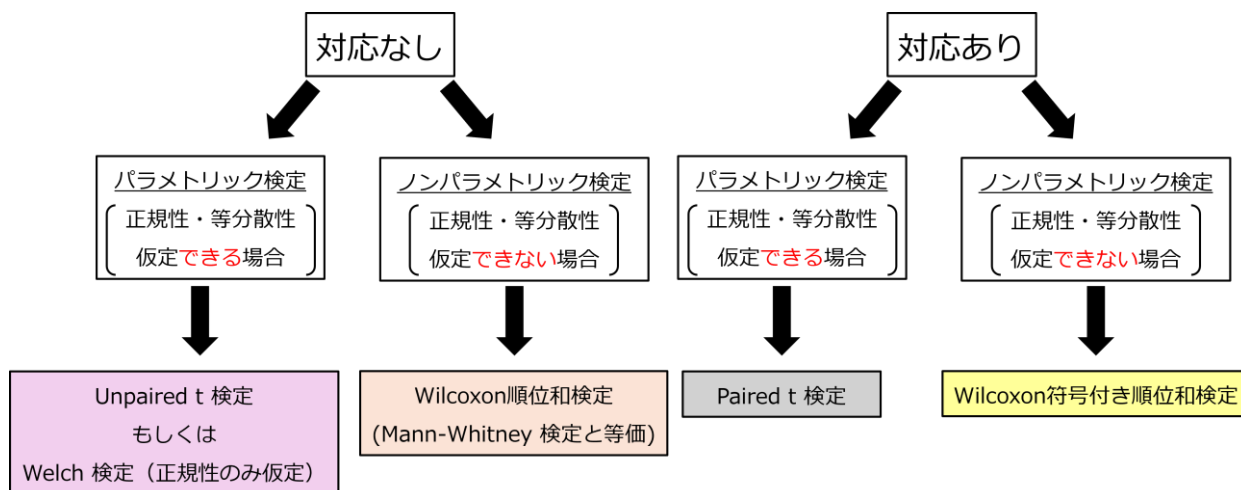


Fig.6 検定方法の選択フローチャート

4. おわりに

本稿では、信号と雑音の確率分布（両正規分布）やカットオフ値に基づく判断基準、ならびにROC解析における統計学的有意差検定について解説した。本稿で述べたROC解析や統計的検定法は、観察者実験のみならず、様々な研究成果を定量的に評価するうえで極めて有用な指標となる。本稿が、医用画像をはじめとする医学分野の研究や学術報告において、その一助となれば幸いである。

最後に、本稿の執筆に際し、学校法人原田学園 鹿児島医療技術専門学校 言語聴覚療法学科の小牧祥太郎先生（医学博士）には、最終稿の丁寧なご査読と有益なご助言を賜った。ここに心より感謝の意を表す。

5. 参考文献

- ・ 馬場 祥吾 (2024) : ROC解析の理論とAUCによる統計的検定, 画像通信 Vol.42 No.1 (通巻92).
- ・ 白石 順二, 他 (2020) : 標準 医用画像の視覚評価法, オーム社.
- ・ 藤田 広志, 他 (1994) : ROC解析の基礎と応用 日本放射線技術学会専門委員会 デジタル画像のROC解析検討班編, 社団法人 日本放射線技術学会 第50回総会記念出版.
- ・ 福士 政広, 他 (2020) : 実践! 医用画像情報学 基礎から実験・演習まで, メジカルビュー社.
- ・ 柳川 堯, 他 (2010) : バイオ統計の基礎, 近代科学社.
- ・ 市原 清志 (1990) : バイオサイエンスの統計学, 南江堂.
- ・ 臨床評価研究会 (ACE) 基礎解析文科会 (2017), 新版 実用 SAS 生物統計ハンドブック, サイエンスティスト社.

臨床画像の評価および読影における課題への取り組み

福岡大学筑紫病院 放射線部 田畑 成章

1. はじめに

医用画像の評価は、MTF や NPS などを用いた客観的評価と、ROC 解析や一対比較法などを用いた臨床条件に近い主観的評価の 2 種類に大別される。2 種類の評価を実施することで、従来法と比較した場合の改善案や新手法の利点・欠点を明らかにすることができる。これらの評価を実施するためには、臨床における現状を把握しておくことが最も重要である。実際の臨床画像は、主治医または放射線科医によって読影が行われている。読影の現状を理解し、それを適切に再現することで、臨床画像の妥当な評価が可能となる。加えて、読影環境のみならず、読影デバイスや最新ソフトウェアについても理解を深めることで、現状の課題を把握し、新たな研究テーマの創出につなげることが可能である。本稿では、はじめに読影における環境整備の基礎について述べ、次に読影デバイスについて概説し、最後に AI の参入と課題について言及する。

2. 読影環境

読影における周辺環境(以下、読影環境)は、診断の質に直結している。読影環境は、大まかに周囲の環境と読影デバイスに分けることができる。この 2 つを理解することで、実際の臨床現場で有用であるほか、視覚評価などの研究においても正しい評価が可能となる。

読影環境を意識することは主治医や放射線科医の領域と考えられがちであるが、私たち診療放射線技師にも深い関係がある。読影環境を理解することは、後述する STAT 画像所見報告の精度にも影響を及ぼす。本稿では、読影室における環境整備、読影時の画像処理の有用性、診療放射線技師による読影の補助(STAT 画像所見報告)について述べる。

2-1 読影室における環境

読影室において特に重要な要素の 1 つが周囲照明である。現在ではモニタによる読影が行われており、モニタ画面と周囲の明るさの差(コントラスト)を適切に設定することで、眼精疲労の軽減や作業効率の向上を図ることができる。日本において読影室の周囲照明について言及した明確なガイドラインは存在しない。欧米では、米国放射線学会(ACR)、米国医学物理学会(AAPM)、医用画像情報科学協会(SIIM)が共同で策定したガイドライン”ACR-AAPM-SIIM Technical Standard for Electronic Practice of Medical Imaging”¹⁾があり、周囲照明について以下のように示されている。

- ・ 周囲照明を制御し、目の疲労を最小限に抑えるように最適化すべきである
- ・ 最適化は周囲照度を 25～75 lux に設定することで達成できる
- ・ 反射特性の改善されたモニタでは、より明るい周囲照度が可能な場合があるが、考慮すべきである。

lux とは単位面積あたりにどれだけの光が当たっているかを表す単位である。具体例を Table 1 に示す。

Table 1 照度の目安と具体例

照度 (lux)	明るさの目安	具体例
0.1 ~ 0.3	非常に暗い	月明りのみの夜
1	かなり暗い	蝋燭 1 本で 1m 離れた位置
10	薄暗い	夜の街灯下
50	やや暗い室内	映画館の通路
100	落ち着いた室内	家庭の廊下・寝室
300	一般的な作業環境	オフィス・教室
500	やや明るい作業環境	精密作業を実施する作業場

周囲照度は、日々読影を実施する読影室では達成されていることが多いが、主治医が閲覧を実施する外来や診療放射線技師の作業場(撮影室、処理室など)では十分に考慮されていない場合もある。本来は、統一できることが望ましいが、実際の環境では困難である。視覚評価を実施する場合は、これらを考慮して目的の周辺照度を把握し、近づけることで、より臨床に即した条件で実験を行うことができる。

2-2 画像処理

一言に画像処理といっても様々な種類が存在する。本稿における画像処理とは、特別なアプリケーションや機器を必要とせず、画像閲覧ソフトに標準的搭載されている機能を指す。視覚評価を実施するには観察環境のみならず、観察条件についても考える必要がある。実際の読影環境では、画像処理を実施することで、疾患の判別が容易になる場合がある。画像処理による効果の 1 例を Table 2 に示す。

Table 2 画像処理とそれにより得られる効果の例

画像処理	得られる効果
Window Level・Window Width を絞る	コントラスト強調による出血や梗塞の鮮明化 エッジ強調効果によるデバイスや腫瘍の判別
Window Level・Window Width を広げる	Free Air などの判別の容易化 高吸収体の内部把握
画像反転	異常病変の判別容易化

視覚評価の際には、画像処理を制限する一方で、評価目的や対象によっては適切な画像処理を許可することで結果が変わる場合がある。これらを踏まえて観察条件は設定すべきである。

2-3 STAT 画像所見報告

昨今では、「生命予後にかかわる緊急性の高い疾患の所見がある画像」を報告する STAT 画像所見報告が注目されている。STAT 画像所見報告の詳細については、日本医学放射線学会、日本放射線科専門医会・医会、日本放射線技師会の共同編集のガイドライン²⁾を参照されたい。画像所見を見る際には、装置に付属している汎用型モニタを使用する場合がある。しかしながら、本来は医用モニタを使用し、前述した読影環境を構築したうえで確認することが望ましいと考える。緊急性の高い場面も多いが、可能な限り環境を整えて画像を見る習慣をつけるべきである。

3. 読影デバイス

読影デバイスとは高い性能と精度を持つ専門的なコンピュータシステムとディスプレイの総称である。中でも本稿で注目したいのはディスプレイ(医用モニタ)である。医用モニタは一般的なモニタと異なり、Table 3 に示すような違いが存在する。

Table 3 一般的なモニタと医用モニタの違い

項目	一般的なモニタ	医用モニタ
色	カラー	カラー&モノクロ
階調特性	ガンマ 2.2	GSDF
最大輝度	~300 cd/m ²	~1000 cd/m ²
解像度	XGA (1024x768)	1MP (1280x1024)
	UXGA (1600x1200)	2MP (1200x1600)
	Full HD (1920x1080)	3MP (1536x2048)
	4K (3840x2160) など	5MP (2048x2560) など
キャリブレーション	ソフトウェアキャリブレーション	対応(医用画像表示要求に準拠)
輝度安定化回路	一部搭載	搭載
外光(照度)センサー	なし	一部機種に搭載

様々な点で違いがあるが、最大の違いは階調特性にある。ガンマ 2.2 は一般的な階調特性であり、中輝度から高輝度において階調が急激に変化する特徴がある。一般的なモニタを医用画像表示に使用した場合には、血管影や微小骨折などの微細なコントラストの識別を必要とする疾患では検出能が低下する恐れがある。一方、GSDF は、DICOM Part14³⁾で規定された階調であり「人間が知覚可能な輝度変化を関数にしたもの」と定義される。暗部から明部まで知覚的に均等な輝度変化となるように設計されており、微細なコントラストの識別が可能となる。(Fig. 1)

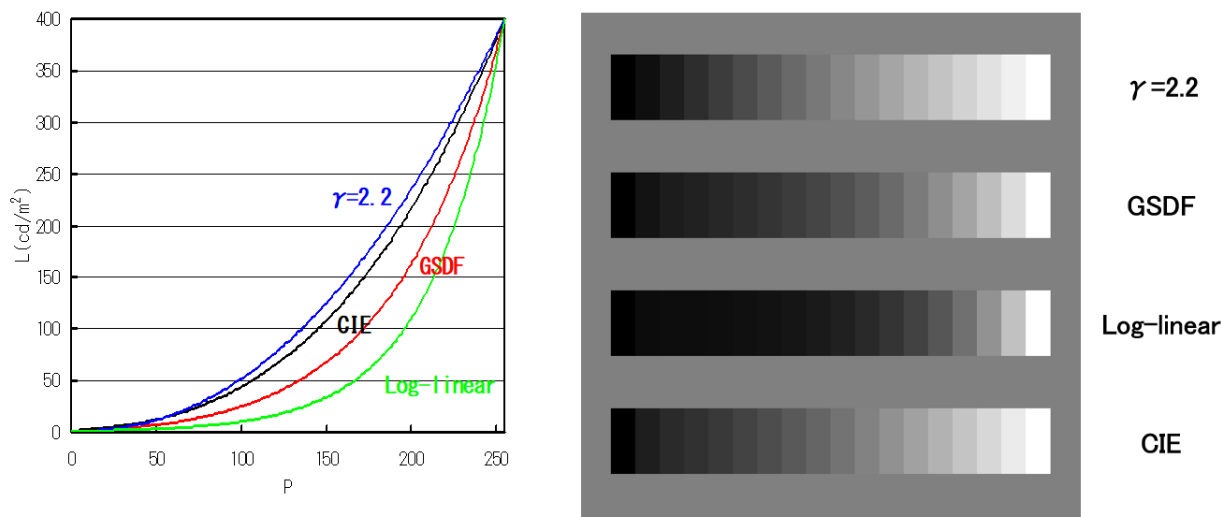


Fig. 1 ガンマ 2.2 と GSDF の違い (EIZO 株式会社 提供)

また、これまで医用モニタは医療機器としては扱われてこなかったが、令和 6 年 7 月より医療機器として告示された(厚生労働省告示第 240 号 - 242 号)。これにより、医用モニタは保守管理の対象となり、精度管理をより適正に行っていかなければならない。これを受けて、品質管理ガイドラインとして新たに JESRA TR-0049 が 2024 年に制定された⁴⁾。しかしながら、現在は移行期間であるため、従来通り JESRA X-0093⁵⁾にて運用された非医療機器として扱うことも可能とされている。このように、医用モニタの特性を把握し、適切に管理することが読影の質の向上につながる。

4. 読影における AI 参入と課題

昨今では、医療の DX 化によって AI の導入が進んでいる。放射線領域においても画像再構成やポジショニング、読影の補助まで様々な領域で活用されている。各ベンダーによって様々な特徴が存在するが、紙面の都合上、詳細は割愛する。本稿では、読影領域における AI(機械学習, 深層学習)の関わりについて述べ、今後の課題について述べる。

まず初めに機械学習と深層学習の違いについて概説する。(Table 4) この 2 つの大きな違いは特徴の抽出が手動もしくは、自動で行うかという点にある。機械学習では特徴を人間が定義することで、定義

に基づいた解答へのプロセスを構築できる。深層学習では、特徴を自動的に抽出し、解答へのプロセスを生成する。どちらも学習データの質や量への依存性や根拠説明の不透明さが課題として挙げられる。また、機械学習では、処理が特定タスクのみであり限定的である点、深層学習では、機械学習以上のデータ量と質が求められ、さらに運用に高いコストが必要である点など、それぞれの課題も存在する。

Table 4 機械学習と深層学習の違い

項目	機械学習	深層学習
特徴	手動設定	自動設定
共通課題	学習データの質や量への依存性 根拠説明の不透明さ	
個別課題	使用範囲が限定的	データ量と質への高い依存性 運用コストの高さ

放射線読影領域への参入は、コンピュータ支援診断(Computer Aided Diagnosis: CAD)に始まる。1960年にはコンピュータを用いた胸部 X 線の正常・異常の自動分類の提案がされている⁶⁾。そこから CAD の研究は進み、1998年には米国企業 R2 Technology 社(現 Hologic 社)の開発した検診マンモグラフィ専用の CAD システムが製品として発売された⁷⁾。日本においては、マンモグラフィの CAD のみ葉機法にて承認されており、第2の意見として読影の補助に使用されている。これまでの CAD の歴史に関しては藤田の文献⁸⁾が分かりやすいため、参照されたい。これまで CAD は機械学習で構築されたものであるが、昨今では深層学習を用いた CAD (AI-CAD) も開発されている。その他、深層学習は臓器のセグメンテーションなどの画像認識から、入力時に医師の次の言葉を予測する自然言語処理まで幅広く活用されている。

放射線領域において AI を活用していくためには、依然として多くの課題がある。中でも質の良い大量の構造化データを揃えることが課題となっている。特に放射線読影レポートでは、所見や病変の様子、サイズなど多くの情報が存在する。これらの膨大なデータの入力規則は一定の枠組みが存在するが、大半はレポートシステムに依存している。そのため、施設間や汎用的な研究にデータを使用したい場合は、データの整形作業が必要となる。これらは人によって行われることが一般的であるが、その作業負担は極めて大きい。現状の AI を臨床応用していくにあたり、複数の課題もあるが、今後の AI 学習および研究を円滑に進めるためには、レポートシステムの標準化が重要な課題の1つであると考えられる。

5. さいごに

読影環境およびデバイスに関する基礎的知識と現状について整理した。実際の臨床現場の状態を把握し、意識することで、新たな観点からの研究テーマの創造や研究条件の設定が可能となる。今後も改定

や新たな技術が登場すると考えられるが、その根幹は変わらない。今後もより臨床に即した研究の発展が期待される。

6. 参考文献

1. Norweck, James T, et al. ACR-AAPM-SIIM technical standard for electronic practice of medical imaging. Journal of digital imaging 2013; 26(1): 38-52.
2. 日本医学放射線学会, 日本放射線科専門医会・医会, 日本診療放射線技師会 共同編集. 放射線科医から診療放射線技師へのタスク・シフト/シェアのためのガイドライン集. 2024.
3. DICOM Standards Committee. Digital Imaging and Communications in Medicine (DICOM) Part 14: Grayscale Standard Display Function. PS3.1; 2026a.
4. 日本画像医療システム工業会技術資料. JESRA TR-0049-2024, 医用画像表示用モニタの品質管理に関するガイドライン. 2024.
5. 日本画像医療システム工業会技術資料. JESRA X-0093*C-2024, 医用画像表示用モニタの品質管理に関するガイドライン. 2024.
6. L. B. Lusted. Logical analysis in Roentgen diagnosis. Radiology 1960; 74(2), 178-193.
7. 長谷川玲. 世界で初めて商品化されたマンモグラフィ用 CAD-ImageChecker-, 日本放射線技術学会雑誌 2000; 56(3). 355-358.
8. 藤田広志. 放射線医学・技術学領域における AI 応用 CAD の新時代. JIRA テクニカルレポート 2018; 54(28-1), 36-42.

臨床画像評価の未来 ～AIを使った画質評価法～

滋賀医科大学医学部附属病院 大橋 昂平

1. はじめに

医用画像の画質評価は、画像診断の質を担保するうえで極めて重要な要素であり、その目的や評価対象に応じて多様な手法が用いられてきた。代表的な手法として、modulation transfer function (MTF) や noise power spectrum (NPS) などの物理評価指標が挙げられる。これらの指標は、画像が有する物理的特性を空間周波数領域で定量的に評価できる点に優れており、医用画像システムの性能評価において中心的な役割を果たしている。一方で、これらの物理評価手法は、評価指標に応じた測定用のファントムを必要とする場合が多く、臨床画像や人体ファントムのような複雑な構造を有する画像に対して直接適用することは必ずしも容易ではない。そのため、臨床画像の画質評価においては、観察者による視覚評価が広く用いられている。

自然画像分野に目を向けると、人が知覚する視覚的品質を定量的に評価することを目的とした image quality assessment (IQA) という研究分野が発展してきた¹⁾。自然画像分野で開発された IQA 手法は、医用画像の画質評価にも利用されている。本稿では、IQA 手法の基本概念および分類について概説したうえで、深層学習を用いた IQA 手法に焦点を当てる。特に CT 画像を対象とした研究動向を紹介し、AI (artificial intelligence) が臨床画像の画質評価にもたらす可能性について考察する。

2. IQA とは

IQA とは、画像の視覚的品質、すなわち人が画像を観察した際に知覚する画質の良否を定量的に評価することを目的とした研究分野であり、主に自然画像を中心に発展してきた。IQA には、大きく分けて主観的 IQA と客観的 IQA の二つのアプローチが存在する。主観的 IQA は、人間の観察者が知識や経験に基づいて画像品質を判断する手法であり、人の知覚を最も直接的に反映することから、IQA におけるゴールドスタンダードと位置づけられている。しかしながら、評価に多大な時間と労力を要する点や、観察者間差、再現性の問題といった課題も存在する。一方、客観的 IQA は、数理モデルや計算アルゴリズムを用いて知覚品質を定量的に予測する手法であり、主観的評価に代わる効率的かつ一貫性のある評価手段を提供する。客観的 IQA は、参照画像の有無に基づき、主に full-reference IQA (FR-IQA: 完全参照型画質評価) と no-reference IQA (NR-IQA: 非参照型画質評価) に大別される。

2.1. FR-IQA (完全参照型画質評価)

FR-IQA は、高品質な参照画像と評価対象画像を比較することで画質を評価する手法である。代表的な指標として、peak signal-to-noise ratio (PSNR) や structural similarity (SSIM)²⁾ が挙げられる。SSIM は、画像内を移動する局所ウィンドウごとに、「輝度」「コントラスト」「構造」の3成分に基づいて参照画像と評価画像の類似度(局所 SSIM)を算出し、それらを平均することで画像全体の SSIM 値を得る手法である (Fig. 1)。SSIM 値は 1 に近いほ

ど参照画像との類似度が高く、画質が良好であることを示す。SSIM は MATLAB や Python などのプログラミング環境において容易に利用可能であり、ImageJ においても専用プラグインを用いることで graphical user interface (GUI) 操作のみで算出できる。このような利便性から、医用画像分野においても広く利用されている。一方で、FR-IQA は幾何学的条件が一致した高品質な参照画像を必要とする特性上、人体ファントムなどには適用しやすいが、臨床画像の評価においては、その適用が制限される場合がある。

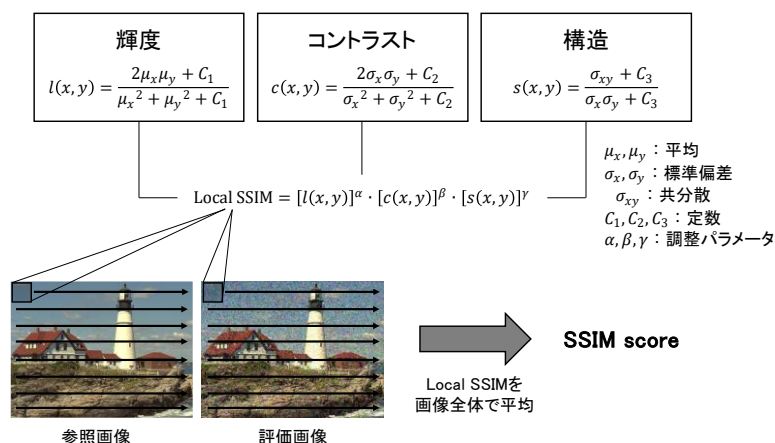


Fig. 1 SSIM 算出アルゴリズムの概要

(例示画像は TID2013 database³より引用)

2.2. NR-IQA (非参照型画質評価)

NR-IQA は、参照画像を用いず、評価対象画像のみから画質を推定する手法である。FR-IQA のように参照画像に基づく評価基準を設定できないため、よりチャレンジングなタスクである。臨床画像の画質評価においては参照画像の取得が困難な場合が多く、NR-IQA はそのような状況に適した枠組みといえるものの、医用画像分野において標準的な手法が十分に確立されているとはいえない。伝統的な NR-IQA 手法としては、BRISQUE (blind/referenceless image spatial quality evaluator)⁴、NIQE (naturalness image quality evaluator)⁵、PIQE (perception-based image quality evaluator)⁶などが挙げられる。これらは自然シーン統計 (natural scene statistics: NSS) に基づく手法であり、「歪みのない自然画像は統計的に共通した規則性を有し、劣化によりその規則性が崩れる」という仮定に基づき画質を推定する。BRISQUE は、平均減算コントラスト正規化 (mean subtracted contrast normalized: MSCN) 係数に基づく NSS 特徴量を抽出し、主観評価スコアを教師とした回帰モデルにより未知画像の画質を推定する。NIQE は主観スコアを用いず、事前に自然画像群から抽出した MSCN 係数に基づく NSS 特徴量を「自然さの規範」としてモデル化し、評価画像がその規範からどの程度乖離しているかを品質指標として算出する。PIQE は学習済みモデルを使用せず、画像をブロック単位で解析し、知覚的に重要な領域における歪みを MSCN 係数に基づく局所統計量から定量化することで画質を推定する。ただし、BRISQUE および NIQE は自然画像を用いて構築されたモデルを用いるため、医用画像に直接適用した場合、性能が十分に発揮されない可能性がある。この課題に対し、医用画像で回帰モデルを再学習する試みも報告されている⁷。さらに、自然画像分野では、深層学習を用いた NR-IQA が発展し、高精度な知覚品質推定が可能であることが示されている⁸。次章では、深層学習ベース NR-IQA の基本的枠組みと、医用画像 (特に CT) への応用研究例を概説する。

3. 深層学習ベースの NR-IQA

深層学習を用いた NR-IQA 手法は、入力画像から画質推定に有用な特徴量を自動的に抽出し、多層構造を通じて高次元の特徴表現を学習できる点に特徴がある。Fig. 2 に、深層学習ベース NR-IQA の基本的な枠組みを示す。学習には、画像とそれに対応する主観的画質スコアから構成される IQA データセットが用いられる。画像群には多様な画質が含まれることが望ましく、主観スコアとしては複数の評価者から得られた mean opinion score (MOS) が一般的に用いられる。IQA は主観スコアを予測する回帰問題として定式化されることが多く、出力は最終的に単一のスカラー値に集約される。ネットワーク出力と主観スコアとの誤差を最小化するように学習が行われ、学習済みモデルに未知画像を入力することで推定画質スコアが得られる。

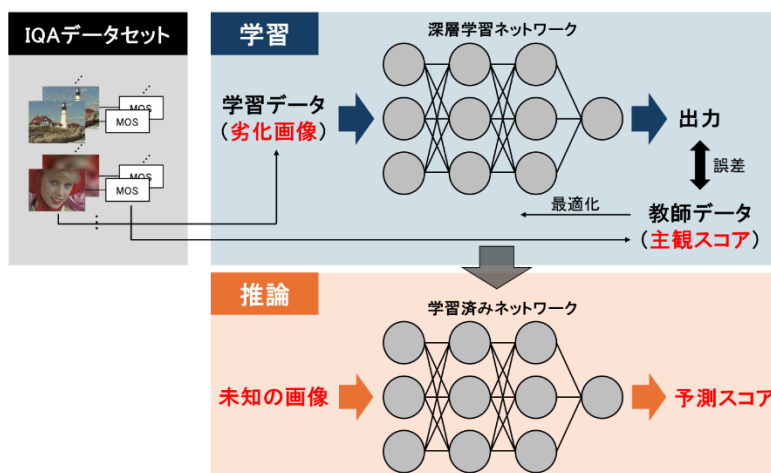


Fig. 2 深層学習ベース NR-IQA 手法の基本的枠組み

(例示画像は TID2013 database³より引用)

自然画像分野では、Kang らが 2014 年に、畳み込みニューラルネットワーク (convolutional neural network: CNN) を用いて NR-IQA をエンドツーエンドで学習する初期の研究を報告した⁹⁾。局所正規化した画像パッチを入力とし、畳み込み層、プーリング層、全結合層からなる CNN により各パッチの画質を推定し、それらを統合して画像全体の画質を算出する手法である。その後、ネットワーク構造や学習方法の改良を通じて、多数の深層学習ベース NR-IQA 手法が提案され、性能の向上が図られている⁸⁾。近年では、vision transformer (ViT) や大規模マルチモーダルモデル (large multimodal models: LMMs) を基盤とする手法も提案されている¹⁰⁾。医用画像分野においても深層学習を用いた NR-IQA の研究が進んでおり、次節では CT 画像を対象とした代表的な研究動向を紹介する。

3.1. CT 分野における初期報告

CT 分野における深層学習ベース NR-IQA の初期的取り組みとして、2018 年に Li らが報告した研究が挙げられる¹³⁾。本研究では、300 枚の高品質臨床 CT 画像に対して 4 段階のノイズを人工的に付加し、計 1200 枚の劣化画像を作成した。これらに対し放射線科医師が 5 段階で主観評価を行い、IQA データセットを構築した。データセットは学習用 1350 枚とテスト用 150 枚に分割され、CNN モデルである AlexNet を用いて学習が行われた

(Fig. 3). 本研究は、実験条件が比較的単純である点から結果の解釈には慎重さが求められるものの、深層学習によるCT-NR-IQAの可能性を示した先駆的研究として意義深い。

一方で、深層学習を用いたCT-NR-IQAには、いくつかの本質的な課題が存在する。第一に、深層学習には大量の学習データが必要であるが、画質の主観評価には専門的知識を有する評価者が不可欠であり、評価者の負担が大きいため、大規模なIQAデータセットの構築が容易ではない。第二に、多様な画質を含むデータセットを効率的に作成するため、高画質画像に人工的な劣化を付加する方法が一般的に用いられているが、この場合、学習が人工劣化に依存しやすく、実臨床で生じる複雑かつ現実的な画質劣化に対する汎用性が低下する可能性がある。第三に、IQAは人の知覚に基づく画質を評価する枠組みであり、病変検出能や診断能といった診断的品質を直接的に表すものではない点にも留意が必要である。第四に、多くの手法では画質を単一の数値として出力するため、ノイズや空間分解能など、どの要因が画質低下に寄与しているかを把握しにくいという課題がある。次章からは、これらの課題に対する解決を試みた深層学習ベースCT-NR-IQAの研究を紹介する。

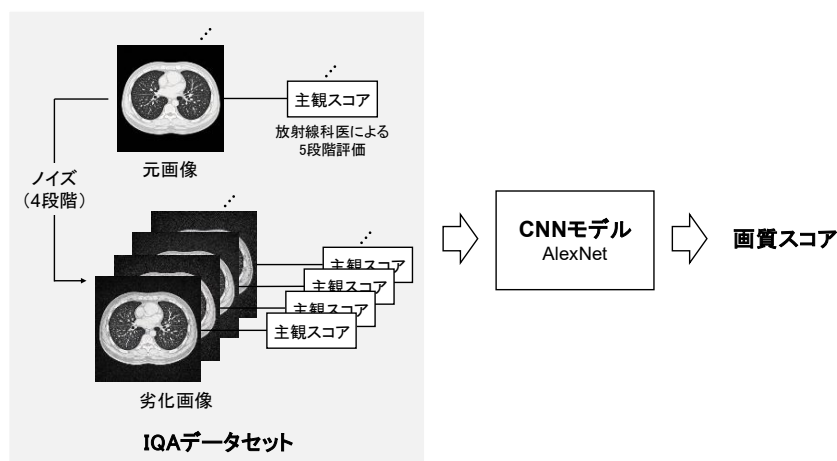


Fig. 3 深層学習ベースのCT-NR-IQAの初期検討

(文献13を基に著者作成. 画像も実際に使用されているものとは異なる.)

3.2. FR-IQAを教師とした自動ラベリング手法

IQAデータセットでは、主観評価に基づく画質スコアをラベルとして用いる必要があるため、大量に用意することが困難である。この点は、多数の学習データを必要とする深層学習において、モデルの汎用性や安定性を制限する要因となり得る。Gaoらはこの課題に対する解決策として、FR-IQA手法を活用した自動ラベリングに基づく学習手法を提案した¹⁴⁾。本手法では、十分な線量で取得されたCT画像に対して複数段階のノイズを付加することで、学習用の劣化画像群を作成している。これらの劣化画像には対応する高画質な元画像が存在するため、参照画像を用いたFR-IQAスコアの算出が可能となる。本研究では、PSNRおよびSSIMといったFR-IQA指標を統合したスコアをground truthとして用いることで、主観評価を必要としない自動ラベリングを実現している(Fig. 4)。FR-IQAを主観スコアの代替として利用するアプローチは、自然画像分野においても広く用いられており、学習データ不足を補う実用的な手法として有効である。

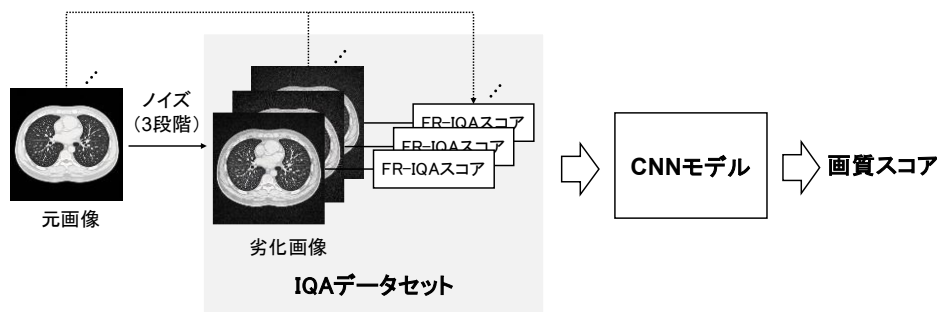


Fig. 4 FR-IQA を教師とした自動ラベリング手法

(文献 14 を参考に著者作成. 画像も実際に使用されているものとは異なる.)

3.3. RadImageNet 事前学習モデルを用いた手法

IQA データセットは、多様な画質を含む画像を大量に用意する必要があるため、一般に高品質な画像に人工的な劣化を複数段階で付加することで構築される。その結果、学習が人工劣化画像に依存しやすく、実臨床画像にみられる複雑な画質劣化に対する汎用性が低下する可能性がある。この課題に対する一つのアプローチとして、RadImageNet により事前学習された CNN モデルを IQA タスクに転移学習する手法が提案されている (Fig. 5)¹⁵⁾。RadImageNet は、CT, MRI, 超音波画像など約 135 万枚の病変ラベル付き画像から構成される大規模医用画像データセットである¹⁶⁾。RadImageNet を用いた事前学習により、モデルは実臨床画像に由来する多様な特徴表現をあらかじめ獲得することが可能となる。この事前学習済みモデルに対して IQA データセットを用いた転移学習を行うことで、実臨床環境における画質劣化への適応能力を備えた CT-NR-IQA 手法が期待される。また、大規模データで事前学習されたモデルを利用することは、IQA データセット不足への対策としても有効である。本研究では、人工劣化を加えていない実臨床画像のみから構成されるデータセットでの評価において、RadImageNet 事前学習モデルは ImageNet 事前学習モデルと比較して、大きな性能向上がみられた。この結果は、RadImageNet による事前学習が、実臨床画像における「リアルな劣化」への適応性向上に寄与している可能性を示唆している。

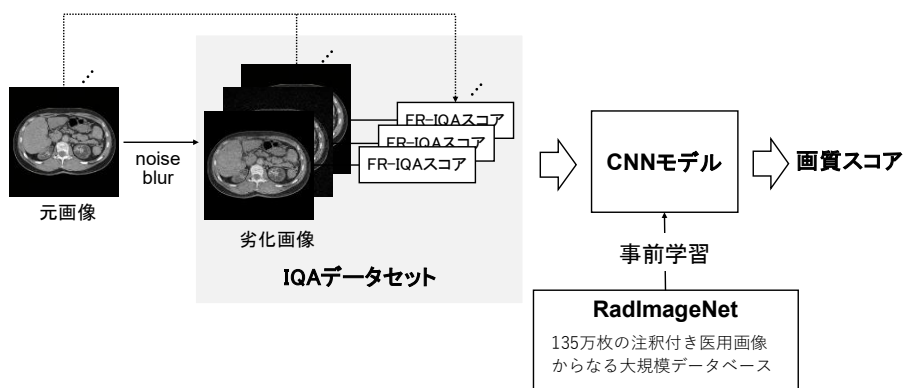


Fig. 5 RadImageNet 事前学習モデルを用いた手法

(文献 15 を参考に著者作成. 画像も実際に使用されているものとは異なる.)

3.4. 病変検出能に基づく手法

多くの CT-NR-IQA 研究は、人が知覚する視覚的品質の定量化を目的としているが、視覚的品質は診断能や病変検出能を直接的に示すものではない。この点は医用画像特有の課題であり、自然画像分野とは異なる重要な観点である。Lee らは、この問題に対し、病変検出能を反映したアプローチを提案した (Fig. 6)¹⁷⁾。具体的には、CT 画像内に仮想的な低コントラスト物体を挿入し、それらを同定する物体検出モデルを学習する。評価対象画像に対しても同様に仮想物体を挿入し、モデルによる仮想物体の検出精度を画質スコアとして定義する手法である。本研究は、画質評価を診断タスクと結び付けた点で意義深く、臨床的価値を反映した IQA の方向性を示す重要な報告である。

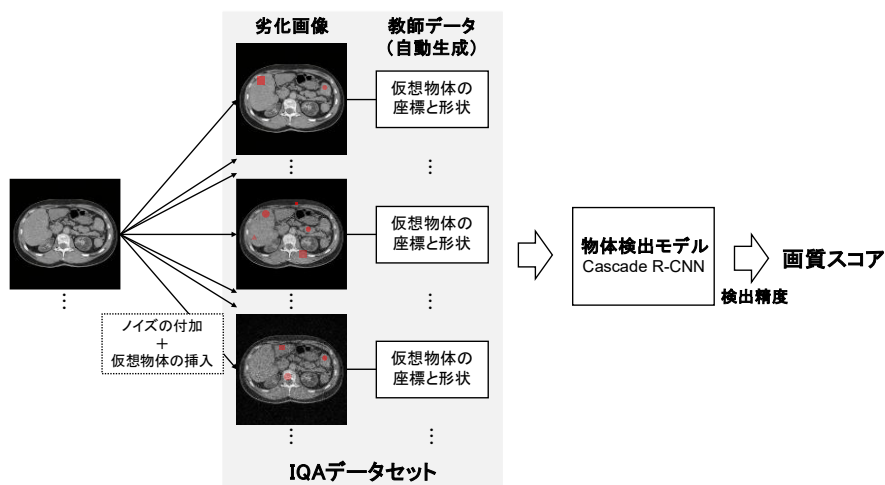


Fig. 6 病変検出能に基づく手法

(文献 17 を参考に著者作成。画像も実際に使用されているものとは異なる。)

3.5. 説明可能な IQA への展開

深層学習ベース NR-IQA は高い推定精度が期待される一方で、出力が単一の数値であるため、「なぜその画質スコアなのか」を理解しにくい。近年、大規模言語モデル (large language model: LLM) や、画像と言語を統合的に扱う視覚言語モデル (vision-language model: VLM) の発展により、この課題に対する新たな可能性が示されている。Chen らは、CT-NR-IQA における説明可能性の向上を目的として、LLM および VLM を用いた画質評価フレームワーク (IQAGPT) を提案した (Fig. 7)¹⁸⁾。本手法では、放射線科医が CT 画像を「画像ノイズ」「微細構造の視認性」「病変の視認性」「診断信頼度」など複数の観点から評価し、それを文章表現として付与する。これらの画像と言語のペアを用いて VLM を学習させることで、CT 画像から画質に関する言語的な説明 (画質キャプション) を生成可能とした。さらに、生成された画質キャプションを LLM に入力することで、画質スコアの推定と品質レポートの自動生成を実現している。本手法は、画質を数値だけでなく言語によって説明可能とした点で非常にユニークであり、臨床現場における意思決定の支援や評価プロセスの透明化を促進するアプローチとして注目される。

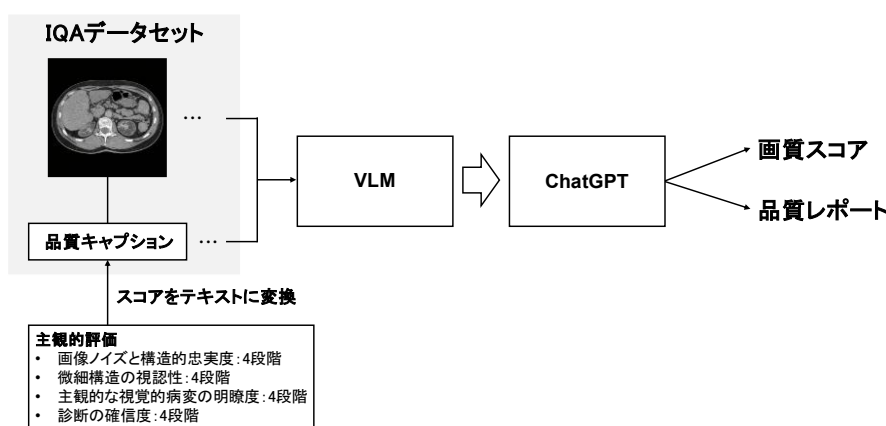


Fig. 7 説明可能な CT-NR-IQA

(文献 18 を参考に著者作成. 画像も実際に使用されているものとは異なる.)

3.6. 課題と今後の展望

以上のように, CT 画像における深層学習ベース NR-IQA 研究は多様な観点から活発に進められており, 本稿で紹介した論文はその一部である. CT に限らず, 一般撮影画像, 超音波画像, MRI など他のモダリティにおいても, 同様に深層学習ベース NR-IQA に関する研究報告がみられる¹⁹⁾. しかし現時点では, その多くが研究段階にとどまっており, 臨床現場で広く実用化されている手法はまだ確立していない. そのため, 現状における臨床画像の画質評価の第一選択は, 依然として視覚評価であると考えられる. 本分野の発展と臨床実装を進める上での大きな課題の一つは, 医用画像におけるオープンソースの IQA データセットが不足している点である. 自然画像分野では複数の公開 IQA データセットが整備されている一方, 医用画像では共通基盤となるデータセットが限られており, 手法間の性能比較や外部検証が困難である. こうした中, 2023 年に開催された国際会議 MICCAI 2023 において CT-IQA チャレンジ (low-dose CT perceptual image quality assessment grand challenge: LDCTIQAC 2023) が実施され, 専門医の主観評価スコアを付与した IQA データセットが公開された²²⁾. 今後, さらに利用可能な大規模医用画像 IQA データセットや評価プロトコルが整備されれば, 医用画像における NR-IQA 研究の信頼性と再現性が向上し, 臨床画像評価の未来に寄与していくことが期待される.

参考文献

- 1) Athar S, Wang Z. A comprehensive performance evaluation of image quality assessment algorithms. IEEE Access 2019; 7: 140030-140070.
- 2) Wang Z, Bovik AC, Sheikh HR, et al. Image quality assessment: from error visibility to structural similarity. IEEE Trans Image Process 2004; 13(4): 600-612.
- 3) Ponomarenko N, Jin L, Ieremeiev O, et al. Image database TID2013: Peculiarities, results and perspectives. Signal Process Image Commun 2015; 30: 57-77.
- 4) Mittal A, Moorthy AK, Bovik AC. No-reference image quality assessment in the spatial domain. IEEE Trans Image Process 2012; 21(12): 4695-4708.

- 5) Mittal A, Soundararajan R, Bovik AC. Making a “completely blind” image quality analyzer. *IEEE Signal Process Lett* 2013; 20(3): 209-212.
- 6) Venkatanath N, Praneeth D, Maruthi Chandrasekhar Bh, et al. Blind image quality evaluation using perception based features. *Twenty First National Conference on Communications (NCC) 2015*; 1-6.
- 7) Chow LS, Rajagopal H. Modified-BRISQUE as no reference image quality assessment for structural MR images. *Magn Reson Imaging* 2017; 43: 74–87.
- 8) Yang X, Li F, Liu H. A survey of DNN methods for blind image quality assessment. *IEEE Access* 2019; 7: 123788-123804.
- 9) Kang L, Ye P, Li Y, et al. Convolutional neural networks for no-reference image quality assessment. *2014 IEEE Conference on Computer Vision and Pattern Recognition* 2014; 1733-1740.
- 10) Yang S, Wu T, Shi S, et al. MANIQA: Multi-dimension attention network for no-reference image quality assessment. *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW) 2022*; 1191-1200.
- 11) Ke J, Wang Q, Wang Y, et al. MUSIQ: Multi-scale image quality transformer. *IEEE/CVF International Conference on Computer Vision (ICCV) 2021*; 5148-5157.
- 12) Wu H, Zhang Z, Zhang W, et al. Q-Align: Teaching LMMs for visual scoring via discrete text-defined levels. *Proceedings of the 41st International Conference on Machine Learning (ICML) 2024*.
- 13) Li S, He J, Wang Y, et al. Blind CT image quality assessment via deep learning strategy: initial study. *Proc SPIE* 2018; 10577A.
- 14) Gao Q, Li S, Zhu M, et al. Blind CT image quality assessment via deep learning framework. *2019 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC) 2019*; 1-4.
- 15) Ohashi K, Nagatani Y, Yamazaki A, et al. Development of a no-reference CT image quality assessment method using RadImageNet pre-trained deep learning models. *J Imaging Inform Med* 2026; 39: 46-58.
- 16) Mei X, Liu Z, Robson PM, et al. RadImageNet: An open radiologic deep learning research dataset for effective transfer learning. *Radiol Artif Intell* 2022; 4(5): e210315.
- 17) Lee W, Cho E, Kim W, et al. No-reference perceptual CT image quality assessment based on a self-supervised learning framework. *Mach Learn Sci Technol* 2022; 3(4): 045033.
- 18) Chen Z, Hu B, Niu C, et al. IQAGPT: computed tomography image quality assessment with vision-language and ChatGPT models. *Vis Comput Ind Biomed Art* 2024; 7(20).
- 19) Guan M, Lyu Y, Cao W, et al. Perceptual quality assessment of chest radiograph. *Lect Notes Comput Sci* 2021; 12907: 315-324.
- 20) Zhang S, Wang Y, Jiang J, et al. CNN-based medical ultrasound image quality assessment. *Complexity* 2021; 9938367.
- 21) Stępień I, Obuchowicz R, Piórkowski A, et al. Fusion of deep convolutional neural networks for no-reference magnetic resonance image quality assessment. *Sensors* 2021; 21(4): 1043.
- 22) Lee W, Wagner F, Galdran A, et al. Low-dose computed tomography perceptual image quality assessment. *Med Image Anal* 2025; 99: 103343.

骨密度測定装置の技術解説

GE ヘルスケア・ジャパン株式会社 Women's Health & X-ray 部 矢ヶ部 真理

【はじめに】

日本は超高齢化社会を迎え、それに伴う医療費の増加が深刻な社会課題となっている。骨粗鬆症患者数も増加の一途をたどっており、骨粗鬆症に起因する骨折を未然に防ぐことは、今後の日本の医療を考える上では重要な課題と思われる。

現在、骨粗鬆症の診断や治療における判断基準として、二重エネルギーX線吸収測定(DXA: Dual Energy X-ray Absorptiometry)法による腰椎および大腿骨の骨密度測定がゴールドスタンダードとして確立している。これは、DXA法が他の測定方法に比べ、骨粗鬆症の兆候を確実にとらえる高い測定精度と、継続的な治療効果を評価するための安定した計測(再現性)を持つことに起因するからである。

このような臨床的要求に加え、増加の一途をたどる骨密度測定検査を効率的に行うために、DXA装置は長い歴史の中で様々な技術的進化を遂げてきた。本稿では当社のDXA装置 Lunar iDXA に搭載されている技術について紹介する。

【照射方式について】

DXA装置に用いられている照射方法には、大きく分けて3つの方式がある(図1)。初期のDXA装置は、X線を線状に細く絞ったペンシルビーム方式という照射方式が採用されていた。第一世代となるペンシルビーム方式は被ばくが少なく、対象骨の投影像に拡大誤差が生じにくいいため、高い測定精度を実現できるというメリットがあった。しかし、そのビームサイズからスキャン時のスイープ回数が多くなるため、測定時間が非常に長くなるというデメリットもあった。このペンシルビーム方式の持つ課題に対して次に生み出されたのが、第二世代となるワイドファンビーム方式である。幅広い検出器全体に扇状のX線を照射して1回のスイープで測定を行うワイドファンビーム方式は、測定時間が短縮できるため一般的にも広く普及した。当社でも古くはワイドファンビーム方式を採用していた(図2)。

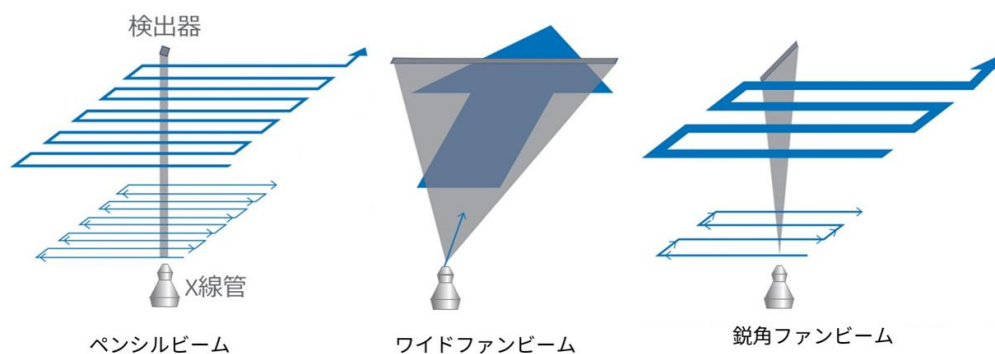
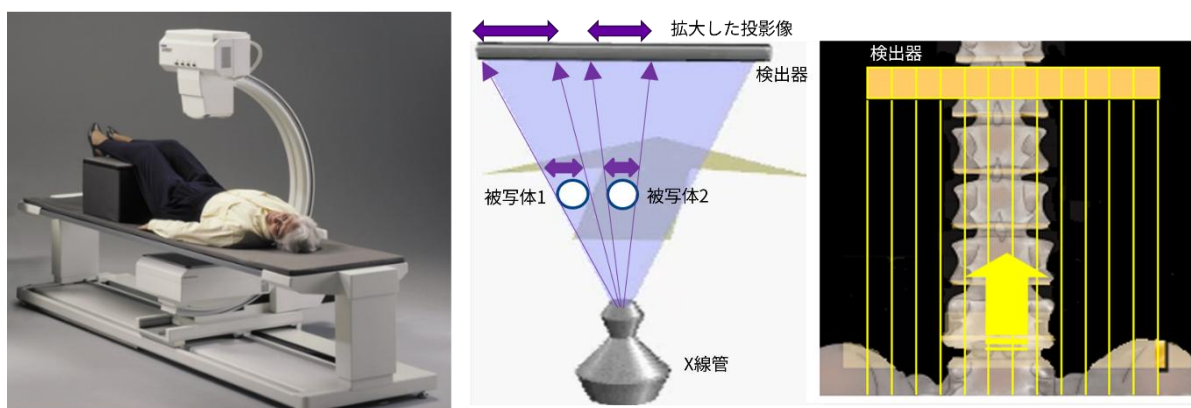


図 1. DXA 装置に用いられている照射方式

しかしながら、ワイドファンビーム方式は扇状の X 線の広がりを持つ特性上、投影される骨面積に幾何学的な拡大誤差が生じてしまい、結果として、骨塩量(BMC)÷骨面積(cm²)から算出される骨密度(BMD)の過小評価につながってしまうという大きな課題が生じた(図2(b))。その骨面積の拡大誤差の程度は、被写体の位置、例えばベッド中心からの左右方向へのずれや、被写体の体厚変動に起因するベッド面からの高さ方向の違いによって、投影される骨面積が大きく変化することになる。つまりこれは、1回1回の測定精度だけではなく、検査毎のポジショニングや被検者の体格変化等によって、経過観察中の再現性にも影響がでることを意味する。また、ワイドファンビーム方式では幅広い X 線束を覆う多くの検出素子が必要となるため、素子毎の感度の不均一性のリスクや検出素子の故障リスクが上がること、幅広い X 線照射によって不必要な部分への被ばくを増やし、かつ、周囲への散乱線も多くなるといった側面もあった(図2(c))。



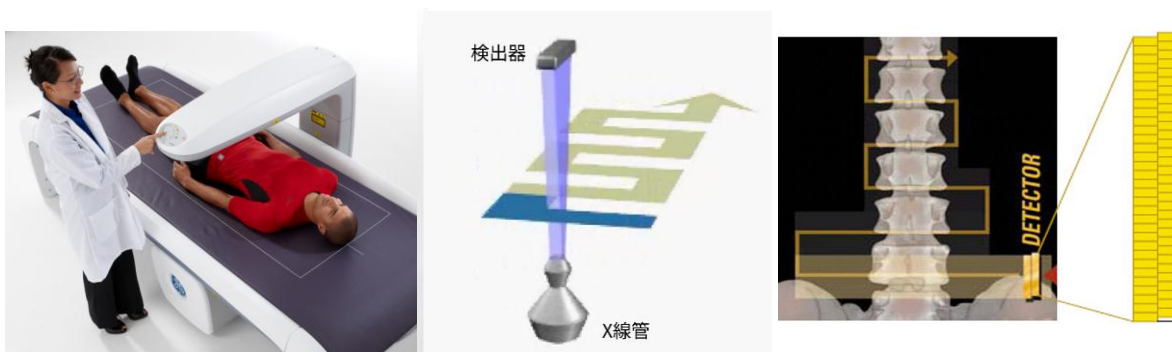
(a)装置の外観

(b)レイアウトと走査方向

(c)検出器並びと走査方向

図2. ワイドファンビーム方式

そこで、ワイドファンビーム方式のもつ課題を打破し、なおかつ、測定時間も短縮させるために生み出されたのが、第三世代となる鋭角ファンビーム方式である(図3)。現在当社の DXA 装置には、この鋭角ファンビーム方式が採用されている。



(a)装置の外観

(b)レイアウトと走査方向

(c)検出器並びと走査方向

図3. 鋭角ファンビーム方式

鋭角ファンビーム方式では、体軸横方向に対してペンシルビームのように細く絞られた X 線を使用することで、投影される骨面積の幾何学的な拡大誤差を抑制することができる。つまり、側弯症の腰椎撮影時や大腿骨近位部の撮影時等、被写体が X 線の線束中心からずれている場合でも、測定誤差を抑えることができる。

また鋭角ファンビーム方式では、体軸縦方向には検出器を複数配列することで幅を持たせ、対象骨を自動追従しながら必要最低限の範囲だけをスイープさせることで、測定時間や被ばく量も低減させることができる(図3(b)(c))。この時、体軸方向に生じた各スイープ間の歪みをなくすために、各スイープはオーバーラップした形で行われ、最終的に各スキャンで得られた画像に対して画像再構成が行われる。CT 等にも使用されているこの再構成技術は MVIR (Multi-View Image Reconstruction) と呼ばれ、体軸方向に対しても正確な骨面積を提供することができる(図4)。

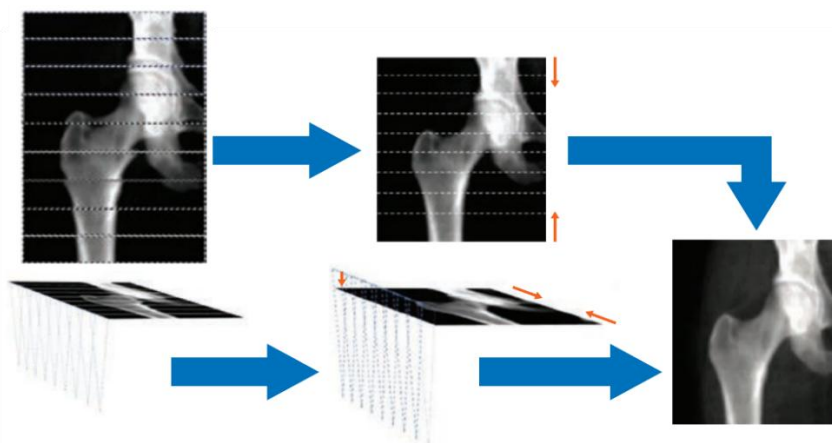


図4. 再構成技術 MVIR(Multi-View Image Reconstruction)

MVIR を併用した第三世代の鋭角ファンビーム方式は、検査効率も大きく改善させた。従来のワイドファンビーム方式による腰椎撮影では、幾何学的拡大誤差の影響を避けるために腰椎撮影時は膝を屈伸して撮影することが必須であった。つまり、検査途中、腰椎と大腿骨撮影の間に一度ポジショニング変更を行う必要があった。対して、鋭角ファンビーム方式では、膝を伸展したまま腰椎と大腿骨を同じポジショニングで計測することができる(図5)。検査途中で操作者がポジショニング変更のために介入する手間がなく、高い測定精度を維持したまま、検査効率を大幅に向上させることができるようになった。高齢の被検者にとっても、ベッドに横たわるだけの安定した体位で検査を受けることができるため、医療従事者および被検者ともに検査負荷が軽減できるわけである。

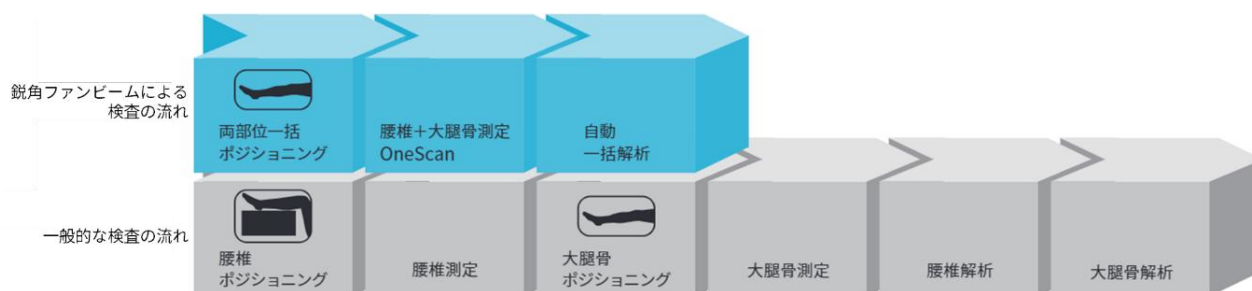


図5. 鋭角ファンビームによる検査の流れ

【検出技術について】

DXA 法では、その名のとおり高エネルギーと低エネルギーの2種類の X 線を用いて撮像する。高エネルギーの X 線は透過力があるが、低エネルギーの X 線は透過力が弱いために骨部で大きな減衰が起きる。その2つの差分をもとに、骨部での X 線減衰の度合いから骨量を求めることができる。実際の X 線の照射部位には骨の前後に軟部組織が含まれるため、骨を含まない近接部位から得た軟部組織の X 線減衰の情報をベースラインとして補正に用いている。最終的に、このベースラインとなる軟部組織と骨部の減衰の合計値をさし引くことで、骨量を計算している。

当社の DXA 装置では、高エネルギーと低エネルギーの2種類の X 線を取り出すために、K-エッジフィルターを使用している(図6)。連続したエネルギーを持つ X 線を、K-エッジフィルターに透過させることで余分な中エネルギー領域の X 線をブロックし、必要な高エネルギーと低エネルギーの X 線だけを被検者へ照射することができるため、他の方法に比べて被ばくを低下させることができるメリットがある。また、高エネルギーと低エネルギーの2つの X 線を得るために高速で電圧を切り替える必要がないため、切り替え時のエネルギーの揺らぎを抑えることもでき、定量測定に適した常に安定した X 線照射を行うことで計測精度もあがり、タイムラグなく2種類の X 線を照射することができるのも、K-エッジフィルターの利点である。

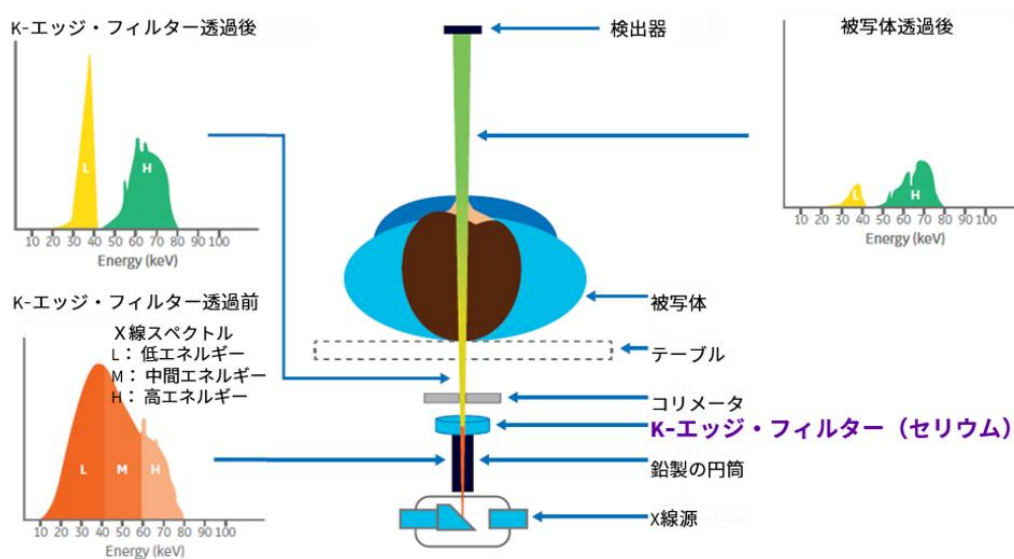


図6. K-エッジフィルター

この K-エッジフィルター透過後の X 線の受け手となる検出器に、当社はフォトンカウンティング(光子計測法)を採用している。フォトンカウンティングとは、ある一定値以上のエネルギーを持った X 線光子数をカウントすることでデジタル信号化することのできる検出技術である。近年、被ばく低減に取り組むCTやPETなどにも商用化されてきた検出効率の高い検出技術である。フォトンカウンティングでは、検出する高エネルギー帯と低エネルギー帯にそれぞれ上限と下限となるしきい値が設けられており、そのしきい値間のエネルギーをもつ光子数をそれぞれデジタルカウントする(図7)。そのため、検出される光子数には高/低のエネルギー混在が起きることはない。

X 線の発生方式と検出方式には様々な方法があり、各方式には利点と欠点があるが、当社では発

生器側の K-エッジフィルターと検出器側の光子カウンティングの技術を組み合わせることで、それぞれがもつ利点を生かし、被ばく低減、測定精度の向上、高画質という臨床的効果を最大化している。

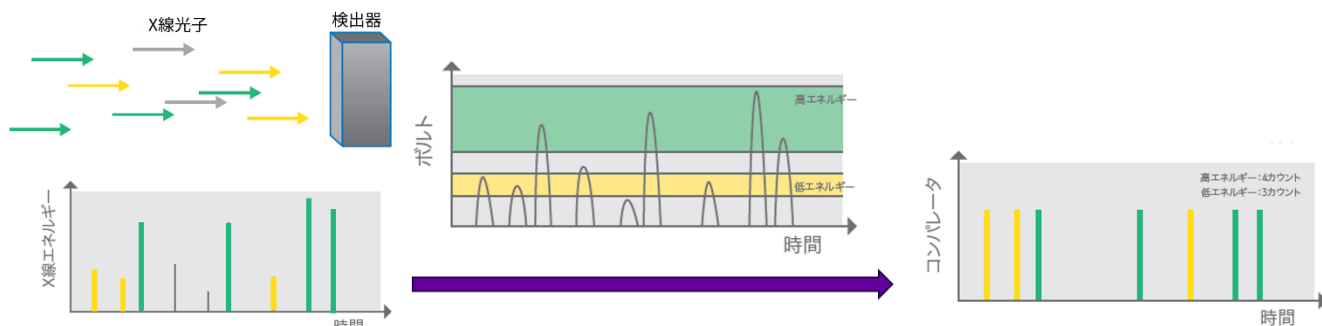


図7. フォトンカウンティング技術

高エネルギーの光子(緑)と低エネルギーの光子(黄)数をデジタルカウントする

また、当社の DXA 装置の Lunar iDXA では32個の素子をもつ検出器を2列(計64個)配列している(図8)。その際に、1列目の素子間の隙間を埋めるような形で、2列目は少しずらした形で配列されている。これにより、高解像度な臨床画像の提供と、各画素の骨部と軟部組織の認識力を向上させることで測定精度をより高めるような配列の工夫がなされている。

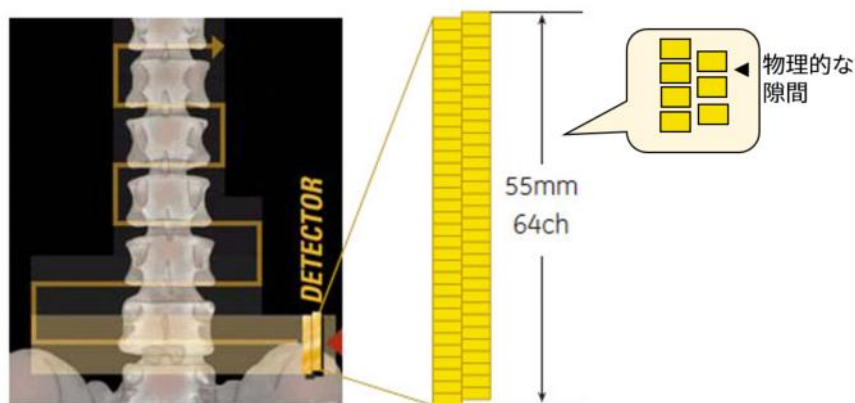


図8. 検出器の配列の工夫

検出器間の物理的隙間を埋める二重配列により、高画質かつ高精度な測定が可能

【DXA 検査の応用: 海綿骨構造指標 TBS】

DXA 装置の進化はハードウェアだけにとどまらず、アプリケーションソフトの分野にも及んでいる。その代表例が、海綿骨構造指標 (Trabecular Bone Score: TBS) ソフトウェアである。TBS は、Medimaps Group 社(スイス)が開発したソフトウェアで、X線骨密度測定装置で計測された腰椎画像の画素濃度をもとに、海綿骨の微細構造を数値化するソフトである。TBS と BMD は同じ DXA スキャンデータを元に解析するが、その計算方法は異なり、BMD とは独立した指標である(図9)。

BMD は骨画像の各ピクセルの輝度を総計したものと関連する一方、骨画像内のピクセル間の値の変動は考慮されていない。対して、TBS のアルゴリズムではピクセルの輝度の空間的変動を解析していく。具体的には、隣接するピクセルの輝度の差を求め、その差の 2 乗を計算することで、ピクセルの輝度の空間的変動を解析している。そのため、TBS は Trabecular Number (Tb.N: 骨梁数)、Trabecular Separation (Tb.Sp: 骨梁間隙)、Connectivity density (Conn.D: 連結性密度)、Structure model index (SMI) のような骨微細構造の指標と相関する¹⁾といわれている。また、骨粗鬆症の予防と治療のガイドライン 2025年版²⁾には、DXA の応用方法の一つとして取り上げられ、「TBS のみで骨粗鬆症治療開始の決定を行うことは推奨されないが、BMD と組み合わせた判定が有用である」と記載されている。

従来の骨密度評価に付加情報を提供するものとして期待されているソフトウェアではあるが、その運用が日常診療においての負荷になってはいけなない。GE ヘルスケア社の DXA 装置では、この TBS ソフトウェアが DXA 装置のユーザーインターフェースに統合されており、独立した TBS ソフトのアプリケーションを立ち上げる必要はない。通常の腰椎骨密度撮影と同様の手順で、撮影、解析からレポート出力まで進めば、自動的に BMD の結果と一緒に TBS の解析結果という付加情報を得ることができる。臨床上での使いやすさや長期的な保守管理の目線でも DXA 装置との一貫性を重視した技術開発がなされている。

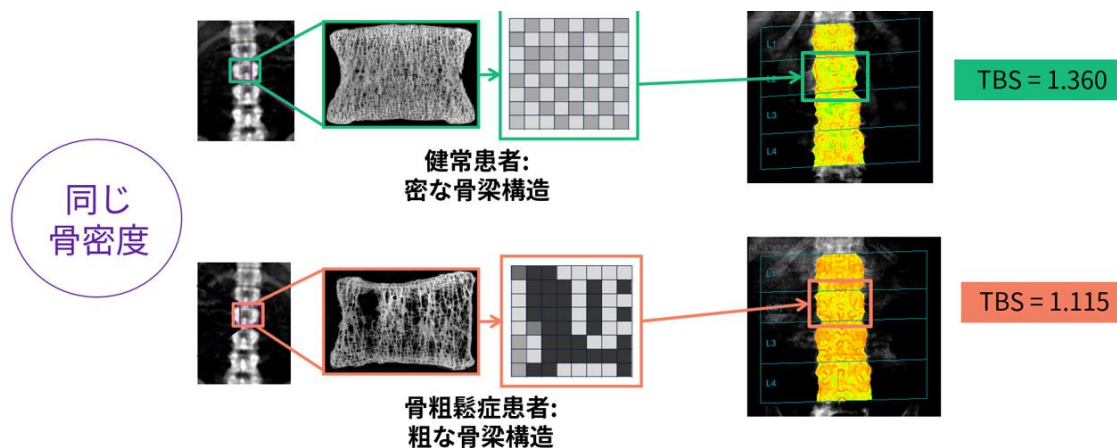


図9. 海綿骨構造指標 (Trabecular Bone Score: TBS)

【おわりに】

増加の一途をたどる骨粗鬆症に対して、骨密度測定ゴールドスタンダードである DXA 装置には、正確な測定精度、常に安定した計測 (高い再現性)、そして検査の効率性が求められている。それらの臨床的要求に対して、各社の DXA 装置は長い歴史の中で様々な進化を遂げてきた。一概に同じ DXA 装置といえども、各社の DXA 装置で使用されている技術は異なるため、装置に搭載されている技術特性を理解したうえで、臨床的な継続観察も維持しながら、新しい技術を取り入れていくことは、よりよい検査の実現につながると思われる。

本稿では、長きにわたる骨密度測定装置の開発・製造の歴史をもつ当社が、その膨大な納入実績から生み出した技術とその変遷を紹介した。当社も引き続き業界をリードする技術開発を通じて、臨床に貢献していきたい。

【参考文献】

- 1) Didier Hans et al. Bone microarchitecture assessed by TBS predicts osteoporotic fractures independent of bone density: the Manitoba study J Bone Miner Res. 2011 Nov;26(11):2762-9. doi: 10.1002/jbmr.499.
- 2) 骨粗鬆症の予防と治療のガイドライン 2025年版 ライフサイエンス社

【薬事情報】

X線骨密度測定装置 Lunar iDXA 医療機器認証番号 21800BZX10007000
JB13410JA

画像の再現性向上をアシストする光学カメラソリューション の最前線

キヤノンメディカルシステムズ(株) XR 事業部 XR 開発部 鈴木省吾

国内営業本部 X線営業部 菊地航希

【はじめに】

当社は、様々な画像診断機器・ソリューションの開発・販売を通じて、医療現場の撮影業務の効率化および質の高い診断への貢献を目指したソリューションの提供に努めている。一般撮影装置においても改善を目指すべく、2024年8月に、X線可動絞りの側面に光学カメラを搭載し、得られた映像情報を Deep Learning などの技術を活用してリアルタイムで処理し、操作者にフィードバックする光学カメラソリューション Camera Assist*¹の販売を開始した。

受診者のワークフロー上の確認の一部を担う役割として、当社の Flat Panel Detector (FPD)である CXDI シリーズと対応する X線可動絞り*²の組み合わせで実現される機能であり、対応していれば既設の装置であっても後付けできる点も好評をいただいている。

既に複数の 400 床以上の施設で運用が開始されており、導入された施設からのフィードバックを基に、2026年に CXDI コンソールのソフトウェアバージョンアップにより機能の追加・改善を計画している。本稿では光学カメラソリューションの導入メリットを、ワークフロー改善および安全性向上の観点で紹介する。

*1 販売名: 「Camera Assist」は CXDI シリーズのオプション

*2 販売名: 医用 X線可動絞り BLR-500A、届出番号: 09B1X00003000078

【特長】

■ Camera Assist : ライブビュー

X線可動絞りの側面に光学カメラ (図 1) を搭載したことで、撮影コンソール上で光学カメラから取得したライブ映像を確認 (図 2) することが可能とした機能である。撮影室確認窓では得られない視野を提供することで、撮影前の受診者の体勢、撮影部位のフレーミングなどを視認でき、髪の毛の映り込みやポジショニング時から動いていないか、両手の挙上具合などを撮影前のポジショニングの再確認をサポートすることができる。

従来本機能の表示にはコンソールとは別のディスプレイが必要であったが、新バージョンより CXDI コンソールと発生装置のパネルの情報が一体となったコンソールの右下部に表示 (図 3) する表示方法も選択可能である。本方式では、撮影時の視線移動を最小限に抑えつつ、操作室のスペースを有効活用でき、使いやすさとワークフローの向上を実現している。



図 1 専用光学カメラ (X線可動絞り部)



図 2 CXDI コンソール表示イメージ



図 3 一体型モニタ (スイッチボックス装着)

■ Camera Assist : プロトコル選択アシスト

光学カメラで得られるライブ映像から人体を検出し、キャノン IT ソリューションズ社製 AI*³ プラットフォーム「LaiGHT」を活用してリアルタイムの姿勢推定を行うことで、選択されたプロトコルとの整合性を比較し、部位/方向、左右の相違を検出した場合には、操作者へ確認を促す機能である。

姿勢推定は大規模データセット*⁴を用いて基本的な姿勢を学習したモデルをベースにして、X線撮影特有の姿勢を追加学習することで認識精度を向上させている。ヒューマンエラーによる再撮影の防止をサポートし、再撮影に伴う受診者の不要な被ばく低減が期待される (表 1)。

表 1 判定対象となるプロトコル*⁵

撮影部位	視野位置	側性
胸部 / 腹部	AP / PA / RL / LL	
胸椎 / 腰椎		
肩 / 鎖骨 / 肩甲骨		L / R
手 / 足 / 肘 / 膝		L / R

*³ 基本的な撮影姿勢の学習・姿勢推定モデルに AI 技術を用いており、本システム自体に自己学習機能は有していない。

*⁴ COCO2017 データセット約 120,000 枚

*⁵ プロトコル選択アシスト機能はすべてのプロトコルに対応しているものではありません。

また、受診者入室からポジショニング完了までの間に導線が行き来する臨床現場において、操作者が本機能の表示内容を一目で判断できるよう、背景に明色を使用して視認性を向上させる表示方法に変更することもできる。特に操作者に注意を促す表示として、部位違い・方向違い・左右違いの 3 パターンがあり、選択されたプロトコルと実際のポジショニングが異なる場合は、直ちに修正を行うことができるため、フェイルセーフとしても有用である。

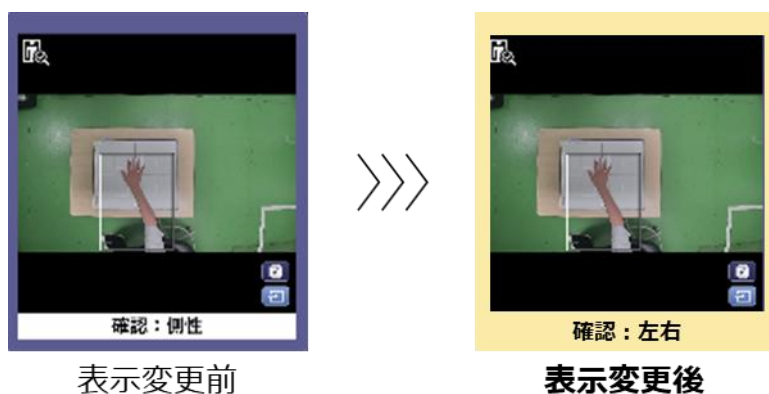


図 4 表示変更前後の比較

■ Image Transfer

光学カメラの画像転送機能として、撮影時の光学画像を DICOM セカンダリキャプチャ形式で転送する機能である。X 線画像撮影後、コンソール側で X 線画像がばく射されたタイミングに最も近い映像フレームを切り出し、光学画像の形で X 線画像と紐づけ、セットで院内サーバー（PACS など）へ転送することができる（図 5）。*⁶

*⁶ CXDI の設定により手動および自動での保存が設定可能



図 5 Image Transfer 機能の動作イメージ図

前述した各機能と組み合わせて期待される主な活用案は、①やむを得ず再撮影する場合に、写損となったポジショニングの画像から再調整の際の支援として活用する、②定期フォローアップの症例におい

て、事前に確認した前回の情報を今回のポジショニング時に応用する、③エビデンスとしての観点で、X線の照射向きが通常と逆の場合に、反転処理操作の確認として利用する、④ポジショニングに関する分析に利用し、撮影の質の向上、学習などに活用する、がある。

■ 体動検知機能

受診者のポジショニング完了時からの体動を確認する技術である。「患者様に正しいポジショニングを完了した」と判断したものの、ハンドスイッチを押そうとする撮影直前に、患者様の体位をみて、「ポジショニング完了時と異なるかも」、「動いたかも」と知覚するシーンは日常的に生じており、本ツールはこの動きに対する不安を確信にする狙いがある。ハンドスイッチを1段押しすると、その時点から本機能が開始される。あらかじめ設定された条件（X秒前など）に基づき準備されたポジショニング完了時点の参考画像と、現在のライブ映像、とを切替えて表示することにより、体動の差分を確認する（図6）。撮影のワークフローを損ねることなく、スピーディに確認が可能である。

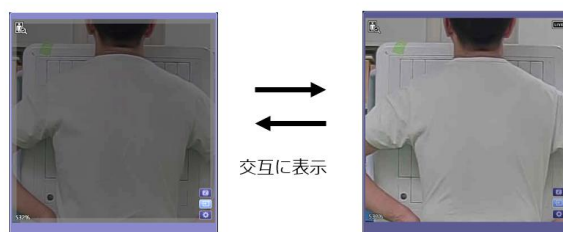


図6 体動検知機能の動作イメージ図

【おわりに】

本稿では、臨床の現場に円滑なワークフローの提供、医療安全に貢献する機能を搭載した当社の光学カメラソリューションについて紹介した。操作者のワークフロー改善にて得られるメリットが、医療の質に直結するポジショニング、画質を担保しつつ不必要な被ばくを削減するための再撮影への配慮、そして医療安全への貢献につながることを期待する。

第 46 回医用画像処理プログラミングセミナーに参加して

横浜栄共済病院 放射線技術科 柿原 妃菜子

このたび、2025年9月7日（日）にオンラインで開催された「第46回医用画像処理プログラミングセミナー」に参加させていただきましたので、その報告をいたします。

私は以前から機械学習や深層学習に興味を持っていましたが、実装に必要なプログラミングに対して強い苦手意識がありました。参考書で基礎を学んでも、いざ自分でコードを書こうとすると「何から始め、どう組み立てればよいのか」が分からず、学習が長続きしないことが悩みでした。そんな折、機械学習の流れを体系的に学べる本セミナーを紹介いただき、この機会に苦手を克服したいと考え、参加を申し込みました。

当日はWebexとSlackを活用し、午前10時から午後3時までのスケジュールで実施されました。事前に配布されたファイルをGoogle Driveに保存し、Google Colaboratory上で実行するという環境で学習が進められました。

セミナーの構成は、①イントロダクションとデータ準備、②探索的データ分析（EDA）、③データ前処理と特徴量エンジニアリング、④機械学習モデルの構築、⑤モデル評価とまとめ、という5つのステップに分かれており、「データの準備」から「モデルの評価」までの一連の流れを通して学ぶことができました。各段階では、専門用語や処理の丁寧な解説ののち、講師の方が用意されたコードを実行して動作を確認する形式がとられ、初心者の方でも安心して取り組むことができました。

特に印象に残ったのは、コード作成に生成AI（Gemini）を活用した点です。「一からすべてをコーディングする」のではなく、「プロンプトを与えてコードを生成させる」という実践的な手法を体験し、プログラミングへのハードルが大きく下がりました。生成AIを活用することで、ハイパーパラメータの調整や統計的手法を用いたモデル評価などもスムーズに行え、自分なりに工夫しながら楽しく学ぶことができました。

また、参加者へのサポートの手厚さも本セミナーの魅力でした。機械学習の専門用語や統計知識を体験的に学べるウェブサイトが用意されており、数値を操作してデータの変化を視覚的に理解できた点は非常に有益でした。進行中に不明点や遅れが生じた際には、Webexの別室機能（ブレイクアウトルーム）にて講師の先生方が個別に対応してくださるなど、オンラインながらきめ細やかな配慮がなされていました。終了後には復習用動画の配布やSlackでの質問受付もあり、フォローアップ体制も万全でした。

受講後、セミナーで学んだ知識を活かし、自力で簡単な「病変の有無の分類タスク」に取り組みました。データの読み込みから前処理、モデル構築、精度評価までの一連の流れを、生成AIの助けを借りながら一人で完遂できたことは大きな自信となりました。一方で、構築したモデルの精度は十分とは言えず、その原因として、モデル構造や学習設定に対する私自身の理解不足、コーディングの甘さを痛感しました。また、エラー発生時に自力で原因を特定し修正する力も不可欠であると学びました。

今後は、生成 AI に頼るだけでなく、処理の一つひとつを深く理解し、より正確で効果的な実装ができるよう研鑽を積んでいきたいと思えます。最後になりますが、丁寧にご指導いただいた講師の先生方をはじめ、運営に携わられたすべての皆様に心より感謝申し上げます。今後もこのような実践的なセミナーが継続されることを願っております。

第 46 回医用画像処理プログラミングセミナーを受講して

順天堂大学大学院保健医療学研究科 隈本神太郎

1. はじめに

この度、2025 年 9 月 7 日（日）に開催された第 46 回医用画像処理プログラミングセミナーに参加させていただきました。私は学部生の頃から深層学習を用いた医用画像解析に興味を持ち、継続的にプログラミングの学習を行ってきました。近年注目されている Radiomics についても自学を始めたところで、本セミナーが「Radiomics 研究にも応用可能な機械学習に関心のある初心者」を対象としていることを知り、既存の知識を整理するとともに、今後の研究基盤を強化する良い機会になると考え、参加いたしました。セミナーは Webex と Slack を用いたオンライン形式で行われ、10:00~15:00（休憩 1 時間）の中で、Python の基礎からデータ分析手法、機械学習の基礎までを、Google Colaboratory 上でのコード実装を通して体系的に学ぶことができました。

2. セミナーの事前準備

セミナーに先立ち、当日使用するプログラムデータや Google Colaboratory の準備に関する資料が配布されました。資料には Google Colaboratory の利用方法や Google Drive との連携手順が丁寧に記載されており、初学者でも戸惑うことなく準備が進められる内容でした。また、Notebook 1~5 に分かれたプログラムデータは、各 Notebook の学習目的が明確かつ段階的に示されており、予習の際にもスムーズに理解を深めることができました。

3. セミナーの内容

セミナー当日は、Notebook 1 から 5 まで順に進めながら実際に手を動かして学習しました。Notebook 1 では分類用と回帰用のサンプルデータを作成し、Notebook 2 ではそれらを用いて、統計量の確認、分布の可視化、欠損値・外れ値の検出などの探索的データ分析（EDA）を行いました。Notebook 3 では欠損値補完、標準化、カテゴリ変換などの前処理に加え、主成分分析（PCA）による次元削減も実施しました。EDA で得られた知見をもとにデータを適切に整える一連の流れを体験でき、これらのプロセスがモデルの精度や信頼性を左右する重要な工程であることを深く理解することができました。Notebook 4 では、分類と回帰の 2 種類のタスクに対して、複数の機械学習モデルを構築しました。資料にはロジスティック回帰、決定木、ランダムフォレスト、SVM といった代表的な機械学習アルゴリズムの概要やそれぞれの長所・短所がまとめられており、それぞれのアルゴリズムの特徴が大変分かりやすかったです。最後の Notebook 5 では、学習したモデルをテストデータで評価しました。Notebook 1 から 5 内のコードは一から自分で書く形式ではなく、用意されたコードを確認しながら実行していくスタイルで、必要に応じて生成 AI（Gemini）にプロンプトを入力してコードを生成する場面もありました。Google Colaboratory 上で生成 AI を活用しての実装は初めての経験で、その効率性に非常に驚きました。また、作業につまずいた際には Webex や Slack を通じて講師の先生方に気

軽に質問でき、個別に丁寧なサポートを受けられるなど、学習環境の整備にも細やかなご配慮がなされており、こうした点にも大いに支えられながら安心して受講することができました。

4. 最後に

本セミナーに参加して、「データを理解し、適切に前処理を行い、モデルを構築・評価する」という一連の考え方が Radiomics を含む医用画像解析において極めて重要であることを改めて認識することができました。今後の研究活動においては、特にデータの理解と前処理の妥当性を丁寧に確認する姿勢を大切にし、より信頼性の高い解析を実践していきたいと思えます。最後になりますが、本セミナーの企画運営をはじめ、資料の作成や質問しやすい環境づくりなど、多方面にわたりご尽力くださいました先生方に深く感謝申し上げます。次回以降の開催も心より楽しみにしております。

第 15 回 ROC セミナー参加記

「わからない」から「少しわかる！」へ、ROC の見え方が変わった体験

山口労災病院 中央放射線部 **皆川 圭太**

放射線画像の評価に携わる中で、ROC 解析や ROC 曲線を目にする機会は少なくありません。しかし私自身、ROC 解析に対して「難しい」「統計が分からないと理解できない」といった苦手意識を持ち、十分に理解しないまま文献を読んでいたのが実情でした。そのような立場で参加した第 15 回 ROC セミナーは、ROC 解析に対する見え方を大きく変える貴重な機会となりました。

このセミナーでは、ROC 解析に対する心理的ハードルを下げ、初心者にも分かりやすい形で解説が行われていた点が印象的です。統計学に十分な自信がなくても、配布される解析ソフトを用いることで実際に ROC 解析を体験でき、理論だけでなく実感を伴って理解を深めることができました。講義の中では「ここを誤解している人が多い」といった、よくある勘違いについても丁寧に説明され、理解が曖昧なままになりがちなポイントを整理することができました。

また、物理評価と臨床画像評価の関係についても多くの気づきがありました。物理評価を行っていても、必ずしも臨床における画像の見え方と一致しない場面は少なくありませんが、その大きな要因が視覚評価にあることを改めて認識しました。これまで「なんとなく見やすい」「少し違和感がある」といった抽象的な判断にとどまっていた画像評価について、なぜそのように感じるのかを ROC 解析という枠組みの中で考えることができるようになったように感じます。

セミナー受講後は、文献に掲載されている ROC 曲線の捉え方にも少し変化がありました。ROC 曲線は単に「なければ通らない解析結果」ではなく、著者がどのような評価視点を持って、何を示したいのかという意図が反映された重要な情報であると感じるようになりました。全くわからなかった ROC 解析が、完璧ではないにせよ「少しわかる」状態になっただけで、その見え方は大きく変わったと感じました。本セミナーは、研究を始めたばかりの方や、評価方法に悩んでいる方、ROC を文献で見たことはあるものの理解に自信がない方にこそぜひ参加を勧めたいです。ROC 解析が少しでも気になっているのであれば、参加して損はなく、自分自身もっと早く受講しておけばよかったと感じたセミナーでした。

第 15 回 ROC セミナーに参加して

社会医療法人北晨会恵み野病院 千葉 峻介

今回、第 15 回 ROC セミナーに参加しました。形式はオンラインで毎週 1 時間程度の講義が全 5 回行われました。業務終了後の疲れている時間ではありますが、勉強するには適度な長さの講義であり、また北海道在住の私にはオンライン講習は参加しやすい形式でした。講義間隔も十分とられていたので、前回の復習や次回の予習を行うことができました。半面、オンラインで周囲の顔が見えない状況でしたので、勇気が出ずに質問できなかつたことが悔やまれます。

内容としては第 1 回から第 4 回までを白石順二先生が視覚評価法の基礎、ROCKIT を使用しての ROC、FROC、対比較法と講義して頂き、第 5 回で原武史先生が統計的処理について講義して頂きました。私は ROCKIT 自体を以前より使用していましたが、ファイルの形式で画像が表示できない、解析が進まないといった経験もあったので、使用方法を再確認できました。また、講義の合間や質問コーナーで漏れ出てくる白石先生の体験談が面白く、研究の参考になりました。研究の目的によってどの評価方法を使用するのか、資料の選定によるバイアスについて（読影医は画像を覚えている）、予備実験の重要性、統計量だけでなく個々の ROC カーブを解析することなど多くを学ぶことができました。

原先生の講義では研究の事前分析、事後分析に必要な統計量とその計算法について学ぶことができました。特に、サンプルサイズの設定に関して、いつもどの程度にすればいいのか悩んでいましたが、事前分析で効果量を算出することで解決できることを教えていただきました。事後分析では検出力を算出することにより、実験で求めた有意差の信用性を確認できることを学びました。今までの解析では p 値の算出で終わっていたので、今後は効果量と検出力も解析に含めていこうと思います。

今回の講義で視覚評価法に対する理解が深まりました。講義中に白石先生が何度も仰っていたように、実際に ROCKIT を自分で使用して講義を受けることで、内容の理解がより深まると思います。画像の視覚評価を行おうと考えている方は、ぜひオーム社の「医用画像の視覚評価法」を購入し、ROCKIT をダウンロードして、この ROC セミナーに参加することをお勧めします（利益相反事項はありません）。最後になりますが、本講義の開催にあたりご尽力して下さった講師の先生方やスタッフの皆様にご心より感謝を申し上げます。

第 12 回臨床画像評価セミナー ～マンモグラフィの画像評価の基礎～ に参加して

愛知県がんセンター 放射線診断部 小林 杏花

この度、2025 年 12 月に名古屋市立大学病院で開催された「第 12 回臨床画像評価セミナー～マンモグラフィ画像評価の基礎～」に参加いたしました。当院にはメーカーの異なるマンモグラフィ装置が 2 台あり、装置ごとの画質差や特性をどのように把握すべきか悩んでいたところ、本セミナーの開催を知りました。マンモグラフィ特有の評価事象を踏まえつつ、講義と演習を通じて効率的に学べる初級セミナーであったため、画像評価について基礎から学びたいと考えていた私にとって大変有意義な機会となりました。

セミナーは、事前の e-learning で物理特性や視覚評価の基礎知識を習得し、当日はより実践的な講義と演習が行われる構成でした。まず、マンモグラフィにおける代表的な評価項目について整理する講義があり、微小石灰化や低コントラスト腫瘍といった重要所見を描出するために必要な画質要素を改めて理解することができました。

物理評価の講義および演習では、配布された画像を用いて、ImageJ と Excel 解析シートによる MTF・NNPS 評価の手法を学びました。オーム社のテキストと演習資料に沿って進められ、一人でテキストを読むだけでは理解が難しかった部分も深く学ぶことができました。当初は演習についていけるか不安もありましたが、進行に遅れた際には講師の先生方が個別に丁寧にサポートしてくださり、質問しやすい環境で最後まで取り組むことができました。これまで画像評価に触れる機会がほとんどなかった私にとって、評価の流れを実際に体験できたことは大きな収穫でした。また、学んだ手順を簡略化できるマクロも用意されており、自施設での画像評価に応用する具体的なイメージを持つことができました。

ROC 解析の講義と演習では、試料の評価、解析を行いました。評定値の読み取り、並べ替え、集計、そして ROC 曲線のプロットまで、一連の流れを自分で実施しました。実際に手を動かしながら進めたことで、解析手順の理解がより確かなものになりました。

1 日という限られた時間ではありましたが、講義内容をすぐに演習で確認できる構成は非常に実践的で、理解の定着に大いに役立ちました。同時に、自施設での画質改善に向けた具体的な取り組みをイメージすることができました。日々の業務の中で画像評価の必要性を感じつつも、知識や技術が不足している状況でしたが、今回のセミナーで得た経験を、より診断能の高いマンモグラフィ画像の提供につなげていきたいと考えております。

最後になりますが、本セミナーを開催いただいた日本放射線技術学会教育委員会、画像部会、中部支部スタッフの皆様、ならびに講師の方々に心より御礼申し上げます。

第12回臨床画像評価セミナー参加記

～マンモグラフィの画像評価の基礎～ を受講して

医療法人志成会 阪本胃腸・外科クリニック 平松 佐和子

もし「マンモグラフィの画像評価を下さい」と突然言われたら、何をどう評価しましょう？

私事ですが、私は大阪府豊中市のマンモグラフィ検診導入初期から画像評価を担当しておりました。

当時はアナログフィルム全盛期。現像処理から撮影条件まで画質は技師の技量に左右されます。どんな工夫をしたらコントラストが得られるのか。撮影条件は適切か。画像を見て兆候をとらえて不具合を指摘する、撮影技師ごとにポジションのポイントを伝えるという役割を、私は昨年まで担っていました。そして、その「評価」とは、画像を観察して判断する「視覚」が主でした。

マンモグラフィ装置はデジタル化され、X線の出し方も画像の作り方もオートが主流に。技師が工夫してきた撮影条件も組み込まれ、精度管理的に口を出すところがないと感じて私は役目を降りました。さらに精中機構のおかげで各メーカーの画質はここ数年で大きな変貌を遂げたな…と感じる「今」になって、画像評価セミナーを「マンモグラフィ向け」に開催されるのを見て。私は吸い寄せられるように申し込みをしました。

講習内容は、基礎と言いつつ、熟練した評価者にもぜひ体験していただきたい内容でした。

一言でいうと、勉強になることが洪水のように押し寄せるセミナーと表現したい。

画像から簡単に得ることができる画素値について、その画素値はそのまま評価していいのか？他メーカー間での比較はどうするのか？といった、一筋縄ではいかないことをデジタルの歴史を絡めて基本からわかりやすく解説し、マンモグラフィで画質の指標になっている CNR の仕組みや画像の何を評価しているかにも触れつつ、画像のなかにある情報をどう取り出すかという切り口から、デジタル画像には欠かせない ImageJ マクロを使った実習をおこなう。そのあとには、ROC 解析を被検者と評価者を同時に体験する実習もありました。

事前学習で習ったことを、講義で詳しく解説しつつ、そのまま実習に移る授業形式なので、習ったことが定着しそうな、ImageJ のマクロが優秀でこのまま使いこなせそうだなという感覚につながります。また、実習中に何かつまづいた時には先生方がすぐに助けてくださるので、安心して実習を進められたのも非常にありがたかったです。

ImageJ を使いこなしている人や、自分の研究にすぐにつなげたい受講生は、実習後に講師の先生に直接質問されているのも目にしました。ImageJ に不慣れな私は新しい知識の連続についていくのが精一杯

でしたが、もしもデジタルで画像評価をするとなった時には、この講習で得たことで間違いなく進められるかもしれない。というか、自分で試行して先生方に尋ねる機会を得る意味でもう一度受講するというのも有りかなと思いました。

事前準備についても充実していて、動画視聴による事前学習がありました。実習に使用するエクセルをフーリエ変換できる仕様にする、ImageJのインストールとプラグインのコピーをしておく課題がありました。マニュアルも読みましたが、この事前準備、実は自分の行いが正しいのかがわからない。

メールでリマインドなどの案内があり、最後のリマインドメールを読んでやっと謎が解けるといった、ドキドキハラハラで準備を進めました。早いうちから準備を始める、特に動画視聴は私のように前日の晩とかではなくて、何回も見返せるように早めに視聴されることを強くお勧めします。

情報交換会にも参加させていただきました。講師の先生方のセミナーに対する思いや背景などを詳しく拝聴できたことが、なにより良かったと感じました。ここでさらに質問を投げて議論を深めることもできそうで、そのような空気を歓迎してくださる感じが、私には強く伝わってきました。

今回はマンモグラフィの画像評価に参加しましたが、自身の研究を解析する手段を探す意味で、ROCセミナーなどほかのセミナーも受講してみたいと思いました。

最後になりましたが、本セミナーの企画・運営をいただいた日本放射線技術学会画像部会をはじめ、講師の諸先生方に心より感謝申し上げます。

X線画像の視覚評価と試料数の決定について

和歌山県立医科大学附属病院 中央放射線部 山崎 純

はじめに

視覚評価はオーバーオールな画像の評価法として用いられている。例えば、撮影条件や画像処理が異なる画像について、その有効性を評価する場合に、物理的な評価だけで結論付けることはできない。最終的に、視覚的な評価を用いてより臨床に近い環境で画像を判断する必要がある。ここで、物理評価は、様々な物理指標において空間周波数特性を数値的に評価できるため定量的である。対して、視覚評価は、観察者の主観的な評価として、出力された画像の総合的な判断である。しかし、視覚評価ではバイアスが影響しやすいほか、評価法の種類、評価基準、観察試料などの影響も受けやすい。

実際に視覚評価された画像間において統計的な有意な差があるかどうかで判断される。その際、統計解析が適正に行われていない場合は、正しく結論付けられない可能性がある。そのため、視覚評価法の種類、統計手法の選択が重要となる。一般的に、統計手法では様々な分布の性質が利用されているが、特に正規分布の性質を利用した統計手法が用いられる。例えば、t検定におけるt分布は自由度の増加で正規分布と近似する。これより、標本データ数(試料数)が多いほど、正規性を利用した統計の精度は高くなると言える。試料数の増加は抽出される標本のバラつきを抑え、母集団の推定精度を上げることができると、試料数の決定は重要となる。

本講義では、視覚評価の基礎とともに、検証に際する統計手法と、試料数の決定について解説する。

1. 視覚評価の基礎

X線画像の診断過程は、露出(X線照射)、画像記録、観察(検出)、認識の4つの過程に分類される[1]。その過程を経て様々な条件下で取得された画像により診断がなされる。その画像の良し悪しを評価するために、直接判断できる手段として視覚評価が多く用いられている。それは、物理特性の評価だけでは把握できない「診断のしやすさ」を検証できる点に意義がある。しかし、人間の知覚・認知過程への依存性が強くバイアスがおきやすい。それには、観察者の経験年数や専門性、医学知識、判断閾値などや、症例の難易度、有病率の設定など、そして、提示順序、学習効果などによる認知のしやすさがある。そのため、堅実な計画を立てて観察環境を整えると共に、バイアスを抑えることが精度の高い視覚評価の検証に繋がる。そのため、研究計画を策定するには、評価目的、対象、検証方法、評価指標などを明確にして、「何をもちいて良い画像とするのか」を定義する必要がある。それを観察者の共通認識とすることで、バラつきを考慮した評価に繋がる。

臨床画像における異常の判断は、4つの反応として刺激-反応マトリクス(図1)で表すことができる。診断の対象として、異常がある(病気がある)ものと正常である(病気がない)ものに対する診断の結果が、異常(病気がある)もしくは正常(病気がない)と判断される4通りの診断結果がある。それは、真陽性(True Positive: TP)、偽陰性(False Negative: FN)、偽陽性(False Positive: FP)、真陰性(True Negative: TN)で表される。ここで「診断の正確さ」を考えた場合に、病気が存在している医用画像を正しく「病気あり」と診断した確率(感度)、病気が存在しない医用画像を正しく「病気

なし」と診断した確率（特異度）によって定量化され、ともに高いことが診断の正確さが良いとされる。

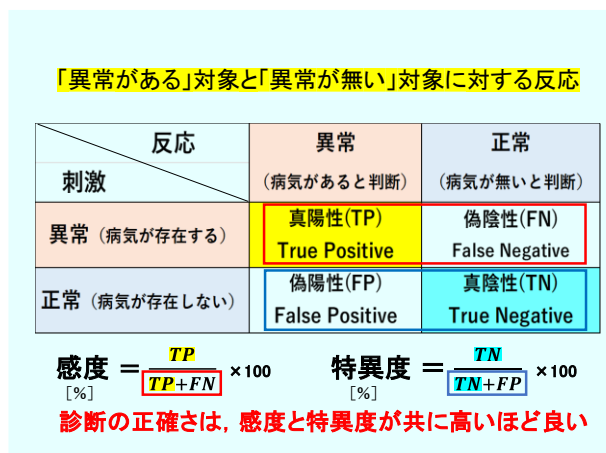


図1 刺激-反応マトリクス

視覚評価手法は心理学の分野より応用されており、視覚刺激に対する感覚、知覚、認知、感性などの心理量の変化を測定するものである。それは、物理的刺激と心理的反応の関係を定量化して評価する精神物理学的測定法と、物理的刺激に対して一定の尺度を構成し評価する尺度構成法に分類できる。精神物理学的測定法ではそれぞれの刺激に対して信号検出率を求めることで良し悪しを判定する。対して、尺度構成法では、主観的な感覚や印象に対して心理的な尺度 (scale) を作成して、それを測定した結果で判定するものである。ここで、精神物理学的測定法の代表的なものを上げる。極限法を利用したハウレットチャート法や C-D ダイアグラム法、恒常法による強制選択法 (AFC) などで信号検出率を求める。また、信号検出理論を応用した ROC 解析では、「信号があるもの」と「信号が無いもの」に対する反応分布より ROC 曲線が作成される。「感度」と「特異度」の両方の評価が可能であることから、診断の正確さの点から有効性が高い測定法である。次に尺度構成法では、心理的尺度を反映させたスケールをもとに評価される。このスケールは物理量ではなく直接数量化されたデータとして表される。尺度の種類では、質的データの尺度として名義尺度と順序尺度があり、量的データの尺度として距離尺度 (間隔尺度) と比例尺度がある。質的データは、分類やカテゴリー化することを目的としたデータであるが、量的データは、数値で表される連続または離散的なデータである。尺度構成法には、2 個ずつ一対を比較する一対比較法や、特定の記述に当てはまる程度をカテゴリー判断して順序尺度を構成する評定尺度法、反対の意味をもつ極性のどちらに近いかに判断させる方法を用いて順序尺度構成させる SD 法などがある [2,3,4]。

主観的評価の対象として、様々な模擬ファントムや物理評価用ファントムにより作成した画像や臨床画像が用いられる。その画像から信号を判断するために視覚と視知覚のどちらかを選択することになる。一般的な視覚が必要とされるか、病気 (異常) に関する専門的な知識が必要とされるかによって、適切な観察者が選出される必要がある。

医用画像を評価する目的は画像の質の良否であり、検出、鑑別、比較などの方法で判断される。その

結果で得られた評価データは統計的に有意な差があるかを判定し、統計量とともに母集団の優越性が推論される。診断の正確さでは、バイアスをできるだけ抑えるために、適正な視覚評価手法と統計手法を選択する必要がある。

1.1 ROC 解析について

視覚評価手法の一つである ROC 解析は、雑音の中から信号を検出する過程を扱った信号検出理論を応用している。刺激に対する反応の関係を受診者動作特性 (ROC : Receiver Operating Characteristic) として表された ROC 曲線は、「病気があるものに対する反応の分布 (TP(x))」と「病気が無いものに対する反応の分布 (FP(x))」より成り立つ曲線で、それぞれの曲線は正規分布の関数であると仮定して作成される。縦軸は病気の存在する判断基準として「真陽性率 (感度)」を、横軸は病気の存在する判断基準として「偽陽性率 (1-特異度)」とした分布になる。ここで、信号ありと信号なしの2つの反応分布は、正規性を仮定した分布であるため、それぞれに正規性を示す程度の試料数が必要とされる。そして、標本データの難易度や観察者の能力に強く影響されることより(図2)、複数の観察者における視覚評価の結果を平均化した ROC 曲線を用いることで精度が高くなる。

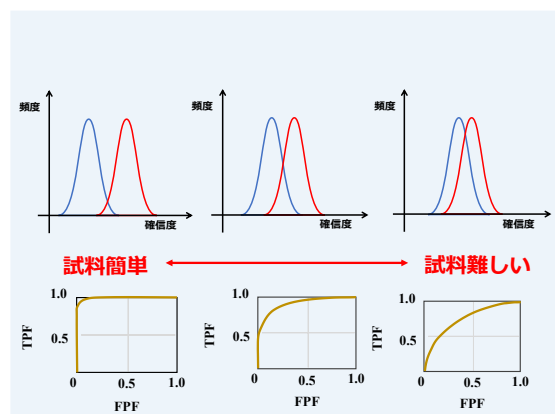


図2 試料の難易度が異なる ROC 曲線

2. 統計について

統計の利用により、堅実な研究計画を事前に策定することができ、研究の精度を示すことができる。解析により得られた統計量を用いて、標本の説明や要約が行われ、母集団の推定が可能になる。統計手法は、検証する目的によって優越性、同等性もしくは非劣性を調べる検定に分類される。放射線画像における視覚評価の検証では、しばしば優越性の統計が行われる。そして、優越性を検定するためには、検証したい対象群の間に「差が無い」とする帰無仮説と、「差がある」とする対立仮説を設定する必要がある。一般的に「差がある」ことを証明することは難しいため、「差が無い」ことの矛盾性より帰無仮説を棄却して対立仮説を採択する。すなわち、対象間に差がない確率 (p 値) が著しく低ければ、差が無いという仮説は成り立たず、差があるとする仮説が採択されることで、対象間の優位性が証明される。この際の判断基準は優位水準といい、一般的には 0.05 が用いられる [5]。

例えば、対象 A 群と対象 B 群より抽出された標本データに関して、平均値の差の検定である t 検定が用いられる。ここで、t 分布は連続データの統計手法で、試料数が多い場合に正規分布に近似すること

を利用して、標本間の平均値に差がない確率（p 値）の算出や母集団の平均値が存在する可能性のある区間（信頼区間）の推定を行う。すなわち、標本データを要約し抽出元の母集団の特性について推定が行われる。倫理審査では、統計手法に関する記載が必要であり、観察対象数（必要な試料数）及びその設定根拠（統計学的な根拠によらずに研究対象者数を設定する場合を含む）、統計解析の方法など、人を対象とした医学研究における倫理指針に示されている[6]。統計解析の役割では、検証で得られたデータの要約と説明、発生元の推測や新しく得られるデータを予測することであり、それぞれを要約統計、記述統計、および推測統計と呼ぶ。

対象となる発生元は母集団であり、標本（データ）を抽出して解析することで検証をおこなう。現実的に母集団全てを調べることは困難であることが多く、抽出された標本の統計量より母集団を推測する[図 3, 4]。

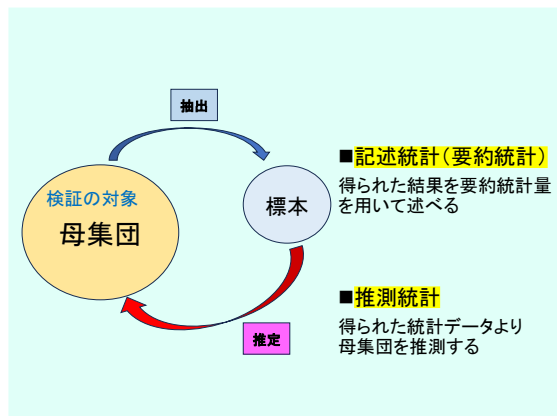


図 3 母集団と標本の関係

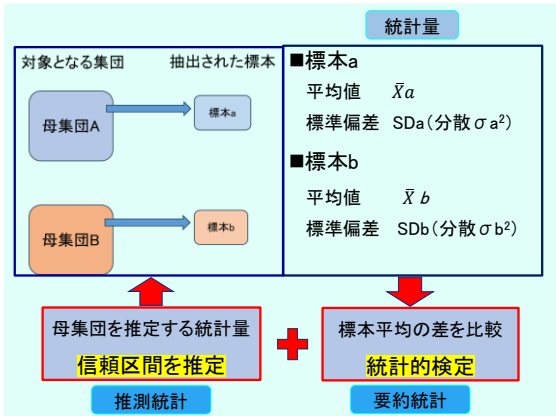


図 4 統計的記述と推論および要約統計量

統計手法の選択では、母集団から抽出された標本に対して、優位水準を決めたのち、母集団から抽出された標本の数（サンプル数）を調べ、連続数であるパラメトリックなもの（正規性あり）か、連続数でないノンパラメトリックなもの（正規性なし）か、対応の有無（同一母集団から抽出した標本かどうか）などを判断して決定される(図 5)。通常は分散が等しい（等分散）かどうかはわからないことが多く、その場合は分散を仮定しない検定を選択する。また、等分散かどうかを調べる F 検定があるが、多重性の問題があるため、使用には注意が必要である。

ここで、統計検定では P 値が優位水準を下回ることで有意差の有無を判定するが、P 値のみで結論付けることは推奨されていない[7,8]。それは、母集団から抽出される標本に限定される点推定のみでは、母集団の平均を正しく判断できないためである。そこで、母集団の平均が存在する可能性の高い信頼区間を推定することが推奨されている[9]。また、データが正規性を示すには、試料データ数（サンプルサイズ）が多く必要になる。計画時にどの程度必要かを設定された検出力、効果量、優位水準の関係より計算する必要がある。試料データ数の設定は統計的な判定に対して特に影響が大きいといえる。

統計解析の種類

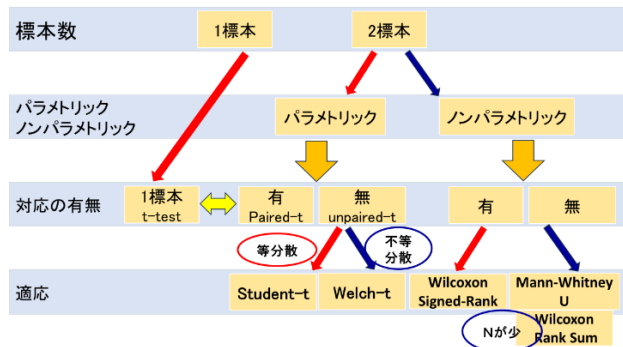


図5 統計手法の選択

2.1. 正規分布と t 分布

多くの自然現象は様々な分布で仮定される。統計学でもっともよく用いられる確率分布は正規分布（ガウス分布）である。連続型の確率変数に対して定義され、縦軸に確率密度（値が生じる確率）、横軸に確率変数（観測データ）をとる確率密度関数の変化である。正規分布は、平均（ μ ）とバラつきの指標である標準偏差（ σ ）で表され、 σ を二乗した分散（ σ^2 ）を用いて（ μ, σ^2 ）と表現される。平均（ μ ）を中心とした釣り鐘型で左右対称な分布となり、平均（ μ ）付近で値の頻度が最も多く、離れるほど低くなる。

正規分布の特徴として、平均値を中心としたバラつきが 1σ 、 2σ 、 3σ の範囲において、それぞれ68%、95.5%、99.7%となる確率で分布が起こる性質がある(図6)。これより、正規分布は平均と標準偏差（または分散）が分かれば分布の形状・位置が定まり、各値がどの程度の確率で現れるかを知ることができる。優位水準を0.05（5%）とした場合の信頼区間は95%信頼区間（95%CI: confidence interval）といい、 1.96σ （約 2σ ）の範囲をとる。通常は、母集団の平均やバラつきは不明であることが多いため、標本より得られた統計量を要約、推定することで算出される（図7）。

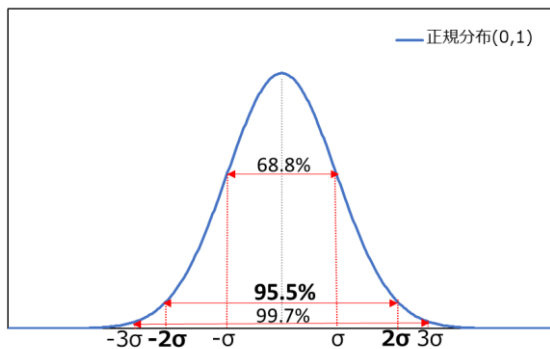


図6 正規分布と信頼区間の関係

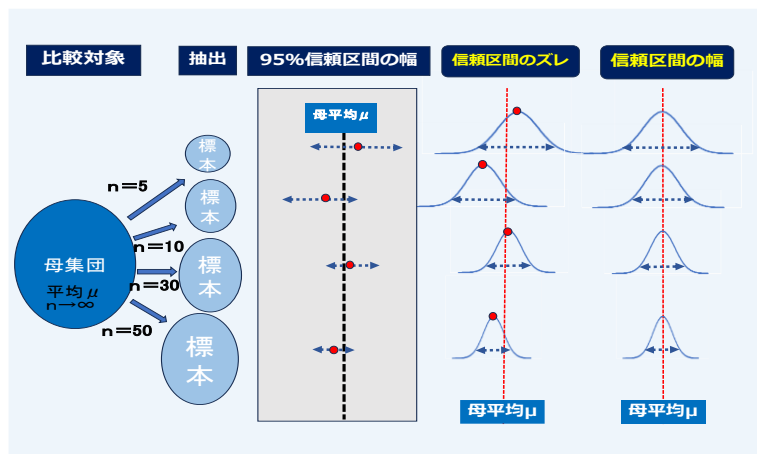


図7 標本と95%信頼区間の関係

統計手法を用いた標本間の差の検定では、標本平均の間に統計的な差が認められるかを検証する。その時、標本平均のデータが連続数であれば t 分布を仮定した t 検定が用いられる。t 分布は試料数が増加するにつれて正規分布に近づく性質があり、比して分布の裾野が長いことが特徴である。その唯一のパラメータは自由度であり、試料数を [n] としたときの自由度は [n-1] で表される。そして t 分布表を用いた値によってどの程度正規分布に近づいているかを確認することができる。しかし、試料数が少ない場合には近似されないため注意が必要である。特に裾野の広がり正規分布のそれとは異なってくる(図8)。そのため、試料数が大きい場合は正規分布の性質を利用し、試料数が少ない場合は t 分布の性質を利用することになる。ここで、95%信頼区間の範囲を以下に示す。これは、母平均が100回中95回は存在する可能性がある区間となる。

◎正規分布に対して、正規分布の平均(μ)、標準誤差(σ)である場合

$$[\mu - 1.96\sigma < \bar{\mu} < \mu + 1.96\sigma]$$

◎標本に対して、標本の平均(\bar{x})、標準偏差(s)、試料数(n)である場合は

$$[\bar{x} - 1.96\sqrt{\frac{s^2}{n}} < \bar{x} < \bar{x} + 1.96\sqrt{\frac{s^2}{n}}] \text{ または } [\bar{x} - 1.96\frac{s}{\sqrt{n}} < \bar{x} < \bar{x} + 1.96\frac{s}{\sqrt{n}}]$$

以上の範囲で表される。試料数が少ない場合に、標準偏差が大きくなることと、t 分布の裾野の広がりの影響のため、誤差が大きくなる。余談であるが、標準誤差(SE: standard error)と標準偏差(SD: standard deviation)の違いに触れておくが、標準誤差は母集団の性質として表される精度であり、母集団を推定する場合に用いられる。その際は正規分布の性質を利用した信頼区間などが用いられる。また、標準偏差はデータに対するバラつきを示す統計量となる。それを区別するには、標準誤差は母集団の精度や推定する意味合いで、標準偏差はデータ全てに対するものとして判断すると良い。

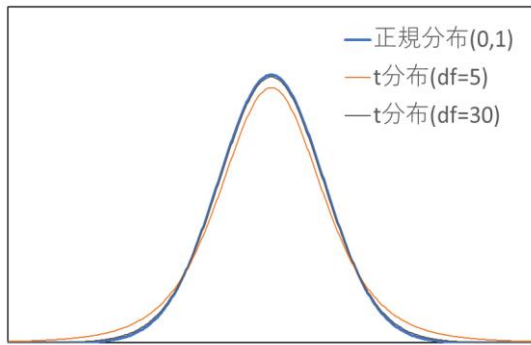


図8 正規分布と t 分布（自由度 5,30）の比較

2.2. 中心極限定理と大数定理

母集団の平均と標本平均の間には、標本データ数に関係した大数の法則と中心極限定理が利用されている。まず、標本データ数を大きくしていくと標本平均 (\bar{x}) が母集団平均 (μ) に限りなく近づく大数定理が成立する。そして、母集団がどのような分布であっても、標本データ数が大きいときに標本平均 (\bar{x}) の分布が正規分布に近似できる中心極限定理が成立する。すなわち、2つの定理の成立条件は試料数が多い場合となる。正規分布に近似させることで、その性質を利用した母集団の推定が可能になる。

3. 試料数について

研究計画時に収集される標本内の試料数（サンプルサイズ）を設定する必要がある。前向き研究では収集を予定するデータ数、後ろ向き研究では使用する既存のデータ数を具体的に示さなくてはならない。試料数の算出には算定ソフトを利用すると便利である。引用実績のある無料統計ソフトでは、ドイツのデュッセルドルフ大学（Heinrich Heine University Düsseldorf）による G-power を紹介する。統計学における検定力分析（Power Analysis）とサンプルサイズの算出について、事前および事後の検証に使用できる。また、自治医科大学の EZR（easy R）は簡易に扱い易く初心者であってもわかりやすい。

ICH-E9 では、臨床研究における研究の質について明言されている。そのためには、優位水準、効果量、検出力を決めて、ばらつきを抑えるための試料数を算定する必要がある。ここで、優位水準 (α) は帰無仮説を棄却する限界値より外側、棄却域になる確率を示す。一般的に 0.05 (0.01) が用いられ、p 値の判断基準となる。0.05 は帰無仮説が正しい時に、平均的に 20 回に 1 回程度、帰無仮説を採択してしまう誤りを許容すること表している。0.01 はより厳しい基準の場合に使用される。効果量 (effect size) は標本間の平均値の差の程度や値の関連の強さを表し、標本間の効果の大きさを知る指標となる (図 9)。算出には Cohen's d 式 (①) がよく用いられる。検出力 ($1-\beta$) は標本間の有意差の出やすさである (図 10)。これは、標本間に差がある仮説 (対立仮説) に正しく差があると判定する確率である。ここで、「 α 」は第一種の過誤を起こす確率で、差が無い仮説 (帰無仮説) に対して差がある仮説 (対立仮説) を正しいとした場合のエラー、「 β 」は 2 種の過誤を起こす確率で、差がある仮説 (対立仮説) に対して差が無い仮説 (帰無仮説) を正しいとした場合のエラーである。ここで効果量 (ES : effect size) について、標本 A の試料数 n_a 、平均値 \bar{X}_a 、標準偏差 S_a 、とし、標本 B の試料数 n_b 、平均値 \bar{X}_b 、標準偏差 S_b とした場合の式を以下に示す。

効果量 $D = \frac{|\bar{x}_b - \bar{x}_a|}{s}$

◎2 標本におけるサンプルサイズが同じ標準偏差 $s = \sqrt{\frac{Sa^2 + Sb^2}{2}}$

◎2 標本におけるサンプルサイズが異なる標準偏差 $s = \sqrt{\frac{(na-1)Sa^2 + (nb-1)Sb^2}{na + nb - 2}}$

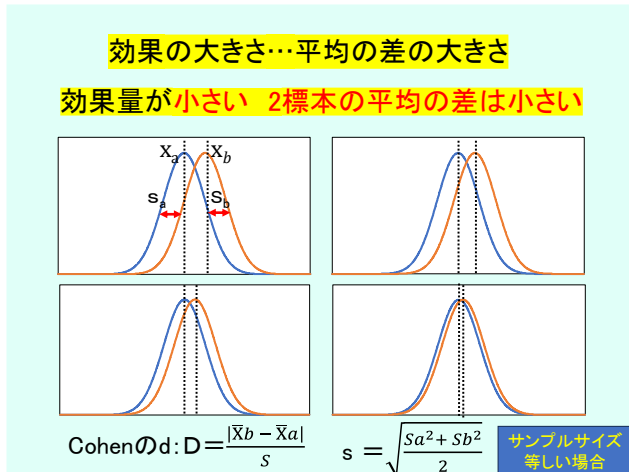


図9 効果量と標本平均の差の関係

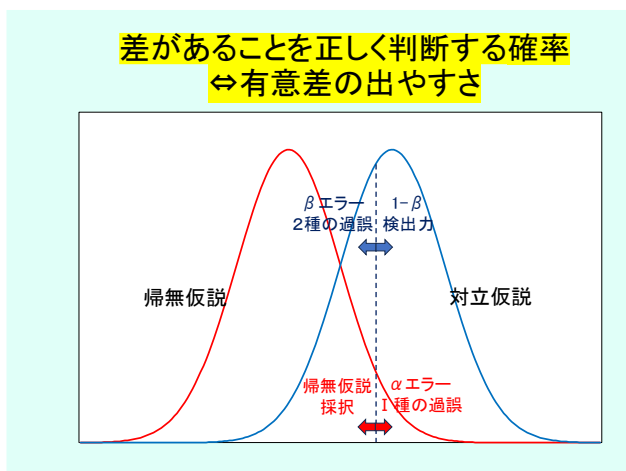


図10 検出力 (1-β) と第2種の過誤 (βエラー)

3.1. 試料数決定の実際

臨床研究における統計の扱いはICH-e9に示されているように、堅実な研究計画が必要になる。そのため、研究計画を策定する段階で、試料数を算定する必要がある、倫理審査においても必要とされている。試料数の算定ではパラメータが複雑であるため、一般的に統計ソフトが利用される。ここで、G*power3というフリーソフトを紹介する。これはドイツのデュッセルドルフ大学により制作されたもので、検定力分析 (Power Analysis) やサンプルサイズの算出について、いろいろな統計パターンで試算することができる。ここで、2つの対象の比較や評価によく用いられる頻度の高い2標本における両

側 t 検定を用いて試算を行った。優位水準 0.05，検出力 0.8，このレベルを 0.3（低い），0.5（中），0.8（高い）の 3 つに設定した場合に，必要とされる試料数はそれぞれ 176，64，26 であった。効果量の 3 つのレベルは標本の差の大きさであるが，試料の難易度と考えることができる。このソフトでは，事前の検証に必要な試料数がどれくらいになるか試算する以外に，研究後に行う検証（Post Hoc）や効果量の算定にも使用することができる。

例として，シミュレーションデータを用いた試料数の違いにおける検出力と効果量の関係を図 11 に示す。

視覚評価の結果は t 分布で表すことができる。試料数の増加とともに正規分布に近似される性質より，t 分布の臨界値 $t_{0.975,df}$ と標準正規分布の臨界値 $z_{0.975} = 1.96$ との差異について，正規性がどの程度近似されているかを確認する。正規分布 $z_{0.975} = 1.96$ との差異を考えた場合に，自由度が 60 の $t_{0.975,df}$ は 2.000 となり差異は +0.040，自由度が 29 の $t_{0.975,df}$ は 2.045 となり差異は 0.085 となる。この値が正規性を示す一つの目安と考えると，二つの分布の差は 0.5~1% と考えることができる。そのため，試料数 30（自由度 29）以上であれば，誤差はあるものの正規分布とみなして使用することが可能と考える。ただし，30 未満の標本では母平均の推定をする際に発生する誤差の大きさを小さく見積もってしまう危険性があることに注意しなくてはならない[10]。そのため，統計の不確かさという点では問題は残るが，正規性を示す一つの目安として試料数 30 を臨界点と考える。

臨床研究では，診断の正確さを示す点で有効な ROC 観察者実験が用いられる。その解析で使用される「異常があるもの」と「異常が無いもの」への反応の分布は正規分布であると仮定される。即ち観察結果の試料データは正規性を持つデータであることが条件となる。ここで，観察試料が正規性をもつデータであるためには，試料数が多く必要になる。しかし，希少な対象では症例数の収集が非常に困難になることが予想される。ここで，過去に実施された ROC または FROC 観察者実験の 85% 以上では試料数が 40 以上であったことが示されている。「信号あり」と「信号なし」をそれぞれ 20 以上として，同数にすることで標準誤差を最小化できるため効果的である。そして，観察者数は観察者間の変動を抑えるため 5 名以上が良いとされる。臨床症例を使用した検証では様々な限界があるのも事実であるが，その根拠を明確にして計画することが重要になる[11]。

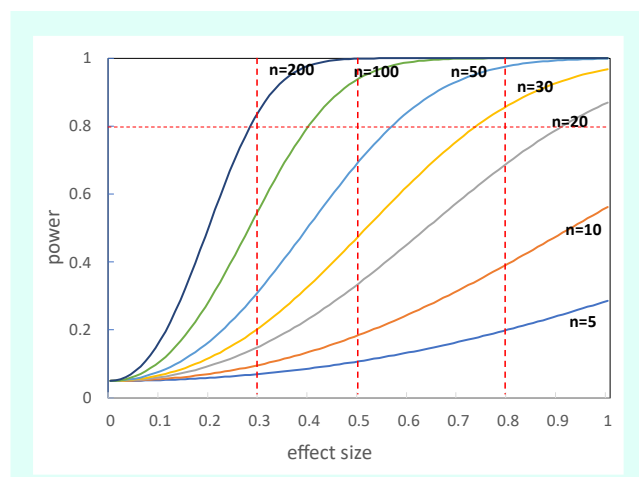


図 11 試料数の違いにおける検出力と効果量の関係

最後に

本講義では、X線画像の視覚評価に用いられている評価法と試料数の関係について、統計的な結びつきを解説した。母集団の推定ではP値による点推定と95%信頼区間を交えた区間推定を説明して、有意水準、検出力、効果量の関係から試料数の算出を述べた。検証の精度において試料数は最も影響がある因子で、研究計画を策定する段階から根拠と共に示す必要がある。研究統計に基づき精度を示すことは国際的にも注目されている。

確率的な起りやすさと効果の大きさは異なる事象であることを踏まえ、検証の意義と臨床における価値がともに高い研究デザインを目指されることを願います。この講義が問題を解決する一助となれば幸いです。

参考文献

- [1] Rossmann K., Radiology, 1968; 90(1), 1-13
- [2] 日本放射線技術学会, 研究方法論と基礎統計学, メディカルレビュー, 2015
- [3] 君山由良, 統計解説書シリーズ C-05 一対比較法のモデル, データ分析研究所, 2010
- [4] 君山由良, 統計解説書シリーズ C-07 一対比較法のモデル 最大値選択のモデル, データ分析研究所, 2016
- [5] 阿部真人, データ分析に必須の知識・考え方 統計学入門, ソシム, 2021
- [6] 人を対象とする生命科学・医学系研究に関する 倫理指針 ガイダンス, 厚労省, 2023
<https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/hokabunya/kenkyujigyou/i-kenkyu/index.html>
- [7] Wasserstein RL & Lazar NA., The American Statistician, 2016;70(2):129-133.
- [8] Scientists rise up against statistical significance Nature, 2019.
- [9] 臨床試験のための統計的原則, 医薬審 第1047号 平成10年11月30日, 1998
- [10] 栗原伸一, 入門統計学, オーム社, 2011
- [11] 白石順二, 標準 医用画像の視覚評価法, オーム社, 2020
- [12] 石田隆行, よくわかる医用画像工学 改訂2版, オーム社, 2015
- [13] 放射線技術学叢書 ICRU レポート 79 (日本語本宅) 医用画像のROC解析, 日本放射線技術学会, 2012
- [14] 宮川公男, 基本統計学 (第5版), 有斐閣, 2022
- [15] 足立堅一, 実践統計学入門, 篠原出版新社, 2001
- [16] 永田靖, サンプルサイズの決め方, 朝倉書店, 2003

Python 入門とバイブコーディングの可能性

岡山大学 学術研究院 保健学域 福井 亮平

はじめに

近年, Python はその扱いやすさと高い汎用性から, 研究, 教育, 産業のあらゆる分野で急速に存在感を高めている. シンプルな文法と豊富なライブラリに支えられ, 初心者から専門家まで幅広く利用されている. 本稿では, Python の歴史や特徴, 代表的なツールを概観するとともに, 近年注目される「バイブコーディング」という新しいコーディングスタイルを紹介する. さらに, その実現を支えるツールや, 鍵となるプロンプトエンジニアリングについても概説し, これからのプログラミングとの向き合い方を解説する.

1. Python 入門

1.1 Python とは

Python は高水準の汎用インタプリタ言語である. オランダ出身のプログラマである Guido van Rossum 氏によって開発

された. Python は 1991 年にバージョン 0.9.0 として初めて公開され, その後も継続的に改良が重ねられ, 現在の最新バージョンは 3.14.3 (執筆時) に至っている. 「同じ動きをするためのコードは, できるだけ 1 つの方法でしか書けないようにする」という設計思想によりコードの可読性を高め, さらに, 他の言語に比べ短いコードで目的を達成できる. Figure 1 に示したのは「Hello World!」と出力するコードを C 言語と Python で比べたものである. これだけで, 初心者が Python を学ぶ理由としては十分である. 近年では, 機械学習やデータサイエンス分野の発展を背景に, Python の注目度は飛躍的に高まっている. Google 検索においてもプログラミング言語関連の検索数で 1 位を獲得している [1]. なお, Python という名称はイギリスのテレビ局 BBC のコメディ番組「空飛ぶモンティ・パイソン」に由来する.

C言語	Python
<pre>#include <stdio.h> int main(void){ printf("Hello World !%n"); return 0; }</pre>	<pre>print("Hello World!")</pre>
Output: Hello World!	

Fig. 1 C 言語との比較

1.2. Python の特徴

1.2.1 オブジェクト指向

Python はオブジェクト指向を中核に据えたプログラミング言語である. 数値や文字列といった基本的なデータ型も含め, ほぼすべてがオブジェクトとして扱われる点が特徴で, 大規模なソフトウェア開発においても構造化された設計が可能となる. 一方で, 記述が過度に冗長にならないよう配慮されており, オブジェクト指向に不慣れな初学者でも段階的に理解できる設計となっている.

3.2 オフサイドルール

オフサイドルール (文頭にインデントを挿入すること) は, Python において最も特徴的な仕組みの 1 つである. コードブロックを波括弧ではなくインデントによって表現する (Fig. 2). これにより, 見た目の

構造と実行上の構造が一致し、可読性の高いコードが自然に書かれる。インデントの揺れは構文エラーとして検出されるため、曖昧な記述を防ぐ効果もある。一方で、厳格な構造化ルールにより、プログラマがコーディングスタイルを選ぶ自由度を狭めているという批判的意見もある。

C言語	Python
<pre>#include <stdio.h> int main(void){ int i, j; int sum=0; for(i = 1; i < 4; i++){ for(j = 1; j < 4; j++){ printf("Number = %d\n" , i * j); sum = sum + i * j; } printf("sum = %d\n" , sum); return 0; } }</pre>	<pre>sum = 0 for i in range(1, 4): for j in range(1, 4): print("Number = ", i * j) sum = sum + i * j print("sum = ", sum)</pre>
Output: Number = 1, Number = 2, Number = 3, sum = 6	

Fig. 2 オフサイドルール

1.2.2 動的な型付け

Python は動的型付けを採用しており、変数の型を事前に明示する必要がない。他言語を知っている人間には驚異的なことである。実行時に型が決定されるため、コードは簡潔になり、試行錯誤を伴う開発やプロトタイピング（実働モデルを早期に構築）に適している。この柔軟性は生産性を高める一方、実行時エラーが発生しやすいという側面も持つ。そのため、型ヒントや静的解析ツールを併用することで、可読性と安全性を両立させる運用が一般的となっている。

1.2.3 Battery Included

「Battery Included」と呼ばれることがあり、Python をインストールした時点で標準ライブラリが非常に充実していることを指す。ファイル操作、正規表現、ネットワーク通信、並列処理など、多くの機能が最初から利用可能であり、外部ライブラリに依存せずとも実用的なプログラムを構築できる。この思想は、環境構築の手間を減らし、学習者や開発者が本質的な問題解決に集中できるよう配慮されたものである。

1.2.4 ガベージコレクション

Python は自動的なメモリ管理機構としてガベージコレクションを備えており、使用されなくなったオブジェクトは自動的にメモリから破棄される。したがって、一部を除けば、開発者が明示的にメモリ解放を行う必要はない。これにより、メモリ管理に起因するバグを減らし、安全性の高い開発が可能となる。一方で、処理タイミングを完全に制御できない点には留意が必要である。

1.2.5 マルチパラダイム

Python はオブジェクト指向をベースとしつつ、手続き型や関数型プログラミングの性質も兼ね備えたマルチパラダイム言語である。関数を第一級オブジェクトとして扱えるほか、高階関数やラムダ式をサポートしているため、極めて柔軟なコード記述が可能だ。問題の性質に応じて最適な設計手法を使い分けられる柔軟性こそが、Python が幅広い分野で支持されている大きな理由の一つといえる。

1.2.6 Python のバージョン

Python には 2.x 系と 3.x 系が存在する。しかし、2.x 系は 2020 年 1 月のバージョン 2.7 をもってサポートが終了している。これから Python を学ぶ際は間違いなく 3.x 系を学ぶことになるが、現在でも 2.x 系で動いているシステムは存在する。しかし、3.x 系に後方互換性が無いため、2.x 系は 3.x 系で動かない。欠点らしい欠点の無い Python において、この問題は特にシステムの保守などで顕在化している。

1.3 Python の導入

Python を自身の PC で使い始めるには、Python の公式サイトから環境に合ったインストーラーをダウンロードし、インストールすれば完了である。しかし、Python 単体で使用することは効率的でなく、統合開発環境（Integrated Development Environment, IDE）を用いることが多い。IDE を用いることでより効率的に Python のコーディングや管理を行うことが可能となる。以下に、Python を活用する上で知っておきたい IDE やコードエディタについて紹介する。

1.3.1 Anaconda

Anaconda は、Python を中心としたデータサイエンス向けの開発環境統合プラットフォームである。Python 本体に加え、NumPy や Pandas, Matplotlib, Scikit-learn など主要なライブラリがあらかじめ同梱されており、インストール直後から実用的な開発環境を構築できる点が特徴である。後述する Jupyter Notebook/Jupyter Lab, Visual Studio Code といった IDE やエディタも Anaconda 環境下で使用可能である。また、conda コマンドによる仮想環境管理機能を備えており、プロジェクトごとにライブラリや Python のバージョンを分離して管理できる。これにより、環境依存によるトラブルを低減し、安定した開発を行うことが可能となるため、多くのユーザーを獲得している。2020 年 4 月以降の利用規約変更により、商用利用や 200 人以上の企業における使用が有料となった。Python により研究環境などを立ち上げる場合、他の環境と干渉しないように仮想環境を設定することは必須である。Anaconda 以外の選択肢としては、例えば Python 標準ライブラリである venv の使用や、Docker を使用する方法がある。しかし、ライブラリを手動でインストールする手間やコンテナの操作方法を覚えるといった学習コストが発生するため、現在でも Python の環境構築において Anaconda は第一選択として挙げられる。また、Anaconda の軽量版である Miniconda は引き続き無料で使用できるため、こちらも選択肢に入る。

1.3.2 Jupyter Notebook/Jupyter Lab

Jupyter Notebook は、ブラウザ上で Python コードを対話的に実行できる IDE である。「対話的」とは、セルと呼ばれる入力欄にコードを入力すると、その直下に実行結果またはエラーが出力されることである (Fig. 3)。この機能があることで、Python 初学者

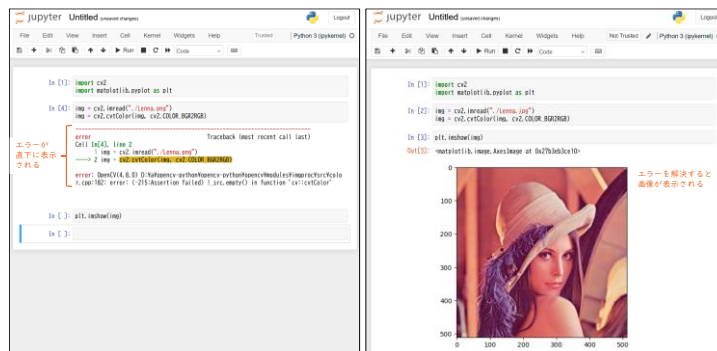


Fig. 3 Jupyter Notebook の実行画面

であってもコードの記述・出力・修正を容易に行うことが可能である。コード、実行結果、数式、文章を一つのノートブック内にまとめて記述できる。データ解析や可視化との相性が良く、研究や教育現場において広く利用されている。Jupyter Lab は Jupyter Notebook の拡張版で、さらに使い易い機能が追加されている。Jupyter Notebook/Lab は Anaconda へ同梱されているため、Anaconda をインストールすれば使用可能である。しかし、保存形式が汎用性の低いノートブック形式 (.ipynb) であること、コードがセルごとに分かれているためにコードの実行する順番により結果が変わる危険性があるなど、これらを理解して使用する必要がある。

1.3.3 Google Colaboratory

Google Colaboratory (Colab) は、クラウド環境で使用可能な Jupyter Notebook である。Google アカウントがあれば使用可能で、Google ドライブからアクセスできる。Colab の最大の利点は環境構築が不要な点である。Colab でノートブックを開くだけで、主なライブラリを使用することが可能である。また、機械学習や深層学習などの効率的な学習には GPU が必要であるが、Colab では計算リソースとして GPU を選択可能である。ただし、無料枠では 90 分間ブラウザ操作が無い場合には処理が停止し、1 つのノートブックの処理は継続して 12 時間で停止するため、これらを考慮したコーディングやブラウザの設定が必要となる。しかし、これらの欠点を加味しても十分な機能を有するため、Python の勉強を始めるには最適なツールである。

1.3.4 Visual Studio Code

Visual Studio Code (VScode) は 2015 年に Microsoft が発表した数多くのプログラミング言語をサポートする IDE である。IDE またはコードエディタとしては後発で、Notepad++, Atom, Emacs など著名なエディタが存在する中で、VScode は OS や言語に依存しない IDE として、現在では高いシェアを獲得している。拡張機能が充実しており、様々な開発環境シーンへ柔軟に対応可能である。また、開発されたコードのアーカイブである GitHub との連携もシームレスに行えるため、組織内のコード管理やアップロードされたコードの実行も容易に行うことが可能である。拡張機能により Jupyter Notebook を使用可能で、2025 年 11 月からは Google Colab も VScode 内で実行可能となった。Anaconda をインストールすると VScode も使用可能となるが、軽量な開発環境を望む場合は VScode 単体でインストールしても良い。

1.3.5 marimo

marimo (マリモ) は、Python コードを対話的に記述・実行できる次世代型のノートブックエディタである。Jupyter Notebook が抱えてきた実行順序の依存や状態管理の不透明さといった課題に対し、marimo はリアクティブな実行モデルを採用している。各セル

は依存関係に基づいて自動的に再実行され、常にコードと実行結果の整合性が保たれる点が特徴である。例えば Fig. 4(a)に示すように変数 x と y を定義した際、 y の値を変更すると、依存関係にある「`print("X+Y=", x+y)`」のセルも更新される (Fig. 4(b))。大きなコードになると変数や関数同士の関係が見えにくくなりエラーの原因になることから、このようなリアクティブ機能は開発の頑強性を向上させる。また、ノートブックは Python スクリプト (.py 形式) として保存されるため、py 形式であれば GitHub との連携も容易である。

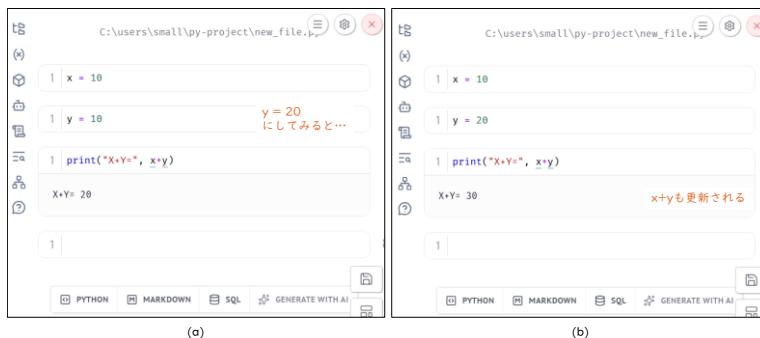


Fig. 4 marimo の実行画面

1.3.6 marimo の導入

Anaconda や VScode の導入方法については多くの良書や Web サイトがあるため、ここでは marimo の導入について簡単に説明する。Python の環境設定において、ライブラリの追加・管理・削除に最も頻繁に使用されるパッケージ管理マネージャーは pip である。しかし、ライブラリの依存関係の競合や、conda などの他ツールを併用すると開発環境を壊してしまうことがあり、これらが問題視されている。近年、注目されているパッケージ管理マネージャーに uv がある。uv は Rust というプログラミング言語で記載された高速な Python 管理ツールであり、同時に仮想環境マネージャーでもある。

以下の操作は Windows を想定している。uv の導入は PowerShell を「管理者として実行」し、以下のコマンドを打って Enter キーを押す。

```
>>powershell -ExecutionPolicy Bypass -c "irm https://astral.sh/uv/install.ps1 | iex"
```

このコマンドにより uv の最新版がインストールされる。また、Python のインストールは、例えば uv.exe ファイルが存在するフォルダへ移動し、

```
>>uv python install 3.11 3.12 3.13
```

とすることで、複数のバージョンの Python をインストール可能である。

```
>>mkdir test-project
```

#test-project というフォルダを作成

```
>>cd test-project
```

#test-project フォルダへ移動

```
>>uv init -p 3.12
```

#仮想環境を作成、Python 3.12 を使用

上記のコマンドでプロジェクト用のフォルダ作成、フォルダに仮想環境を構築する作業が終了する。ライブラリのインストールには「uv add」コマンドを使用する。

```
>>uv add marimo
```

上記だけで marimo のインストールは完了する。作成した仮想環境は Python の標準ライブラリしか含まれていないため、必要なライブラリはこのコマンドでインストールを繰り返す必要がある。しかし、uv のインストールは pip に比べ数倍早く、大規模なライブラリでも短時間でインストール可能である。また、作成したフォルダの中には「pyproject.toml」と「uv.lock」というファイルが自動生成される。作成した「test-project」と同じ環境を別 PC に再現したい場合には、この 2 つのファイルを別 PC の「test-project」に相当するフォルダ内にコピーしておく、

```
>>uv sync
```

を実行するだけで、元の開発環境と同じ環境を再現可能である。

marimo を起動するには以下のコマンドを打つ。

```
>>uv run marimo edit new_notebook.py
```

#空のノートブック(new_notebook.py)が開く

.py ファイル名を既存のファイル名にすれば、既存のファイルが marimo で開く。使用感は Jupyter Notebook と同じで、各セルにコードを入力し、実行すれば直下に実行結果かエラーが返される。marimo の便利な機能として、仮想環境下に存在しないライブラリを import しようとする、GUI 上にインストールを促すウィンドウが現れ、シームレスにライブラリのインストールが可能である (Fig. 5(a))。また、生成 AI によるコード補間やコード生成機能を有している (Fig. 5(b))。

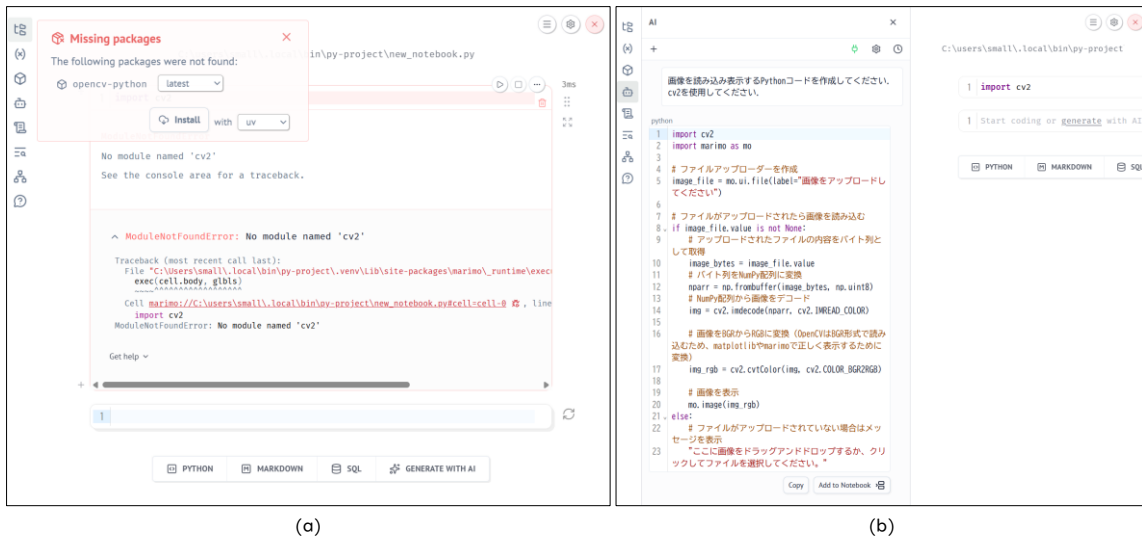


Fig. 5 ライブラリのインストール(a)と生成AIによるコード生成(b)

1.4 Python を使う

1.4.1 Python のライブラリなど

すでに標準ライブラリという言葉が出ているが、Python で多彩なコードを実現するためにモジュールやライブラリという機能の使用をコード内で宣言する必要がある。コード冒頭に、

>>import ライブラリ名

を記載することで、その機能を導入できる。モジュールは関数などが記載された 1 つの Python ファイルであり、import 可能な最小単位である。例えば、よく使う複雑な関数などを Python ファイルにして、その都度 import することができる。なお、import 不要で使用できる関数を組み込み関数と呼ぶ。モジュールを複数まとめたものをパッケージと呼ぶ。また、特定の目的でモジュール、パッケージなどをまとめたものをライブラリと呼び、Python と一緒にインストールされる標準ライブラリと、サードパーティが作成した外部ライブラリに分かれる。これらライブラリは膨大な種類が存在し、例えば Python ライブラリのリポジトリである Python Package Index (PyPi) には執筆時点でプロジェクトが 733,456 個立ち上がっており、このライブラリの豊富さが Python の拡張性を高めている。Table 1 にごく一部であるが、主な標準・外部ライブラリを示す。

Table 1 ライブラリと主な機能

標準/外部	ライブラリ名	用途・機能
	os	OSに依存した機能を使用できる。パスの取得やファイル名の取得などに使用する。
標準ライブラリ	glob	ファイルまでのパスを一括取得できる。大量のデータアクセス時に重宝する。
	shutil	高水準のファイル操作ライブラリ。ファイルの移動やコピーを使用する。
	random	ランダムな数値出力が可能。シード値など乱数が必要な場面で使用される。
	tkinter	GUIアプリを開発するためのライブラリ。標準ではあるが、簡単なコードで構築可能。
	Numpy	高速な数値計算を目的に使用する。特に多次元配列の操作に特化しており、科学技術計算、データ分析、機械学習分野では必須の基礎ライブラリ。
	Pandas	データ分析ライブラリ。データフレームという構造を用いてデータの読み込み、クリーニング、集計、可視化などを効率的にサポートする。
外部ライブラリ	Matplotlib	グラフの描画ライブラリ。折れ線グラフ、棒グラフ、散布図、ヒストグラムなど多様なグラフを、NumPyなどのライブラリと連携しながら作成可能。
	OpenCV (cv2)	コンピュータビジョンや画像処理を実行できるライブラリ。画像や動画の読み込みや表示だけでなく、フィルタ処理、特徴抽出、物体認識などの機能がある。
	Pydicom	DICOM画像を扱うためのライブラリ。DICOMヘッダの読み込みおよび編集が可能。ピクセル値はNumPy形式で読み込まれる。
	Tensorflow	Googleが開発した機械学習や深層学習のためのライブラリ。多次元配列を用いた数値計算を効率的に行え、深層学習のネットワークを簡単に実装可能である。

1.4.2 Python の文法

Python の記述方法を全て網羅するには紙面が足りないため、ここではコードを短く簡単に記述可能であることを確認したい。例えば、OpenCV を使って Lenna.jpg 画像を読み込み、Matplotlib で表示するには、次のようなコードになる。

```

>>import cv2                                #OpenCV をインポート
>>import matplotlib.pyplot as plt          #matplotlib を plt という名前でインポート
>>img = cv2.imread("./Lenna.jpg")         #cv2 の機能 imread で変数 img へ読み込み
>>img_rgb = cv2.cvtColor(img, cv2.COLOR_BGR2RGB) #カラーモデルを BGR から RGB へ変換
>>plt.imshow(img_rgb)                       #img_rgb を plt で表示

```

このコードの出力は Fig. 6(a)のようになる。数行で画像表示が可能であった。Pydicom を import すれば DICOM 画像も取り扱い可能である。また、グラフを表示するコードは以下のようになる。

```

>>import matplotlib.pyplot as plt          #matplotlib を plt という名前でインポート
>>x = []                                    #変数 x を初期化
>>y = []                                    #変数 y を初期化
>>for i in range(10):                       #for 以下の処理を 10 回繰り返す
>>    x.append(i)                            #変数 x に i の値を追加
>>    y.append(i**2)                         #変数 y に i の 2 乗の値を追加
>>plt.plot(x, y)                            #x 軸に変数 x, y 軸に変数 y を入れ散布図化
>>plt.show()                                #グラフを表示する

```

コードの実行結果は Fig. 6(b)のようになる。for 文は「range(n)」の時、0～n-1 までがループ変数 i に代入され、その都度、for 文のブロック内の処理が繰り返される。for 文のブロックはインデント（オフサイドルール）で示されるため、「for i in～」から「y.append(i**2)」の行までが for 文のブロックである。上記のように画像やグラフの描画は短いコードで達成できる。また、Python を知らなくても意味のある程度予測可能なシンプルな文法である。



Fig. 6 コードの実行結果

2. バイブコーディングについて

2.1 バイブコーディングとは

バイブコーディングとは、開発者が細かな実装や構文に過度に意識を向けるのではなく、「やりたいこと」や「雰囲気 (vibe)」を自然言語で AI に伝えながらコードを生成・修正していく新しいコーディングスタイルである。これは、従来のコーディングの重心が「コードを書くこと」であったのに対し、「AI が生成したコードを管理する」ことへ移行したことを示している。この概念は、OpenAI の共同設立者である Andrej Karpathy によって 2025 年 2 月に提唱された。例えば、イスラエル発のスタートアップ「Base44」は、2024 年にバイブコーディングを活用し独力で起業、プロダクト開発を行った例として注目を集めている。Base44 は 2025 年にホームページ作成プラットフォームを提供する Wix に 118 億円で買収されている。

2.2 バイブコーディングツール

バイブコーディングを支えるツール群は、機能や関与の度合いに応じて大きく分類できる。まず、生成

AIの代表格である ChatGPT や Gemini, Copilot は汎用性の高い対話型 AI であり、パイプコーディングにも十分対応可能である。基本的に無料で使用できることも利点であるが、生成されたコードを検証する機能は有していない。次に、GitHub Copilot や Claude Code, Cursor, Gemini CLI はエディタや CLI (Command Line Interface) と統合されたコード特化型であり、既存コードを高速に書き進めたい実務開発者に適している。一方、Devin, Lovable, Manus はタスク遂行やプロジェクト全体を自律的に進めるエージェント型であり、少人数や個人でプロダクト開発を行う起業家に向く。AI (コーディング) エージェントは目標達成のための計画立案、コードの生成、検証、修正、ブラウザ・ファイル操作などを自律的に実行可能なソフトウェアの総称である。もう一人の相棒として優れた結果を出力可能であるが、十分に使用するには課金が必要である。そんな中、2025年11月にリリースされた Google Antigravity は、執筆時点では Gemini 3 pro を無料で使用可能で、AI エージェントを試してみたい方には最適である。以上のように、パイプコーディングに使用可能なツールは多種多様、群雄割拠の時代であり、自身の開発環境や開発目的、開発スピードおよびスキルに応じたツールを選択すべきである。また、複数の AI モデルを横断的に使用する手段をマルチエージェントと呼び、Antigravity にも実装されている。マルチエージェントを Python で組み上げるためのライブラリも登場しており、例えば OpenAI の Swarm などが有名である。

2.3 バイプコーディングによる Web アプリ作成

実際に簡単な Web アプリケーションをパイプコーディングで作成してみよう。最新の論文を日々検索するのは骨が折れるため、今回は掲載された論文を定期的にお知らせする Web アプリ作成を目指した。生命科学論文のデータベースである PubMed にアップされた論文のうち、任意のキーワードがタイトルに入った論文を新しい方から 3 つピックアップし、これらの Abstract を日本語訳して表示する仕様とする。また、同じ内容を同時に Gmail で受信できればスマホでも確認できるため、この機能も追加する。Web アプリ構築は Streamlit という Python のライブラリを使用する。これらの

内容をプロンプトへ落とし込むと、Fig. 7 のようになる。コードの作成には Antigravity を使用した。プロンプトをコピー&ペーストし実行すると計画立案とコード生成がスタートする。この間、筆者が行ったのは許可を求められた際に 3 回ほどクリックしただけである。Python ファイルが作成されたため、これを実行すると無事に Web アプリが表示され、キーワードへ入力すると 3 つの論文の Abstract が表示された (Fig. 8)。繰り返し時間も設定可能である。ここまで必要な時間は 10 分程度であった。実際には Gmail で日本語訳が届くという機能でエラーが出たが、これは Gmail API の設定の問題で、これを解決することで Gmail へも Abstract が届いた。ここまでで 40 分程度しか経過しておらず、開発速度が劇的に向上することを確認できた。なお、PubMed 内の論文をソフトウェア等で自動取得する場合、1 度に取得する文献数などに制限があるため事前に確認が必要である [2]。このような Web スクレイピングには利用規約等に抵触する場合もあるため、検索対象個々に確認して頂きたい。

```
### 目標
- PubMedに掲載された論文のタイトルと要約がGメールで届く

### 動作
- PubMedに掲載された最新の論文を抽出するWebアプリケーションを作成する
- ユーザーが入力したキーワードが入った論文のタイトルを抽出する
- 投稿日時が最も最近の論文3本を選出する
- 選出された論文の要約を日本語訳する
- 論文のタイトルと日本語訳された要約をダッシュボードに表示する
- この動作をユーザーが指定した繰り返し時間に1回繰り返し実行(デフォルトは24時間)
- 抽出された論文タイトルと日本語訳された要約をGメールで送信する

### 仕様
- 記述言語はPython
- 実行環境はローカルPC(GPUなし)
- Webアプリケーション構築はStreamlitを使用

### GUIの装飾
- 3本の論文を表示できるダッシュボードを用意
- 全体的に青がベースで知的な色合い
- 以下のウィンドウ内に表示する文字は全てゴシック体
- 「キーワード」を入力可能な検索ボックスを左端サイドバーに配置
- 繰り返し時間を入力できる「Repeat」ウィンドウを配置
- 次の検索までの残り時間を表示する「Timer」ウィンドウを配置
- 任意のタイミングで検索開始を実行する「Search」ボタンを配置
- ダッシュボード内に論文タイトルを表示する「Title」ウィンドウを配置
- ダッシュボード内に論文の要約を表示する「Abstract」ウィンドウを配置
- 要約はHTMLコードが表示されないよう注意

### Gメール
- GメールはGmail APIを使用
- Gメールの送信先は「〇〇〇@gmail.com」のみ
- Gmail APIについての情報はJSONファイルをフォルダ内に配置済み
- Gメールのスコープは「https://www.googleapis.com/auth/gmail.send」
```

Fig. 7 作成したプロンプト例

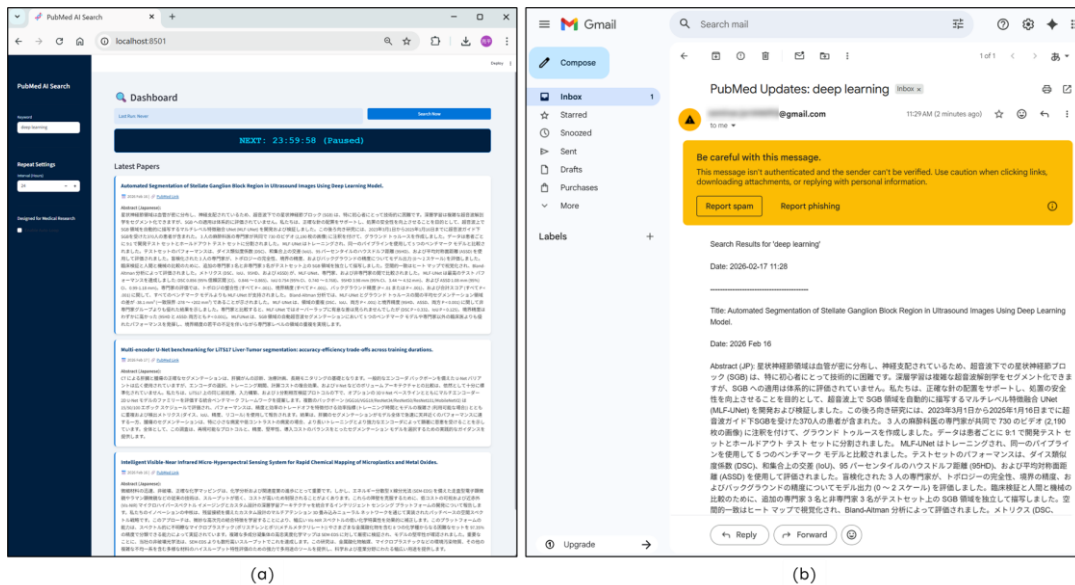


Fig. 8 作成された Web アプリ (a) と Gmail への受信結果 (b)

3. プロンプトエンジニアリング

バイブコーディングを効果的に実践するうえで、生成 AI への指示=プロンプトの作成は欠かせない要素である。バイブコーディングは「雰囲気」を伝えて開発を進めると前述したが、生成 AI は与えられた指示の内容や表現に強く依存して振る舞うため、適切なプロンプト設計は、生成されるコードの品質や再現性を大きく左右する。生成 AI の能力を最大限に引き出すためのプロンプト作成技術をプロンプトエンジニアリングと呼ぶ（他にプロンプトプログラミング、プロンプトデザインなど）。すでに様々なセミナーが企画され、関連書籍が発刊されており、認定資格も存在する。以下では代表的なプロンプトエンジニアリング手法を紹介し、実践的な指示の与え方について整理する。

3.1 基本的な手法

プロンプトエンジニアリングの基本から発展的な手法については、これらを解説した Prompt Engineering Guide を見ると詳しい [3]。Table 2 に示すように、具体的な指示や提供可能なデータ、最終的な出力形式などを整理して生成 AI にプロンプトとして渡すことが必要である。したがって、指示・目標が具体的であることは前提で、「Markdown 記法による箇条書き」や「ロールプレイ」、「制約を加える」といったことが基本的な手法になる。

Markdown 記法は特定の記号などを使ってテキスト形式の文章を手軽に HTML 形式の文章へ変換するために開発されたものである。インターネット上の文章は HTML で表現されており、Web 上のデータを

Table 2 プロンプトに含める内容

要素	説明
Instruction (命令・指示)	生成モデルが実行する指示やタスク
Context (背景・文脈)	出力の質を高めるために生成 AI に考慮させる背景・外部情報
Input Data (入力)	実際に生成 AI に渡すデータやテキスト
Output Indicator (出力形式)	出力したい形式、フォーマット

<Markdown記法>	<出力結果>								
# 大見出し	大見出し								
- リスト1									
- リスト2	• リスト1								
## 中見出し	• リスト2								
- リスト3									
- **強調**	中見出し								
### 小見出し	• リスト3								
要素 説明	• 強調								

要素1 説明1	小見出し								
要素2 説明2									
要素3 説明3									

	<table border="1"> <thead> <tr> <th>要素</th> <th>説明</th> </tr> </thead> <tbody> <tr> <td>要素1</td> <td>説明1</td> </tr> <tr> <td>要素2</td> <td>説明2</td> </tr> <tr> <td>要素3</td> <td>説明3</td> </tr> </tbody> </table>	要素	説明	要素1	説明1	要素2	説明2	要素3	説明3
要素	説明								
要素1	説明1								
要素2	説明2								
要素3	説明3								

Fig. 9 Markdown 記法

教師データにして学習した生成 AI とは親和性が高く、Markdown 記法でプロンプトを書くことで的確に意思を生成 AI に伝えることができる。Markdown 記法の一例を Fig. 9 に示す。Markdown 記法により箇条書きすることで、伝えたい情報を明確に分けた「一文一意」の指示が可能である。また、ロールプレイは生成 AI に役割や立場を与えて、専門性や



Fig. 10 ロールプレイの1例

応答の方向性を制限する手法である。Figure 10 は (a) X 線について単純に尋ねた場合と、(b) 「診療放射線技師」のロールを指定して回答を得た例である。前者では X 線について一般的な特性を回答している一方、後者では X 線管の構造など、より踏み込んだ回答を得ることができている。また、「制約を加える」とは期待する出力の範囲や内容を制限することで、よりユーザーに効率的な回答を得る方法である。例えば「200 字でまとめて」、「例を 5 つ出して」、「です・ます調で」といった制約を加えることで出力フォーマットの調整や、生成 AI の冗長な思考を防ぐことでメモリの節約にもなる。

3.2 Zero/One/Few-Shot プロンプト

生成 AI は膨大なデータで学習しているため、質問やタスクに対し明確な回答を得られる場合が多い。しかし、質問の前に回答のサンプルを与えると、回答の精度向上や出力の制御が可能となる。サンプルを 1 つも与えないものを Zero-Shot、1 つ与えるものを One-Shot、複数個与えるものを Few-Shot プロンプトと呼ぶ。Figure 11 はいちごの色を回答させる例である。質問が「いちご→」だけだと人間でも正しく回答することは難しいが、サンプルを与えることで、無駄なく回答できることが確認できる。

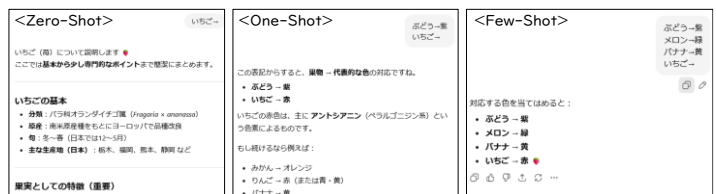


Fig. 11 Zero/One/Few-Shot プロンプトの1例

3.3 Chain-of-Thought プロンプト

Chain-of-Thought (思考の連鎖, CoT) プロンプトは、Few-Shot プロンプトに中間推論ステップを組み込むことで、より高度なタスクを達成するための手法である。例えば Zero-Shot と CoT でカニとイカの足の数を計算させた結果が Fig. 12(a) である。タラバガニはエビの仲間なので足は 8 本というところがミソであるが、Zero-Shot では数を間違えている。中間推論ステップを与えた CoT にすることで、正しく足の数を計算できている (Fig. 12(b))。

また、注目されている手法に Zero-Shot CoT がある。まず、サンプルを与えない Zero-Shot プロンプトでタスクを与え、文末に「step by step で考えてください」というような文言を加えることで、CoT を誘発する手法である [4]。同じ足の数を計算させるタスクでも、Zero-Shot CoT では正解を導き出せている (Fig. 12(c))。

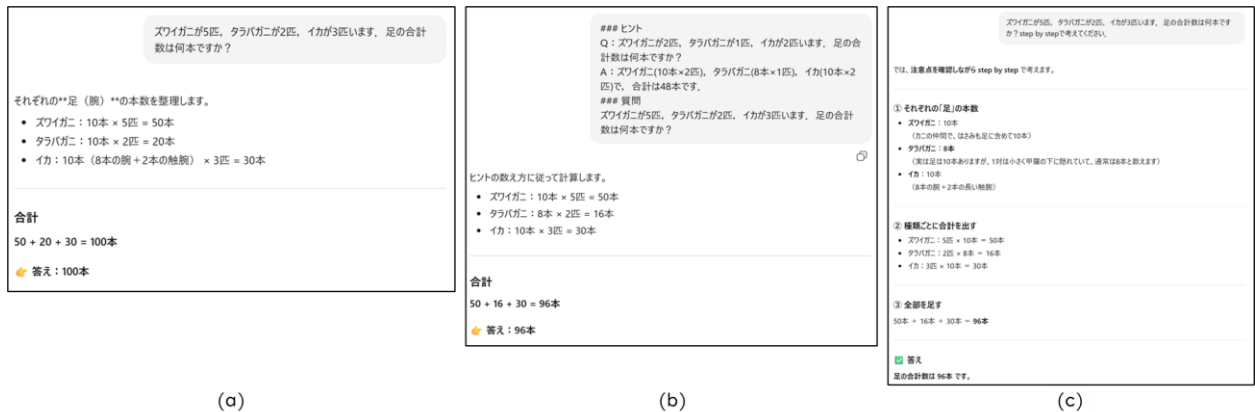


Fig. 12 同じ問題に対する (a)Zero-Shot, (b)CoT および (c)Zero-Shot CoT の回答例

3.4 Self-Consistency/Tree-of-Thought

Self-Consistency (自己整合性) は生成 AI の出力が確率的であることに着目した手法である。1つの質問に対して複数回推論させ、最も出現頻度の高い回答をこの質問の回答として採用する手法である。また、Tree-of-Thought (思考の木) は「木構造」のように思考を選択肢として分岐して並列で推論させ、最終的に推論が完結した回答を採用する手法である。これらの推論の違いについて Fig. 13 に示す。なお、Tree-of-Thought において各分岐にルールを設定して、各ロールの立場から得た回答を統合して最終的な回答として採用する手法もある。例えば「経営者・営業担当・技術担当」のロールで推論させ、三者の意見をすり合わせた回答を出力させることで、より多角的な意見を得られる。この手法を MAGI システムと呼ぶことがある。名称はアニメ「新世紀エヴァンゲリオン」に出てくるスーパーコンピュータから取られている。

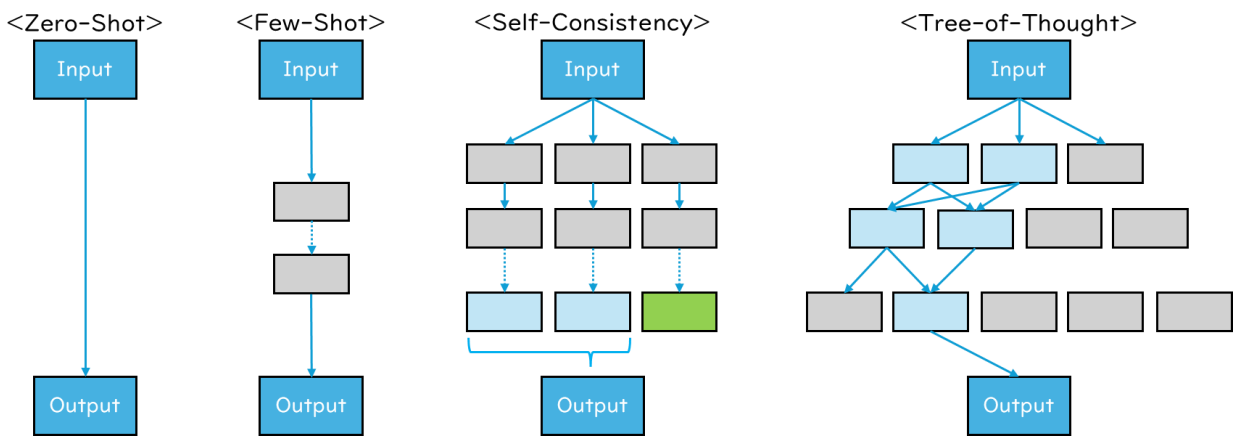


Fig. 13 各推論プロセスの比較

3.5 その他の手法

Zero-Shot CoT では「step by step」という言葉が推論精度を向上させたが、他にも推論を刺激するフレーズがあり、一部ではマジカルワードと呼ばれている。また、「深呼吸して」といった精神論のようなフレーズも効果があると言われており、Microsoftなどはエモーションプロンプトと呼んでいる。これら

の一部を Table 3 に示す。

さらに最近の研究では、質問を単純に 2 回続けて入力するだけでも回答の精度が上がると報告されている [5]。この手法を Prompt Repetition (プロンプト反復) と呼ぶ。これにより、生成 AI の読み取り方に由来する取りこぼしを防ぐ効果があると報告されている。このように、プロンプトエンジニアリングの研究は日進月歩で、より効果的なフレーズがこれからも発見されると考えられる。難しい手法ではないため、生成 AI を用いた開発に 1 つでも多く取り入れることを推奨する。

Table 3 推論を刺激するフレーズ

マジカルフレーズ	エモーションプロンプト
<ul style="list-style-type: none"> 多角的な視点で考えてください 水平思考で考えてください デザイン思考で考えてください 問題解決思考で考えてください 仮説思考で考えてください 批判的思考で考えてください アナロジー思考で考えてください アダプティブ思考で考えてください 構造化思考で考えてください ホリスティック思考で考えてください 反証的思考で考えてください 	<ul style="list-style-type: none"> 深呼吸して 自分を信じて限界を超えてください 成長の機会だと捉えて挑戦してください これは私のキャリアにとって非常に重要です 努力は報われます

3.6 トークン/クレジットと言語について

生成 AI の利用において重要な概念として「トークン」と「クレジット」がある。トークンとは、AI が処理する文字や単語を細分化した単位であり、入力文と出力文の双方で消費される。文章が長いほど使用トークン数は増加する。一方、クレジットはトークン使用量に応じて消費される利用枠や料金単位を指す。これらを理解することで、コストを意識した効率的な生成 AI 活用が可能となる。また、トークン数は日本語より英語の方が少なくカウントされる。

市場にあるほとんどの生成 AI は多言語対応しているため、多くのユーザーは母国語でプロンプトを作成している。しかし、生成 AI の学習データの多くが英語であることから、英語で質問した方が推論の精度が高いことが知られている (Fig. 14) [6]。したがって、トークン数と精度を考えるとすれば、英語でプロンプトを作成する方がクレジットの消費を考えても有益である。英語のプロンプト作成は困難ではなく、日本語のプロンプトを生成 AI に英訳してもらうだけである。

まとめ

本記事では人気のあるプログラミング言語 Python の紹介や導入方法、バンプコーディングおよびプロンプトエンジニアリングについて概説した。適切なプロンプトを作成してバンプコーディングするのであれば、Python の学習コストは不要に思える。実際、非エンジニアが Web アプリを開発するなどの記事は多く見られる。しかし、前述したように「AI が生成したコードを管理する」ためにコードを正しく理解し、意図通りに修正・発展・保守するためには、言語仕様や基本構造の理解は不可欠である。Python の文法を知っていれば曖昧な生成結果にも的確な指示を与えることが可能である。バンプコーディングは学習の代替ではなく、Python の理解を前提とした自身の能力を拡張する手段と考えるべきである。また、生成 AI によるコード出力は「無から有 (0→1)」のタスクには向くが、既存のシステムの改善などには現状では力不足であり、この点からも Python を学ぶ利点はあると考えている。本記事の内容が会員のこれからの研究開発の一助となれば幸いである。

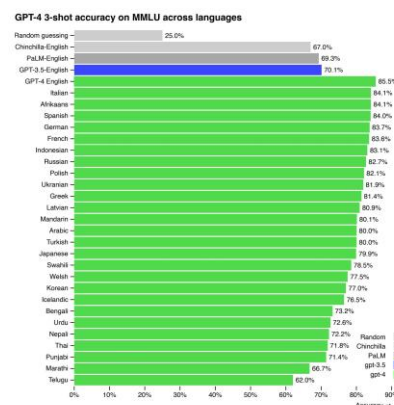


Fig. 14 言語による推論精度 (出典: 参考文献 6)

参考文献

- [1] PYPL Popularity of Programming Language. <http://pypl.github.io/PYPL.html>. Accessed Jan 18th, 2026.
- [2] PubMed A General Introduction to the E-utilities. https://www.ncbi.nlm.nih.gov/books/NBK25497/#chapter2.Usage_Guidelines_and_Requiremen. Accessed Jan 19th, 2026.
- [3] Prompt Engineering Guide. <https://www.promptingguide.ai/>. Accessed Jan 19th, 2026.
- [4] Zhou Y, Muresanu AI, Han Z, et al. Large Language Models Are Human-Level Prompt Engineering. Conference paper at ICLR2023. <https://arxiv.org/pdf/2211.01910>.
- [5] Leviathan Y, Kalman M, Matias Y. Prompt Repetition Improves Non-Reasoning LLMs. arXiv. 2025. <https://arxiv.org/abs/2512.14982>.
- [6] OpenAI. GPT-4 Technical Report. arXiv. 2024. <https://arxiv.org/pdf/2303.08774>.

RSNA2025 の参加記

金沢大学医薬保健研究域保健学系 川嶋 広貴

1. はじめに

本稿は、2025 年 11 月 30 日から 12 月 4 日に米国イリノイ州のシカゴで開催された RSNA 111th Scientific Assembly and Annual Meeting の参加記である（以下、この学術集会のことを RSNA と呼ぶ）。RSNA は、放射線医学に関する世界最大規模の学術集会であり、参加者は、放射線科医師を中心に医学物理士や医用画像に関わる専門家が一堂に会する。会場は、シカゴの中心部（ダウンタウンやリバーノース）から、南へ約 4 km に位置する McCormick Place である。筆者は、今回が 7 回目の現地参加であり、同大学の市川教授とその研究室の一般院生 2 名、社会人院生 1 名の 5 名で行動を共にした（図 1）。自分自身の経験を伝えることで、これから RSNA を目指す読者の一助となることを期待し、筆者の私見も交えつつ概要を記す。



図 1 学会会場での写真。毎年このようなオブジェが設置され、写真スポットとなっている。

2. シカゴについて

シカゴは全米第 3 位の人口を擁する大都市である。羽田からシカゴのオヘア国際空港まで、約 11 時間 20 分のフライトを終えると、15 時間の時差が生じており、毎年のように時差ボケによる体調管理に悩まされる。シカゴは“windy city”とも呼ばれ、気温がマイナスの中、風に切られるような思いをしながらダウンタウンを歩くことになるので、しっかりとした防寒対策が必要である。今年は到着時に珍しく雪が降っており、フード付きのコートや耳当てが重宝した。

学会とはいえ、せっかくの渡航であれば、その土地の文化を体感したいものである。シカゴの近代的な摩天楼は魅力的で、個人的には、ジョンハンコックセンターからの夜景をお勧めしたい。摩天楼の夜景は、圧巻の光景で、“またこの夜景を見られるように頑張ろう！”と研究へのモチベーションとなっている。また、食文化も豊かで、極厚のシカゴピザ（stuffed deep-dish pizza）やクラフトビール、ホットドッグ、イタリアンビーフサンドなどは学会時の楽しみの 1 つである。

2. 参加に至るまで

RSNA の演題募集期間は、2月に始まり、ゴールデンウィークごろが締め切りとなる。抄録は、応募する形式によって異なり、筆者は、“Scientific presentation”と“Education exhibit”の経験がある。Scientific presentationは、学術的な研究内容の発表（仮説に基づく研究）であり、採択された場合には口述もしくは電子ポスターの発表が割り当てられる。数年前までは、発表形式を希望できたが、ここ最近では希望する形式の選択肢はなく、採択通知時に形式が分かるようになっている。Education exhibitは、その名のとおり教育的な内容を含んだものであり、電子ポスターでの発表もしくは展示を行う。

RSNA における演題の採択率は決して高くないため、洗礼された抄録づくりが重要となる。筆者の考えるポイントは次の2つである。①適切な英語で、定量値（計測結果や統計）を示し、理にかなった説明がなされていること、②適切な投稿分野を選択すること。我々のような放射線技術学の研究では、技術系と医学系の研究が混在しているが、研究内容の訴求点が技術的な部分なのか、もしくは、医学的な部分なのかで、演題登録の入り口である投稿分野の選択や抄録の書き方が大きく変わってくるため慎重に選択されるべきである。

採択通知は、7月末頃にメールで届く。ちなみに今年は、私が関わった研究から5演題を応募し、最終的には3演題がポスター発表（1演題が筆頭、2演題は大学院生が筆頭）、1演題が共同研究グループによる口述発表、1演題が不採択という結果であった。RSNAにこれだけの演題が採択されたことは、素直に嬉しく思った。

採択通知を受け取ってからは、旅程を立て、飛行機、ホテル、米国の入国に必要な手続き、参加登録を行うこととなる。円安の影響で、飛行機やホテルの価格が高騰していることは言うまでもなかった。また、以前は発表者の参加登録料が無料であったが、数年前から既定額の半額となった。早期登録による割引期間において、診療放射線技師かつ発表者であれば既定の475ドルの半額といった具合である。

3. RSNA2025 での関心事項

3.1 展示会場

RSNA の機器展示は、2つの大きな建物（North と South building）で行われており、広い会場には多くの機器が並び、その規模の大きさが印象に残った。RSNA の注目度が高いことから、展示会場を回ってみると、各社の最新の機器や技術に関する情報収集を行うことができる。

今年の展示の中で、最も印象的だったのは、GE HealthCare 社から発表された“Photonova Spectra”という新たなフォトンカウンティングCTであった（RSNAの時点では、FDA 510K pending）。GE社のブースで行われていたunveil event（除幕式）では、多くの聴衆が集まっており、その期待の高さがうかがえた。今回新たに発表されたフォトンカウンティングCTは、検出器に比較的低原子番号のsiliconを用いており、主流であるCdTe（テルル化カドミウム）やCZT（テルル化カドミウム亜鉛）とは異なる特徴を有する。Silicon検出器は、X線阻止能が低いため、十分な検出効率を得るために厚さのある検出器構造（deep silicon検出器と呼ばれる所以）となっており、優れたエネルギー分解能やfluorescence emissionの発生確率が低いことから、スペクトルの利用において有利となる可能性がある。今後、silicon検出器ならではの利点を活かした臨床的有用性が数多く報告されることを期待したい。

また、展示会場の一角には、“AI SHOW CASE”と呼ばれる区画が設けられ、100社を超える関連企業が集まっていた。筆者は専門外であるため、1-2社の説明を聞くにとどまったが、最新のAI製品に関心のある本誌の読者にとっては新しい発見があるかもしれない。

3.2 研究発表

今回のRSNAでは、フォトンカウンティングCTに関する情報収集を目的としていたため、それに関連する演題を多く聴講した。CT画像からDQEを測定する手法を提示し、フォトンカウンティングCTと従来のCTの比較を行うことで、空間周波数ごとのPCCTの性能的利点を示した研究は、非常に理にかなったものだと感じた。また、人体構造を忠実に再現したファントムやマルチエナジーCT解析に関する演題は非常に興味深く、聴講する中で新しい研究のアイデアも次々と湧いてきた。また、日本の研究者の方が口述発表されている姿もたびたび目撃し、刺激を受けるとともに、いつか自分もその舞台に立ちたいと思った。

筆者の研究グループの電子ポスター発表では、臨床用のフォトンカウンティングCTやディーブラーニングCT画像再構成の性能についてファントム実験を行った結果を報告した。電子ポスターの発表は、指定された日時に30分の時間が割り当てられ、会場に設置されている大型ディスプレイに自分の演題を表示し、質問者に対して質疑をするという形式である。口述発表に比べて、比較的フランクに意見を交換できるため、個人的には好ましいと思っている。ここで、よくある流れの一例を記す：自分のポスターの前に質問者が立ち寄り→タイトルを見て、興味があれば話しかけられる→簡単に挨拶をする→研究内容を簡単に説明する→質疑に移る。これに対応するため、学生には2-3分で研究内容を簡単に説明できるように準備することを指導している。短いプレゼンが終わった後は、勇気を出して、「ご質問はありますか？」や「この解釈は妥当でしょうか？」などと一歩踏み込んで問いかけると、関心のある方からは具体的な質問が寄せられ、さらに有益な議論につながる。筆者の発表時間は幸いにも質問者が途切れず、非常に有意義な時間を過ごすことができた。若手研究者が多かったが、中には著名な研究者もおり、厳しい質問もあったが、自分の意見をしっかりと伝えることができた点はポジティブにとらえている。

5. おわりに

RSNAでは、放射線医学・技術分野を牽引する研究者の発表や質疑を間近で体験でき、改めてその魅力を実感した。自分たちの研究成果を発表し、好意的な反応や建設的なフィードバックを得られたことは大きな励みである。また、個人的には、共同研究で滞在していたウィスコンシン大学のメンバーとの再会も非常に印象深いものであった。このように、国際学会の参加には多様で刺激的な魅力が詰まっている。学術的・経済的ハードルは決して低くは無いが、そこにしか得られない経験が確かにある。本稿が一人でも多くの読者のモチベーション向上につながり、RSNAでの発表を目指すきっかけになれば幸いである。

6. 謝辞

RSNAでの研究発表の成果は、富山大学附属病院の河原康宏先生、伊藤貞則先生、野口京先生との共同研究である。研究へのサポートに心から感謝申し上げます。

画像部会「研究情報サイト」のご案内

岐阜大学教育学部技術教育講座 福岡 大輔

1. はじめに

画像部会の部会ホームページでは、これまでに本誌「画像通信」において紹介されたトピックスや、各種医用画像データベースに関する情報、プログラミングに関する情報などを集約し情報提供を行う「研究情報サイト (<http://imgcom.jsrt.or.jp/research/>)」を2017年11月に開設いたしました(図1)。

研究情報サイトは、画像研究に携わる研究者や、研究を始めてみようとする初学者の、情報提供や情報交換の場となることをめざしています。

2. 研究情報サイトの概要

研究情報サイトでは、図2のように画像研究に役立つ情報として、本稿の執筆時点では、以下の3つのテーマについて情報提供を行っています。

(1) 医用画像データベースに関する情報

近年話題の深層学習をはじめとする画像研究においては、多くの画像から共通する特徴を抽出することによって成り立っており、画像データベースの必要性が高まっている。しかしながら、CAD(Computer-aided Diagnosis)システムの開発や性能評価の研究においては、研究に用いる医用画像データの収集や、付随する所見データの入手は、倫理上や個人情報の観点からも近年では難しくなっている。そこで、研究情報サイトでは、国内外の大学や政府機関、各種学会など公的な機関が公開するデータベースを紹介している。

例えば、胸部X線画像においては、日本放射線技術学会標準デジタル画像データベースをはじめ、NIH



図1 画像部会 HP の「研究情報サイト」



図2 「研究情報サイト」の掲載情報

Clinical Center の Ronald M. Summers 氏らの ChestX-ray8 データベースなどを紹介している。ChestX-ray8 データベースにおいては、症例数が患者数 30,000 で画像数も 100,000 画像と非常に豊富であるため、深層学習を用いた画像研究に有用なデータベースとなっている。また、画像通信 40(1)「Radiomics に関係した文献とデータベース等の紹介」において紹介された Radiomics 研究用のデータベースとして、The Cancer Imaging Archive (TCIA) の医用画像と遺伝子の情報セット (The Cancer Genome Atlas (TCGA)) を研究情報サイト上で紹介している。

(2) プログラミングに関する情報

画像処理プログラミングに関する情報として、画像通信 35(2)に紹介された GUI (Graphical User Interface) による画像処理ソフトウェア開発の基本となる「C#プログラミングによる画像表示ソフトの作成 (超初心者編)」のソースコード一式や、画像通信 35(2)において紹介されている OpenCV (Open source Computer Vision library)、画像通信 32(1)「CAD のための統計解析の資料」に紹介されている統計解析ソフト R (The R Project for Statistical Computing) を掲載している。また、医用画像を取り扱う上で必要となる DICOM フォーマットに関する情報として、ソフトウェア開発に有益な DCMTK (Dicom ToolKit) に関する情報と、日本画像医療システム工業会 (JIRA) が公開している DICOM 規格書 (日本語版) へのリンクを紹介している。

(3) 論文作成 (文献検索など) に関する情報

画像通信 31(2)「論文作成に役立つサイトやソフト、書籍の紹介」において紹介された米国立医学図書館内の NCBJ がインターネット上で提供している生命科学分野の文献検索システム PubMed や、Google Scholar について掲載している。

3. 今後の活用と情報提供のお願い

「研究情報サイト」は、本誌「画像通信」の誌面に掲載される記事を、インターネット上のデジタルコンテンツとして補足する機能を担い、誌面上では提供できないソフトウェアの配布や、ソースコードなどの提供など、画像研究に役立つ情報を幅広く配信してゆきたいと考えています。また、医用画像データベースに関する情報など、医用画像に関する研究者の情報共有・提供の場として今後活用してゆきたいと考えています。

現在、サイト上に掲載されている掲載内容のほかにも有益な情報がありましたら、「研究情報サイト」のページ下部にあるコメント欄に情報をお寄せください。

2025 年度 画像部会事業報告

1) 第 96 回画像部会の開催(第 81 回総会学術大会時)

開催日:2025 年 4 月 11 日(金)15 時 30 分~18 時 00 分 会場:パシフィコ横浜(横浜市)F201+202 室

(1) 教育講演 司会 鈴鹿医療科学大学 東出 了
「DR システムの非線形画像処理を評価するためのキーポイント」

昭和大学病院 後藤 光範

(2) 第 96 回画像部会

「DR システムにおける非線形画像処理の画質評価について考える」

司会 岐阜医療科学大学 篠原 範充・小倉記念病院 佐保 辰典

① DR における最新画像処理とデータ

聖マリアンナ医科大学病院 田沼 隆夫

② DR システムにおける非線形画像処理評価の問題点

東海大学医学部附属八王子病院 由地 良太郎

③ Task transfer function (TTF)評価における解析方法と課題 鈴鹿医療科学大学 東出 了

④ 胸部 X 線画像における非線形画像処理の評価 東北大学病院 小野寺 崇

2) 第 97 回画像部会の開催(第 53 回秋季学術大会時)

開催日:2025 年 10 月 18 日(土)14 時 00 分~17 時 00 分 会場:札幌コンベンションセンター(札幌市)第 3 会場 大ホール C

(1) 教育講演 司会 鈴鹿医療科学大学 東出 了
「放射線技師の業務を支援する AI 研究」 名城大学 寺本 篤司

(2) 第 97 回画像部会

「放射線診療業務を支える AI 研究」

司会 神戸常盤大学 山崎 明日美・小倉記念病院 佐保 辰典

① 業務効率の向上を目指した放射線技術支援システムの研究

日本文理大学 衛藤 路弘

② 胸部 X 線撮影の吸気状態に着目した再撮影自動判定

信州大学医学部附属病院 松原 尚輝

③ 頭頸部 3D-CTA 撮影補助のための画像変換 AI 研究

市立四日市病院 倉谷 洋佑

④ 放射線治療×AI の研究

九州大学医学部附属病院 廣瀬 貴章

3) DR(デジタルラジオグラフィ)セミナーの開催

第 26 回 DR セミナー (教育委員会, 関東支部共催)

開催日:2025 年 8 月 9 日(土) 9:00 ~ 18:00, 10 日(日) 9:00 ~ 13:00

会 場:中央医療技術専門学校(東京都葛飾区)

受講生人数:22 名

4) 医用画像処理プログラミングセミナーの開催

第 46 回医用画像処理プログラミングセミナー(教育委員会共催)

2025 年 9 月 7 日(日) 10:00 ~ 15:00

会 場:Cisco Webex によるオンライン・小倉記念病院(本部)

受講生人数:49 名

5) ROC セミナーの開催

第 15 回 ROC セミナー(教育委員会共催)

開催日:2025 年 9 月 9, 23, 30 日, 10 月 7, 14(水)の5日間

会 場:Cisco Webex によるオンライン(50 分単位で 5 回開催)

受講生人数:62 名

6) 臨床画像評価セミナーの開催

第 12 回臨床画像評価セミナー(教育委員会, 中部支部共催)

日時:2025 年 12 月 21 日(日)

会場:名古屋市立大学病院(愛知県名古屋市)

受講生人数:24 名

7) 部会誌(画像通信)の発行

Vol.48 No.1(通巻 94 号)(2025 年 4 月), Vol.48 No.2(通巻 95 号)(2025 年 10 月)

9) 部会委員会の開催

第1回委員会	2025年	4月11日	横浜
第2回委員会	2025年	5月26日	WEB
第3回委員会	2025年	10月8日	WEB
第4回委員会	2025年	11月6日	WEB
第5回委員会	2025年	11月30日	京都

2026 年度 画像部会事業計画

1) 第 98 回画像部会の開催(第 82 回総会学術大会時)

開催日:2026 年 4 月 17 日(金) 会場:パシフィコ横浜(F201+202 室)

(1) 教育講演

司会 鈴鹿医療科学大学 東出 了

「非線形再構成時代の CT 画像評価:普遍性の喪失にどう向き合うか?」

茨城県立医療大学大学院 瓜倉 厚志

(2) 画像部会

「パラダイムシフト?! 臨床画像評価の未来」

司会:神戸常盤大学 山崎 明日美・小倉記念病院 佐保 辰典

① 「臨床画像評価のための物理評価」

東北大学病院 小野寺 崇

② 「ゼロから学ぶ ROC 解析～判断基準とカットオフ値で理解する「観察者実験」の仕組み～」

西南学院大学 馬場 祥吾

③ 「臨床画像の評価および読影における課題への取り組み」

福岡大学筑紫病院 田畑 成章

④ 「臨床画像評価の未来 ～AI を使った画質評価法～」

滋賀医科大学医学部附属病院 大橋 昂平

2) 核医学部会・画像部会 合同企画シンポジウムの開催(第 54 回秋季学術大会時)

開催日:2026 年 10 月 24 日(土) 午前(予定) 会場:G メッセ群馬(高崎市)

(1) 教育講演

司会 北海道大学病院 孫田 恵一・鈴鹿医療科学大学 東出 了

「画像の成り立ちと核医学・画像情報の基礎(仮)」

国際医療福祉大学 木村 信篤

(2) シンポジウム

「AI による画像技術の変革と臨床応用」(予定)

司会 帝京平成大学 平原 大助・島根大学医学部附属病院 矢田 伸広

① 「PET 臨床機における AI 技術の応用:画質改善とセグメンテーション(仮)」

福島県立医科大学 山尾 天翔

② 「医用画像(主に CT)への画像生成モデル(Diffusion model, GAN など)の応用(仮)」

新潟大学 市川 翔太

③ 「臨床現場の課題を起点とした AI 実践:プログラミング未経験からの画像処理コード作成(仮)」

横浜栄共済病院 江上 桂

3) 画像部会・核医学部会 合同企画シンポジウムの開催(第 54 回秋季学術大会時)

開催日:2026 年 10 月 24 日(土) 午後(予定) 会場:G メッセ群馬(高崎市)

- (1) 教育講演 司会 鈴鹿医療科学大学 東出 了・北海道大学病院 孫田 恵一
「大規模言語モデル(LLM)の最新動向とアカデミックな視点(仮)」 北海道科学大学 谷川原 綾子
- (2) シンポジウム
「生成 AI の多角的活用の実践と未来像」(予定)
司会 岡山大学 福井 亮平, 弘前大学 奥田 光一
- ① 「生成 AI の技術的背景の理解(仮)」
滋賀医科大学 福井 悠介
- ② 「生成 AI×プログラミングによる業務・研究エンジニアリング(仮)」
小倉記念病院 佐保 辰典
- ③ 「臨床業務への生成 AI 展開:被ばく相談支援とローカル LLM の活用(仮)」
市立角館総合病院 千葉 大志
- ④ 「研究・論文作成における生成 AI:文章添削から統計処理・グラフ作成まで(仮)」
熊本大学 椎葉 拓郎
- ⑤ 「学生教育・臨床業務効率化への応用(仮)」
藤田医科大学 辻本 正和

4) 医用画像処理プログラミングセミナーの開催(教育委員会共催)

第 47 回医用画像処理プログラミングセミナー(オンライン)

開催日時:2026 年 7 月 26 日(日)

会場:オンライン実施 セミナー本部:小倉記念病院

5) 臨床画像評価セミナーの開催(教育委員会, 近畿支部共催)

第 13 回臨床画像評価セミナー

開催日時:2026 年未定

会場:未定

6) 部会誌(画像通信)の発行

Vol.49 No.1(通巻 96 号)(2026 年 4 月), Vol.49 No.2(通巻 97 号)(2026 年 10 月)

7) 部会委員会の開催 年 5 回(横浜, WEB, WEB, WEB, WEB の開催予定)

画像部会入会のご案内

医療には、X線画像、CT、MRI、US、核医学画像、そして放射線治療用画像など、様々な画像が利用されています。画像部会は、これら全ての画像を対象とし、そのイメージング技術、画像評価、画像処理・解析、コンピュータ支援診断（CAD）に関する新しい知識や技術の習得を目的とした活動を行っています。よって、診断、治療、核医学などの専門領域を問わず、多くの学会員の皆さまに入会し、参加して頂くことができます。

春と秋に開催される学術大会においては、話題性が高く学術的価値の高い教育講演とシンポジウムを開催しています。このシンポジウムでは、企画されたテーマの第一線の研究者らに問題提起や話題提供をして頂いた上で、会員の皆さんを交えた討論を行い、新しい知識や技術の有用性や問題点を共有しています。その他、地方部会の協力を得ながら、医用画像処理プログラミングセミナー、DRセミナー、ROCセミナーおよび臨床画像評価セミナーを開催して、必要な基本知識と技術の普及を図っています。こういった活動を通して、会員の皆さんが画像研究の新しい風を肌で感じたり、学術レベルの向上や技術の臨床への還元をして頂けるようになります。また、活動案内や情報は、学術大会前にお届けしている画像部会雑誌「画像通信」（学術雑誌 ISSN コード付）に掲載しています。学術雑誌である「画像通信」には、教育講演やシンポジウムの内容、注目されている技術の紹介、専門家による文献紹介、日本各地の研究室や研究会の紹介、国際会議出席者の体験記など、参考になる記事が多数掲載されており、画像に興味を持つ会員にとって非常に魅力的な専門雑誌となっています。

画像部会は医療で広く利用されている画像に関する理解を深め、医療の進歩に寄与したいと考えている方に入会して頂き、共に学んでいきたいと思っています。また、すでに会員の方も、是非画像にこだわりを持つ周辺の方々に声をかけて入会を促して頂けるようお願いいたします。

【入会資格】 日本放射線技術学会の会員であること。

【入会方法】 Web 上 (<https://www.jsrt.or.jp/data/procedure/bunka-01/>) から、お申し込み下さい。

【年会費と会員特典について】

1. 専門部会の会員登録システムと年会費の変更について

2022年より専門部会の入会については、正会員・シニア会員は1つ目の年会費は無料ですが、2つ目以降は各1,000円が別途必要です。学生会員は会費免除で全ての部会に自動的に登録されます。

2. 専門部会誌の電子化について

現在、専門部会員の皆様に冊子体で届けている専門部会誌を、2015年3月発刊分より全面的に電子化します。電子化により印刷製本費や郵送費が軽減できますので、専門部会活動の充実に充てたいと考えています。なお、電子版の閲覧方法については、追ってお知らせします。

編集後記

私が最初に「画像通信」の執筆をしたのが、2007年10月のvol. 30 No. 2（通巻59号）でした。編集後記の執筆にあたり、自分の原稿が初めて「画像通信」に掲載され、冊子を手にして喜んだことを思い出しました。当時はその後に画像部会の委員や部会長を務めることになるとは想像もしていませんでした。

時代の変化に伴い、紙の冊子から電子版となりましたが「画像通信」は今回で通巻96号となり、画像部会の歴史を表す重要な象徴であるといえます。一方、JRC2026では第98回画像部会が開催され、JRC2027は第100回画像部会と大きな節目を迎えます。これまで画像部会にご尽力された先生方への感謝と共に、若い研究者へ画像研究の面白さが伝わる企画や開催を意識しています。AIなどコンピュータの進化と共に「画像」は成長し、評価や解析は常に進化を求められています。老舗の専門部会ですが、「画像」の成長に合わせて「画像部会」も成長していると感じています。是非とも若い方と一緒に画像部会に足を運んでください。様々な年代の画像研究に関わる皆さんと一緒に議論し、次世代の研究者にとって新たなモチベーションを得る場となることを期待しています。

堅苦しい後記となりましたが、ミラノ・コルティナ冬季五輪の白銀の世界で戦う選手を応援しつつ、辛い花粉症の症状にて春の訪れを感じています（涙）。

(RH 記)

画像通信 Vol.49 No.1(通巻'96)

発行日 2026年4月1日

発行所 公益社団法人 日本放射線技術学会

〒600-8107 京都市下京区五条通新町東入東鋸屋町167

TEL075-354-8989 FAX075-352-2556 e-mail: office@jsrt.or.jp

発行者 公益社団法人 日本放射線技術学会 画像部会 部会長 東出 了

画像通信 通卷九六号

VOL. 49 No.1 SPRING 2026