

乳腺密度計測への人工知能技術の
応用に関する研究

**A study on application of deep learning technology
in volumetric breast density measurement**

令和 5 年 2 月
February 2023

山室 美佳
Mika Yamamuro

目次

第 1 章 緒論	1
1.1 はじめに	1
1.1.1 乳癌とマンモグラフィ	1
1.1.2 乳癌検診	3
1.2 本研究の歴史的背景	4
1.3 本論文の目的と構成	6
参考文献	9
第 2 章 ROC 解析を用いた乳腺密度と病変検出感度の関係の定量化	12
2.1 緒言	12
2.2 乳腺密度可変型乳房等価ファントムの開発	13
2.2.1 ファントムの使用素材	13
2.2.2 ファントムの作成	15
2.3 病変検出感度の乳腺密度依存性	18
2.3.1 ROC 解析のための画像作成および観察方法	18
2.3.2 ROC 解析結果	20
2.3.3 ファントムのノイズパワースペクトル	22
2.4 考察	24
2.5 結言	26
参考文献	26
第 3 章 乳腺密度計測への人工知能技術の応用	30
3.1 緒言	30
3.2 U-NET を用いた乳腺領域の自動抽出	33
3.2.1 乳腺領域の定義	33
3.2.2 データ収集	33
3.2.3 Ground Truth の作成	34
3.2.4 U-net モデルの構築	34

3.2.5	Dice 係数による U-net モデルの評価	36
3.3	乳腺密度の算出	37
3.4	平均乳腺線量の算出	40
3.5	乳腺密度および平均乳腺線量による U-NET モデルの評価	40
3.5.1	評価方法	40
3.5.2	評価結果	41
3.5.3	乳腺領域の抽出例	43
3.6	考察	45
3.7	結言	48
	参考文献	49
第 4 章	マンモグラフィの X 線露光条件を用いた乳腺密度の推定	55
4.1	緒言	55
4.2	身体特性および生育歴情報を用いた乳腺密度予測（先行研究）	56
4.3	乳腺密度推定モデルの作成方法	59
4.3.1	データ収集	59
4.3.2	乳腺領域の画素値の決定法	59
4.3.3	新しい脂肪領域画素値の決定法	60
4.3.4	乳腺密度の算出	60
4.3.5	モデルの作成と評価方法	61
4.4	乳腺密度推定モデルの精度評価	61
4.4.1	Improved threshold method 用いた脂肪領域画素値の精度検証	61
4.4.2	乳腺密度推定モデルの決定	62
4.4.3	モデルの精度検証	63
4.5	考察	66
4.6	結言	68
	参考文献	69
第 5 章	結論	72
	謝辞	76

研究業績.....	77
付録	81

第 1 章 緒論

1.1 はじめに

1.1.1 乳癌とマンモグラフィ

W. C. Röntgen が 1895 年に X 線を発見してから 120 余年が経過し、その間 X 線は医学や工学等の分野で広く利用され、各分野の発展に貢献してきた^{1, 2)}。医学領域では、画像診断やがんの治療に利用され成果を挙げている。世界で最初に撮影された X 線写真は、Röntgen 自身が夫人の手を撮影したもので、当初はこのように被写体が薄く、かつ X 線減弱の大きい骨の撮影を中心に、骨折や金属片等の残留異物の確認に用いられていた。X 線に関する研究が進むにつれ、しだいに人体の広い範囲に対して X 線写真が適用されるようになってるとともにアナログ画像からデジタル画像へ移行し、現在も日進月歩で新しい技術が開発されている。

本論文は、X 線を用いて撮影した乳房画像（以下、マンモグラムと呼ぶ）における乳癌を中心とした病変の検出特性に焦点を当てたものである。乳癌とは乳腺を形成する乳管や小葉から発生した悪性腫瘍を指し、そのほとんどは乳管由来の乳管癌であるが、まれに小葉から発生することもある。国立がん研究センターがん統計³⁾によると、乳癌は女性のがん罹患率第 1 位、死亡率第 4 位であり、このうち罹患率第 1 位は 1995 年以降現在まで継続されている。一方、2009 年から 2011 年にかけて調査された乳癌の 5 年相対生存率ⁱ は 92.3% と極めて高い。ただし、この数値は病期に大きく依存しており、0-1 期であれば 90% 以上であるが 4 期の場合は約 10% まで低下する。このことから、早期発見が乳癌から女性を救うための極めて重要な因子であると言える。本研究は乳房撮影に特化した X 線撮影システムであるマンモグラフィの臨床応用領域拡張の可能性を追求し有益性の向上を目指すものである。

まず、マンモグラフィの歴史を簡単に振り返る。乳癌に対して最初に X 線写

ⁱ がんと診断された場合に、治療でどのくらい生命を救えるかを示す指標。がんと診断された人のうち 5 年後に生存している人の割合が、日本人全体で 5 年後に生存している人の割合に比べてどのくらい低いかを表す。100% に近いほど治療で生命を救えるがん、0% に近いほど治療で生命を救い難いがんであることを表している。

真を適用したのは A. Salomon で 1913 年のことである。Salomon は乳癌手術患者約 3000 例について切除標本を作製し X 線撮影を行った。そして肉眼所見や病理所見と X 線写真を比較し、X 線写真を使って癌の範囲を説明できること、さらに浸潤性と限局性に分類できることを示した⁴⁾。図 1.1(a)と図 1.1(b)に 1930 年頃の乳房 X 写真撮影の様子と正常乳房の X 線写真の一例を示す⁵⁾。1951 年、R. Leborgne はマンモグラムの所見において微小石灰化が重要であることに着目した。この時、撮影法として頭尾方向と側面方向の 2 方向撮影、さらには照射野を必要最小限に絞るための照射筒を用いるなど、現在へも通じる撮影法が用いられた。Leborgne はマンモグラムに描出された多くのタイプの微小石灰化を腫瘍の鑑別に適用し、十分な経験を積めば微小石灰化の X 線画像を使って良性腫瘍と悪性腫瘍の鑑別診断を比較的容易に行うことができると結論付け、「家族歴にがんがある全女性に対して予防的に X 線写真を撮影することで、乳癌に対する新たなステージが展開されるであろう」と乳癌検診の必要性についても言及している⁶⁾。

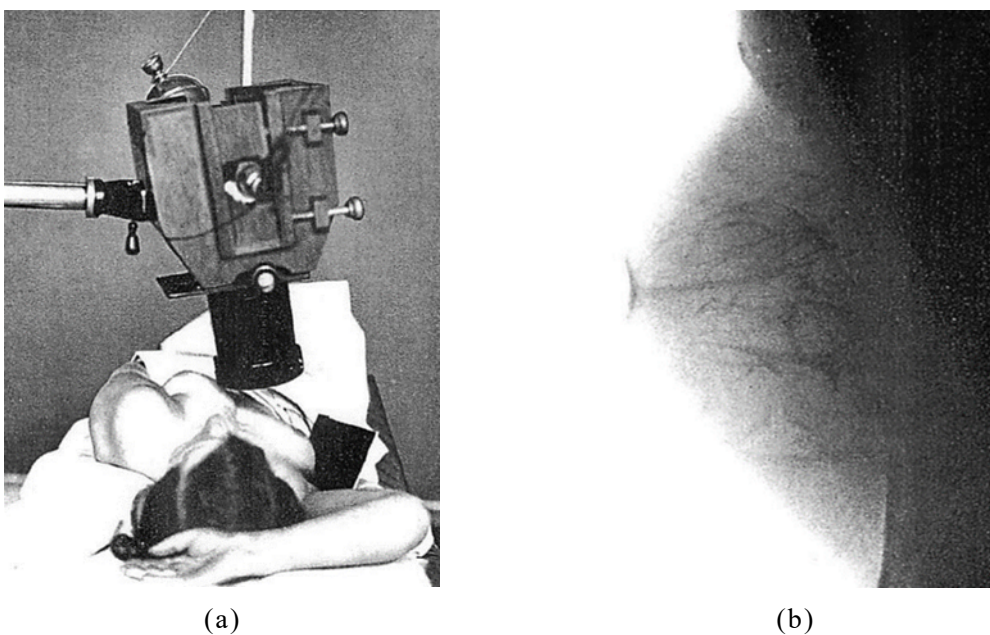


図 1.1 1930 年頃のマンモグラフィ。(a)Vogel らによる乳房撮影の様子、
(b)当時撮影された乳房 X 線写真の正常例 (文献 5 から引用)。

乳癌診断に対するマンモグラフィの有用性を臨床例で示し、近代のマンモグラフィの基礎を築いたのは、1960 年に R. L. Egan が発表した「Experience with

「Mammography in a Tumor Institution」と題する論文である。原文によると1956年から1959年にM. D. Anderson病院で撮影された1000例のマンモグラフィを検討し、術前マンモグラフィで乳癌と診断できた症例が238例であったのに対し偽陰性は2例のみであり、陽性的中率は99.2%であった。撮影方向は内外斜位方向・頭尾方向・腋窩部撮影の3方向である。フィルムはコダック社の工業用片面乳剤フィルムを用い増感紙を使っていないため、基本的な撮影条件は26 kV~28 kV、300 mA、6秒と非常に多くのX線量を要しているものの、X線画像の有効性およびその必要性を示すのに十分な成果を示したと言える⁷⁾。この時Eganは論文の結論で、「Experience has shown mammography to be a simple method of demonstration of a breast lesion with a high accuracy in prediction of its type」と記し、それまで無かった「マンモグラフィ」という言葉を生み出した。

1.1.2 乳癌検診

ここまで述べてきたように、乳癌の診断に対するマンモグラフィの臨床応用は着実に成果を挙げてきたと言える。しかし、米国人口動態統計⁸⁾によると、25歳以上の女性を対象としたがん全体の年齢調整死亡率は1935年から1963年

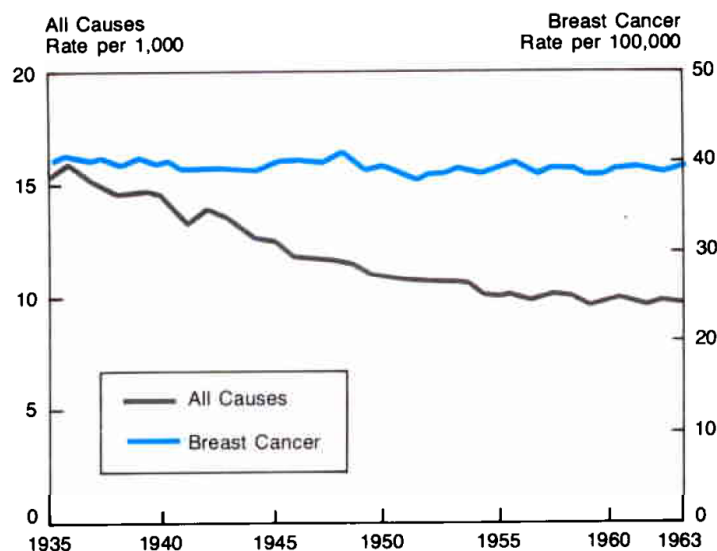


図 1.2 1935 年から 1963 年の米国における 25 歳以上の女性の年齢調整死亡率の変化。青線は乳癌の死亡率，黒線はすべてのがんの死亡率（文献 8 から引用）。

の 28 年間で約 39%も低下しているのに対し、乳癌についてはほとんど変化していないという結果となっている(図 1.2)。そこで 1963 年、the Health Insurance Plan of Greater New York (HIP) が中心となり「マンモグラフィを定期的に受けている女性はそうでない女性と比較して乳癌の早期発見が可能であるため死亡率が低下する」という仮説を立て、40 歳から 65 歳の女性を対象にスタディグループおよびコントロールグループ各々 30,000 人を登録しての大規模な調査が立ち上げられた。その追跡調査は S. Shapiro らによって 1966 年から 1982 年まで継続され 4 編の論文にまとめられた⁹⁻¹²⁾。16 年間にわたる追跡調査の結果、マンモグラフィを定期的に受けた女性はそうでない女性と比べて 30%以上の死亡率減少が認められることが示され、マンモグラフィによる乳癌検診が普及するきっかけを作った。

一方、我が国における乳癌検診は 1987 年に視触診によって開始された。2000 年に 50 歳以上の女性だけに限って視触診に加えてマンモグラフィの追加が認められたが、50 歳未満には認められなかった。その後 2004 年になって「乳癌検診はしこりが触知可能となる前の自覚症状がない段階での早期発見が重要である」との考えから、40 歳以上を対象にマンモグラフィを用いた 2 年に 1 回の検診を推奨することが提案され現在に至っている¹³⁾。

以上、マンモグラフィおよび乳癌検診の歴史を振り返ったが、この間、散乱 X 線除去用グリッドや幾種類にも及ぶ X 線管ターゲット/フィルタの開発、アナログ画像からデジタル画像への移行、有効なポジショニングの検討を含めた認定精度による総合的技術担保がなされたことに加え、近年では人工知能技術の導入等、多くの技術革新を経て現在のマンモグラフィを用いた乳癌検診のスタイルが確立された。さらに、将来的には被検者の乳房特性や罹患リスクに応じた検診モダリティの推奨を行うなど個別化乳癌検診の実践が期待されている。

1.2 本研究の歴史的背景

2003 年、米国コネチカット州に住む Nancy M. Cappello 氏は主治医の触診によって乳房にしこりを指摘された。彼女はそれまで 10 年以上にわたって毎年定期的にマンモグラフィによる乳癌検診を受けていたが、一度も異常を指摘さ

れたことは無かった。しこりを指摘された6週間前にもマンモグラフィ検診を受け「異常なし」の通知を受けたばかりであった。指摘された3日後に再度のマンモグラフィと超音波検査を受けたところ、マンモグラフィではやはり異常を指摘されなかったものの、超音波検査によって直径25mmの腫瘍と13個のリンパ節転移が見つかった。これはCappello氏が高濃度乳房ⁱⁱであったことが影響していると考えられる。医師たちは経験的に高濃度乳房ではマンモグラフィで病変を発見しにくいことを知ってはいたものの、検診においては「異常あり」か「異常なし」のいずれかの通知義務しかなく、結果の信頼性、すなわち病変が正常乳腺に隠されることの可能性についてまで言及することはなかった。Cappello氏は医師たちへ、高濃度乳房の女性については検査の信頼性や他のモダリティの選択等、がんの早期発見につながる情報を知らせるようにしてほしいと訴えたが「標準的な方法ではない」との理由で受け入れられなかった。しかし、このような訴えが少しずつ広がり始め、2005年コネチカット州では高濃度乳房の女性に対しては乳癌検診に超音波検査を追加し、その費用を保険で負担する法律が成立した。2008年Cappello氏はNPO法人「Are you dense? (denseは高濃度乳房のこと)」を創設しマンモグラフィ検診における高濃度乳房のリスクを訴えた。そして2009年、コネチカット州において初めて乳癌検診の際に被検者が高濃度乳房であるか否かの判定や診断の確からしさの通知を義務付ける法令が定められた。当初はコネチカット州だけであったが2015年には24州、2019年には全50州において同様の法令が定められ、米国で世界に先駆けて個別化乳癌検診が始まった。

わが国では2016年10月、全国32の乳癌患者団体代表が乳癌検診結果の通知方法を見直すよう厚生労働省へ要望書を提出したことを受け、2017年6月、厚生労働省は対策型検診を受けたすべての女性に対して高濃度乳房に関連する病変が隠される危険性等について通知・説明する体制を整備し、2018年3月までにガイドラインを作成すると発表した^{14, 15)}。しかし現状は、高濃度乳房かどうかの判定精度が十分ではないことや、追加モダリティとして期待される超

ⁱⁱ 乳房に存在する乳腺の割合を4段階に分類する際、最も乳腺が多いクラスと2番目に多いクラスを合わせて高濃度乳房という。

音波検査には死亡率減少のエビデンスがまだないなど、現時点では国費を投入する対応には至っていない。結局「全国の市町村で一律に乳房構成の判定（1.3参照）を通知することは時期尚早」とされ、従来どおり「異常あり・異常なし」と通知するのみにとどまっている¹⁶⁾。ただし、厚生労働省は各自治体の判断で被検者へ高濃度乳房に関連する個々の情報やそのリスクを説明することを妨げるものではないとしており、2017年7月現在、全国の自治体の約13%が独自の判断により対策型乳癌検診を受診した女性へ乳房構成の判定結果を通知している。また、聖マリアンナ医科大学ブレスト&イメージングセンターが2016年から2017年にかけて乳癌検診受診者約1000人に対して実施したアンケート調査によると「自分の乳房構成の判定を知りたい」と答えた女性は全体の85%にも上っている。2021年10月「がん予防重点健康教育及びがん検診実施のための方針」の一部改正により「ブレスト・アウェアネス(乳房を意識する生活習慣)」が盛り込まれたこともあり、自身の乳房の状態を知るという点からも、今後、より一層乳房構成判定の通知やそのリスクの説明を取り入れた個別化乳癌検診への国民の関心と期待が高まってくると予想される。

1.3 本論文の目的と構成

本論文は我が国における個別化乳癌検診の導入と実践に向け、人工知能技術を応用した医工連携の革新的手法を用いて高精度な乳腺密度ⁱⁱⁱ（皮膚および皮下脂肪を除いた乳房の質量に対する乳腺組織の質量の割合）の計測技術および予測技術を開発するものである。個別化乳癌検診を実現するにあたり現在の課題は、前頁で述べた保険制度に関する問題と技術的問題に大別されるが、このうち保険制度の問題については本論文では扱わない。個別化乳癌検診の実施にあたって未解決の技術的な問題点としては、乳腺密度の観点からは以下の3点である。

ⁱⁱⁱ 乳腺密度の表示法として、本論文で用いた乳房と乳腺の質量比以外に体積比や面積比を用いる場合もあり、どれを用いるかは目的により任意である。ただし、本論文のように平均乳腺線量の算出まで拡張する場合は必ず質量比を用いねばならない（私信、David R. Dance、Eメール、2017年2月9日）。

1) 1.2 で述べたように、検診受診者の乳腺密度を通知することの重要性が広く認識されるようになってからまだ 20 年程度であるが、その一方で乳腺密度が高ければマンモグラムで検出感度が低下することはそれ以前から経験的に知られていた。しかし、この経験則を高精度かつ客観的に検証することは困難であり現在に至るまで解決されていない。臨床画像を用いた検証法として中間期癌を利用する手法が用いられる場合があるが、中間期癌は発症時期が不詳であるため精度に問題があることがわかっている¹⁷⁾。検診受診者へ乳腺密度に応じたリスクを説明するためには、まずこの経験則を定量的に検証しておく必要があり、そのためには X 線減弱特性が実乳房と等価であるなど臨床画像を再現でき、かつ、乳腺密度と病変を任意に調整することができるファントムの開発が必要となる。

2) 臨床で撮影されたマンモグラムから乳腺密度を計測する技術は現在まだ完成されているとは言えない。これまでは、マンモグラムは 2 次元かつアナログであり、画像から乳房を構成する乳腺および脂肪各組織の質量を求めることは非常に困難で、定量的計測は不可能であった。そのため、従来は乳腺のエキスパートが目視で画像を観察し、主観に基づいて①きわめて高濃度、②不均一高濃度、③乳腺散在、④脂肪性、の 4 段階に分類してきた。これを「乳房構成判定」と呼び、国際的には American College of Radiology (ACR) が提唱する Breast Imaging Reporting & Data System (BI-RADS)¹⁸⁾ の 4 分類が用いられ、我が国では日本乳癌検診精度管理中央機構（以下、精中機構という）が定める 4 分類（BI-RADS の 4 分類を基本としている）が用いられている。しかし、両者とも乳腺エキスパートの目視によって決定されるため、主観の違い等に起因した結果のバラツキが問題となっており、近年、客観的な定量化の必要性が論じられている。

現在、精中機構の検診認定を受ける施設のほとんどはデジタルマンモグラムを採用している。デジタルマンモグラムの画素値は、乳房を透過してきた X 線の強度を表しており、これに加えて乳房の厚さや乳房を構成する脂肪組織および乳腺組織の線減弱係数、および X 線の光子エネルギースペクトルがわかれば、理論的には乳房中の脂肪組織と乳腺組織の割合、すなわち乳腺密度を算出することが可能である。しかし、実際の乳房の構造は複雑で、単純化された

モデルのように乳腺領域や脂肪領域の画素値を容易に決定することはできない。van Engeland¹⁹⁾が2006年にデジタルマンモグラムの画素値を利用して乳腺密度を算出して以降、いくつかのメーカーが同様の手法で臨床マンモグラムから乳腺密度を自動算出する装置を開発し市販している²⁰⁻²²⁾。しかし、乳腺密度算出にとって重要な因子である脂肪領域の画素値や乳腺領域は各メーカーが独自の手法で決定しているため、乳腺密度のゴールドスタンダードが定まらな
いまま今日にいたっている。このため乳腺密度の比較検討には同一メーカーの装置を用いざるを得ないのが実状である²³⁻²⁵⁾。一方、自施設で独自に必要な領域を抽出し乳腺密度を計測することも可能であるが、撮影されるすべての乳房画像について限られたエキスパートがそれらの作業を担うのは多大な労力を要することとなり現実的ではない。

3) 乳腺密度は年齢とともに変化することが以前から知られているが、その変化の程度には個人差がある。近年、乳腺密度の経時変化を調べ、新しく開発された治療薬の効果や乳癌手術後の予後予測に利用しようとする研究が行われており、いわゆるバイオマーカーとしての役割も期待されるようになってきた。マンモグラフィによる乳癌検診を定期的を受診していれば、そのデータはバイオマーカーとしての研究に役立つうえ、自身が乳癌になった場合の治療方針にも利用できるかも知れない。これもまた個別化乳癌検診が必要となる因子のひとつである。この研究はまだ始まったばかりで、データが少ないことに起因して十分な精度が得られていない²⁶⁾。しかし最も重要な問題は、現在行われている乳腺密度の計測手法ではマンモグラムの生データを用いねばならないことである。一般的に診断に用いられることのない生データは容量が大きいこともあり長期保存されない。上述した市販の乳腺密度解析ソフトウェアでも生データを用いての計測となる。したがって過去に遡って乳腺密度の経時変化を解析する課題は未解決のままである。

本論文は上述した3つの課題に焦点を当て、我が国での個別化乳癌検診の実現を支援する。第1章では、本論文の根幹をなすマンモグラフィや乳腺密度通知に関する歴史的背景をまとめ、現在未解決の問題点を明確にした。第2章では、乳腺密度を任意に調整可能な乳房等価オリジナルファントムを開発し、Receiver Operating Characteristic (ROC) 解析を用いて乳腺密度と病変検出率の

関係を定量化する。これにより、個別化乳癌検診で医師が受診者へ個々の乳腺密度や乳腺密度に応じた検診の信頼度およびリスクを説明するためのエビデンスに基づいたデータを提供する。第3章では、人工知能技術のひとつである U-net を用いて、乳腺エキスパートがマンモグラムから抽出した乳腺領域を自動抽出するシステムを開発する。またその結果を先行研究²⁷⁾ 著者らが提案したマンモグラム中の脂肪領域を決定する手法と組み合わせて乳腺密度および平均乳腺線量を高精度に自動算出し、臨床への適用の可能性を検証する。第4章では、過去に遡って乳腺密度の時系列解析を行うことができるように、マンモグラムを用いない乳腺密度の新しい算出法を提案する。これが高精度に達成されれば、乳腺密度をバイオマーカーとして用いるための研究を支援することができる。第5章では、本研究の成果を総括する。

このように、本研究は我が国における乳腺密度を考慮した個別化乳癌検診の実現にむけ、乳腺密度に関する技術的諸問題を独自の手法で解決し、保健学はもとより放射線技術学や放射線医学の発展に貢献するものである。

参考文献

- 1) W. C. Röntgen: Über eine neue Art von Strahlen: Erste Mittheilung, Sitzungsber. Physikmed. Ges. Würzburg,132-141(1895)
- 2) W. C. Röntgen: Über eine neue Art von Strahlen I , II , Ann. d. Physik.,64,1-17 (1898)
- 3) 国立がん研究センターがん情報サービスがん統計,
“https://ganjoho.jp/reg_stat/statistics/stat/summary.html”
- 4) A. Salomon: Beiträge zur Pathologie und Klinik der Mammacarcinome. Arch Klin Chir,101,573-668(1913)
- 5) W. Vogel: Die Röntgendarstellung von Mammatumoren. Arch Klin Chir,171:618-626(1932)
- 6) R. Leborgne: Diagnosis of Tumors of the Breast by Simple Roentgenography: Calcifications in Carcinomas. Am. J. Roentgenol.,65,1-11(1951)
- 7) R. L. Egan: Experience with Mammography in a Tumor Institution. Evaluation

- of 1000 Studies. *Radiology*,75,894-900(1960)
- 8) Vital Statistics of the United States. National Cancer for Health Statistics, Public Health Service,(1935-1963)
 - 9) S. Shapiro, P. Strax, L. Venet: Evaluation of periodic breast cancer screening with mammography methodology and early observations. *JAMA*,195,731-738(1966)
 - 10) S. Shapiro, P. Strax, L. Venet: Periodic breast cancer screening in reducing mortality from breast cancer. *JAMA*,215,1777-1785(1971)
 - 11) S. Shapiro, W. Venet, P. Strax, L. Venet, R. Roeser: Ten- to Fourteen-year Effect of Screening on Breast Cancer Mortality. *J. Natl. Cancer Inst.*,69,349-355(1982)
 - 12) S. Shapiro, W. Venet, P. Strax, L. Venet, R. Roeser: Selection, Follow-up, and Analysis in the Health Insurance Plan Study: A Randomized Trial with Breast Cancer Screening. *Natl. Cancer. Inst. Monogr.*,67,65-74(1985)
 - 13) 遠藤登紀子:乳癌早期発見の画像診断(1), マンモグラフィ検診を中心に. *明日の臨床* 16(2), 11-16 (2004)
 - 14) 宮崎日日新聞 : 2017年7月30日
 - 15) 日本経済新聞 : 2017年7月4日
 - 16) 平成30年5月24日付け健発第0331058号厚生労働省健康局長通知
 - 17) A. B. Hollingsworth: Redefining the sensitivity of screening mammography: A review. *Am J Surg.*, 218,411–418(2019)
 - 18) ACR BI-RADS atlas: Breast imaging reporting and data system: mammography, ultrasound, magnetic resonance imaging, follow-up and outcome monitoring, data dictionary, 5th ed., Reston, Va., American College of Radiology, c2013
 - 19) S. van Engeland, P. R. Snoeren, H. Huisman, C. Boetes, N. Karssemeijer: Volumetric breast density estimation from full-field digital mammograms, *IEEE Trans. Med.Imaging*,25,273–282(2006)
 - 20) R. Highnam, X. Pan, R. Warren, M. Jeffreys, G. Davey Smith, M. Brady: Breast composition measurements using retrospective standard mammogram form (SMF), *Phys. Med. Biol.*,51,2695-713(2006)
 - 21) R. Highnam, M. Jeffreys, V. McCormack, R. Warren, G. Davey Smith, M. Brady:

- Comparing measurements of breast density, *Phys Med Biol.*,52,5881-5895(2007)
- 22) S. Pahwa, S. Hari, S. Thulkar, S. Angraal: Evaluation of breast parenchymal density with QUANTRA software, *Indian J. Radiol. Imaging*, 25,391-396(2015)
 - 23) C. Schmachtenberg, S. Hammann-Kloss, U. Bick, F. Engelken: Intraindividual comparison of two methods of volumetric breast composition assessment, *Acad. Radiol.*,22,447-452(2015)
 - 24) O. W. E. Morrish, L. Tucker, R. Black, P. Willsher, S. W. Duffy, F. J. Gilbert: Mammographic breast density: comparison of methods for quantitative evaluation, *Radiology*, 275,356-365(2015)
 - 25) D. van der Waal, G. J. den Heeten, R. M. Pijnappel, K. H. Schuur, J. M. H. Timmers, A. L. M. Verbeek, M. J. M. Broeders: Comparing Visually Assessed BI-RADS Breast Density and Automated Volumetric Breast Density Software: A Cross-Sectional Study in a Breast Cancer Screening Setting, *PLoS One*,10(9):e0136667(2015)
 - 26) E. C. Atakpa, M. A. Thorat, J. Cuzick, A. R. Brentnall: Mammographic density, endocrine therapy and breast cancer risk: a prognostic and predictive biomarker review, *Cochrane Database of Systematic Reviews*, Issue 10. Art. No.: CD013091(2021)
 - 27) M. Yamamuro, Y. Asai, K. Yamada, Y. Ozaki, M. Matsumoto, Takamichi Murakami: Prediction of glandularity and breast radiation dose from mammography results in Japanese women, *Med. Biol. Eng. Comput.*,57,289-298(2019)

第 2 章 ROC 解析を用いた乳腺密度と病変検出感度の関係の定量化

2.1 緒言

第 1 章で述べたように、乳癌は世界的に見て女性の癌罹患率の上位を占める最も一般的な癌の一種であるが、早期発見によって死亡率を下げることができる癌としても知られている。マンモグラフィは、乳癌のスクリーニングに広く使用されている 2 次元の X 線画像診断法で、死亡率が有意に低下するという根拠がある唯一のモダリティであり、乳癌の画像診断にとって不可欠である¹⁻⁴⁾。しかし、マンモグラフィでは乳腺密度が病変の検出に影響し、高濃度乳房と呼ばれる脂肪の少ない乳房では注意が必要なことが知られている。これは正常な乳腺組織が密集していると、画像上ではマスキング効果によって病変検出感度が低下するためである⁵⁻¹⁰⁾。そのため、画像読影時にはマンモグラフィ検査結果に正常乳腺の影響を考慮できるよう、1.3 で述べたような乳房構成判定結果を記録することが必要である。近年、米国では乳房構成判定結果を被検者に通知することが求められ、2019 年にはすべての州で通知が義務付けられた。今後は、個別の乳房構成に基づいて、乳房エコーなどの適切な第 2、第 3 のモダリティの追加が推奨されるような個別の検診プログラムが準備されることが期待されている¹¹⁻¹⁵⁾。日本では 40 歳から 69 歳まで、隔年でマンモグラフィ検診を受けるよう推奨されており、通知については検討されている段階である。しかし乳房構成判定は目視で行われることが多く、BI-RADS¹⁶⁾ に準拠した定量化アルゴリズムが確立されていないのが現状である。

X 線画像がアナログからデジタルに移行したことで、近年、フルフィールドデジタルマンモグラフィ (FFDM) のデジタル値を利用した volumetric breast density measurement (VBDM) が開発された。これによって乳腺密度の定量的計測が可能になりつつある。本研究では FFDM によって生成された画素値を使用して乳腺密度を正確に決定するための新しい方法を提案した¹⁷⁾。しかし、マンモグラフィ検診で「乳腺密度」と「乳腺組織によって病変が隠されるリスク」との定量的な関係はまだ明確にされていない。Chiu らはレビュー論文で「長期間にわたる追跡調査で、乳腺密度が乳癌発症率・ステージ・死亡率・

マスクング効果と関連している検診感度へ及ぼす影響について明らかにした研究はほとんどない」と述べている¹⁸⁾。また、Ekpoら¹⁹⁾は多くの文献を調査した結果、「脂肪性乳房に対する高濃度乳房のマンモグラフィ感度はすべての文献で一貫して低下してはいるものの、その低下率は22.1%から70.0%までさまざまである」と報告している。このように、著者の知る限り乳腺密度とマンモグラフィで病変が隠されるリスクの関係はまだ定量的に明らかにされていないとは言えない。したがって、マンモグラフィ検診が乳腺密度に基づいた、より効果的な個別化検診へステップアップするには、FFDMのデジタル値から計測された乳腺密度と、病変が隠されるリスクの関係を明確に定量化することが求められる。

マンモグラフィの検出感度の指標として中間期癌の件数が用いられることがある。中間期癌は、前回のマンモグラフィ検診の未検出病変を反映していると仮定し、その件数を偽陰性として用いてきたが、1.3で述べたように中間期癌の真の発症時期は不明であるため、この仮定は厳密には成り立たない²⁰⁻²²⁾。また、この方法は定期的に検診を受けている女性にしか対応していないことも鑑み、Hollingsworthはマンモグラフィ感度の指標として中間期癌を使用することは適切ではないと述べている²³⁾。

このように、臨床画像を用いてマンモグラフィの検出感度を高精度に決定することは困難である。そこで第2章では、マンモグラフィで一般的に用いられているX線エネルギーの全範囲に対してX線減弱特性が実乳房と等価となる乳腺密度可変型オリジナルファントムを開発してReceiver Operating Characteristic (ROC) 実験を行い、乳腺密度がマンモグラフィの検出感度へ及ぼす影響を定量化する。

2.2 乳腺密度可変型乳房等価ファントムの開発

2.2.1 ファントムの使用素材

一般にマンモグラフィ装置で用いられる主な乳房ファントム、例えばポリメチルメタクリレートファントム^{24, 25)}や、米国Computerized Imaging Reference System (CIRS)社から提供されているACR認定FFDMファントム²⁶⁾は、マンモグラフィで使用されるX線光子エネルギー領域の一部（厳密には一点）でし

か実乳房に対して同じ X 線減弱特性を示さない。すなわち、実際の乳房組織とは異なる組成の材質をファントムに用いているため、コントラスト特性が実際の乳房画像とは異なっている。病変検出感度と乳腺密度を高精度に関連付けるためには、真の乳腺密度がわかっており、かつ、臨床でのコントラスト特性を正確に再現した X 線画像を用いる必要がある。そこで実乳房と同等の組成から成る素材を用いて、乳腺密度を任意に調整可能な乳房ファントムを開発した(以下、オリジナルファントムという)。このファントムはマンモグラフィ撮影で用いられる全 X 線光子エネルギー範囲で実乳房と同じ X 線減弱特性を持つため、臨床でのマンモグラフィの画像特性を再現することができる。

本章で開発したオリジナルファントムは、脂肪等価素材=C:72.0%、O:16.4%、H:9.2%、N:2.4%、乳腺等価素材=C:69.5%、O:17.3%、H:8.9%、N:2.3%、Ca:1.4%、P:0.6%で構成されており、(株)京都科学と共同開発したものである。この構成は国際放射線単位測定委員会(ICRU) Report 44²⁷⁾に掲載されている Adipose tissue および Breast tissue の構成とほとんど同じである。

オリジナルファントムに使用する脂肪等価素材と乳腺等価素材の X 線減弱特性を確認するため、各素材の線減弱係数を米国標準技術局が提供するデータベースシステム XCOM で計算し、ICRU Report 44 の値と比較した。ここで、質量減弱係数を線減弱係数に変換するための物理密度は同 Report 44 の値(脂肪組織=0.950、乳腺組織=1.020)を用いた。図 2.1 に示すように、両素材ともマンモグラフィで使用される X 線光子エネルギー範囲全域についてオリジナルファントムと ICRU の線減弱係数はよく一致していることがわかる。したがって、VBDM を目的とする場合、これらの素材は乳房ファントムを作成するのに理想適な素材であり、今回作成したオリジナルファントムと実乳房の乳腺密度の変化に対する画像特性は同等で、臨床画像と同様に評価することが可能である。

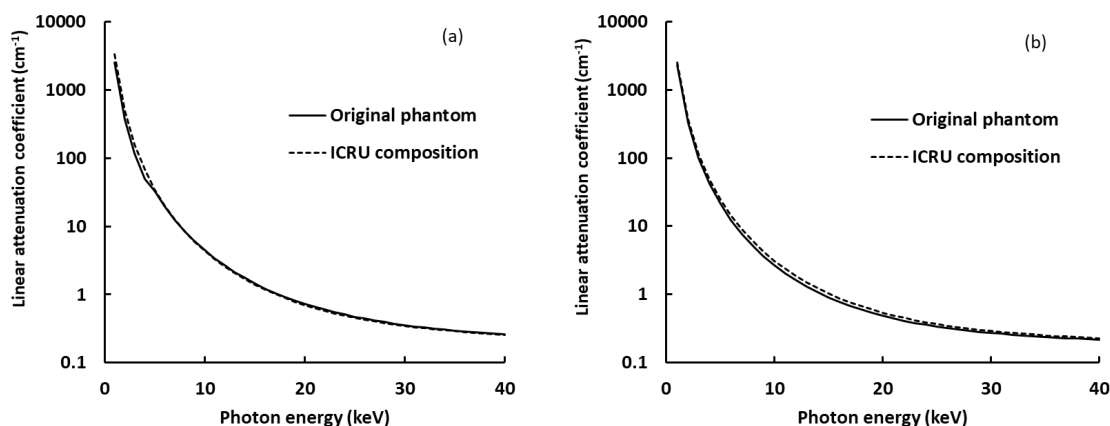


図 2.1 本研究で開発したオリジナルファントムの線減弱係数（実線）と ICRU Report 44 に掲載されている乳房の線減弱係数（点線）の比較. (a), (b)は各々乳房および脂肪の結果を示しており，両者とも非常によく一致している.

2.2.2 ファントムの作成

オリジナルファントムに用いた脂肪等価素材および乳房等価素材の製造方法を以下に示す。

- ① 脂肪等価素材と乳房等価素材のベース樹脂を、活性剤を含む電子天秤を用いて 0.1 グラム単位で個別に秤量する。
- ② 十分に攪拌した後、真空消泡剤で樹脂を脱気する。
- ③ 樹脂を専用の型に流し込み 24 時間放置する。
- ④ 硬化した樹脂を型から外し成形する。

上記のすべてのステップを、湿度 40%、温度 23°C に調整された一定の室内条件下で実行した。

作成した脂肪等価素材を大きさ 120 mm × 120 mm、厚さ 5 mm（5 枚）、10 mm（3 枚）、20 mm（3 枚）、30 mm（2 枚）のスラブ状に加工した。また、乳房等価素材は実際の乳房の走行に近づけるため、直径 0.3 mm から 2 mm、長さ約 50 mm の枝状に加工し、総質量を 250 g とした。

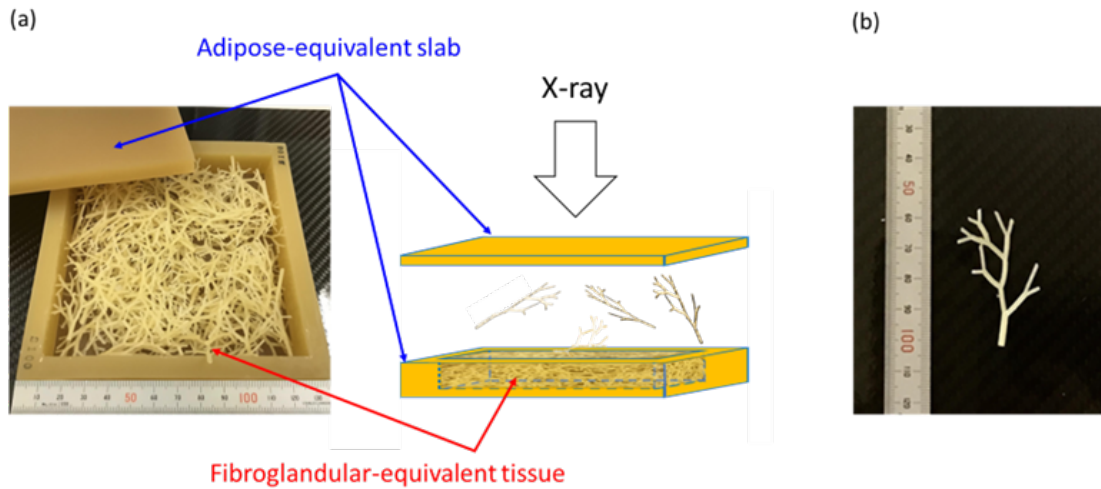


図 2.2 オリジナルファントムの説明図. (a) 組み立て図, (b) 乳腺等価素材のピースの一例. これと同類のピースを総質量 250g 作成した.

図 2.2 にオリジナルファントムの組み立て手順を示す。まず、(a)のようにスラブ状の脂肪等価素材を使用して、底厚 5 mm、幅 120 mm、奥行き 120 mm 高さ 10 mm の立体ボックスを作成した。次に、(b)に示すような枝状の乳腺等価素材を電子天秤で計量し目的の乳腺密度（25%、50%、および 75%）に相当する量（表 2.1 参照）を立体ボックスに充填した後、目的の乳腺密度に相当する厚さのスラブ状脂肪等価素材で蓋をし、圧迫乳房厚（compressed breast thickness : CBT）20 mm に相当するオリジナルファントムを作成した。

表 2.1 に本研究で用いた脂肪および乳腺等価素材の厚さおよび質量一覧を示す。本章において乳腺密度とは、第一章で述べた定義に従い、脂肪および乳腺等価素材の総質量に対する乳腺等価素材の質量の割合を指す。図 2.3 に乳腺密度 25%、50%、および 75% に調整したオリジナルファントムの X 線画像を示す。使用装置はキャノンメディカルシステムズ Pe-ru-ru、ターゲット/フィルタ : Mo/Mo、フラットパネルディテクタ (FPD) の画素サイズ : 85 μm \times 85 μm 、出力レベル : 13 bit、撮影条件は CBT 20 mm に対して臨床で用いているのと同じ（管電圧 29 kV、管電流 18 mA、auto exposure control : AEC 使用）で、画像処理は臨床画像と同様に自動設定とした。

表 2.1 各乳腺密度に対する脂肪等価素材と乳腺等価素材の厚さおよび質量. いずれも CBT 20 mm に相当する.

Breast density (%)	Adipose-equivalent tissue			Fibroglandular-equivalent tissue			Total	
	Area	Thickness	Weight	Area	Thickness	Weight	Thickness	Weight
	(mm)	(mm)	(g)	(mm)	(mm)	(g)	(mm)	(g)
25	120 × 120	15.0	216.0	120 × 120	4.8	72.0	19.8	288.0
50	120 × 120	10.0	144.0	120 × 120	9.5	144.0	19.5	288.0
75	120 × 120	5.0	72.0	120 × 120	14.3	216.0	19.3	288.0

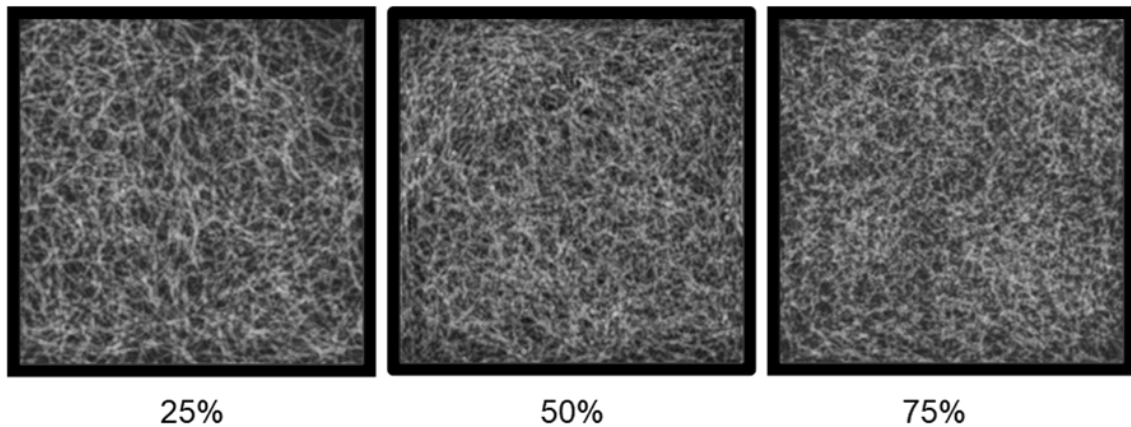


図 2.3 表 2.1 に示す構成に調整したオリジナルファントムを臨床画像と同じ露光条件で撮影した X 線画像. 左から順に乳腺密度 25%, 50%, 75% の例を示しており, 乳腺密度が高くなるほど白く見える乳腺等価組織が密になっている様子がわかる.

2.3 病変検出感度の乳腺密度依存性

2.3.1 ROC 解析のための画像作成および観察方法

図 2.3 に示す 3 種類の乳腺密度に調整したオリジナルファントムに模擬病変を挿入し、ROC 解析のための観察用 X 線画像を作成した。本研究で用いた模擬病変は微小石灰化（炭酸カルシウム、直径 $100\ \mu\text{m}$ × 6 個）、腫瘍性病変（塩化ポリビニル樹脂、直径約 5 mm）、およびスピキュラ状病変（模擬乳腺と同一素材、表面に凹凸状の突起（角状）をもつ直径約 10 mm の小球形）の 3 タイプである。図 2.4 に ROC 解析のための観察画像作成の様子を示す。乳腺密度 25%、

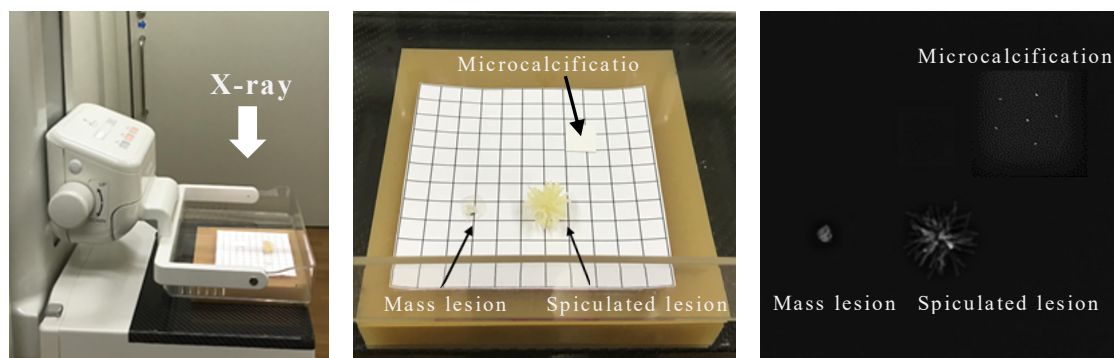


図 2.4 ROC 観察試料の作成。オリジナルファントムをマンモグラフィ装置にセットし、表面に模擬病変を置いて X 線撮影を行った（左図）。中央の図はファントム表面に置いた模擬病変の座標の一例。右図は形状を見やすくするために乳腺素材を入れずに撮影した模擬病変の X 線画像で、微小石灰化のみ拡大して表示している。

50%、75%の 3 種類についてそれぞれポジティブ画像とネガティブ画像を計 100 枚ずつ作成した。ポジティブ画像には病変タイプが 1 種類のみ挿入された画像、2 種類挿入された画像および 3 種類すべて挿入された画像を混在させ、どの病変タイプもポジティブ画像が 50 枚となるようにした。各病変は、110 mm × 110 mm の撮影範囲を 10 mm × 10 mm 毎に分割した 121 座標を設定し、乱数にしたがって配置した。使用機器、撮影条件、画像処理条件は 2.2.2 で述べたとおりである。図 2.5 に模擬病変を挿入したオリジナルファントム（乳腺密度 50%）の X 線写真の一例を示す。研究を実施した近畿大学病院は、2015 年以降デジタルマンモグラフィ施設認定を継続して取得しており、マンモグラフィに関する X 線出力や画質は適切に担保されている。

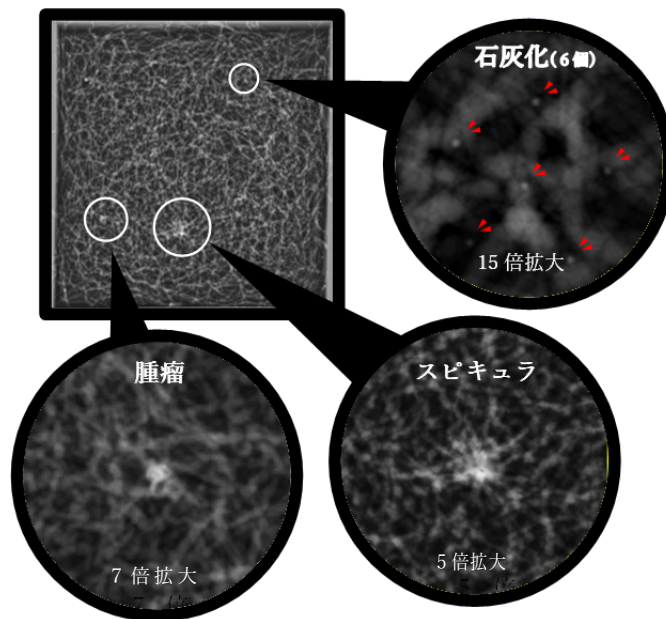


図 2.5 ROC 解析において観察に用いたポジティブ画像の一例（乳腺密度 25%）。ここでは 3 タイプの模擬病変がすべて挿入されているが、1 タイプのみ又は 2 タイプのみ挿入された画像も混在する。

作成した観察試料をマンモグラフィ検診認定診療放射線技師 5 名（認定歴：1 年、5 年、15 年、18 年、20 年）が各模擬病変毎に（1）間違いなく病変はない、（2）おそらく病変はない、（3）どちらとも言えない、（4）おそらく病変がある、（5）間違いなく病変がある、の 5 段階で評価した。観察条件は以下のとおりである。使用モニタ：解像度 5M ピクセル・観察輝度 600 cd/m²・画面サイズ 21 インチ・モノクロ液晶ディスプレイモニター、周辺照度：20 lx、観察時間および観察距離：フリー、window width および window level は固定、画像拡大および白黒反転は不可とした。観察は乳腺密度の低い順に行い、各乳腺密度における一連の観察時間を 2 時間以内とし、20 分以上のインターバルを挟んで次の乳腺密度における一連の観察を行った。観察者は一枚ずつランダムに表示された画像について、微小石灰化、腫瘍またはスピキュラが存在する確信度の判定を行った。ROC 解析には Web-based Calculator for ROC Curves (Engeland Morgan, Department of Radiology and Radiological Science, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA: <http://www.jrocfite.org>) を用いた。

乳腺密度の変化と病変検出率に統計的有意差があるかどうかを判断するために Microsoft Excel for Windows ver. 2016 (Microsoft Corporation、Redmond、Washington、USA) を用いて曲線下面積 (AUC) に対する *Student's t-test* (両側検定、Bonferroni 補正) を行った。本研究では $p < 0.05$ の場合、統計的に有意であると判断した。AUC は、ランダムに選ばれた一組のネガティブ画像とポジティブ画像のペアに対する観察者の評価において、ポジティブ画像の方が信号存在の確信度が高い確率を表しており、値が大きいほどネガティブ画像とポジティブ画像を適切に区別できていることを表している。

2.3.2 ROC 解析結果

表 2.2 および図 2.6 に各乳腺密度に対する AUC を示す。全体として乳腺密度の増加とともに病変の検出感度が低下する傾向が見られた。例えば、乳腺密度 25% を基準とした場合、乳腺密度 50% および 75% に対する AUC の低下率は、

表 2.2 病変毎の AUC. すべての病変で乳腺密度 25% (本研究で用いた中の最小値) の時 AUC が最も大きい. () 内の数値は 25% に対する低下率を示す.

Type of lesion	Breast density		
	25%	50%	75%
Microcalcification	0.860	0.656 (23.7)	0.574 (33.3)
Mass lesion	0.857	0.703 (18.0)	0.533 (37.8)
Spiculated lesion	0.910	0.778 (14.5)	0.770 (15.4)
Average	0.876	0.712 (18.7)	0.626 (28.5)

微小石灰化の場合それぞれ 23.7% および 33.3%、腫瘍の場合それぞれ 18.0% および 37.8% であった。しかし、スピキュラについては乳腺密度 25% に対する 50% および 75% の AUC の低下率はそれぞれ 14.5% および 15.4% とほとんど同じであった。この現象については 2.3.3 で詳しく述べる。表 2.3 に AUC の乳腺密度依存性に対する *Student's t-test* の結果を示す。乳腺密度が 25% から 50% へ変化した場合はどの病変も有意に AUC が低下したが、50% から 75% へ変化した場合、腫瘍は有意に AUC が低下したが、スピキュラと微小石灰化は統計的に有意

な低下を示さなかった。

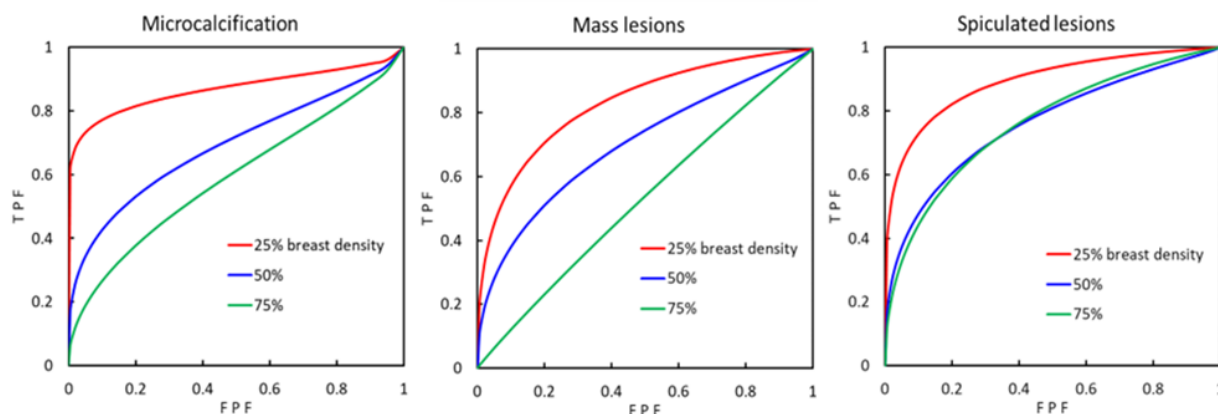


図 2.6 各病変における ROC 曲線.

表 2.3 AUC に対する乳腺密度間の t 検定の結果. 乳腺密度 25% と 50% 間ではすべての病変について有意差が認められるが, 50% と 75% 間では腫瘍を除いて $p > 0.05$ となり, 微小石灰化とスピキュラは AUC に統計的有意差があるとは言えない.

Type of lesion	Between 25% and 50%		Between 50% and 75%		Between 25% and 75%	
	<i>P</i>	95% CI	<i>P</i>	95% CI	<i>P</i>	95% CI
Microcalcification	0.0144*	0.0765 – 0.3331	0.1787	-0.0576 – 0.2212	0.0003*	0.2167 – 0.3565
Mass lesion	0.0001*	0.1258 – 0.1822	0.0058*	0.0822 – 0.2582	0.0003*	0.2453 – 0.4031
Spiculated lesion	0.0465*	0.0033 – 0.2603	0.7595	-0.0553 – 0.0701	0.0093*	0.0570 – 0.2214
Average	0.0039*	0.0880 – 0.2390	0.0113*	0.0325 – 0.1405	0.0000*	0.2144 – 0.2856

CI ; 信頼区間

図 2.7 は乳腺密度 25% におけるすべての病変の ROC 曲線を平均し、観察者毎にまとめたものである。便宜的に専門認定歴 1 年未満を ‘junior’、1 年以上 6 年未満を ‘semi-senior’、6 年以上 15 年未満を ‘senior’、15 年以上を ‘expert’ として区分した。図からわかるように、expert 2 名（観察者 A と B）はほとんど同じ AUC となり、ともに極めて高い検出能を示した。逆に、専門認定歴が 1 年未満の観察者 E は最も検出能が低く、他の観察者についても専門認定歴が長いほど検出能が高くなる傾向を示した。

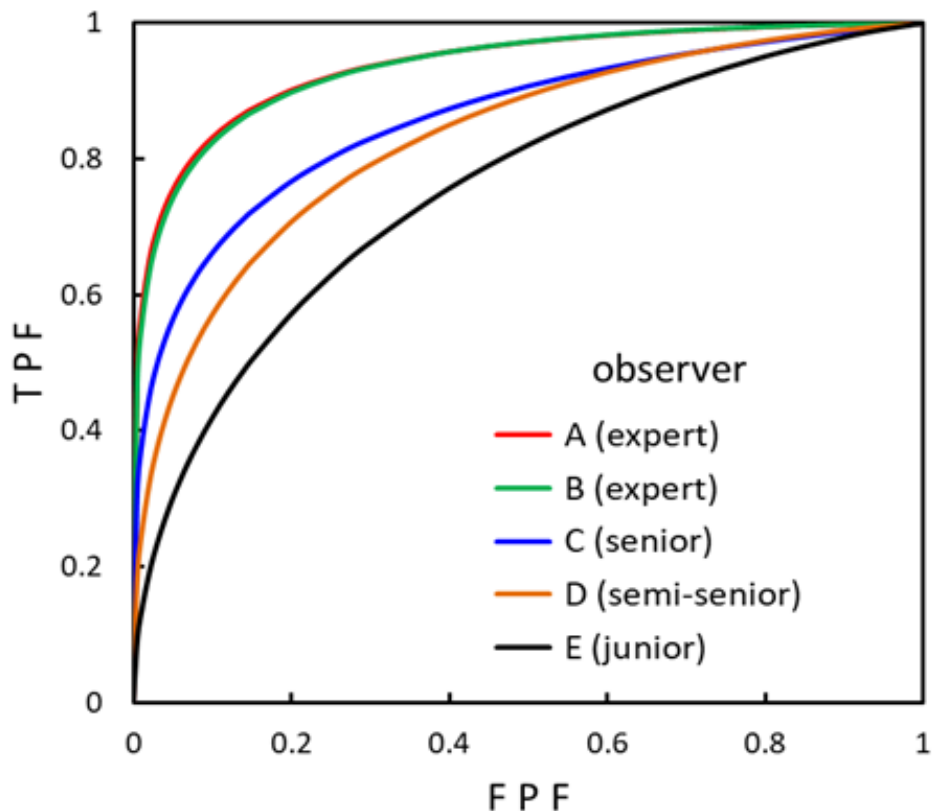


図 2.7 3 タイプの模擬病変に対する観察者毎の平均 ROC 曲線.
 乳腺専門認定歴が長いほど高い検出能を示す傾向がみられる.

2.3.3 ファントムのノイズパワースペクトル

模擬病変の空間周波数特性が検出感度に及ぼす影響を明らかにするために、オリジナルファントムのノイズパワースペクトル (NPS) を 2 次元高速フーリエ変換法を用いて乳腺密度 25%、50%、および 75% について測定した。本章で扱う「ノイズ」とは、図 2.3 のようなファントム内の模擬乳腺組織の分布を意味する。ファントム画像の中央領域から 1024×1024 ピクセルのサブ領域を抽出し、二次多項式による面近似補正によりヒール効果等の低空間周波数バックグラウンドを除去したあと、 256×256 ピクセルの ROI を設定し、x、y 各方向に 128 ピクセルずつオーバーラップしながら乳腺密度毎に 64 個の ROI を設定して 2 次元 NPS を計算し、アンサンブル平均を求めた。次に、2 次元 NPS の中心軸を挟む 7 ラインずつ計 14 ライン (中心軸は含めない) のデータを平均する

ことよって 1 次元 NPS を求めた²⁸⁾。

図 2.8 に各乳腺密度に調整したオリジナルファントムの NPS を示す。今回用いた模擬病変のひとつであるスピキュラの空間周波数帯域 (約 0.05 mm^{-1}) の値はグラフを内挿して求めた。図 2.8 からわかるように、乳腺密度 25%、50%、75% に調整したオリジナルファントムの 0.05 mm^{-1} でのスペクトル値 (ノイズ強度) は、乳腺密度 25% : 0.17 mm^2 、50% : 0.29 mm^2 、75% : 0.33 mm^2 であり、乳腺密度 25% のノイズ強度と比べると 50% および 75% のノイズ強度はどちらも約 1.8 倍以上高かったが、一方 50% と 75% では顕著な差が見られなかった。

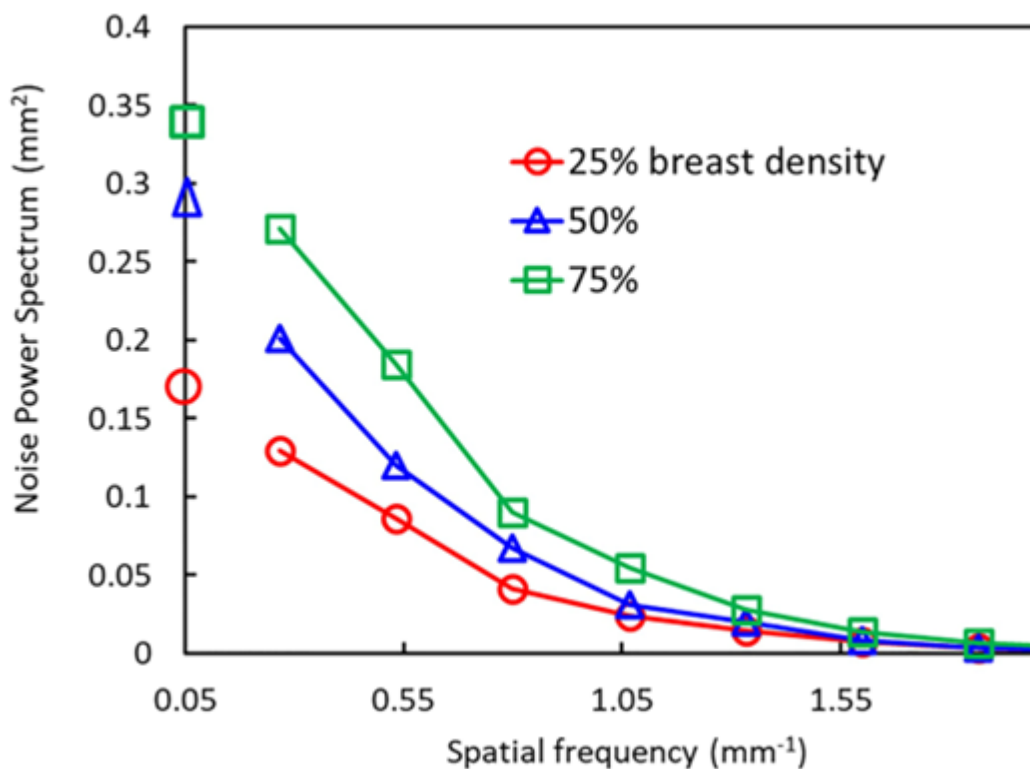


図 2.8 各乳腺密度におけるオリジナルファントムの NPS (赤：乳腺密度 25%、青：50%、緑：75%)。スピキュラの空間周波数帯域 (約 0.05 mm^{-1}) では乳腺密度 50% と 75% のノイズ強度は接近しており、どちらも 25% と比べて顕著に高い。

2.4 考察

正確な乳腺密度計測方法が確立されたとしても、マンモグラフィで正常乳腺に病変が隠されるリスクに対する乳腺密度の影響が解明されない限り、乳腺密度に応じた個別の対応を確立することは困難である。臨床画像を用いた検出感度の評価では、乳腺密度に加え、乳腺組織の分布、ポジショニングや乳房圧迫などの撮影技術、圧迫された乳房の厚さ、被検者の年齢、読影者の集中力、人口あたりの有病率、画像処理条件など、多くの要因が複雑に重なり合って影響すると報告されており¹⁹⁾、これらの影響が研究者間で検出感度が異なる結果を導き出す原因になることは明らかである。したがって、精度良く乳腺密度だけの影響を評価するためには、個々の要因を分離したアプローチが必要となる。

本章では、マンモグラフィの検出感度に対する乳腺密度の影響だけを調べるために、乳腺密度を任意に変えられるオリジナルファントムを作成し、これを用いて ROC 実験を行った。観察者は全員、精中機構が実施する乳癌検診の品質保証に関するマンモグラフィ検診の技術、読影、品質管理に関する特別プログラムを用いて訓練と試験を受け、認定を取得した診療放射線技師であった。

まず、本章で行ったオリジナルファントムによる ROC 実験が適切に臨床マンモグラフィを再現できているかどうかについて述べる。図 2.7 に示すように専門認定歴が長いベテラン観察者ほど AUC が大きく検出能が高い傾向にあった。ただし、今回の実験では各観察者につき 3 回の観察実験（微小石灰化・腫瘍・スピキュラ各 1 回）の平均しか取ることができないため検定を適用することが困難であり、観察者の検出能の差が統計的に有意であることを検証することができなかった。しかし、認定歴が長い、すなわちマンモグラフィの臨床に携わっている期間が長いほど検出能が高い傾向にあることが実験結果から推定された。したがって本研究で開発したオリジナルファントムおよび模擬病変を用いた実験によって、臨床マンモグラフィでの病変検出の状況を適切に再現できることが示唆されたと考える。

病変検出能の乳腺密度依存性については、よく知られている「乳腺密度が高いほど病変検出感度が低下する」という経験則とほぼ一致した。しかし、表 2.2 に見られるように、本研究では乳腺密度 75% の AUC は 25% の AUC に対し最

大 37.8%の低下率であったが、Mandelson らの研究²⁹⁾では脂肪性乳房に対する高濃度乳房の検出感度の低下率は最大で 62%であり、本研究の結果とは 20%以上の違いがあった。これは先に説明した乳腺密度以外のいくつかの要因に加え、中間期癌を用いて調査を行ったことが影響していると考えられる。この他にも脂肪性乳房に対する高濃度乳房のマンモグラフィ感度の低下率が本研究の結果より高い報告が散見される (Mousa et al.³⁰⁾; 34.3~62.5%、Ekpo et al.¹⁹⁾; 22.1~70.0%)。マンモグラフィの検出感度低下への対策を行うためには、絡み合った要因それぞれの影響を解明することが必要であるが、他の要因の影響については今後の課題である。

表 2.3 および図 2.6 において腫瘍性病変の検出結果では乳腺密度が高いほど検出能は有意に低下したが、一方、サイズが大きくコントラストの低いスピキュラ病変や、サイズが非常に小さくコントラストも低い微小石灰化では、乳腺密度 25%と 50%間では有意な差が認められたものの 50%と 75%の間には有意な差がなかった。特にスピキュラについてはほとんど検出能に差がなかった。オリジナルファントムの正常乳腺組織の空間周波数特性と模擬病変の空間周波数帯域を調べた結果、病変タイプ (サイズやコントラスト) によっては乳腺密度の影響を受けにくい空間周波数帯域があると推察できる。図 2.8 に示したオリジナルファントムの各乳腺密度の NPS とスピキュラ病変の空間周波数帯域の関係は図 2.6 の ROC 曲線の結果と一致しており、本章で用いたオリジナルファントムの場合には直径約 10 mm の被写体の検出が乳腺密度の影響を受けにくいということが示唆された。また微小石灰化の検出感度は、乳腺密度が 25%から 50%の間で急激に低下したが、それ以上乳腺密度が高くなっても顕著な変化がなかったことから、微小石灰化の検出には閾値が存在する可能性も示唆された。これらのことから、マンモグラフィの検出感度に対する乳腺密度の影響を論じる場合、病変のタイプに応じて乳腺密度と検出感度の非線形な関係を考慮する必要があると考える。

ここで本章の limitation について述べる。第一に、ROC 実験で用いた模擬病変には、マンモグラフィで良悪性鑑別が必要な陰影として一般的である腫瘍、微小石灰化、スピキュラの 3 種類の病変タイプ各 1 つずつとしたが、臨床画像では病変の形状やサイズ、およびコントラストなどが多様である。今後は、よ

り多くの病変タイプを用いた実験や、同じ乳腺密度についても正常乳腺の走行状態の変化をファントムに取り入れるなど、評価画像をより一層臨床環境に近づけた検出感度の定量化が必要である。第二に、本章では5人の認定診療放射線技師がROC実験の観察を行ったが、さらに多くの観察者を追加することで、より信頼度の高い結果を導くことができると考える。今後、これらを取り入れた研究により、乳腺密度と病変検出感度の関連がより詳細に示されれば、現在、乳房構成分類の世界基準となっているBI-RADSの4分類に追加や修正を加えた新しい分類基準の作成につなげることができるものと考ええる。

2.5 結言

乳腺密度と病変検出感度の関係を定量化するために、マンモグラフィで用いられるX線光子エネルギーの全範囲について脂肪および乳腺組織と同等のX線減弱特性を有する乳腺密度可変型乳房等価ファントムを開発し、3種類の模擬病変と3段階の乳腺密度に対してROC実験を行った。従来の「病変検出感度は乳腺密度の増加とともに低下する」という経験則を定量化し検定することで、その傾向は病変のタイプにも依存することを見出した。本章の結果は、乳腺密度と病変検出感度の間に非線形の関係があることを示しており、特に検出感度が低い高濃度乳房被検者へのスクリーニングマンモグラフィのリスク説明に有用である。

参考文献

- 1) M. G. Marmot, D. G. Altman, D. A. Cameron, J. A. Dewar, S. G. Thompson, M. Wilcox: The benefits and harms of breast cancer screening, an independent review, *Lancet*, 380,1778-1786(2012)
- 2) M. M. Schapira, R. A. Hubbard, H. H. Seitz, E. F. Conant, M. Schnall, J. N. Cappella, et al: The impact of a risk-based breast cancer screening decision aid on initiation of mammography among younger women: report of a randomized trial. *MDM Policy Pract.*,4,1-13 (2019)
- 3) A. L. Siu: Screening for breast cancer: US Preventive Services Task Force

- recommendation statement, *Ann Intern Med*,164,279-296(2016)
- 4) K. C. Oeffinger, E. T. Fontham, R. Etzioni, A. Herzig, J. S. Michaelson, Y. C. Shih, et al: Breast cancer screening for women at average risk: 2015 guideline update from the American Cancer Society, *JAMA*,314,1599-1614(2015)
 - 5) M. Von Euler-Chelpin, M. Lillholm, I. Vejborg, M. Nielsen, E. Lynge: Sensitivity of screening mammography by density and texture: a cohort study from a population-based screening program in Denmark, *Breast Cancer Res.*,21, 111(2019)
 - 6) M. Posso, J. Louro, M. Sánchez, M. Román, C. Vidal, M. Sala, et al: Mammographic breast density: how it affects performance indicators in screening programmes? *Eur. J. Radiol.*,110,81-87(2019)
 - 7) S. J. Vinnicombe: Breast density, why all the fuss? *Clin. Radiol.*,73,334-357(2018)
 - 8) I. Théberge, M. H. Guertin, N. Vandal, G. Côté, M. P. Dufresne, É Pelletier, et al: Screening sensitivity according to breast cancer location, *Can. Assoc. Radiol. J.*,70,186-192(2019)
 - 9) S. Weigel, W. Heindel, J. Heidrich, H. W. Hense, O. Heidinger: Digital mammography screening, sensitivity of the programme dependent on breast density, *Eur. Radiol.*,27,2744-2751(2017)
 - 10) I.Saarenmaa, T. Salminen, U. Geiger, P. Heikkinen, S. Hyvärinen, J. Isola, et al: The effect of age and density of the breast on the sensitivity of breast cancer diagnostic by mammography and ultasonography, *Breast Cancer Res. Treat.*,67, 117-123(2001)
 - 11) L. B. Spalluto, C. L. Roumie, K. R. Bonnett, D. G. Schlundt, C. M. DeBenedectis, C. H. Wilkins: Women's response to state-mandated language in dense breast notification, *Breast J.*,24,1046-1050(2018)
 - 12) N. Houssami, C. I. Lee: The impact of legislation mandating breast density notification - review of the evidence, *Breast*,42,102-112(2018)
 - 13) M. Horný, M. Shwartz, R. Jr. Duszak, C. L. Christiansen, A. B. Cohen, J. F. Jr. Burgess: Characteristics of state policies impact health care delivery: an analysis

- of mammographic dense breast notification and insurance legislation, *Med. Care.*,56,798-804(2018)
- 14) C. M. Gunn, N. R. Kressin, K. Cooper, C. Marturano, K. M. Freund, T. A. Battaglia: Primary care provider experience with breast density legislation in Massachusetts, *J. Womens Health*,27,615-622(2018)
 - 15) S. L. Chau, A. Alabaster, K. Luikart, L. M. Brenman, L. A. Habel: The effect of California's breast density notification legislation on breast cancer screening, *J. Prim Care Community Health*,8,55-62(2017)
 - 16) ACR BI-RADS atlas: Breast imaging reporting and data system: mammography, ultrasound, magnetic resonance imaging, follow-up and outcome monitoring, data dictionary, 5th ed., Reston, Va., American College of Radiology, c2013
 - 17) M. Yamamuro, Y. Asai, K. Yamada, Y. Ozaki, M. Matsumoto, T. Murakami: Prediction of glandularity and breast radiation dose from mammography results in Japanese women, *Med. Biol. Eng. Comput.*,57,289-298(2019)
 - 18) S. Y. H. Chiu, S. Duffy, A. M. F. Yen, L. Tabár, R. A. Smith, H. H. Chen: Effect of baseline breast density on breast cancer incidence, stage, mortality, and screening parameters, 25-year follow-up of a Swedish mammographic screening, *Cancer Epidemiol Biomarkers Prev.*,19,1219-1228(2010)
 - 19) E. U. Ekpo, M. Alakhras, P. Brennan: Errors in mammography cannot be solved through technology alone, *Asian Pac J Cancer Prev.*,19,291-301(2018)
 - 20) S. Hofvind, B. M. Geller, J. Skelly, P. M. Vacek: Sensitivity and specificity of mammographic screening as practised in Vermont and Norway, *Br. J. Radiol.*,85: e1226-e1232(2012)
 - 21) M. T. Mandelson, N. Oestreich, P. L. Porter, D. White, C. A. Finder, S. H. Taplin, et al: Breast density as a predictor of mammographic detection, comparison of interval- and screen-detected cancers, *J. Natl. Cancer Inst.*,92: 1081-1087(2000)
 - 22) D. S. M. Buist, P. L. Porter, C. Lehman, S. H. Taplin, E. White: Factors contributing to mammography failure in women aged 40-49 years, *J. Natl. Cancer Inst.*,96,1432-1440(2004)

- 23) A. B. Hollingsworth: Redefining the sensitivity of screening mammography, A review. *Am. J. Surg.*,218,411-418(2019)
- 24) R. W. Bouwman, R. E. Van Engen, K. C. Young, G. J. den Heeten, M. J. M. Broeders, S. Schopphoven, et al: Average glandular dose in digital mammography and digital breast tomosynthesis: comparison of phantom and patient data, *Phys. Med. Biol.*,60,7893-7907(2015)
- 25) R. W. Bouwman, R. E. Van Engen, K. C. Young, W. J. Veldkamp, D. R. Dance: Dose assessment in contrast enhanced digital mammography using simple phantoms simulating standard model breasts, *Phys. Med. Biol.*,60, N1-N7(2014)
- 26) K. Izdihar, K. C. Kanaga, V. Krishnapillai, T. Sulaiman: Determination of tube output (kVp) and exposure mode for breast phantom of various thicknesses/glandularity for digital mammography, *Malays J. Med. Sci.*,22,40-49(2015)
- 27) ICRU Report 44: Tissue substitutes in radiation dosimetry and measurement, ICRU (International Commission on Radiation Units and Measurements), Bethesda, MD (1989)
- 28) IEC (International Electrotechnical Commission): IEC 62220-1-1:2015, Medical electrical equipment — Characteristics of digital X-ray imaging devices — Part 1-1: Determination of the detective quantum efficiency — Detectors used in radiographic imaging. Geneva, Switzerland: IEC (2015)
- 29) M. T. Mandelson, N. Oestreicher, P. L. Porter, D. White, C. A. Finder, S. H. Taplin, E. White: Breast density as a predictor of mammographic detection: comparison of interval- and screen-detected cancers, *J. Natl. Cancer Inst.*,92, 1081-1087(2000)
- 30) D. S. A. L.Mousa, E. A. Ryan, C. M. Thoms, P. C. Brennan: What effect does mammographic breast density have on lesion detection in digital mammography? *Clin. Radiol.*,69,333-341(2014)

第3章 乳腺密度計測への人工知能技術の応用

3.1 緒言

Global Burden of Disease Cancer Collaboration¹⁾によると、乳癌は世界の女性のがん罹患率と死亡率で第1位にランクされている。乳癌死亡率を低下させることが証明されている唯一のモダリティとして、マンモグラフィが広く使用されているが²⁾、一方でマンモグラフィの特徴として、病変検出感度が乳腺密度の影響を受けやすく、高濃度乳房では病変が隠されてしまう可能性が高いことについては第2章で述べたとおりである。さらに高濃度乳房では乳癌発症リスクが増加することは確実であり³⁻¹⁰⁾、乳癌の発症リスクモデルや予防介入モデルの確立が必要である。もし、乳腺密度を客観的、かつ、高精度に数値化することができれば、上記リスクモデルや介入モデルの確立へつなげることが可能となる。このような理由からマンモグラフィ検診では乳房構成を評価することが重要視されており、乳房構成に基づく個別化された検診プログラムは乳癌検診の有効性をより一層向上させることが期待できる。Hollingsworth¹¹⁾は、「この個別化医療の時代では、数学モデルを使用し個別に乳癌リスクを日常的に計算するのと同じように、個別の検診感度を計算することは価値のある目標と見なされるべきである」と述べている。また、個別化されたマンモグラフィ検診では、放射線科医は被検者に乳房構成（理想的には個別化された定量的計測値）および Dance¹²⁾ が提唱するマンモグラフィの被ばく線量を表す平均乳腺線量（mean glandular dose: MGD）を通知することが望ましい¹³⁻¹⁶⁾。しかし、従来は乳腺密度を正確に定量化することが困難であったため、MGD 算出には一律50%の乳腺密度が用いられてきた。今後は個々の乳腺密度で補正した正確な平均乳腺線量（以下、corrected MGD と定義する）を通知することが期待される。

米国は、2019年に任意保険プランを利用して、すべての州のすべての受検者に個々の乳房組成の通知を義務付けた最初の国である。ただし、現在、乳腺密度を計算するために使用できるソフトウェアは複数あるもののアルゴリズムは独自仕様であり、計算された値はソフトウェアによって大きく異なっている^{17, 18)}。信頼できる個別化された検診プログラムに必要な正確なデータを提供するには、さらなる技術的な研究が必要である。したがって、本章で述べる研究と

比較できる先行研究は存在しない。

従来、乳房構成判定は乳腺専門認定を持つ放射線科医や診療放射線技師によって目視により決定されており、その結果、観察者間の変動が顕著であることがわかっている¹⁹⁾。近年、これらの問題を解決するために、コンピューターにより自動化された VBDM が導入されつつある²⁰⁻²³⁾。VBDM では、乳房全体を計測対象とすることがよくあるが²⁰⁻²²⁾、この場合、乳腺組織が存在する位置の偏りによって乳腺密度が過小評価されることに注意する必要がある。著者は VBDM の計測対象領域として、現在乳腺組織が存在する領域 (**dense region**) と、さらに過去に乳腺組織が存在した領域も含めた領域 (**diffuse region**) の 2 つの領域タイプ²⁴⁾について計測する必要があると考えている。**Dense region** は病変が正常乳腺に隠される危険性と直接関連している非常に重要な領域である。一方、**diffuse region** は乳腺組織を目視では認識できなくても、過去に存在していたと考えられる領域も含めていることから、乳癌発症リスクと関連する領域である。図 3.1 に VBDM ではこれら 2 つのタイプの乳腺領域を計測対象とすることの重要性を示す²⁵⁾。乳房全体を計測対象とした場合は同じ乳腺密度であっても、それぞれの限定された領域では乳腺密度が顕著に異なっていることがわかる。

近年、信頼性の高い VBDM を実現するために、人工知能技術の一種であるディープラーニングを用いた乳腺領域の自動抽出が試みられている²⁶⁻²⁸⁾。乳腺密度の高精度な定量化は MGD の高精度な算出につながる。しかし、乳腺領域の自動抽出は、現時点ではまだ十分には確立されているとは言えない。なぜなら、その抽出精度は単に抽出領域の形状について検証されているだけで、VBDM にとって必要不可欠な乳腺密度や MGD を評価尺度とした検証はまったく行われていないからである。さらに、先行研究では **dense region** だけが計測対象とされており、**diffuse region** は乳癌発症リスクと関係するにもかかわらず計測対象とされることはなかった²⁶⁻²⁸⁾。

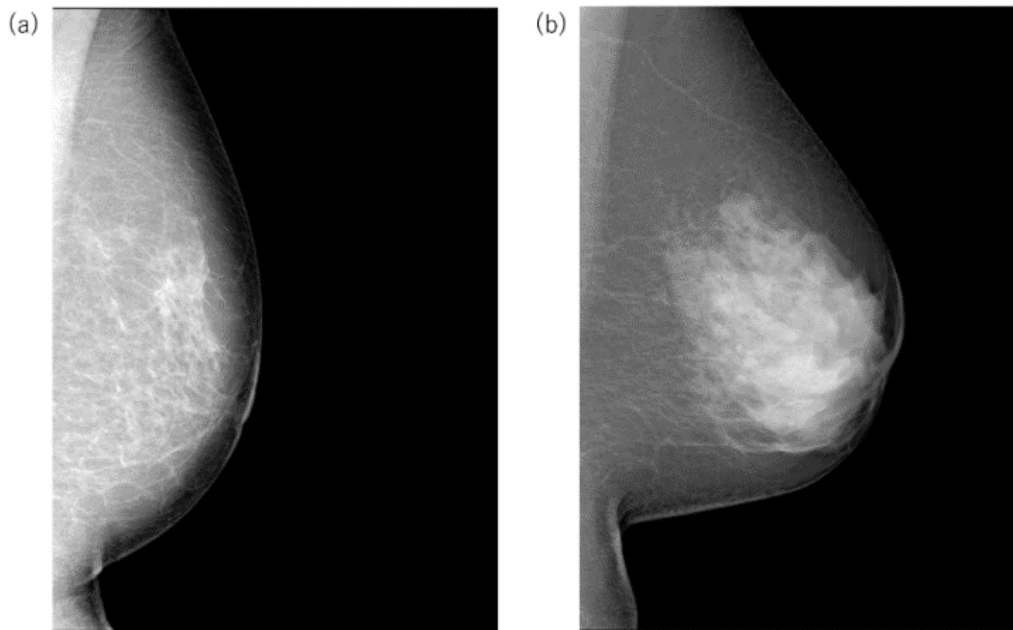


図 3.1 乳腺密度計測のための乳腺領域抽出の必要性. 乳房全体に対する乳腺密度は(a), (b)とも同じ27%であるが, (b)の乳房は(a)の乳房よりも乳腺組織が密集しているため病変を見逃すリスクが高くなる.

現在、精中機構では乳腺領域に **diffuse region** を含めることを義務付けている。しかし、このことはディープラーニング技術の実践にとっては非常に困難なタスクである。なぜなら **diffuse region** は領域の辺縁を視覚的にはっきりと認識することができないからである。著者の知る限り、VBDM のために **diffuse region** を計測対象としてディープラーニング技術を適用した研究は行われていない。

マンモグラフィを用いた高精度な VBDM 実現のためには、上述した2つのタイプの乳腺領域を再現性良く高精度に自動抽出すること、およびマンモグラム上の脂肪のみの領域を高精度に決定することが必要となる。これらのうち、脂肪のみの領域の高精度な決定法については、すでに著者らが先行研究においてマンモグラムの解剖学的知識に基づく新しい手法を提案した²⁹⁾。本章では、信頼性の高い個別化検診マンモグラフィプログラムの実現を目指し、ディープラーニング技術を用いてマンモグラムから **dense region** と **diffuse region** を自動抽出する手法、およびその抽出精度を乳腺密度と **corrected MGD** の観点から分析する手法について述べる。

3.2 U-net を用いた乳腺領域の自動抽出

3.2.1 乳腺領域の定義

マンモグラフィを用いた VBDM の計測対象領域となる dense region および diffuse region の 2 つの領域タイプの定義について説明する。3.1 で述べたよう

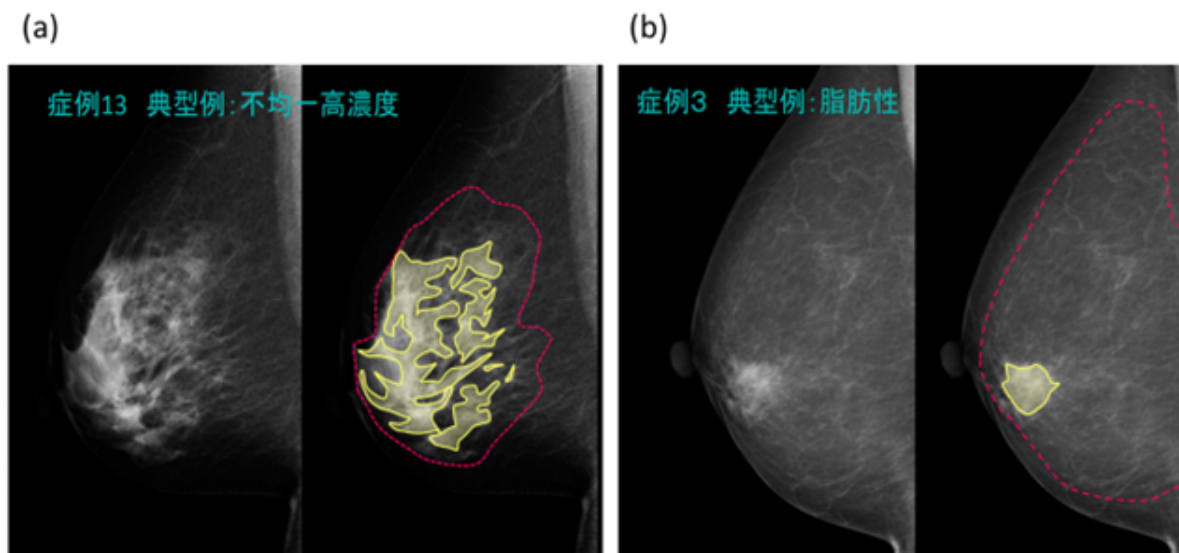


図 3.2 Dense region (黄色で囲まれた領域) と diffuse region (赤で囲まれた領域) の抽出例. (a)は不均一高濃度, (b)は脂肪性乳房の一例で, 左側は原画像.

— 参考文献 24), 乳房構成判定アトラス (笠原, 他) より引用 —

に dense region は正常乳腺が現在存在している領域で、乳腺組織と脂肪組織では X 線に対する減弱特性が異なるため、画像濃度の差によって目視で領域を判別することができる。一方、diffuse region は目視では乳腺組織を認識できなくても、過去に存在していたと考えられる領域を含めるため画像濃度の差によって判別するのが困難である。また、diffuse region は皮下脂肪や大胸筋部分を除き、乳腺後隙のように明らかに乳腺組織が存在しない領域は含めない。結果として乳頭から胸壁側に向かって扇状に広がるような形になる。図 3.2 に笠原らによる両乳腺領域の典型的な抽出例を示す²⁴⁾。

3.2.2 データ収集

本章で用いた画像は、近畿大学病院で 2021 年 4 月から 10 月までに取得された計 703 枚の病変を含まない medio-lateral oblique (MLO) view マンモグラム

で、年齢は 22 歳から 80 歳、CBT 16 mm から 82 mm である。圧迫圧が 90 N 未満およびポジショニング不良の画像は除外した。使用装置は第 2 章と同じであり、撮影には AEC を使用した。マンモグラフィシステムは、撮影毎に生データおよび臨床診断で用いられる画像処理（階調処理、周波数処理およびダイナミックレンジ圧縮処理）を施したデータで構成される 2 種類のマンモグラムを自動的に作成する。そのうち生データマンモグラムは乳腺密度計算用として、また画像処理後のマンモグラムは乳腺領域抽出用として外部メディアへ記録した。なお、本研究は近畿大学医学部倫理審査委員会の承認（R02-237）を得て、研究に参加したすべての被検者から書面によるインフォームドコンセントを取得して実行された。

3.2.3 Ground Truth の作成

読影診断で用いる画像処理後のマンモグラムを使用して、専門認定歴 20 年、10 年、1 年の 3 名の診療放射線技師が dense region および diffuse region の ground truth (GT) を独立して作成した。1 枚のマンモグラムにおいて、dense region と diffuse region の両方を認識できなくても、どちらかが認識できれば GT として採用した。3 名の診療放射線技師は、作業の前に 20 年以上の経験を持つ乳腺専門放射線科医から Breast Imaging Reporting and Data System: BI-RADS³⁰⁾に基づいた dense region および diffuse region について 6 か月間隔で 2 度のレクチャーを受け、その後 3 名の間でディスカッションを行い、全員が内容を理解していることを確認した。このような 2 度にわたるレクチャーの有効性は、先行研究で検証済みである³¹⁾。

3.2.4 U-net モデルの構築

3.2.3 で作成した GT を使って学習するのに先立ち、収集したすべての元画像（画素サイズ 2016 × 2816、コントラスト解像度 13 ビット、最大画素値 7500）を画素サイズ 256 × 256、コントラスト解像度 8 ビット、最大画素値 256 にリサイズした。コンピュータプラットフォームとして MATLAB 2020a (MathWorks, Natick, USA) を、ネットワークの学習には adaptive moment estimation (Adam) optimizer を用いた。図 3.3 に本研究で使用した U-net の構造を示す。設定条件

は以下のとおりである。活性化関数：ReLU、損失関数：交差エントロピー、ストライド：1、ディレーション：1、学習率：0.0001、エポック数：50、ミニバッチサイズ：8。これらの条件を dense region と diffuse region の抽出に適用した。

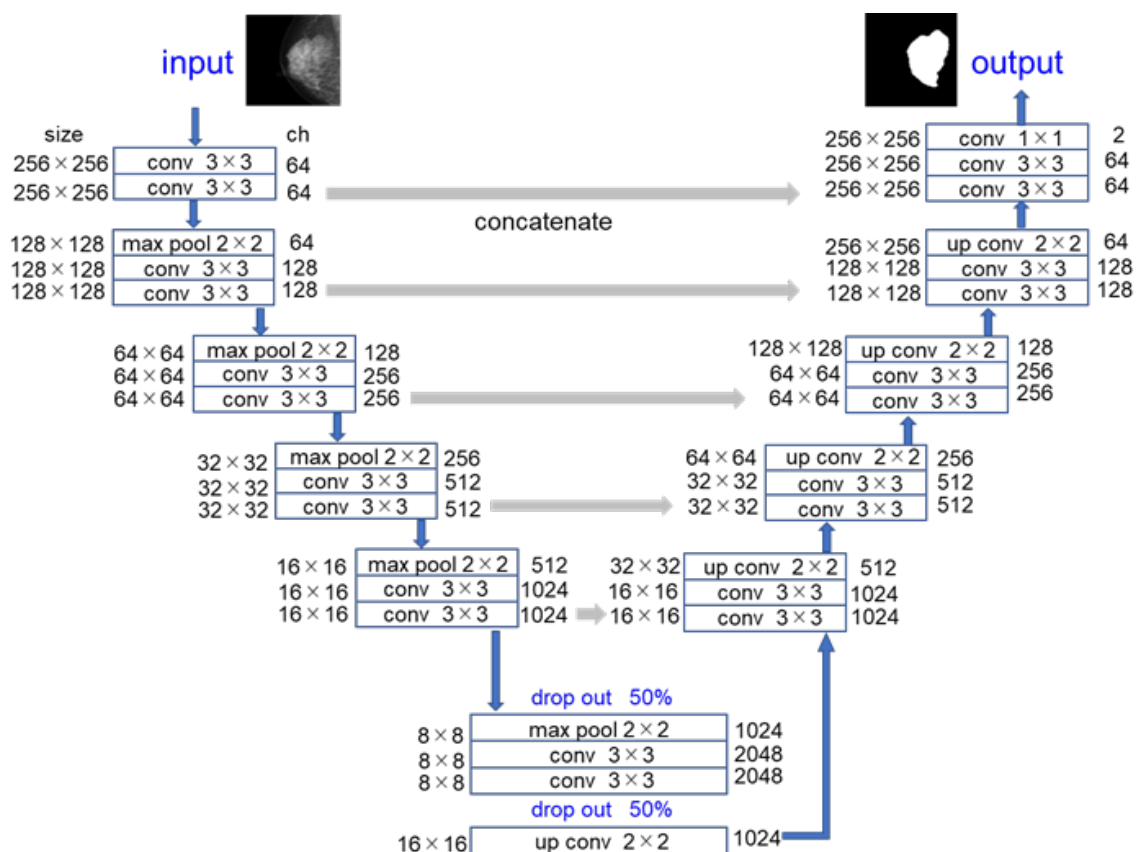


図 3.3 本研究で用いた U-net の構造。

U-net モデルを構築する前に、収集した学習画像に偏りが無いことを確かめるため、各 fold の学習画像を 400 枚、検証画像を 100 枚とした 5 分割交差検証を実施した。評価指標は Dice 係数³²⁾、各 fold 間の変動の許容基準は標準偏差 (SD) < 0.05 とした。5 分割交差検証の後、学習用データ 500 画像を用いて乳腺領域抽出モデルを構築し、独立して収集したテスト用データ 203 画像に対してモデルを適用した。

3.2.5 Dice 係数による U-net モデルの評価

表 3.1 に 5 分割交差検証の結果を示す。各 fold の平均 Dice 係数は、dense region では 0.857~0.893、diffuse region では 0.920~0.939 の範囲であった。SD はそれぞれ 0.013 および 0.007 といずれも許容基準とした 0.05 を下回っており、収集したマンモグラムに偏りが無いことがわかる。

表 3.1 5 分割交差検証における各 fold の Dice 係数

Region	1	2	3	4	5	Mean	SD
Dense	0.893	0.857	0.879	0.892	0.877	0.880	0.013
Diffuse	0.929	0.936	0.926	0.920	0.939	0.930	0.007

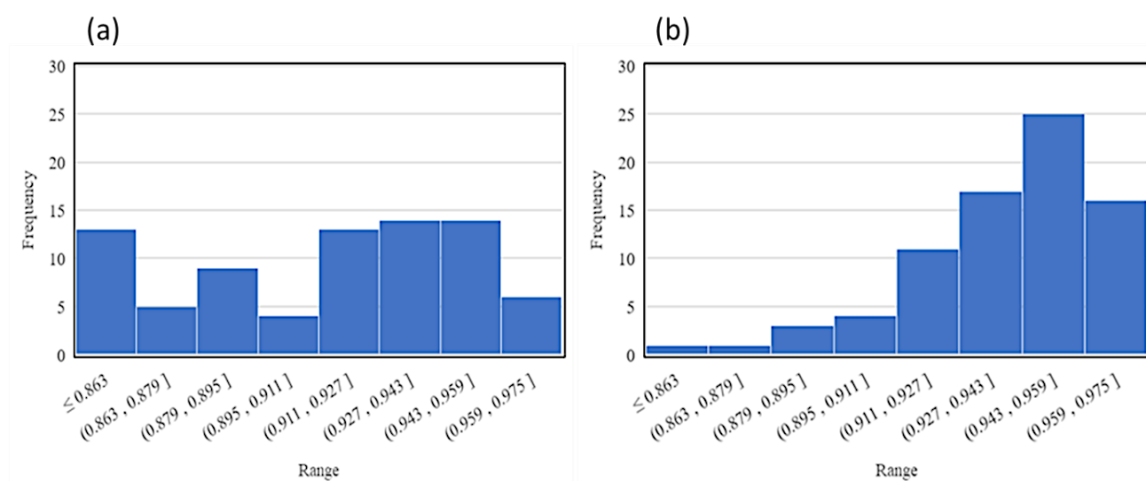


図 3.4 U-net が抽出した領域と GT の Dice 係数のヒストグラム. (a) dense region, (b) diffuse region

図 3.4 に U-net が抽出した領域と GT の Dice 係数のヒストグラムを示す。(a)、(b)は各々 dense region、diffuse region の結果である。Dice 係数が 0.90 以上となった割合は、dense region で全画像の 62.7%、diffuse region で 91.2%であり、良好な抽出結果であった。

3.3 乳腺密度の算出

乳腺密度を計算するために、画像処理後のテスト画像 203 枚に対して U-net が自動抽出した dense region と diffuse region および 3 名の認定診療放射線技師が作成した GT をフリー画像処理ソフトウェア ImageJ (National Institutes of Health, Maryland, USA) を用いて同じ座標軸にある生データマンモグラム上に転写した。これは、乳腺密度の計算には画素値と X 線強度の線形関係が要求されるためである。生データマンモグラムにおける dense region、diffuse region および GT の乳腺密度を式 (3.1) を使って計算した³³⁾。

$$T_{gla} = \frac{\ln\left(\frac{P_{gla}}{P_{adi}}\right)}{\mu_{adi} - \mu_{gla}} \quad (3.1)$$

ここで、 T_{gla} は乳腺組織の厚さ、 P_{gla} および P_{adi} は各々乳腺領域および脂肪領域の画素値、 μ_{gla} および μ_{adi} は各々乳腺組織および脂肪組織の実効線減弱係数である。 μ_{gla} および μ_{adi} は CdTe 検出器を使って測定した X 線エネルギースペクトルから式 (3.2) と式 (3.3) によって計算した。

$$\mu_{gla} = \frac{\int_0^{E_{max}} I(E)\mu_{gla}(E)dE}{\int_0^{E_{max}} I(E)dE} \quad (3.2)$$

$$\mu_{adi} = \frac{\int_0^{E_{max}} I(E)\mu_{adi}(E)dE}{\int_0^{E_{max}} I(E)dE} \quad (3.3)$$

ここで、 $I(E)$ はエネルギー E での X 線光子数であり、脂肪組織と乳腺組織の線減弱係数は第 2 章と同様 XCOM³⁴⁾ を使って算出した。脂肪組織の厚さは式 (3.4) により求めた。

$$T_{adi} = T_{bre} - T_{gla} \quad (3.4)$$

ここで、 T_{bre} は CBT で、マンモグラフィ装置が撮影時に表示する数値を用いた。

次に、式 (3.1) に代入する乳腺領域および脂肪領域の画素値の求め方について

て述べる。乳腺領域の画素値 P_{gla} は、生データに転写された U-net による自動抽出領域（dense region と diffuse region）および GT の平均画素値を用いた。一方、マンモグラム内で CBT が表示値に達しており、かつ、乳腺組織を含まない脂肪組織だけの領域を探すのは極めて困難であり、自動計測ソフトを提供するメーカーはどれもその求め方および得られた画素値を開示していない。脂肪領域の画素値は式 (3.1) を解くための基準となる値であり、この値に整合性がないことが、同じマンモグラムを解析しても各メーカー間で乳腺密度が顕著に異なる要因となっている。著者らは先行研究²⁹⁾において ImageJ の threshold 機能を利用し、マンモグラムの解剖学的知見に基づいた脂肪領域の決定法を考案したので以下に説明する。

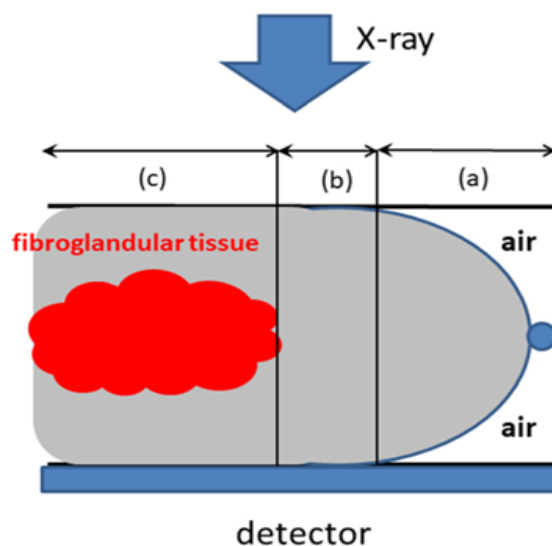


図 3.5 乳房の断面模式図.

図 3.5 は乳房の断面模式図である。領域 (a) は乳房の形状にともない CBT が表示値に達しておらず X 線透過方向に空気の領域が存在する。また、(c) の領域は脂肪組織と乳腺組織が混在する領域である。領域 (b) は CBT が表示値に達しており、かつ、乳腺組織が存在しない領域であり、このような領域が式 (3.1) の P_{adi} に相当する。著者は先行研究²⁹⁾で ImageJ の threshold 機能を使いこのような領域を探す手法を提唱した。ImageJ の threshold 値（設定画素値）を最大値（乳房が存在しない空気だけの領域に相当し、本研究で用いた装置では 7500 と決められている）に設定し、その後、幅をもたない単一の設定画素値を徐々に

低くしていくと、図 3.6 (a)-(c)の赤色部分のような設定画素値の分布を見ることが出来る。設定画素値が乳房内に入ると、まず(a)のような線状の領域が見られる。これは図 3.5 の(a)に相当する領域で、ここでは位置が内外方向に少しでもズレルと乳房の厚さが異なり、それに伴い画素値も異なるために単一の画素値の分布は線状で表されることになる。設定画素値が十分に低くなると図 3.6 (c)のように赤色部分が広がり、大胸筋の一部や乳腺組織の分布が見られるようになる。この過程において、図 3.5 (b)に示すような CBT が表示値に達しており（すなわち内外方向に移動しても一定の厚さとなっている）、かつ、乳腺組織が存在しない領域に設定画素値が定められた場合は、図 3.6 (b)のように U 領域を中心に広がりをもつ分布を見ることが出来る。この時の設定画素値が脂肪領域の画素値 P_{adi} である。

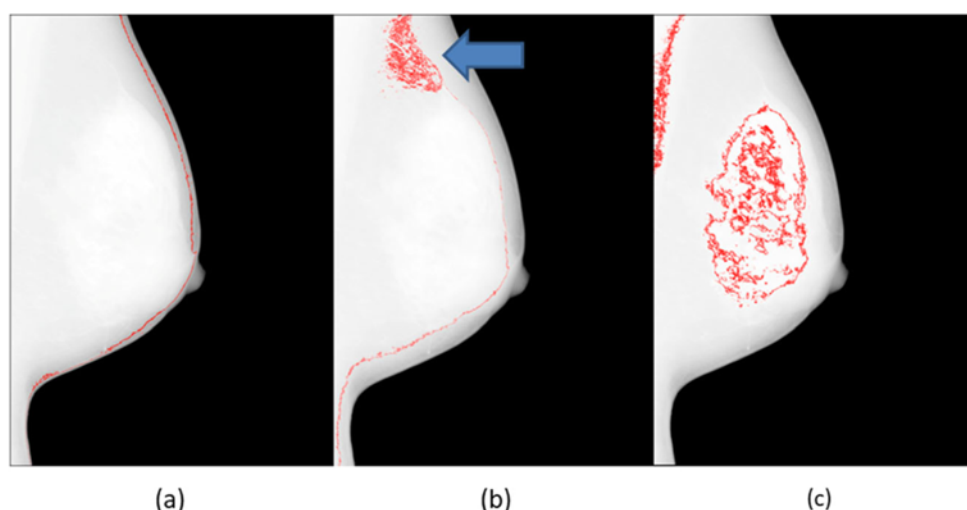


図 3.6 脂肪領域の画素値 P_{adi} を決定するための *threshold* 法の説明図。指定する単一の画素値 (*threshold* 値) を連続的に変化させながらその分布 (赤色部分) を目視で確かめ P_{adi} を決定する。(a)は CBT が表示値に達していない領域の分布, (b)は脂肪だけの領域を適切に表している分布, (c)は大胸筋や乳腺領域まで広がった分布を表している。

Dance ら³⁵⁾によると、乳腺密度は乳腺組織と脂肪組織の質量比として定義されている。したがって、式 (3.1) および式 (3.4) で計算した乳腺組織および脂肪組織の厚さに各々の物理密度を乗じて質量比を求めた。さらに Dance ら³⁵⁾は、圧迫された乳房は 5 mm の皮下脂肪層に囲まれており、乳腺密度はこの皮下脂肪層を除いた中央領域である *central region* について計算せねばならないと

している。しかし、日本人女性の乳房は欧米人と比べ比較的薄いため³⁶⁾、本研究では皮下脂肪層の厚さを 2.5 mm と想定して乳腺密度を算出した。

3.4 平均乳腺線量の算出

マンモグラフィの X 線被ばく線量は一般に式 (3.5) によって求まる平均乳腺線量 (mean glandular dose: MGD) を用いて表される¹²⁾。

$$D = K \cdot g \cdot c \cdot s \quad (3.5)$$

ここで、 D は平均乳腺線量、 K は乳房表面における空気カーマ、 g は CBT に対する補正係数、 c は乳腺密度に対する補正係数、 s はターゲット/フィルタに対する補正係数である。本研究では CdTe 検出器を備えた EMF123 X 線スペクトロメータ (EMF-JAPAN、兵庫、日本) を用いて K を測定した。従来、乳腺密度の正確な値を計測できなかったため、便宜的に乳腺密度 50% に対応する c ファクタ (= 1.0) が MGD の計算に使用されてきた。しかし本研究では、乳腺密度に対応する c ファクタを式 (3.5) に代入することで個々の正確な MGD (3.1 で述べた corrected MGD) を算出し、dense region、diffuse region および GT に対して適用した。 c ファクタの数値は参考文献 12) の表 6 の数値を CBT 1 mm 間隔、乳腺密度 1% 間隔で補間して用いた。

3.5 乳腺密度および平均乳腺線量による U-net モデルの評価

3.5.1 評価方法

構築した U-net モデルの評価として、乳腺密度および corrected MGD の両者について U-net が抽出した領域と GT 間の相関係数を求めるとともに、Bland-Altman (B-A) 解析³⁷⁾を使って、U-net の臨床適用の可能性を評価した。B-A 解析では乳腺密度および corrected MGD を指標として U-net が抽出した領域と GT の同等性または互換性を以下に示す 4 項目に基づきノンパラメトリックに評価した³⁸⁾。(1) 95% 一致限界 (95% LOA) が十分に狭い、(2) 95% LOA の上限と下限がゼロを挟む、(3) プロットに対する回帰直線が有意な傾きをもたない、(4) 回帰直線の傾きの 95% 信頼区間 (CI) がゼロを挟む。項目 (1) を満たす

基準として、U-net と GT の測定値の差が平均値の±20%以内である頻度が 75% 以上である場合を「十分に狭い」と定義した³⁹⁾。項目 (3) については $p < 0.05$ であれば統計的有意性があるとした。上記 4 項目がすべて満たされている場合、U-net が抽出した領域と GT は乳腺密度または corrected MGD について同等であると判定し、4 項目すべてが満たされていないが少なくとも項目 (1) が満たされている場合は互換性があると判定した。ここで、同等性とは U-net が抽出した領域と GT をそのまま置き換えることができることを意味し、互換性とは四則演算による補正を加えることで両者の入れ替えが可能であることを意味している。

3.5.2 評価結果

図 3.7 および図 3.8 に乳腺密度および corrected MGD に対する U-net が抽出した領域と GT 間の相関図を示す。乳腺密度に対する U-net が抽出した領域と GT の相関係数は、dense region が 0.994、diffuse region が 0.986 と良好であった。また、corrected MGD に対する相関係数は dense region が 0.997、diffuse region が 0.994 とこちらも良好であった。

図 3.9 および図 3.10 は各々乳腺密度および corrected MGD に対する B-A プロットを示している。また、表 3.2 に dense region、diffuse region の 95% LOA およびプロットに対する回帰直線の解析結果を示す。図 3.9 に示すように、乳腺密度において U-net と GT の測定値の差が平均値の±20%以内に入る頻度は、dense region で 98%、diffuse region で 75.4%であった。また図 3.10 に示すように、corrected MGD においては dense region および diffuse region の両方で、測定値の差はすべて (100%) 平均値の±20%以内に入った。これらの結果は、乳腺密度および corrected MGD の両者とも U-net が抽出した領域と GT 間の同等性または互換性を判定するための項目(1) を満たしていることを示している。このように、乳腺密度については U-net が抽出した領域と GT 間の同等性が dense region および diffuse region の両方について証明された (表 3.2、図 3.9)。しかし、corrected MGD については回帰直線に有意な傾きが認められたため、dense region および diffuse region の両方について U-net が抽出した領域と GT 間の互換性のみが証明された (表 3.2、図 3.10)。

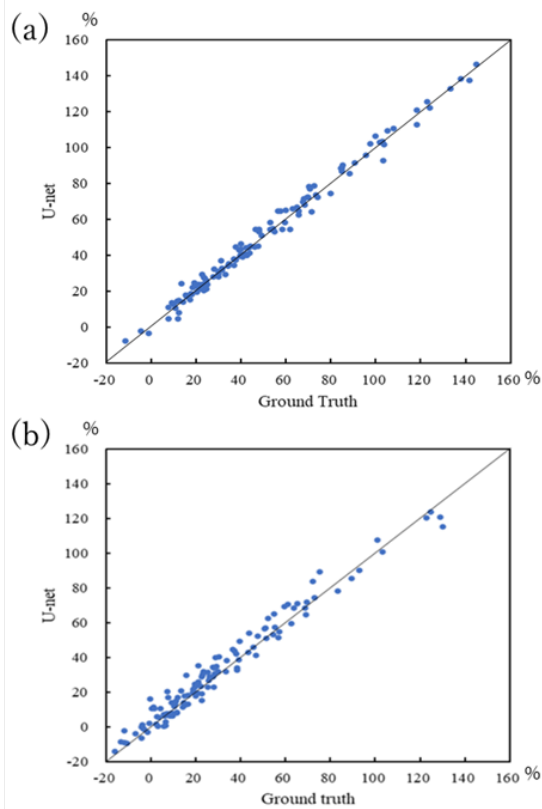


図 3.7 Dense region (a)と diffuse region (b)における乳腺密度の U-net が抽出した領域と GT 間の相関図. 相関係数は dense region: 0.994, diffuse region: 0.986.

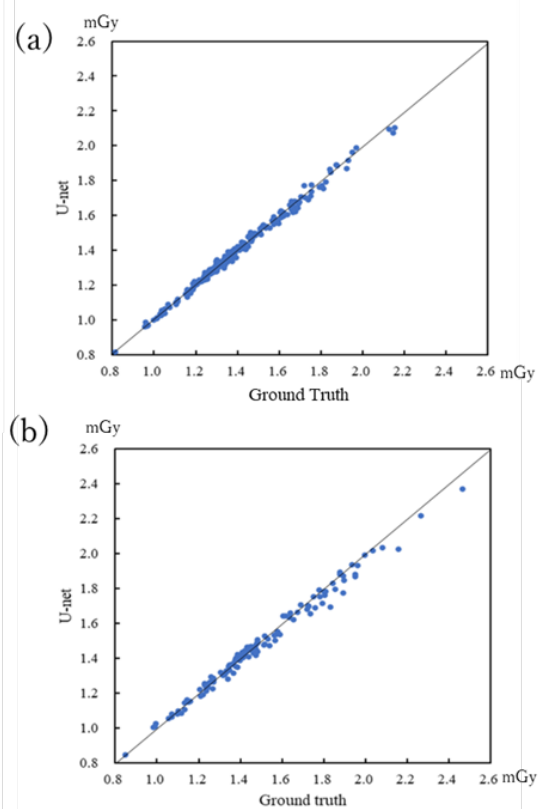


図 3.8 Dense region (a)と diffuse region (b)における corrected MGD の U-net が抽出した領域と GT 間の相関図. 相関係数は dense region: 0.997, diffuse region: 0.994.

表 3.2 乳腺密度および corrected MGD に対する U-net が抽出した領域と GT 間の B-A 解析結果.

Parameter		Breast Density		Corrected Mean Glandular Dose	
		Dense	Diffuse	Dense	Diffuse
95% Limits of Agreement	Lower	-6.7	-8.7	-0.04	-0.08
	Upper	7.3	11.2	0.04	0.06
	Width	14.0	19.9	0.08	0.14
Slope of Regression Line		0.0124	-0.0022	-0.0299*	-0.0443*
95% CI of the Slope	Lower	-0.0023	-0.0255	-0.0407	-0.0592
	Upper	0.0272	0.0211	-0.0190	-0.0295

* 回帰直線に有意な傾きあり ($P < 0.05$)

CI: 信頼区間

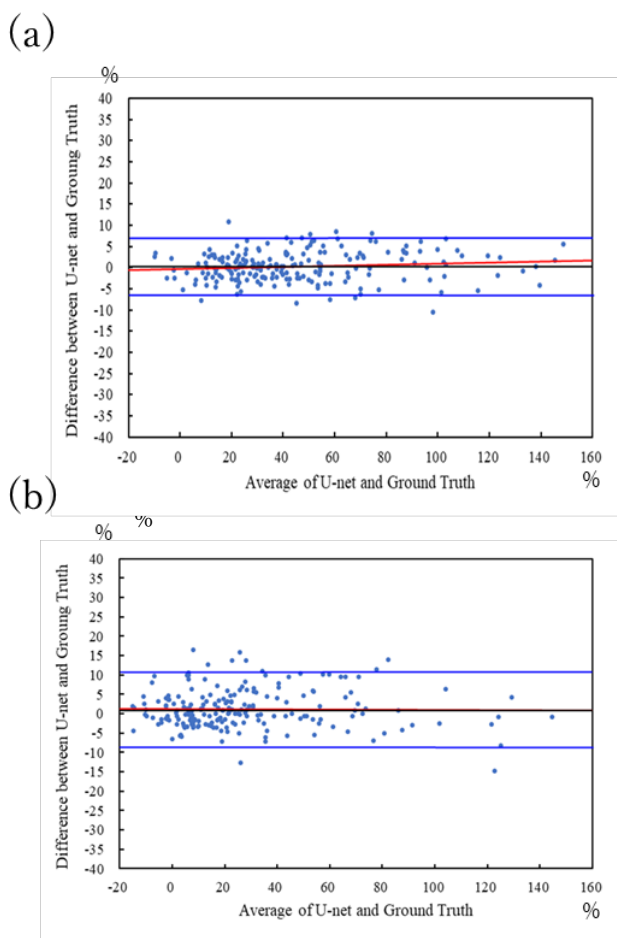


図 3.9 Dense region (a) と diffuse region (b) における乳腺密度の B-A プロット. 黒線は U-net が抽出した領域と GT の差の平均, 青線は 95%LOA, 赤線はプロットに対する回帰直線. (a), (b)とも系統誤差は見られない.

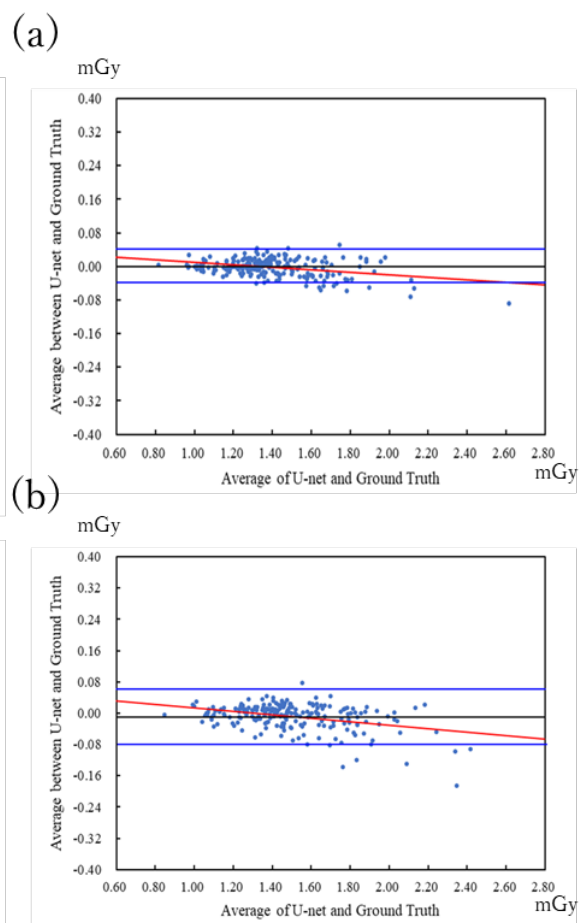


図 3.10 Dense region (a) と diffuse region (b) における corrected MGD の B-A プロット. 黒線は U-net が抽出した領域と GT の差の平均, 青線は 95%, 赤線はプロットに対する回帰直線. (a), (b)とも回帰直線に有意な傾きが見られる. ($p < 0.05$)

3.5.3 乳腺領域の抽出例

図 3.11–図 3.14 に U-net が抽出した乳腺領域と GT の比較例を示す。図 3.11 と図 3.13 は Dice 係数を指標として用いた場合の代表的な抽出例で、(a) Dice 係数が最大となった例、(b) 平均値に最も近かった例、(c) 最小となった例を示している。図 3.12 と図 3.14 は乳腺密度を指標として用いた場合の代表的な抽出例で、(a) U-net が抽出した領域と GT 間で乳腺密度の差が最も小さかった例、(b) 平均的であった例、(c) 最も大きかった例を示している。

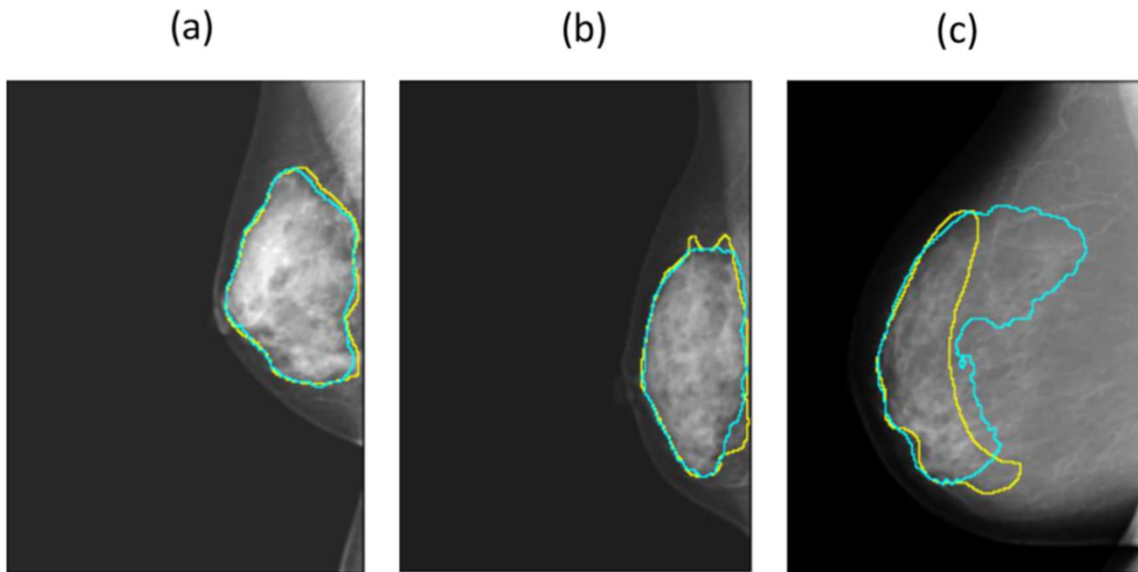


図 3.11 Dense region に対する U-net が抽出した領域と GT 間の Dice 係数の比較例. (a)~(c)は各々Dice 係数が最大値 (0.975), 平均値 (0.895), 最小値 (0.682) の一例. 青線は U-net が抽出した乳腺領域, 黄色は GT を表している.

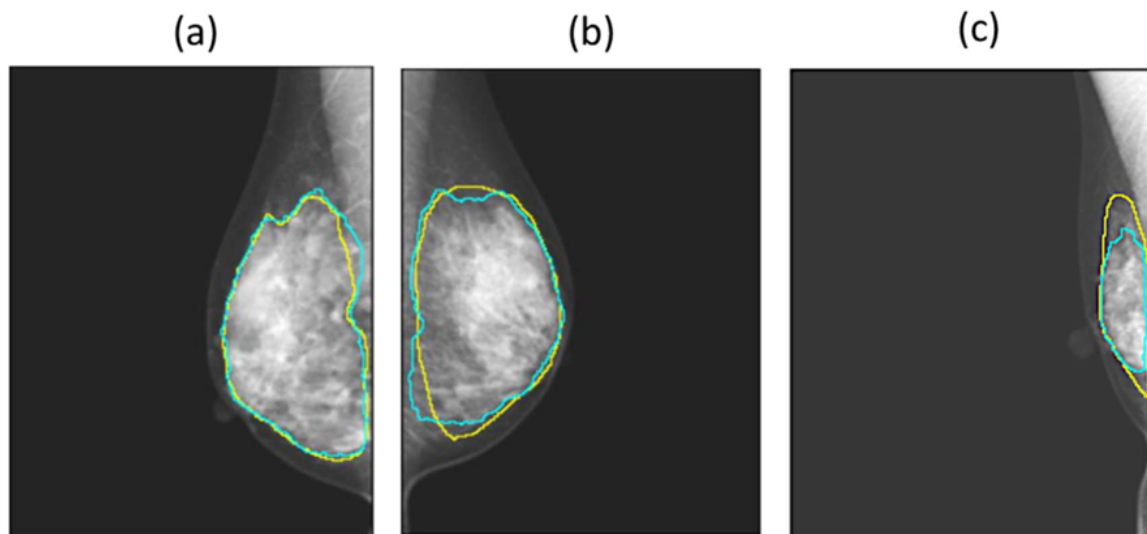


図 3.12 Dense region に対する U-net が抽出した領域と GT の乳腺密度の比較例. (a)~(c)は各々乳腺密度の差が最小 (0.1%), 中程度 (2.8%), 最大 (10.8%) となった一例. 青線は U-net が抽出した乳腺領域, 黄色は GT を表している.

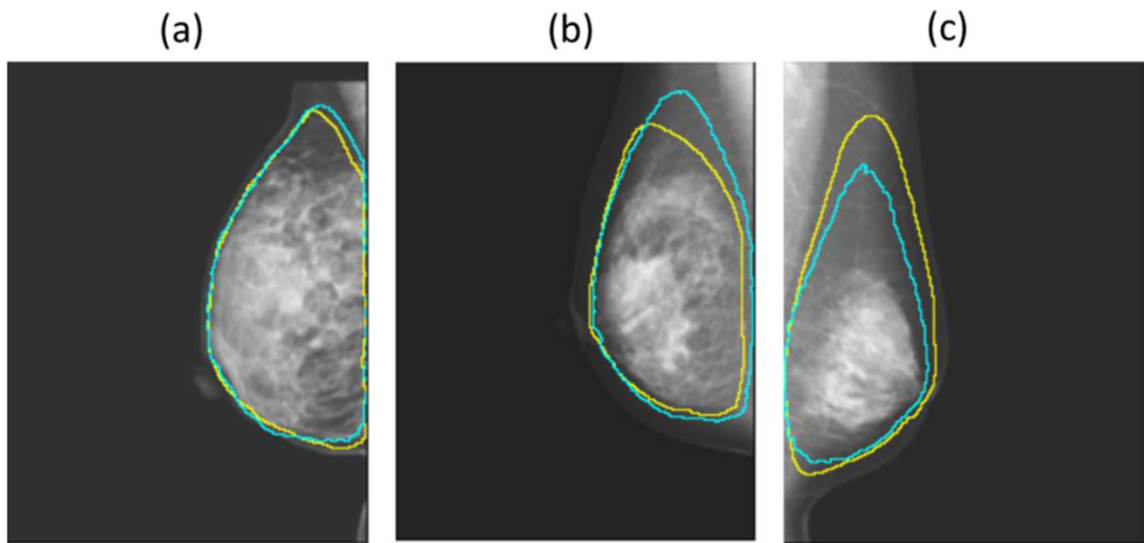


図 3.13 Diffuse region に対する U-net が抽出した領域と GT 間の Dice 係数の比較例. (a)~(c)は各々Dice 係数が最大値 (0.975), 平均値 (0.939), 最小値 (0.863) の一例. 青線は U-net が抽出した乳腺領域, 黄色は GT を表している.

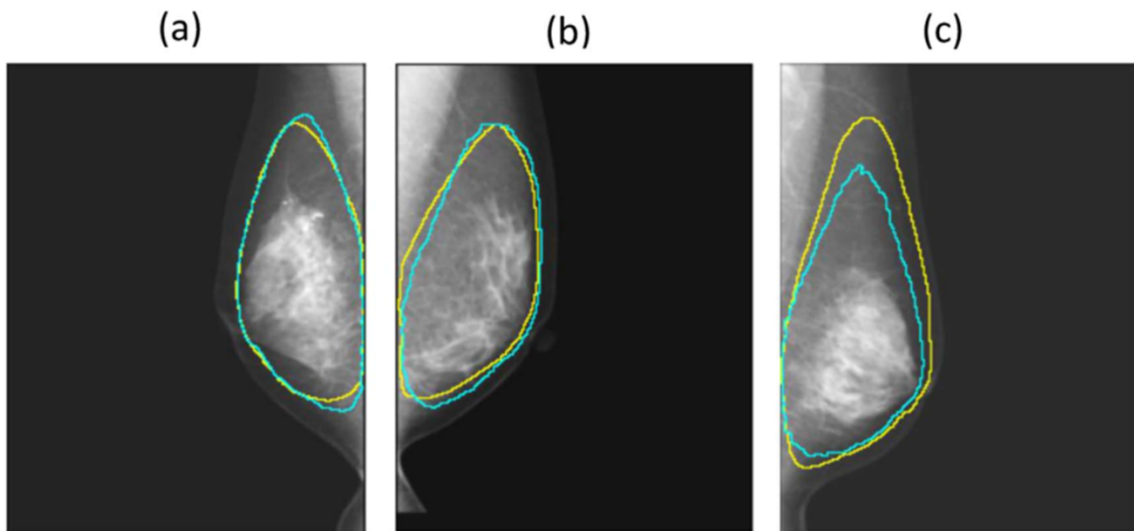


図 3.14 Diffuse region に対する U-net が抽出した領域と GT の乳腺密度の比較例. (a)~(c)は各々乳腺密度の差が最小 (0.0%), 中程度 (3.6%), 最大 (16.4%) となった一例. 青線は U-net が抽出した乳腺領域, 黄色は GT を表している.

3.6 考察

将来の個別化スクリーニングマンモグラフィプログラムおよび個人への乳腺密度と被ばく線量の通知の信頼性を向上させるために、乳腺エキスパートによる目視判断に代えて画素値を用いた安定的手法が広まりつつあるが、この手法

に必要であるマンモグラム中の乳腺領域自動抽出へディープラーニング技術を応用できる可能性を検討した。近年、さまざまな研究でディープラーニング技術が利用され、臨床マンモグラフィや乳房 MRI における領域抽出に優れた精度と再現性を提供しているが、これまでの研究は面積ベースによる 2 次元での領域抽出^{26-28, 40)}や Dice 係数を用いた精度評価⁴¹⁻⁴³⁾に焦点を当てていた。しかし、本研究では形状だけでなく、抽出された乳腺領域に対して算出した乳腺密度や、乳腺密度で補正した平均乳腺線量を GT とともに評価することが重要であると考えている。ここでは、乳腺密度および被ばく線量の観点から U-net による乳腺領域自動抽出の臨床適用の可能性を考察する。

図 3.4 に示すように、U-net が抽出した領域と GT 間の平均 Dice 係数は dense region よりも diffuse region の方が高かった。これは図 3.11-図 3.14 からわかるように diffuse region の単純な形状によって説明することができる。しかし、図 3.7 と図 3.8 に示すように U-net が抽出した領域と GT 間の乳腺密度および corrected MGD の相関係数は diffuse region よりも dense region の方が高かった。この理由は、diffuse region は dense region と比べて皮膚面に近い領域を多く含んでおり、そのような領域には非圧迫領域、すなわち適切な乳腺密度が得られていない領域が存在するためであると考えられる。非圧迫領域は乳腺密度の過小評価および平均乳腺線量の過大評価につながり、乳腺密度計測の信頼性に影響を及ぼす。このような現象は、2 次元マンモグラムでは非圧迫領域を視覚的に観察できないため Dice 係数では評価することができない。

次に、B-A 解析の結果に基づき U-net を、従来の目視による乳腺領域抽出に代えて臨床適用できる可能性について述べる。Dice 係数は diffuse region よりも dense region の方が低かったが（図 3.4）、このような Dice 係数の差は乳腺密度の計算に実質的な影響を与えていない。なぜなら、図 3.9 および表 3.2 に示すように、乳腺密度に対する U-net が抽出した領域と GT 間の B-A 解析結果は dense region および diffuse region と同等性または互換性を保証するための 4 項目すべてを満たしているからである。

一方、表 3.2 および図 3.10 からわかるように、corrected MGD に対する B-A 解析では、回帰直線に統計的に有意な傾き、すなわち比例誤差が認められた。しかし、同等性または互換性を保証するための 4 項目のうち項目 1 を含む 2 項

目を満たしているため、corrected MGD に対しては U-net が抽出した領域と GT 間には互換性があると結論できる。Corrected MGD における比例誤差は dense region よりも diffuse region の方が顕著であった。しかし、95% LOA は図 3.10 の X 軸の範囲に比べて十分に狭く、dense region ではわずか 4.4%、diffuse region でも 7.8%であることから、この場合は除算によって U-net が抽出した領域と GT 間に同等性を持たせることができると考える。B-A 解析において乳腺密度に対しては同等性が認められた一方で、corrected MGD に対しては互換性に限定された理由として、MGD は乳腺密度だけでなくマンモグラフィに使用された撮影 X 線量にも依存することが一因であると考えられる。撮影 X 線量が増加すると corrected MGD も増加し、その結果として U-net が抽出した領域と GT 間の corrected MGD の差も大きくなる。したがって、図 3.10 の回帰直線に統計的に有意な傾きが存在するのは U-net の精度だけが原因ではないと考える。

Dice 係数または乳腺密度のどちらかだけを用いた乳腺領域抽出精度の評価では不十分である。たとえば図 3.12(a) では、U-net によって自動抽出された乳腺領域の Dice 係数は 0.893 であり平均 (0.897) をわずかに下回ったが、同じ領域における乳腺密度は U-net: 95.8%、GT: 95.7%と非常によく一致していた。その理由は、反対方向へのミスアライメントが効果を相殺するためである。すなわち、乳腺領域全体を見た場合、ある部分では U-net の方が乳腺密度の高い領域を含み、別の部分では逆に乳腺密度の低い領域を含んでいることが原因である。このような現象は、Dice 係数で評価することの重要性を示している。一方、図 3.13(a)に示すように乳腺領域の形状は U-net が抽出した領域と GT 間でよく一致しているが (Dice 係数: 0.975)、乳腺密度はあまり一致しているとは言えないケース (U-net: 85.4%、GT: 89.6%) も散見された。その理由は非圧迫領域の影響によるものであると考える。このような現象は、乳腺密度による評価の重要性を示している。これらの結果から、U-net の精度評価には Dice 係数および乳腺密度、さらには乳腺密度に依存する corrected MGD を用いることが重要であると言える。

本研究にはいくつかのリミテーションがある。第一に、本研究では非圧迫領域について考慮していない。マンモグラフィでは皮膚面近傍に非圧迫領域が生じ、この領域の画素値は圧迫板と乳房の間に存在する空気の影響で実際よりも

高くなる。そのため、乳腺領域に非圧迫領域が含まれる場合は乳腺密度が過小評価されることに注意が必要である²⁹⁾。第二に、3.3 で説明したように本研究では乳房内皮下脂肪層の厚さを 2.5 mm と仮定した。乳腺密度や MGD を計算するための既存の乳房モデルは乳房の厚さとは無関係に 5 mm を仮定している¹²⁾。著者はこのような一定値は現実的ではなく、皮下脂肪層は CBT に関連し欧米人女性と比較すると日本人女性の方が薄いと考えている。西出らは、福井県立病院でマンモグラフィを受けた 7566 人の日本人女性の CBT の平均が 31.7 mm であり、全体の 44.8% が 30 mm 未満であったことを報告している³⁶⁾。しかし、現在のところ皮下脂肪層の厚さを正確に測定する方法は著者の知る限り存在しない。計測対象乳房の皮下脂肪層厚が 2.5 mm よりも薄い場合は乳腺密度を過大評価することになり、図 3.7 および図 3.9 において全体の約 6% の症例で乳腺密度が 100% を超えて計算された一因となっていると考える。第三に、多くの研究者がディープラーニング技術の実践にはできるだけ多くの GT が必要であると指摘している⁴⁴⁻⁴⁷⁾。Yi らは、ディープラーニングを用いた乳腺密度分類は乳房の左右識別よりもはるかに大きなデータセットが必要であると述べている⁴⁷⁾。乳腺組織の分布パターンは個人毎に異なるため、より多くのパターンを含む大規模なデータセットを使用した抽出モデルの構築が望まれる。

本章では、画像分割のための deep convolutional neural network (DCNN) として最もシンプルな構造で汎用性の高い U-net を用いて、U-net が抽出した領域がエキスパートにより作成された乳腺領域と同等性または互換性があることを確かめた。一方で新しい人工知能技術が日進月歩で開発されており⁴⁸⁾、最新の DCNN 技術を組み込むことによって VBDM の精度をさらに向上させることが期待される。

3.7 結言

本章では、マンモグラムから乳腺領域を自動抽出するためにディープラーニング技術の一種である U-net を導入し、従来の Dice 係数に加えて乳腺密度および平均乳腺線量を指標として U-net の有用性を検証した。そして B-A 解析により、dense region および diffuse region の両方について U-net が抽出した領域と ground truth の間には乳腺密度の同等性および corrected MGD の互換性が成立す

ることを示した。本章の成果は、U-net を用いて算出した乳腺密度や平均乳腺線量を臨床の場へ適用することが可能であることを示唆している。人工知能技術の導入は、信頼性の高い個別化マンモグラフィ検診プログラムの実現に貢献するものである。

参考文献

- 1) Global Burden of Disease Cancer Collaboration et al: Global, regional, and national cancer incidence, mortality, years of life lost, years lived with disability, and disability-adjusted life-years for 29 cancer groups, 1990 to 2017: a systematic analysis for the global burden of disease study, *JAMA Oncol.* 5 1749–1768(2019)
- 2) D. L. Monticciolo, M. S. Newell, L. Moy, B. Niell, B. Monsees, E. A. Sickles: Breast cancer screening in women at higher-than-average risk: recommendations from the ACR, *J. Am. Coll. Radiol.*, 15 408–414(2018)
- 3) S. Y. H. Chiu, S. Duffy, A. M. F. Yen, L. Tabár, R. A. Smith, H. Chen: Effect of baseline breast density on breast cancer incidence, stage, mortality, and screening parameters: 25-year follow-up of a Swedish mammographic screening, *Cancer Epidemiol. Biomarkers Prev.* 19 1219–1228(2010)
- 4) J. M. Bae and E. H. Kim: Breast density and risk of breast cancer in Asian women: a meta-analysis of observational studies, *J. Prev. Med. Public Health* 49 367–375(2016)
- 5) A. M. Jeffers, W. Sieh, J. A. Lipson, J. H. Rothstein, V. McGuire, A. S. Whittemore, D. L. Rubin: Breast cancer risk and mammographic density assessed with semiautomated and fully automated methods and BI-RADS, *Radiology* 282 348–355(2017)
- 6) L. Soguel, F. Durocher, A. Tchernof, C. Diorio: Adiposity, breast density, and breast cancer risk: epidemiological and biological considerations, *Eur. J. Cancer Prev.* 26 511–520(2017)
- 7) A. R. Brentnall, J. Cuzick, D. S. M. Buist, E. J. Aiello Bowles: Long-term

- accuracy of breast cancer risk assessment combining classic risk factors and breast density. *JAMA Oncol.* 4 e180174(2018)
- 8) K. E. McLean, J. Stone: Role of breast density measurement in screening for breast cancer, *Climacteric* 21 214–220(2018)
 - 9) M. Rebolj, O. Blyuss, K. S. Chia, S. W. Duffy: Long-term excess risk of breast cancer after a single breast density measurement, *Eur. J. Cancer* 117 41–47(2019)
 - 10) A. Gastounioti A et al: Evaluation of LIBRA software for fully automated mammographic density assessment in breast cancer risk prediction, *Radiology* 296 24–31(2020)
 - 11) A. B. Hollingsworth: Redefining the sensitivity of screening mammography: a review, *Am. J. Surg.* 218 411–418(2019)
 - 12) D. R. Dance, C. L. Skinner, K. C. Young, J. R. Beckett, C. J. Kotre: Additional factors for the estimation of mean glandular breast dose using the UK mammography dosimetry protocol, *Phys. Med. Biol.* 45 3225–3240(2000)
 - 13) N. Houssami, C. I. Lee: The impact of legislation mandating breast density notification - review of the evidence, *Breast* 42 102–112(2018)
 - 14) J. S. Haas, C. S. Giess, K. A. Harris, J. Ansolabehere, C. P. Kaplan: Randomized trial of personalized breast density and breast cancer risk notification, *J. Gen. Intern. Med.* 34 591–597(2019)
 - 15) M. Manning, T. L. Albrecht, S. O'Neill, K. Purrington: Between-race differences in supplemental breast cancer screening before and after breast density notification law. *J. Am. Coll. Radiol.* 16 797–803(2019)
 - 16) W. V. Ingman, B. Richards, J. M. Street, D. Carter, M. Rickard, J. Stone, P. Dasari: 2020 Breast density notification: an Australian perspective, *J. Clin. Med.* 9 681(2020)
 - 17) K. R. Brandt, et al: Comparison of clinical and automated breast density measurements: implications for risk prediction and supplemental screening, *Radiology* 279 710–719(2016)
 - 18) K. Rahbar, A. Gubern-Merida, J. T. Patrie, J. A. Harvey: Automated volumetric

- mammographic breast density measurements may underestimate percent breast density for high-density breasts, *Acad. Radiol.* 24 1561–1569(2017)
- 19) B. H. Østerås, A. C. T. Martinsen, S. H. B. Brandal, K. N. Chaudhry, E. Eben, U. Haakenaasen, R. S. Falk, P. Skaane: Classification of fatty and dense breast parenchyma: comparison of automatic volumetric density measurement and radiologists' classification and their inter-observer variation, *Acta Radiol.* 57 1178–1185(2016)
 - 20) R. Highnam, S. M. Brady, M. J. Yaffe, N. Karssemeijer, J. Harvey: Robust breast composition measurement - Volpara™ Digital Mammography. *IWDM 2010 (Lecture Notes in Computer Science, vol 6136)* eds Martí J, Oliver A, Freixenet J and Martí R (Berlin, Heidelberg: Springer) pp 342–349(2010)
 - 21) A. Gubern-Mérida, M. Kallenberg, B. Platel, R. M. Mann, R. Martí, N. Karssemeijer: Volumetric breast density estimation from full-field digital mammograms: a validation study. *PLoS ONE* 9 e85952(2014)
 - 22) S. Destounis, L. Johnston, R. Highnam, A. Arieno, R. Morgan, A. Chan: Using volumetric breast density to quantify the potential masking risk of mammographic density, *Am. J. Roentgenol.* 208 222–227(2017)
 - 23) A. Fieselmann, D. Förnvik, H. Förnvik, K. Lång, H. Sartor, S. Zackrisson, S. Kappler, L. Ritschl, T. Mertelmeier: Volumetric breast density measurement for personalized screening: accuracy, reproducibility, consistency, and agreement with visual assessment. *J. Med. Imaging (Bellingham)* 6 031406(2019)
 - 24) 乳房構成判定アトラス. 厚生労働行政推進調査事業費補助金がん対策推進総合研究事業「乳癌検診の適切な情報提供に関する研究」研究代表者 笠原善郎. 令和 2 年度総括・分担研究報告書. 2021 年 5 月.
<https://brestcs.org/study/achievement/page1.html>
 - 25) M. Yamamuro, Y. Asai, N. Hashimoto, N. Yasuda, K. Sakaguchi, T. Konishi, K. Yamada, Y. Ozaki, K. Ishii, Y. Lee: Deep learning-based segmentation of mammary gland tissue region in digital mammograms of scattered mammary glands and fatty breasts, *Proc. SPIE* 11513 115131V(2020)
 - 26) A. Hamidinekoo, E. Denton, A. Rampun, K. Honnor, R. Zwiggelaar: Deep

- learning in mammography and breast histology, an overview and future trends. *Med. Image Anal.* 47 45–67(2018)
- 27) J. Lee and R. M. Nishikawa: Automated mammographic breast density estimation using a fully convolutional network, *Med. Phys.* 45 1178–1190(2018)
 - 28) L. Li, J. Wei, H. P. Chan, M. A. Helvie, M. A. Roubidoux, Y. Lu, C. Zhou, L. M. Hadjiiski, R. K. Samala: 2018 Computer-aided assessment of breast density: comparison of supervised deep learning and feature-based statistical learning, *Phys. Med. Biol.* 63 025005(2018)
 - 29) M. Yamamuro, Y. Asai, K. Yamada, Y. Ozaki, M. Matsumoto, T. Murakami: Prediction of glandularity and breast radiation dose from mammography results in Japanese women. *Med. Biol. Eng. Comput.* 57 289–298(2019)
 - 30) C. D’Orsi, E. A. Sickles, E. B. Mendelson, E. A. Morris: *ACR BI-RADS Atlas: Breast Imaging Reporting and Data System* 5th ed, 2013 (Reston, VA: American College of Radiology)
 - 31) M. Yamamuro, Y. Asai, H. Hashimoto, N. Yasuda, Y. Kondo, et al: How to select training data to segment mammary gland tissue region using a deep-learning approach for reliable individualized screening mammography. *Proc. SPIE* 11597 115972V(2021)
 - 32) L. R. Dice: Measures of the amount of ecologic association between species, *Ecology* 26 297–302(1945)
 - 33) S. van Engeland, P. R. Snoeren, H. Huisman, C. Boetes, N. Karssemeijer: Volumetric breast density estimation from full-field digital mammograms. *IEEE Trans. Med. Imaging* 25 273–282(2006)
 - 34) M. J. Berger, J. Hubbell, S. M. Seltzer, J. Chang, J. S. Coursey, R. Sukumar, D. S. Zucker, K. Olsen: *XCOM: Photon Cross Sections Database*, NIST Standard Reference Database 8, (XGAM), 2010, NIST, PML, Radiation Physics Division. Updated November 2010. <https://dx.doi.org/10.18434/T48G6X>
 - 35) D. R. Dance, R. A. Hunt, P. R. Bakic, A. D. A. Maidment, M. Sandborg, G. Ullman, G. Alm Carlsson: Breast dosimetry using high-resolution voxel phantoms. *Radiat Prot Dosimetry* 114 359–363(2005)

- 36) H. Nishide, K. Ohta, Cui, K. Murata, Y. Kodera: Exposure conditions according to breast thickness and glandularity in Japanese women. *Breast Imaging. IWDM 2016 (Lecture Notes in Computer Science, vol 9699)* eds Tingberg A, Lång K and Timberg P (Cham.: Springer) pp 408–414(2016)
- 37) D. G. Altman and J. M. Bland: Measurement in medicine: the analysis of method comparison studies, *Statistician* 32 307–317(1983)
- 38) P. L. Myles and J. Cui: Using the Bland-Altman method to measure agreement with repeated measures, *Br J Anaest* 99 309–311(2007)
- 39) L. A. Critchley and J. A. Critchley: A meta-analysis of studies using bias and precision statistics to compare cardiac output measurement techniques, *J. Clin. Monit. Comput.* 15 85–91(1999)
- 40) A. Ciritsis, C. Rossi, I. Vittoria De Martini, M. Eberhard, M. Marcon, A. S. Becker, N. Berger, A. Boss: Determination of mammographic breast density using a deep convolutional neural network, *Br. J. Radiol.* 92 20180691(2019)
- 41) R. Ha, P. Chang, E. Mema, S. Mutasa, J. Karcich, R. T. Wynn, M. Z. Liu, S. Jambawalikar: Fully automated convolutional neural network method for quantification of breast MRI fibroglandular tissue and background parenchymal enhancement. *J. Digit. Imaging* 32 141–147(2019)
- 42) H. Fashandi, G. Kuling, Y. Lu, H. Wu, A. L. Martel: An investigation of the effect of fat suppression and dimensionality on the accuracy of breast MRI segmentation using U-nets. *Med. Phys.* 46 1230–1244(2019)
- 43) H. Jiao, X. Jiang, Z. Pang, X. Lin, Y. Huang, L. Li: Deep convolutional neural networks-based automatic breast segmentation and mass detection in DCE-MRI, *Comput. Math. Methods Med.* 2020 2413706(2020)
- 44) R. K. Samala, H. P. Chan, L. Hadjiiski, M. A. Helvie, J. Wei, K. Cha: Mass detection in digital breast tomosynthesis: deep convolutional neural network with transfer learning from mammography, *Med. Phys.* 43 6654(2016)
- 45) A. Rajkomar, S. Lingam, A. G. Taylor, M. Blum, J. Mongan: High-throughput classification of radiographs using deep convolutional neural networks. *J. Digit. Imaging* 30 95–101(2017)

- 46) P. H. Yi, T. K. Kim, J. Wei, J. Shin, F. K. Hui, H. I. Sair, C. D. Hager, and J. Fritz: Automated semantic labeling of pediatric musculoskeletal radiographs using deep learning. *Pediatr. Radiol.* 49 1066–1070(2019)
- 47) P. H. Yi, A. Lin, J. Wei, A. C. Yu, H. I. Sair, F. K. Hui, G. D. Hager, and S. C. Harvey: Deep-learning-based semantic labeling for 2D mammography and comparison of complexity for machine learning tasks. *J. Digit. Imaging* 32 565–570(2019)
- 48) S. Minaee, Y. Y. Boykov, F. Porikli, A. J. Plaza, N. Kehtarnavaz, D. Terzopoulos: Image segmentation using deep learning: a survey. *IEEE Trans. Pattern Anal. Mach. Intell.* (Early Access) doi: 10.1109/TPAMI.2021.3059968(2021)

第 4 章 マンモグラフィの X 線露光条件を用いた乳腺密度の推定

4.1 緒言

乳癌の死亡率は依然としてさまざまな国で増加している¹⁻⁴⁾。乳腺密度は、マンモグラフィにおける乳癌罹患率や病変検出率に影響を及ぼす重要な因子として注目されており、それらの研究は着々と進められている⁵⁻⁷⁾。一方、患者の乳腺密度の変化を調べることで、乳腺密度が乳癌発症の予測バイオマーカーやがん治療後の予後バイオマーカーとして使用できる可能性についてもこれまでにいくつかの研究が行われている⁸⁻¹³⁾。しかし、これらの研究は乳腺密度の時系列分析を必要とするため、その研究は着実に進歩しているとは言えない。

現在、個人の乳腺密度を求める方法には、大きく分けて次の 3 種類がある。

①計測；画像処理を施していない生データマンモグラムを使って、その乳腺領域内の画素値に基づいて乳腺密度を算出する手法¹⁴⁻¹⁶⁾。乳腺領域を高精度に抽出するためにディープラーニング技術が導入されるなど、近年は計測精度が向上しつつある¹⁷⁻¹⁹⁾。この手法は計測に生データマンモグラムを用いることで X 線強度との線形性を担保しているが、生データは容量が多いうえ画像診断に有用でないため各施設で長期間保存されることはない。このため過去に遡った計測ができない。

②予測；被検者の身体特性や生育歴情報などの個々の要因に基づいて乳腺密度を予測する手法。この方法の最大の利点は、マンモグラフィを撮影しないため X 線被ばくが生じないことであるが、一方、算出精度は低くなる。著者ら行った先行研究¹⁵⁾では、被検者の年齢・CBT・身長・体重・出産回数・閉経の有無を説明変数とした重回帰分析による予測精度は約 62%であった。現在のところ、この研究以外に外的因子を用いて乳腺密度を予測する研究は行われていない。

③推定；マンモグラムを用いず、マンモグラフィで使用された管電圧や mAs 値（管電流露光時間積）といった X 線露光データに基づき乳腺密度を推定する手法。Beckett らは、厚さと乳腺密度が異なる多くのファントムを X 線撮影し、ファントム撮影に要した mAs 値とマンモグラフィ検診を受けた被検者の mAs 値を比較するという実験的手法により乳腺密度を推定した²⁰⁾。この手法は、乳腺組織の方が脂肪組織よりも X 線減弱がわずかに大

きいというその差に着目したもので、乳腺密度の決定に X 線露光条件を利用した最初の研究である。しかし、当時はまだ乳腺密度の高精度な計測技術が確立されていなかったため、実験的手法によって推定された Beckett 等による乳腺密度の精度検証は行われていない。①～③に述べた研究は、それぞれの長所と短所を理解したうえで適切に利用することで、いろいろなシチュエーションにおける乳腺密度の決定に役立つことが期待できる。

これからの乳腺密度に期待されるのは、バイオマーカーとしての役割であると考え、その役割に着目したこれまでの研究においては、乳腺密度の変化と乳癌の再発リスクまたは乳癌治療の有効性との間には整合性のある結論が得られていない。すなわち、決定的な有効性はまだ結論付けられていない。その理由のひとつは、生データマンモグラムが残されていないために、同一患者の過去画像の乳腺密度変化を十分に追跡できないことである。Atakpa らは、乳腺密度の経時変化をバイオマーカーとして使用することについて、「効果の信頼性については多くの不確実性がある。これは、研究に参加した女性の数が少ないため十分な結果が得られておらず、乳腺密度がどのように変化するかなど、多くの研究に制限が課せられたためである」と結論付けている⁸⁾。

本章では、「予測」よりも高い精度が期待でき、「計測」よりも汎用性が高く、かつ、生データマンモグラムを必要としない乳腺密度の「推定」に焦点を当てる。そして乳腺密度をバイオマーカーとして利用した乳癌発症の予測や乳癌治療の予後研究の発展を支援するために、先行研究¹⁵⁾において乳腺密度と有意な相関が認められた CBT や年齢以外に、管電圧、mAs 値といったマンモグラフィで用いられた X 線露光データを説明変数に追加した重回帰分析により被検者の乳腺密度の推定を試みた。たとえ過去に撮影されたマンモグラフィにおいて生データマンモグラムが残されていなくても、撮影時の X 線露光データはどの施設でも確実に記録されているため、「推定」による乳腺密度の時系列解析が可能となる。

4.2 身体特性および生育歴情報を用いた乳腺密度予測（先行研究）

著者らは 2017 年から 2018 年にかけて、被検者の身体特性（年齢、CBT、Body Mass Index : BMI）および生育歴情報（出産歴、閉経の有無）を使って乳腺密度

を予測する研究を行った [近畿大学医学部倫理審査委員会承認 (29-056)]。本章はその延長にあたる研究となるため、簡単に内容を振り返る。

当然であるがマンモグラフィは X 線被ばくをともなう。もし、マンモグラフィを撮影せずに個人の乳腺密度をある程度予測できれば、個人に応じた乳癌検診のためのモダリティ選択に役立てられる。また、現在推奨されている 40 歳以上・2 年に一度という一律の基準を、各個人に応じて調整することも可能となる。これらは不要な X 線被ばくの低減にもつながるものである。このような観点から、重回帰分析を用いた乳腺密度の予測を試みた。

重回帰分析の目的変数は乳腺密度、説明変数は年齢・CBT・BMI・閉経の有無・出産歴である。ここでは大胸筋と皮膚面近傍の非圧迫領域を、目視および ImageJ の threshold 機能により乳房領域から除去し、残りの領域を対象として乳腺密度を算出した。用いたマンモグラフィは正常乳房の MLO 画像のみで、学習用 457 画像、テスト用 200 画像である。

表 4.1 身体特性および生育歴情報を用いた乳腺密度の重回帰モデル

Regression model					
Regression model	Partial regression coefficient	Standard partial regression coefficients	Standard error	t value	P value
Constant term	165.2213	-	6.9312	23.8373	0.0000
Age (years)	-0.7295	-0.3153	0.0972	-7.5055	0.0000
CBT	-1.1628	-0.5061	0.1019	-11.4068	0.0000
BMI	-0.9358	-0.1082	0.3891	-2.4052	0.0166
Menopause	-7.3069	-0.1154	2.6362	-2.7717	0.0058
Parity	-4.1910	-0.1429	0.9076	-4.6178	0.0000
Determination coefficient	Degree-of-freedom-adjusted determination coefficient				
0.6172	0.6129				
Analysis of variance					
Factors of variance	Degree of variance	Sum of squares	Mean square	F value	P value
Regression variation	5	243194.0572	48638.8114	145.4265	< 0.0001
Residual variation	451	150839.8441	334.4564	-	-
Total variation	456	394033.9012	-	-	-

CBT; Compressed breast thickness

BMI; Body mass index

表 4.1 に重回帰分析で得られた乳腺密度推定モデルを示す。有意水準 5% で年齢・CBT・BMI・出産歴・閉経の有無の 5 項目が有意に乳腺密度に影響を及ぼすことがわかった。また、影響の強さは標準化偏回帰係数より CBT が最も強く BMI の約 5 倍であった。実際には飲酒、喫煙、動物性脂肪摂取量などさらに

多くの因子を検討していたが、事前に実施したロジスティック回帰分析により有効な因子を上記 5 項目に絞った結果である。表 4.1 下段の分散分析では F 値が 145 と顕著に高く、モデルの説明力の高さがわかる。作成された予測式を式 (4.1) に示す。このモデルにより乳腺密度の全変動の 61.7%を説明できることがわかった。図 4.1 にテスト画像に対して予測式を適用した結果を示す。(a)の yy プロットでは概ね対角線近辺にプロットが集まっており、(b)の残差プロットでは特定の傾向はみられず、均一に分布している様子がわかる。

$$\begin{aligned} \text{Breast density (\%)} = & 165.2213 - (0.7295 \times \text{age}) - (1.1628 \times \text{CBT}) - (0.9358 \times \text{BMI}) \\ & - (7.3069 \times \text{menopause}) - (4.1910 \times \text{parity}) \end{aligned} \quad (4.1)$$

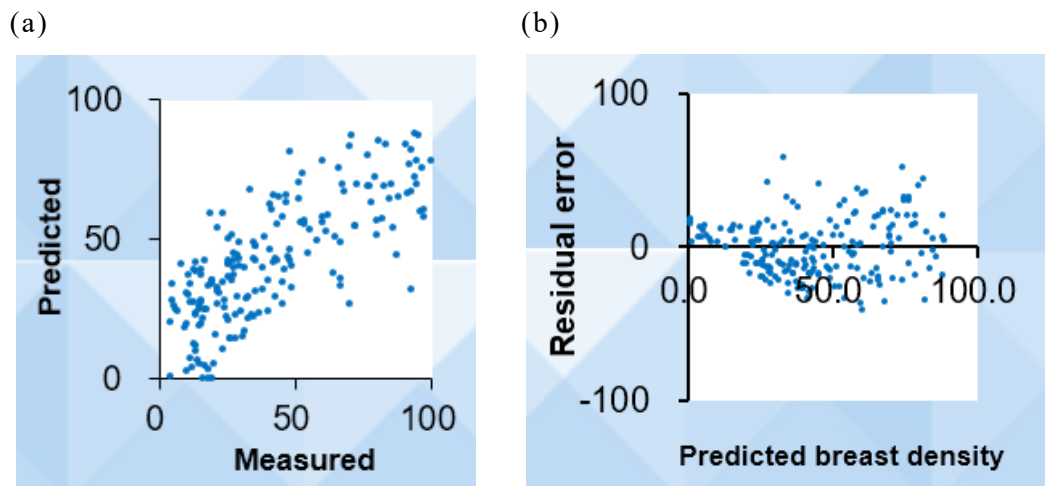


図 4.1 身体特性および生育歴情報から作成した重回帰モデルをテスト画像に対して適用した結果. (a) yy プロット (乳腺密度の計測値 vs 予測値), (b) 残差プロット.

先行研究の成果として、乳腺密度に影響を及ぼす因子は複数存在し、有効な因子を見つけることができれば重回帰分析等を用いてある程度は予測可能であることがわかった。しかし、説明率 (決定係数) 61.7%では臨床で用いるにはまだまだ不十分であり、更なる有効因子の解明が必要である。

4.3 乳腺密度推定モデルの作成方法

4.3.1 データ収集

本章では、第3章で収集した703画像に2021年11月から2022年4月にかけて近畿大学病院でマンモグラフィ検査を受けた被検者のMLOマンモグラム617画像を加えた計1320画像を使用した。すべての乳房は所見の無い正常乳房で、圧迫圧90N以上で撮影されたマンモグラムである。使用装置は第2章および第3章で用いたのと同じPe-ru-ru（キヤノンメディカルシステムズ）、ピクセルサイズは $85 \times 85 \mu\text{m}^2$ 、出力グレーレベルは13ビットで、撮影にはAECを使用した。なお、本研究は近畿大学医学部倫理審査委員会の承認（R02-237）を得て、研究に参加したすべての被検者から書面によるインフォームドコンセントを取得して実行された。

4.3.2 乳腺領域の画素値の決定法

第3章で述べたとおり、乳腺領域には乳腺組織が現在密集して存在するdense regionと、過去に乳腺組織が存在したと考えられる領域も含めたdiffuse regionがある。Diffuse regionは乳腺組織の大部分を含んでいるため、乳癌罹患リスクや乳癌発症率と密接に関連している。このことから、本章で扱う乳腺領域はdiffuse regionとした。

本章で扱うすべてのマンモグラムの乳腺領域は、第3章と同じ3名の専門認定診療放射線技師が独立して手動で抽出し、領域内の平均画素値をImageJで計測してExcelシート（Excel 2016、Microsoft Corporation、Redmond、WA、USA）に出力した。ただし、本章では抽出された乳腺領域内に4.3.3で述べるimproved threshold methodによって決定された脂肪領域の画素値よりも高い画素値が含まれる場合はその領域を除外して平均画素値を求めた。なぜなら、そのような領域は圧迫板と乳房支持台の間に、脂肪や乳腺よりもX線減弱の小さい空気を含む領域（圧迫板が乳房に接していない領域）であり、さらにはその割合も不明であることから、乳腺密度およびその算出精度を低下させるためである。これらの一連のプロセスは自作プログラム（MATLAB 2020a、Mathworks、Natick、USA）を使って実行した。

4.3.3 新しい脂肪領域画素値の決定法

第3章で述べたように、乳腺密度を高精度に計算するために最も重要なことは、乳房内の脂肪のみの領域を正確に見つけることである。先行研究では、ImageJ ソフトウェアに組み込まれている threshold 機能を使用して脂肪のみの領域を決定する独自の方法 (threshold method) を提案した¹⁵⁾。しかし、threshold method は脂肪組織が著しく少ない乳房や一般的ではない形状の乳房に適用することが困難な場合があった。以下に、threshold method を改良し、すべての乳房に対して脂肪領域の画素値を適切に決定できる手法を提案する。

著者は、脂肪領域の単位 mAs あたりの画素値は管電圧やターゲット／フィルタの組合せといった X 線エネルギーに関連する条件が同じであれば、乳房の形状や組成に依存せず、同じ CBT の乳房については同じ値となるはずであることに気付いた。以下、この考えに基づいた具体的な手法について述べる。まず、認定歴 20 年の診療放射線技師が threshold method を安定して確実に適用できる 510 枚の生データマンモグラムを選択した。そしてすべてのマンモグラムに対して threshold method を実施し、各画像について脂肪領域の画素値を決定した。その画素値を各マンモグラムの撮影に用いた mAs 値で除した後、CBT 16 mm から 82 mm まで 2 mm 間隔でグループ化し、最小二乗近似によって平滑化することで各 CBT における脂肪領域の単位 mAs あたりの画素値としてテーブル化した。この手法を improved threshold method と呼ぶ。このようなテーブルを作成することで、以後、乳腺密度を計測する場合マンモグラム毎に threshold method を実施する必要はなくなり、CBT に対応したテーブルの値にマンモグラフィで使用された mAs 値を乗じることで、乳房の脂肪量や形状に依らず正確な脂肪領域の画素値を迅速に得ることができるようにした。

4.3.4 乳腺密度の算出

第3章で算出済みの 703 画像の乳腺密度に加え、4.3.1 で追加収集した 617 画像についても 4.3.2 および 4.3.3 に示した方法で乳腺領域および脂肪領域の画素値を決定し、式 (3.1) および式 (3.4) に代入して乳腺密度の計測値とした。乳腺組織と脂肪組織の線減衰係数および物理密度は第3章で用いた値と同じである。

本章においても 2.5 mm の皮下脂肪層に囲まれた乳房モデルを想定し²⁴⁾、皮下脂肪層を除いた central region 内の乳腺組織と脂肪組織の質量比を乳腺密度と定義した²⁴⁾。

4.3.5 モデルの作成と評価方法

画像処理後のマンモグラムから乳腺密度を推定するために重回帰分析を行った。目的変数は 4.3.2 で求めた乳腺密度、説明変数はマンモグラフィで使用された管電圧、mAs 値、CBT および被検者の年齢である。すべての説明変数データは、DICOM ヘッダー情報からマンモグラフィの撮影直後に取得可能である。

用いた目的変数または説明変数に偏りがないかを確認するために、1320 枚のマンモグラムを使って、学習画像 1056 枚、テスト画像 264 枚とした 5 分割交差検証を実施し、各 fold の決定係数を考慮して最終的な重回帰モデルを決定した。重回帰モデルの精度は以下の 7 項目によって検証した。①乳腺密度に対する計測値 vs 推定値プロット (yy プロット)、②残差プロット、③決定係数、④二乗平均平方根誤差 (RMSE)、⑤平均絶対誤差 (MAE)、⑥二乗平均平方根誤差率 (RMSPE)、⑦平均絶対誤差率 (MAPE)。

本章で求める乳腺密度推定モデルは CBT 16 mm から 82 mm までの広い範囲をカバーするため、z 検定 (Microsoft Excel for Windows ver. 2016、Microsoft Corporation、WA、USA) を用いてテスト画像に対するモデル精度の CBT 依存性を調べた。CBT の分類は、①18 mm~36 mm (84 枚、薄い乳房)、②38 mm~46mm (91 枚、標準乳房)、③48 mm~78 mm (89 枚、厚い乳房) の 3 グループとした。比較対象が 3 種類であるため、Bonferroni の多重補正により、 $p < 0.0167$ を統計的に有意であると見なした。

4.4 乳腺密度推定モデルの精度評価

4.4.1 Improved threshold method 用いた脂肪領域画素値の精度検証

Improved threshold method を使用して決定したマンモグラム中の脂肪領域の単位 mAs あたりの画素値を図 4.2 に示す。画素値は CBT の増加とともに減少していくが、22 mm、32 mm、42 mm、52 mm、62 mm、72 mm、82 mm において○印で示すような不連続点が認められる。これらの不連続点は、CBT の増加に

ともない撮影装置が自動的に管電圧を増加させる CBT に相当している。各不連続点における管電圧の増加は CBT 62 mm のみ 3 kV で、それ以外はすべて 1 kV である。この現象は improved threshold method がわずかな管電圧の増加にともなう脂肪領域の画素値のわずかな増加を的確に捉えていることを表している。このように信頼できる脂肪領域の画素値をテーブル化したことは乳腺密度計測の高精度化に大きく貢献するものである。

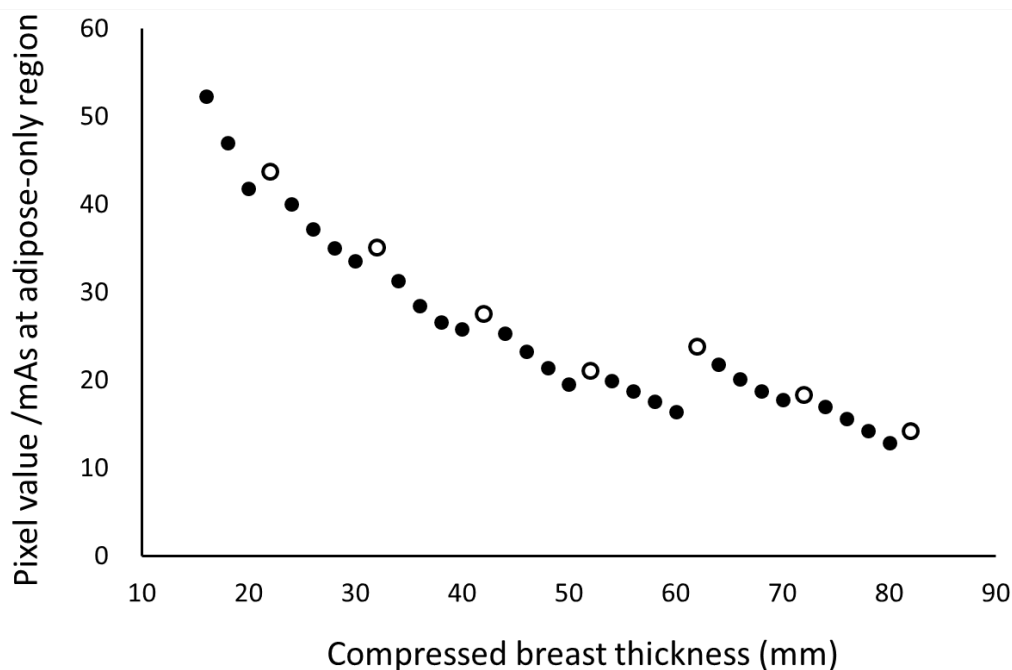


図 4.2 Improved threshold method で求めた CBT (X 軸) に対する脂肪領域の単位 mAs あたりの画素値 (Y 軸). CBT 10 mm 間隔で管電圧が 1 kV 増加するごとに画素値もわずかに増加する現象 (open circle) が的確に捉えられている.

4.4.2 乳腺密度推定モデルの決定

表 4.2 は 5 分割交差検証における乳腺密度の決定係数を示している。全 fold において決定係数は学習データで 0.85 から 0.87、テストデータで 0.84 から 0.89 の範囲内に収まり安定していた。本研究では fold 4 における重回帰モデルを本章で使用する推定モデルと決定した。その理由は学習データとテストデータ双方の決定係数が高く、かつ、両者の差が小さいためである。

表 4.2 5 分割交差検証の結果

	Data set	Determination coefficient R ²					Average
		Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	
Breast density	Training	0.867	0.864	0.864	0.860	0.853	0.862
	Testing	0.856	0.861	0.840	0.868	0.884	0.862

4.4.3 モデルの精度検証

表 4.3 に示すように、CBT・mAs 値・年齢・管電圧は、乳腺密度と有意な相関を示した。重回帰モデルから得られた乳腺密度の計算式を式 (4.2) に示す。

$$\begin{aligned} \text{Log [breast density (\%)]} = & 1.146 - [3.239 \times \log (\text{CBT})] + [2.214 \times \log (\text{mAs})] \\ & + (0.073 \times \text{tube voltage}) - (0.002 \times \text{age}) \quad (4.2) \end{aligned}$$

表 4.3 X 線露光条件を用いた乳腺密度の重回帰モデル

Regression model for Log breast density			Standard partial regression coefficients				VIF
Independent variables	Partial regression coefficients	Standard error	(in linear scale)	t value	p value		
Age	-0.00198	0.11268	-0.00090	-5.92505	4.22702E-09	1.21098	
X-ray tube voltage	0.07252	0.00537	0.00466	13.50164	1.99758E-38	6.22552	
Log X-ray tube current exposure time (mAs)	2.21431	0.04880	0.74422	45.37433	7.9342E-250	2.08890	
Log CBT (mm)	-3.23892	0.07760	-1.54365	-41.73871	2.6634E-225	8.27052	
Constant term	1.14614	0.11268	-	10.17132	3.04579E-23	-	

Determination coefficient : 0.868

Analysis of variance

Factors of variance	Degree of variance	Sum of squares	Mean square	F value	p value
Regression variation	4	81.51330	20.37833	1304.40822	0.00000
Residual variation	1051	16.41941	0.01562	-	-
Total variation	1055	97.93272	-	-	-

Log; Logarithm

CBT; Compressed breast thickness

VIF; Variance inflation factor

乳腺密度へ特に強く影響を及ぼす因子は CBT と mAs 値であり、CBT および mAs 値の影響の強さは管電圧の影響度と比べ各々 331 倍および 158 倍であった (表 4.3、standard partial regression coefficient : 標準化偏回帰係数)。年齢が乳腺密度へ及ぼす影響はこれら 2 つの因子より極めて小さいものの有意な相関が認められた ($p = 4.227 \times 10^{-9}$)。本章で用いた 4 つの因子は、表 4.3 に示されてい

るように乳腺密度の全変動の 86.8%を推定することができた。また、表 4.3 下段の分散分析結果からわかるように、この重回帰モデルは上記 4つの因子を与えなかった場合の 1304 倍の説明力を持ち（表 4.3 下段右から 2 列目 F 値）、このような説明力が偶然に発生する確率はほとんどゼロ（同最右列 p 値）である。図 4.3 に乳腺密度の計測値と推定値の関係（yy プロット）を示す。図 4.3(a)、図 4.3(b)とも特に顕著な異常値は見られず、どちらも対角線付近にプロットが分布している。図 4.4 は図 4.3 の残差プロットを表しており、(a)および(b)は各々図 4.3 の(a)と(b)に対応している。残差プロットにおいても特に顕著な傾向は見られず、(a)、(b)ともおおむねランダムな分布を示している。

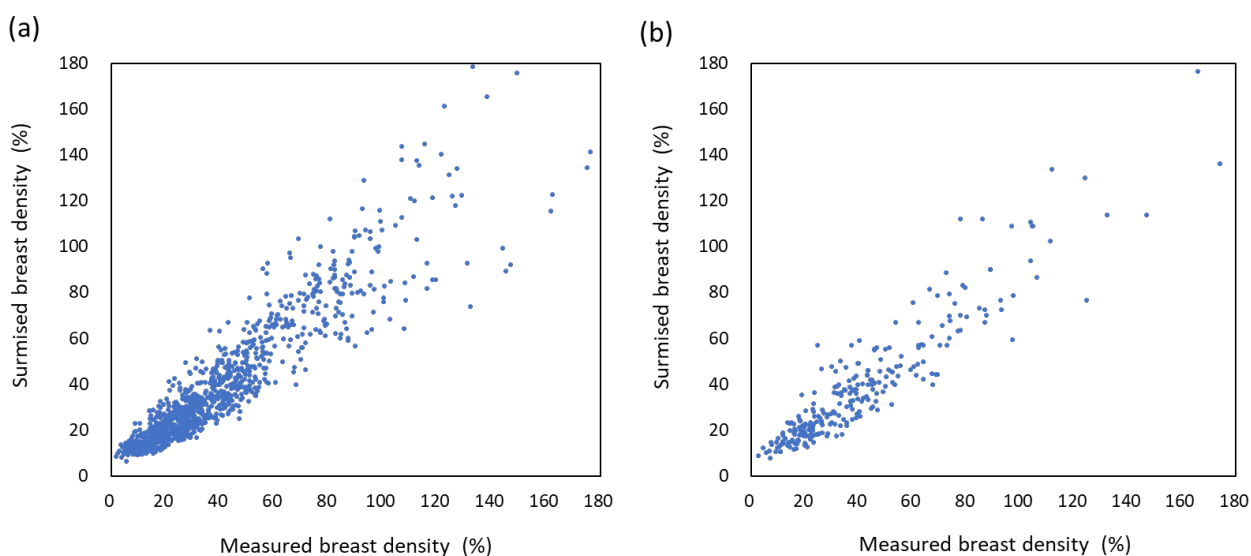


図 4.3 乳腺密度の yy（計測値 vs 推定値）プロット。(a)、(b)は各々学習データ、テストデータの結果を示す。

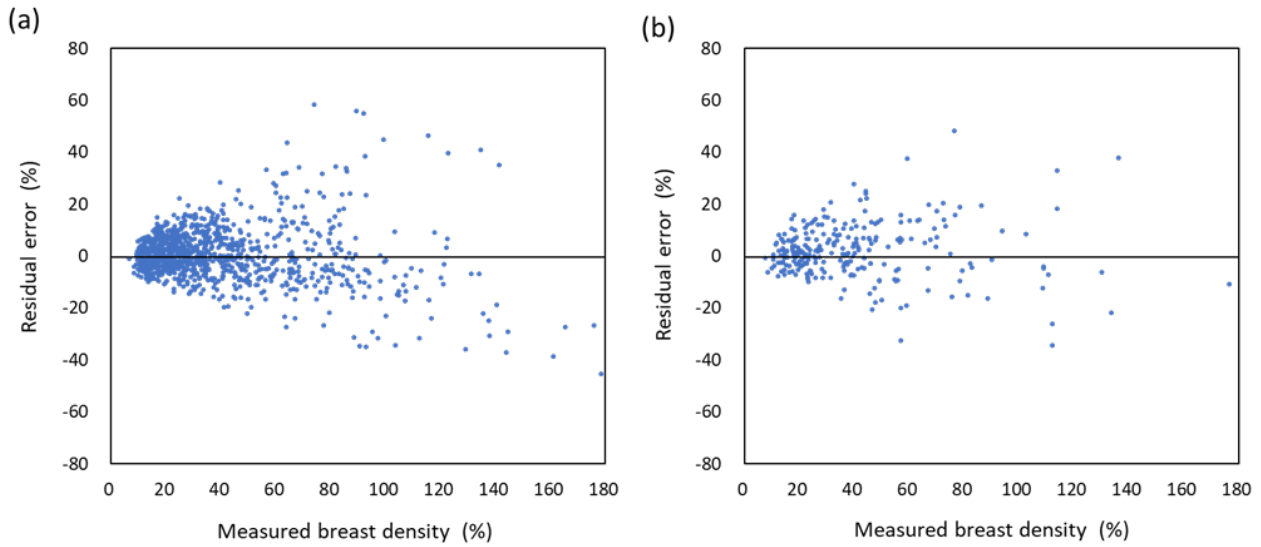


図 4.4 乳腺密度の残差プロット. (a), (b)は各々学習データ, テストデータの結果を示す.

推定モデルの検証項目を表 4.4 に示す。左から RMSE、MAE、RMSPE、および MAPE である。各項目とも特に顕著な誤差は見られない。表 4.5 は 3 種類の CBT グループ間の相関係数の差に対する z 検定の結果を示している。平均的な CBT の集団である 38 mm~46 mm における相関係数は 0.945 と最も高い値を示し、日本人女性にとっては厚い乳房となる 48 mm~78 mm の相関係数は 0.867 と最も低い値を示した。結果としてそれら 2 種類の CBT グループ間で相関係数に有意差が見られた ($p = 0.00231$)。

表 4.4 乳腺密度推定モデルの精度検証

	Data set	RMSE	MAE	RMSPE	MAPE
Breast density	Training	10.581	7.445	0.374	0.236
	Testing	10.942	7.943	0.332	0.224

RMSE; 二乗平均平方根誤差, MAE; 平均絶対誤差,

RMSPE; 二乗平均平方根誤差率, MAPE; 平均絶対誤差率

表 4.5 z 検定による推定精度の乳房厚依存性

Breast thickness range	Sample size	Correlation coefficient
18 mm～36 mm	84	0.913
38 mm～46 mm	91	0.945
48 mm～78 mm	89	0.867

Z-test target	Test statistic	p value
18 mm～36 mm vs. 38 mm～46 mm	1.542958	0.12284
38 mm～46 mm vs. 48 mm～78 mm	3.046717	0.00231*
18 mm～36 mm vs. 48 mm～78 mm	1.449232	0.14727

* p < 0.0167 due to Bonferroni correction

4.5 考察

X 線マンモグラフィに関する 4 つの因子（CBT、mAs 値、管電圧、年齢）を使用して、乳腺密度を推定するための新しい式を提案した。先行研究では、さまざまな身体特性と生育歴情報を用いた予測式を提案したが、乳腺密度の全変動の 61.7%しか説明できなかった。これに対し本章では、表 4.3 に示す *determination coefficient*（決定係数）から、乳腺密度の全変動の 86.8%が、マンモグラフィを用いた X 線露光条件を主とした 4 つの因子だけを使用して説明できることを示した。これらの露光条件や被検者情報の記録は、精中機構によるマンモグラフィ検診施設認定取得の条件となっている。

4.1 で述べたように、mAs 値と CBT の関係から臨床マンモグラムの乳腺密度を実験的に推定した報告もあるがその精度は検証されておらず、アクリルフェントムを用いた実験的手法では、限定された条件下での乳腺密度しか算出できないため、誰もが使えるような乳腺密度推定式の導出には至ることができない。本章では、高精度に測定された乳腺密度を目的変数とする重回帰分析により誰でも簡単に乳腺密度を算出できる推定式を確立した。

表 4.3 の 3 列目に示した標準化偏回帰係数からわかるように、CBT と mAs 値は管電圧と比べると乳腺密度に対して 150 倍以上の大きな影響を与えており、これら 2 つの因子が乳腺密度のほとんどを決定していることがわかる。ただし、CBT および管電圧については表 4.3 最右列の *variance inflation factor*（VIF）が

示すように多重共線性が疑われ、管電圧では 6.23、CBT では 8.27 であった。マンモグラフィでは CBT に基づいて管電圧が設定されるためこの関係を避けることはできない。VIF の許容範囲は一概には定められていないが、10 を超える場合は回帰モデルの信頼性が問題となり、関連の強い説明変数を削除することが望ましいとされている。今回のモデルでは 10 には至っておらず、さらに 5 分割交差検証における各 fold の決定係数が、学習データおよびテストデータの両方に対し安定して高い値を示しているため、得られた乳腺密度推定式は信頼できるものであると考える。ただし、この点に関する今後の課題として人工知能技術などを用い、多重共線性が問題にならないような推定式を構築することが望まれる。表 4.4 に示すように、乳腺密度の推定値と計測値の誤差は、テストデータにおいて RMSE が 10.942、MAE が 7.443 とほぼ 10% 程度又はそれ以内に収まっており、本推定式は臨床での実用性を備えていると考える。

表 4.5 に見られるように、推定値と計測値間の相関係数は厚い CBT 群が中間の CBT 群と比べて有意に低い結果であった ($p = 0.00231$)。その理由として以下の二点が考えられる。①本論文では、先行研究同様、乳房内の皮下脂肪層厚を 2.5 mm と仮定した。しかし、厚い乳房については皮下脂肪層が 2.5 mm よりも厚い場合が多いと考えられる。一方で、乳腺密度は“皮下脂肪層を除いた central region 内における脂肪と乳腺の質量比”と定義されているため²⁵⁾、皮下脂肪層が 2.5 mm よりも厚い乳房では乳腺密度が過小評価されることになり、乳腺密度の計測誤差が増加する。その結果、回帰モデルの推定精度が低下したと考える。将来的に個人毎の皮下脂肪層の厚さを適切に設定できるようになり、乳腺密度計測技術がさらに向上すれば重回帰モデルの推定精度もさらに向上することが期待できる。②撮影時に乳房を圧迫する圧迫板は撮影支持台と平行ではなくわずかに傾きをもっているため、胸壁から乳頭方向への距離が遠くなるほど撮影支持台と圧迫板の距離から計測される CBT の表示値に誤差を生じやすくなる。乳腺密度は装置が表示する CBT が正しいとして計算されるため、このような現象は乳腺密度の計測誤差を増加させ、重回帰モデルの精度低下につながるかと考える。乳腺密度計測対象となる乳腺領域の位置における正確な CBT の表示を期待したい。

本章で述べた研究にはいくつかのリミテーションがある。第一に、乳腺密度

の計測領域として **diffuse region** を対象としたことである。この領域は乳癌罹患率や乳癌発症率のリスクと深く関連しているため、バイオマーカーとして乳腺密度を利用するために有効な領域であると考えるが、もし病変検出感度への影響を調べるための研究であれば、**dense region** に対する乳腺密度を推定する必要がある。第二に、この推定式はマンモグラフィ撮影に使用される AEC の精度が十分に高く、かつ、AEC の X 線検出領域が計測対象の乳腺領域と一致していることを前提としている。AEC の精度については近年普及している FPD を用いたデジタルマンモグラフィ装置においては十分に達成されていると言える。後者については、特徴のある乳房（X 線減弱率が大きい所見があるなど）に対して AEC を用いると乳腺領域と X 線検出領域が一致していない可能性が高くなる。したがって、左右で顕著に異なる mAs 値が記録されているときは注意する必要がある。しかし、ほとんどの装置メーカーは AEC の X 線検出領域を開示していないため、マンモグラム毎に両者のズレの程度を確認することができない。本研究で用いたマンモグラフィ装置を含め、ほとんどの装置では本撮影の直前にプレ照射として微量の X 線を乳房へ照射し、乳腺が存在する領域をおおむね確認したうえで AEC の検出領域を設定しているので両者が顕著に異なることはないと考えるが、より高精度な推定モデルを作成するために、マンモグラム毎に AEC の設定領域を表示するシステムの提供を期待する。

4.6 結言

マンモグラフィ撮影の直後に得られる 4 つの因子（CBT、mAs 値、管電圧、年齢）を用いて個々の乳腺密度を推定するためのシンプルなモデルを構築した。このモデルにより、乳腺密度の全変動の 86.8% を説明することができた。この結果は、たとえマンモグラフィ画像が残されていなくても、X 線の露光条件が記録されていれば、乳腺密度を高精度に推定することができることを示している。本章の成果は、マンモグラフィ撮影前のプレ照射を利用した被検者個々の乳腺密度に対応する最適な画像処理条件の迅速設定や乳腺密度の経時変化を追跡することによるバイオマーカーとしての研究の発展に貢献するものである。

参考文献

- 1) C. R. Castellano, P. M. Aguilar Angulo, L. C. Hernández, P. S. González-Carrato, R. G. González, J. Alvarez, J. I. Chacón, J. Ruiz, M. Á. Fuentes Guillén, G. Gutiérrez Ávila: Breast cancer mortality after eight years of an improved screening program using digital breast tomosynthesis, *J. Med. Screen*,28:456–463(2021) doi: 10.1177/09691413211002556. Epub 2021 Mar 29
- 2) J.S. Torres-Roman, J. F. Martinez-Herrera, G. Carioli, J. Ybaseta-Medina, B. Valcarcel, J. A. Pinto, A. Aguilar, K. A. McGlynn, C. La Vecchia: Breast cancer mortality trends in Peruvian women, *BMC Cancer* 20:1173(2020) doi: 10.1186/s12885-020-07671-x
- 3) C. Fitzmaurice, et al: Global Burden of Disease Cancer Collaboration. *JAMA Oncology*, 5,1749–1768(2019)
- 4) N. Azamjah, Y. Soltan-Zadeh, F. Zayeri: Global Trend of Breast Cancer Mortality Rate: A 25-Year Study, *Asian Pac. J. Cancer Prev.*,20,2015–2020(2019) doi: 10.31557/APJCP.2019.20.7.2015
- 5) M. Román, J. Louro, M. Posso, R. Alcántara, L. Peñalva, M. Sala, J. Del Riego, M. Prieto, C. Vidal, M. Sánchez, X. Bargalló, I. Tusquets, X. Castells: Breast density, benign breast disease, and risk of breast cancer over time, *Eur. Radiol.*, 31:4839–4847(2021) doi: 10.1007/s00330-020-07490-5
- 6) L. L. Reimers, M. Goldberg, P. Tehranifar, K. B. Michels, B. A. Cohn, J. D. Flom, Y. Wei, P. Cirillo, M. B. Terry: Benign breast disease and changes in mammographic breast density, *Breast Cancer Res.*, 23,49(2021) doi: 10.1186/s13058-021-01426-7
- 7) J. Lian, K. Li: A Review of Breast Density Implications and Breast Cancer Screening, *Clin. Breast Cancer*, 20,283–290(2020) doi: 10.1016/j.clbc.2020.03.004
- 8) E. C. Atakpa, M. A. Thorat, J. Cuzick, A. R. Brentnall: Mammographic density, endocrine therapy and breast cancer risk: a prognostic and predictive biomarker review, *Cochrane Database Syst. Rev.*,10:CD013091(2021)
- 9) I. Skarping, D. Förnvik, U. Heide-Jørgensen, H. Sartor, P. Hall, S. Zackrisson, S. Borgquist: Mammographic density as an image-based biomarker of therapy response in neoadjuvant-treated breast cancer patients, *Cancer Causes Control*

- 32,251–260(2021) doi: 10.1007/s10552-020-01379-w
- 10) J. H. Porembka, J. Ma, H. T. Le-Petross: Breast density, MR imaging biomarkers, and breast cancer risk, *Breast J.* 26,1535–1542(2021) doi: 10.1111/tbj.13965
 - 11) M. S. Shawky, H. Martin, H. J. Hugo, T. Lloyd, K. L. Britt, A. Redfern, E. W. Thompson: Mammographic density: a potential monitoring biomarker for adjuvant and preventative breast cancer endocrine therapies, *Oncotarget*,8,5578–5591(2017) doi: 10.18632/oncotarget.13484
 - 12) S. Azam, M. Eriksson, A. Sjölander, R. Hellgren, M. Gabrielson, K. Czene, P. Hall: Mammographic Density Change and Risk of Breast Cancer, *JNCI J. Natl. Cancer Inst.* 112,391–399(2020) doi: 10.1093/jnci/djz149
 - 13) E. Y. Kim, Y. Chang, J. Ahn, J. S. Yun, Y. L. Park, C. H. Park, H. Shin, S. Ryu: Mammographic Breast Density, Its Changes, and Breast Cancer Risk in Premenopausal and Postmenopausal Women, *Cancer*,126:4687–4696(2020) doi: 10.1002/cncr.33138 Epub 2020 Aug 7
 - 14) E. Serwan, D. Matthews, J. Davies, M. Chau M: Mechanical standardisation of mammographic compression using Volpara software Radiography (Lond), 27,789–794(2021) doi: 10.1016/j.radi.2020.12.009
 - 15) M. Yamamuro, Y. Asai, K. Yamada, Y. Ozaki, M. Matsumoto, T. Murakami: Prediction of glandularity and breast radiation dose from mammography results in Japanese women, *Med. Biol. Eng. Comput.*,57,289–298(2019) doi: 10.1007/s11517-018-1882-4
 - 16) G. Richard-Davis, B. Whittemore, A. Disher, V. M. Rice, R. B. Lenin, C. Dollins, E. R. Siegel, H. Eswaran: Evaluation of Quantra Hologic Volumetric Computerized Breast Density Software in Comparison with Manual Interpretation in a Diverse Population. *Breast Cancer (Auckl)*, 12,1178223418759296(2018) doi: 10.1177/1178223418759296
 - 17) N. Saffari, H. A. Rashwan, M. Abdel-Nasser, V. Kumar Singh, M. Arenas, E. Mangina, B. Herrera, D. Puig: Fully Automated Breast Density Segmentation and Classification Using Deep Learning. *Diagnostics (Basel)*,10:988(2020) doi: 10.3390/diagnostics10110988

- 18) H. P. Chan, M. A. Helvie: Deep Learning for Mammographic Breast Density Assessment and Beyond, *Radiology*, 290, 59–60 (2019) doi: 10.1148/radiol.2018182116 Epub 2018 Oct 16
- 19) J. Lee, R. M. Nishikawa: Automated mammographic breast density estimation using a fully convolutional network, *Med. Phys.*, 45,1178–1190(2018) doi: 10.1002/mp.12763
- 20) J. R. Beckett, C. J. Kotre: Dosimetric implications of age-related glandular changes in screening mammography, *Phys. Med. Biol.*, 45,801–13(2000) doi: 10.1088/0031-9155/45/3/316
- 21) C. D’Orsi, E. A. Sickles, E. B. Mendelson, E. A. Morris: Breast Imaging Reporting and Data System: ACR BI-RADS breast imaging atlas, 5th ed., Reston, VA: American College of Radiology (2013)
- 22) M. Yamamuro, Y. Asai, N. Hashimoto, N. Yasuda, T. Yamada, M. Nemoto, et al: How to select training data to segment mammary gland region using a deep-learning approach for reliable individualized screening mammography. *Medical Imaging: Computer-Aided Diagnosis*, edited by Maciej A. Mazurowski, Karen Drukker, Proc of SPIE Vol 11597, 115972V (2021) doi: 10.1117/12.2581424
- 23) S. van Engeland, P. R. Snoeren, H. Huisman, C. Boetes, N. Karssemeijer: Volumetric breast density estimation from full-field digital mammograms, *IEEE Trans. Med. Imaging*, 25,273–282(2006) doi: 10.1109/TMI.2005.862741
- 24) M. Yamamuro, Y. Asai, N. Hashimoto, N. Yasuda, T. Yamada, M. Nemoto, et al: Utility of a U-Net for the objective segmentation of the mammary gland region on clinical digital mammograms, *Biomed. Phys. Eng. Express*, 8,4(2022) Doi: 10.1088/2057-1976/ac7ada
- 25) D. R. Dance, R. A. Hunt, P. R. Bakic, A. D. A. Maidment, M. Sandborg, G. Ullman, G. A. Carlsson: Breast dosimetry using high-resolution voxel phantoms, *Radiation Protection Dosimetry*, 114,359–363(2005)
- 26) N. Pandis: Comparison of 2 means (independent z test or independent t test), *Am. J. Orthod. Dentofacial Orthop.*, 148:350–351(2015)

第 5 章 結論

乳癌は早期発見・早期治療による根治が見込まれるがんのひとつである。そのためには国が推奨する対策型乳癌検診の定期的な受診が重要となる。マンモグラフィは、現在、死亡率を減少させる効果が認められている唯一のモダリティであり世界中で乳癌検診プログラムに組み入れられている。しかし、乳腺の割合が多い乳房では「高濃度乳房問題」と呼ばれるマンモグラフィの感度低下や乳癌発症リスクの増加が認められている。この問題について知識を持つ被検者は少なく、マンモグラフィ検診結果の“異常なし”に絶対的な信頼が寄せられているのが現状である。乳腺密度の通知は、被検者自身のリスクや高濃度乳房問題を知るきっかけとなり、満足度の高い個別化検診実現への最初のステップと言える。

本論文は、我が国においても近い将来実現するであろう個別化乳癌検診の実現を技術的側面から支援するものである。まず、乳腺密度と病変検出感度のあいだに有意な相関があることを検証し、続いて 2 次元のマンモグラムから高精度に 3 次元乳腺密度を計測する手法を提案した。その過程において人工知能技術のひとつである U-net を 3 次元乳腺密度算出のための乳腺領域抽出に利用できる可能性を検証した。さらに、乳腺密度を新薬の効果予測や乳癌術後の予後予測のバイオマーカーとして利用する研究を支援するために、乳腺密度を計測できる画像が残されていない過去のマンモグラフィについても、記録された X 線露光条件を使って乳腺密度を推定するモデルを構築した。

本論文は 5 章から成っている。

第 1 章では、本研究の歴史的背景をまとめ、乳癌診療における乳腺密度の重要性および本論文の目的と内容を述べた。

第 2 章では、乳腺密度と病変検出感度のあいだには統計的に有意な相関があることをファントム製造メーカーと共同開発したオリジナルファントムを使って検証した。開発したオリジナルファントムはマンモグラフィで用いられる全 X 線光子エネルギー範囲に対して実乳房と等価な線減弱係数を有しており、その画像特性は臨床マンモグラムと同等である。

乳腺密度を 25%、50%、および 75% に調整したオリジナルファントムに、マ

ンモグラフィにおける代表的な病変である微小石灰化、腫瘤、およびスピキュラを模した模擬病変を挿入し、乳腺専門認定を持つ5名の診療放射線技師によるROC実験を行った。結果を病変タイプ別にみると、腫瘤は乳腺密度の増加とともに単調に病変検出率が低下した。微小石灰化は乳腺密度が25%から50%へ増加したとき有意に病変検出率が低下したが、50%から75%へ増加しても病変検出率は有意な低下を示さなかった。スピキュラについては乳腺密度25%と50%では病変検出能に有意な差が認められたが、50%と75%における病変検出率はほとんど同じであった。スピキュラに見られた現象を解明するために、オリジナルファントムの各乳腺密度におけるノイズパワースペクトルを測定した結果、本実験で用いたスピキュラの空間周波数帯域である 0.05 mm^{-1} におけるノイズパワーは25%と50%間で顕著な差を有していたが、50%と75%間では顕著な差がなかった。このことから、スピキュラのような低空間周波数で低コントラストの病変の検出能はその背景に存在する正常乳腺の分布に依存することがわかった。乳腺密度と病変検出率の間には病変のタイプ毎に特徴が見られたものの、3病変を総合的に評価した場合、3種類の乳腺密度のどの組み合わせにおいても、有意水準5%において乳腺密度が高くなると病変検出率が低くなる結果となった。このように、従来経験則として語られてきた現象の正当性を、オリジナルファントムを用いたROC実験によって統計的に検証し、乳腺密度の増加にともなう病変検出能の低下率を定量的に示した。

第3章では、ディープラーニング技術の一つであるU-netを用いた乳腺領域自動抽出システムを開発し、自動抽出された乳腺領域に対して算出した乳腺密度および平均乳腺線量の臨床適用の可能性について述べた。

ディープラーニング技術を用いて乳腺領域を自動抽出する研究はここ数年多くの研究者が取り組んでいるが、その精度はDice係数のような領域一致度だけで評価されている。しかし、乳腺領域の一致度と乳腺領域について算出した乳腺密度や平均乳腺線量の一致度は必ずしも強い相関関係にあるとは限らない。臨床において重要となるのは乳腺密度や平均乳腺線量そのものの一致度である。そこで、U-netが自動抽出した乳腺領域と乳腺エキスパートが抽出した乳腺領域の各々について、Engelandの式を用いて乳腺密度を、Danceの式を用いて平均乳腺線量を算出し、それらを評価指標としたB-A解析を行うことで両乳腺領

域の一致度を評価した。この一連の過程において、Engeland の式を解くために要求されるマンモグラム中の脂肪領域の決定法について、マンモグラムの解剖学的知見に基づく独自の手法を示すとともに、これまで Dance の式の中で利用することができなかった乳腺密度に対する補正項 c ファクタを取り入れた高精度な平均乳腺線量の算出を実現した。

B-A 解析の結果、U-net が自動抽出した乳腺領域と乳腺エキスパートが抽出した乳腺領域各々に対して算出した乳腺密度には同等性が、また、平均乳腺線量には互換性があることを見出し、統計的に互いに入れ替え可能であること、すなわち臨床適用が可能であることを示した。

第 4 章では、近年行われつつある乳腺密度の経時変化を新薬の効果予測バイオマーカーや乳癌術後の予後予測バイオマーカーとして用いる研究を支援するために、第 3 章で述べたような乳腺密度計測を実施することができない過去のマンモグラフィについて乳腺密度を推定する手法について述べた。過去に実施されたマンモグラフィにおいて生データマンモグラムが残されていない場合は乳腺密度を計測することが困難である。しかし、画像がなくてもマンモグラフィに用いた X 線露光条件（具体的には管電圧と mAs 値）が記録されていれば、それらの数値に年齢や CBT を加味した重回帰分析を実施することにより、撮影された当時の乳腺密度を推定することが可能であることを示した。

X 線が被写体を通過する際の減弱過程が指数関数を用いて表されることは周知の事実である。したがって、撮影に用いる X 線のエネルギー（管電圧）と X 線の量（mAs 値）に加え被写体の厚さ（CBT）がわかれば、線減弱係数が既知である脂肪組織と乳腺組織で構成されるシンプルな乳房モデルにおける各組織の質量比を容易に決定できることは誰もが考え及ぶことであろう。しかし、これまで上記理論に基づいて乳腺密度を推定した報告はなかったため、本研究では乳腺密度の時系列解析が要求される研究への支援とすべく乳腺密度推定モデルの構築を試みた。

重回帰モデルにおける標準化偏回帰係数から、乳腺密度の決定に最も顕著な影響を及ぼす因子は CBT であり、続いて mAs 値、管電圧、年齢の順であった。これら 4 つの因子を用いることで乳腺密度の全変動の 86.8% を説明できることがわかった。本研究では対数変換を組み入れた重回帰式を使って推定モデルを

構築したが、将来的には同様の説明変数に対して人工知能技術を利用して目的変数を推定することで更なる精度の向上が期待できる。

第5章では、本研究の成果を総括した。本研究の目的は、マンモグラフィを用いた乳腺密度の高精度な計測手法を確立し、個別化乳癌検診などの臨床分野へ応用することである。そのために、まず、従来経験則であった“乳腺密度が増加すればマンモグラフィの感度が低下する”という両者の定性的関係を、乳房等価オリジナルファントムを使った ROC 実験を行うことで定量化した。また、VBDM のために乳腺エキスパートが抽出した乳腺領域を忠実、かつ、簡便に再現するために、ディープラーニング技術の一種である U-net を用いた乳腺領域自動抽出モデルを構築し、このモデルを用いて抽出した乳腺領域に対して算出した乳腺密度および平均乳腺線量は、乳腺エキスパートが抽出した乳腺領域に対して算出したそれらの結果と同等性又は互換性を有することを示した。さらに、過去のマンモグラフィを遡り、乳腺密度の時系列変化を分析するため、撮影に用いた X 線露光条件などから乳腺密度を推定するモデルを提案した。

本論文は、2003 年に米国から広まったマンモグラフィにおける高濃度乳房問題に対応する技術体制の確立を見据え、今後導入が期待されている個別化乳癌検診に必要な被検者個々の乳腺密度や被ばく線量を高精度に計測または推定するための手法をまとめたものである。

本研究の成果が、我が国の個別化乳癌検診における乳腺密度やそれに対応するリスクの通知およびマンモグラフィス検診プログラムの最適化に貢献することを期待する。

謝辞

本研究の遂行と本論文の作成にあたり、主たるご指導ならびにご鞭撻を賜った新潟大学大学院保健学研究科放射線技術科学分野の近藤世範教授に衷心より感謝をささげる。

本論文を審査していただき、有益なご助言とともに叱咤激励をいただいた新潟大学大学院保健学研究科放射線技術科学分野の高橋直也教授ならびに小林公一教授に深甚なる謝意を表す。

本研究の遂行に際し、終始、ご指導ならびに有益なご助言、ご討論をいただいた近畿大学病院中央放射線部の浅井義行副技師長、マンモグラムの画像解析に多大なご協力をいただいた橋本直美氏、安田奈生氏、木村浩都氏に深く感謝する。また、人工知能技術に関して有益なご指導を賜った近畿大学高度先端医療センターPET分子イメージング部の山田誉大博士に深甚なる謝意を表す。

本研究遂行上の便宜を図っていただいた近畿大学医学部放射線診断学部門の石井一成教授、近畿大学病院中央放射線部の南部秀和技師長、ならびに診療放射線技師の方々に深く感謝する。

本研究は、一部、文部科学省科学研究費助成事業[基盤研究C，課題番号：18K07736 および 21K07657]の援助を受けて行ったものである。ここに謝意を表す。

研究業績

本博士論文は以下の学術論文 2 編、国際会議 proceeding 5 編、および学会発表 10 件についてまとめた内容である。

学術論文 2 編

1. **Mika Yamamuro**, Yoshiyuki Asai, Naomi Hashimoto, Nao Yasuda, Yoshiaki Ozaki, Kazunari Ishii, and Yongbum Lee. The effect of breast density on the missed lesion rate in screening digital mammography determined using an adjustable-density breast phantom tailored to Japanese women. PLoS ONE, Vol. 16(1), 2021, e0245060. <https://doi.org/10.1371/journal.pone.0245060>
2. **Mika Yamamuro**, Yoshiyuki Asai, Naomi Hashimoto, Nao Yasuda, Hiorto Kimura, Takahiro Yamada, Mitsutaka Nemoto, Yuichi Kimura, Hisashi Handa, Hisashi Yoshida, Koji Abe, Masahiro Tada, Hitoshi Habe, Takashi Nagaoka, Seiun Nin, Kazunari Ishii, and Yohan Kondo. Utility of U-Net for the objective segmentation of the fibroglandular tissue region on clinical digital mammograms. Biomed. Phys. Eng. Express, Vol. 8(4), 2022, 045016. <https://doi.org/10.1088/2057-1976/ac7ada>

このうち、第 2 章は 1. の内容について記した。また、第 3 章は 2. の内容について記した。

国際会議 proceeding 5 編

1. **Mika Yamamuro**, Yoshiyuki Asai, Naomi Hashimoto, Nao Yasuda, Kenta Sakaguchi, Tatsuo Konishi, Koji Yamada, Yoshiaki Ozaki, Kazunari Ishii, Yongbum Lee. Deep learning-based segmentation of mammary gland region in digital mammograms of scattered mammary glands and fatty breasts. Proc. SPIE

- 11513, 15th International Workshop on Breast Imaging (IWBI2020), 115131V (22 May 2020); doi:10.1117/12.2562180
2. **Mika Yamamuro**, Yoshiyuki Asai, Naomi Hashimoto, Nao Yasuda, Koji Yamada, Yoshiaki Ozaki, Kazunari Ishii, Yongbum Lee. Effectiveness of high-luminance display monitors in digital mammography. Proc. SPIE 11513, 15th International Workshop on Breast Imaging (IWBI2020), 115131T (22 May 2020); doi:10.1117/12.2555276
 3. **Mika Yamamuro**, Yoshiyuki Asai, Naomi Hashimoto, Nao Yasuda, Takahiro Yamada, Mitsutaka Nemoto, Yuichi Kimura, Hisashi Handa, Hisashi Yoshida, Koji Abe, Masahiro Tada, Hitoshi Habe, Takashi Nagaoka, Yoshiaki Ozaki, Seiun Nin, Kazunari Ishii, Yongbum Lee. How to select training data to segment mammary gland region using a deep-learning approach for reliable individualized screening mammography. Proc. SPIE 11597, Medical Imaging 2021: Computer-Aided Diagnosis, 115972V (15 February 2021); doi:10.1117/12.2581424
 4. **Mika Yamamuro**, Yoshiyuki Asai, Naomi Hashimoto, Nao Yasuda, Yoshiaki Ozaki, Seiun Nin, Kazunari Ishii, Takahiro Yamada, Mitsutaka Nemoto, Yuichi Kimura, Hisashi Handa, Hisashi Yoshida, Koji Abe, Masahiro Tada, Hitoshi Habe, Takashi Nagaoka, Yongbum Lee. Can radiological technologists create training data for automatically segmenting the mammary gland region in non-dense breasts? CARS 2021: Computer Assisted Radiology and Surgery Proceedings of the 35th International Congress and Exhibition. Munich, Germany, June 21–25, 2021. Int J CARS (2021) 16 (Suppl 1): S84–8
 5. **Mika Yamamuro**, Yoshiyuki Asai, Naomi Hashimoto, Nao Yasuda, Hiroto Kimura, Takahiro Yamada, Mitsutaka Nemoto, Yuichi Kimura, Hisashi Handa, Hisashi Yoshida, Koji Abe, Masahiro Tada, Hitoshi Habe, Takashi Nagaoka, Seiun Nin, Kazunari Ishii, Yongbum Lee. Robustness of a U-net model for different image processing types in segmentation of the mammary gland region. SPIE Proceedings, Vol.12286, 16th International Workshop on Breast Imaging (IWBI2022), 122860T (13 July 2022); doi10.1117/12.2624139

このうち、第2章は2.の内容について記した。また、第3章は1.3.4.5.の内容について記した。

学会発表

1. **Mika Yamamuro**, Yoshiyuki Asai, Naomi Hashimoto, Nao Yasuda, Kenta Sakaguchi, Tatsuo Konishi, Koji Yamada, Yoshiaki Ozaki, Kazunari Ishii, Yongbum Lee. Deep learning-based segmentation of mammary gland region in digital mammograms of scattered mammary glands and fatty breasts. 15th International Workshop on Breast Imaging. May 25–27, 2020, Leuven
2. **Mika Yamamuro**, Yoshiyuki Asai, Naomi Hashimoto, Nao Yasuda, Koji Yamada, Yoshiaki Ozaki, Kazunari Ishii, Yongbum Lee. Effectiveness of high-luminance display monitors in digital mammography 15th International Workshop on Breast Imaging. May 25–27, 2020, Leuven
3. **山室美佳**, 浅井義行, 橋本直美, 安田奈生, 尾崎吉明, 任誠雲, 石井一成, 山田誉大, 根本充貴, 木村裕一, 吉田久, 半田久志, 李鎔範. Deep learning segmentation での複数の放射線技師による教師データ作成の有用性. 第30回日本乳癌検診学会学術総会. 2020年11月22日–23日. 仙台
4. **山室美佳**, 浅井義行, 橋本直美, 安田奈生, 尾崎吉明, 任誠雲, 石井一成, 山田誉大, 根本充貴, 木村裕一, 吉田久, 半田久志, 李鎔範. Deep learning を用いた非高濃度乳房に対する乳腺領域の自動抽出. 第30回日本乳癌検診学会学術総会. 2020年11月22日–23日. 仙台
5. **Mika Yamamuro**, Yoshiyuki Asai, Naomi Hashimoto, Nao Yasuda, Takahiro Yamada, Mitsutaka Nemoto, Yuichi Kimura, Hisashi Handa, Hisashi Yoshida, Koji Abe, Masahiro Tada, Hitoshi Habe, Takashi Nagaoka, Yoshiaki Ozaki, Seiun Nin, Kazunari Ishii, Yongbum Lee. How to select training data to segment mammary gland region using a deep-learning approach for reliable individualized screening mammography. SPIE Medical Imaging 2021 Feb. 14–18, 2021. San Diego

6. **Mika Yamamuro**, Yoshiyuki Asai, Naomi Hashimoto, Nao Yasuda, Yoshiaki Ozaki, Seiun Nin, Kazunari Ishii, Takahiro Yamada, Mitsutaka Nemoto, Yuichi Kimura, Hisashi Handa, Hisashi Yoshida, Koji Abe, Masahiro Tada, Hitoshi Habe, Takashi Nagaoka, Yongbum Lee. Can radiological technologists create training data for automatically segmenting the mammary gland region in non-dense breasts? 35th International Congress and Exhibition on Computer Assisted Radiology. June 21–25, 2021. Munich
7. **山室美佳**, 浅井義行, 橋本直美, 安田奈生, 木村浩都, 任誠雲, 石井一成, 山田誉大, 根本充貴, 木村裕一, 半田久志, 吉田久, 阿部孝司, 多田昌裕, 波部斉, 永岡隆, 李鎔範. 画像処理条件の違いがディープラーニングによる乳腺領域自動抽出及び乳腺密度算出精度に及ぼす影響. 第 191 回医用画像情報学会学術大会. 2021 年 10 月 9 日. 新潟
8. **山室美佳**, 浅井義行, 橋本直美, 安田奈生, 任誠雲, 石井一成, 山田誉大, 根本充貴, 木村裕一, 吉田久, 半田久志, 李鎔範. 3 次元乳腺密度計測の高精度化に向けたマンモグラム上での脂肪領域画素値の決定法 第 30 回日本乳癌検診学会学術総会. 2021 年 11 月 26 日–27 日. 京都
9. **Mika Yamamuro**, Yoshiyuki Asai, Naomi Hashimoto, Nao Yasuda, Hiroto Kimura, Takahiro Yamada, Mitsutaka Nemoto, Yuichi Kimura, Hisashi Handa, Hisashi Yoshida, Koji Abe, Masahiro Tada, Hitoshi Habe, Takashi Nagaoka, Seiun Nin, Kazunari Ishii, Yongbum Lee. Robustness of a U-net model for different image processing types in segmentation of the mammary gland region. 16th International Workshop on Breast Imaging. May 22–25, 2022. Leuven
10. **山室美佳**, 浅井義行, 橋本直美, 安田奈生, 木村浩都, 任誠雲, 石井一成, 山田誉大, 根本充貴, 木村裕一, 吉田久, 半田久志, 近藤世範. 3 次元乳腺密度を指標とした U-net の乳腺領域自動抽出精度の評価. 第 32 回日本乳癌検診学会学術総. 2022 年 11 月 11 日–12 日. 浜松

このうち第 2 章は 2. の内容について記した。また、第 3 章は 1. 3. 4. 5. 6. 7. 9. 10. の内容について記した。第 4 章は 8. の内容について記した。

付録

本論文で用いた略語一覧

略語	完全表記	意味
ACR	American College of Radiology	米国放射線専門医会
AEC	Auto Exposure Control	自動露出制御器
AUC	Area Under the Curve	曲線下面積
B-A 解析	Bland-Altman 解析	ブランド-アルトマン解析
BI-RADS	Breast Imaging Reporting and Data System	乳がん画像診断総合ガイドライン
BMI	Body Mass Index	ボディ・マス指数
CBT	Compressed Breast Thickness	圧迫乳房厚
CI	Confidence Interval	信頼区間
DCNN	Deep Convolutional Neural Network	階層型畳み込みニューラルネットワーク
DICOM	Digital Imaging and Communications in Medicine	医療データ通信の国際標準規格
FFDM	Full Field Digital Mammography	フラットパネル搭載型デジタルマンモグラフィ
FPD	Flat Panel Detector	フラットパネルディテクタ
FPF	False Positive Fraction	偽陽性率
GT	Ground Truth	グラウンドトゥルース（本論文では正解乳腺領域を指す）
ICRU	International Commission on Radiation Units and Measurements	国際放射線単位測定委員会
LOA	Limits of Agreement	一致限界
Log	Logarithm	対数
MAE	Mean Absolute Error	平均絶対誤差
MAPE	Mean Absolute Percentage Error	平均絶対誤差率

mAs	Tube Current–Exposure Time Product	管電流時間積
MGD	Mean Glandular Dose	平均乳腺線量
MLO	Medio-Lateral Oblique	内外斜位方向
MRI	Magnetic Resonance Imaging	磁気共鳴画像診断装置
NPS	Noise Power Spectrum	ノイズパワースペクトル
RMSE	Root Mean Squared Error	二乗平均平方根誤差
RMSPE	Root Mean Squared Percentage Error	二乗平均平方根誤差率
ROC	Receiver Operatorating Characteristic	受信者動作特性
ROI	Region of Interest	関心領域
SD	Standard Deviation	標準偏差
TPF	True Positive Fraction	真陽性率
VBDM	Volumetric Breast Density Measurement	3次元の乳腺密度解析
VIF	Variance Inflation Factor	分散拡大係数