

論文 / 著書情報
Article / Book Information

題目(和文)	免疫組織化学における画像解析の臨床利用に向けたWhole Slide Imageの色補正手法の研究
Title(English)	
著者(和文)	大西千絵
Author(English)	Chie Ohnishi
出典(和文)	学位:博士(工学), 学位授与機関:東京工業大学, 報告番号:甲第12706号, 授与年月日:2024年3月26日, 学位の種別:課程博士, 審査員:山口 雅浩,中本 高道,鈴木 賢治,小尾 高史,渡辺 義浩
Citation(English)	Degree:Doctor (Engineering), Conferring organization: Tokyo Institute of Technology, Report number:甲第12706号, Conferred date:2024/3/26, Degree Type:Course doctor, Examiner:,,,,
学位種別(和文)	博士論文
Type(English)	Doctoral Thesis

博士論文

免疫組織化学における画像解析の
臨床利用に向けた
Whole Slide Image の色補正手法の研究

指導教員 山口 雅浩

東京工業大学 工学院 情報通信系

大西 千絵

2024 年 2 月

目次

第1章	序論.....	1
1.1	研究の背景と目的.....	1
1.2	本論文の構成.....	2
第2章	研究背景・基礎理論.....	4
2.1	病理検査に関する基礎知識.....	4
2.1.1	乳がんにおける治療.....	4
2.1.2	病理検査.....	6
2.1.3	HER2 検査.....	8
2.1.4	Whole Slide Image の利用と課題.....	10
2.2	基礎理論.....	15
2.2.1	カラーアンミキシング.....	15
2.2.2	非負値行列因子分解.....	16
2.2.3	色の表示法.....	17
2.2.4	色差.....	19
2.2.5	ガンマ補正.....	21
2.2.6	色のキャリブレーション.....	22
2.3	関連研究.....	24
第3章	キャリブレーションスライドを用いた色味・濃さの補正手法.....	26
3.1	概要.....	26
3.2	IHC キャリブレーションスライドを用いた染色の補正.....	29
3.2.1	全体像.....	29
3.2.2	Step1：色素成分ベクトルの取得.....	30
3.2.3	Step2,4：キャリブレーション画像からの濃さの取得.....	32
3.2.4	Step3：閾値の算出.....	32

3.2.5	手法1：閾値の補正.....	33
3.2.6	手法2：画像の補正.....	34
3.3	カラーチャートスライドを用いたスキャナの色補正	35
3.3.1	全体像.....	35
3.3.2	カラーチャートスライドの基準色の取得	36
3.3.3	スキャナの色取得.....	37
3.3.4	逆ガンマ補正.....	37
3.3.5	色補正関数の取得.....	38
3.4	まとめ	39
第4章	染色によるばらつきの補正効果の検証	40
4.1	概要.....	40
4.2	実験内容.....	40
4.2.1	実験データ.....	40
4.2.2	IHC キャリブレーションモデルの散乱項の影響の検証.....	41
4.2.3	HER2の自動解析における提案手法の効果の検証.....	41
4.2.4	基準データと異なる組織標本の補正効果の検証	43
4.3	実験結果.....	43
4.3.1	IHC キャリブレーションモデルの散乱項の影響の検証結果.....	43
4.3.2	HER2の自動解析における提案手法の効果の検証結果	44
4.3.3	基準データと異なる組織標本の補正効果の検証結果.....	50
4.4	考察.....	51
4.5	まとめ	55
第5章	スキャナによるばらつきの補正効果の検証	57
5.1	概要.....	57
5.2	実験内容.....	57
5.2.1	実験データ.....	57
5.2.2	色補正関数算出手法および多項式変換の次数の検討	57

5.2.3	色補正効果の検証：スキャナ間の色差の算出.....	58
5.2.4	カラーチャートの色のシミュレーション	58
5.2.5	HER2 の自動解析における提案手法の補正効果の検証	59
5.2.6	基準データと異なる組織画像での補正効果の検証.....	61
5.3	実験結果.....	61
5.3.1	色補正関数算出手法および多項式変換の次数の検討.....	61
5.3.2	色補正効果の確認.....	66
5.3.3	カラーチャートの色のシミュレーション	69
5.3.4	HER2 の自動解析における提案手法の補正効果の検証結果	73
5.3.5	基準データと異なる組織画像の補正効果の検証結果.....	76
5.4	考察.....	77
5.4.1	カラーチャートを用いたスキャナの色補正	77
5.4.2	HER2 の自動解析への補正手法の効果.....	80
5.5	まとめ	84
第 6 章	HER2 IHC 検査の高度化に向けた検討.....	85
6.1	概要.....	85
6.2	IHC 法と ISH 法の比較実験.....	85
6.2.1	実験データ.....	85
6.2.2	HER2 評価で得られるパラメータを用いた ISH ステータス分類.....	86
6.2.3	組織の DAB の濃さの分布からの ISH ステータス分類	87
6.3	実験結果.....	89
6.3.1	HER2 評価で得られるパラメータを用いた ISH ステータス分類結果.....	89
6.3.2	組織の DAB の濃さの分布からの ISH ステータス分類結果	91
6.4	考察.....	92
6.4.1	HER2 評価で得られるパラメータを用いた ISH ステータス分類.....	92
6.4.2	組織の DAB の濃さの分布からの ISH ステータス分類	95
6.5	社会実装に向けた検討.....	96

6.6	まとめ	99
第7章	結論.....	101
7.1	本論文のまとめ.....	101
7.2	研究の発展性	102
参考文献	105
研究業績	111
謝辞	113

[略語の説明]

IHC (Immunohistochemistry) : 免疫組織化学染色

HER2 (Human epidermal growth factor type 2) : 乳がんの増殖に関わるタンパク

ASCO/CAP (American Society of Clinical Oncology/College of American Pathologists) : 米国病理学会。病理検査に関するガイドラインなどを定める

FDA (Food and Drug Administration) : アメリカ食品医薬品局

WSI (Whole Slide Imaging) : デジタル病理画像

DAB 染色 (3,3' -Diaminobenzidine 染色) : IHC 法で用いられる褐色の染色色素

ISH (*in situ* hybridization) : 特定の遺伝子の分布状態を検出する染色方法

CISH (Chromogenic *in situ* hybridization) : 明視野で観察可能な染色方法を利用した ISH 法

第1章 序論

1.1 研究の背景と目的

がんの治療法の一つとして、近年、分子標的治療薬が選択されるようになってきている。分子標的治療薬はがん細胞の増殖や転移に関わるタンパク質やその生成遺伝子を標的に作用する効果があり、正常な細胞も同時に攻撃してしまう抗がん薬に比べ、副作用による身体への負担を軽減した治療が期待できる。ただし、抗がん薬とは異なる副作用が生じることや、対象となるタンパク・遺伝子が増幅していない場合には効果がみられないことから、病理検査によって分子標的治療薬の適応性を判断したうえで治療方針が決定される。病理検査では患者から採取した細胞や組織から病理標本スライドを作製し、検査目的に合わせて染色したスライドを顕微鏡で観察することで診断を行う。

女性における部位別のがんで罹患数が最も多い乳がんでは、代表的な分子標的治療薬として HER2 タンパクを対象としたトラスツズマブが挙げられる。HER2 タンパクは乳がんの増幅や悪性化に関係し、過剰発現が確認された患者は転移や再発の危険性が高いことが知られている。がんの転移・再発の予測やトラスツズマブの適応性判定のため、すべての浸潤性の乳がん患者で HER2 検査を行うことが決められている[1][2]。免疫組織化学染色 (Immunohistochemistry : IHC)法は、抗体を用いて病理組織内のタンパク質発現を検出・解析する染色手法であり、HER2 検査の第一段階として用いられる。IHC 法では細胞膜を染色する DAB 染色の濃さと腫瘍細胞の割合を評価指標として HER2 スコアを判定し、スコアが陽性だった場合にトラスツズマブの適応性ありと判断される。DAB 染色の濃さは「かすかな染色」「強い染色」といった言葉によるガイドラインの定義を、病理医が経験と主観に基づき顕微鏡観察下で判断している。そのため、判定する病理医によって結果のばらつきや誤りが生じる恐れが指摘されている[3]–[5]。

判定の客観性を向上させるため、従来の顕微鏡での診断に代わり、病理標本スライド全体をバーチャルスライドスキャナによって高倍率でデジタル画像化した、Whole Slide Image (WSI)及び WSI用の自動画像解析ソフトウェアの利用ニーズが高まっている[6]。自動画像解析の臨床診断への利用は検査の迅速化や効率化につながり、近年の検査数の増大や病理医不足といった課題への対応になると期待される。しかしながら、IHC 法における WSI 利用時の課題に画像の色や DAB の濃さのばらつきがあり、その要因として染色とスキャンの2つが挙げられる。アメリカでは臨床利用が認可された自動染色装置と試薬のセットを用いて病理標本スライドを染色するように決められているが、そのような管理された環境下であっても日々の染色で組織標本スライドの DAB の濃さが変わってしまうことがある[7][8]。さらに、WSI の場合、病理標本スラ

イドのデジタル化に利用するバーチャルスライドスキャナによっても画像の色や明るさが変わってしまう[7][9]。DABの濃さを判定の指標とするHER2検査ではこのような画像の違いが画像解析に影響し、判定スコアの変動につながってしまう。臨床診断への利用が認可されたHER2検査用の自動解析ソフトウェアでは、画像の色が変わらない条件での解析に限定するため、利用できる自動染色装置や試薬、スキャナの組み合わせを限定している。このような解析ソフトウェアの導入には既存の病理検査のプロトコルの変更が必要となる場合もあり、HER2検査における画像解析の臨床利用は進んでいない。

本研究では、臨床のHER2検査において自動画像解析を導入しやすくなるように、WSIの色と濃さのばらつきを補正することを目的とする。色の違いに特に影響を与える染色、スキャンのそれぞれに原因を分けて対応するため、IHC法の反応の程度を把握するためのキャリブレーションと、スキャナの特性を把握するためのキャリブレーションの2種類のスライドを利用する。また、実用化されている画像解析ソフトウェアでHER2の評価を行うことを想定し、それらと併用可能な補正手法を提案する。染色状態が異なる画像や、どのようなスキャナで撮影した画像であっても、病理医の診断結果と一致し、一貫した判定結果が自動解析で得られるようにすることを目標とする。

1.2 本論文の構成

図1に本論文の構成を示す。本論文は7章の構成である。

第2章では、病理検査に関する基礎知識を記載する。本研究で対象とした乳がんにおける治療法や検査法、さらに、病理分野で応用が期待されるWSIの概要についてまとめる。また、提案手法や検証実験で用いる光学の基礎理論についてまとめ、最後に関連する先行研究を記載する。

第3章では、2種類のキャリブレーションスライドを用いた画像の補正手法について記載する。提案手法では、染色状態を把握するための指標として、IHC法に対応したキャリブレーションスライドを染色の工程に追加する。キャリブレーションスライドから組織標本に対応したDABの濃さや色味を取得することで、スコア判定対象となる標本の染色のばらつきの補正を行う。スキャンの工程にはWSIスキャナの色校正用のカラーチャートスライドを導入し、異なるスキャナ間の画像の色のキャリブレーションに利用する。

第4章では、染色による画像のばらつきを対象とした手法の有効性検証実験について記載する。病理医によってHER2スコアの判定が行われた乳がんの組織標本スライド、IHCキャリブレーションスライドのセットを異なる日に染色し、日々の染色のばらつきを反映したデータセット

を取得して実験に用いた。すべてのデータセットを1台のスキヤナでデジタル画像化し、病理医の判定結果と自動解析ソフトウェアの結果を比較することで、補正を行わない場合と提案手法を適用した場合の自動解析への影響を検証した。

第5章では、スキヤナによるばらつきを対象とした手法の有効性検証実験について記載する。ここでは第4章で用いたデータセットを2台のWSIスキヤナで撮影し、実験に用いた。はじめに、カラーチャートスライドによる色補正効果を検証するために、2台のスキヤナの色補正前後の画像のRGB値の違いを比較した。次に、2枚のキャリブレーションスライドの組み合わせを変えて補正を行い、自動解析への影響を検証した。

第6章では、HER2 IHC検査の高度化に向けた検討をまとめる。HER2検査ではIHC法での判断が難しい場合に、ISH法という染色法で追加検査が行われる。提案手法の適用によってIHC法における自動解析結果の一貫性が上がることで、本来ISH法で判断されるステータスをIHC法の画像解析から判断できるのではないかと考えた。また、ISH法と比較することでIHC法の信頼性が高まる可能性もあると考えた。その検討内容を記載する。さらに、第4、5章での検証内容を臨床に応用するための検討を記載する。

第7章では、結論と今後の研究の発展性を述べる。

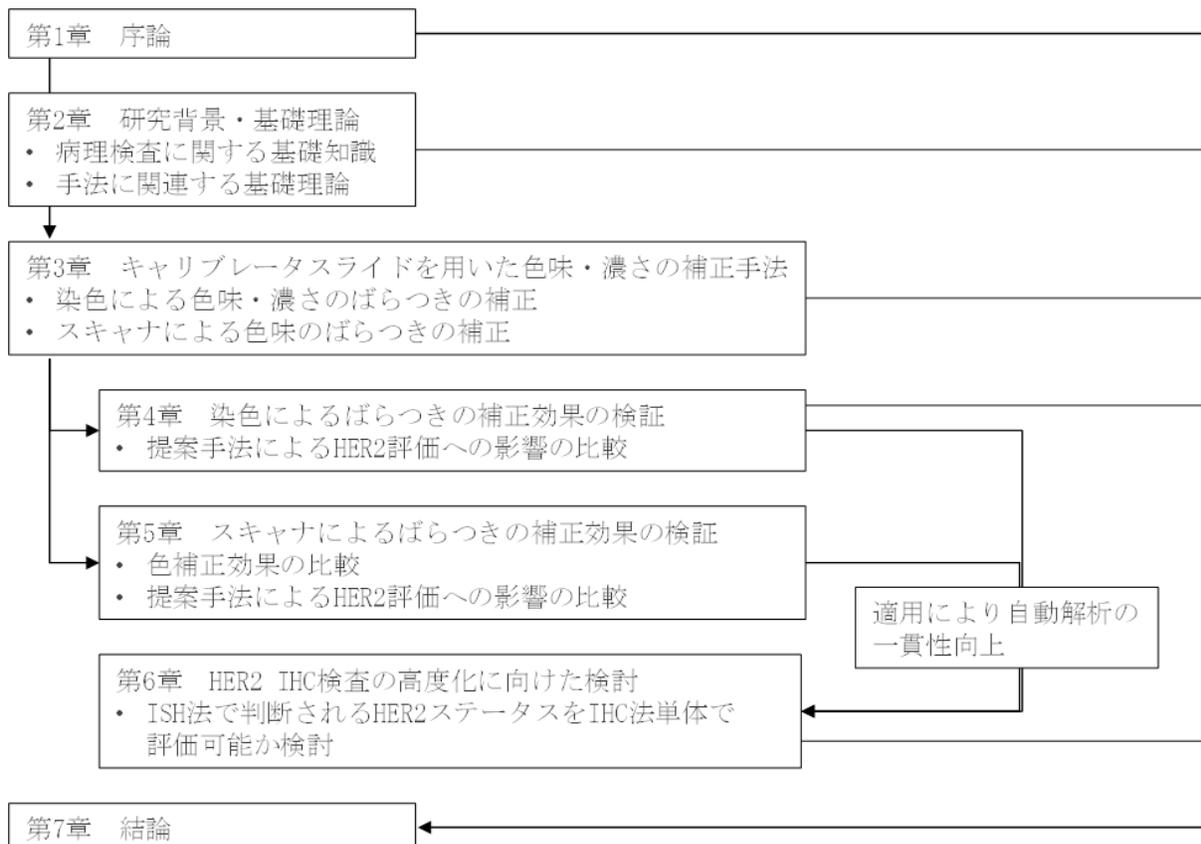


図1 本論文の構成

第2章 研究背景・基礎理論

2.1 病理検査に関する基礎知識

2.1.1 乳がんにおける治療

乳がんは乳腺の組織にできるがんで、女性における部位別のがんで最も罹患数が多いがんである。がん細胞が発生個所である乳管・乳腺小葉にとどまっている「非浸潤がん」であれば、がんを取り切ることによって完治が見込まれるが、がん細胞が乳管の外にまで増殖している「浸潤がん」では、リンパ管や血管から他の臓器へがん細胞が侵襲することもあり、転移や再発の危険性が高まる。視診、触診、マンモグラフィ検査、超音波検査が初期検査として選択され、乳がんの可能性がある場合には病理検査による診断が行われる。病理検査の目的は大きく 2 つに分けられ[10]、初めに症状の原因を診断する目的で行われる。乳がんと確定診断された後は、がんの進行度や性質などの診断を目的とし、結果は治療方針の決定に利用される。治療は主に手術、放射線治療、薬物療法の組み合わせで行われる[11]。

薬物療法には、手術前のがんを小さくする、再発の危険性を下げる、手術が困難な進行がんや再発に対して延命や症状の緩和を行う等の目的があり、抗がん薬やホルモン療法薬、分子標的治療薬が主に用いられる。中でも分子標的治療薬は次世代の医薬品として近年注目を集めている。がん細胞の増殖には特定の因子が関係するとされており、分子標的治療薬はその因子に対してピンポイントに作用するように作られている。抗がん薬が正常な細胞も含めて増殖の盛んな細胞を攻撃するのに対し、特定の因子のみを標的とする分子標的治療薬は患者の副作用を抑えることができる。分子標的治療薬の標的対象は、がん細胞の増殖に関わるタンパク質や遺伝子、血管やがんを攻撃する免疫に関わるタンパク質などが挙げられる。表 1 に乳がん治療に用いられる分子標的治療薬の一例を示す[10]。多くの場合は抗がん薬と組み合わせて用いられる。抗 HER2 薬であるトラスツズマブは、乳がんにおける分子標的治療薬として最も代表的なものである。HER2 は乳がんの増殖に関わるタンパク質であり、細胞の表面に存在する(図 2)。核内の HER2 生成遺伝子の増殖によって HER2 タンパクが過剰に発現し、がん細胞の増殖につながると考えられている。浸潤性乳がんの 15~25%で HER2 生成遺伝子の増幅またはタンパクの過剰発現が認められたとの報告がある[12]。HER2 タンパクの過剰発現が確認された乳がんは転移や再発の危険性が高いことが知られており、HER2 タンパクを標的とするトラスツズマブを抗がん薬と併用することによって、ある一定の期間を経過した時点で生存している人の割合を示す全生存率が 37%相対的に改善、10 年生存率が 75.2%から 84%に上昇し[13]、患者の予後

改善に有望な治療薬となっている。これらの報告から見ても分子標的治療は効率的な治療薬ではあるが、薬の種類によって特徴的な副作用が出現することや、費用が高額であるなどの理由から、全ての患者が選択できるわけではなく、病理検査によって適応可能な患者を判断したうえで選択される。

表 1 乳がん治療に用いられる分子標的治療薬の例 [10]

#	タイプ	標的	治療薬	効果
1	抗 HER2 薬	HER2 タンパク	トラスツズマブ	乳がんの増殖に関わる HER2 タンパクに作用し、抗がん薬との併用で再発の危険性を抑制
2			トラスツズマブ エムタンシン	トラスツズマブ+抗がん薬では効果がみられなくなった転移乳がんを使用することで生存期間を延長
3	血管新生阻害薬	血管新生	ベバシズマブ	がん細胞に栄養や酸素を運ぶ血管の新生を防ぐことでがんの進行を抑制
4	免疫チェックポイント阻害薬	PD-L1、PD-1 タンパク	アテゾリズマブ	がんを攻撃する免疫反応にブレーキをかける PD-1/PD-L1 タンパクに作用し、免疫細胞の機能を強めることでがん細胞を制御

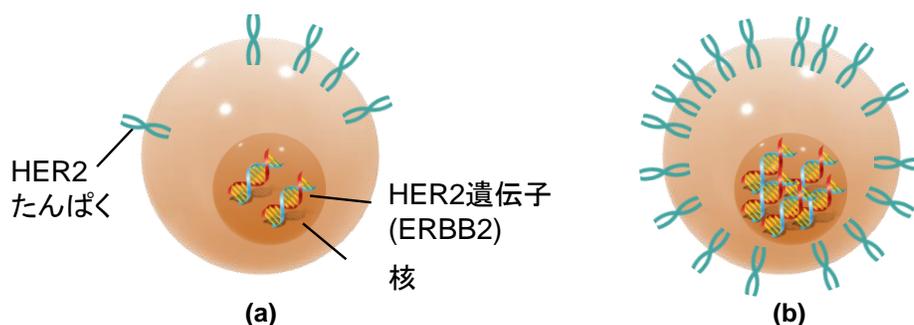


図 2 HER2 発現のイメージ；(a)正常細胞；(b)HER2 増幅細胞

2.1.2 病理検査

病理検査は、患者から採取した組織や細胞を染色し、顕微鏡で観察することで、病気の診断、治療方針の決定、治療効果の評価、予後の判定などを行う重要な検査方法である。症状のみでは確定診断が難しい場合も多く、病理検査は医療現場において不可欠な役割を果たしている。

図3に病理検査のおおまかな流れを示す[14]。検査ではがんの疑いのある箇所の一部を針で採取する「細胞診」と、ある程度広範囲の細胞を含む組織を採取する「組織診」があり、本論文では組織診を対象とする。患者から採取された組織は、腐敗防止や構造の安定化等の目的ですぐにホルマリン固定が行われる。組織の大きさや種類に応じて6~72時間固定される。固定後の組織は一定の大きさに切り出され、パラフィンの浸透が行われる。パラフィンを浸透させることで組織が適度な硬さと粘り気を持つようになり、後に薄く切り出すことが可能となる。パラフィンを浸透させた組織は包埋皿に移され、パラフィンの中で冷やし固められる。この工程で完成したものを組織標本ブロックと呼ぶ。ブロックから約3~4 μmの厚みに薄切された標本をガラススライドに固定して組織標本スライドを作成し、診断の目的に応じた染色法で染色する。染色されたスライドを病理医が顕微鏡で観察することで診断が行われる。各工程には、それぞれに注意すべき点があり、例えば、固定時間が短すぎると組織の形態が変化し、診断に影響を与える可能性がある。



図3 病理検査の流れ

組織標本ブロック一つに対して一つの検体を含むのが基本的だが(図4(a))、複数の組織を一つのブロックに集約した組織マイクロアレイ (Tissue Microarray: TMA)ブロックが使用されることもある(図4(b))。組織標本ブロックから検索対象となる部分を選定し、小さな円柱(コア)状に抜き取ったものを一つのブロックに集約・整列させて再包埋することでTMAブロックが出来上が

る。TMA ブロックから薄切した組織標本スライド 1 枚には多数の小组織片が存在するため、複数の臓器や異なる症例を一度に染色して分析、比較できる利点がある。

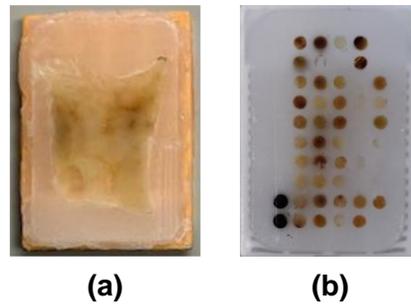


図 4 組織標本ブロックの例; (a) 一検体のブロック; (b) 組織マイクロアレイブロック

図 5 に代表的な染色法における画像の例を示す。通常は一つの組織標本ブロックから複数枚の組織標本スライドを作成し、診断の目的に応じて複数の染色法が選択される。ヘマトキシリン・エオジン(H&E)染色は病理検査において最も基本的な染色方法として、組織の形態を観察するのに利用される。ヘマトキシリンが細胞核や骨組織などを青紫色に、エオジンが細胞質や赤血球などを赤色に染色する。乳がんではがんの浸潤部を特定するために H&E 染色が必ず利用される。IHC 法は抗原抗体反応を利用し、抗体を用いて組織標本上のタンパク質 (抗原)を検出する。反応の標識には茶褐色の DAB 色素が用いられることが多い。ホルモン療法薬の適応性を判断したい場合には細胞核を、分子標的治療薬の適応性を判断したい場合には細胞膜を標識する方法がとられる。他にも粘液や脂肪、遺伝子を染色する方法や、蛍光色素を用いた染色法などがある。

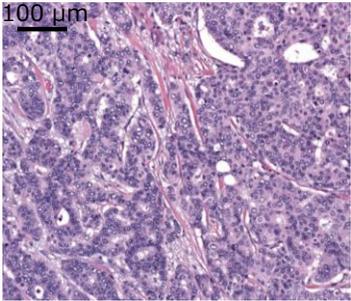
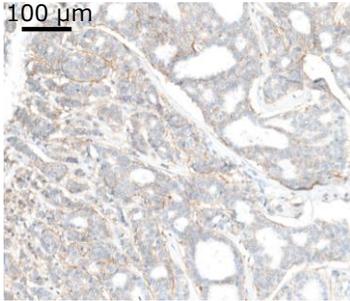
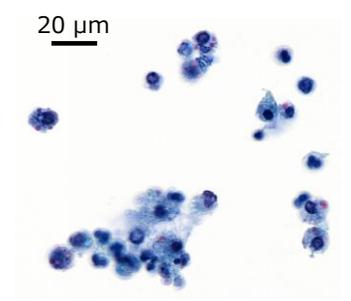
種類	ヘマトキシリン・エオジン(H&E)染色	免疫組織化学染色	特殊染色
染色例			
概要	化学反応によって細胞・組織の構成要素を標識	抗原抗体反応を利用して組織標本上のたんぱくを標識	化学反応によって細胞、粘液、脂肪、細菌等を標識

図 5 染色法の例

2.1.3 HER2 検査

がんの転移・再発の危険性の予測や、トラスツズマブの適応性判定のため、全ての浸潤性乳がんに対して HER2 検査を行うことがガイドラインで決められている。HER2 検査は細胞表面に存在する HER2 タンパクを標識する IHC 法と、HER2 生成遺伝子を標識する ISH 法の 2 種類の染色方法が用いられる。

図 6 に HER2 検査の流れを示す。ASCO/CAP が発表した HER2 検査ガイドライン[1][2]に準拠した検査ガイド[15][16]が日本病理学会により定められ、これに基づいた検査が実施されている。検査では HER2 の状態を評価する第一段階として IHC 法が適用される。IHC 法では細胞膜の染色強度と陽性細胞の割合をもとに、0、1+、2+、3+の 4 段階のスコアに分類を行う。0 または 1+ を HER2 陰性、3+ を HER2 陽性、2+ を HER2 equivocal (未確定) と判定し、2+ の場合はさらに ISH 法によって再判定が行われる。ISH 法では核内の遺伝子発現の数を計測することで ISH 陰性、ISH 陽性を判定する。HER2 陽性、及び、ISH 陽性と判定された場合、トラスツズマブの適応ありと判断される。

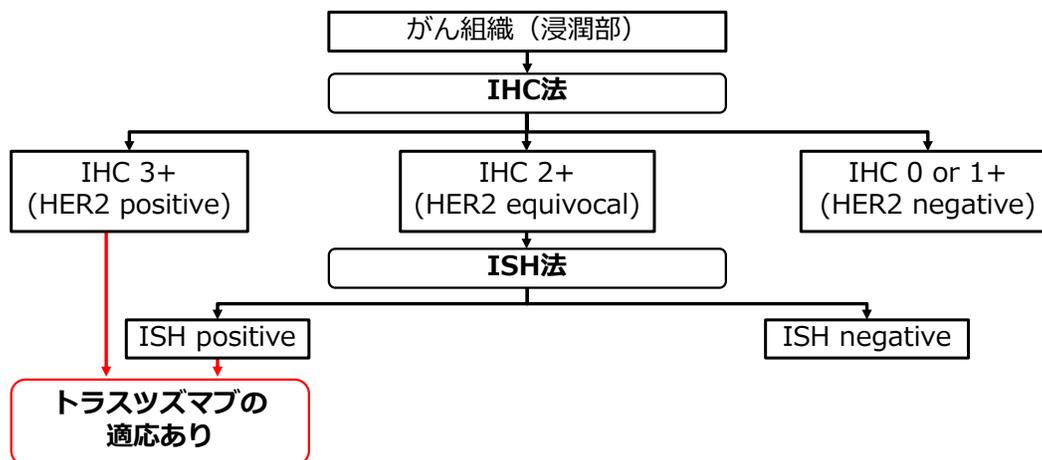


図 6 ガイドラインに基づく HER2 検査の流れ

HER2 IHC 法では、抗体を用いて細胞表面のタンパク質 (抗原) を検出する。抗体に反応した細胞膜の HER2 タンパクを DAB 色素で茶褐色に標識したのち、対比染色としてヘマトキシリンで細胞核を青色に標識する。図 7 に IHC 法による HER2 検査の具体的なアルゴリズムを示す。まず第一に H&E 染色スライドを観察してがん浸潤部を判断し、浸潤部における DAB 染色の濃さ、膜染色の連続性、陽性細胞の割合を観察する。「細胞膜に陽性染色なし」または「細胞膜の染色がある陽性細胞が 10% 以下」の場合はスコア 0、「かすかな細胞膜の染色がある陽性細胞が 10% より多い」場合はスコア 1+、「弱～中程度の完全な細胞膜の染色がある陽性細胞が 10% より多い」場合はスコア 2+、「強い完全な細胞膜の染色がある陽性細胞が 10% より多い」場合は

スコア 3+と判定される。HER2 IHC 検査では顕微鏡の 20 倍の対物レンズで 5 視野観察することが決められている。

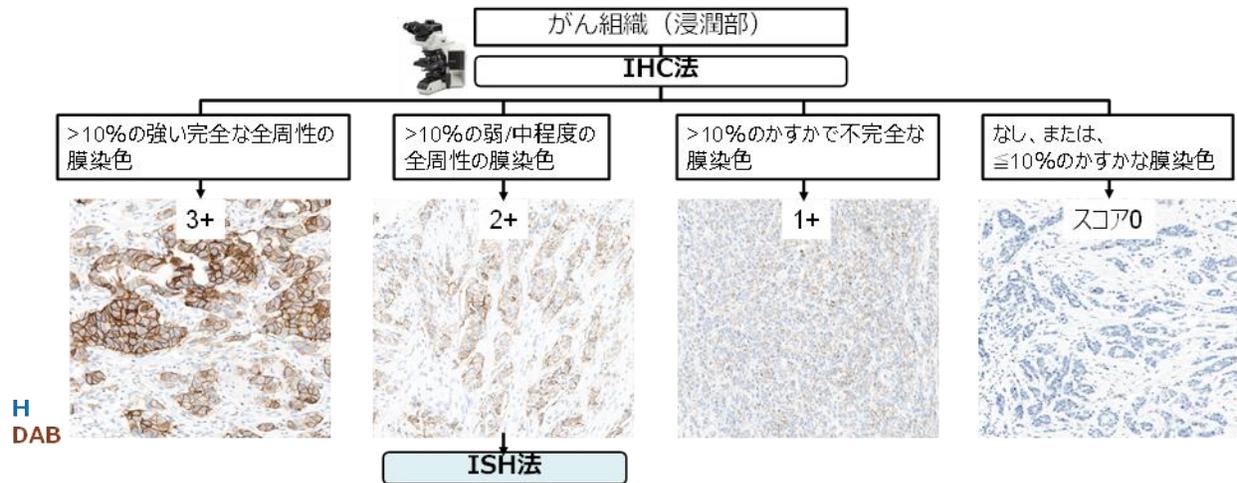


図7 HER2 IHC 検査のアルゴリズム

IHC 法では組織標本とともに、陰性、陽性コントロールとなる組織がスライドに付与され染色される (図 8)。コントロールは染色が適正に行われているかどうかの判断や、対象組織の染色の濃さを判定する際の比較のために用いられる。

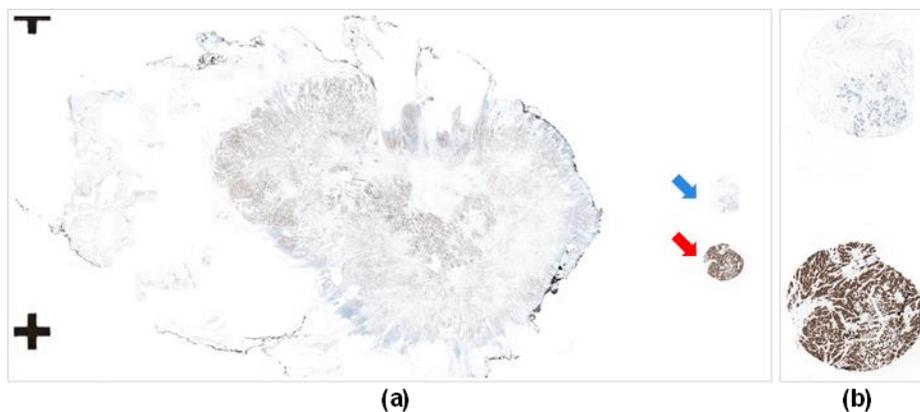


図8 スライド全体像；(a)組織標本とコントロール(青矢印：陰性、赤矢印：陽性)；(b)コントロールの拡大像

図9に ISH 法による HER2 検査の具体的なアルゴリズムを示す。ISH 法には蛍光色素を用いる FISH 法と明視野観察用の CISH 法の 2 種類がある。ここでは CISH 法を例として記載する。CISH 法では細胞核、HER2 遺伝子、HER2 遺伝子が位置する第 17 番染色体 (CEP17) の 3 つをそれぞれ青色、黒色、赤色に標識し、細胞ごとの HER2、CEP17 のシグナル数を計数する。正常細胞では HER2 シグナル、CEP17 シグナルは 2 つずつ存在するが、陽性細胞では HER2 シグナルの増

幅がみられるようになる。がん浸潤部において最低 20 個の核のシグナル数を計数し、HER2/CEP17 のシグナル比率が 2.0 以上の場合、ISH 陽性と判断される。

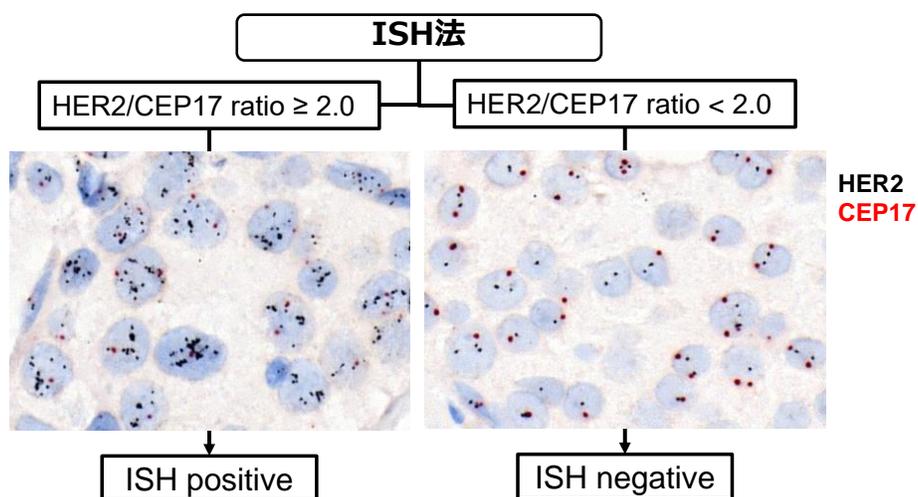


図9 HER2 ISH 検査のアルゴリズム

表 2 に IHC 法と ISH 法の比較を示す。IHC 法は染色費用や診断までに要する日数の面でメリットがあるが、病理医が経験に基づき判断するため解釈が主観的な側面もある。ISH 法では遺伝子の増幅を計数するため客観的かつ定量的な検査とされる。通常は IHC 法で 2+ となった症例に限って ISH 法による検査が実施される。

表 2 IHC 法と ISH 法の比較 [17]

	IHC	ISH
対象	HER2 タンパク	HER2 遺伝子
基準	濃さ、割合	遺伝子数
費用	¥	¥ ¥
日数	1 日	1~2 日
解釈	経験と主観	客観的
評価	半定量的	定量的

2.1.4 Whole Slide Image の利用と課題

デジタル病理は、病理診断における画像処理技術とコンピュータ技術の応用分野であり、病理診断の効率化、精度向上、診断の標準化などを目的として発展が期待される技術である [18]。WSI はバーチャル顕微鏡画像とも呼ばれ、染色した組織標本スライドをバーチャルスライドスキャナによってデジタル画像化して得られる。専用のビューアソフトウェア上での病理標本全

体の高倍率での画像観察や、画像解析の利用を可能とする (図 10)。2010 年代からデジタル病理診断システムが開発され、臨床現場での導入も始まっている。日本では病理医の不足が深刻な問題となっており、病理医不在の病院では通いの病理医が診断を行ったり、患者が別の病院へ足を運ぶこともある。WSI はネットワークを利用して遠隔地からの病理診断を可能とし、患者、病理医双方の負担軽減につながると期待される。

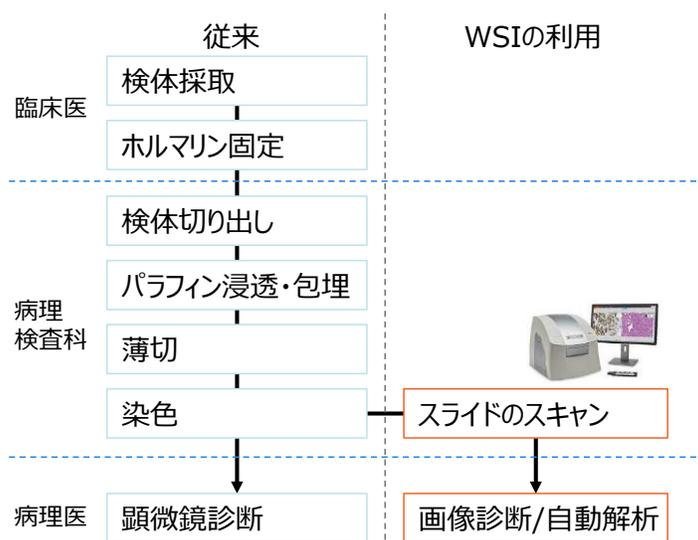


図 10 Whole Slide Image 導入の比較

ベンダーによって様々な分解能や機能を持つバーチャルスライドスキャナが販売されている。スキャナは 40 倍の対物レンズ相当で取り込みを行った場合の画像分解能が $0.25\mu\text{m}/\text{pixel}$ 以下であること、RGB 三原色等による拡大像のカラー撮影ができること等の機能要件が定められている[19]。臨床利用用のスキャナは、スキャンが高速であることが重視されていたが、近年は 1 時間または 1 日にスキャンできるスライドの枚数により重点が置かれている[20]。スキャナに一度にセットできるスライドの枚数、スキャンプロセスを中断することなくスライドをロード・アンロードする連続スキャン機能、スキャン範囲の自動検出機能やオートフォーカス機能などが評価される。

WSI は画像の読み込みを効率的に行うために、複数の解像度の画像で構成されたピラミッド構造で表現される。ベース画像レイヤーはオリジナルの解像度を持ち、各レイヤーは通常 1/4、1/2 等の係数でダウンサンプリングされる。この構造とすることで通常 25%程度ファイルサイズが抑えられ、ビューアでの観察効率が向上する。低倍率で WSI を観察する場合、解像度はコンピュータのモニターによって制限されるため、画像を最高解像度で読み込む必要がなく、ダウンサンプリングされた画像が表示される。高倍率で WSI を観察する場合、モニターには常に

小さな視野しか表示されない。そのため、画像を全体的に読み込む必要がないとして、高解像度の画像の一部を表示するように設計されている。

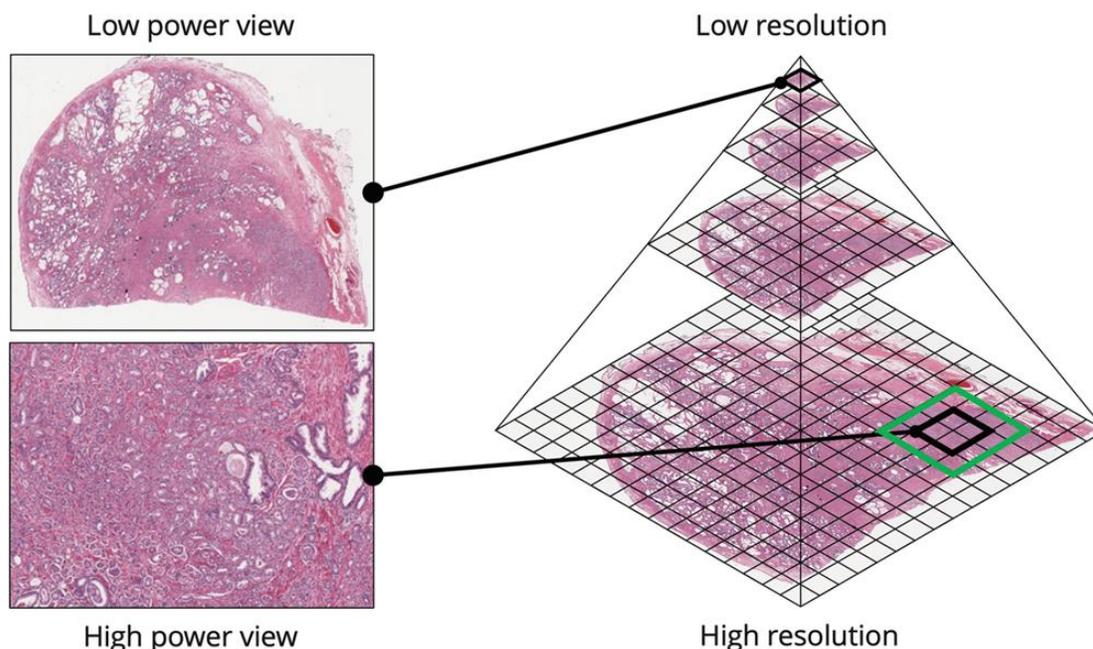


図 11 Whole Slide Image のデータ構造 [20]

HER2 IHC 検査は通常スライドの顕微鏡観察により行われるが、WSI を利用した自動解析ソフトウェアも近年選択できるようになってきている。FDA によって臨床診断への利用が認可されたものから研究用途のオープンソースのものまで、様々な自動解析ソフトウェアが存在する [6][21]。図 12 に HER2 IHC 検査用の自動解析ソフトウェアの一般的なアルゴリズムを示す。始めに入力画像にカラーアンミキシングが適用され、DAB とヘマトキシリンの 2 つの色素成分画像に分離される。続いて、分離された色素成分画像を利用して細胞検出が行われる。ヘマトキシリン画像から細胞核を検出し、検出された細胞核を起点として DAB 画像から細胞膜が検出される。次に、検出されたすべての細胞は、細胞膜の DAB の濃さの平均値に基づき、Immunoscore と呼ばれる 4 段階の濃さに分類される。Immunoscore は染色の濃さを表し、図 7 のガイドラインの定義と比較したとき、「染色無し」が 0、「かすかな染色」が 1、「中程度の染色」が 2、「強い染色」が 3 に該当する。最後に、それぞれの Immunoscore に分類された細胞の割合を式(1)により計算する。

$$P_m = T_m / (\sum T_m) \quad (1)$$

ここで、 P は陽性細胞の割合、 T は分類された細胞の数、 m は Immunoscore (0-3) をそれぞれ示す。HER2 スコアは分類された細胞の割合からガイドラインに基づき決定される。自動解析では HER2 スコアとは別に、陽性細胞の割合と Immunoscore から計算される H-score という指標も

算出される (式(2))。セミオートタイプの自動解析ソフトウェアでは、カラーアンミキシングのための色素成分ベクトルや、Immunoscore を分類するための DAB の濃さの閾値をユーザー自身で設定できる。カラーアンミキシングについては 2.2.1 節で記述する。

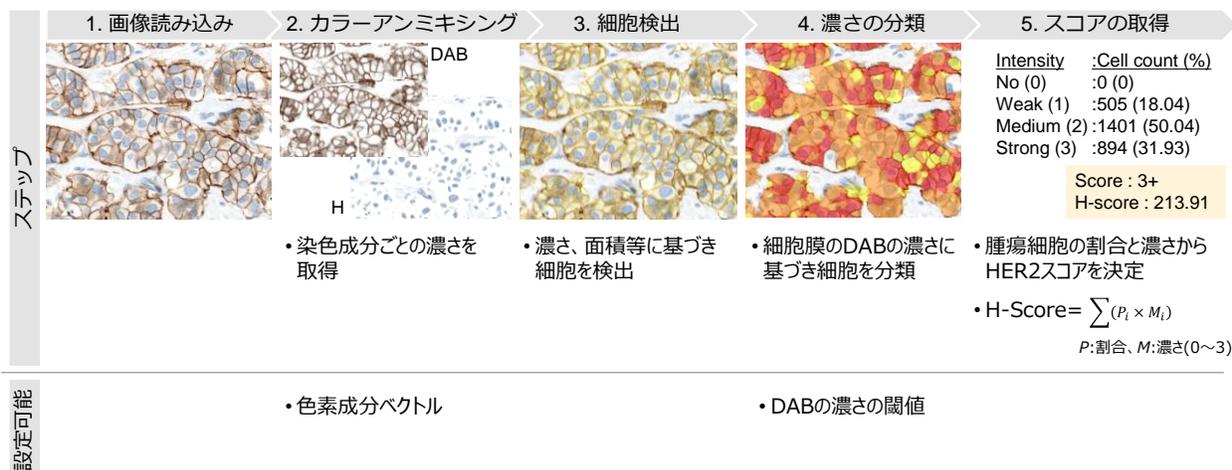


図 12 HER2 IHC 検査の自動解析アルゴリズム

$$H - score = \sum (P_m \times m) \quad (2)$$

検査の迅速化や効率化、観察者による判定のばらつきの軽減の面で利用が期待される自動解析ソフトウェアだが、普及に向けた課題として DAB の色や明るさのばらつきが挙げられる。ばらつきの要因には染色とスキャンの 2 つが挙げられる。図 13 に異なる日に染色した連続切片スライドの比較を、図 14 に 3 種類の異なるベンダーのスキャナで同一のスライドをスキャンした画像の比較を示す。まず染色工程において、アメリカでは臨床利用が認可された自動染色装置と試薬のセットを用いて病理標本スライドを染色するように決められているが、そのような管理された環境下であっても日々の染色で組織標本スライドの染色の結果が変わってしまうことが報告されている[7][8]。また、FDA 認可済みの試薬は一種類ではなく、複数のベンダーから販売されており、染色感度や特異性がそれぞれ異なるため、染色色素の色味や濃さも異なってくる。他にも、組織標本ブロック作成の手順や薄切する際の厚みなど、技術的なアーチファクトが染色性に影響するとも言われている。

さらに、WSI の場合、利用するスキャナによる色の違いも画像に現れる[7][9]。WSI の色域および階調は、各ベンダーが設定した品質基準を満たしていることが機能要件とされており[19]、ベンダー間で共通化はされていない。カメラの出力は対象物体の分光透過率、光源の分光分布、カメラセンサの分光感度によって決まり、光源、対物レンズ、カメラセンサなどのハードウェアの構成要素はベンダーや機種によって異なるため、スキャナごとの画像の RGB 値の違いにつながっている。また、ソフトウェア面でもガンマ補正やベンダー特有のカラープロファイルの

設定などによって画像の RGB 値や明るさの違いが生じる。ガンマ補正については 2.2.5 節に記載する。

DAB の染色の濃さを指標とする HER2 検査では、このような画像の違いが画像解析に影響し、判定スコアの変動につながる。そのため、臨床診断への利用が認可された自動解析ソフトウェアでは、画像の色や明るさが変わらない条件での解析に限定するために、利用できる自動染色装置や試薬、スキャナの組み合わせを限定している。このような解析ソフトの導入には臨床の病理検査ワークフローの変更が必要となる場合もあり、自動解析の臨床利用は進んでいない。また、導入に向けた各種装置の変更が多く、病理診断に影響するのに対し、解析ソフトウェアは一部のタンパク質の解析機能に限られることも導入の妨げとなっている。

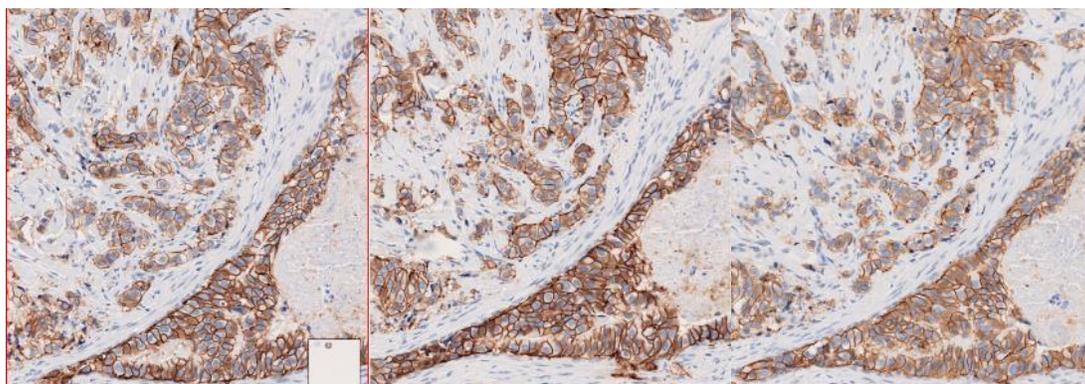


図 13 異なる日に染色した3枚のスライドの画像の比較

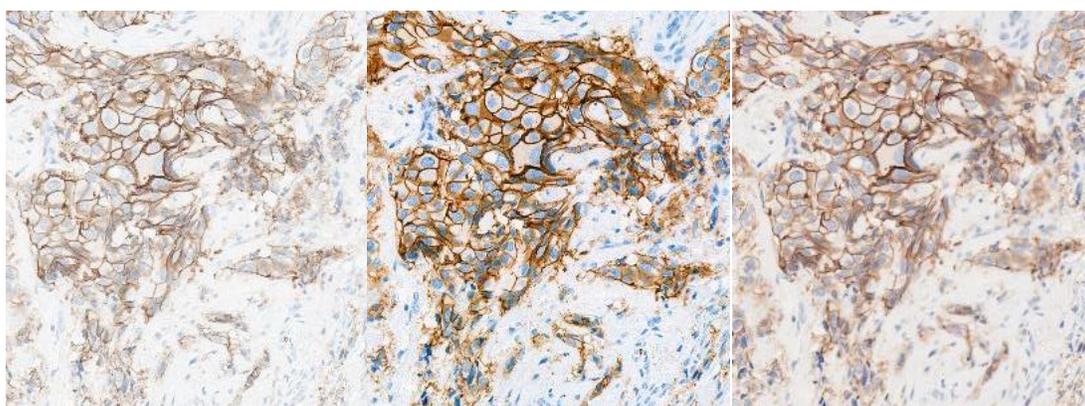


図 14 3種類のスキャナで同一のスライドをスキャンした画像の比較

2.2 基礎理論

2.2.1 カラーアンミキシング

多重染色された組織標本スライドをデジタル化した RGB 画像を色素ごとの染色強度の成分に分離するために、光学モデルに基づくカラーアンミキシング法がよく使われる。標本を透過した光の検出強度は、均質な吸収媒体を通過する光の輸送式として知られるランバート・ベールの法則で表される。

$$I = I_0 \exp(-VC) \quad (3)$$

ここで、 $I_0 = (I_{0R} I_{0G} I_{0B})^t$ は入射光の RGB 強度ベクトル、 $I = (I_R I_G I_B)^t$ は標本を通過した光の RGB 強度ベクトル、 V は色素ごとの RGB 吸収係数の単位ベクトルを列とする色素成分行列、 C は染色の濃さをそれぞれ表す。なお、ここでのベクトル行列表記では、簡略化のため空間座標は省略している。RGB 各チャンネルの強度 I_0 は染色の濃さとは非線形の関係であるため、染色成分の分離に使用することはできない。

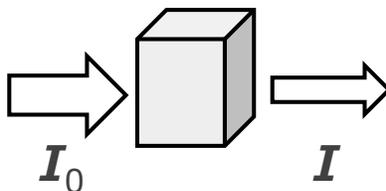


図 15 ランバート・ベールのイメージ

一方で、RGB 各チャンネルの光学濃度 $O = (O_R O_G O_B)^t$ は、標本を透過した光の対数に相当し、染色の濃さに比例する式として定義される。

$$O = -\log_{10}(I \oslash I_0) = VC \quad (4)$$

ここで、 \oslash はアダマール演算子を表す。各波長の光学濃度は、標本のないガラス領域から選択した I_0 で正規化し、対数をとることで、染色の濃さと線形の関係となる。 V が得られている場合、 C は式(5)より算出される。

$$C = V^+ O \quad (5)$$

ここで、 \cdot^+ はムーア・ペンローズの擬似逆行列演算子を表す。染色成分の数は 3 つ未満と仮定するとき、一意な解を得ることが出来る。染色成分が 3 つの場合、式(5)は式(6)で表現される。

$$C = V^{-1} O \quad (6)$$

複数 (i 個)の吸収物体がある場合、式(4)は式(7)で表される。

$$\mathbf{O} = \sum_i^p \mathbf{o}_i = \sum_i^p \mathbf{V}_i \mathbf{C}_i \quad (7)$$

標本が DAB とヘマトキシリンの 2 つの成分で染色された IHC 法による組織標本の場合、式(7)は式(8)で表現される。組織標本画像の各ピクセルの光学濃度は 2 つの純粋な染色成分の足し合わせとして特徴づけられる。

$$\mathbf{O}_{\text{tissue}} = \mathbf{V}_{\text{DAB}} \mathbf{C}_{\text{DAB}} + \mathbf{V}_{\text{H}} \mathbf{C}_{\text{H}} \quad (8)$$

ここで、 \mathbf{V}_{DAB} は 3×1 の DAB の色素成分ベクトル、 \mathbf{V}_{H} は 3×1 のヘマトキシリンの色素成分ベクトル、 \mathbf{C}_{DAB} 、 \mathbf{C}_{H} はそれぞれの染色の濃さを表す。

2.2.2 非負値行列因子分解

非負値行列因子分解 (Non-negative matrix factorization : NMF) は、非負の値のデータを要素とする行列を、基底行列と係数行列に分解する手法である [22][23]。データを行列で表現することができれば適用可能な手法であるため、文書データや音声分離、画像の分離など、様々な分野に応用されている。

ある $I \times J$ の観測行列 \mathbf{X} について、 $I \times K$ の非負の基底行列 \mathbf{T} 、 $K \times J$ の非負の係数行列 \mathbf{V} の積によって式(9)により近似を行う。

$$\mathbf{X} \cong \mathbf{TV} \quad (9)$$

近似の際には、 \mathbf{X} と \mathbf{TV} の距離を目的関数として定義し、これを最小化することを考える。NMF の目的関数として、 β ダイバージェンスの利用が考案されている。 β ダイバージェンスは式(10)で表される。

$$D_{\beta}(y|x) = \frac{y^{\beta}}{\beta(\beta-1)} + \frac{x^{\beta}}{\beta} + \frac{yx^{\beta-1}}{\beta-1} \quad (10)$$

$D_{\beta}(y|x)$ は y の x に対する疑距離を表し、 $\beta=0$ のとき板倉斎藤擬距離、 $\beta=1$ のとき I ダイバージェンス、 $\beta=2$ のときユークリッド距離となる。 β ダイバージェンスを規範とした NMF の目的関数は式(11)で表される。

$$D_{\beta}(\mathbf{X}|\mathbf{TV}) = \sum_{i,j} D_{\beta}(x_{i,j} | \sum_k t_{i,k} v_{k,j}) \quad (11)$$

ここで、 x 、 t 、 v は行列 \mathbf{X} 、 \mathbf{T} 、 \mathbf{V} の要素を表す。 $D_{\beta}(\mathbf{X}|\mathbf{TV})$ が最小となるように \mathbf{T} 、 \mathbf{V} を推定する。

れ、表色系の統一基準として決められている[26]。図 17(a)に CIE 測色標準観察者の XYZ 等色関数を示す。RGB 表色系から XYZ 表色系への変換は RGB から XYZ への変換行列 \mathbf{N} を用いて式 (14)で得られる。

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \mathbf{N} \begin{pmatrix} R \\ G \\ B \end{pmatrix} \quad (14)$$

試料から反射される光の分布が得られたとき、三刺激値 X 、 Y 、 Z は式(15)によって求められる。

$$X = k \int_{340}^{830} T(\lambda)S(\lambda)\bar{x}(\lambda)d\lambda$$

$$Y = k \int_{340}^{830} T(\lambda)S(\lambda)\bar{y}(\lambda)d\lambda \quad (15)$$

$$Z = k \int_{340}^{830} T(\lambda)S(\lambda)\bar{z}(\lambda)d\lambda$$

$$k = 1 / \int_{340}^{830} S(\lambda)\bar{y}(\lambda)d\lambda$$

ここで、 $T(\lambda)$ は波長 λ における試料から反射される光の分光分布、 $S(\lambda)$ は照明光源の分光分布、 $\bar{x}(\lambda)$ 、 $\bar{y}(\lambda)$ 、 $\bar{z}(\lambda)$ はそれぞれ XYZ 等色関数を示す。係数 k は三刺激値を正規化するもので、完全な白色拡散板では Y は 1.0 の値を持つことになる。光源 $S(\lambda)$ には CIE の基準光である D65 光源 (図 17(b)) や D50 光源が用いられることが多い。

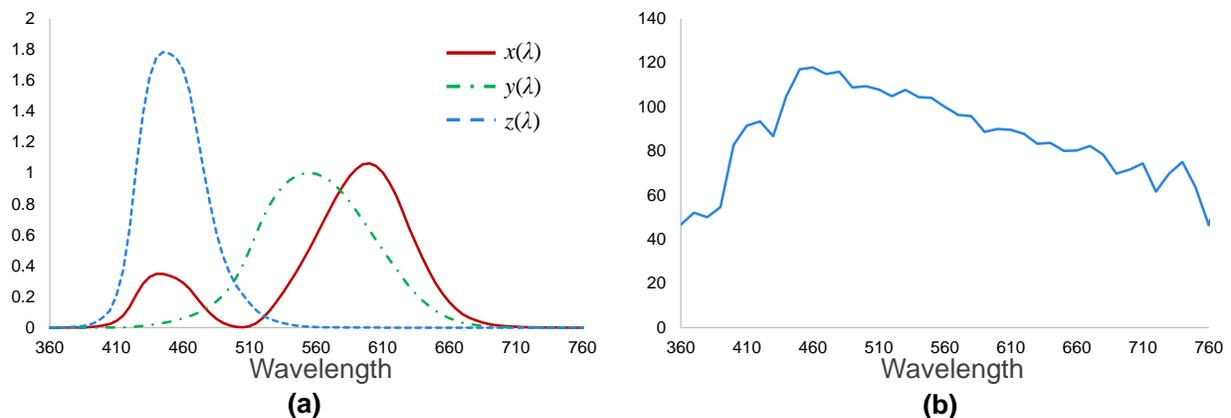


図 17 スペクトルデータ；(a) XYZ 等色関数；(b) D65 光源

XYZ 表色系では、Y 成分が輝度になるように設計されている。三刺激値は虚色の混色量のため、XYZ の値を見てどのような色かを判断することは難しい。そのため、三刺激値の比率を用いた x 、 y で色度が表現される。 x 、 y は式(16)で得られる。図 18 に x 、 y で色度座標を表現した xy 色度図を示す。

$$x = \frac{X}{(X + Y + Z)} \quad (16)$$

$$y = \frac{Y}{(X + Y + Z)}$$

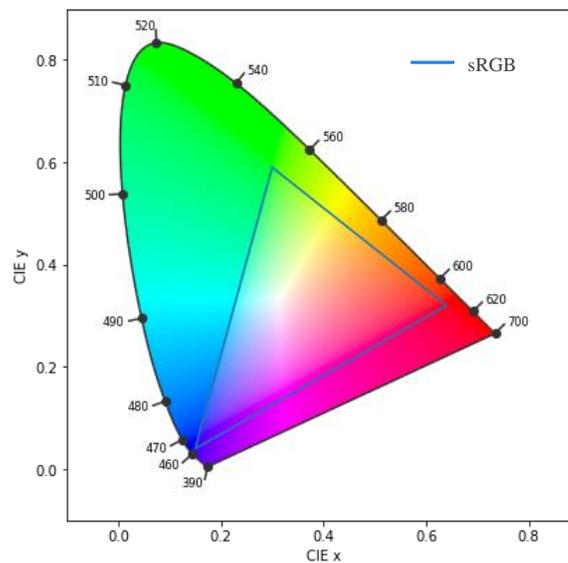


図 18 xy 色度図

図 18 中の青色の三角で示す空間は sRGB 表色系を示す。sRGB は CIE が定めた国際標準規格の表色系であり、ディスプレイやデジタルカメラ、プリンタなどで sRGB 表色系に準拠した色で統一することで、どの機器でも一貫した色を表現できるようにしている。sRGB 表色系の範囲外の色は再現ができない。

2.2.4 色差

画像間の色の違い(色差)を表現するためには、 $L^*a^*b^*$ 表色系が用いられる。 $L^*a^*b^*$ 表色系では、 L^* は明度、 a^* と b^* は色相と彩度を表す。 a^* 、 b^* は色の方向を示し、 a^* は赤方向、 $-a^*$ は緑方

向、 b^* は黄方向、 $-b^*$ は青方向を示す(図19)。 $L^*a^*b^*$ 表色系はXYZ三刺激値を用いて式(17)で表される。

$$L^* = 116f\left(\frac{Y}{Y_0}\right) - 16$$

$$a^* = 500\left[f\left(\frac{X}{X_0}\right) - f\left(\frac{Y}{Y_0}\right)\right]$$

$$b^* = 200\left[f\left(\frac{Y}{Y_0}\right) - f\left(\frac{Z}{Z_0}\right)\right] \quad (17)$$

$$f(x) = \begin{cases} x^{1/3} & x > 0.00886 \\ 7.787x + \frac{16}{116} & x \leq 0.00886 \end{cases}$$

ここで、 X_0, Y_0, Z_0 は基準白色点の三刺激値に相当する。 $L^*a^*b^*$ 表色系での2つの色 (L_1^*, a_1^*, b_1^*) と (L_2^*, a_2^*, b_2^*) があったとき、色差 dE^* はユークリッド距離に比例する。

$$dE^* = \sqrt{(L_1^* - L_2^*)^2 + (a_1^* - a_2^*)^2 + (b_1^* - b_2^*)^2} \quad (18)$$

得られた dE^* が1.0から2.0のとき、経験豊富な観察者だけが知覚的な違いを認識でき、1.0以下では観察者が違いを認識できないとされる[28]。

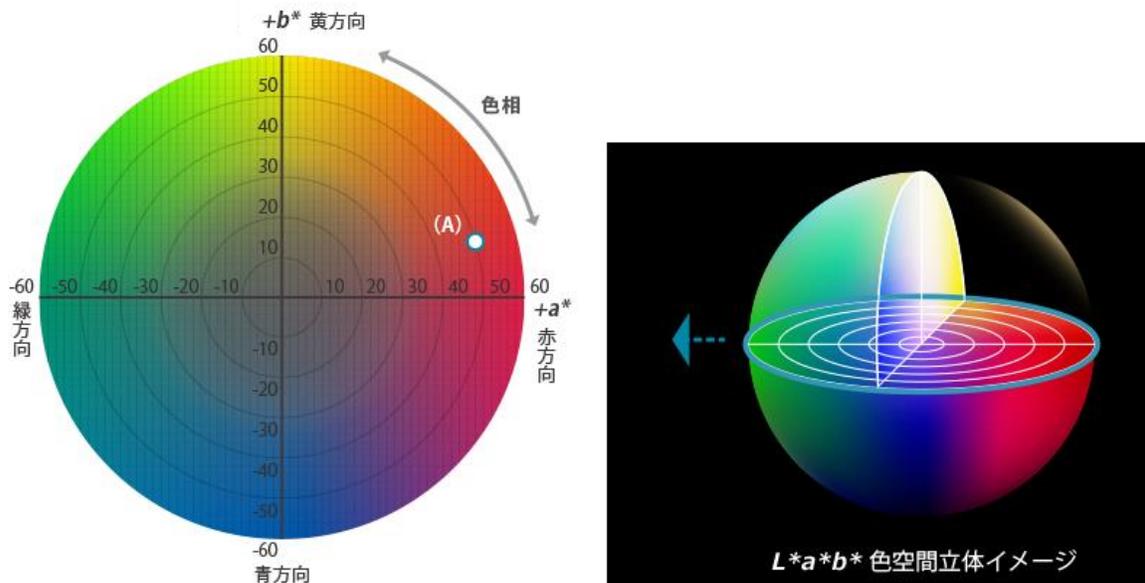


図 19 L*a*b*表色系色度図 [27]

2.2.5 ガンマ補正

画像の色がディスプレイで正しく表示されるように、ディスプレイのガンマ値に応じた色の補正を行うことをガンマ補正という。画素の色をそのまま入力値としてディスプレイに表示すると、ディスプレイの特性に応じて色にひずみが生じ、表示される明るさや彩度が変わってしまう。これを避けるため、ディスプレイの持つガンマ値を打ち消すようにデータに補正を加え、表示された状態が本来の色に近づくようにする。ガンマ補正は式(19)で定義される。

$$V_{out} = (V_{in})^\gamma \quad (19)$$

ここで、 $V^{in} = (R^{in}, G^{in}, B^{in})$ は入力値、 $V^{out} = (R^{out}, G^{out}, B^{out})$ はガンマ補正後の出力値、 γ はガンマ補正値をそれぞれ表す。ガンマ値が 1.0 のときはリニア変換を意味する。ガンマ値が 1.0 より小さい場合は画面全体が暗くなり、輝度が高い部分の違いが協調される。1.0 より大きい場合は画面全体が明るくなり、輝度が低い部分の違いが強調される。

ディスプレイのガンマ値は 2.2 が使用されることが多い。リニア RGB 画像をそのままディスプレイに表示すると本来よりも中間調が暗く沈んだ表現となってしまうが、ガンマ補正によって中間調を明るくした画像を入力して入出力の関係をリニアに近づけることで、本来の色に近い表示像が得られる。

2.2.6 色のキャリブレーション

カラーカメラやスキャナなどのデバイスで環境の色を正確に表現するには、測色表色系の観点からカメラの RGB 出力をキャリブレーションする必要がある。色のキャリブレーションは、デバイスで測定された RGB 値をデバイスに依存しない知覚的な色空間にマッピングするのに必要な色空間変換を計算するものである。一般的には、測色値が既知であり一定数のカラーパッチを持つカラーチャートを撮影することで実行される。カラーチャートの各パッチの真の XYZ 値とデバイスの RGB 値とをマッピングする係数 a_{11} - a_{33} の推定が必要となる。

$$\begin{aligned} X &= a_{11}R + a_{12}G + a_{13}B \\ Y &= a_{21}R + a_{22}G + a_{23}B \end{aligned} \quad (20)$$

$$Y = a_{31}R + a_{32}G + a_{33}B$$

カラーチャートのパッチが n 個ある場合、式(20)は式(21)で表される。

$$\mathbf{G}_R = \mathbf{G}_S \mathbf{M} \quad (21)$$

ここで、 \mathbf{G}_R はカラーチャートの基準となる三刺激値を要素とする $n \times 3$ の行列、 \mathbf{G}_S はデバイスで取得した画像の RGB 値を要素とする $n \times 3$ の行列、 \mathbf{M} は 3×3 の色変換行列を表す。色変換行列 \mathbf{M} は式(22)で得られる。

$$\mathbf{M} = \mathbf{G}_S^{-1} \mathbf{G}_R \quad (22)$$

ここで、 \cdot^{-1} は逆行列を表す。 \mathbf{M} が 3×3 の正方行列の場合、逆行列を計算することで一意の解が得られる。それ以外の場合は、ムーア・ペンローズの擬似逆行列を計算し、 $\mathbf{M} = \mathbf{G}_S^{\dagger} \mathbf{G}_R$ と変換することで解が得られる。

式(20),(21)の線形変換は、多項式変換[29][30]の次数が 3 の場合を意味する。実際に色のキャリブレーションを行う際には、より高い次元が選択されることが多いとされる。多項式変換を用いる場合、式(21)は式(23)で表現される。

$$\mathbf{G}_R = [\theta_m \mathbf{G}_S] \mathbf{M} \quad (23)$$

ここで、 θ_m は多項式変換の変換行列を表す。例として、次数 10 までの変換行列は表 3 で定義される。

表3 多項式変換の次数と変換行列

次数	変換行列 θ_m
3	[R G B]
4	[R G B 1]
5	[R G B RGB 1]
6	[R G B RG RB GB]
7	[R G B RG RB GB 1]
8	[R G B RG RB GB RGB 1]
9	[R G B RG RB GB R ² B ² G ²]
10	[R G B RG RB GB R ² B ² G ² 1]

式(22)を解くに当たり、色変換では白とグレースケールを保持することが重要であるとの主張のもとで、white-point preserved least-square (WPPLS)が提案されている[31][32]。WPPLSでは式(21)に制約を追加した式(24)で表現する。

$$\mathbf{G}_R = \mathbf{G}_S \mathbf{M} \text{ and } \mathbf{uM} = \mathbf{u} \quad (24)$$

ここで、 $\mathbf{u}=(1,1,1)^t$ であり、色変換行列 \mathbf{M} で変換した後も白色性が維持されることを制約とする。WPPLS はカラーカメラの色のキャリブレーションを目的に開発された手法であり、色変換後の白色性を維持することを優先しつつ、目的とする三刺激値と測定値の間の二乗誤差を小さくする効果を報告している。

本技術は、携帯電話用カメラのカラーキャリブレーションを行う研究[33]や、ノイズの少ない画像再現を目的としてデジタル画像センサーのカラーフィルターの配列パターンを設計する研究[34]、ハイパースペクトルカメラで撮影した画像のデモザイク処理を目的とした研究でのカメラの内部処理への活用[35]などの応用が報告されている。

式(24)を解くには、まず式(21)を解くこととなる。目的とする三刺激値と測定値の色を合わせることが、式(25)を最小とする変換行列 \mathbf{M} を求めることを意味する。

$$I = \|\mathbf{G}_S \mathbf{M} - \mathbf{G}_R\| \quad (25)$$

ここで、 $\|\cdot\|$ はL2ノルムを示す。 I^2 は式(26)で表される。

$$J = I^2 = \mathbf{M}^t \mathbf{G}_S^t \mathbf{G}_S \mathbf{M} - 2\mathbf{M}^t \mathbf{G}_S^t \mathbf{G}_R + \mathbf{G}_R^t \mathbf{G}_R \quad (26)$$

式(26)を \mathbf{M} について微分し、式(28)により、式(25)を最小とする \mathbf{M} を求める。

$$\frac{\delta J}{\delta \mathbf{M}} = 2\mathbf{G}_S^t \mathbf{G}_S \mathbf{M} - 2\mathbf{G}_S^t \mathbf{G}_R = 0 \quad (27)$$

$$\mathbf{M} = [\mathbf{G}_S^t \mathbf{G}_S]^{-1} \mathbf{G}_S^t \mathbf{G}_R \quad (28)$$

次に、WPPLS について考える。WPPLS では式(24)のように、白色性を維持する制約を加えたうえで、式(25)を最小化することを目指す。そこで、ラグランジュの未定乗数法を利用し、式(29)で表現する。

$$J = \mathbf{M}^t \mathbf{G}_S^t \mathbf{G}_S \mathbf{M} - 2\mathbf{M}^t \mathbf{G}_S^t \mathbf{G}_R + \mathbf{G}_R^t \mathbf{G}_R + \lambda(\mathbf{M}^t \mathbf{u} - 1) \quad (29)$$

ここで、 \mathbf{u} は背景色を表す列ベクトルであり、白色を表現するために $(1,1,1)^t$ に等しいとする。式(29)をラグランジュ乗数 λ について微分すると、式(30)で表される。

$$\frac{\delta J}{\delta \lambda} = \mathbf{M}^t \mathbf{u} - 1 = 0 \rightarrow \mathbf{M}^t \mathbf{u} = 1 \quad (30)$$

これは、 J の定常点を見つけるとき、 $\mathbf{M}^t \mathbf{u} = 1$ となり、白色を維持しながら最小二乗を解く制約となる。また、式(29)を \mathbf{M} について微分すると、式(31)で表される。

$$\frac{\delta J}{\delta \mathbf{M}} = 2\mathbf{G}_S^t \mathbf{G}_S \mathbf{M} - 2\mathbf{G}_S^t \mathbf{G}_R + \lambda \mathbf{u} = 0 \quad (31)$$

式(29)、式(31)より、 \mathbf{M} について解くと、式(32)が得られる。

$$\mathbf{M} = [\mathbf{G}_S^t \mathbf{G}_S]^{-1} \mathbf{G}_S^t \mathbf{G}_R + \frac{(1 - \mathbf{G}_R^t \mathbf{G}_S [\mathbf{G}_S^t \mathbf{G}_S]^{-1} \mathbf{u})}{(\mathbf{u}^t [\mathbf{G}_S^t \mathbf{G}_S]^{-1} \mathbf{u})} [\mathbf{G}_S^t \mathbf{G}_S]^{-1} \mathbf{u} \quad (32)$$

2.3 関連研究

2013–2015 年にかけて、主要なスキャナベンダーや有識者を集めたワーキンググループを International Color Consortium (ICC) が主催し、WSI の色の共通化に向けて協議が行われた [36]。ワーキンググループでは、WSI の再現性や安定性を高めるために、カラーチャートを使ったキャリブレーションを各ベンダーとともにスキャナ内部で行うことが FDA のガイドラインとしてまとめられた [37]。しかしながら色の共通化に関しては、ベンダーごとにユーザの求める色に合わせたい、統一化を図ろうとするとデバイスの値段が上がってしまうためそれを避けたいなどの意見があり、現在まで統一には至っていない。

スキャナによる WSI の色のばらつきの補正に関しては、H&E 染色を対象に、WSI 用のカラーチャートスライドを用いて画像の色をキャリブレーションする方法を我々の研究グループの先行研究として報告している[38]。他にも、IHC 法に特化した色を選択した独自のカラーチャートスライドによる補正効果の検証[39]や、商用のカラーチャートスライドを用いた効果の検証[40]が報告されている。これらの研究では通常最小二乗法を解くことで補色変換行列を算出しており、2.2.6 節に記載した WPPLS のデジタル病理分野での利用は報告されていない。また、スキャナの色補正に限定しており、染色とスキャナの両方のばらつきがある場合の有効性は検証されておらず、自動解析を適用した場合の結果についても検討は行われていない。

染色のばらつきの補正に関しては、H&E 染色の色の標準化が活発な研究の対象となっている[41]–[44]。H&E 染色は最も一般的に使われる染色方法であり、色が評価指標となることが理由である。IHC 法を対象としては、DAB 成分とヘマトキシリン成分の染色成分ごとの画像の分離を目的とした研究[45]–[47]は行われているものの、DAB 成分の濃さの補正まではターゲットになっていない。IHC 画像のヘマトキシリン成分の濃さをテンプレートに合わせて正規化する研究[48]が報告されているが、DAB 成分の濃さを補正するには、組織画像の染色状態を把握するための指標や基準となる濃さが必要となる。組織画像の染色状態を把握するためのアプローチとして、基準となるスライドを準備する研究が行われている。Eycke ら[49]は、TMA スライドを基準スライドとして準備し、バッチごとに染色した基準スライドの色空間でのプロットを合わせることで画像の色補正を行う手法を報告している。ただし、画像のプロットを合わせる方法は物理モデルに従っているわけではないため、濃さが正しく補正されているかどうかは分からない。また、組織スライドは検体や薄切位置による変化があるため、基準スライド同士が必ずしも一致するとは限らない。この組織スライドの弱点を補うものとして、化学素材を材料とした基準スライドを利用した研究の報告がある。化学素材の染色状況をみることでラボごとの染色の傾向を把握・管理することが可能となるとの考えから、H&E 染色用にはバイオポリマーフィルムを利用したスライド[50]が、IHC 法用には抗原でコーティングされたマイクロビーズを利用したスライド[51]がそれぞれ提案されている。ただし、これらのスライドは染色の品質管理を目的に作成・利用されており、濃さの補正や、スコア評価への利用は報告されていない。

第3章 キャリブレーションスライドを用いた色味・濃さの補正手法

3.1 概要

本章では、本研究にて提案する染色とスキャナによる画像のぼらつきの補正手法について記載する。提案手法では、下記の2種類のキャリブレーションスライドを用いる。

1. IHC キャリブレーションスライド (図 20)

キャリブレーションスライド (IHC Calibrator, Boston Cell Standards)は、施設やラボごとの染色の品質管理を目的に設計された[51]。ガラススライド上に 10 個のドットが存在し、各ドットは 2 つの大きさのガラス製マイクロビーズで構成される。大きいマイクロビーズ (約 7~8 μm)は抗原 (タンパク質)でコーティングされ、ドットごとに異なる抗原の濃度を有している。マイクロビーズは DAB 成分によって 10 段階の濃さに染色される。小さいマイクロビーズ (約 4 μm)はあらかじめ茶色に色づいており、大きいマイクロビーズと DAB の色味を対比する用途で存在する。本研究では大きいマイクロビーズを提案手法に利用することとし、以降、大きさの記載がないマイクロビーズは、大きいマイクロビーズを指すものとする。

2. カラーチャートスライド (図 21)

カラーチャートスライド (IAM-9C-WSI, Applied Image Inc.)は、画像解析に向けたスキャナの色キャリブレーションや標準化のために設計された。ガラススライド上に包埋された 12 色の半透明のカラーパッチと背景部分の RGB 値をスキャナの色補正に用いる。この 12 色は我々のグループの先行研究[38]に基づき、H&E 染色や IHC 法の補正に適した色が選択された。

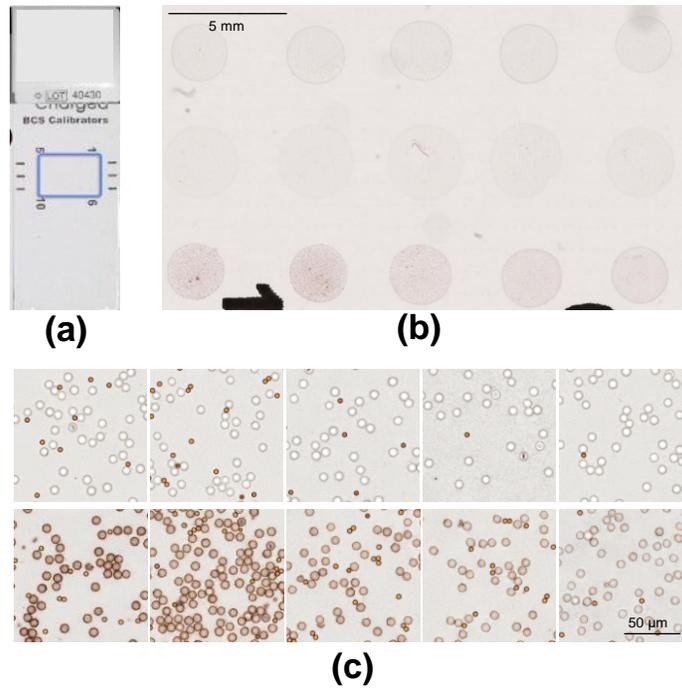


図 20 IHC キャリブレーションスライド；(a)スライド像；(b)WSI 像(低倍率)；(c) WSI 像(高倍率)レベル 1-5(上段右から左)、6-10(下段右から左)

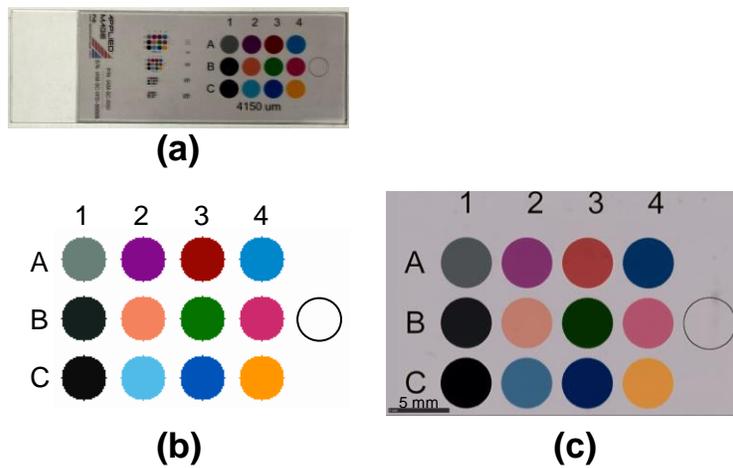


図 21 カラーチャートスライド；(a)スライド像；(b)各パッチのスペクトルデータ、CIE 等色関数、D65 光源のデータから計算した標準 RGB 画像；(c)WSI 像

図 22 は本研究の臨床のワークフローへの導入イメージを示す。FDA 認可済みの自動解析ソフトウェアを利用するためには、染色装置や試薬、スキャナの変更など、臨床のワークフローの大幅な変更が必要となることも想定される。本研究では、染色プロセスに IHC キャリブレーション

スライドを、スキャンプロセスにカラーチャートスライドを追加することで対応できるため、臨床の病理検査ワークフローの軽微な変更で導入可能な手法となっている。

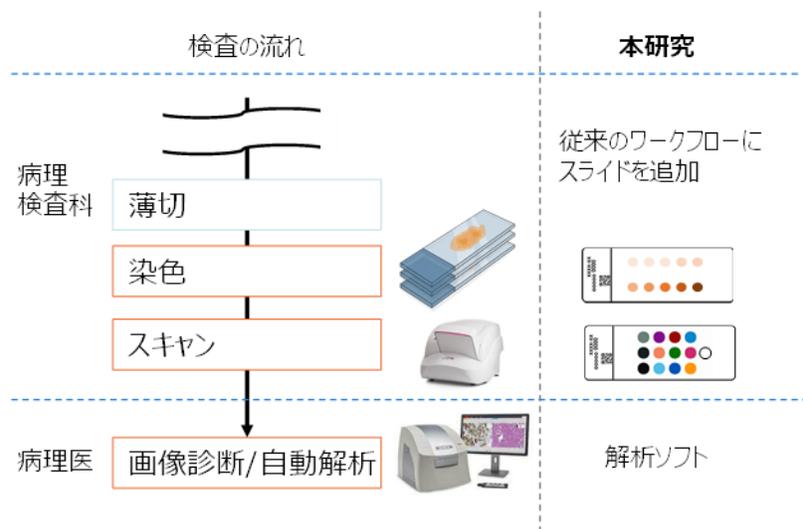


図 22 検査の流れの変更イメージ

提案手法の全体像を図 23 に示す。手法は準備ステップと評価ステップに分かれている。準備ステップでは、IHC キャリブレーションスライドと HER2 スコアが既知の組織スライドを基準データセットとして準備する。評価ステップでは、評価対象となる組織スライドと一緒に IHC キャリブレーションスライドを染色し、染色の補正に利用する。カラーチャートスライドを用いたスキヤナの色補正を行う場合は、染色したすべてのスライドとともにカラーチャートスライドをスキヤナする。スキヤナの色補正は、染色の補正の前に実施する。評価ステップで用いるスキヤナは準備ステップと同じものでも異なるものでも対応できる。

3.2 節では IHC キャリブレーションスライドを用いた染色の補正手法を、3.3 節ではカラーチャートスライドを用いたスキヤナの色補正手法を説明する。

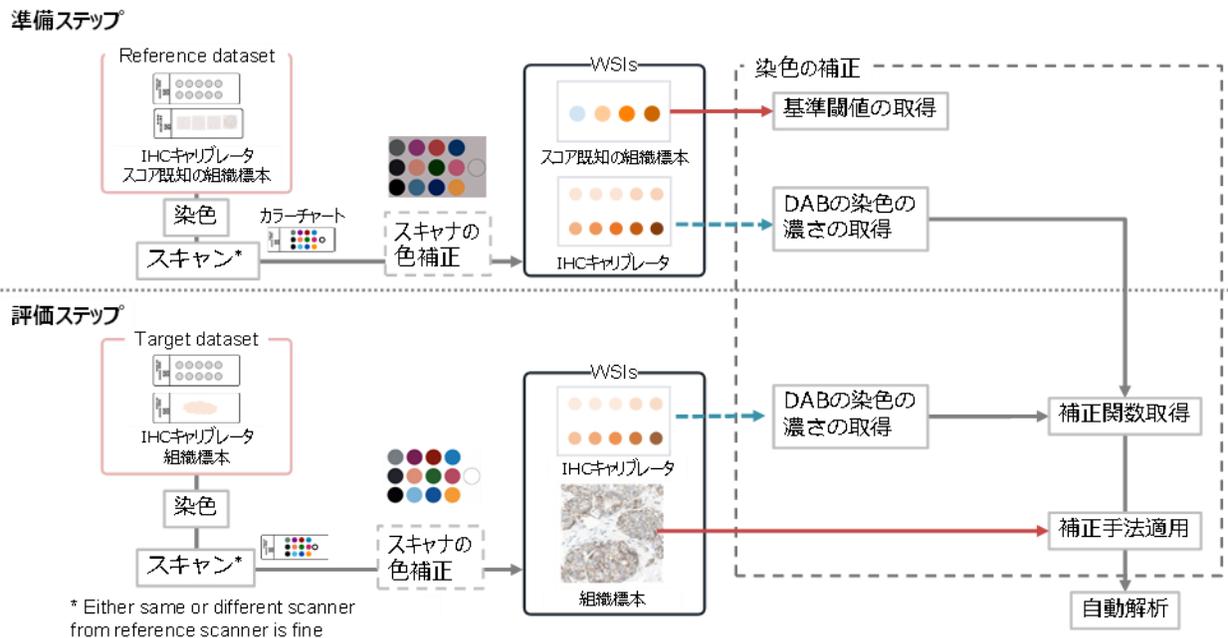


図 23 提案手法の全体イメージ

3.2 IHC キャリブレーションスライドを用いた染色の補正

3.2.1 全体像

図 24 に IHC キャリブレーションスライドを用いた染色の補正手法の全体像を示す。準備ステップでは、基準データセットとして IHC キャリブレーションスライドと各スコアを代表する組織標本スライドを準備し、WSI を取得する。Step1,2 では、IHC キャリブレーション画像を利用して DAB の色味と濃さを取得する。Step3 では組織標本画像から基準閾値を算出する。2.1.4 節で記載したように、セミオートタイプの解析ソフトウェアでは、細胞膜の濃さを分類するための閾値を設定することができる。本研究では、解析ソフトウェアで用いる閾値の算出も行った。IHC キャリブレーションスライドはもともと染色の品質管理を目的に作られていることから、キャリブレーションのレベルと HER2 スコアとの関係性は決まっておらず、キャリブレーションから閾値を決めることは出来ない。そこで、スコアが既知の組織標本を基準データセットとして準備し、組織画像から得られる DAB の濃さの分布を閾値決定に利用することとした。ここまでは評価の前の準備ステップとして事前に行う。

評価ステップとして、Step4 では評価対象となる組織標本スライドと一緒に染色した IHC キャリブレーションスライドの画像から DAB の色味と濃さを取得する。2 つのキャリブレーションの濃さの関係性を解析することで補正関数を取得し、補正に利用する。本研究では 2 つの補正手法を提案する。手法 1 では、準備ステップで求めた基準閾値を評価対象画像に合わせて補正する。手

法 2 では、評価対象画像を基準に合うように補正する。提案手法は判定に用いる解析ソフトウェアの種類によって使い分けることを想定している。内部処理がオープンで閾値の設定が可能な解析ソフトウェアでは手法 1 を、内部処理がブラックボックスな解析ソフトウェアでは手法 2 を適用する。以降の節で step ごとに分けて説明する。

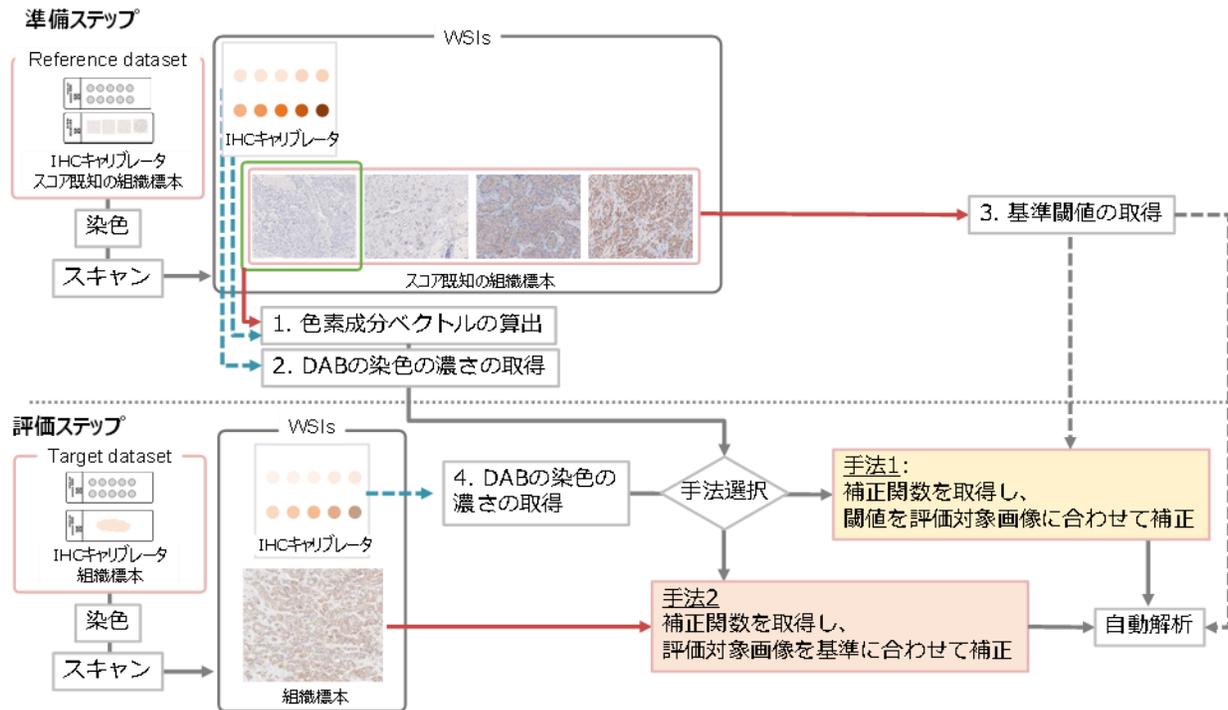


図 24 IHC キャリブレーションを用いた染色の補正の流れ

3.2.2 Step1：色素成分ベクトルの取得

画像のカラーアンミキシングに必要な色素成分ベクトル V_{DAB} 、 V_H を取得する。始めに、2.2.1 節に記載したランバート・ベールの法則に基づき、得られた RGB 画像を光学濃度に変換する。組織画像の光学濃度 O_{tissue} は式(8)のモデルで表現される。

V_H の取得には、スコア 0 の組織や、図 8 に示す陰性コントロールなど、ヘマトキシリンのみで染色されている組織画像を利用する。組織画像の平均光学濃度から単位ベクトルを算出することで V_H とした。

V_{DAB} の取得には IHC キャリブレーション画像を利用した。はじめに、ハフ変換[52][53]を用いて画像からマイクロビーズを検出する。図 25(a)のように、暗く撮影されるエッジ部分は解析から除外するため、検出したマイクロビーズごとに中央の領域の平均値を計算に用いた。IHC キャリブレーションは DAB の 1 色のみで染色されているが、染色強度が低いマイクロビーズは画像ではグレーがかかった色になっていることや (図 25(b))、光学濃度 $O_{calibrator}$ の分布に 2 方向の特徴がみら

れることから(図 26)、散乱による影響があるものと想定した。そこで、DAB とビーズによる散乱成分の 2つの項から成る式(33)のモデルで IHC キャリブレーションを表現した。

$$\mathbf{O}_{\text{calibrator}} = \mathbf{V}_{\text{DAB}}\mathbf{C}_{\text{DAB}} + \mathbf{V}_{\text{scatter}}\mathbf{C}_{\text{scatter}} \quad (33)$$

ここで、DAB 成分と散乱成分の 2つを分離する必要がある。式(33)のいずれの要素も負の値を取らないと考えられることから、2.2.2 節に記載した SSNMF を分離に利用した。 $\mathbf{V}_{\text{scatter}}$ はグレー成分を表現するために、R,G,B 同一の値を持つ 3×1 の単位ベクトル $\mathbf{V}_{\text{scatter}}=(0.58,0.58,0.58)^t$ とした。この項は染色によらず固定とする。式(33)と式(12)を比較したとき、 $\mathbf{O}_{\text{calibrator}}$ を観測行列 \mathbf{X} 、 $\mathbf{V}_{\text{scatter}}$ を既知の基底行列 \mathbf{H} 、 \mathbf{V}_{DAB} を分離の目的とする基底行列 \mathbf{T} とし、セミブラインドで信号分離を行い \mathbf{V}_{DAB} を取得した。 β ダイバージェンスは $\beta=2$ とし、ユークリッド距離を採用した。

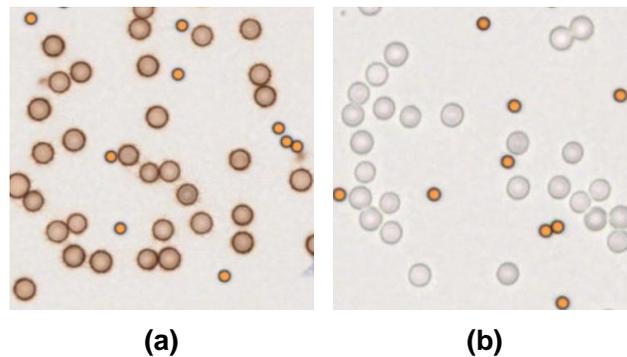


図 25 画像でのマイクロビーズの見え方；(a) 強い染色強度；(b) 弱い染色強度

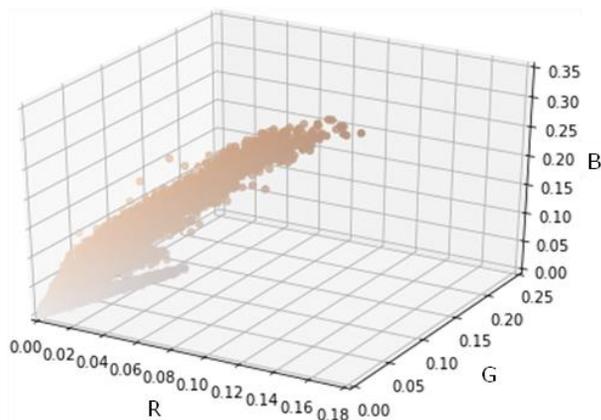


図 26 光学濃度空間での IHC キャリブレーションの分布

3.2.3 Step2,4：キャリブレーション画像からの濃さの取得

DAB の染色の濃さ C_{DAB} は、step1 で取得した V_{DAB} を用いて、キャリブレーション画像をカラーアンミキシングすることで取得する。準備ステップ、評価ステップで同様の処理を行い、 $V_{DAB,R}$, $V_{H,R}$, $C_{DAB,R}$, $V_{DAB,T}$, $V_{H,T}$, $C_{DAB,T}$ を取得した。ここで、添え字 R は基準データを、T は評価対象データを表す。

3.2.4 Step3：閾値の算出

図 27 に閾値算出の流れを示す。基準となる閾値は、濃さ 0、1、2、3 を分ける $th1$ 、 $th2$ 、 $th3$ と定義する。step1 で取得した $V_{DAB,R}$, $V_{H,R}$ を用いて、カラーアンミキシングを基準組織画像に適用し、分離された DAB 成分の濃さからスコアごとの確率密度 $p0$ 、 $p1$ 、 $p2$ 、 $p3$ を取得する。閾値の決定方法としては、各確率密度の交点値を用いるのがわかりやすいと考えた。 $p0$ と $p1$ は正規分布に近いので、閾値 $th1$ は $p0$ と $p1$ の交点値として決定することが出来る。しかし、 $p2$ と $p3$ は範囲が広く分布の重複が多いため、交点値を決定することが難しい。そこで、確率密度の差をとることで、上位の一つの濃さに対応する濃さのみを含むと想定される新しい分布を生成する。

スコア 0 の分布は他のスコアの濃さを含まないため、もともとの $p0$ で表される。スコア 1+ は $p1$ から $p0$ を減算することで表される。同様に、スコア 2+ を表すには $p2$ から $p0$ と $p1$ の差分を取り、スコア 3+ を表すには $p3$ から $p0$ 、 $p1$ 、 $p2$ の差分を取る。上位の濃さを含む分布を取得する目的においては、濃さが濃い部分が重要であるため、負の値となった場合は 0 として切り捨てて処理した。得られた各分布はやや非対称性があることから、式(34)のジョーンソン SU 分布 [54] で近似を行い、新しい分布を 4 つ取得した。そして、得られた分布の交点値を閾値 $th1$ 、 $th2$ 、 $th3$ として設定した。

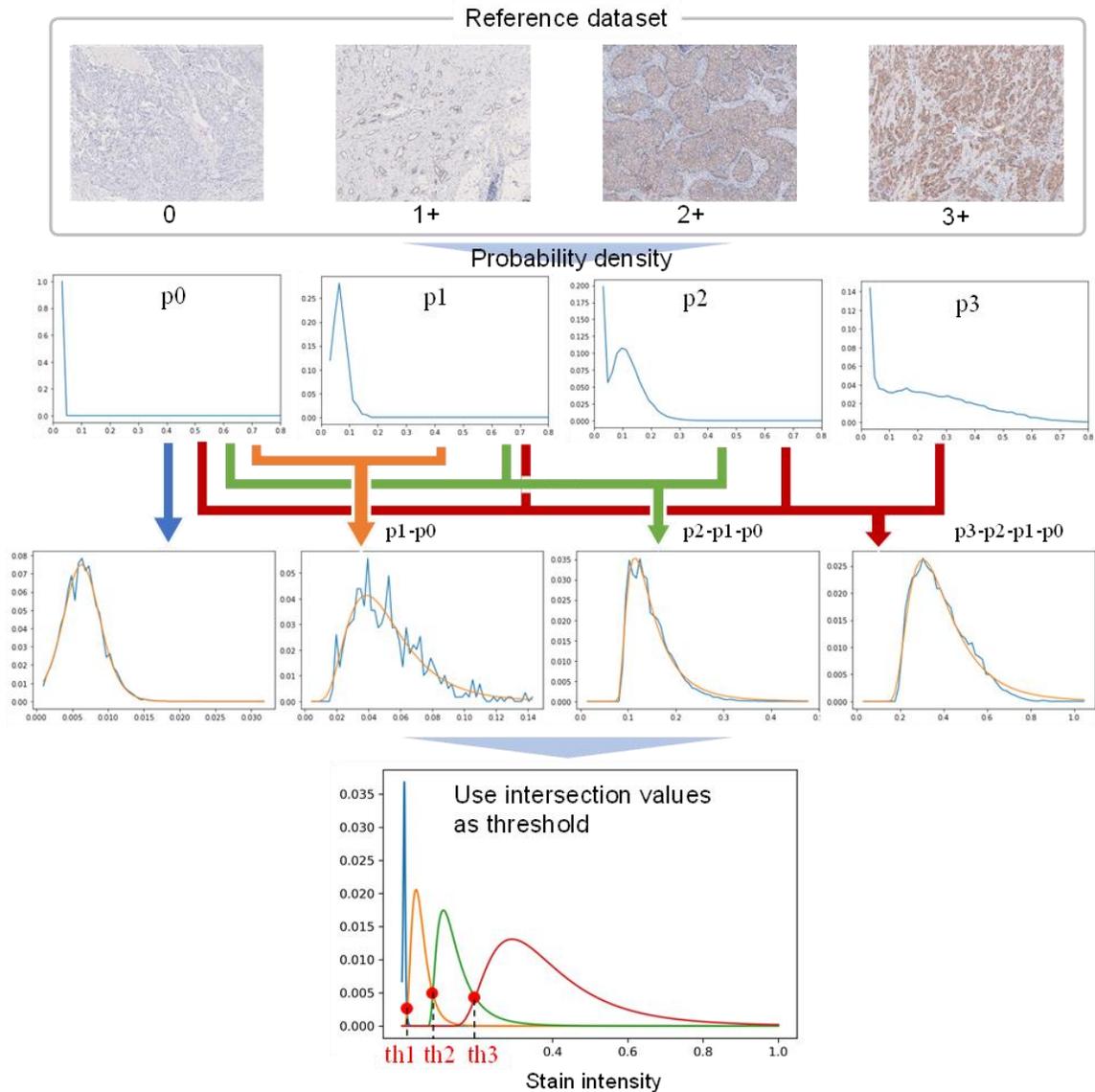


図 27 閾値の算出の流れ

$$f(x) = \lambda \exp \left[-\frac{1}{2} (\gamma + \delta \sin^{-1} x)^2 \right] \quad (34)$$

3.2.5 手法 1：閾値の補正

図 28 に手法 1 の流れを示す。手法 1 では、評価対象となる組織画像に合うように基準閾値の補正を行う。Step2,4 で得られた $C_{DAB,R}$, $C_{DAB,T}$ の関係性を解析することで補正関数 $f(C)$ を取得する。IHC キャリブレーションのレベルごとに、中央値前後の 50 個のビーズを選択し、そのビーズの濃さを用いた。リファレンスと対象のキャリブレーションそれぞれで同様の処理を行い、新たな

DAB の濃さの行列 $C'_{DAB,R}$ 、 $C'_{DAB,T}$ を得た。そして、線形回帰によって、 $|C'_{DAB,T} - f(C'_{DAB,R})|^2$ を最小とする補正関数 $f(C)$ を取得した。

得られた補正関数 $f(C)$ は、基準閾値を補正するために用いる。例えば、評価対象用の閾値 $th1'$ は $th1' = f(th1)$ で得られる。補正後の閾値 $th1'$ 、 $th2'$ 、 $th3'$ は、補正を行っていない元画像を評価する際に解析ソフトウェアで用いられる。また、画像解析ソフトのカラーアンミキシング用の色素成分ベクトルは対象データに対応した $V_{DAB,T}$ 、 $V_{H,T}$ を利用する。

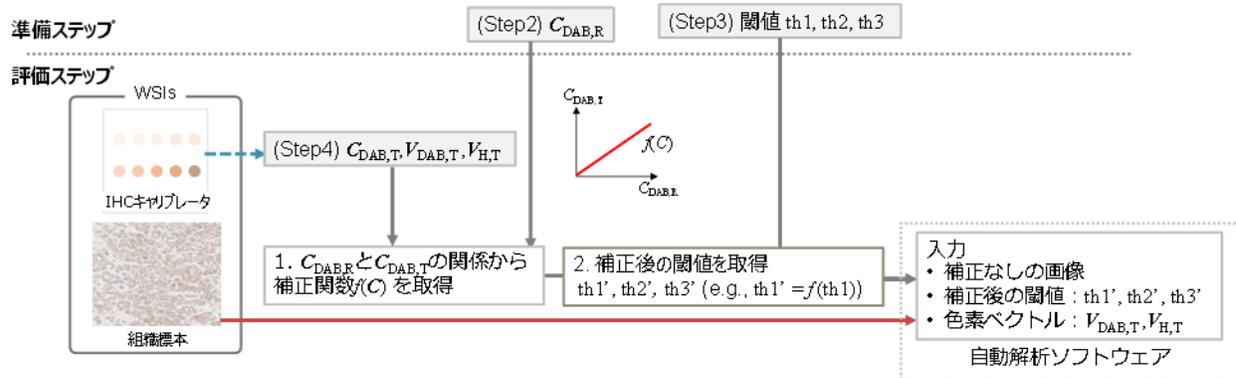


図 28 手法 1 の流れ

3.2.6 手法 2：画像の補正

図 29 に手法 2 の流れを示す。手法 2 では、評価対象の組織画像の DAB の色味と濃さを基準に合うように補正する。手法 1 と同様の手順で $C'_{DAB,R}$ 、 $C'_{DAB,T}$ を取得し、線形回帰によって、 $|C'_{DAB,R} - g(C'_{DAB,T})|^2$ を最小とする補正関数 $g(C)$ を取得する。

次に、評価対象画像の色補正を行う。評価用のデータセットから取得した $V_{DAB,T}$ 、 $V_{H,T}$ を用いて組織画像のカラーアンミキシングを行い、DAB とヘマトキシリンの濃さを取得する。HER2 IHC 検査では DAB の濃さが評価指標であるため、得られた DAB の濃さに補正関数 $g(C)$ を適用し、濃さの補正を行う。カラーミキシングの際には基準データセットから取得した $V_{DAB,R}$ 、 $V_{H,R}$ を用いることで、DAB の色味も合わせて補正する。補正後の画像は基準閾値 $th1$ 、 $th2$ 、 $th3$ とともに解析ソフトウェアで評価される。

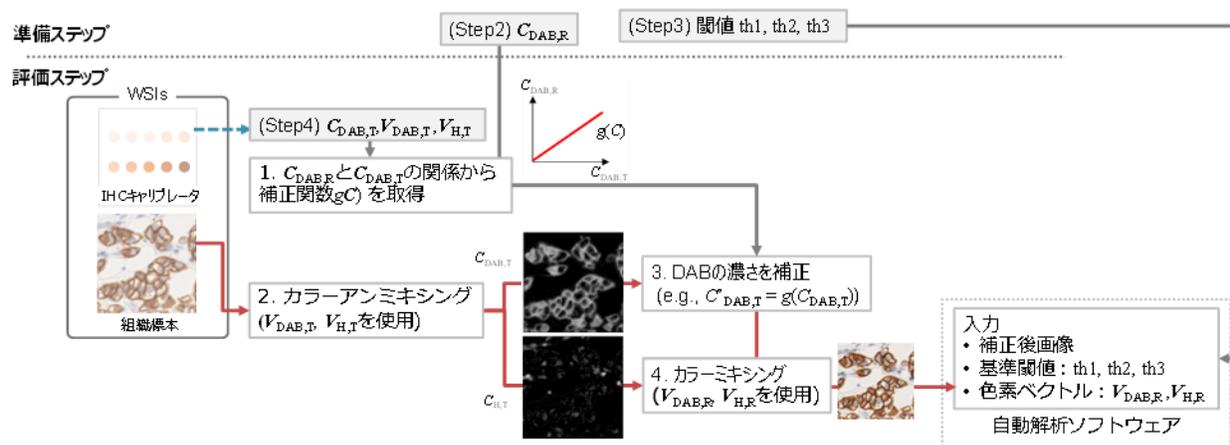


図 29 手法 2 の流れ

3.3 カラーチャートスライドを用いたスキャナの色補正

3.3.1 全体像

図 30 にカラーチャートスライドを用いたスキャナの色補正の流れを示す。本研究ではスキャナ間の色のばらつきをなくすために、カラーチャートスライドを利用してスキャナの色のカリブレーションを行う。カラーチャートスライドの基準となる色のデータと、スキャンした画像から取得した RGB 値を用いる。

画像の補正は、逆ガンマ補正と色補正の 2 つのステップで進める。2.2.5 節に記載したように、ディスプレイ上の画像の見え方と顕微鏡下で観察したスライドの見え方が近づくように、スキャナでガンマ補正が設定されることが想定される。そこで、最初にデバイス依存の RGB 値をリニア RGB 画像に変換するためのガンマ値を取得し、逆ガンマ補正を行う。続いて、リニア RGB に変換されたスキャナの値を用いて、カラーチャートスライドの基準値に合わせるための色補正関数を算出する。得られた 2 つのパラメータを組織画像とキャリブレーション画像に適用することで、スキャナの色のカリブレーションを行った画像が取得できる。以降の節でステップごとに分けて説明する。

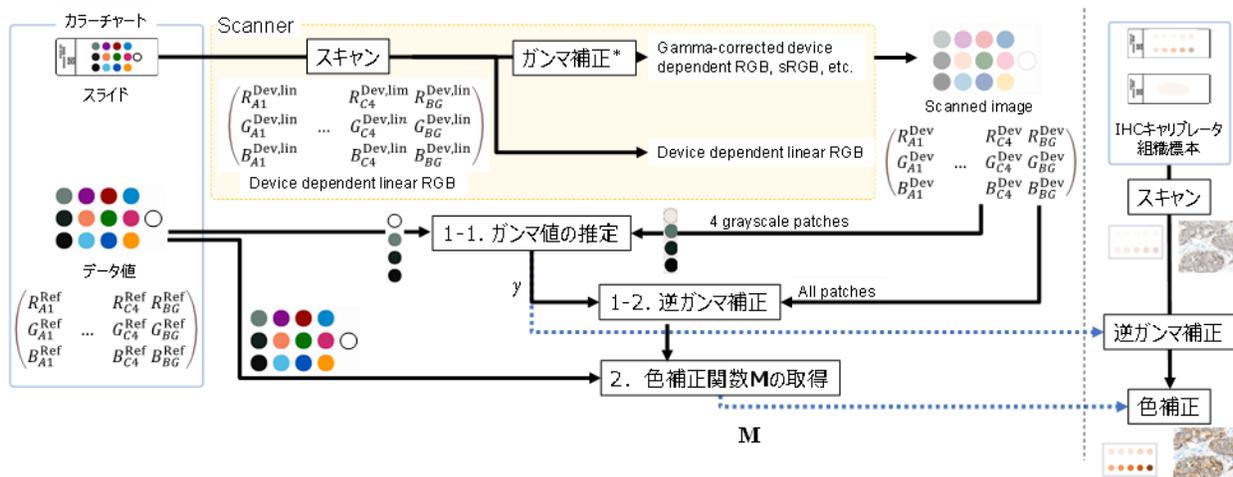


図 30 カラーチャートスライドを用いたスキャナの色補正の流れ

3.3.2 カラーチャートスライドの基準色の取得

始めに、カラーチャートスライドの基準となる色を取得する。背景を含む 13 色のパッチは NIST トレサブルな分光分布のデータシートがメーカーから提供されており、任意の表色系への変換が可能となっている。図 21(a)のスライド像がやや暗く見えるように、カラーチャートスライドの撮影部分には光量を減らすための ND フィルタが貼付されている。提供されている分光分布のデータシートも ND フィルタを介して取得されたスペクトルデータのため、このまま使用すると色補正後の画像も ND フィルタがかかった分、暗い画像として変換されてしまう。そこで、全てのカラーパッチのスペクトルデータを背景部分のスペクトルデータで除算し、背景部分の分光分布が 1 になるように調整を加えた。図 31 に調整後のパッチの分光分布データ $T(\lambda)$ の例を示す。

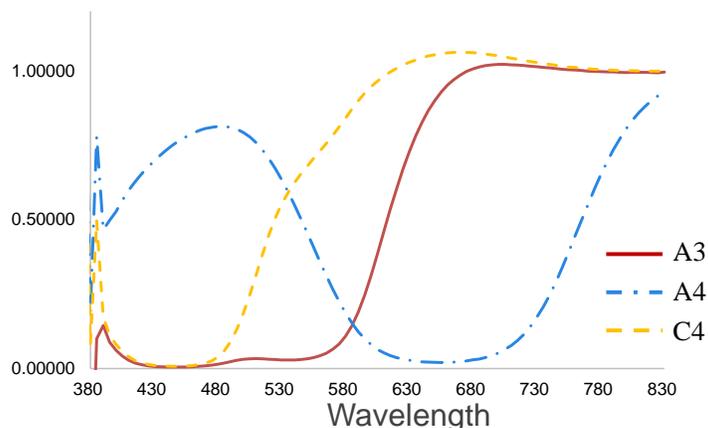


図 31 カラーパッチのスペクトルデータの例

次に、パッチの分光分布 $T(\lambda)$ 、光源の分光分布 $S(\lambda)$ 、XYZ 等色関数を用いて、式(15)に基づき各パッチの XYZ 値を取得した。光源には D65 光源を使用した。さらに、式(14)の逆変換によって RGB 値に変換し、12個のカラーパッチと背景部分の RGB 値を含む 3×13 の行列 \mathbf{G}_R を求めた。ここで、添え字 A1, ... C4 は図 21 に示すカラーパッチのインデックスを表し、BG は背景パッチを表す。また、上付き文字 Ref は、リファレンスの RGB 値を示す。図 21(b)は、算出された基準色を sRGB 表色系に変換して示している。

$$\mathbf{G}_R = \begin{pmatrix} R_{A1}^{\text{Ref}} & & R_{C4}^{\text{Ref}} & R_{BG}^{\text{Ref}} \\ G_{A1}^{\text{Ref}} & \dots & G_{C4}^{\text{Ref}} & G_{BG}^{\text{Ref}} \\ B_{A1}^{\text{Ref}} & & B_{C4}^{\text{Ref}} & B_{BG}^{\text{Ref}} \end{pmatrix}, \quad (35)$$

3.3.3 スキャナの色取得

ここでは、スキャナごとの色を取得する。スキャンした画像はリニアな RGB 値、または、式(19)によってガンマ補正されたデバイス依存の RGB 値を持つ。いずれにしても、光学濃度 $\mathbf{R}^{\text{Dev}} = (R^{\text{Dev}} \ G^{\text{Dev}} \ B^{\text{Dev}})^t$ を取得するため、画像の各画素の RGB 値を背景部分の RGB 値で除算する。上付き文字 Dev は、スキャン画像の RGB 値を示す。カラーチャートスライドのスキャン画像から各パッチの中心部 50×50 pixel の平均値を算出し、 3×13 の行列 \mathbf{G}_s を求めた。

$$\mathbf{G}_s = \begin{pmatrix} R_{A1}^{\text{Dev}} & & R_{C4}^{\text{Dev}} & R_{BG}^{\text{Dev}} \\ G_{A1}^{\text{Dev}} & \dots & G_{C4}^{\text{Dev}} & G_{BG}^{\text{Dev}} \\ B_{A1}^{\text{Dev}} & & B_{C4}^{\text{Dev}} & B_{BG}^{\text{Dev}} \end{pmatrix}. \quad (36)$$

3.3.4 逆ガンマ補正

逆ガンマ補正は、べき乗則で定式化された式(19)のガンマ補正モデルに基づく。リニア RGB 値 $D^{\text{Dev,Lin}} = R^{\text{Dev,Lin}}, G^{\text{Dev,Lin}}, B^{\text{Dev,Lin}}$ は、式(19)よりスキャナ内でガンマ補正が行われ、デバイスに依存する RGB 値 $D^{\text{Dev}} = R^{\text{Dev}}, G^{\text{Dev}}, B^{\text{Dev}}$ に変換される。逆ガンマ補正は、ガンマ補正值 γ の逆数 ($1/\gamma$) を用いて式(37)で表される。

$$D^{\text{Dev,Lin}} = (D^{\text{Dev}})^{\frac{1}{\gamma}} \quad (37)$$

画像からガンマ補正值 γ を推定するには、カラーチャートスライドのグレースケールのパッチ 4 色 (A1、B1、C1、BG) を使用し、輝度を表す XYZ 表色系の Y 成分を計算に用いる。基準の

グレースケールの Y 値 $\mathbf{Y}^{\text{Ref}} = (Y_{BG}^{\text{Ref}}, Y_{A1}^{\text{Ref}}, Y_{B1}^{\text{Ref}}, Y_{C1}^{\text{Ref}})$ 、スキャン画像のグレースケールの Y 値 $\mathbf{Y}^{\text{Dev}} = (Y_{BG}^{\text{Dev}}, Y_{A1}^{\text{Dev}}, Y_{B1}^{\text{Dev}}, Y_{C1}^{\text{Dev}})$ をそれぞれ取得する。基準はリニアな値を持つため、 $|\mathbf{Y}^{\text{Ref}} - (\mathbf{Y}^{\text{Dev}})^\gamma|^2$ を最小とする γ を求めることで、デバイス依存の RGB 値をリニア RGB に変換するためのガンマ値を取得する。スキャン画像のカラープロファイルが線形 ($\gamma=1$) であれば、スキャン値と基準値は等しくなり、直線の関係となる。そうでない場合は、スキャン画像の画素値を、得られたガンマ値で補正することで逆ガンマ補正を行う。図 32 の緑色のプロットはガンマ補正された画像と基準との関係を、青色のプロットは得られたガンマ値で逆ガンマ補正を行った画像と基準との関係を示す。

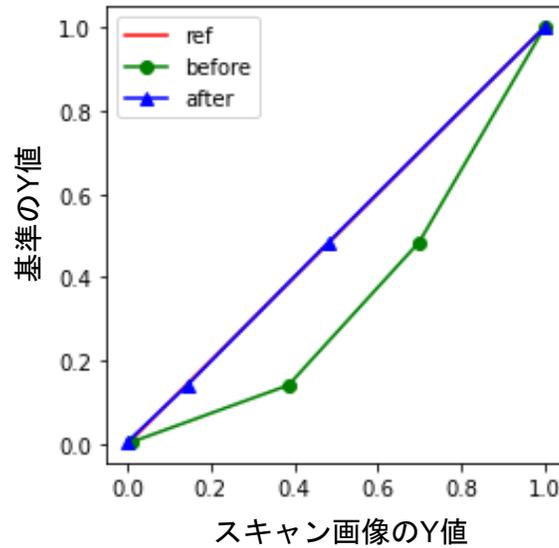


図 32 Y 値のプロットの例

3.3.5 色補正関数の取得

最後にスキャン画像の色を基準に合わせるための色補正行列 \mathbf{M} を算出する。式(21)の \mathbf{G}_R 、 \mathbf{G}_S に式(35)(36)で取得した行列をそれぞれ用いる。本研究では、多項式変換を利用した。次数 3-8 で変換した際の 2 つのスキャナ間の色差を確認し、IHC 画像において小さい色差が確認された次数 5 を選択した。次数が 5 の時、式(23)は式(38)で表される。

$$\mathbf{G}_R = \begin{bmatrix} \left(G_s(R_{A1}^{\text{Dev}}, G_{A1}^{\text{Dev}}, B_{A1}^{\text{Dev}}, R_{A1}^{\text{Dev}} G_{A1}^{\text{Dev}} B_{A1}^{\text{Dev}}, 1) \right)^t \\ \vdots \\ \left(G_s(R_{BG}^{\text{Dev}}, G_{BG}^{\text{Dev}}, B_{BG}^{\text{Dev}}, R_{BG}^{\text{Dev}} G_{BG}^{\text{Dev}} B_{BG}^{\text{Dev}}, 1) \right)^t \end{bmatrix} \mathbf{M} \quad (38)$$

また、色補正関数 \mathbf{M} の算出には 2.2.6 節に記載した WPPLS を利用した。疑似逆行列を計算して通常通りに解いた場合、色補正後の画像の背景に赤みや青みが生じることがあった。そこで、カラーパッチの背景色を白色とみなし、色補正後も背景パッチの白色性が維持されるように制約を設けたうえで色補正関数 \mathbf{M} を算出した。多項式変換の次数の選択、通常の変換と WPPLS の変換の比較については、5.2.2 節および 5.3.1 節に記載する。

ここまでの手順は事前に行い、スキャナごと、または、同一のスキャナでも撮影条件ごとにガンマ値 γ 、色補正関数 \mathbf{M} を取得する。得られたガンマ値 γ 、色補正関数 \mathbf{M} をキャリブレーション画像、組織標本画像に適用することでスキャナの色キャリブレーションが行われ、スキャナ間の色の違いを補正した画像が取得される。

3.4 まとめ

本章では、WSI の色や明るさのばらつきに影響する要因を染色、スキャナに分けたうえで、それぞれを補正するための手法を提案した。染色のばらつきの補正には、染色時の反応の程度を把握するための指標として、抗原でコーティングされたマイクロビーズからなる IHC キャリブレーションスライドを導入した。スキャナの補正には、スキャナの特徴を把握するための指標としてカラーチャートスライドを導入した。臨床利用が認可されている解析ソフトウェアは染色装置から試薬、スキャナまでトータルのシステムで導入する必要があるのに対し、本手法は染色プロセスとスキャンプロセスにそれぞれスライドを追加することで対応できることから、臨床のワークフローの軽微な変更で導入することができる。

IHC キャリブレーションスライドの解析では、無染色のビーズがグレーがかった色となっていたことから、ビーズの散乱成分と DAB の染色成分の 2 つからなるモデルを想定し、DAB 染色の色味と濃さを取得した。自動解析結果の標準化に向けては、2 つの補正手法を提案した。診断の自動化の観点では画像の色補正を行う手法 2 が有効だが、病理医が画像観察を行う観点では手法 1 が有効と考える。画像補正への抵抗感がある場合や、画像処理を加えていないオリジナルの画像や顕微鏡の見え方に近い画像を見ながら評価したい場合、オリジナルの画像と補正画像を比較したい場合などには手法 1 が選択できる。

カラーチャートスライドによる補正は、逆ガンマ補正と色補正の 2 ステップから成る手法を提案した。また、通常の方法で色補正行列を求めた際に、背景の RGB 値が不自然な色味となっていたことから、他分野で利用されていた背景の白色性を維持する手法を WSI の色補正に初めて取り入れ、背景の色補正後差を少なくするよう工夫を加えた。

第4章 染色によるばらつきの補正効果の検証

4.1 概要

ここでは、染色によるばらつきの補正効果に限定し、IHC キャリブレーションスライドを用いた補正手法の有効性検証を行った内容について記載する。

4.2 実験内容

4.2.1 実験データ

実験には匿名化された患者の組織標本ブロックを使用した。病理医が目視診断にて HER2 スコアを決定した乳がんの組織標本ブロック 4 種類 (スコア 0、1+、2+、3+) を準備し、それぞれのブロックから 4 μm の厚みで薄切した連続切片を用いて組織標本スライドを作成した。各スコアのスライド 1 枚ずつと IHC キャリブレーションを追加した 5 スライドを 1 つのデータセットとみなした (図 33)。日々の染色のばらつきを反映したデータを取得するため、各データセットを異なる日に同一のラボで染色した。すべての IHC 法の染色には同一の抗体 (ベンタナ PATHWAY anti-HER2/neu (ロシュ・ダイアグノスティクス株式会社)) が用いられた。実験 4.2.3 では、染色済みの 16 データセットを解析に利用した。

4.2.4 節の実験には、浸潤性乳がんから成る TMA スライドを 1 枚準備し、IHC キャリブレーションスライドとともに染色を行った。TMA スライドは病理医 1 名によってコアごとに HER2 スコアが判定された。判定対象となる乳がんの浸潤領域を含まないコアは除外し、17 個のコアを解析対象とした。

スキャンには WSI スキャナ S60 (浜松ホトニクス) を使用し、分解能 0.23 $\mu\text{m}/\text{pixel}$ で撮影を行った。解析には、WSI から bmp 形式でエクスポートした画像を使用した。

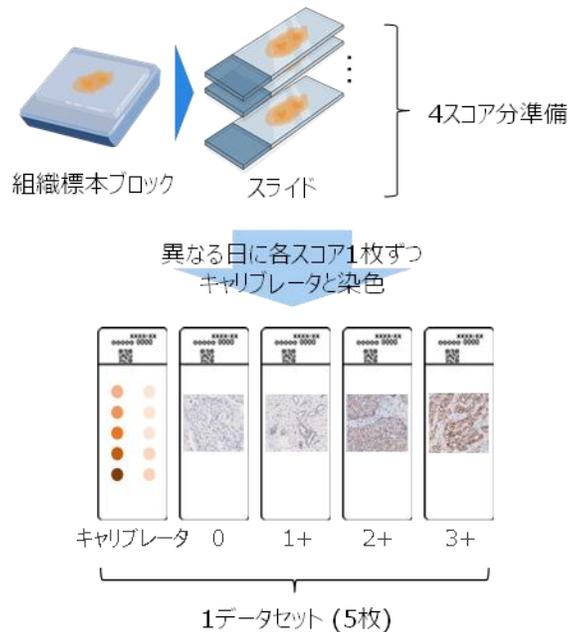


図33 実験で用いたデータセットのイメージ

4.2.2 IHC キャリブレーションモデルの散乱項の影響の検証

本研究ではキャリブレーションから DAB の濃さを推定する際に、式(33)に示す散乱項を考慮したモデルを使用している。そこで、散乱項を考慮した場合と考慮しなかった場合に推定された濃さのプロットを比較し、散乱項の効果を検証した。

4.2.3 HER2 の自動解析における提案手法の効果の検証

ここでは、リーブワウンアウト形式で交差検証を行った。16 データセットのうちの 1 データセットを基準データ、15 データセットを対象データとみなし、提案手法による染色のばらつきの補正効果、および、HER2 の自動解析結果への影響を検証した。4つの異なる基準データと対象データの組み合わせについて、テストセット TS1-TS4 とする。基準データセットは、TS1 では病理医が染色状態を見て決定し、TS2-TS4 はランダムに選択した。

HER2 の自動解析には、オープンソースソフトウェアの QuPath を用いた。QuPath では図 12 に示した解析手順とほぼ同様の手順で HER2 のスコア判定が行われる。また、カラーアンミキシングのための色素成分ベクトルの設定や、濃さの閾値の設定が可能となっている。実験では表 4 に示す 3つの組み合わせについて自動解析を行い、結果を比較した。

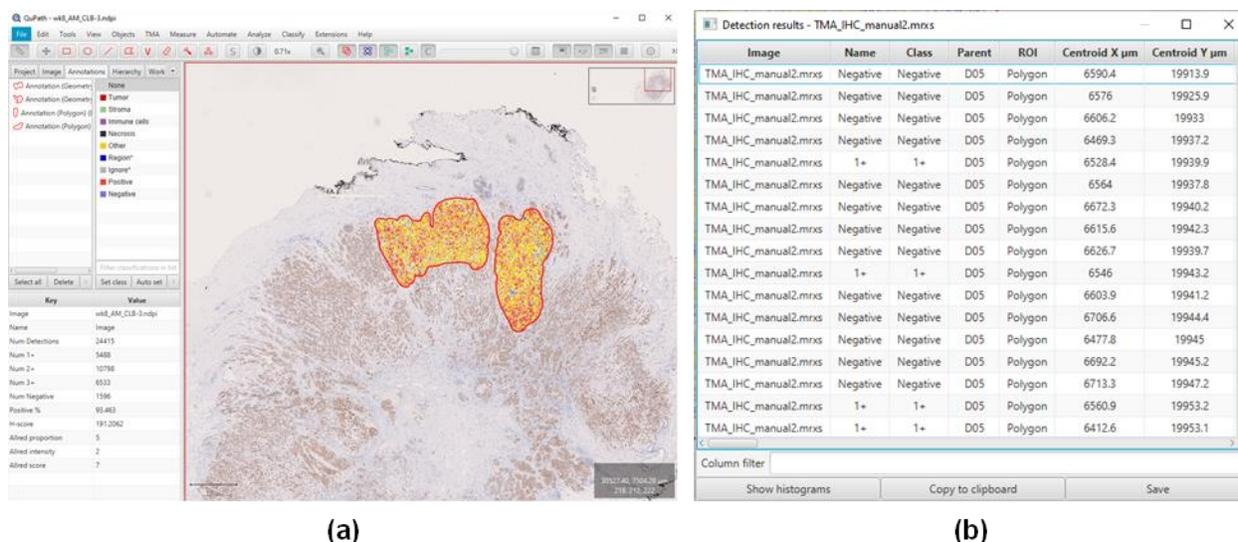


図 34 QuPath での HER2 評価の例；(a)解析画面；(b)検出された細胞ごとの DAB の濃さの出力情報

表 4 自動評価で比較する手法

	補正なし	提案手法 1	提案手法 2
画像	元画像	元画像	補正画像
閾値	固定	補正	固定

評価指標は、HER2 スコアに関して、病理医の診断結果と自動解析で得られた結果との一致率を比較した。また、式(2)で計算される H-score の標準偏差を染色のばらつきの指標として比較した。実験に用いたデータは同一の組織標本から薄切したスライドを用いており、染色のばらつきがない場合には H-score は組織標本のケースごとに近い値が得られると想定されことから、ばらつきの指標として選択した。さらに、H-score の標準偏差に対して一要因の分散分析 (ANOVA)[55]を使い、補正を行わなかった場合と提案手法を適用した場合で有意差が見られるか検定を行った。ANOVA で有意差が見られた際の事後検定法には Tukey 法を用いた[56]。それぞれの検定において、 p 値が 0.05 以下の際に統計的な有意差があるとした。

また、手法 2 の画像の補正効果の検証として、DAB の濃さのヒストグラムの比較を行った。QuPath では検出された細胞それぞれについて、細胞部位ごとの DAB の濃さを出力することができる (図 34(b))。出力された細胞膜の DAB の濃さを利用してヒストグラムを作成し、基準データと補正前後の対象データの分布を比較した。また、類似度の評価として、基準データと補

正前の対象データ、基準データと補正後の対象データの組み合わせについて相関係数とカイニ乗値[57]を計算した。

4.2.4 基準データと異なる組織標本の補正効果の検証

実験 4.2.3 節では同一の組織標本を基準データと対象データに分けて実験を行った。本節では、基準データは 4.2.3 節と同様の組織標本を利用し、対象データとして異なる組織標本である TMA スライドを利用した。表 4 に示す 3 つの組み合わせについて QuPath で自動解析を行い、病理医の診断結果と自動解析で得られた HER2 スコアとの一致率を比較した。

4.3 実験結果

4.3.1 IHC キャリブレーションモデルの散乱項の影響の検証結果

式(33)のキャリブレーションモデルにおいて、ビーズの散乱項の影響を図 35 に示す。レベルごとに推定された濃さのプロットと、各レベルの平均の濃さから取得したカラーバーを示している。散乱を考慮せず、ビーズ部分の光学濃度から算出した単位ベクトルを V_{DAB} として濃さを推定した場合(図 35(a))、レベル 4 と 5 において濃さの反転がみられることがあった。散乱項を考慮したモデルとすることで(図 35(b))、設計通りに段階的な濃さが取得できていることが確認できた。また、レベル内でビーズの濃さのばらつきはあるが、四分位の範囲で見るとレベル間の重複なく濃さの推定が行われていることがわかった。

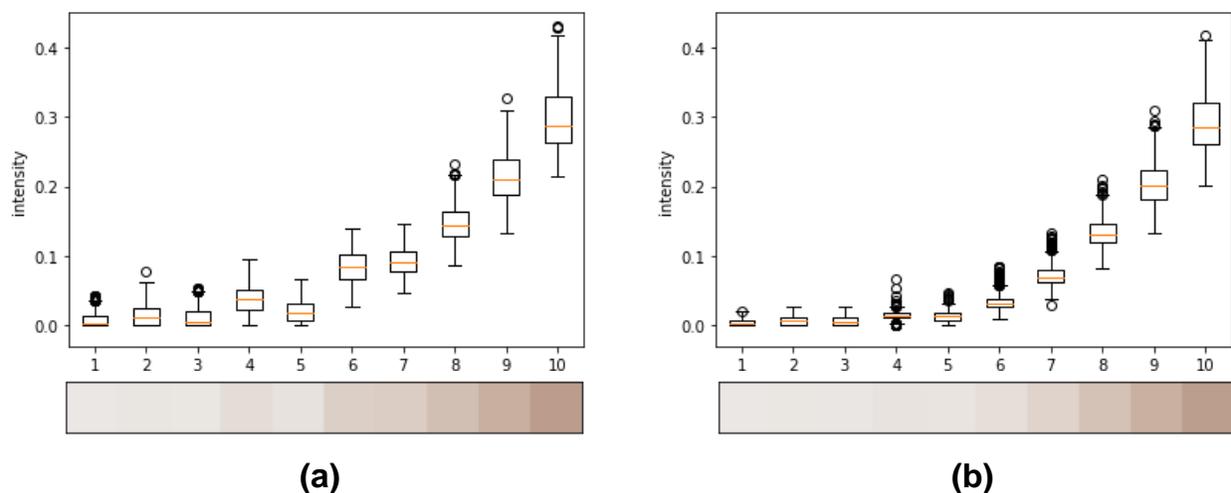


図 35 キャリブレーションから推定された DAB の濃さのプロットおよび濃さのカラーバーの比較；(a) 散乱を考慮しないモデル；(b) 散乱を考慮したモデル

4.3.2 HER2の自動解析における提案手法の効果の検証結果

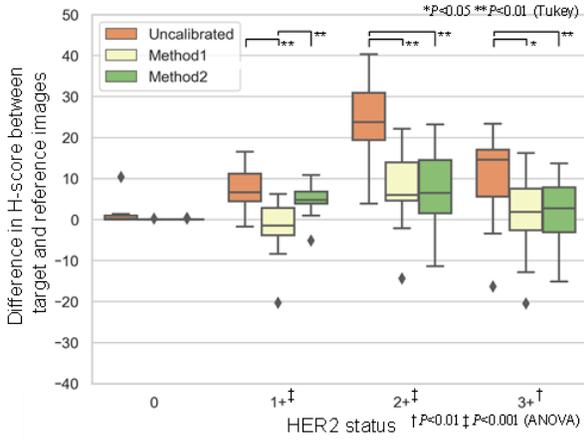
表 5 に HER2 評価結果を示す。補正を行わなかった場合、スコア 1+、2+の組織で病理医の診断と自動解析の結果の不一致がみられた。TS1 では 60 スライドのうち 7 スライド、TS2 と TS3 では 1 スライド、TS4 では 4 スライドであった。提案手法を用いることで、不一致となったスライドが 1 枚にまで減少し、一致率の改善が確認された。提案手法 1 では 100%の一致が得られ、提案手法 2 では TS1-TS4 で各 1 スライドの不一致が認められた。TS1 の 1+に提案手法 1 を適用した場合を除き、提案手法のいずれかを適用することで H-score の標準偏差が補正なしの場合に比べて小さくなっていた。

図 36 は基準データと対象データの H-score の差分をボックスプロットとして作成したものである。TS1 について、補正なし、手法 1、手法 2 のそれぞれの平均値を比較すると、0 のケースでは 1.07、0.04、0.04、1+のケースでは 7.25、-1.66、4.91、2+のケースでは 23.32、7.98、7.48、3+のケースでは 10.90、0.86、1.96 となっていた。手法を適用することによって H-score の差分が 0 に近づいていることが確認された。また、H-score の標準偏差に対して ANOVA を行った結果、1+のケースで $F(2, 42) = 10.75, P = 0.000$ 、2+のケースで $F(2, 42) = 11.79, P = 0.000$ 、3+のケースで $F(2, 42) = 5.14, P = 0.010$ と手法間で有意な差が得られた。

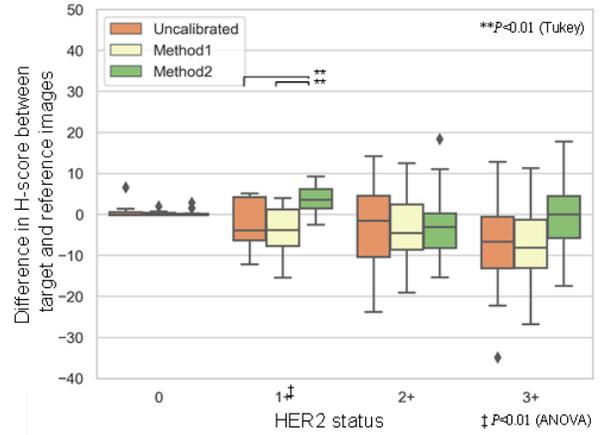
表 6 は有意差が確認されたケースについて、Tukey 法で事後検定を行った結果を示す。TS1 について、1+のケースでは、手法 1 と手法 2 ($P = 0.006, 95\% \text{ C.I.} = [1.73, 11.41]$)、手法 1 と補正なし ($P = 0.001, 95\% \text{ C.I.} = [4.07, 13.75]$)、2+のケースでは、手法 1 と補正なし ($P = 0.001, 95\% \text{ C.I.} = [6.34, 24.36]$)、手法 2 と補正なし ($P = 0.001, 95\% \text{ C.I.} = [6.83, 24.85]$)、3+のケースでは、手法 1 と補正なし ($P = 0.015, 95\% \text{ C.I.} = [1.70, 18.39]$)、手法 2 と補正なし ($P = 0.033, 95\% \text{ C.I.} = [0.60, 17.28]$)の間で有意な差が確認された。TS2 では、1+のケースにおいて補正を行わなかった場合と提案手法の間で有意差が認められた。TS3 では、すべてのスコアの標準偏差が提案手法により改善された。TS4 では、提案手法 2 により、2+と 3+の H-score の差が 0 に近づいた。

表5 自動解析ソフトウェアによる解析結果と病理医の評価結果の比較

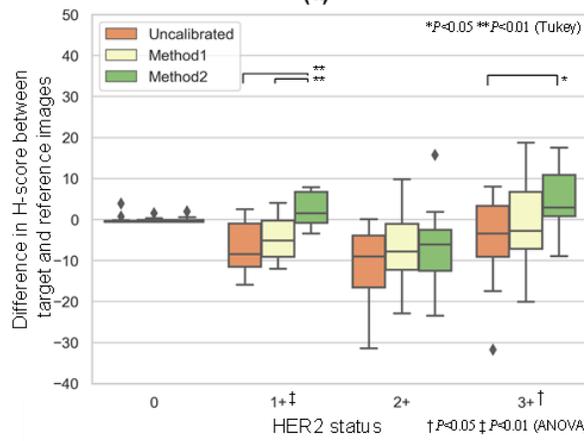
Test set ID	HER2 Status	Uncalibrated		Proposed method1		Proposed method2	
		Concordance (Rate (%))	<i>SD</i> of H-score	Concordance (Rate (%))	<i>SD</i> of H-score	Concordance (Rate (%))	<i>SD</i> of H-score
TS1	0	15 (100.0)	2.5	15 (100.0)	0.1	15 (100.0)	0.1
	1+	12 (80.0)	5.3	15 (100.0)	6.4	15 (100.0)	3.7
	2+	11 (73.3)	10.8	15 (100.0)	9.3	14 (93.3)	9.2
	3+	15 (100.0)	10.4	15 (100.0)	9.1	15 (100.0)	7.5
TS2	0	15 (100.0)	1.7	15 (100.0)	0.5	15 (100.0)	0.9
	1+	15 (100.0)	5.4	15 (100.0)	5.4	15 (100.0)	3.6
	2+	14 (93.3)	10.5	15 (100.0)	9.3	14 (93.3)	8.6
	3+	15 (100.0)	11.6	15 (100.0)	9.2	15 (100.0)	8.7
TS3	0	15 (100.0)	1.1	15 (100.0)	0.6	15 (100.0)	0.9
	1+	15 (100.0)	6.1	15 (100.0)	4.9	15 (100.0)	4.8
	2+	14 (93.3)	9.5	15 (100.0)	8.7	14 (93.3)	9.3
	3+	15 (100.0)	10.4	15 (100.0)	9.7	15 (100.0)	7.3
TS4	0	15 (100.0)	1.7	15 (100.0)	0.5	15 (100.0)	0.1
	1+	13 (86.7)	7.3	15 (100.0)	6.2	15 (100.0)	4.3
	2+	13 (86.7)	10.4	15 (100.0)	8.7	14 (93.3)	9.2
	3+	15 (100.0)	11.1	15 (100.0)	8.7	15 (100.0)	7.8



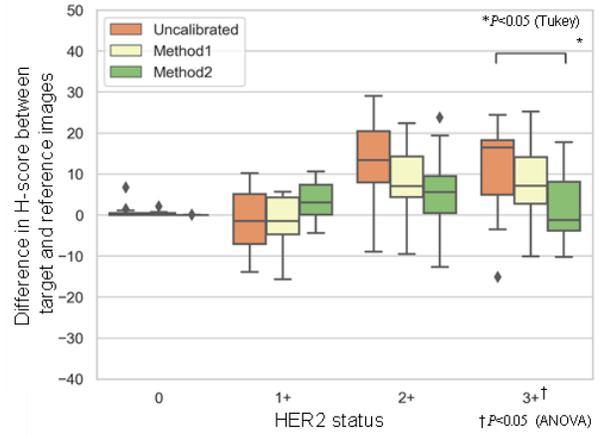
(a)



(b)



(c)



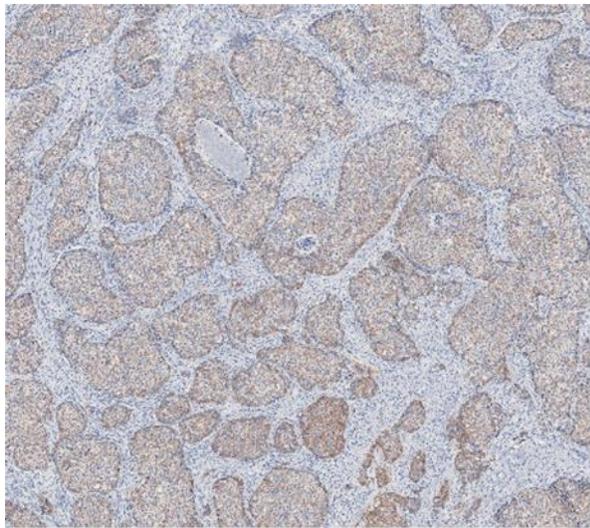
(d)

図 36 ケースごとの基準画像と評価対象画像の H-score の標準偏差と有意差の比較；(a)–(d) テストセット TS1–TS4

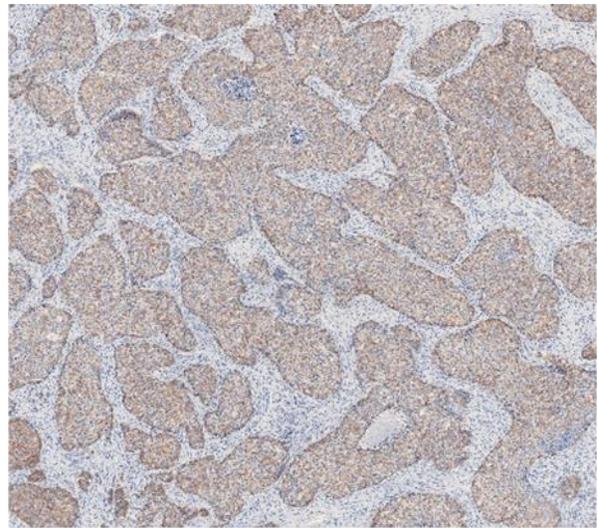
表 6 Tukey 法による有意差の検定

Test set ID	HER2 status	group1	group2	Mean diff.	P	95% confidence interval	
						Lower	Upper
TS1	1+	Method1	Method2	6.57 **	.006	1.73	11.41
		Method1	Uncalibrated	8.91 **	.001	4.07	13.75
		Method2	Uncalibrated	2.34	.476	-2.50	7.18
	2+	Method1	Method2	-0.49	.900	-9.50	8.52
		Method1	Uncalibrated	15.35 **	.001	6.34	24.36
		Method2	Uncalibrated	15.84 **	.001	6.83	24.85
	3+	Method1	Method2	1.10	.900	-7.24	9.45
		Method1	Uncalibrated	10.04 *	.015	1.70	18.39
		Method2	Uncalibrated	8.94 *	.033	0.60	17.28
TS2	1+	Method1	Method2	7.14 **	.001	1.47	12.81
		Method2	Uncalibrated	-5.87 **	.007	-11.54	-0.21
TS3	1+	Method1	Method2	6.63 **	.003	0.75	12.51
		Method2	Uncalibrated	-8.57 **	.001	-14.45	-2.69
	3+	Method2	Uncalibrated	-9.95 *	.478	-12.47	4.35
TS4	3+	Method2	Uncalibrated	10.08 *	.017	1.54	18.62

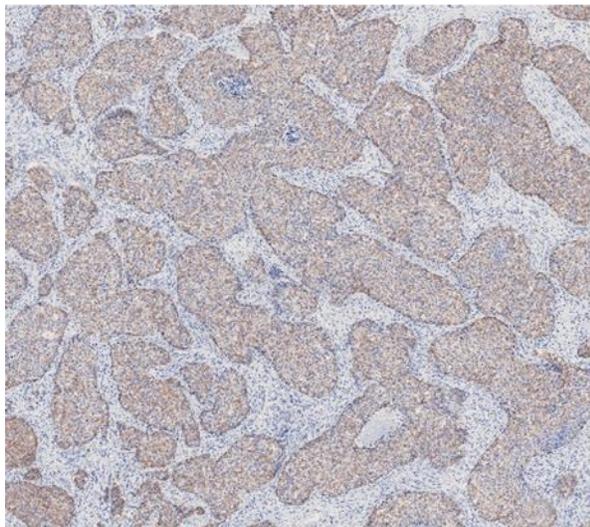
TS1 について、手法 2 で補正した画像の例を図 37、図 38 に示す。対象データの DAB とヘマトキシリンの色味は基準データに近づくように補正された。また、各図(d)に示す細胞膜の DAB の濃さのヒストグラムより、濃さも基準に近づくように補正されたことを確認した。表 7 はヒストグラムの類似度を評価した結果を示す。相関係数は 1 に近いほど類似することを意味し、カイ二乗値は値が小さいほど類似することを意味する。相関係数を確認すると、補正前でも高い類似度があることがわかる。カイ二乗値を確認すると、補正によって値が小さくなったことが確認できた。得られた相関係数、カイ二乗値に T 検定を行った結果、2+ のケースで相関係数 ($t = -4.90$, $P = 0.000$)、カイ二乗値 ($t = 3.58$, $P = 0.002$)それぞれで有意な差が確認された。3+ のケースでは有意差は確認されなかった。



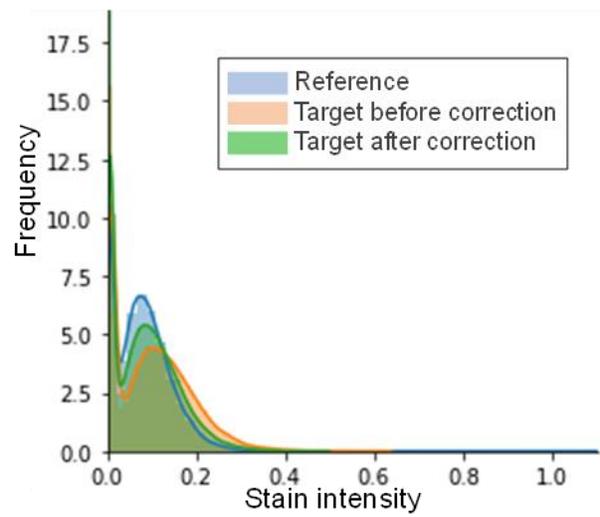
(a)



(b)

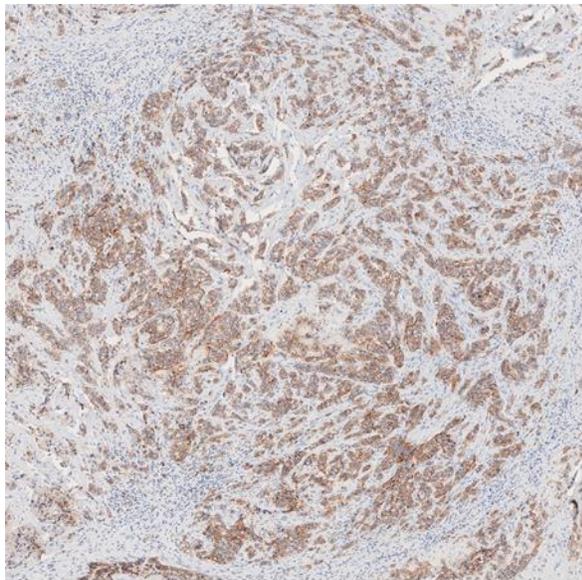


(c)

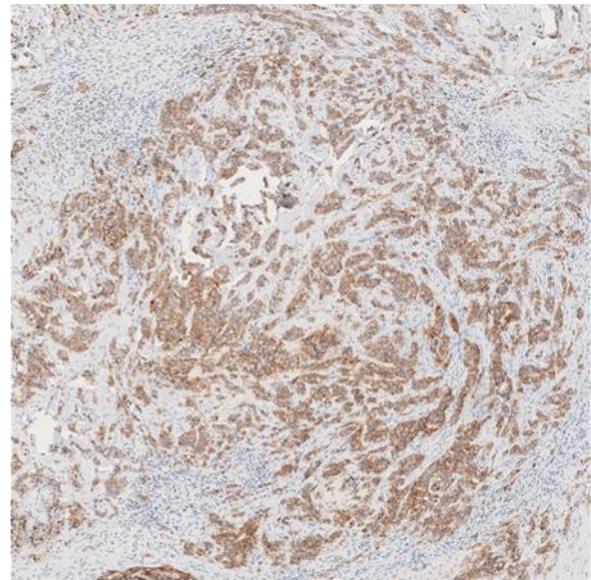


(d)

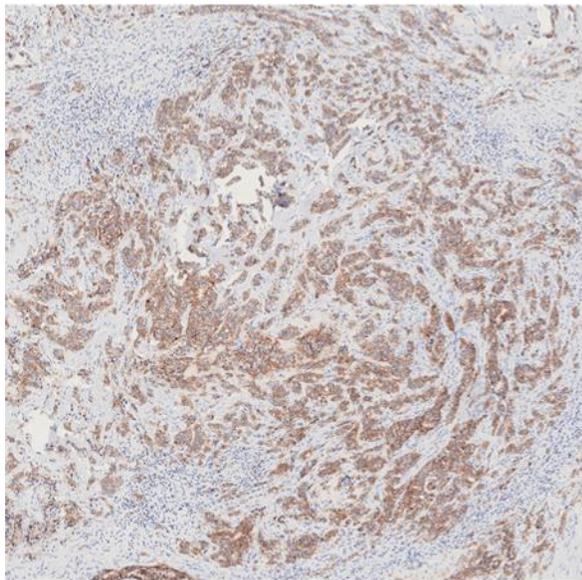
図 37 2+のケースにおける画像の色補正効果の比較；(a) 基準画像；(b) 評価対象画像(補正前)；(c) 評価対象画像(補正後)；(d) 細胞膜における DAB の濃さの分布の比較(縦軸：細胞の頻度、横軸：細胞膜の DAB の濃さ)



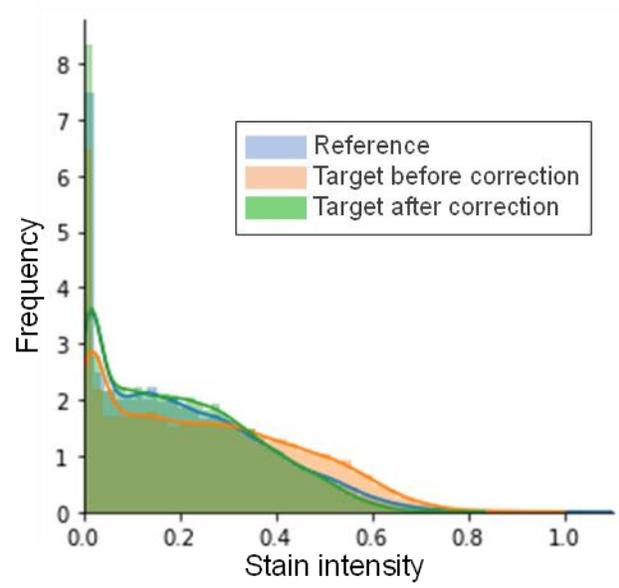
(a)



(b)



(c)



(d)

図 38 3+ のケースにおける画像の色補正効果の比較；(a) 基準画像；(b) 評価対象画像(補正前)；(c) 評価対象画像(補正後)；(d) 細胞膜における DAB の濃さの分布の比較(縦軸：細胞の頻度、横軸：細胞膜の DAB の濃さ)

表7 DABの濃さのヒストグラムの類似度の比較

HER2 Score	Similarity (DAB)			t-test	
	Method	Reference vs Before	Reference vs After	<i>t</i>	<i>p</i>
1+	Correlation	0.98±0.02	0.99±0.02	-0.30	.766
2+		0.95±0.03	0.98±0.02	-4.90 ***	.000
3+		0.98±0.01	0.98±0.02	1.16	.265
1+	Chi-square	0.14±0.19	0.05±0.06	1.52	.151
2+		0.43±0.36	0.08±0.09	3.85 **	.002
3+		0.04±0.04	0.05±0.04	-0.63	.542

4.3.3 基準データと異なる組織標本の補正効果の検証結果

表8は基準データと異なる組織標本を評価データとした場合のHER2の評価結果を示す。補正を行わなかった場合でも、ほとんどのケースで病理医の評価と一致するスコアが得られているが、2スライドで結果の不一致がみられた。提案手法を用いることで、100%の一致率が確認された。図39に手法2の画像補正の一例を示す。補正により、対象画像のDABの色味と濃さが基準画像に近づいたことを確認した。

表8 HER2の自動解析結果と目視評価との一致率の比較

	Uncalibrated					Method1					Method2				
	Auto					Auto					Auto				
	0	1+	2+	3+	total	0	1+	2+	3+	total	0	1+	2+	3+	total
manual	0	3	2		5	0	5			5	0	5			5
	1+		2		2	1+		2		2	1+		2		2
	2+			5	5	2+			5	5	2+			5	5
	3+				5	3+				5	3+				5
total		3	4	5	17		5	2	5	17		5	2	5	17
	Concordance				88.2%	Concordance				100.0%	Concordance				100.0%

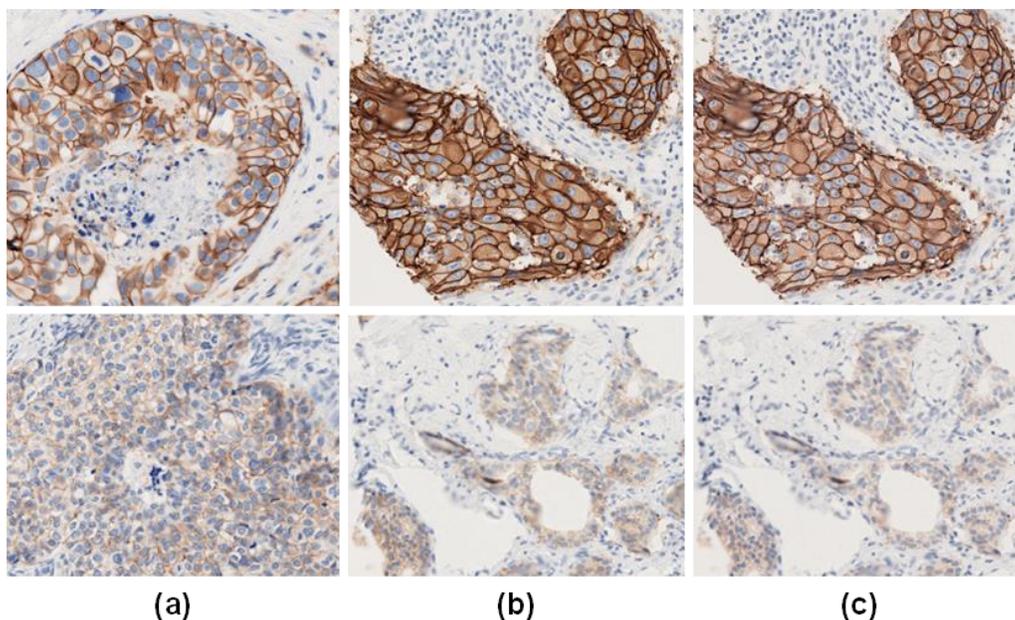


図 39 画像の比較（上段：3+、下段：2+）(a) 基準画像；(b) 評価対象画像(補正前)；(c) 評価対象画像(補正後)

4.4 考察

本章では、IHC キャリブレーションスライドを導入した提案手法による、日々の染色のばらつきの補正効果を評価した。アメリカでは臨床診断のための IHC 法については、染色の安定性を保証するために FDA によって承認された染色抗体と自動染色装置を用いることが決められている。本研究のデータもすべて FDA の承認を得た環境下で染色されたスライドを利用しているため、染色のばらつきとして評価した H-score の標準偏差は、補正を行わなかった場合であっても小さい値が得られていた(表5)。それにもかかわらず、染色の補正を適用せずに自動解析ソフトウェアで画像解析を行った際には病理医の判定と一致しない結果が 60 スライド中 7 スライド確認された。この結果より、日々のわずかな染色の変動が自動解析結果に影響してしまうことがわかった。提案手法を適用することで、表 5 に示すように判定の不一致は 1 スライドにまで減少した。

H-score の標準偏差も 1+ のケースに提案手法 1 を適用した場合を除き、小さくなることを確認した。今回実験に利用した 1+ のケースの形態上、HER2 の評価対象となる領域が 2+、3+ のケースと小さいため、Immunoscore の分類が変わると H-score が変動しやすいためと考える。H-score の標準偏差は大きくなったものの、一致率は両提案手法を適用することで 100% に改善された。2+ と 3+ のケースでは、補正を行わなかった場合と提案手法との間で統計的に有意な差を確認することができた。提案した 2 つの方法を比較すると、一致率は提案手法 1 が良く、H-

score の標準偏差については提案手法 2 で小さな値が確認された。交差検証ではすべての組み合わせをテストしたわけではないが、TS1-TS4 のいずれにおいても提案手法によって一致率と H-score の標準偏差の改善が見られた。TS1-TS4 の間に有意差が見られるかどうか H-score の標準偏差を ANOVA で検定したところ、いずれの手法においても有意差は見られなかった。そのため、後述する 1 つのデータセットを基準とした場合を除き、他の組み合わせでも同様の結果が得られると想定される。

手法2を適用した際に、病理医の判定と不一致となったスライドが1つ確認された。これは、IHC キャリブレーションの染色性が原因と考える。IHC キャリブレーションはレベルが上がるごとに段階的に濃く染色されるように設計されている。実験に使用した 16 枚の IHC キャリブレーションスライドのうち、15 枚は設計通りに染色されていたが (図 40(a)-(b))、判定が不一致となったスライドではレベル 9 と 10 の染色強度の差がわずかであった (図 40(c))。レベル 9 と 10 で推定された DAB の濃さの差を数値で見ると、15 データでは平均で約 0.07 の差があったが、問題のスライドでは約 0.01 であった。補正関数の係数を比較すると、レベル 10 を計算に含めた場合と、含めずにレベル 9 までを使って計算した場合とで約 1.2 倍の差があった (図 40(d))。そのため、通常よりも DAB の濃さが薄いレベル 10 を含めて計算した補正関数を使用した場合、通常の補正よりも補正結果が暗くなってしまったと考える。提案手法 1 では正しい HER2 ステータスが得られたものの、提案手法 2 ではごく僅かな差で不一致となっていた。比較として、レベル 9 までを用いて補正関数を求め、補正をかけた結果を図 41 に示す。色補正後の画像の色強度は参照画像に近くなり、自動解析の結果も病理医と一致するスコアが得られた。この結果は、製造上のばらつきが色強度補正に影響する可能性を示している。より安定した色補正のためには、キャリブレーションの製造上のばらつきの改善が望まれる。

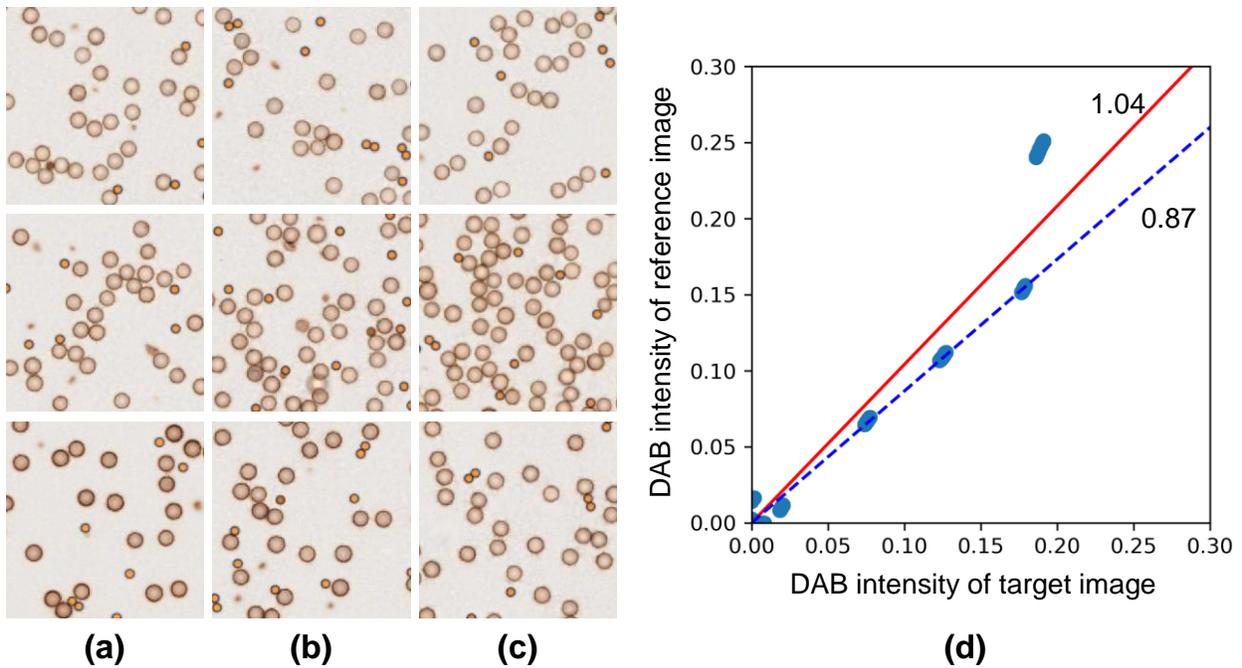


図 40 レベル 8-10 のキャリブレーション画像と DAB の濃さの関係性の比較；(a) 基準データセット；(b) 判定が一致した評価データセットの例；(c) 判定を誤った評価データセット；(d) 基準キャリブレーションと判定を誤った評価キャリブレーションの DAB の濃さの関係性。赤線はレベル 10 まで含めて得られた補正関数、青線はレベル 9 までを使って得られた補正関数を示す

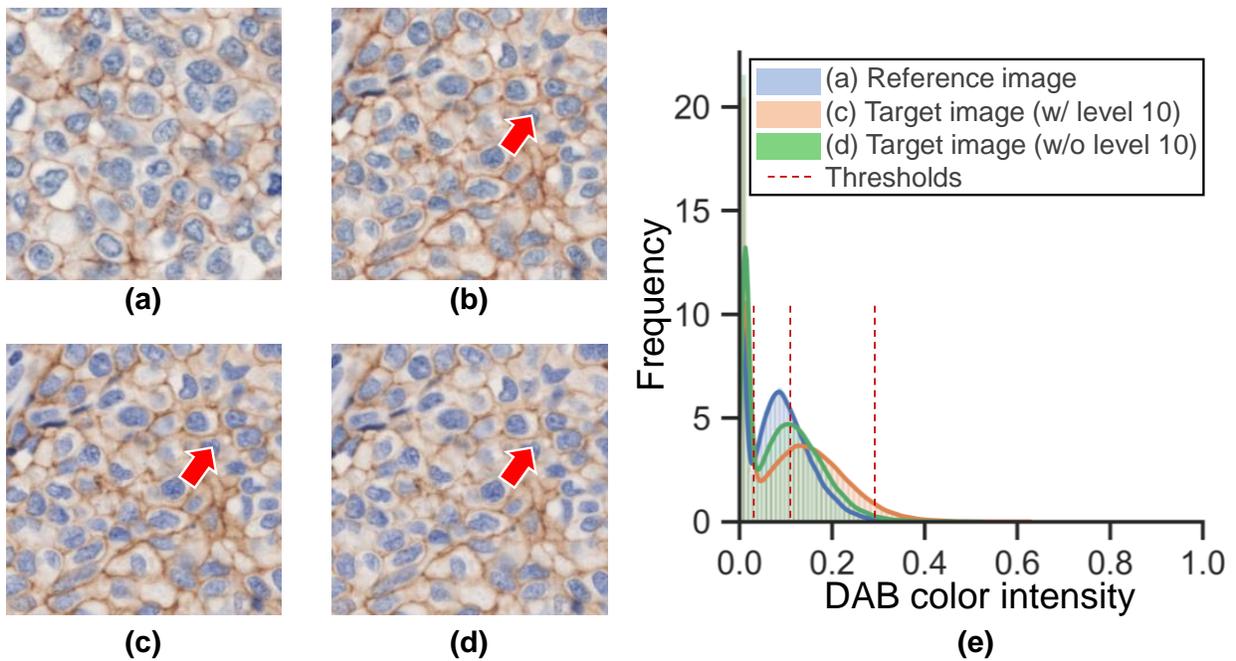


図 41 (a) 基準組織画像；(b) 対象組織画像(補正前)；(c) レベル 10 まで含めて得られた補正関数で色補正した対象組織画像；(d) レベル 9 までで得られた補正関数で色補正した対象組織画像。

赤矢印は濃さの違いがわかりやすい箇所を指す；(e) 細胞膜の DAB の濃さのヒストグラムの比較 (画像(a),(c),(d))

提案手法では IHC キャリブレーションのタンパク量そのものは解析には利用せず、ランバート・ベール則に基づいて画像から求めた DAB の濃さを基準としている。病理医がもともと濃さを基準とした評価に慣れていることも一つではあるが、タンパク量と DAB の濃さの関係性が単調増加にあることしかわかっていないことが理由にある。タンパク量と画像から求めた DAB の濃さを比較したとき、レベル 10 に当たるタンパク量で濃さの傾きが変わっていることがわかった (図 42(a))。IHC キャリブレーションのタンパク量の範囲でもタンパク量と DAB の濃さの関係が線形の関係になく、どのような関係式での表現が妥当か判断が難しい。また、さらにタンパク量が増えた際に濃さが増加傾向となるのか、頭打ちとなるのかが明らかでなく、よりタンパク量の多い組織標本に対しても IHC キャリブレーションから求めた関係式が適用できるかわからない。別日に染色した 2 つの IHC キャリブレーションから求めた DAB の濃さの関係は図 42(b) のようにほぼ線形の関係であることから、DAB の濃さを基準とすることでタンパク量によらず一つの係数で染色の増減を補正することができる。

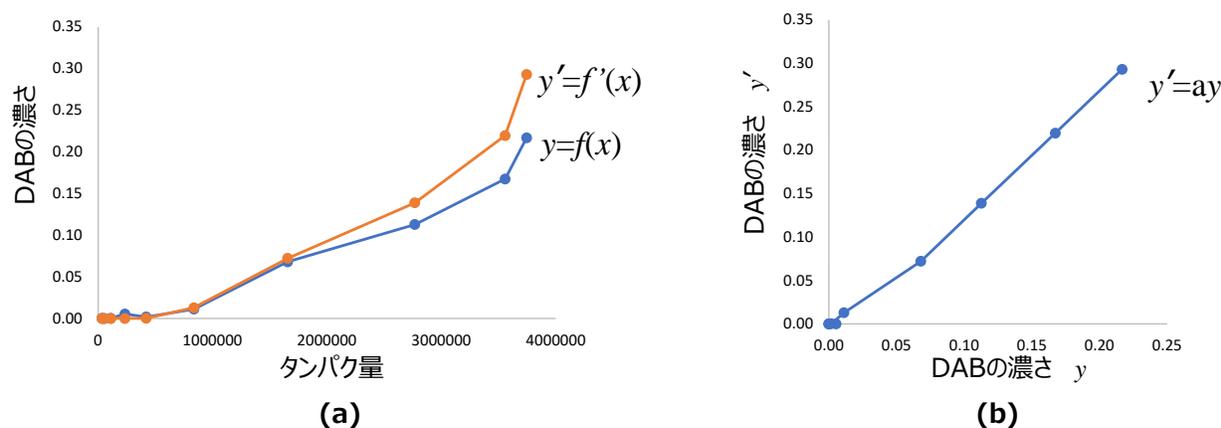


図 42 別日に染色した 2 つの IHC キャリブレーションの関係性の比較；(a) タンパク量と DAB の濃さの関係；(b) DAB の濃さの関係

本研究では、色補正手法と併せて HER2 スコアが既知の組織画像から自動解析ソフトウェアで用いる閾値を計算する方法を提案した。HER2 の自動解析を行っている研究のひとつ[5]では、病理医に解析結果を確認してもらいながら閾値を調整したとの記載があったが、多くの研究では閾値の設定方法は言及されていなかった。IHC キャリブレーション単体では HER2 スコアと DAB の濃さとの関係性を知ることはできないため、本研究では経験を積んだ病理医が選択した組織標本を閾値決定に利用している。これにより、病理医の診断を反映しつつ、自動での閾値決定

を実現した。得られた閾値と補正手法を併用することで、手法 2 を適用した際に病理医の判定と不一致となった 1 スライドを除き、正しい HER2 スコアを得ることができた。さらに、基準データセットとは異なる組織標本を対象とした 4.3.3 節の実験では、評価対象となるすべての組織標本が基準データセットとは異なる組織検体から作成されたにもかかわらず、88.2%の症例が補正なしでも正常に評価された。ただし、現在は一つのデータセットのみを使用しているため、後述する 6.3.1 節の実験では例外も見られた。これらの結果は、今後より多くの組織標本を用いて閾値決定方法の頑健性を評価することで、完全自動化された閾値決定が達成されることを示唆している。また、確率密度分布の差分を取ることは数学的な観点では正しいと言えないかもしれないが、段階的な濃さを代表する分布を取得して閾値算出に利用するという今回の目的では有効に働いていた。今後、組織標本の代わりに、特定の HER2 スコアに該当する細胞を培養して一枚のスライドに配置したコントロールスライド(図 43)を活用することで、確率密度分布の差分をとることなく各スコアを代表する濃さの分布を取得することも可能になると考える。コアごとの分布や、それらの分布を線形結合することで組織標本の分布に近似されるかを確認し、新たな手法の提案につなげたい。

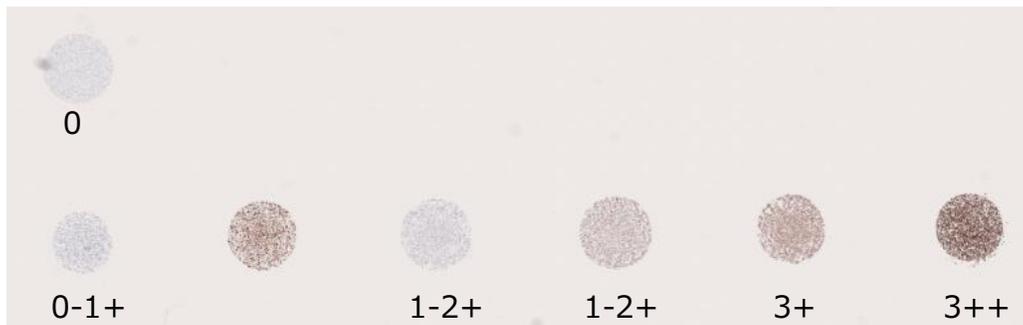


図 43 コントロールスライドの WSI の例

4.5 まとめ

本章では、第 3 章で提案した IHC キャリブレーションスライドを用いた補正手法の有効性の検証を行った。まずは染色のばらつきの補正効果を検証するために、異なる日に染色された乳がんの組織標本スライドと IHC キャリブレーションスライドを 1 台の WSI スキャナで撮影して実験を行った。提案手法を適用していない画像、適用した画像それぞれを自動解析ソフトで評価して HER2 スコアを取得し、病理医の診断結果との一致率を比較した。また、H-score の標準偏差を比較することで染色のばらつきを比較した。

補正を行わなかった場合には病理医の診断と自動解析とで不一致となるデータがみられた。データを取得した施設では染色の品質管理を特に厳しく実施しているが、そのような環境であっても日々の染色の小さなばらつきが自動解析結果に影響することが確認された。提案手法の適用により、病理医の判定結果との一致率や H-score のばらつきが改善されることを確認した。他施設で染色されたデータではもともとのばらつきや補正効果の程度も変わってくると想定される。今後は複数の研究室や施設で染色したデータを用いて、本手法の拡張性を検証することが求められる。

第5章 スキャナによるばらつきの補正効果の検証

5.1 概要

ここでは、染色とスキャナの両方のばらつきがあるデータを対象とし、IHC キャリブレーションスライドによる補正効果、カラーチャートスライドによる補正を組み合わせた場合の補正効果の検証を行った内容について記載する。

5.2 実験内容

5.2.1 実験データ

実験データは 4.2.1 節と同じく、HER2 スコア 0、1+、2+、3+ と診断された 4 種類のスライド 1 枚ずつと IHC キャリブレーションを追加した 5 スライドのデータセットを用いた。染色済みの 6 データセットを実験 5.2.5 に利用した。また、実験 5.2.6 では、浸潤性乳がんから成る TMA スライド 1 枚と IHC キャリブレーションを利用した。

スキャンには 2 台の WSI スキャナを使用した。1 台は 4.2.1 節と同様に S60 を使用した。もう 1 台は Panoramic 250 (3D Histech Ltd.) を使用し、分解能 $0.17 \mu\text{m}/\text{pixel}$ で撮影を行った。本スキャナは以降、P250 と記載する。P250 の画像は tiff 形式でエクスポートし、解析に用いた。

5.2.2 色補正関数算出手法および多項式変換の次数の検討

色補正関数の算出において、白色性維持の制約の有無による色変換後画像の比較を行った。制約を加えない場合を LS 法、制約を加えた場合を WPPLS 法として記載する。また、多項式変換を用いた場合の次数の影響についても調査するため、次数 3~8 で多項式変換を行い、LS 法、WPPLS 法それぞれで色変換行列を算出した。得られた補正関数を組織画像に適用し、補正後画像を比較した。

2 台のスキャナで撮影した画像に対して一連の処理を同様に行い、スキャナ間の画像の色差を計算した。スキャナ間の色差を評価するために、2 つの画像の色差をピクセル単位で計算し、dE*マップを作成した。dE*マップの各画素は、2 つの画像の同じ画素の色差 dE*を表している。2 台のスキャナから得られる画像は画像サイズや分解能が異なるため、dE*マップの作成に当たっては画像位置合わせや倍率補正が必要である。本実験では、画像レジストレーションに、2 枚の画像上のキーポイントを検出しマッチングする AKAZE (Accelerated-KAZE) [58]を採用し

た。S60 の画像を対象画像とし、P250 の画像に AKAZE を適用してスキャナ間の画像の位置合わせを行った。作成した dE^* マップからビン数 64 でヒストグラムを作成し、色差の分布を確認した。補正前、補正後それぞれの色差の分布から dE^* の平均値と標準偏差を比較した。

5.2.3 色補正効果の検証：スキャナ間の色差の算出

カラーチャートスライドを用いた色補正手法を適用することでの、スキャナ間の色のばらつきの補正効果の検証を行った。5.2.3 節の実験にて選択した多項式変換の次数で色変換行列を算出した。ここでは 2 台のスキャナでスライドをスキャンされた H&E 染色の組織標本画像、IHC 法の組織画像および IHC キャリブレーション画像を色補正対象とした。実験 5.2.2 と同様、スキャナ間の dE^* マップを作成して補正効果を比較した。ここで、IHC キャリブレーションについては、図 44(a)に示すように背景部分の占める割合が大きいため、画像全体でヒストグラムを作成するとマイクロビーズ部分の変化がわかりにくくなってしまっていた。そのため、マイクロビーズ領域のみに焦点を当てたマスク画像 (図 44(b)) を手動で作成し、背景部分を除いて色差を比較した。

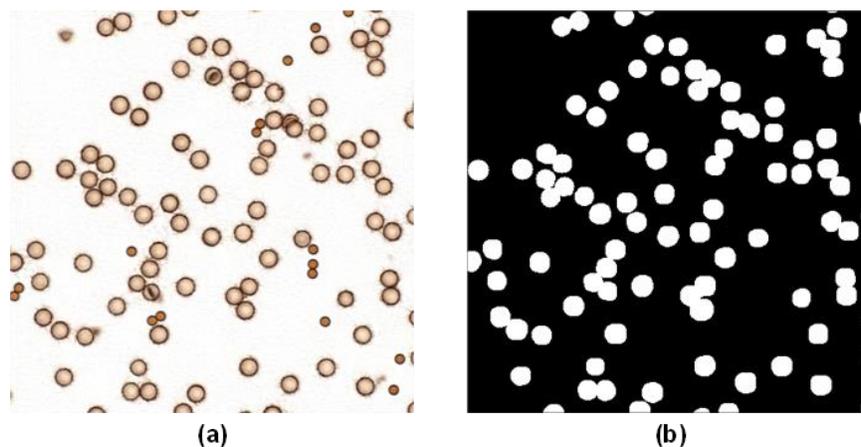


図 44 IHC キャリブレーション画像の例；(a) 元画像；(b) マスク画像

5.2.4 カラーチャートの色のシミュレーション

先行研究[39]では IHC 画像に適したカラーキャリブレーション用スライドとして、茶色や黄色を中心としたカラーチャートスライドを提案していた (図 45)。本研究のカラーチャートスライドには茶色は含まれない。そこで、IHC 法に対応した色を含めた場合の色補正効果を比較するため、シミュレーション実験を行った。オリジナルのカラーチャートスライドのパッチから

得られた 3×13 の行列 \mathbf{G}_R 、 \mathbf{G}_T に茶色を 1 色または 2 色を実験的に追加した 3×14 または 3×15 の行列 \mathbf{G}_R' 、 \mathbf{G}_T' を取得し、色補正行列 \mathbf{M}' を算出した。

基準色 \mathbf{G}_R' に追加する茶色の色には、通常の 13 色を使って色補正を行ったキャリブレーション画像の大きいビーズから取得した色素成分ベクトルを利用した。また、スキャナの色 \mathbf{G}_T' には、各スキャナで撮影したキャリブレーション画像の小さいビーズから取得した色素成分ベクトルを利用した。得られた色素成分ベクトルを利用し、2 種類の濃さに対応する茶色を用意した (図 46)。茶色を追加しない場合 (オリジナルの 13 色使用)、1 種類の茶色を追加した場合 (14 色) を 2 パターン、2 種類の茶色を追加した場合 (15 色) の計 4 パターンで色補正関数を取得した。取得した色補正関数を組織画像およびキャリブレーション画像に適用し、実験 5.2.3 と同様にスキャナ間の dE^* マップを作成して補正効果を比較した。



図 45 先行研究で IHC 法用に提案されたカラーチャート [39]

#	濃さ	基準	S60	P250
1	レベル10相当			
2	小さいマイクロビーズ相当			

図 46 シミュレーションで使用した茶色

5.2.5 HER2 の自動解析における提案手法の補正効果の検証

6 データセットのうち 1 データセットを基準データ、5 データセットを対象データとみなし、提案手法によるスキャナのばらつきの補正効果、および、HER2 の自動解析結果への影響を検証した。基準となる 1 データセットは、4.2.3 節の実験と同様のものを選択した。対象データは臨床では様々な設定でスキャンが行われることを想定し、ガンマ値 1.8 または 2.2 で補正を加えたデータを準備した。データセットとガンマ補正值の組み合わせはランダムに選択し、合計 10 個のデータセットを準備した (表 9)。

表9 対象データのガンマ補正の設定

スキャナ	ガンマ値	データ数
S60	1.0 (リニア)	2
	1.8	2
	2.2	1
P250	1.0 (リニア)	2
	1.8	1
	2.2	2
Total		10

HER2の自動解析には、オープンソースソフトウェアのQuPathを用いた。スキャナと染色のキャリブレーション方法の異なる組み合わせについて、図47に示す6パターンを比較した。評価指標としては、HER2スコアに関して、病理医の診断結果と自動解析で得られた結果との一致率を比較した。また、H-scoreの標準偏差を染色のばらつきの指標として比較した。得られたH-scoreに対して二要因のANOVAで検定を行い、染色、スキャナに関して補正を行わなかった場合と提案手法を適用した場合で有意差が見られるか検定を行った。ANOVAで有意差が見られた際の事後検定法にはTukey法を用いた。それぞれの検定において、 p 値が0.05以下の際に統計的に有意差があるとした。

		CCS	IHC-CS	
				
			Method1: Thresholds	Method2: Images
0-0	No	No	No	No
0-1	No	Yes	No	No
0-2	No	No	Yes	Yes
1-0	Yes	No	No	No
1-1	Yes	Yes	No	No
1-2	Yes	No	Yes	Yes

図47 カラーチャートとIHCキャリブレーションの組み合わせ

5.2.6 基準データと異なる組織画像での補正効果の検証

本節では、基準データは 4.2.3 節の実験と同様の組織標本を利用し、異なる組織標本である TMA スライドを対象データとして評価した。TMA スライドはガンマ補正なしで撮影したリニア RGB 画像を解析に用いた。QuPath で自動解析を行い、図 47 に示す 6 パターンについて病理医の診断結果と自動解析で得られた HER2 スコアとの一致率を比較した。

5.3 実験結果

5.3.1 色補正関数算出手法および多項式変換の次数の検討

図 48、図 49 に LS 法で、図 50、図 51 に WPPLS 法で色変換行列を取得した場合の、各スキャナの色補正後画像を示す。多項式変換の次数は 3~8 で比較している。LS 法を利用した場合、いずれのスキャナでも次数 3 のときに背景の青みが強い画像となっていた。また、S60 では次数 4 以上の場合に背景の赤みが強い画像となっていた。一方で、WPPLS 法を適用することにより、背景の白色が維持された画像が取得された。ただし、S60 では次数が 8 まで増えると背景部分も含めて赤みが強い画像となっていた。H&E 染色では目で見た際の次数による色の違いは大きくはなかった。IHC 画像では次数が大きくなると DAB 部分の黄味がやや強くなる傾向がいずれのスキャナでもみられた。病理医に色補正後の IHC 法の組織画像を確認してもらったところ、WPPLS 法で次数 3~5 で変換した画像が目を見た際に自然な色に感じるとの意見であった。

図 52 にスキャナ間の色差を比較した結果を示す。WPPLS 法で取得された色補正行列で組織画像の色補正を行った結果を示している。S60 において赤みが強かった次数 8 は除き、次数 3~7 で比較した。IHC 法、H&E 染色組織画像それぞれについて、色差マップ (図 52 (a)) を作成し、色差マップの平均値を比較した (図 52 (c))。次数の組み合わせは S60-P250 の順で示している。H&E 染色画像では P250 の次数が上がるほど色差が小さくなる傾向がみられた。また、IHC 画像では P250 の次数が小さいときに比較的色彩差が小さくなる傾向がみられた。特に、次数 5-5 の組み合わせで色差が小さくなっていた。病理医の意見や IHC 画像でのスキャナ間の色差の傾向を踏まえ、本研究では次数 5 を選択した。

次数	カラーチャート	H&E染色	IHC法
元画像			
3			
4			
5			
6			
7			
8			

図 48 スキャナ S60 について白色の制約なしで色補正行列を求めた結果

次数	カラーチャート	H&E染色	IHC法
元画像			
3			
4			
5			
6			
7			
8			

図 49 スキャナ P250 について白色の制約なしで色補正行列を求めた結果

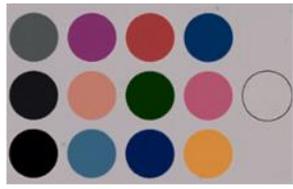
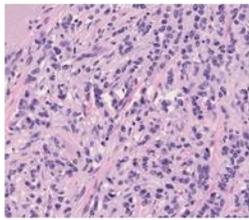
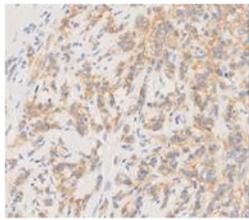
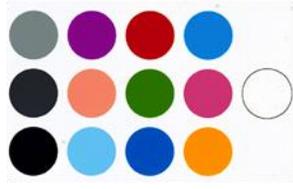
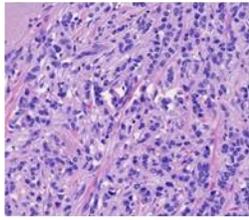
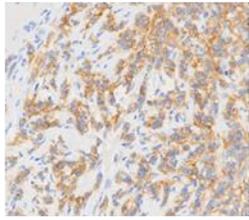
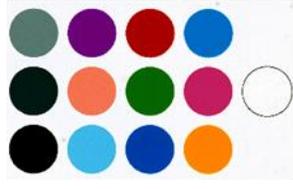
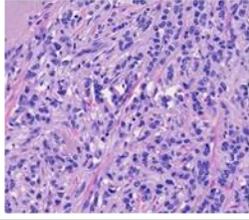
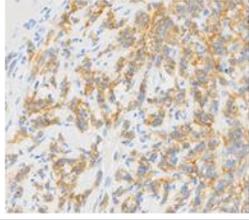
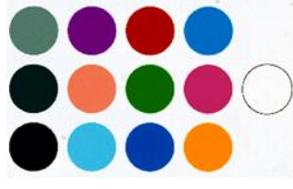
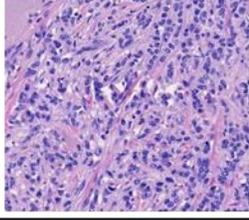
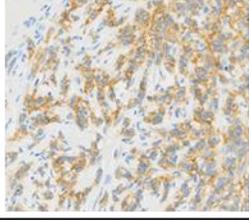
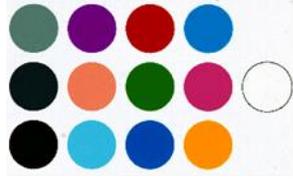
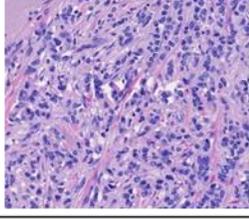
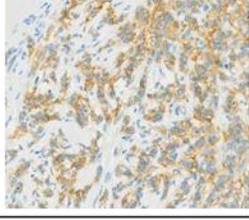
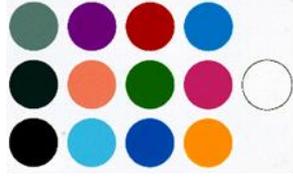
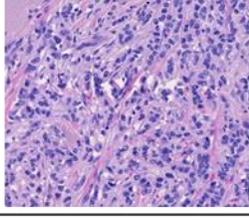
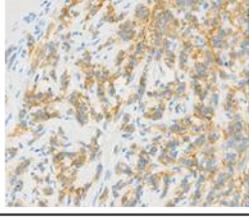
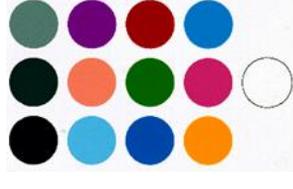
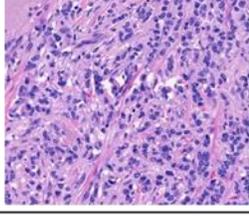
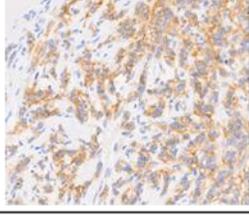
次数	カラーチャート	H&E染色	IHC法
元画像			
3			
4			
5			
6			
7			
8			

図 50 スキャナ S60 について白色の制約ありで色補正行列を求めた結果

次数	カラーチャート	H&E染色	IHC法
元画像			
3			
4			
5			
6			
7			
8			

図 51 スキャナ P250 について白色の制約ありで色補正行列を求めた結果

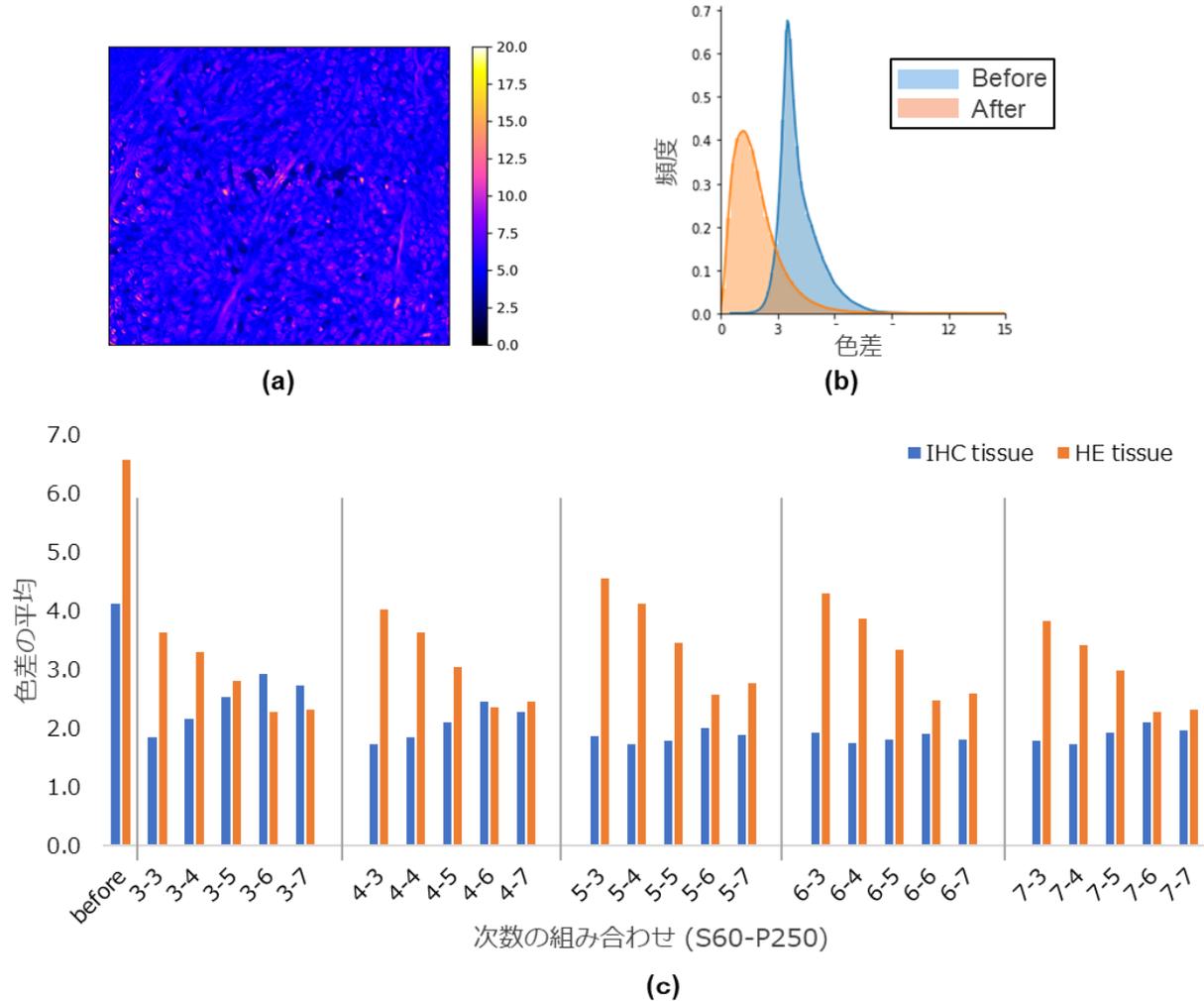


図 52 制約ありで色補正関数を算出する際の次数の選択；(a) 2 つのスキャナの画像間の色差マップ；(b) 色差マップのヒストグラム；(c) 次数の組み合わせによる色差の平均値の比較（次数は S60-P250 の順）

5.3.2 色補正効果の確認

図 53–図 56 に、色補正の例を示す。ここで示す結果は次数 5、WPPLS 法で取得したものである。図 53(b)では、カラーチャートの 13 個のパッチについて、色補正前後のスキャナ間の色差を計算した。補正前は色差の平均値が 15.1、最大値が 44.1 であったが、補正によって平均値が 4.2、最大値が 7.4 となった。H&E 染色の組織画像 (図 54)では、紫色の細胞核部分でスキャナ間の色の違いが顕著に見られていたが、色補正によって色差が小さくなったことが確認された。色差の平均値は 6.2 から 2.5 へと減少した。IHC 法の組織画像 (図 55)では、茶色の細胞膜部分で特にスキャナ間の色の違いが見られていたが、色補正によってほとんどの画素で色差が 4 以下

となった。色差の平均値は 4.1 から 1.6 へと減少した。IHC キャリブレーション画像 (図 56) では、組織画像に比べると色補正の効果は小さかったが、色差の平均値は 4.7 から 3.9 へと減少した。

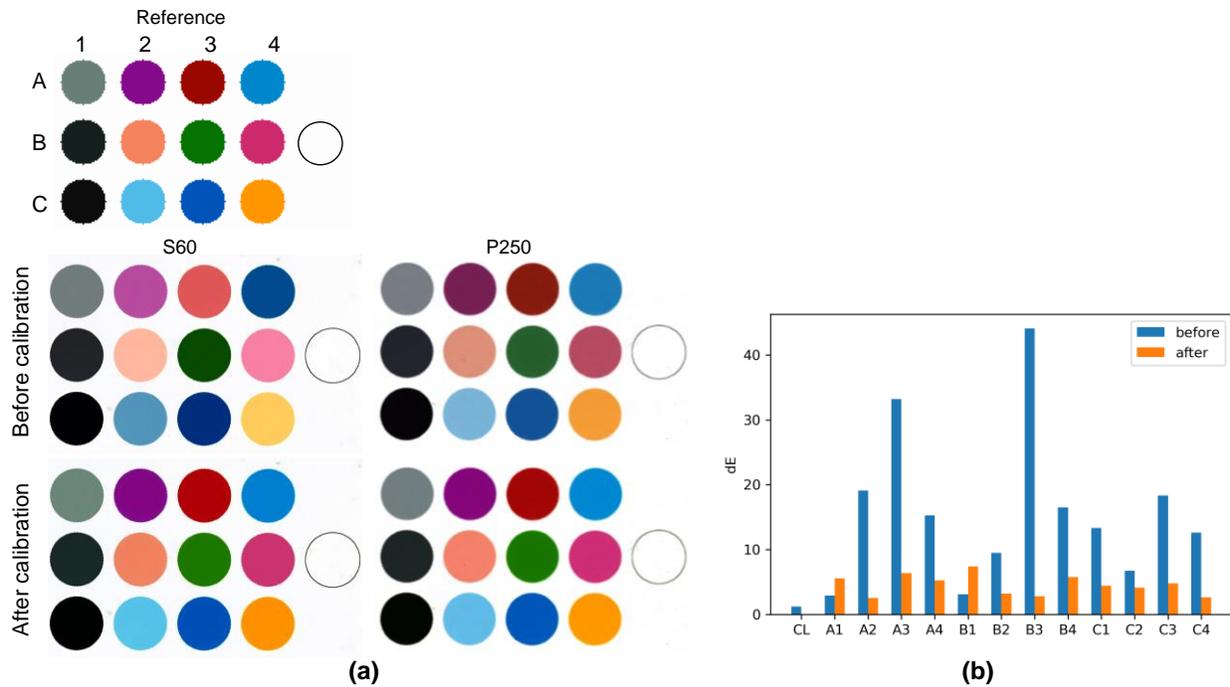


図 53 色補正結果の比較 (カラーチャート画像) ; (a) 上段: カラーチャートの基準の RGB 画像、中段: 色補正前画像、下段: 色補正後画像; (b) パッチごとの色補正前後のスキヤナ間の色差の比較

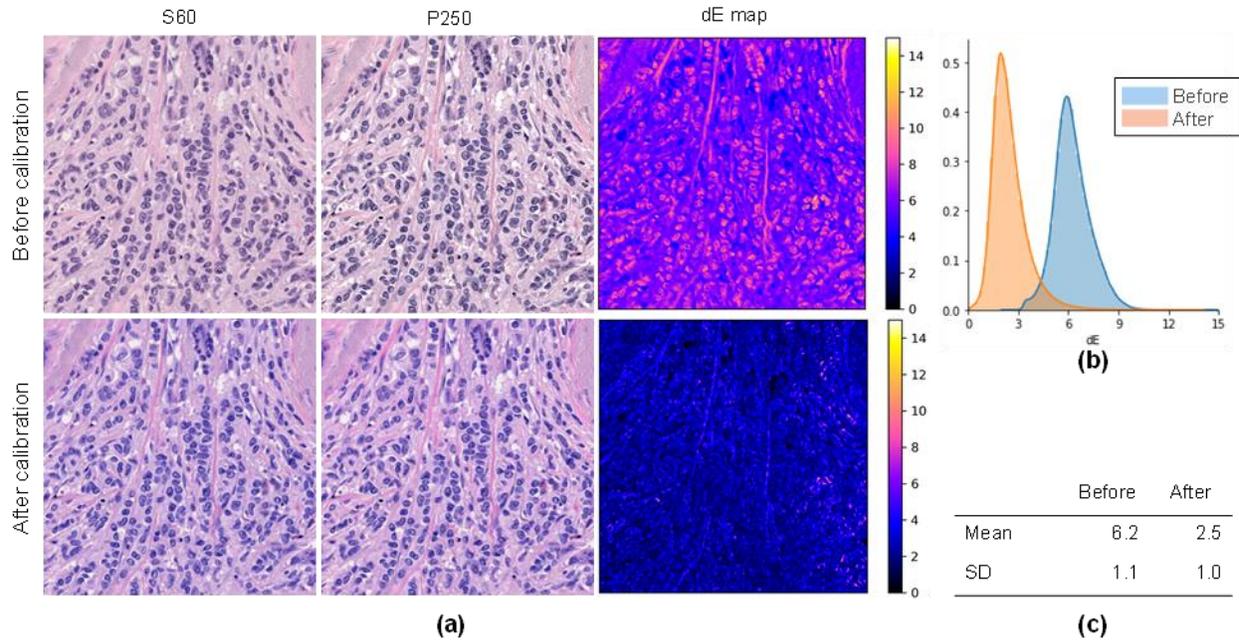


図 54 組織画像におけるスキャナ間の色差の比較 (H&E 染色) ; (a) 上段 : 色補正前画像、下段 : 色補正後画像 ; (b) 色差マップから作成した色差のヒストグラム ; (c) 色差のヒストグラムの平均値と標準偏差

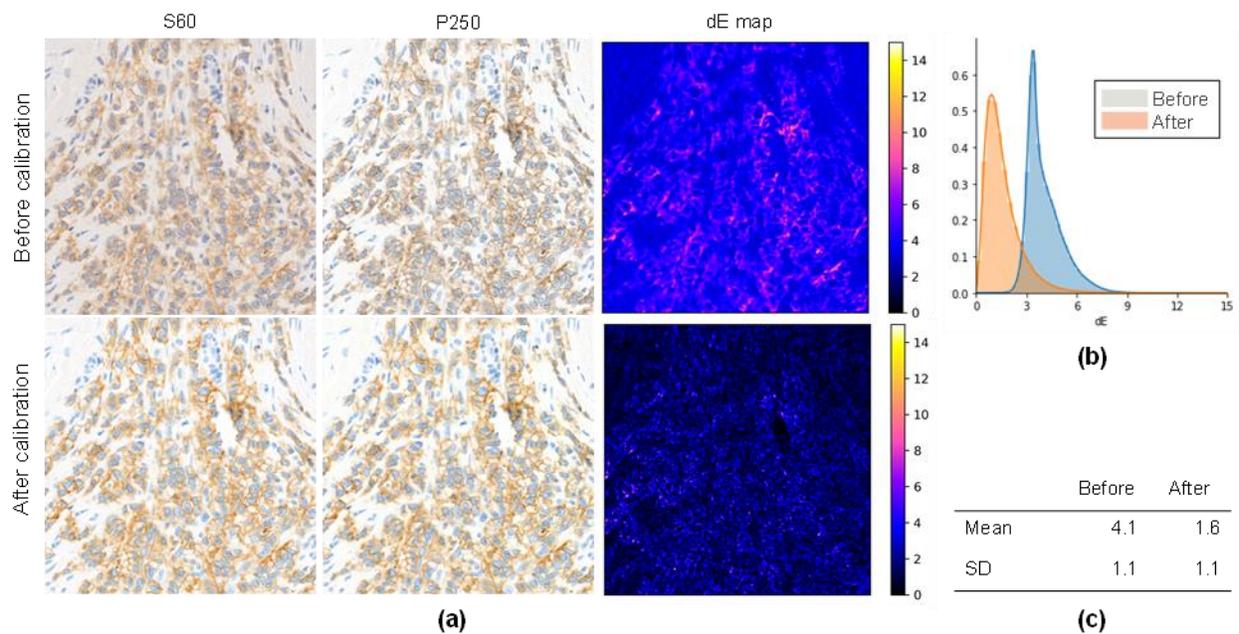


図 55 組織画像におけるスキャナ間の色差の比較 (IHC 法) ; (a) 上段 : 色補正前画像、下段 : 色補正後画像 ; (b) 色差マップから作成した色差のヒストグラム ; (c) 色差のヒストグラムの平均値と標準偏差

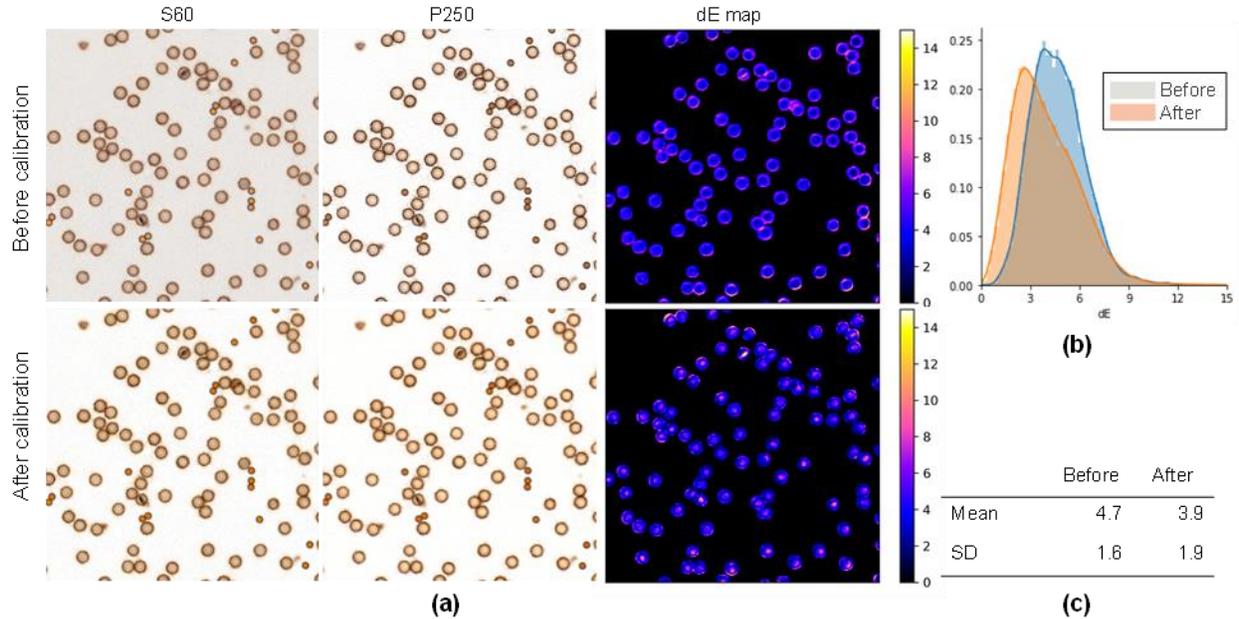


図 56 キャリブレーション画像におけるスキャナ間の色差の比較；(a) 上段：色補正前画像、下段：色補正後画像；(b) 色差マップから作成した色差のヒストグラム；(c) 色差のヒストグラムの平均値と標準偏差

5.3.3 カラーチャートの色のシミュレーション

シミュレーションによってカラーチャートに茶色を追加した場合の、H&E 染色の組織画像、IHC 法の組織画像およびキャリブレーション画像の結果を図 57–図 62、表 10–表 12 に示す。H&E 染色の組織画像では、茶色を追加することで色差の平均が 0.1 小さくなる効果が確認された。ただし非常に小さな差であるため、色変換後画像の比較では見た目での印象はほとんど変わらない結果であった。IHC 法の組織画像では、濃い茶色を追加した場合には色差の数値の変化は見られなかったが、薄い茶色を追加した場合にやや数値が悪くなった。また、キャリブレーション画像では、いずれの茶色を追加した場合でも色差 6 以上の割合が減る効果が見られた。特に濃い茶色を追加した場合に顕著であった。ただし、こちらも小さな差であるため、図 60、図 62 の色変換後画像の比較ではほとんど印象が変わらない結果であった。

同一のキャリブレーションスライドを 2 つのスキャナで撮影した画像について、色補正後のキャリブレーション画像から取得した濃さの関係性を表 13 に示す。同一スライドの濃さの関係性を求めているため、スキャナ間の色の違いがない場合には関係性は 1.0 に近づくと考えられる。茶色を追加しなかった場合でも 1.0 に近い値が得られており、茶色を追加するごとに 1.0 よりも小さくなる傾向がみられた。

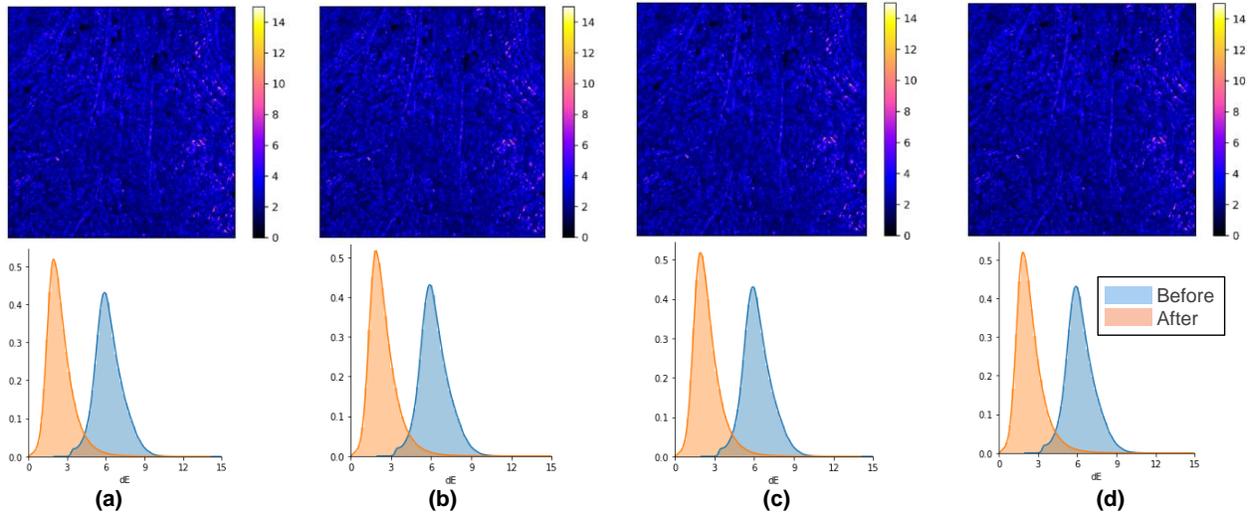


図 57 組織画像でのスキャナ間の色差マップと色差のヒストグラムの比較 (H&E 染色)；(a) 追加無し；(b) 濃い茶色を追加；(c) 薄い茶色を追加；(d) 両方の茶色を追加

表 10 組織画像でのスキャナ間の色差の数値比較 (H&E 染色)

	補正前	補正後			
		追加なし	濃い茶色	薄い茶色	両方
色差 6 以上の割合 [%]	55.2	0.9	0.9	0.9	0.9
色差の平均	6.2	2.5	2.4	2.4	2.4

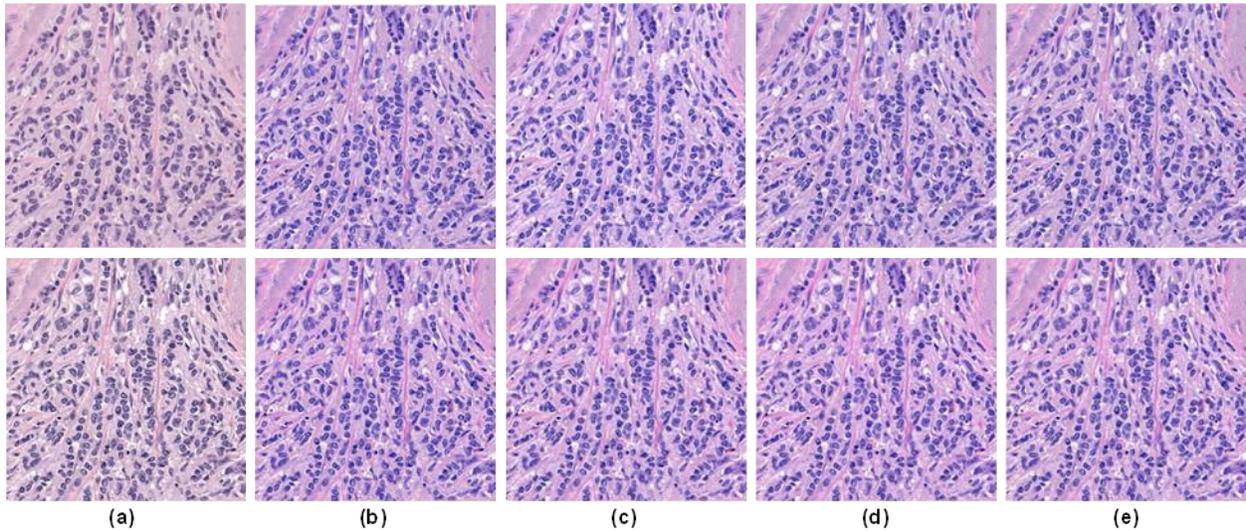


図 58 組織画像の色変換後の比較 (H&E 染色) 上段 S60、下段 P250；(a) 補正前；(b) 茶色の追加なし；(c) 濃い茶色を追加；(d) 薄い茶色を追加；(e) 両方の茶色を追加

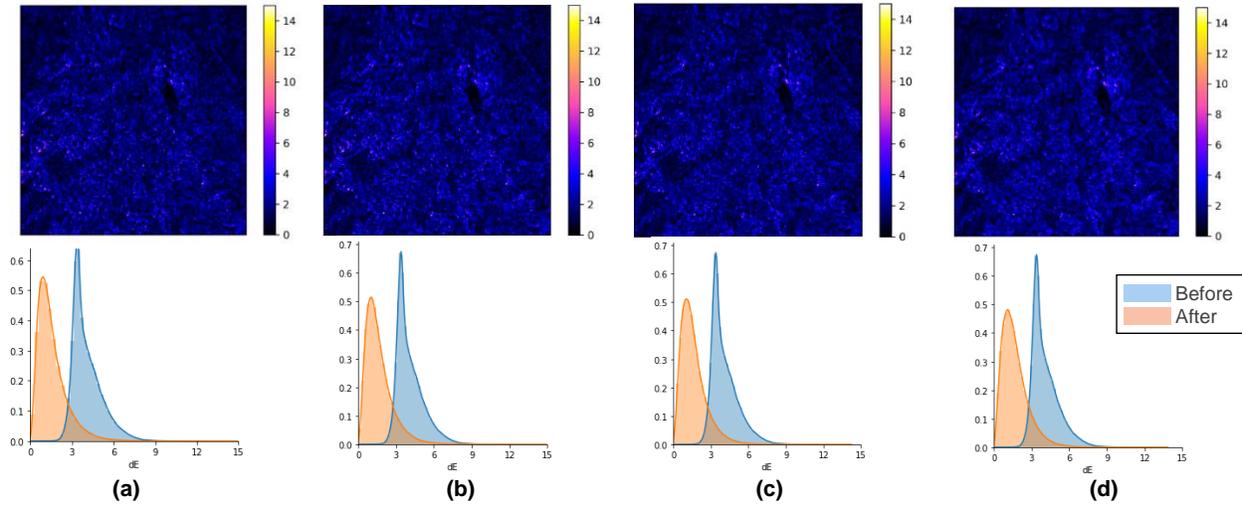


図 59 組織画像でのスキャナ間の色差マップと色差のヒストグラムの比較 (IHC 法) ; (a) 追加無し ; (b) 濃い茶色を追加 ; (c) 薄い茶色を追加 ; (d) 両方の茶色を追加

表 11 組織画像でのスキャナ間の色差の数値比較 (IHC 法)

	補正前	補正後			
		追加なし	濃い茶色	薄い茶色	両方
色差 6 以上の割合 [%]	6.1	0.4	0.4	0.5	0.5
色差の平均	4.1	1.7	1.7	1.8	1.9

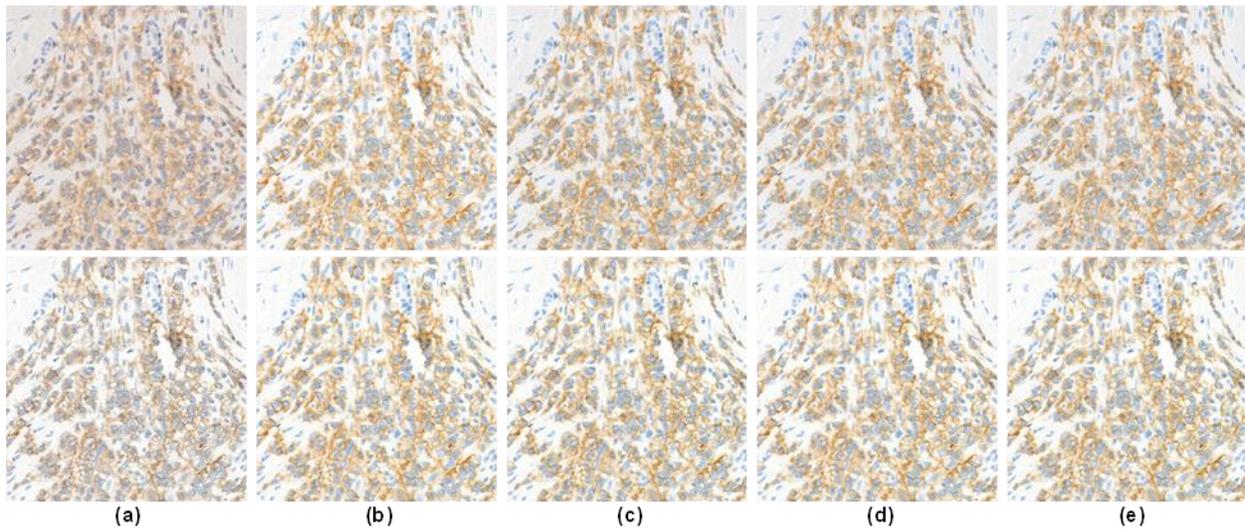


図 60 組織画像の色変換後の比較 (IHC 法) 上段 S60、下段 P250 ; (a) 補正前 ; (b) 茶色の追加なし ; (c) 濃い茶色を追加 ; (d) 薄い茶色を追加 ; (e) 両方の茶色を追加

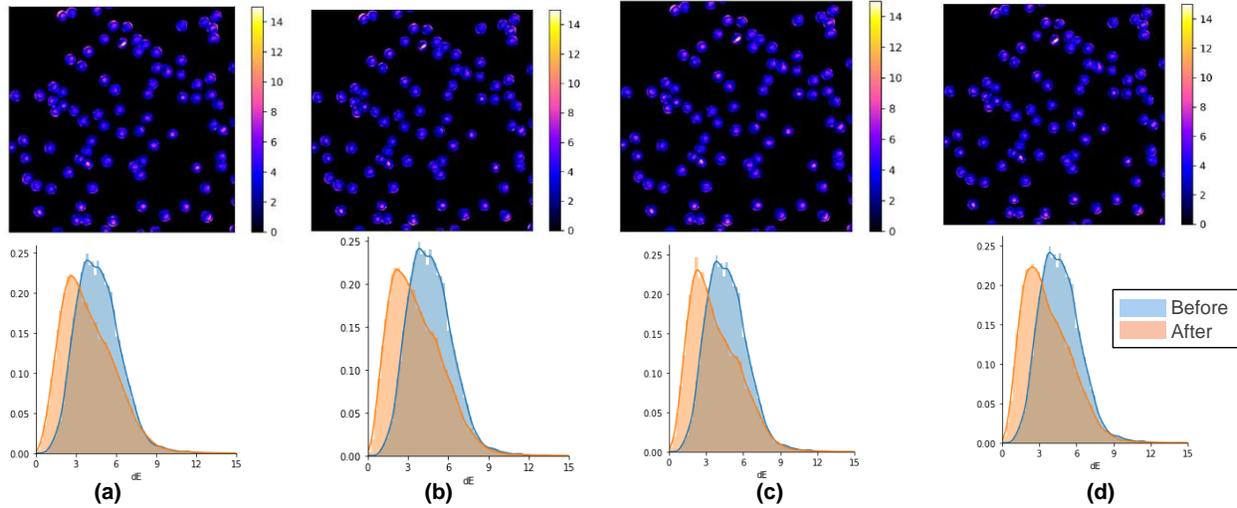


図 61 キャリブレーション画像でのスキャナ間の色差マップと色差のヒストグラムの比較 (IHC 法) ; (a) 追加無し ; (b) 濃い茶色を追加 ; (c) 薄い茶色を追加 ; (d) 両方の茶色を追加

表 12 キャリブレーション画像でのスキャナ間の色差の数値比較

	補正前	補正後			
		追加なし	濃い茶色	薄い茶色	両方
色差 6 以上の割合 [%]	21.2	15.5	12.9	14.1	12.6
色差の平均	4.8	3.9	3.7	3.7	3.7

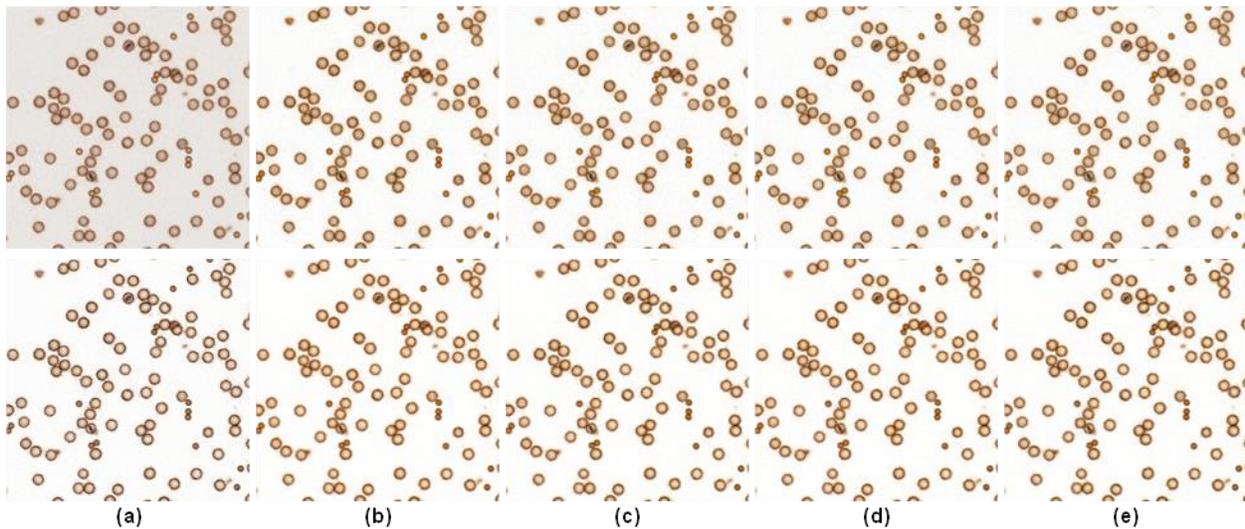


図 62 キャリブレーション画像の色変換後の比較(IHC 法)上段 S60、下段 P250 ; (a) 補正前 ; (b) 茶色の追加なし ; (c) 濃い茶色を追加 ; (d) 薄い茶色を追加 ; (e) 両方の茶色を追加

表 13 同一のキャリブレーションにおけるスキャナ間の濃さの関係

	補正後			
	追加なし	濃い茶色	薄い茶色	両方
濃さの係数	1.01	0.99	0.95	0.93

5.3.4 HER2 の自動解析における提案手法の補正効果の検証結果

表 14 に自動解析ソフトウェアによって得られた HER2 評価結果を示す。カラーチャートによる補正はここまでの実験を踏まえ、次数 5、WPLS 法、もともとの 13 色のパッチを用いて行われた。補正を適用しない場合 (0-0)、1+、2+、3+ それぞれのケースで病理医の診断結果との不一致が見られ、30 スライドのうち 14 スライドで結果の不一致が確認された。IHC キャリブレーションスライド単体で補正した場合 (手法 0-1) には、判定結果の不一致は 3 スライド、カラーチャートスライド単体で補正した場合 (手法 1-0) には、判定結果の不一致は 1 スライドに減少する効果が確認された。さらに、IHC キャリブレーションスライドとカラーチャートスライドの両方で補正を行った場合 (手法 1-1、1-2)、病理医の診断結果との一致率は 100% となった。また、いずれの染色キャリブレーション方法でも、補正を適用しない場合に比べて H-score の標準偏差が小さくなることが確認された。

図 57、図 58 に色補正後画像の例を示す。左の一枚は基準画像 (S60、ガンマ補正なし)、その他の 9 枚は評価対象画像 (上段：S60、 $\gamma 1.8$ でガンマ補正、中段：P250、 $\gamma 2.2$ でガンマ補正、下段：P250、ガンマ補正なし) であり、左から色補正前、IHC キャリブレーションでの色補正後画像 (手法 0-2)、IHC キャリブレーションとカラーチャートの両方で色補正後画像 (手法 1-2) である。IHC キャリブレーションのみで色補正を行った場合 (手法 0-2)、DAB の色と強度のみが補正対象となる。そのため、ガンマ補正された画像が対象の場合、ヘマトキシリンの色と濃さは補正されないため、補正後の画像も依然として明るくなっている。リニア RGB 画像が対象となる場合、DAB の色と濃さを補正することで基準画像に近い画像を得ることができている。両方のスライドを使って色補正を行った場合 (手法 1-2)、始めに逆ガンマ補正によって画像全体の明るさが補正されるため、DAB とヘマトキシリンの両方を補正することができる。

図 59 は、基準画像と対象画像の H-score の差を HER2 のケースごとに箱ひげ図として比較した結果を示す。二要因の ANOVA の結果、1+ のケースでは $F(2,54) = 5.70$ 、 $P = 0.006$ 、2+ のケースでは $F(2,54) = 4.53$ 、 $P = 0.015$ 、3+ のケースでは $F(2,54) = 3.69$ 、 $P = 0.031$ と染色の補正とスキャナの補正の交互作用の有意差が確認された。表 11 は、ANOVA で統計的な有意差がみられたケースにおける Tukey 検定を用いた事後検定の結果を示す。

表 14 キャリブレータの組み合わせごとの HER2 評価結果

Methods	Color chart	IHC calibrator		Concordance (%)			SD of H-score		
		Method1	Method2	1+	2+	3+	1+	2+	3+
0-0	No	No	No	90	30	40	30.3	27.5	47.6
0-1	No	Yes	No	100	100	70	11.0	11.9	27.1
0-2	No	No	Yes	100	100	100	13.3	10.1	16.2
1-0	Yes	No	No	90	100	100	11.2	10.3	8.7
1-1	Yes	Yes	No	100	100	100	7.9	8.4	5.9
1-2	Yes	No	Yes	100	100	100	10.4	7.0	5.3

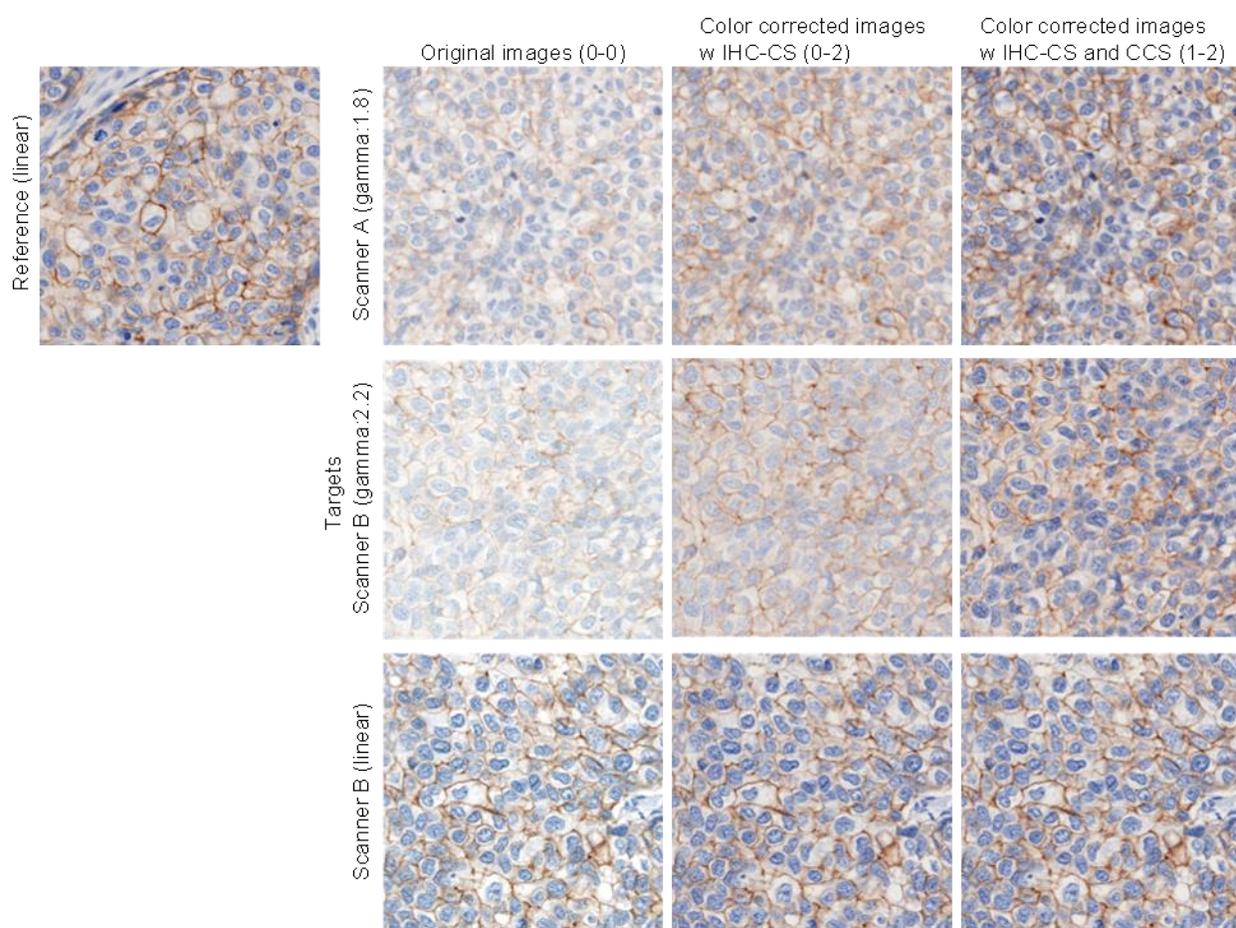


図 63 色補正後画像の比較(2+)

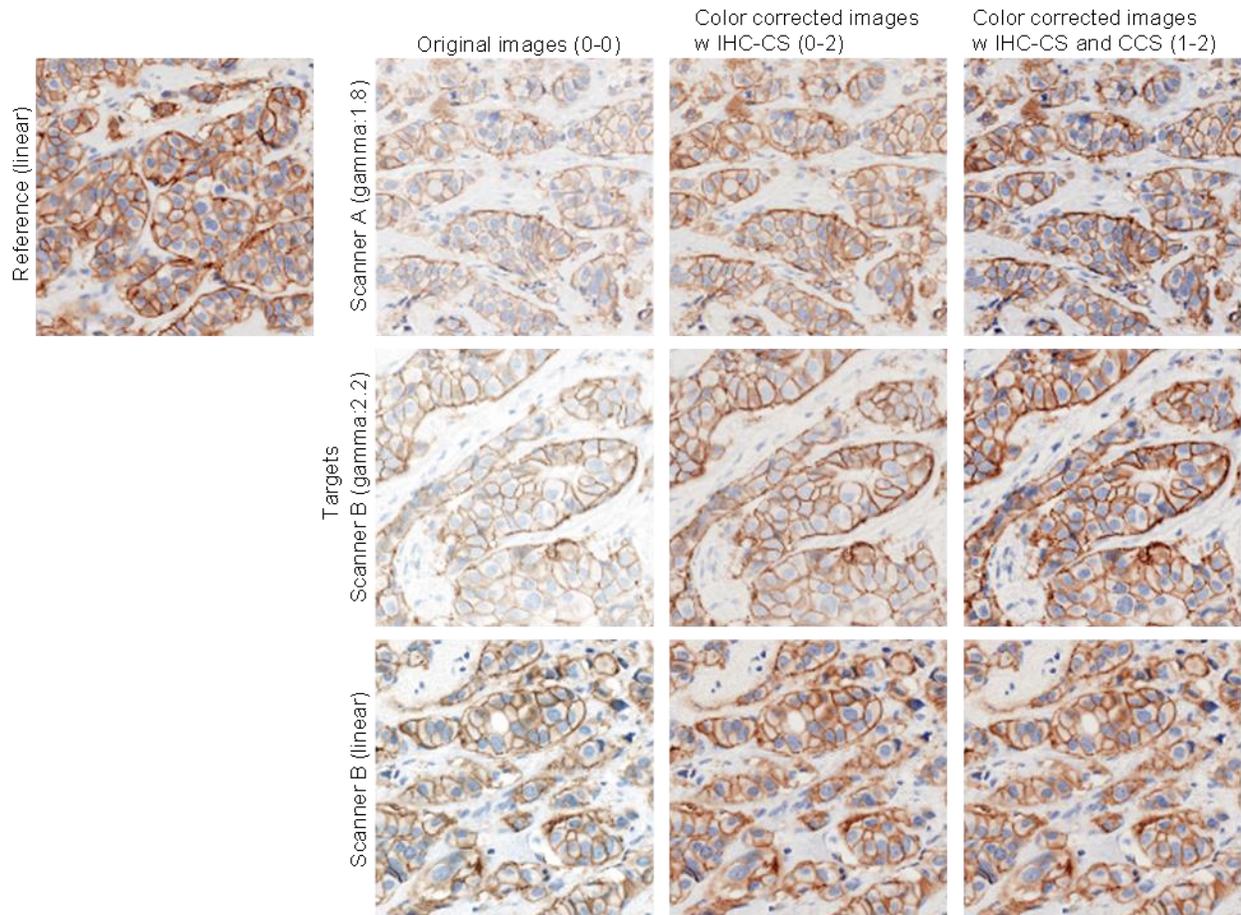


図 64 色補正後画像の比較(3+)

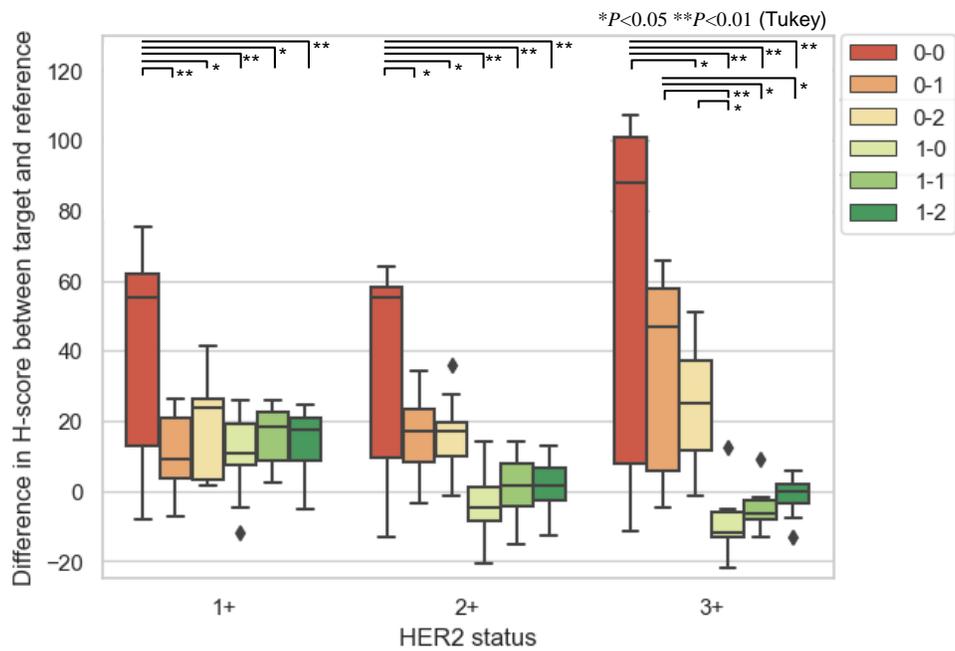


図 65 基準画像と評価対象画像の H-score の標準偏差の比較

表 15 有意差のあるケース Tukey 検定

HER2 score	group1	group2	Mean diff.	p	95% C.I.	
					Lower	Upper
1+	0-0	0-1	-28.53 **	0.003	-54.20	-2.87
	0-0	0-2	-21.82 *	0.043	-43.19	-0.45
	0-0	1-0	-29.25 **	0.002	-54.92	-3.59
	0-0	1-1	-24.45 *	0.016	-45.82	-3.08
	0-0	1-2	-26.62 **	0.007	-52.29	-0.95
2+	0-0	0-1	-20.91 *	0.034	-40.81	-1.01
	0-0	0-2	-20.63 *	0.038	-40.53	-0.73
	0-0	1-0	-39.44 **	0.001	-63.34	-15.53
	0-0	1-1	-35.43 **	0.001	-59.33	-11.52
	0-0	1-2	-35.33 **	0.001	-59.24	-11.43
3+	0-0	0-2	-34.31 *	0.039	-67.46	-1.16
	0-0	1-0	-68.91 **	0.001	-108.73	-29.10
	0-0	1-1	-64.86 **	0.001	-104.67	-25.04
	0-0	1-2	-61.15 **	0.001	-100.96	-21.34
	0-1	1-0	-43.75 **	0.004	-83.56	-3.93
	0-1	1-1	-39.69 *	0.010	-72.84	-6.54
	0-1	1-2	-35.99 *	0.026	-69.14	-2.84
	0-2	1-0	-34.60 *	0.036	-67.75	-1.45

* $P<0.05$ ** $P<0.01$

5.3.5 基準データと異なる組織画像の補正効果の検証結果

表 12 は基準データと異なる組織標本を評価データとした場合の HER2 の評価結果を示す。補正を行わなかった場合でも、ほとんどのケースで病理医の評価と一致するスコアが得られているが、5 スライドで結果の不一致がみられた。提案手法を用いて染色、スキャナの両方の補正を行うことで、100%の一致率が確認された。

表 16 基準データと異なる組織画像を対象として HER2 評価を行った結果

		0-0					0-1					0-2					
		Auto					Auto					Auto					
		0	1+	2+	3+	total	0	1+	2+	3+	total	0	1+	2+	3+	total	
Pathologist	0	6	4			10	0	10			10	0	10			10	
	1+		4			4	1+	4			4	1+	4			4	
	2+			9	1	10	2+		10		10	2+		10		10	
	3+				10	10	3+			10	10	3+			10	10	
total		6	8	9	11	34		10	4	10	10		10	4	10	10	34
		Concordance			85.3%		Concordance			100.0%		Concordance			100.0%		

		1-0					1-1					1-2					
		Auto					Auto					Auto					
		0	1+	2+	3+	total	0	1+	2+	3+	total	0	1+	2+	3+	total	
Pathologist	0	10				10	0	10			10	0	10			10	
	1+		4			4	1+	4			4	1+	4			4	
	2+			9	1	10	2+		10		10	2+		10		10	
	3+				10	10	3+			10	10	3+			10	10	
total		10	4	9	11	34		10	4	10	10		11	4	10	10	34
		Concordance			97.1%		Concordance			100.0%		Concordance			100.0%		

5.4 考察

5.4.1 カラーチャートを用いたスキャナの色補正

LS 法を適用した際には、色変換後画像の背景で青みや赤みが目立っていたが、WPPLS 法を適用することで白色性が維持された色補正後画像を得ることができた。LS 法、WPPLS 法それぞれ回数 3 で補正されたカラーチャート画像の各パッチの RGB 値を変換し、xy 色度図上で比較した結果を図 66 に示す。xy 色度図上で見ると LS 法、WPPLS 法の差はそれほど大きくはないが、背景部分を拡大すると、LS 法ではわずかに青緑方向に寄っていることがわかった(図 66(b))。このようなわずかな差が色補正した際には目に見える差となり、背景の青みとして表れたと考えられる。また、本来は色度図の座標だと(0.3,0.3)周辺にあるはずの黒色が 1 点外れた位置にある。黒は RGB 値いずれも 0 に近い値が得られているが、R 値、G 値が 0.000 程度なのに対し、B

値が0.023程度と比較して大きい値であったため、青方向に寄ったプロットとなっていたと考えられる。

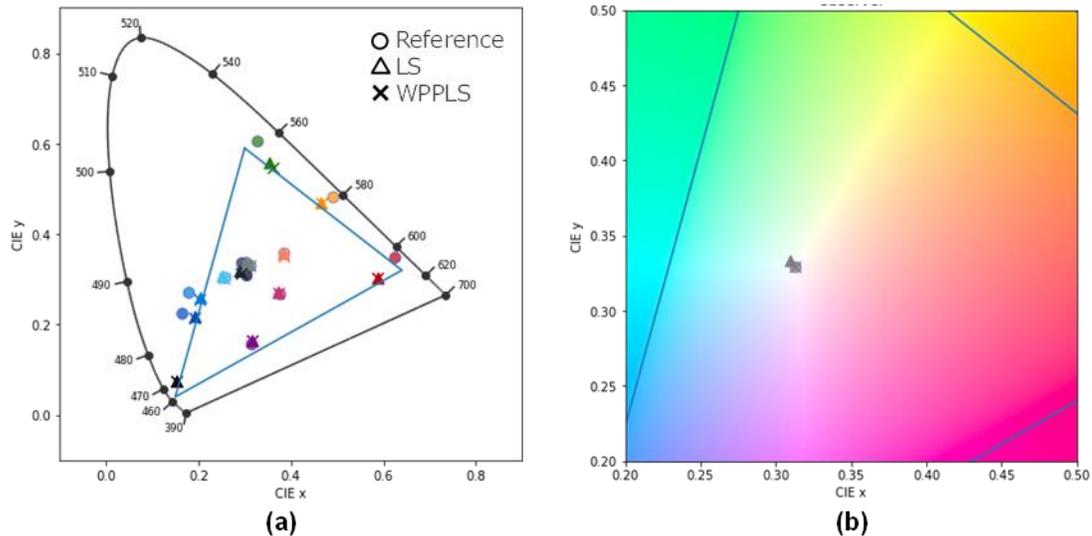


図66 LS法、WPPLS法で補正されたカラーチャートの各パッチの色のxy色度図上での比較。○はリファレンス、△はLS法で補正されたパッチのプロット、×はWPPLS法で補正されたパッチのプロット；(a)全体；(b)背景部分の拡大

カラーチャートを利用した色補正により、WSI スキャナ間で異なる画像の色を合わせる効果があるか検証を行った。2台のスキャナで撮影した画像に対して一連の処理を同様に行い、スキャナ間の画像の色差をピクセル単位で計算した色差マップを作成して検証に用いた。H&E染色、IHC法の組織画像ともに、色補正によってスキャナ間の色差が減少したことを確認した。IHC画像では、色差マップの平均値が2.0以下となった。カラーチャートを用いることでRGB値の補正は行われるが、画像解像度や開口数などの違いは補正することができない。色差マップを作成する際、P250でスキャンした画像はS60の解像度に合わせてダウンスケールしている。しかし、ダウンスケール後であっても、依然としてP250の画像でよりコントラストが優れているような印象を受けた。このような画像解像度の違いにより、図67(c)の緑色の矢印で示す、エッジのような形状の色差が残っていると考えられる。図54、図56(b)に示す色差の分布では、色補正後のヒストグラムは右側に尾を引いていた。これは、前述したようなカラーチャートではキャリブレーションできない要因による色差に関係していると思われる。カラーチャートでは補正できない画像の違いが後続のHER2の自動解析にどれほど影響するかは調査が必要である。

IHCキャリブレーション画像でも色補正によって色差は減少したが、組織画像ほどの大幅な効果は確認されなかった。これはスキャンの際のフォーカスの深さが影響していると考えられる。組織標本スライドは全体が均一な厚みとなるように約4 μ mの厚さでスライスされる。一方で、IHC

キャリブレーションスライドは大きさの異なる2種類のマイクロビーズ(約4 μm 、8 μm)で構成され、フォーカスの深さはスキャナによって自動で決定される。手動でフォーカスの深さを変えてビーズの見え方を比較したところ、小さいマイクロビーズ(約4 μm)にピントが合っている場合、目的とする大きいマイクロビーズ(約8 μm)は正しくフォーカスが合っている場合よりも明るく撮影されることがわかった。このようなフォーカスによる明るさの変化は、カラーチャートスライドでは補正できない可能性がある。組織スライドは場所による厚みの差があったとしてもマイクロビーズほどの差にはならず、画像の明るさの違いにつながるほどのフォーカスの変化はなかったと推測される。S60ではIHCキャリブレーションスライドをオートフォーカスで撮影しても目視で感じるほどの明るさの違いは見られなかったが、P250では画像によって明るさが変わってしまうことがあった。そのため、P250でIHCキャリブレーションスライドをスキャンする際には、大きいマイクロビーズにフォーカスが合うようにフォーカス範囲を手動で限定している。臨床では自動での撮影が好まれるため、臨床応用時に同じような撮影方法をとることは難しいと想定される。今後に向けては、フォーカスを対象のマイクロビーズに合わせやすくするように、小さいビーズを除くなどのIHCキャリブレーションスライドの改良を検討している。また、スキャナのベンダーとも撮影における課題を共有し、改善に向けた検討を図っている。

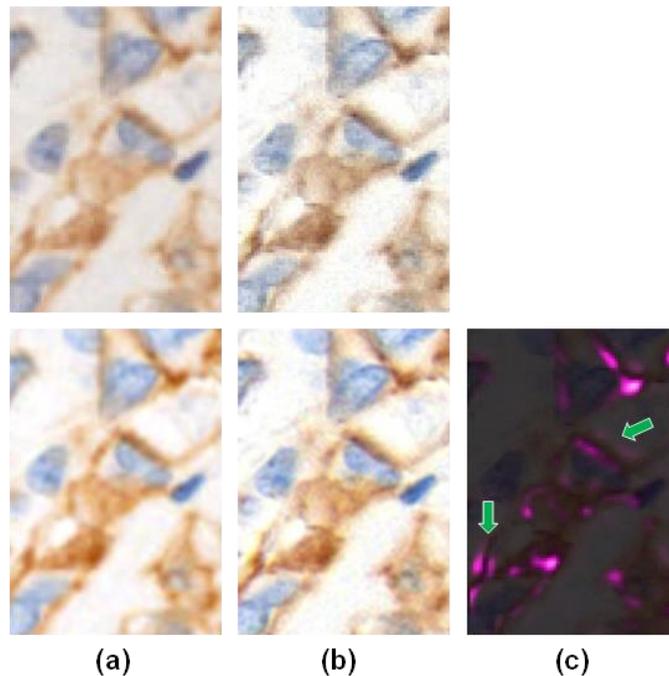


図 67 IHC 画像において色差の大きい箇所の比較

5.2.4 節では、シミュレーションによって IHC 法に対応した茶色をカラーチャートに追加し、従来のカラーチャートを使ったときと補正効果を比較した。H&E 染色ではカラーチャートの色数が増えたことで、色差の平均が小さくなる結果につながったと考える。IHC キャリブレーション画像では、染色されたマイクロビーズの濃さに対応する濃さの茶色を追加することで色差が小さくなるのではないかと考え、2種類の濃さの茶色を準備した。表 12 の結果では、染色されたマイクロビーズに対応する濃さである薄い茶色よりも濃い茶色を追加したときに、より補正効果がみられることがわかった。一方で、組織画像ではいずれの茶色を追加しても効果は見られなかった。これは上記のフォーカスによる明るさの違いが結果の違いに影響している可能性がある。表 13 の結果から、茶色を追加しない場合でもスキャナ間で理想的な濃さの関係性が得られていることから、もともとのカラーチャートのパッチだけでも十分な補正効果が得られていると判断した。ただし、今回はシミュレーションであるため、カラーチャートに実際に茶色のパッチを追加した場合にはまた違った結果が得られることも想定される。

5.4.2 HER2 の自動解析への補正手法の効果

染色、スキャンの補正を行わなかった場合、30 スライド中 14 スライドで目視診断と自動診断の結果の不一致が確認された (表 14)。不一致となった 14 スライドのうち、2 スライドは基準スキャナと同一のスキャナ・設定で取得された評価対象データであったことから、染色によるばらつきが影響して結果の不一致につながったと考える。残りの 12 スライドは、染色のばらつきに加え、スキャンによる色や明るさの違いが結果に影響したと考えられる。評価対象データには、リニア RGB 画像とガンマ補正された画像の両方が含まれており、特にガンマ補正された画像で結果の不一致が多く確認された。ガンマ補正を行うと、画像の暗い部分が明るくなるように補正される。HER2 評価の際の基準データにはリニア RGB 画像を選択しているため、基準データに適した閾値でガンマ補正された画像を評価すると、染色が濃い部分であっても弱い染色と評価されることになり、判定の誤りにつながってしまう。初めにカラーチャートスライドを用いて逆ガンマ補正を適用し、そのうえで色と染色の補正を行うことで、病理医の判定結果と 100% の一致率を得ることができた。また、H-score の標準偏差についても色補正を適用することでほとんどの場合において小さくなった。1+ のケースでは、評価される浸潤領域が他のケースに比べて小さく、Immunoscore の分類が変わると、H-score が大きく変化する可能性があることがわかった。第 4 章でも記載したように、実験で用いたスライドはすべて FDA 認可済みの染色試料と自動染色装置を用いて染色されている。そのため、染色におけるばらつきはスキャナ

によるばらつきよりも小さいものとなり、手法(1-0)と手法(1-1)、(1-2)との差は小さかったと考えられる。

3+のケースについて、IHC キャリブレーションを用いた2つの方法(0-1と0-2)で異なる判定結果が得られた。この違いは、解析ソフトのアルゴリズムに起因するものと思われる。検証のため、3+の組織画像を用いて、入力画像と解析ソフトの内部処理で検出される細胞の数の関係と比較した(図68左)。同一の組織画像について、(a)リニアRGB画像、ガンマ値2.2でガンマ補正された画像を対象として、(b)色補正なし、(c)IHC キャリブレーションスライドで色補正、(d)カラーチャートスライドとIHC キャリブレーションスライドで色補正した、4パターンを示している。各画像を解析ソフトに入力し、同じ設定で細胞検出を行った。図68(e)には検出された細胞の数を比較した。今回使用したQuPathでは図12に示すように、カラーアンミキシングで得られたヘマトキシリンとDABの画像を利用して細胞を検出する。細胞核が検出されていない細胞も一部含まれるが、そのような細胞もHER2の評価には含まれる。(b)の色補正なしのガンマ補正された画像を入力としたとき、(a)のリニアRGB画像を入力したときに比べて検出された細胞の数は半分以下であった。これは、ヘマトキシリンとDABの濃さの画像がガンマ補正された状態であるため、細胞検出するには強度が足りなかったためと考えられる。(c)のIHC キャリブレーションでDABの濃さを補正した画像を入力とすることで、DABの画像から検出される細胞数が増加し、合計の検出細胞数は増加した。しかし、ヘマトキシリンの強度は依然としてガンマ補正された状態であるため、検出される細胞数はリニアRGB画像の半分程度にとどまっている。(d)のカラーチャートとIHC キャリブレーションを使用した画像を入力とすることで、検出される細胞数はリニアRGB画像を入力した場合と近いものとなった。図68右に示すように、2+のケースにおいても同様の傾向が確認された。2+のケースのスコア分類の結果は、方法(0-1)と(0-2)でいずれも正しく得られていた(表2)。しかしながら、細かく画像を確認するとこのような細胞検出精度の違いがあることから、手法(0-1)の結果の信頼性は手法(1-2)に比べて低いものと考えられる。

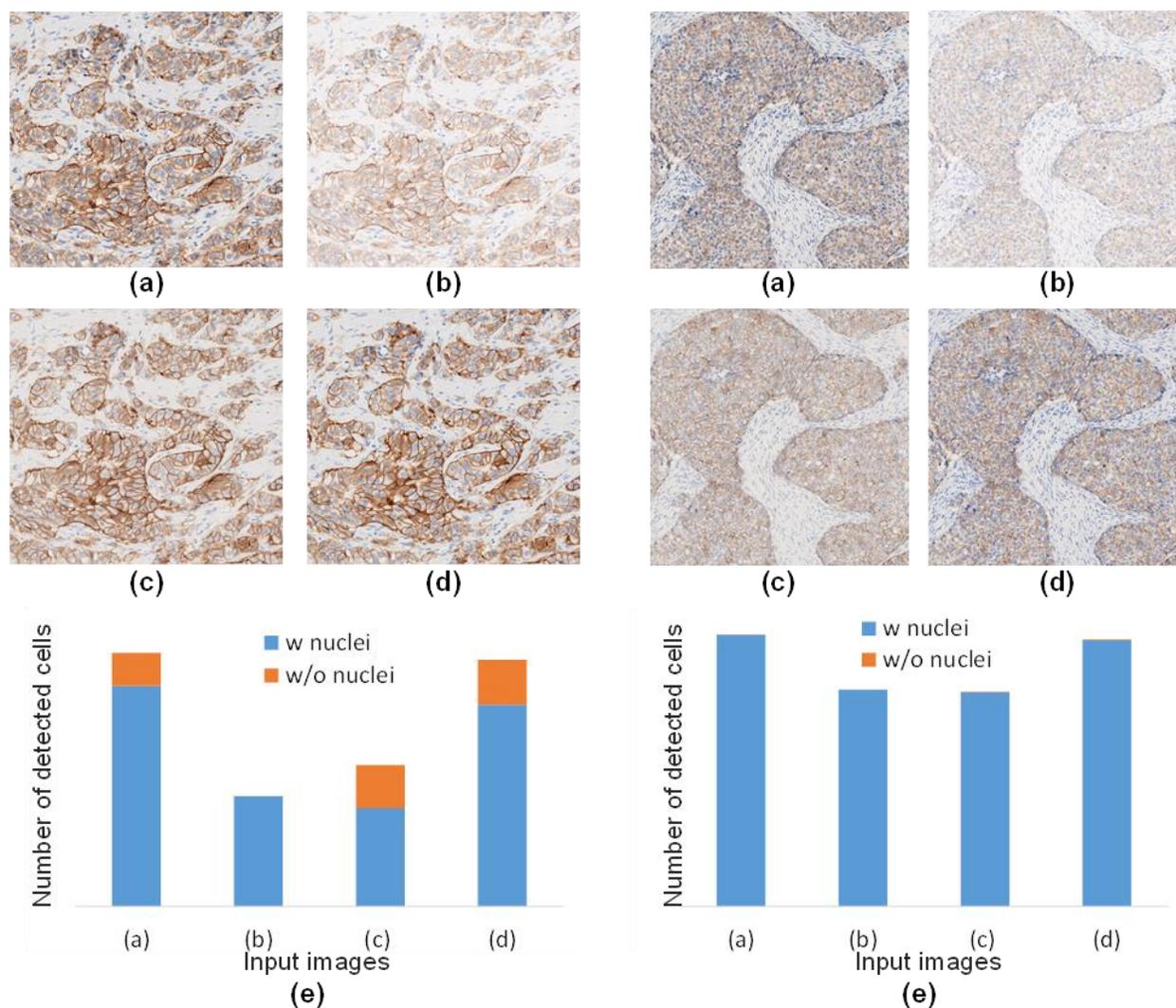


図 68 入力画像の違いによる解析ソフトウェアでの細胞検出数の比較(左：3+、右：2+)；(a) リニア RGB 画像；(b)ガンマ補正画像；(c)ガンマ補正画像の IHC キャリブレーション単体での補正結果；(d)ガンマ補正画像を 2 枚のスライドで補正した結果；(e)解析ソフトウェアで検出された細胞数の比較

手法(0-1)を適用した際に、1+、2+のケースと 3+のケースとの間で一致率に差が見られた。原因として DAB の散乱性が考えられる。濃く染色された DAB は、DAB の散乱特性のためにランバート・ベール則に従わないことが指摘されている[59]。これは、DAB の濃さに応じて色素成分ベクトルが変わることを意味する。この DAB の特性は RGB 画像をカラーアンミキシングする工程に影響する。0~2+のケースなど、DAB がそれほど濃くない場合には線形的な関係性が成り立つため、RGB 画像をカラーアンミキシングによって DAB とヘマトキシリンそれぞれに

分離することができる (図 69(b))。一方で、3+のケースで特に DAB が濃い領域では、DAB 成分の一部がヘマトキシリン成分として分離されてしまうことがある (図 69(d))。3+のケースのリニア RGB 画像でカラーアンミキシングを行った場合、ヘマトキシリン成分として分離される DAB の濃さが真の核領域のヘマトキシリンの濃さよりも小さいため、細胞検出への影響は小さいと考えられる。しかし、ガンマ補正された画像では、たとえ誤って分離される DAB 成分が少量であったとしても、真の細胞核部分のヘマトキシリンの濃さも小さくなっている。そのため、誤った DAB の細胞膜領域の影響が強くなってしまい、ヘマトキシリン画像を使った細胞検出が正しく機能せず、後続の HER2 評価に影響したと考えられる。本手法で使用している閾値は、光学濃度単位で 0.03~0.29 の範囲であった。DAB が特に濃い部分は 1.0 程度であるため、濃い部分の誤差は濃さの分類には影響しないと考えられる。DAB の散乱性については提案手法単体で対応できるものではなく、自動解析ソフトウェアと併せてカラーアンミキシングモデルや HER2 スコアの解析アルゴリズムを改良することが望まれる。

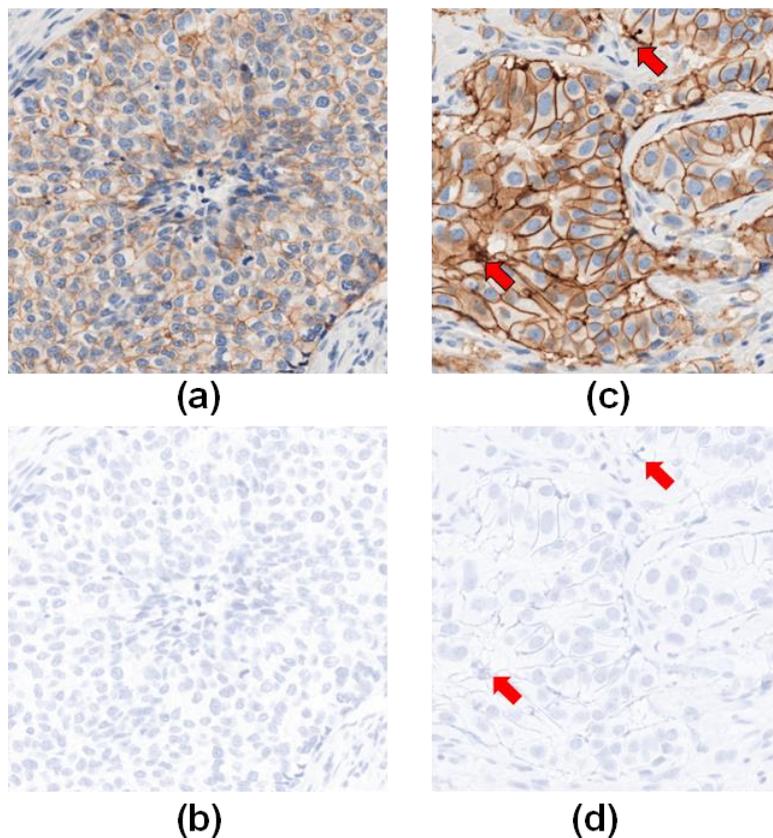


図 69 RGB 画像とカラーアンミキシングで得られたヘマトキシリン画像の比較；(a)(b) 2+のケースの例；(c)(d) 3+のケースの例。赤矢印は DAB が特に濃く、ヘマトキシリン成分として誤って分離されている例を示す

5.5 まとめ

本章では、第3章にて提案した IHC キャリブレーションスライドを用いた補正手法によってスキャナによる画像のばらつきが補正可能か、また、カラーチャートスライドを用いた補正手法との併用による相乗効果があるかどうか検証を行った。異なる日に染色した乳がん組織スライドを2種類の WSI スキャナで撮影し、実験に用いた。

カラーチャートスライドを用いたスキャナの色補正手法により、IHC 法においても色補正の効果が得られ、異なる WSI スキャナ間の色の差を小さくする効果があることを確認した。さらに、染色とスキャナのばらつきがある状態での HER2 の自動評価において、IHC キャリブレーションスライドを用いた補正手法とカラーチャートスライドを併用することで、病理医の判定結果と一致する解析結果を得ることができた。リニア RGB 画像では IHC キャリブレーション単体でも画像の補正効果が確認されたが、ガンマ補正された画像ではカラーチャートでの補正が有効であった。臨床ではスキャンの際にガンマ補正が設定されていることも想定されることから、カラーチャートスライドを併用して補正を行うことが望ましい。

第6章 HER2 IHC 検査の高度化に向けた検討

6.1 概要

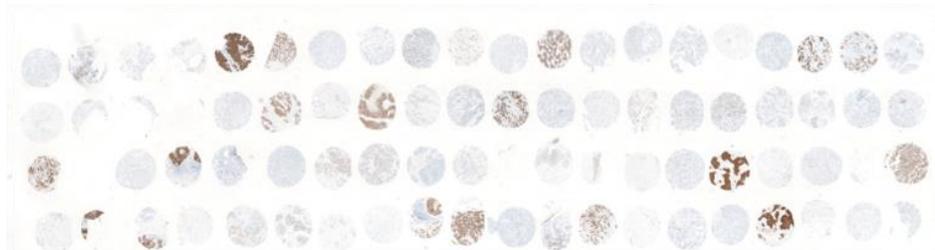
第4章、第5章の実験により、提案手法を適用することで染色、スキャンによる画像のばらつきを補正し、一貫性のある自動解析結果が得られることを確認した。IHC 法の自動解析結果の一貫性が上がることにより、より詳細な HER2 評価が行えるようになる可能性がある。そこで本章では、本来 CISH 法で判断される ISH ステータスを IHC 法の画像解析から分類できるか検討した内容について記載する。また、後半では、臨床での利用に向けた実装に関する検討事項についてまとめる。

6.2 IHC 法と ISH 法の比較実験

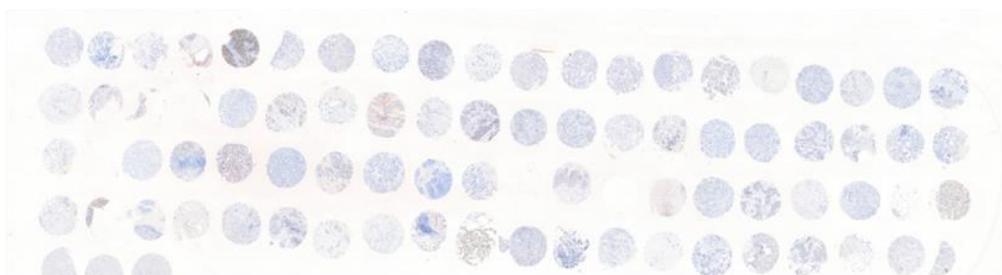
6.2.1 実験データ

実験には浸潤性の乳がん組織から作成した TMA スライドを用いた。TMA ブロックから連続切片を薄切して2枚の TMA スライドを準備し、一枚は IHC 法で、もう一枚は CISH 法で染色した。IHC 法で染色する際には IHC キャリブレーションスライドも併せて染色した。

いずれのスライドもスキャンには5章で使用した WSI スキャナ P250 を使用し、分解能 0.17 $\mu\text{m}/\text{pixel}$ で撮影を行った。解析には、WSI から tiff 形式でエクスポートした画像を使用した。



(a)



(b)

図 70 TMA スライドの画像；(a) IHC 法；(b) CISH 法

6.2.2 HER2 評価で得られるパラメータを用いた ISH ステータス分類

IHC 法で染色したスライドは病理医が WSI を観察することで目視評価を行い、コアごとに HER2 スコアを判定した。評価対象となるがん浸潤部を含まないコアは解析から除外し、計 69 コアを評価対象として選択した。目視評価は病理医 3 名によって実施され、評価結果が分かれた場合は多数決によって HER2 スコアを決定した。自動解析には QuPath を用いた。基準データは第 4 章、第 5 章と同じデータセットを使用し、TMA スライドは評価対象データとした。

CISH 法の評価には、ラボで開発した HER2 CISH 用の自動解析ソフトウェア[60][61]を用いた(図 71)。解析ソフトウェアでは入力画像に対して自動で核を検出し、核ごとの HER2、CEP17 それぞれのシグナル数を計数する。シグナル数はユーザーによって変更も可能である。最終的に入力した画像全体で検出された核の HER2/CEP17 比率を計算したうえで、ASCO/CAP のガイドラインに基づき決定した ISH ステータスを出力する。

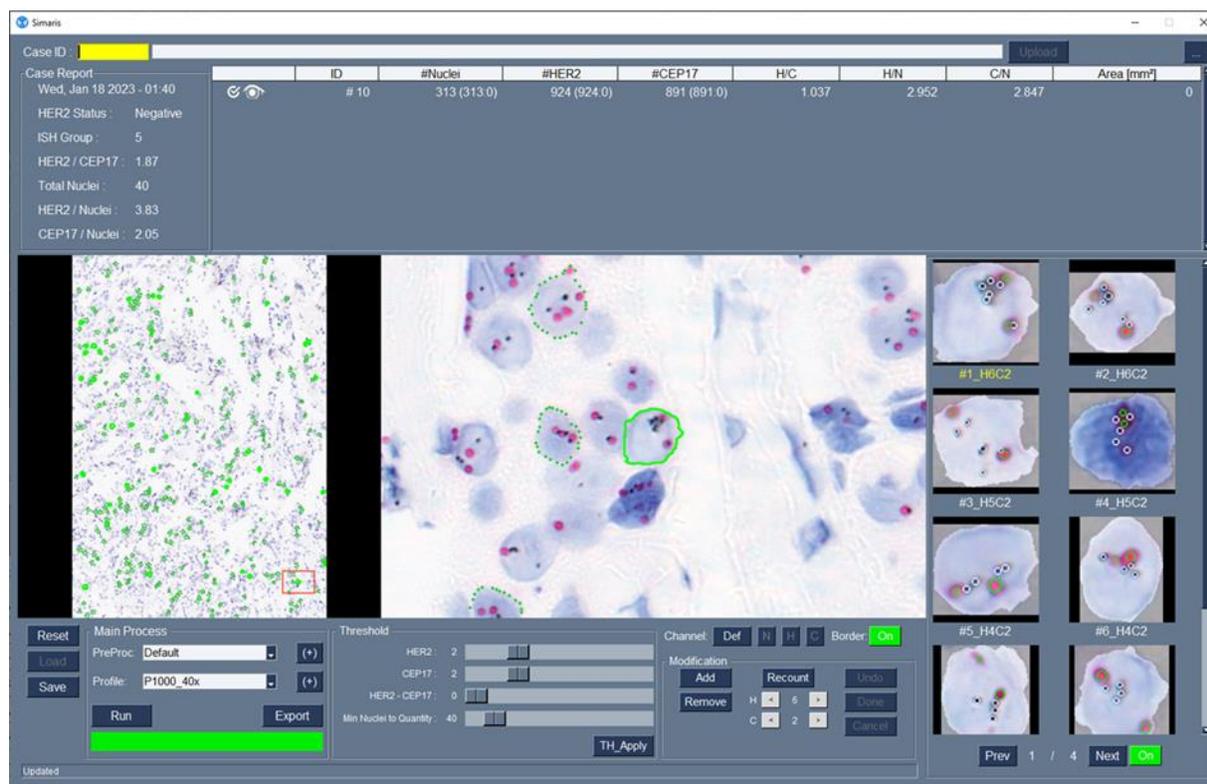


図 71 CISH 画像の自動解析ソフトウェアの画面の例

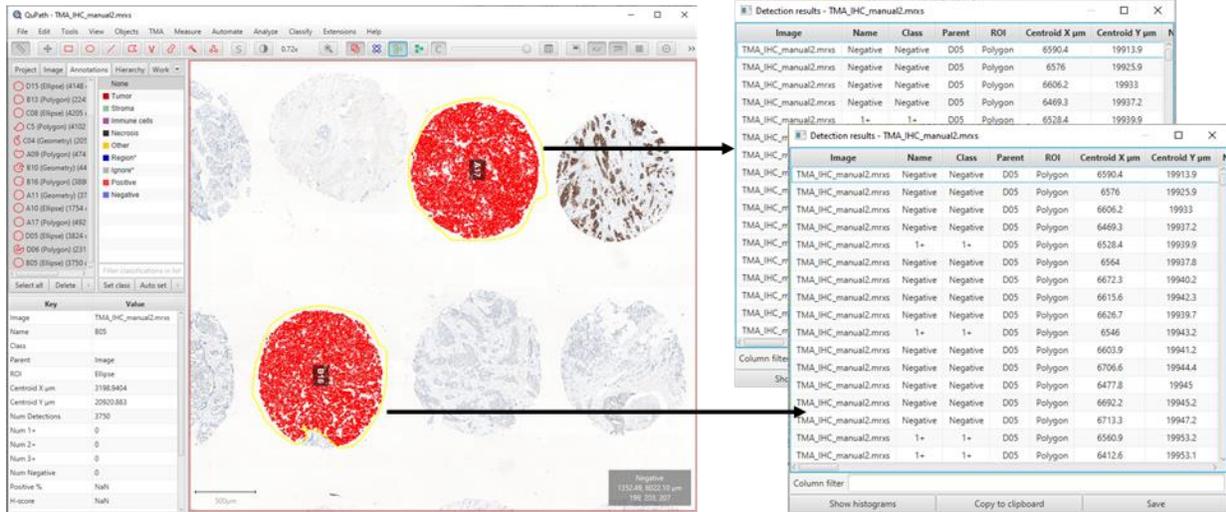
始めに、病理医の目視評価、自動解析のそれぞれで得られた HER2 スコアと、CISH 法の自動解析で得られた ISH ステータスの関係を比較した。次に、IHC 法の自動解析で得られた H-score や各 Immunoscore に分類された細胞の割合から ISH ステータスを分類できるか検討を行った。IHC キャリブレーションで補正を行わなかった場合と、閾値を補正した場合の結果を比較した。

6.2.3 組織の DAB の濃さの分布からの ISH ステータス分類

細胞膜の DAB の濃さの情報を利用し、コアごとの DAB の濃さの分布から ISH ステータスの分類が可能か実験を行った。図 72 に示すように、QuPath の内部処理でカラーアンミキシング、細胞検出を行って得られた細胞膜の DAB の濃さを出力し、コアごとに細胞の数を縦軸、DAB の濃さを横軸とするヒストグラムを作成して解析に利用した。

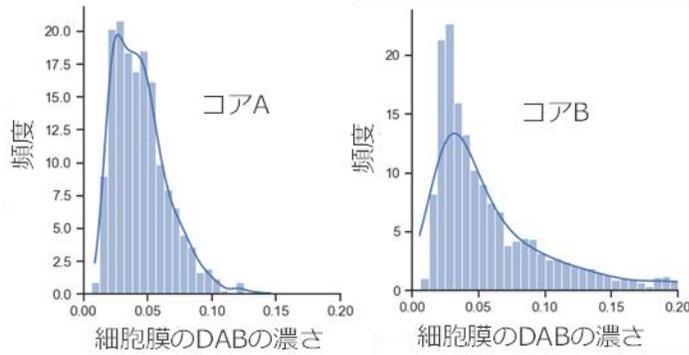
事前の確認として、スコア別の分布の違いを比較した結果を図 73 に示す。図 73(a)は 0、1+、2+/ISH ネガティブ、2+/ISH ポジティブ、3+の 5つのグループについて、グループごとに 5つのコアを合わせた分布を示している。図 73(b)は 2+/ISH ネガティブ、2+/ISH ポジティブの 2つのグループの分布を示している。2+の中ではわずかに ISH ポジティブのケースのほうが濃い割合が多い分布となっており、HER2 スコアが異なるグループ間では明らかな分布の違いが見られていた。1+のグループはいずれも ISH ネガティブ、3+のグループはいずれも ISH ポジティブであるため、2+の 10 ケースを分類の対象とした。

はじめに、コアごとに 0 からコア内の濃さの最大値を範囲として、ビン数 64 でヒストグラムを作成した。得られたヒストグラムは右側にすそ野が広い分布となっていたため、式(39)に示すジョーンソン SU 分布で近似した。得られた 10 ケースの分布について濃さのレンジ(0,0.4)、ビン数 64 で確率密度分布を作成したのち、フィッシャーの線形判別[62]を適用して ISH ステータスによって違いが見られる濃さのレンジを確認した。判別は 2 クラス、64 次元のデータとして行った。



1. QuPathでカラーアンミキシング、細胞検出を行い、コアごとに細胞膜のDABの濃さを取得

2. 取得したコアごとの細胞膜のDABの濃さをcsvファイルとして出力



3. コアごとにヒストグラムを作成し解析

図 72 細胞膜の DAB の濃さの分布の解析イメージ

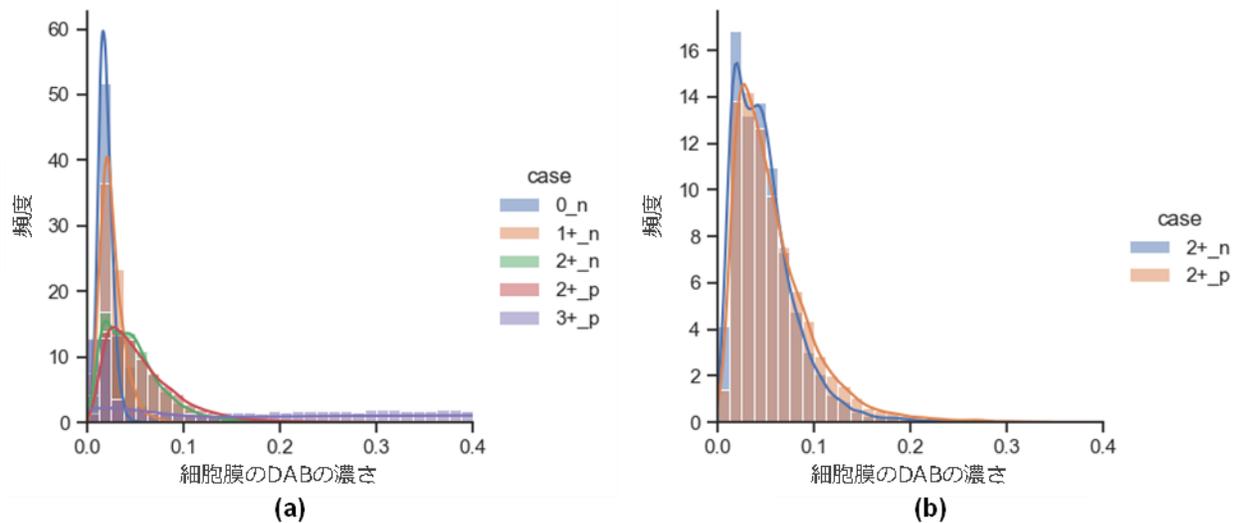


図 73 スコアごとの分布の比較

6.3 実験結果

6.3.1 HER2 評価で得られるパラメータを用いた ISH ステータス分類結果

図 74 に病理医が目視評価した HER2 スコア、自動解析ソフトウェアで得られた HER2 スコアと ISH ステータスの 3 つの関係を表したサンキーダイアグラムを示す。自動解析の結果は IHC キャリブレーションを利用して閾値を補正する提案手法 1 を適用して得られた結果となる。病理医の目視評価と自動解析で得られた HER2 スコアの一致率は 87.5%であった。自動解析で得られた HER2 スコアと ISH ステータスを比較すると、IHC 1+かつ ISH ポジティブと評価されたケースが 2 件みられた (図 74 赤矢印)。CISH 検査は通常、IHC 法で 2+と評価されたケースを対象として実施される。そのため、ISH ポジティブにも関わらず IHC 1+と評価されたケースは見逃しにつながる可能性がある。

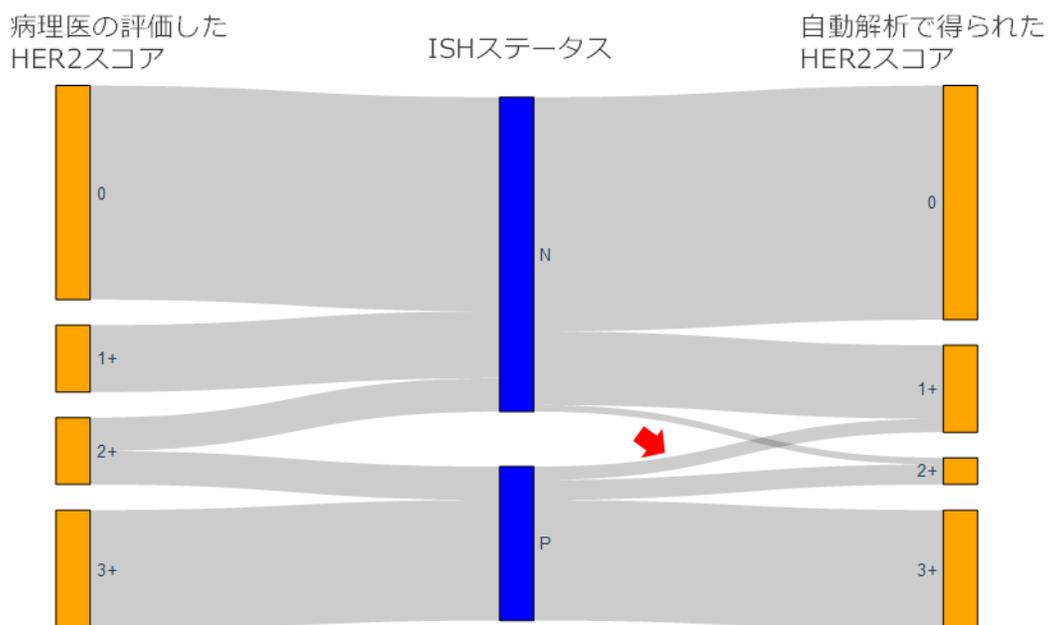


図 74 病理医の評価、自動解析のそれぞれで得られた HER2 スコアと ISH ステータスの関係

図 75 は CISH 画像の自動解析で得られた HER2/CEP17 の比率を縦軸、IHC 法の自動解析で得られたパラメータを横軸とし、ISH ステータスでグループ分けした評価対象 71 データをプロットした結果を示す。図 75 (a)のように H-score を横軸としたとき、50~100 の間で ISH グループが混ざっており、分類が難しいことが分かった。1+、2+、3+に分類された細胞の割合を横軸とした場合(図 75 (b)-(d))にも、分布の重なりが見られる。2+と 3+を合わせた割合を横軸としたと

き(図75(e))、比較的分布が分かれている。ただし、赤矢印で示す2+かつISHポジティブの2ケースがネガティブの分布に寄っていた。

図76は2+と3+を合わせた割合を横軸として、IHCキャリブレーションによる補正なしでHER2評価を行った場合(表4の補正なしに該当)と、IHCキャリブレーションを用いて閾値を補正して評価した場合(表4の提案手法1に該当)の比較を示す。補正を行うことで補正前よりも分布が分かれたようにも見えるが、依然としてISHポジティブの2ケースがネガティブ側に混ざったような分布となっていた。

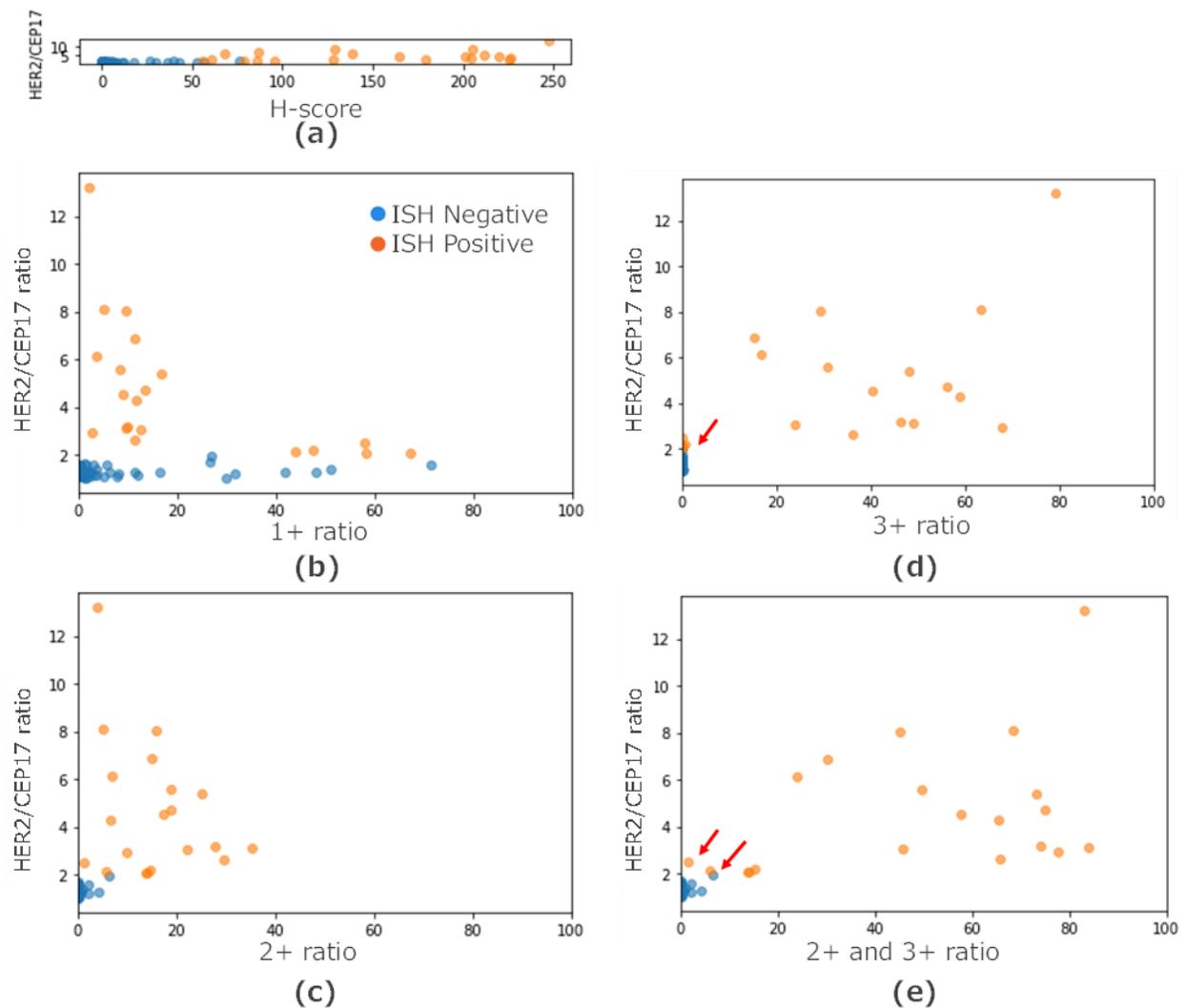


図75 CISH法で得られるHER2/CEP17 ratio(縦軸)とIHC法で得られるパラメータ(横軸)の比較；(a) H-score；(b) 1+に分類された細胞の割合；(c) 2+に分類された細胞の割合；(d) 3+に分類された細胞の割合；(e) 2+、3+に分類された細胞の割合

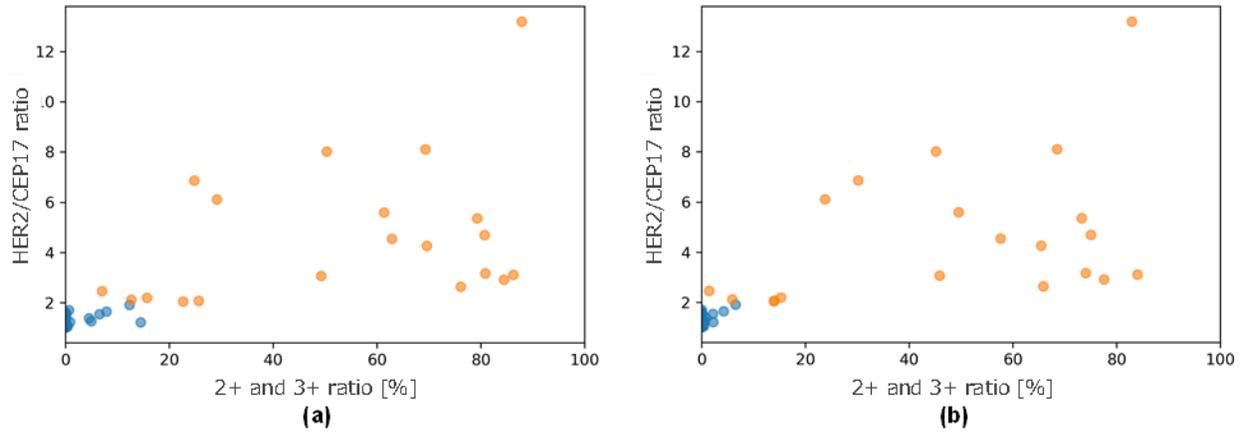


図 76 IHC キャリブレーションにより補正の有無による、HER2 自動解析結果からの ISH ステータスの分類の比較；(a) 補正なし；(b) 閾値の補正あり

6.3.2 組織の DAB の濃さの分布からの ISH ステータス分類結果

図 77(a)は HER2 スコア 2+ の 10 コアの DAB の濃さの分布を示す。図 77(b)は濃さのレンジ 0.08 ~0.18 を抜粋して表示している。コアごとに見た場合も ISH ポジティブのケースがやや濃い割合の多い分布となっていた。図 77(c)はフィッシャーの判別分析で得られた射影関数の重みのプロットを示す。濃さが約 0.02、約 0.10 のときに突出した値をとっていた。0.02 付近ではネガティブの確率が高いが、完全には分離されていない。0.10 付近ではポジティブの確率が高く、ネガティブの確率が低くなっていた。

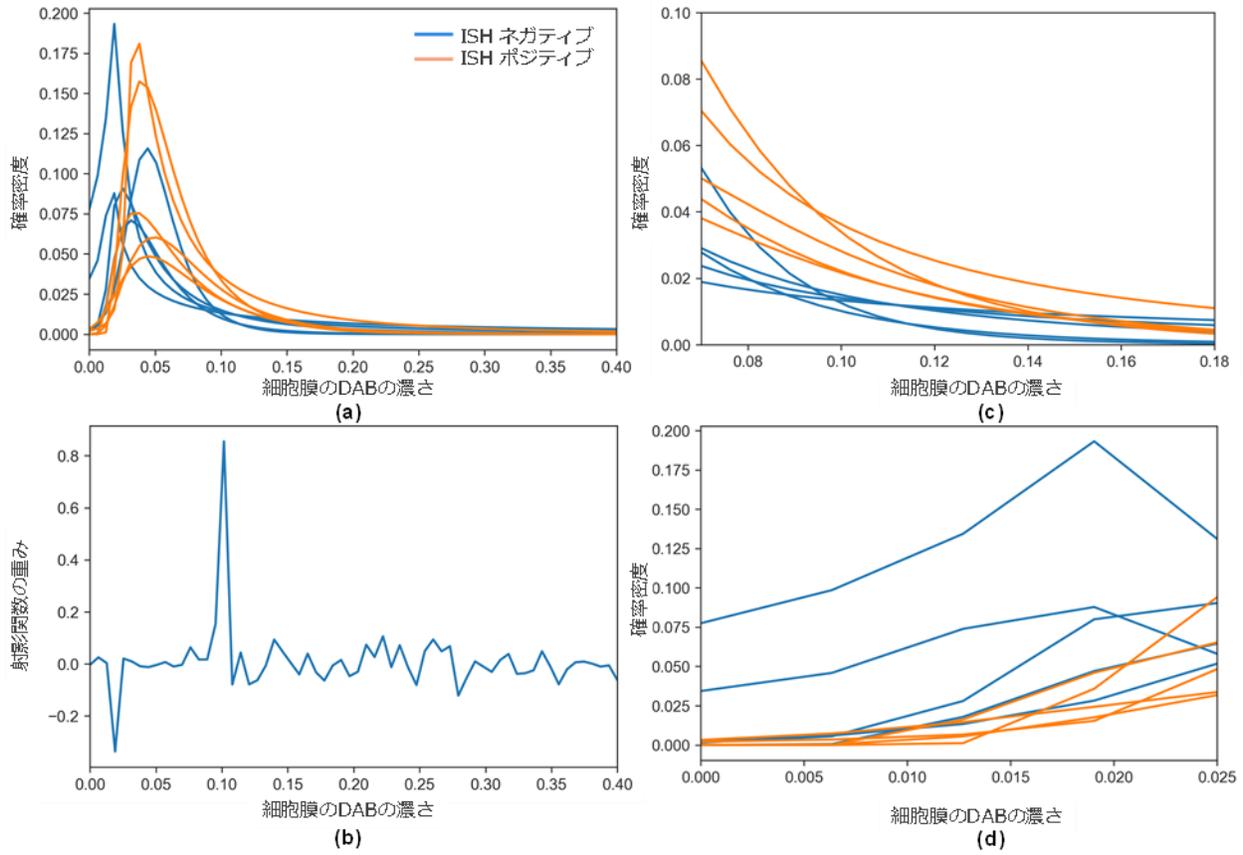


図 77 (a)コアごとの DAB の濃さの確率密度分布(青：ISH ネガティブ、オレンジ：ISH ポジティブ)；(b) フィッシャーの判別分析の結果；(c)(d) 判別分析で特徴が見られたレンジの確率密度分布の抜粋

6.4 考察

6.4.1 HER2 評価で得られるパラメータを用いた ISH ステータス分類

本実験では、特に 1+ と 2+ の組織において病理医の評価と自動解析の結果で HER2 スコアの不一致が見られた。原因として、基準組織の選択に原因があると考えられる。各 HER2 スコアに該当する組織は、そのスコアの中である程度薄いものから濃いものまで幅を持って存在する。本研究での基準組織は各スコアの中で標準的な濃さのものを選択しているため、薄めの 2+ は 1+ として分類されてしまったと考えられる。図 78 に 1+、2+ の基準組織と 2+ の評価対象組織の例を示す。評価対象組織の 2+ は基準組織の 1+ と 2+ の間のような濃さであった。

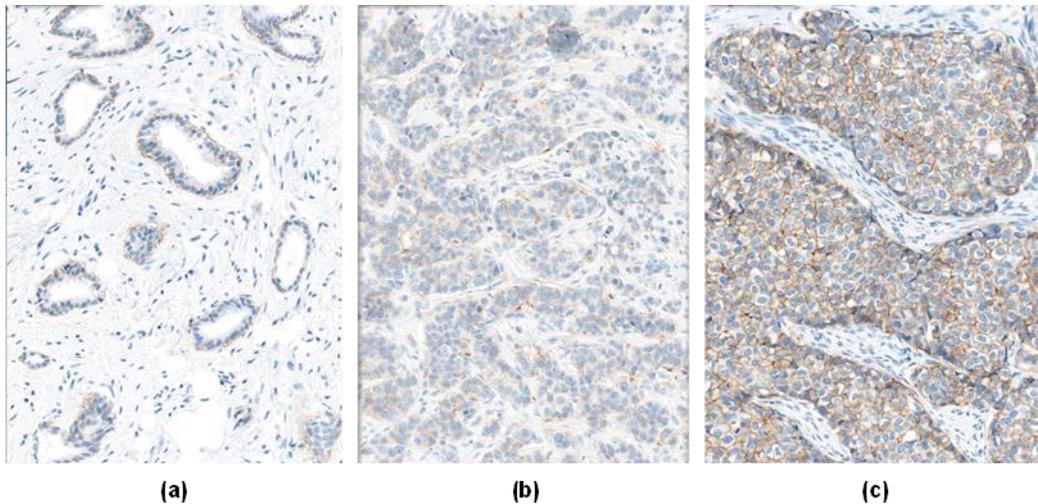


図 78 基準データと対象データの比較；(a) 基準データの 1+；(b) 評価対象データの 2+；(c) 基準データの 2+

この結果を踏まえ、閾値を変えた検証を行った。基準組織はこれまでと同様のデータを用いて、閾値算出に選択する ROI を組織の中でも薄めの箇所に変更して新しい閾値を算出した。新しく算出した閾値を用いて自動解析を行って得られた HER2 スコアと病理医が評価した HER2 スコア、ISH ステータスとの比較を図 79 に示す。病理医の評価と自動解析の一致率は 96.8% となり、図 74 の結果からの改善が見られた。ISH ポジティブのケースは IHC 2+ または 3+ に分類されるようになった。

0、1+ のケースにおいて、病理医の目視評価と自動解析結果とで 3 ケース不一致が確認された。これらのケースは病理医 3 名の中でも 0 と 1+ のどちらで評価するか意見が割れており、判断が難しいケースであった。CISH 法の解析では、病理医が自動解析の結果を参考にして判定を行った場合、参考がなかった場合と比べて正解値との一致率が大幅に改善されたとの報告がある [63]。IHC 法においても、意見が分かれるようなケースでは自動解析の結果を第 4 の意見として合意を図り、最終評価の補助として活用することも今後考えられる。

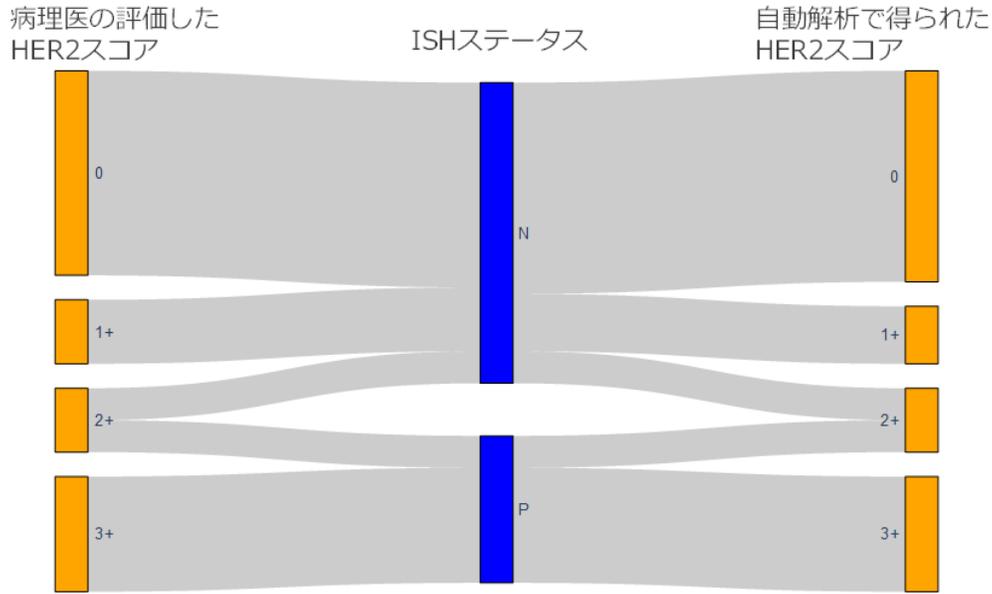


図 79 病理医の評価、新たな閾値を用いた自動解析のそれぞれで得られた HER2 スコアと ISH ステータスの関係

新しく算出した閾値を用いて自動解析を行い、図 76 と同様に 2+と 3+を合わせた割合を横軸としてプロットを行った結果を図 80 に示す。補正なしで HER2 評価を行った場合と、IHC キャリブレーションを用いて閾値を補正して評価した場合の結果を比較している。補正を行わない場合、横軸の割合 40%付近で分布の重なりがみられるが、補正を行うことで 14~15%付近で ISH ステータスの分離が可能に見える結果が得られた。以上の結果より、閾値算出方法については CISH 法の結果と比較することによって精度を上げることができるとの知見が得られた。

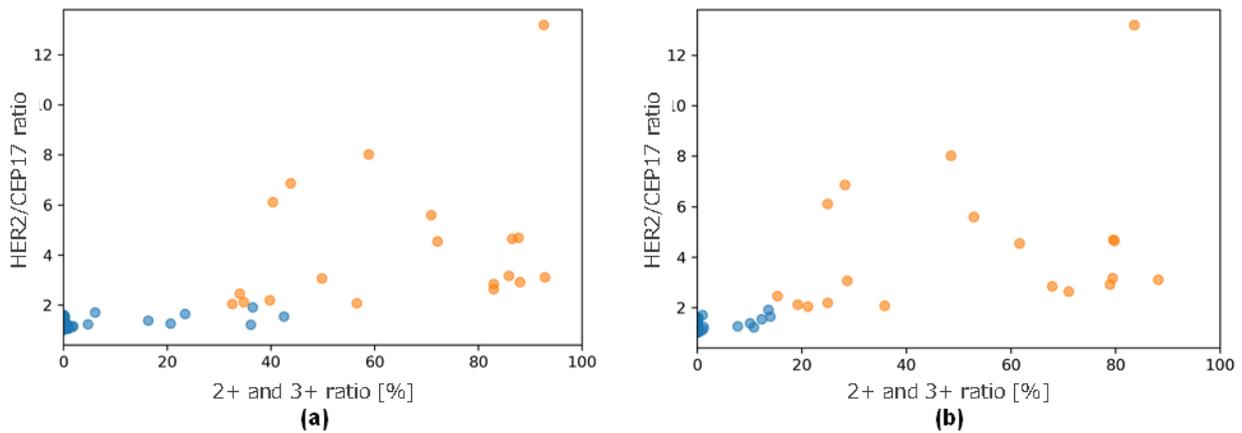


図 80 閾値を変えて評価した際の、IHC キャリブレーションにより補正の有無による HER2 自動解析結果からの ISH ステータスの分類の比較；(a) 補正なし；(b) 閾値の補正あり

6.4.2 組織の DAB の濃さの分布からの ISH ステータス分類

細胞膜の DAB の濃さの分布から ISH ステータスを判別できるか検証した結果、0.1 付近で特徴が見られた。実際に分布を見ても ISH ステータス別に確率密度に差が見られ、分類に利用できる可能性が確認された。ネガティブの分布の中で、ポジティブ寄りの分布が 1 ケースみられた(図 81(a)赤矢印)。図 81(b)にネガティブのコアの例 (図 81(a)緑矢印)と、(c)にポジティブ寄りのコア (図 81(a)赤矢印)の、画像と QuPath での分類結果を示す。(b)のコアと比較すると(c)のコアは DAB の染色はそこまで濃くはなく、大部分が弱い染色として黄色に分類されている。しかし、染色されていない青色の割合が少ないため、分布のピークがポジティブ側に寄った分布になっていると考えられる。中程度/強い染色として分類されている割合は多くないため、ポジティブのケースに比べるとすそ野は狭い分布となっている。

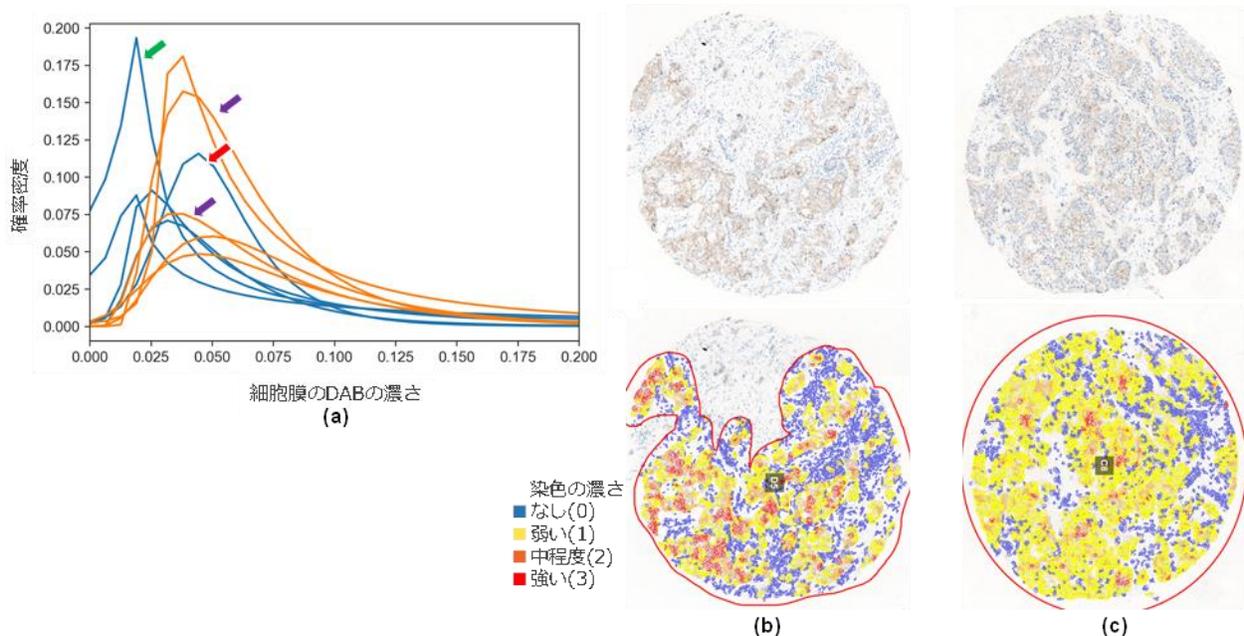


図 81 ISH ネガティブで分布に特徴の見られたコアの比較；(a) コアごとの濃さの確率密度分布 (再掲)(b) ネガティブのコアの例 ((a)緑矢印)；(c) ポジティブ寄りのコア ((a)赤矢印)

図 76 の結果で分類が難しかった 2+ の 2 つのケースは、図 81(a)紫矢印で示す分布となっており、他の 3 つのケースと比べて突出して特徴のある分布を取っているわけではなかった。図 82 に高倍率で表示した画像の比較を示す。問題の 2 ケースはやはり他のケースと比べても染色が薄めであるため、ネガティブ寄りの分類となっていたと推察される。細胞膜の形状を比べたとき、上段のネガティブのケースでは細胞膜の形状が比較的是っきりと確認できるのに対し、問

題の 2 ケースでは境界があいまいになり細胞核が重なっているような形態が見られている。下段の染色の濃いポジティブのケースではその特徴がより顕著にみられ、細胞の異形性として ISH ポジティブの特徴が表れていると考えられる。形態的な特徴に関しては H&E 染色のスライドを併せて観察することで、より詳細な解析が可能と考える。

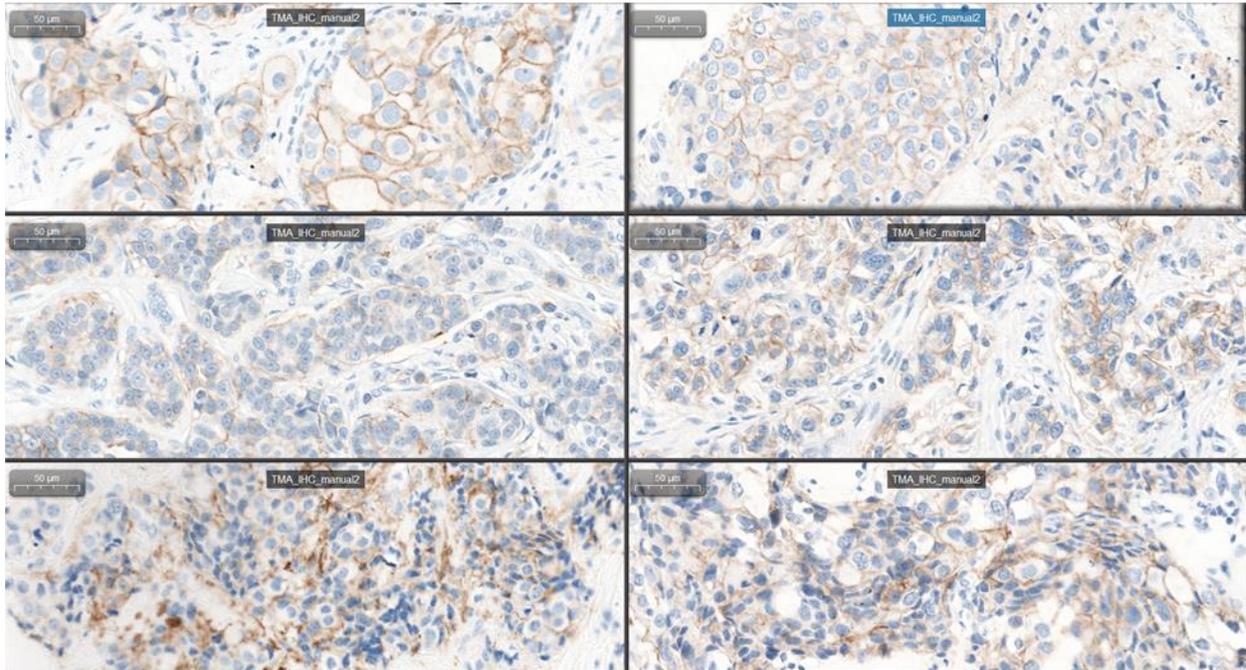


図 82 2+のケースの高倍率での画像の比較。上段：2+/ネガティブの例、中段：2+/ポジティブで分類が難しかったケース、下段：2+/ポジティブの例

6.5 社会実装に向けた検討

手法の有効性検証の次のステップとして、臨床で病理医が利用できるよう、実運用に向けた社会実装を行う必要がある。まず、IHC キャリブレーションスライドの解析のアップデートを行った。これまでの検討にて、WSI からエクスポートした画像を対象として一連の解析手順を構築することができた。実運用を想定したとき、画像のエクスポートは作業者の負担になると想定される。そこで、画像のエクスポートにかかる工数を減らすため、WSI を対象としたマイクロビーズの濃さの推定処理手順を構築した (図 83)。WSI の解析に当たっては、WSI 読み込み用のライブラリとして提供されている OpenSlide [64] を活用した。OpenSlide は主要なスキャナベンダーの WSI フォーマットに対応しており、本研究で利用した浜松ホトニクス、3DHistech いずれの WSI も読み込むことができる。図 11 に示したように、WSI は複数の解像度の画像から成るピラミッド構造を持つ。画像が高解像度であるほど細部を確認できるが、読み込み及び解析には時間を要す。IHC キャリブレーションスライドの画像を確認したとき、マイクロビーズが貼付され

ているレベルごとのドット部分は最も低解像度の画像でも認識することができた。そこで、始めに低解像度で画像全体を読み込み、必要な 10 個のレベルのドット部分の検出を行う。次に、検出した座標情報をもとにドットの中央部分に限定して高解像度画像を読み込み、3.2.3 節に記載したマイクロビーズの解析を行う。ここで、処理時間を短縮するために、ハフ変換で行っていたマイクロビーズの検出を輪郭検出で行うように変更を加えた。スキャナ S60 で撮影した画像を対象にエクスポートした bmp 画像と WSI の解析結果を比較したとき、おおむね近い DAB の濃さが推定されていることを確認した(図 84)。推定された DAB の濃さのずれは、選択された画像領域やビーズが異なることが原因と考えられる。

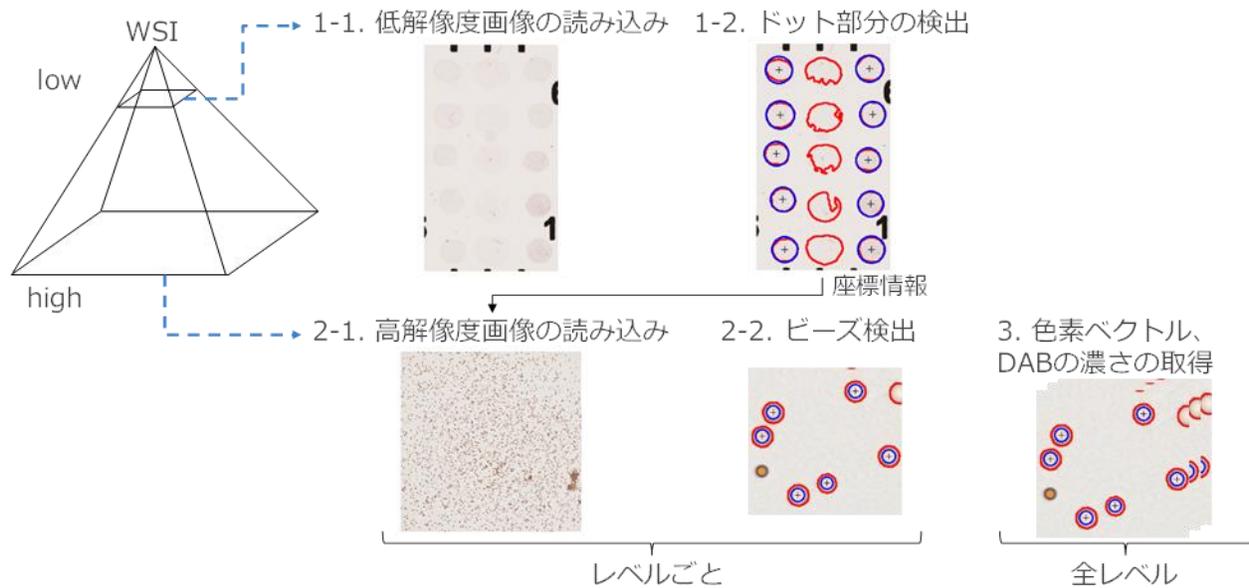


図 83 IHC キャリブレーションの WSI を対象とした解析手順

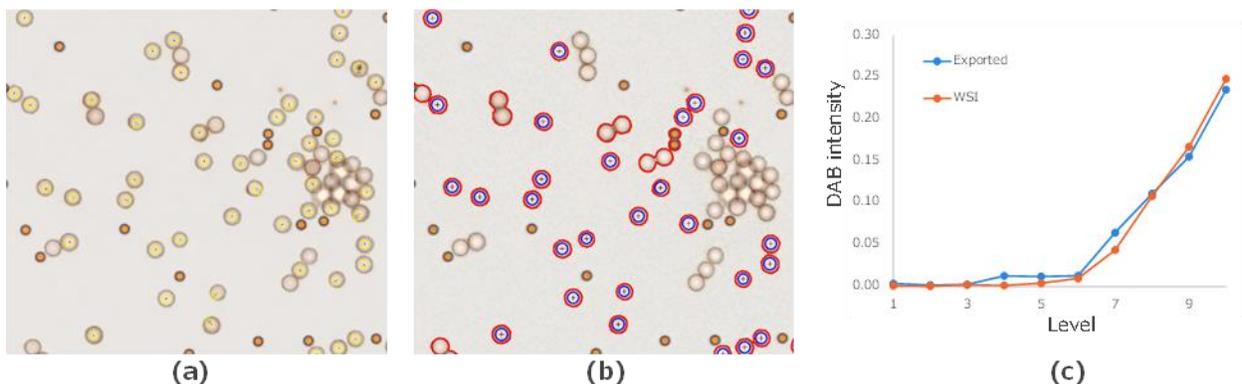


図 84 bmp 画像と WSI の解析結果の比較；(a) bmp 画像のビーズ検出結果(黄色が検出されたビーズ)；(b) WSI のビーズ検出結果(青色が処理用に選択されたビーズ)；(c) 推定した DAB の濃さのプロット

HER2 の評価ステップでは、臨床診断にも利用されている WSI ビューアのひとつである SlideViewer (3DHistech Ltd.)上で使うことができる解析ツール QuantCenter [65]の併用を検討している。QuantCenter では HER2 IHC 画像の解析や HER2 ISH 画像の解析などの 10 種類の解析機能が提供されており(図 85(a))、ビューアで表示している画像に対して自動解析を行った結果が得られる。表 17 に利用予定の機能の例と、提案手法の適用方針を示す。ツールでは MembraneQuant として HER2 IHC 画像の自動解析機能が実装されているが、内部処理がブラックボックスであることから、手法 2 で画像補正した画像を入力とすることで解析結果の標準化が図れる。また、ScriptQuant としてビューアで表示している画像に対して自作の解析機能を適用し結果を出力する機能が提供されている。この機能を活用し、HER2 の自動解析機能を構築することによって、手法 1 の閾値補正によって解析結果の標準化が図れる。

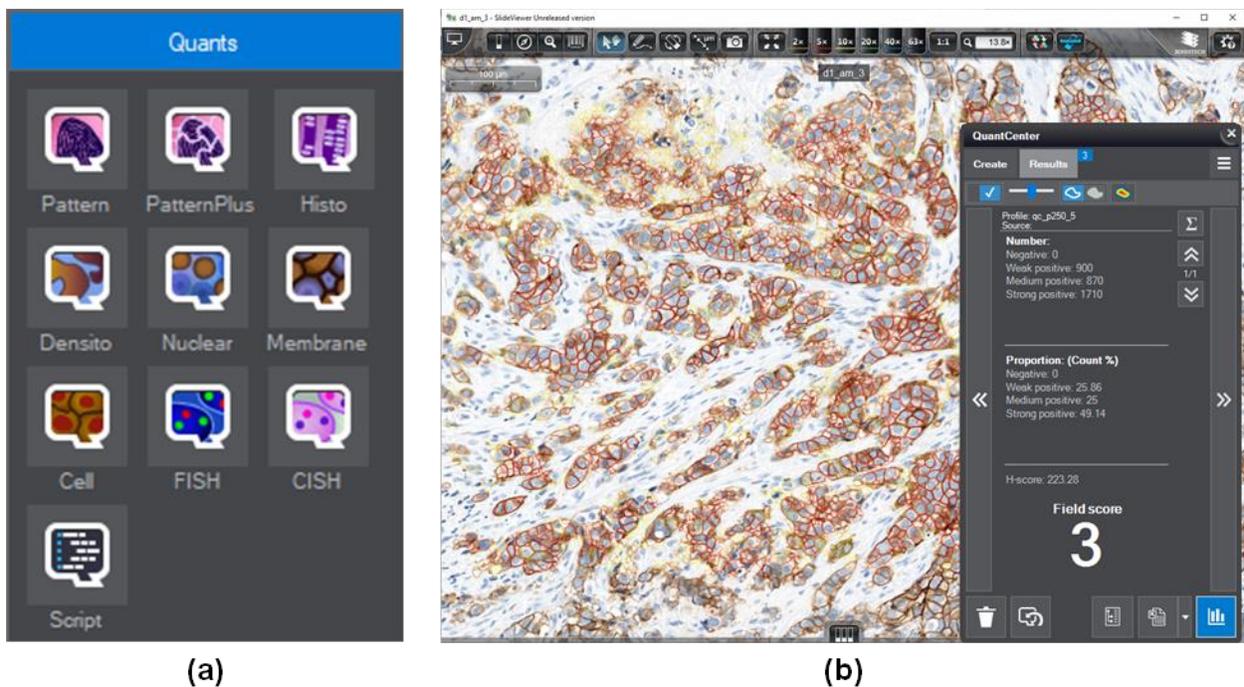


図 85 QuantCenter ; (a)提供されている解析機能一覧 ; (b)HER2 IHC 画像の自動解析機能の結果画面の例

表 17 QuantCenter の提供機能と適用方針

#	機能名	機能概要	提案手法の適用方針
1	MembraneQuant	HER2 の自動解析機能 (内部処理はブラックボックス)	手法 2 : 画像補正を適用した画像を解析
2	ScriptQuant	自作した解析機能を表示画像に適用する機能	HER2 の自動解析機能を構築したうえで手法 1 : 閾値補正を適用

2つ目の方針に関して、自作した HER2 の自動解析アルゴリズムを ScriptQuant で動かした際のスクリーンショットを図 86 に示す。図 12 に示した自動解析アルゴリズムのカラーアンミキシング、細胞検出、細胞膜の DAB の濃さの分類までを内部で処理し、分類された Immunoscore ごとに色を分けた細胞膜の表示まで、一つの処理で動作することを確認した。細胞検出は精度向上に向けたアップデートが望まれるが、まずはベンダーとの協議の材料となるたたき台を構築することを目的とした。臨床利用に向けては、IHC キャリブレーションスライドの解析結果との紐づけや、HER2 の解析結果の表示などに関して協議が必要となる。今後は実装に当たっての制約や要求事項をベンダーに共有し、社会実装に向けた検討を進める。

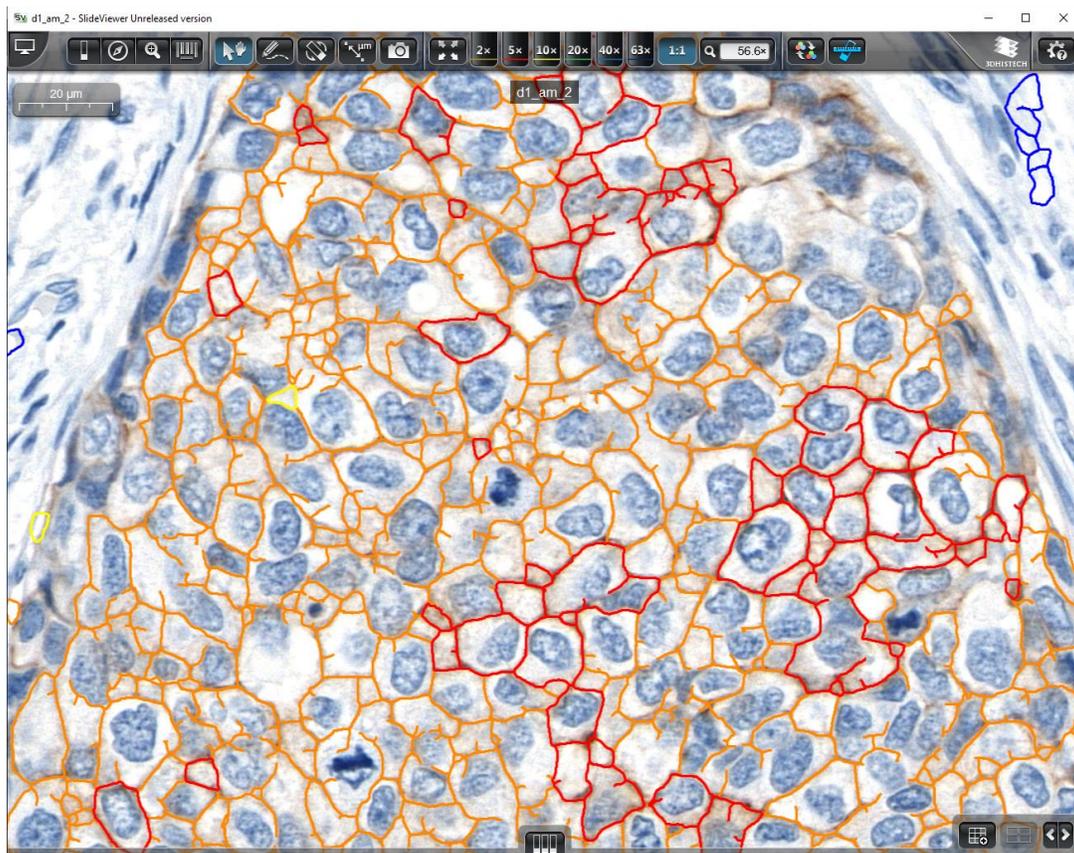


図 86 自作した HER2 の自動解析結果の SlideViewer での表示例

6.6 まとめ

本章では、提案手法によって IHC 法の自動解析結果の一貫性が向上することにより、IHC 法からより詳細な診断が行えないか、また、CISH 法の解析結果と比較すること IHC 法の診断精度をさらに向上させることができないかとの考えの下、検討を行った。CISH 法で診断される ISH ステータスを対象とし、IHC 法と CISH 法で染色したスライドの解析結果を比較した。

6.4.1 節の考察より、3.2.4 節で提案した閾値の設定方法に課題があり、CISH 法の結果と比較して閾値を調整することによってより正確な結果が得られることがわかった。また、2+と 3+の濃さとして分類された細胞の割合を見ることで、IHC 法から ISH ステータスを分類できる可能性が示唆された。また、DAB の濃さの分布から ISH ステータスによる違いがみられるか解析したところ、濃さ 0.1 付近の分布から ISH ステータスを分類できる可能性が示された。ただし、HER22+は 10 ケースと他のスコアに比べて少なく、ISH ネガティブ、ISH ポジティブそれぞれ 5 ケースと限られた症例数での検証結果となっている。また、IHC 法と CISH 法の両方を含むデータセットは 1 セットのみであるため、この結果については IHC キャリブレーションによる補正の影響を検証することはできない。より信頼性の高い検証に向けては、症例数を増やすことや H&E 染色のスライドも準備することが望まれる。

また、社会実装に向けて、市販のビューアの解析機能を利用しての提案手法の適用方針を検討した。画像のばらつきの補正まで行える解析ソフトウェアは現在存在しないため、本手法を組み合わせることはベンダーにとっても大きなアドバンテージとなる。今後は実装に当たっての制約や要求事項を共有し、社会実装を進める予定である。

第7章 結論

7.1 本論文のまとめ

本論文では、HER2 IHC 検査における自動画像解析の普及に向けて、WSI のばらつきの補正手法を提案した。WSI のばらつきに特に影響するとされる染色、スキャンのそれぞれの原因に対応するため、2種類のキャリブレーションを導入した。

染色による DAB の濃さや色味のばらつきの補正には、IHC キャリブレーションスライドを導入した。このスライドが意図している本来の用途は施設ごとの染色の品質管理であり、スライドの観察は顕微鏡下を想定している。そこで初めに、WSI スキャナで取得した IHC キャリブレーション画像から DAB の色味や濃さの取得が可能であるか、基本的な画像処理プロセスを構築して確認を行った。IHC キャリブレーションではビーズの散乱を考慮することで、目的とする DAB の色味と濃さを安定して得られることを確認した。画像のばらつきの補正への応用では、物理モデルに基づく補正手法を提案した。異なる日に染色した乳がん組織標本スライドを用いた検証実験により、提案手法を適用することで自動解析結果と病理医の診断との一致率が向上することを確認した。

スキャンによる画像の色のばらつきの補正には、カラーチャートスライドを導入した。先行研究では H&E 染色画像を対象とし、スキャナの補正のみを目的にカラーチャートスライドが使われていた。本研究では対象を IHC 画像とし、かつ、染色とスキャンの両方のばらつきがある画像に対しての補正効果を検証した。さらに、IHC キャリブレーションを用いた補正手法はスキャンの補正効果も含むことから、併用することで相乗効果が得られるか比較した。異なる 2 種類のスキャナを用いた実験により、IHC 画像においてもカラーチャートスライドによって異なるスキャナで撮影した画像間の色差を小さくする効果を確認した。HER2 の評価では、提案手法を適用して染色、スキャンの両方の色補正を行うことで、自動解析結果と病理医の診断との一致率が向上することを確認した。IHC キャリブレーションスライド単体でもスキャナの補正効果は確認されたが、スキャナごとの設定の違いを補正するためには、カラーチャートスライドの併用が望ましいことがわかった。

提案手法によって HER2 の自動解析結果の一貫性が高まるという結果を踏まえ、IHC 画像の自動解析からより詳細な判定ができないか検証を行った。通常 IHC 法と併せて検査される ISH 法の指標である ISH ステータスを判定対象とし、IHC 法の解析で得られるパラメータとの関係性を調査した。2+と 3+の濃さに分類された割合をみることで、IHC 画像から ISH ステータスの分類が行える可能性が示唆された。また、DAB の濃さの分布においても ISH ステータスによ

て差が見られる可能性を確認した。IHC 画像からの閾値算出方法については、CISH 法の結果と比較して閾値算出を調整することによってより正確な結果が得られることがわかった。

今後の課題として、今回実験で使用したデータ取得環境とは異なる環境で取得したデータにも提案手法が効果的であるか拡張性の検証が必要である。具体的には、異なるベンダーの染色装置や試薬を利用した染色データ、実験で検証を行っていないスキャナで撮影したデータ、他施設のデータなどが挙げられる。実験を行った施設では検証に用いた 2 台のスキャナ以外にも、主要なベンダーのスキャナを導入している。これらのスキャナにて検証を深めた後、他施設にも範囲を広げた検証が望まれる。また、臨床利用・社会実装に向けた検討として、実用化されている WSI ビューアや解析ソフトに組み込むためのプロトタイプの構築や、ベンダーとの協議を検討している。

提案手法を臨床診断に導入する場合、染色、スキャンのそれぞれの工程にスライドを追加することで対応できるため、臨床のワークフローの変更を軽微に止めることが可能と考える。しかしながら、いくら軽微であっても、ワークフローの変更には現場から反対意見が出ることも想像に難くない。診断に直接関わる病理医だけでなく、染色やスキャンを行う現場に対しても手法を導入することのメリットを示すことができれば、より導入への理解を得やすくなるかもしれない。例えば、本実験で利用したデータの日々の染色による違いは、画像解析には影響してしまっていたものの、病理医の視点では許容できるほど小さいものであるとの意見があった。IHC キャリブレーションから取得した DAB の色味や濃さは日々の染色状況がどのようなかを示すわかりやすい指標であり、染色の安定性の確認や、今後の作業の改善にもつながるかもしれない。お互いのメリットを現場に伝えることで、手法の臨床応用にもつながっていくと期待される。

施設へのスキャナの導入や WSI の研究、臨床利用はアメリカやカナダなど国外を中心に進められている。提案手法の臨床応用をアメリカで進めることができれば、これからデジタル化を進める立場にある日本の病理診断の発展にも貢献できるかもしれない。病理診断には豊富な経験と臨床知識が求められ、その技術的なハードルの高さが新しい担い手不足の要因の一つとも言われる。自動解析が病理医の診断の補助や代替となるツールとして選択しやすくなり、社会的課題の解決につながるよう、本研究が寄与することを期待する。

7.2 研究の発展性

本論文は乳がんにおける HER2 検査を対象として実験を行ったが、胃がんや大腸がんなど他の臓器での HER2 検査にも応用できると考える。他にも、IHC キャリブレーションスライドは

HER2 以外のタンパクの検査にも対応できるように、主要な IHC 検査に対応した抗原ごとにスライドを準備している。使用するスライドの種類を変えることで、HER2 検査以外の IHC 法にも提案手法を拡張できる可能性がある。特に、表 1 に示した PD-L1 の検査は HER2 と同じく細胞膜の染色性を評価する検査であるため、手法を応用しやすいと考えられる。

さらに、IHC キャリブレーションスライドは目視評価用にも応用が可能である。図 87 に活用イメージを示す。6.5 節に記載した ScriptQuant の機能を利用し、WSI と同じ画面にカラーバーを表示するプログラムを作成した。中央に示すカラーバーは、キャリブレーションから取得した DAB の色味と濃さを利用して作成した。図 8 に示した陰性コントロールと陽性コントロールは、典型的な 0 と 3+ のケースに対応している。陽性コントロールは明らかに陽性であるケースが選択されることが多い。そのため、評価対象となる 3+ の組織よりも濃いことが多く、判定が難しい 2+ 周辺の濃さの参考には向いていない。一方、IHC キャリブレーションスライドは 10 段階のマイクロビーズで構成され、レベルに応じて徐々に染色が濃くなるように設計されているため、0 から 3+ までの染色の濃さを詳細に観察することができる。濃さの参考としてカラーバーをビューアに表示することで、経験の浅い病理医の診断の補助や、日々の染色性を確認しながらの評価が可能となる。

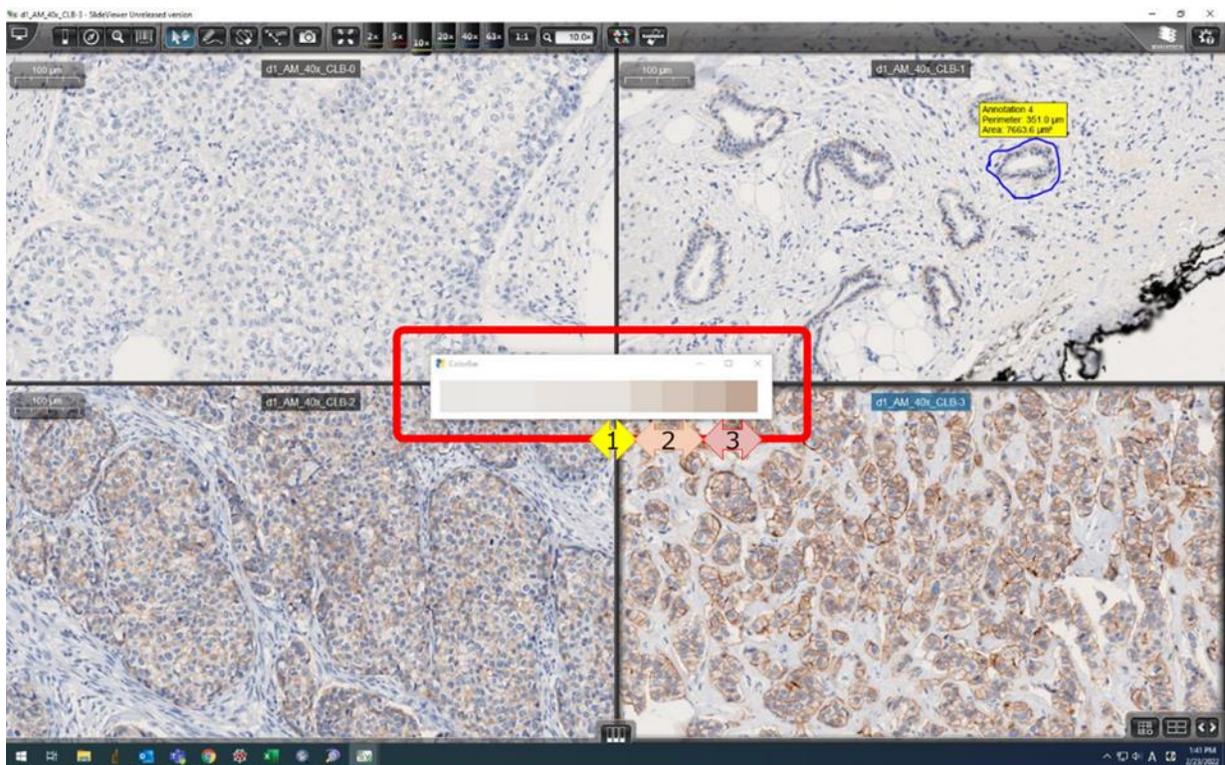


図 87 IHC キャリブレーションの目視評価用における活用イメージ

本手法を適用した HER2 IHC 法の解析は、第 6 章での CISH 法との比較検討のように、他の染色法と組み合わせることでより効果的に作用すると考えられる (図 88)。従来の IHC 法の目視診断では H&E 染色で選択した浸潤領域のうち、20 倍の倍率で 5 視野の領域を選択して評価している。これは目視診断用のガイドラインで定められた視野数ではあるが、目視・自動に依らずこの視野数が最適か、視野数を変えた場合に診断結果がどのように変わってくるかといった関係性は明らかになっていない。自動診断では目視診断と比べ、より柔軟に診断領域を拡大することができる。近年の AI の発展に伴い、H&E 染色画像から浸潤部を自動で選択するような研究も行われている [66]。このような H&E 染色からの情報を IHC 法に組み合わせることで浸潤部全体の解析が可能となり、診断領域とスコアの関係性調査や、がんの不均一性を考慮した診断が可能となる。診断領域を広げることで CISH 法との比較でも新しい知見が得られるかもしれない。現在は自動解析も目視診断用のガイドラインに従って行われているが、自動解析用の新たなガイドラインの確立につながる可能性もある。

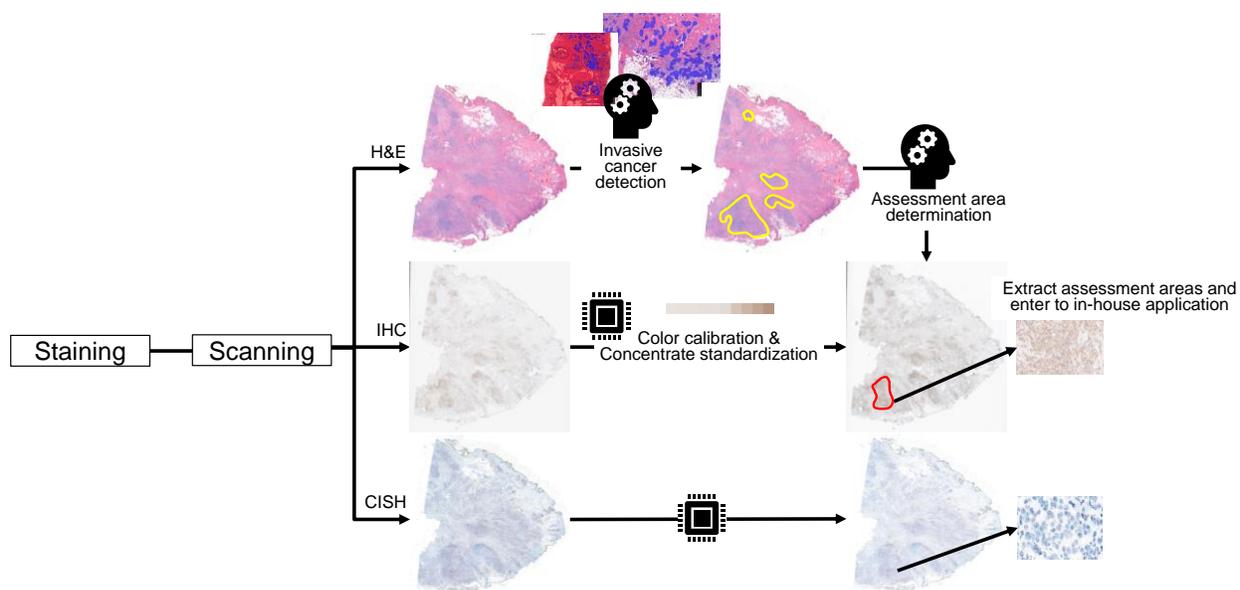


図 88 HER2 の自動診断の発展に向けたフロー

参考文献

- [1] A.C. Wolff, M.E. Hammond, J.N. Schwartz, K.L. Hagerty, D.C. Allred, R.J. Cote, American Society of Clinical Oncology/College of American Pathologists guideline recommendations for human epidermal growth factor receptor 2 testing in breast cancer. *Arch Pathol Lab Med.* 131(1):18–43 (2007).
- [2] A.C. Wolff, M.E. Hammond, K.H. Allison, B.E. Harvey, P.B. Mangu, J.M.S. Bartlett, M. Bilous, I.O. Ellis, P. Fitzgibbons, W. Hanna, R.B. Jenkins, M.F. Press, P.A. Spears, G.H. Vance, G. Viale, L.M. McShane, M. Dowsett, Human Epidermal Growth Factor Receptor 2 Testing in Breast Cancer: American Society of Clinical Oncology/College of American Pathologists Clinical Practice Guideline Focused Update. *Arch Pathol Lab Med.* 142(11):1364–1382 (2018).
- [3] M. Lacroix-Triki, S. Mathoulin-Pelissier, J.P. Ghnassia, G. Macgrogan, A. Vincent-Salomon, V. Brouste, M.C. Mathieu, P. Roger, F. Bibeau, J. Jacquemier, F. Penault-Llorca, L. Arnould, High interobserver agreement in immunohistochemical evaluation of HER-2/neu expression in breast cancer: a multicentre GEFPICS study. *Eur J Cancer.* 42(17):2946–2953 (2006).
- [4] T.A. Thomson, M.M. Hayes, J.J. Spinelli, E. Hilland, C. Sawrenko, D. Phillips, B. Dupuis, R.L. Parker, HER2/neu in breast cancer: interobserver variability and performance of immunohistochemistry with 4 antibodies compared to fluorescent in situ hybridization. *Mod Pathol.* 14(11):1079–1086 (2001).
- [5] M.A. Gavrielides, B.D. Gallas, P. Lenz, A. Badano, S.M. Hewitt, Observer variability in the interpretation of HER2/neu immunohistochemical expression with unaided and computer-aided digital microscopy. *Arch Pathol Lab Med.* 135:233–242 (2011).
- [6] T. Cornish, Clinical application of image analysis in pathology. *Adv Anat Pathol.* 27(4):227–235 (2020).
- [7] A. Gray, A. Wright, P. Jackson, M. Hale, D. Treanor, Quantification of histochemical stains using whole slide imaging: development of a method and demonstration of its usefulness in laboratory quality control. *J Clin Pathol.* 68:192–199 (2015).
- [8] S.A. Bogen, A Root Cause Analysis Into the High Error Rate in Clinical Immunohistochemistry. *Appl Immunohistochem Mol Morphol.* 27(5):329–338 (2019).
- [9] Y. Yagi, Color standardization and optimization in whole slide imaging. *Diagn Pathol.* 6(S1):S15 (2011).
- [10] 患者さんのための乳がん診療ガイドライン2019年版(<https://jbcs.xsrv.jp/guidline/p2019/>)
2023/5/15 訪問

- [11] 国立研究開発法人国立がん研究センター がん情報サービス 乳がん
(<https://ganjoho.jp/public/cancer/breast/index.html>) 2023/5/15 訪問
- [12] D.J. Slamon, G.M. Clark, S.G. Wong, W.J. Levin, A. Ullrich, W.L. McGuire, Human breast cancer : correlation of relapse and survival with amplification of the HER-2/neu oncogene. *Science*. 235(4785):177-82 (1987).
- [13] D. Slamon, W. Eiermann, N. Robert, T. Pienkowski, M. Martin, M. Press, J. Mackey, J. Glaspy, A. Chan, M. Pawlicki, T. Pinter, V. Valero, M.C. Liu, G. Sauter, G. von Minckwitz, F. Visco, V. Bee, M. Buyse, B. Bendahmane, I. Tabah-Fisch, Adjuvant Trastuzumab in HER2-Positive Breast Cancer. *N Engl J Med*. 365(14):1273–1283 (2011)
- [14] 一般社団法人 日本病理学会、病理検体処理ガイドラインワーキンググループ、日本臨床衛生検査技師会、病理検体取扱いマニュアルー病理検体取り違えを防ぐためにー (初版)、2016.
- [15] トラスツズマブ病理学会編、中外製薬、HER2 検査ガイド第 3 版、東京、2009.
- [16] 一般社団法人 日本病理学会、乳癌・胃癌 HER2 病理診断ガイドライン 第 2 版、金原出版、2021.
- [17] J.F. Hechtman, D.S. Ross, The past, present, and future of HER2 (ERBB2) in cancer: Approaches to molecular testing and an evolving role in targeted therapy. *Cancer Cytopathol*. 127(7):428-431 (2019).
- [18] M.D. Zarella, D. Bowman, F. Aeffner, N. Farahani, A. Xthona, S.F. Absar, A. Parwani, M. Bui, D.J. Hartman, A Practical Guide to Whole Slide Imaging - A White Paper From the Digital Pathology Association. *Arch Pathol Lab Med*. 143 (2019).
- [19] 病理診断のためのデジタルパソロジーシステム技術基準 第 3 版、一般社団法人日本病理学会、日本デジタルパソロジー研究会、デジタルパソロジー技術基準検討会、2018.
- [20] D.S. McClintock, J.T. Abel, T.C. Cornish. Whole Slide Imaging Hardware, Software, and Infrastructure. In: Parwani, A.V. (eds) *Whole Slide Imaging*. Springer, Cham. 2022.
- [21] P. Bankhead, M.B. Loughrey, J.A. Fernandez, Y. Dombrowski, D.G. McArt, P.D Dunne, S. McQuaid, R.T. Gray, L.I. Murray, H.G. Coleman, J.A. James, M. Salto-Tellez, P.W. Hamilton, QuPath: Open source software for digital pathology image analysis. *Sci Rep*. 7(16878) (2017).
- [22] D.D. Lee, H.S. Seung, Learning the parts of objects by non-negative matrix factorization, *Nature* 401:788–791 (1999).
- [23] 澤田宏、解説 非負値行列因子分解 NMF の基礎とデータ／信号解析への応用、電子情報通信学会誌 95(9):829–833 (2012)
- [24] P. Smaragdis, B. Raj, M. Shashanka, Supervised and Semi-Supervised Separation of Sounds from Single-Channel Mixtures, In proc. of the 7th International Conference on Independent Component Analysis and Signal Separation, 2007

- [25] D. Kitamura, H. Saruwatari, K. Yagi, K. Shikano, Y. Takahashi, K. Kondo, Music Signal Separation Based on Supervised Nonnegative Matrix Factorization with Orthogonality and Maximum-Divergence Penalties. *IEICE Trans. Fundamentals of Electr.* E97-A(5):1113–1118 (2014).
- [26] K. Nassau. *Color for Science, Art and Technology*. 1st ed. North Holland. Chapter 2 - The Measurement of Color. p.31–96 (1998).
- [27] コニカミノルタ 色々雑学
(<https://www.konicaminolta.jp/instruments/knowledge/color/section2/02.html>)
2023/5/25 訪問
- [28] W. Mokrzycki, M. Tatol, Color difference Delta E – A survey. *Mach Graph Vis.* 20(4):383–411 (2011).
- [29] V. Cheung, S. Westland, D. Connah, C. Ripamonti, A comparative study of the characterisation of color cameras by means of neural networks and polynomial transforms. *Color Technol.* 120(1):19–25 (2004).
- [30] G. Hong, M.R. Luo, P.A. Rhodes, A Study of Digital Camera Colorimetric Characterisation Based on Polynomial Modelling, *Color Res. Appl.*, 26:79–84 (2001).
- [31] G.D. Finlayson, M.S. Drew, White-point preserving color correction. in *Proceedings of the Fifth colour Imaging Conference: colour Standards and colour Measurements*. 258–261 (1997).
- [32] G.D. Finlayson, M.S. Drew, Constrained least-squares regression in color spaces. *J of Electron Imaging*. 6(4):484–493 (1997).
- [33] H. Eliasson, Color calibration of a CMOS digital camera for mobile imaging, *Proc. SPIE 7537, Digital Photography VI*, 753709 (2010).
- [34] L. Anzagira, E.R. Fossum, Color filter array patterns for small-pixel image sensors with substantial cross talk, *J Opt Soc Am A* 32(1):28–34 (2015)
- [35] Mohamed-C. Larabi, S. Süssstrunk, Evaluation of a hyperspectral image database for demosaicking purposes, *Proc. SPIE 7876, Digital Photography VII*, 78760C (2011).
- [36] International Color Consortium, Medical Imaging: Whole Slide Imaging, (https://color.org/groups/medical/histopathology_calibration_slide.xalter) 2023.08.01 訪問
- [37] Technical Performance Assessment of Digital Pathology Whole Slide Imaging Devices, Guidance for Industry and Food and Drug Administration Staff (https://color.org/groups/medical/wsi/Technical_Performance_Assessment_of_Digital_Pathology_Whole_Slide_Imaging_Devices-colour-related_elements.pdf) 2023.08.01 訪問

- [38] P.A. Bautista, N. Hashimoto, Y. Yagi, Color standardization in whole slide imaging using a color calibration slide. *J Pathol Inform.* 5(1):4 (2014).
- [39] A. Korzynska, J. Zak, L. Roszkowiak, D. Pijanowska, T. Markiewicz, "Color standardization for the immunohistochemically stained tissue section images," 2016 IEEE International Conference on Imaging Systems and Techniques (IST), Chania, Greece, 2016, pp. 48-53 (2016).
- [40] E.L. Clarke, C. Revie, D. Brettle, M. Shires, P. Jackson, R. Cochrane, R. Wilson, C. Mello-Thoms, D. Treanor, Development of a novel tissue-mimicking color calibration slide for digital microscopy. *Color Res Appl.* 43(2):184–197 (2018).
- [41] M. Macenko, M. Niethammer, J.S. Marron, D. Borland, J.T. Woosley, X. Guan, C. Schmitt, N.E. Thomas. A method for normalizing histology slides for quantitative analysis, 2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro, Boston, MA, USA, 2009, pp. 1107-1110
- [42] T. Abe, Y. Murakami, M. Yamaguchi, N. Ohyama, Y. Yagi. Color Correction of Pathological Images Based on Dye Amount Quantification. *Opt Rev.* 12, 293–300 (2005).
- [43] Y. Zheng, Z. Jiang, H. Zhang, F. Xie, J. Shi, C. Xue. Adaptive color deconvolution for histological WSI normalization. *Computer methods and programs in biomedicine*, 170, 107–120 (2019).
- [44] T. Kausar, A.Kausar, MA. Ashraf, MF. Siddique, M. Wang, M. Sajid, MZ. Siddique, AU. Haq, I. Riaz. SA-GAN: Stain Acclimation Generative Adversarial Network for Histopathology Image Analysis. *Applied Sciences.* 12(1):288 (2022).
- [45] A.C. Ruifrok, D.A. Johnston, Quantification of histochemical staining by color deconvolution, *Anal Quant Cytol Histol* 23(4), 291–299 (2001).
- [46] J. Xu, L. Xiang, G. Wang, S. Ganesan, M. Feldman, N.N.C. Shih, H. Gilmore, A. Madabhushi, Sparse Non-negative Matrix Factorization (SNMF) based color unmixing for breast histopathological image analysis, *Compute Med Imaging Graph*, 46(1) 20-29 (2015).
- [47] Y. Nie, C. Roessler, E. Andersson, O. Grimm, Color Deconvolution Method with DAB Scatter Correction for Bright Field Image Analysis. *Proc. SPIE 10581, Medical Imaging 2018: Digital Pathology*, 105810K (2018)
- [48] M.V. Sainz de Cea, Y. Nie, Single stain normalization for IHC whole slide images, *Proc. SPIE 10581, Medical Imaging 2018: Digital Pathology*, 1058104 (2018)
- [49] Y.R.V. Eycke, J. Allard, I. Salmon, O. Debeir, C. Decaestecker. 2017. Image processing in digital pathology: an opportunity to solve inter-batch variability of immunohistochemical staining. *Sci Rep.* 7(42964).

- [50] C. Dunn, D. Brettle, M. Cockroft, E. Keating, C. Revie, D. Treanor. The use of a biopolymer film for quantitative H&E stain assessment and quality control in pathology. Pathology Vision 2022, Las Vegas, NV, USA, 2022
- [51] S.R. Sompuram, K. Vani, B. Tracey, D.A. Kamstock, S.A. Bogen, Standardizing Immunohistochemistry: A New Reference Control for Detecting Staining Problem. J Histochem Cytochem. 63:9:681–690 (2015).
- [52] P.V.C. Hough, Method and means for recognizing complex patterns, U.S. Patent 3 069 654, Dec. (1962).
- [53] P.E. Hart, How the Hough transform was invented [DSP History], IEEE Signal Processing Magazine 26(6):18–22 (2009).
- [54] N.L. Johnson, Systems of frequency curves generated by methods of translation. Biometrika, 36, 149–176 (1949)
- [55] 池田郁男、統計検定を理解せずに使っている人のために III、科学と生物、51(7) (2013)
- [56] 浜田知久馬、統計学的な検定の考え方と多重比較、アプライド・セラビューティクス、4(2):55–66 (2013)
(https://www.jstage.jst.go.jp/article/appliedtherapeutics/4/2/4_56/_pdf) 2023/6/8 訪問
- [57] 井口豊、カイ二乗検定（独立性検定）から残差分析へ：全体から項目別への検定 (<https://biolab.sakura.ne.jp/chi-square-residual-analysis.html>) 2023/6/8 訪問
- [58] P.F. Alcantarilla, J. Nuevo, A. Bartoli, Fast explicit diffusion for accelerated features in nonlinear scale spaces. Trans Pattern Anal Machine Intell. 34(7):1281–1298 (2011).
- [59] C.M. Van der Loos, Multiple Immunoenzyme Staining: Methods and Visualizations for the Observation With Spectral Imaging, J Histochem Cytochem. 56(4):313–328 (2008).
- [60] M.S. Hossain, M.G. Hanna, N. Uraoka, T. Nakamura, M. Edelweiss, E. Brogi, M.R. Hameed, M. Yamaguchi, D.S. Ross, Y. Yagi, Automatic quantification of HER2 gene amplification in invasive breast cancer from chromogenic in situ hybridization whole slide images. J of Med Imaging (Bellingham), 6(4):047501 (2019).
- [61] T. Ohnishi, N. Bakoglu, M.G. Hanna, M. Edelweiss, E. Brogi, M.R. Hameed, D.S. Ross, Y. Yagi, AI-Integrated Automated CISH Analysis in Breast Invasive Carcinomas: Evaluation of HER2 Status and ASCO/CAP Group’s Classification & Investigation of Number of Nuclei for the Best Results, USCAP 2023 Annual Meeting, New Orleans Ernest N. Morial Convention Center, New Orleans, LA, USA (2023.3.11–16), Lab Invest. 103(3):S100081 (2023).
- [62] 石垣司、数理統計 補助資料～判別分析～
(http://www2.econ.tohoku.ac.jp/~isgk/lec_material/suri_tokei2021/SuriTokei2021_20.pdf)
2023/6/18 訪問

- [63] C. Palm, C.E. Connolly, R. Masser, B. Padberg Sgier, E. Karamitopoulou, Q. Simon, B. Bode, M. Tinguely, Determining HER2 Status by Artificial Intelligence: An Investigation of Primary, Metastatic, and HER2 Low Breast Tumors. *Diagnostics (Basel)*. 2023 Jan 3;13(1):168.
- [64] OpenSlide (<https://openslide.org/>) 2023/8/9 訪問
- [65] T. Micsik, G. Kiszler, D. Szabó, L. Krecsák, T. Krenács¹, B. Molnár, Is HER2 amplification predictable by digital immunohistochemistry? *Diagn Pathol*. 8 (Suppl 1):S14 (2013).
- [66] N. Bakoglu, E. Cesmecioglu, A. Teplov, C. Ohnishi, M. Yoshida, T. Ohnishi, Y. Yagi, “Artificial Intelligence (AI)-Based Automated Determination of Histomorphological Features and Mitosis Detection in Multiple Organ Systems”, Pathology informatics summit 2023, David L. Lawrence Convention Center and the Westin Convention Center Hotel, Pittsburgh, PA, USA (2023)

研究業績

投稿論文

1. C. Ohnishi, T. Ohnishi, P. Ntiamoah, D.S. Ross, M. Yamaguchi, Y. Yagi.
“Standardizing HER2 immunohistochemistry assessment: Calibration of color and intensity variation in whole slide imaging caused by staining and scanning.”, Applied Microscopy, 53(8) (2023.09) <https://doi.org/10.1186/s42649-023-00091-8>.
2. C. Ohnishi, T. Ohnishi, K. Ibrahim, P. Ntiamoah, D.S. Ross, M. Yamaguchi, Y. Yagi,
“Color standardization and stain intensity calibration for whole slide image-based immunohistochemistry.”, Microscopy and Microanalysis, ozad136 (2023.12) <https://doi.org/10.1093/micmic/ozad136>. *

*Editor’s choice

国際会議での発表

1. C. Ohnishi, K. Ibrahim, P. Ntiamoah, S.A. Bogen, D.S. Ross, M. Yamaguchi, Y. Yagi,
“Initial study to evaluate color and staining intensity for image based IHC evaluations”, Pathology Informatics Summit 2022; David L. Lawrence Convention Center and the Westin Convention Center Hotel, Pittsburgh, PA, USA (2022.01.25-27).*
2. C. Ohnishi, N. Bakoglu, P. Ntiamoah, S.A. Bogen, D.S. Ross, M. Yamaguchi, Y. Yagi,
“Stain and Color Calibration and Standardization for Whole Slide Image Based Automated IHC Assessment”, USCAP 2023 Annual Meeting, New Orleans Ernest N. Morial Convention Center, New Orleans, LA, USA (2023.03.11-16), Lab Invest. 103(3):S100081 (2023). <https://doi.org/10.1016/j.labinv.2023.100081>

* Travel award 受賞

学会・研究会での発表

1. 大西千絵, 山口雅浩, 八木由香子, Steven Bogen, “免疫組織化学染色法における画像の色標準化”, 第 19 回 日本デジタルパソロジー・AI 研究会 総会, p. 52 (2021.08.20-22).
2. 大西千絵, スティーブン ボーゲン, 山口雅浩, 八木由香子, “画像解析技術に基づく HER2 検査の標準化に向けた免疫組織化学染色用キャリブレーションの導入”, メディカルイメージ

ング連合フォーラム, 電子情報通信学会技術研究報告, Vol. 121, no. 347, pp. 133-134 (2022.01.25-27).

3. 大西千絵, 大西峻, P. Ntiamoah, S.A. Bogen, D.S. Ross, 山口雅浩, 八木由香子, “免疫組織化学染色用キャリブレーションを用いた HER2 検査の標準化に向けた Whole Slide Image の解析”, 第 43 回視覚情報基礎研究会研究発表会、CSA-FVI-2022-2, Sept. (2022.09.25).

特許

1. Yagi Y, Ross DS, Ohnishi C, Ohnishi T, “AUTOMATED CHROMOGENIC IN SITU HYBRIDIZATION QUANTIFICATION FROM WHOLE SLIDE IMAGES”, U.S.Patent SK2020-086; F&L 115872-2691. (出願中)

その他の関連業績

1. N. Bakoglu, E. Cesmecioğlu, T. Ohnishi, C. Ohnishi, A. Teplov, D.S. Ross, Y. Yagi, “Optimization of the Scanning Protocols for AI-Integrated Automated CISH Assessment in Breast Invasive Carcinomas on Whole Slide Images”, USCAP 2023 Annual Meeting, New Orleans Ernest N. Morial Convention Center, New Orleans, LA, USA (2023), Lab Invest. 103(3):S100094 (2023). <https://doi.org/10.1016/j.labinv.2023.100094>
2. N. Bakoglu, E. Cesmecioğlu, A. Teplov, C. Ohnishi, M. Yoshida, T. Ohnishi, Y. Yagi, “Artificial Intelligence (AI)-Based Automated Determination of Histomorphological Features and Mitosis Detection in Multiple Organ Systems”, Pathology informatics summit 2023, David L. Lawrence Convention Center and the Westin Convention Center Hotel, Pittsburgh, PA, USA (2023)

謝辞

本研究は東京工業大学工学院情報通信系の博士課程および Memorial Sloan Kettering Cancer Center (MSKCC)に在籍して行ったものです。

本研究を進めるにあたり、多大なるご指導ご鞭撻を賜った山口雅浩教授に心から御礼申し上げます。定期ミーティングとしていつもまとまった時間を割いていただき、的確なアドバイスをいただいたことで、ここまで成果をまとめ上げることができました。また、論文審査等において、貴重なお時間を割いてご指導いただいた中本高道教授、鈴木賢治教授、小尾高史准教授、渡辺義浩准教授に深く感謝いたします。

本研究を始めるきっかけをくださった、MSKCC の八木由香子先生には感謝してもきれません。センターへの入場や研究面でのご指導ご尽力はもちろん、アメリカ生活においても先生のサポートなしには進めることができませんでした。ありがとうございます。実験データの取得や解析にお力添えをいただいた Dr.Ross, Peter, Rene, Nilay, Alexi, Merve, キャリブレータを提供いただいた Boston Cell Standards および Dr.Bogen, Applied Image Inc., WSI スキャナを提供いただいた浜松ホトニクス、3DHistech、皆様のサポートに感謝いたします。

そして、事務手続きなどでお世話になりました研究室秘書の駒井様、八木研究室の Boh, Angela, Tiena, Marco に感謝申し上げます。

最後に、公私にわたってサポートし、励ましてくれた夫の峻に感謝します。