

口腔粘膜炎診断支援システム利用時に生じる
個人の特徴を考慮し、影響を軽減する手法に関する研究

尾崎 南斗

目次

第1章	序論	1
1.1	背景と目的	1
1.2	本論文の構成	4
第2章	口腔画像	7
2.1	はじめに	7
2.2	口腔画像	7
2.3	口腔画像の選定方法	7
2.4	実データを用いた画像選定	10
2.4.1	概要	10
2.4.2	選定結果	10
2.5	切り出し傾向の特定	10
2.6	画像選定で除外した共通画像の調査	11
2.7	おわりに	15
第3章	口腔粘膜疾患診断支援システム	17
3.1	はじめに	17
3.2	診断支援システム	17
3.3	特徴量抽出	18
3.3.1	白斑特徴量	19
3.3.2	白斑形状特徴量	21
3.3.3	発赤特徴量	22
3.3.4	隆起特徴量	22
3.3.5	顆粒特徴量	23
3.4	従来システムを用いた識別率	24
3.4.1	実験概要	24
3.4.2	実験結果	24
3.5	切り出し傾向結果を考慮した従来システムの改良	25
3.6	提案システムを用いた識別率	28
3.6.1	実験概要	28
3.6.2	実験結果	28
3.7	おわりに	30

第 4 章	局所画像サイズ, 枚数パラメータ δ, N 推定方法	33
4.1	はじめに	33
4.2	切り出し場所・範囲におけるデータ化	33
4.3	局所画像サイズ, 枚数パラメータ δ , N 推定方法	35
4.4	重回帰分析	36
4.5	リッジ回帰	36
4.6	Canonical Correlation Analysis	37
4.7	Regularized Canonical Correlation Analysis(RCCA)	38
4.8	Canonical Correlation Analysis SOM	38
4.9	おわりに	41
第 5 章	局所画像サイズ, 枚数パラメータ δ, N 推定実験	43
5.1	はじめに	43
5.2	実験概要	43
5.3	重回帰分析による推定結果	44
5.4	リッジ回帰による推定結果	45
5.5	CCA による推定結果	46
5.6	RCCA による推定結果	47
5.7	CCA-SOM による推定結果	49
5.7.1	推定方法	49
5.7.2	推定結果	49
5.8	5 手法によるパラメータ推定のまとめ	51
5.9	おわりに	53
第 6 章	リッジ回帰を用いた切り出し傾向特定に必要な画像枚数の調査	55
6.1	はじめに	55
6.2	実験条件	55
6.3	実験結果	56
6.4	おわりに	58
第 7 章	結論	62
	謝辞	65

第1章 序論

本論文では、口腔粘膜疾患診断支援システムを使用する前段階である口腔画像内の口腔粘膜疾患に対して範囲指定時に生じる個人の切り出し傾向を考慮し、影響を軽減する手法に関して述べたものである。

1.1 背景と目的

改正医療法の施行以降、臨床医の医療水準の高度化と注意義務の拡大が求められている [1]。歯科医の場合、虫歯治療や歯石除去など歯の治療だけでなく口腔粘膜疾患に対して適切な診断を行う必要に迫られている。口腔粘膜疾患は肉眼で発見できるため、歯科医が患者の歯の治療中に発見を行う。軽度の口腔粘膜疾患の場合は経過観察や歯科医が治療を行うが、重度の口腔粘膜疾患の場合、口腔外科医と呼ばれる専門医に紹介を行い、改めて検査が行われる。口腔は会話するなどのコミュニケーションや日常生活において重要な器官であるため、早期発見及び早期治療が望ましい。そのため、口腔粘膜疾患の発見が遅れると疾患が進行し手術が必要になるが手術後に発話・嚥下障害を引き起こす可能性が高いため、歯科医が早期に口腔粘膜疾患を発見し、適切な治療計画を立案することで Quality Of Life(QOL) を維持することが必要となる。

国立がん研究センターが発表している最新がん統計によると 2018 年の口腔・咽頭癌による罹患数は約 2 万 2 千人、2019 年の口腔・咽頭癌による死亡数は約 8 千人の報告されている。また口腔・咽頭癌は高齢になるにつれて発症する確率が高くなる [2]。口腔癌は、肺や大腸といった臓器と異なり肉眼で確認できるため発見が容易である特徴がある。また、口腔癌は口腔粘膜上皮で過度な酸化刺激時に遺伝子異常が蓄積されることによって多段階発癌機構が働くことから 1 日で発症するというのではなく、5 年以上をかけて発症するといわれている [3]。また、口腔内では様々な病気や病変が生じ、口腔粘膜疾患の中には病変の色・形状が似ている疾患があるため、疾患を適切に診断するのは困難である [4]-[8]。

前述の通り、歯科医は口腔粘膜疾患の早期発見を行うことが望ましいが、全ての歯科医が口腔粘膜疾患に対して強い意識を持っているわけでない。そのため、多くの歯科医は口腔粘膜疾患に対する知識が不十分であり、経験が少ないことが挙げられる。それにより本来、治療の必要がない軽度の口腔粘膜疾患に対し、専門医に紹介することから、患者と専門医の負担が増加するといった問

題が起きている。この問題に対する解決の1例として専門医の紹介に関し、2012年に東京歯科大学口腔顎顔面外科では口腔がんナビゲーションシステム（Navi-system）による歯科医が口腔写真などの画像と共に問診票をインターネットを介して専門医に送ることでアドバイスを頂く医療協力を開始している [9]。また、長崎歯科医師会と長崎大学病院口腔顎顔面外科との間で、電子メールによる口腔粘膜疾患患者の遠隔相談システムという電子メールで送信された口腔画像によって専門医が診断を行う形態を実施している [10]。これらの方法は疾患の早期発見実現へ向けた取り組みだが、専門医は迅速な対応を求められることや2016年時点の専門医は約4千人しかいないことから専門医への負担がさらに増加する可能性が懸念される。しかし、2021年6月に可決された改正医療法に沿った遠隔医療が今後普及されることが期待される [11]。

口腔粘膜疾患の病名を確定するため、侵襲など病理検査や生検による客観的評価を行うことが一般的である。侵襲とは手術、医療処置のような体を傷つけることを指し、病理診断や生検と呼ばれる疑わしい病変の一部を切り取る行為である。しかし、最近の動向として侵襲を行わずに診断を行う研究が多く行われている。光学機器から放たれる青緑光による蛍光画像を撮影し、輝度に注目することで口腔癌、口腔悪性疾患を診断する [12][13]。また、口腔前悪性疾患に対して、高解像度マイクロ内視鏡検査、自家蛍光、巨視による白色光を解析することで生検を行うかどうかの判断および生検部位の場所を特定する研究が行われている [14]。これらの研究を含め、光の作用や性質を利用した画像診断が多く取り組まれている。しかし、遠隔相談システムで用いられている口腔内写真を用いた画像診断は注目されていない。その理由として、口腔内写真は蛍光を照射した画像やCT画像とは異なり撮影環境や撮影条件に依存しておらず規格化が行われていないからである。そのため、関連研究として、スマートフォンのカメラを用いた口腔内画像の撮影角度のアシストを行う口腔内画像撮影支援システムの開発が行われている [15]。研究で取り扱われている口腔粘膜疾患は癌や白板症といった口腔悪性疾患の診断に関するものが多い。また、九州歯科大学の報告では癌、白板症、扁平苔癬のいずれかに当てはまる50例の口腔画像に対して歯科医10-20年目5人、歯科研修医5名、歯学科の学生5名に口腔画像を見せ、どの口腔粘膜疾患か回答を行った結果、平均正答率は歯科医31.8%、歯科研修医58.2%、歯学科の学生は63.2%の正答率であることから、口腔粘膜の知識が全体的に低い傾向や臨床経験が長いほど正答率が低い報告がある [16]。

そこで筆者は遠隔相談システムなどで用いられている口腔画像を用いて歯科院内で利用を想定した口腔粘膜疾患診断支援システムの開発に取り組んでいる。識別する口腔粘膜疾患は癌、白板症、扁平苔癬の3疾患である。癌、白板症、扁平苔癬の例を図1.1に示す。診断支援システムが実現することで専門医の負担を軽減するだけでなく、即時に口腔粘膜疾患の罹患の有無結果を患者に提供できることで早期発見や治療に繋がることを期待される。

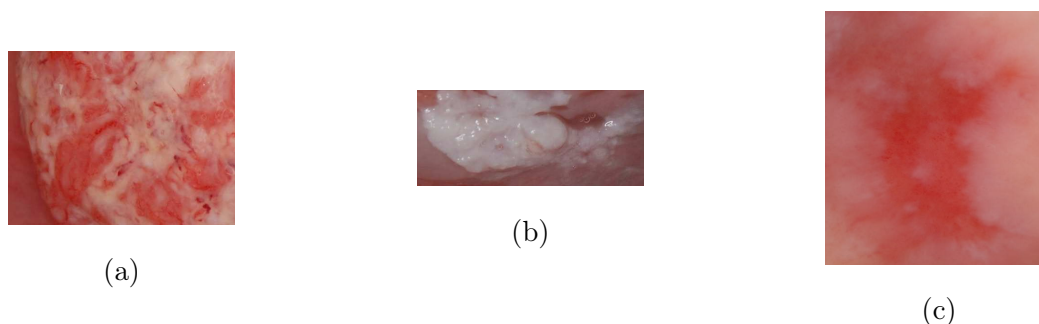


図 1.1: 口腔粘膜疾患 (a) 癌 (b) 白板症 (c) 扁平苔癬

診断支援システムの開発を行うにあたり、口腔画像から疾患場所を検出する方法と疾患場所に対してどのように識別を行うか検討する必要がある。口腔画像から疾患場所を自動で検出する方法として深層学習を用いた研究が行われている。Yolov5l を用いて口腔画像内に疾患があるかどうかを検出した結果、一定の精度が得られている [17]。また、別の報告として Faster RCNN を用いてクラスを 4 つ (専門医の紹介なし、別の理由で専門医に紹介する疾患群、低リスク口腔潜在的悪性疾患 (OPMD)、高リスク癌と OPMD) に分け口腔画像から検出を行った結果、検出率が低い結果となっている [18]。そのため、深層学習では口腔画像のどこに疾患位置があるか大まかな検出は可能ではあるが、複数クラスで設定した場合、疾患の検出は困難になる。また、疾患場所に関してある程度位置検出することは可能だが、具体的な疾患場所は検出できないため、歯などの不必要な情報が含まれる可能性が考慮される。そのため現状ではシステム利用者が手動によって疾患範囲を指定する必要がある。また、疾患範囲に対する識別に関して、2 クラス (例:癌であるかどうか、OPMD であるかどうかなど) の識別では十分な識別精度を得られているが、多クラス (正常、専門医の紹介なし、別の理由で専門医に紹介する疾患群、低リスク口腔潜在的悪性疾患 (OPMD)、高リスク癌と OPMD) の識別に関しては識別が困難である [17][18]。また、これらの実験では複数の疾患を 1 つのクラスとして扱っているため、本論文で取り扱う疾患毎に行われていない。

そこで筆者ら開発している診断支援システムでは口腔画像を用い、口腔画像から疾患部位を手動で切り出し、疾患部位に対して識別を行うシステムを採用した。患者の口腔内を図 1.2 に示すデジタルカメラで撮影した口腔画像を使用し、これまでの研究成果として口腔外科医の診断をモデル化を行い、属性特徴量に基づいて 3 つの口腔粘膜疾患と正常部位の 4 分類を行う手法を提案することで有用性を示している [22]-[24]。さらに属性特徴量の追加及び画像の解像度調整を行うことで診断支援システムの識別率は向上した [25]。

診断支援システム利用者は利用前に口腔画像内から疾患だと思われる場所を指定する必要がある。図 1.3 では疾患だと思われる箇所に対して 3 人の歯科医と専門医が切り出した範囲を示してい



図 1.2: 口腔画像

る。黄色は歯科医 2 年目，水色は歯科医 5 年目，青色は歯科医 30 年目，緑色は専門医が切り出した範囲を示している。図 1.3 より切り出しを行う人によって疾患だと思われる場所やサイズが異なっていることがわかる。切り出し範囲を診断支援システムの入力としているため，疾患に近い場所を切り出している場合でも，切り出し範囲や場所の違いによって識別結果が変化することがある。そのため，識別率を向上させるためには切り出し範囲を考慮したシステム構築を目指す必要がある。

個人の特性を考慮したシステム開発の関連研究として，医薬品の薬物有害反応に関し，クラウドソーシングによってペアワイズ比較によるランク付けを行うことで，医薬品のリスク評価を行った報告がある [19]。また，教育において子供を対象にした学習能力向上に関する研究報告もある。子供は事前に学習テストを行い，その結果をシステムに学習させる。学習モデルを用いて，授業内での子供の学習成果より適切な休憩タイミングを提案したり，長期間において学習成果より隠れマルコフモデルを用いて授業内容の構成を組み立てる研究が行われている [20][21]。

本研究の目的は歯科医が現場で使用できる診断支援システムの実用化に向け，切り出し範囲や位置といった個人の特性を考慮するシステム構築を行う。本研究は九州歯科大学との連携により共同研究を行っている。使用する口腔画像は全て同大学から提供されたものである。

1.2 本論文の構成

本論文は 7 つの章から構成される。

第 1 章は序論について述べた。

第 2 章では本研究で扱う口腔画像について説明を行う。ここでは口腔画像から疾患だと思われる範囲設定の中には疾患範囲を含んでいない画像が含まれる可能性がある。それらの口腔画像は

□ : 歯科医2年目 □ : 歯科医30年
□ : 歯科医5年目 □ : 専門医

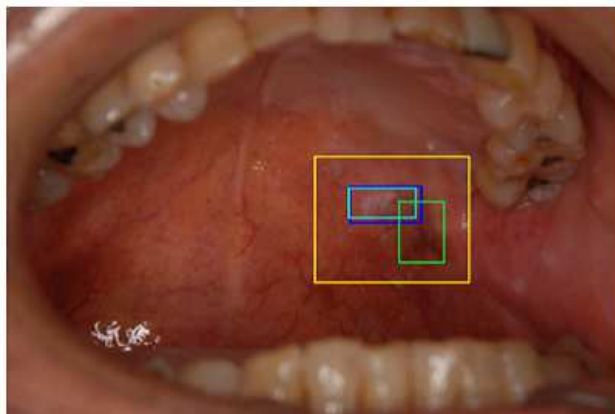


図 1.3: 4名による切り出し範囲の違い

診断支援システムにおいて誤識別を引き起こす可能性や歯科医の切り出し傾向を捉えられなくなる可能性が考えられるため、画像選定方法について述べ、切り出し傾向の調査を行った結果に関して述べる。

第3章では口腔粘膜疾患診断システムについて述べるとともに2章の切り出し傾向の結果より口腔粘膜疾患診断支援システムを改良するとともに識別率に関する実験を行った結果について述べる。

第4章では改良した口腔粘膜疾患診断支援システムで用いる局所画像サイズ・枚数パラメータ推定に用いる推定方法として重回帰分析，リッジ回帰，CCA，RCCA，CCA-SOMについて説明を行い，識別率と切り出し範囲のデータを入力とした推定方法について述べる。

第5章では第4章で述べたパラメータ推定の実験を行い，どの推定方法が適しているかについて述べる。

第6章ではパラメータ推定において切り出し傾向を特定できる必要最小限の枚数を調査したことについて述べる。実際の現場で口腔粘膜疾患診断支援システムを導入した際，新規利用者のことを想定し，事前準備の手間を軽減させるものである。

第7章は結論である。

第2章 口腔画像

2.1 はじめに

本章では、本論文で扱う口腔画像に説明するとともに実際のデータを用いて個人の切り出し傾向について調査した結果を示す。また個人の切り出し傾向を正確に捉えるため、疾患範囲と異なる切り出し範囲を設定しているなど除外する口腔画像について説明する。2.2節では、口腔画像の概要について説明する。2.3節では、除外する口腔画像の種類について説明する。2.4節では、実際のデータを用いた画像選定結果について述べる。2.5節では、画像選定結果の口腔画像を用いて切り出し面積より切り出し傾向を調査した結果について述べる。2.6節では、画像選定結果より総口腔画像枚数の見直しを行った結果について述べる。2.7節では、本章のまとめである。

2.2 口腔画像

筆者の所属研究室で開発している診断支援システムは、歯科医院での使用されることを想定しているため、患者および歯科医にとって負担を軽減するシステムを目指している。したがって、光学機器による蛍光画像や病理画像などは使用せず、デジタルカメラによって撮影された患者の口腔内の画像（以下、口腔画像とする）を用いて識別を行う。口腔画像は、疾患と思われる部位を中心に合わせて撮影される。そのため、撮影における照明の明暗、撮影角度、口腔画像における疾患の大きさが口腔画像毎に異なっている特徴がある。本論文で使用する口腔画像は全て病理検査より疾患が確定したものであり、使用する口腔画像枚数は117枚である。これらの口腔画像は癌、白板症、扁平苔癬の3病変の疾患である。また、口腔画像は撮影されたものを使用しているため、疾患場所が描画されていない。そのため、歯科医は図2.1に示すように疾患だと思われる範囲を手動で設定する。

2.3 口腔画像の選定方法

歯科医が疾患だと思われる範囲を手動で描画することから疾患範囲とは異なる場所を指定する可能性が考えられる。これは、個人の切り出し傾向特定が難しくなると考えられるため、画像選定を行う必要がある。また、2.2節でも述べたが本論文で用いる口腔画像は病理検査結果は明らか

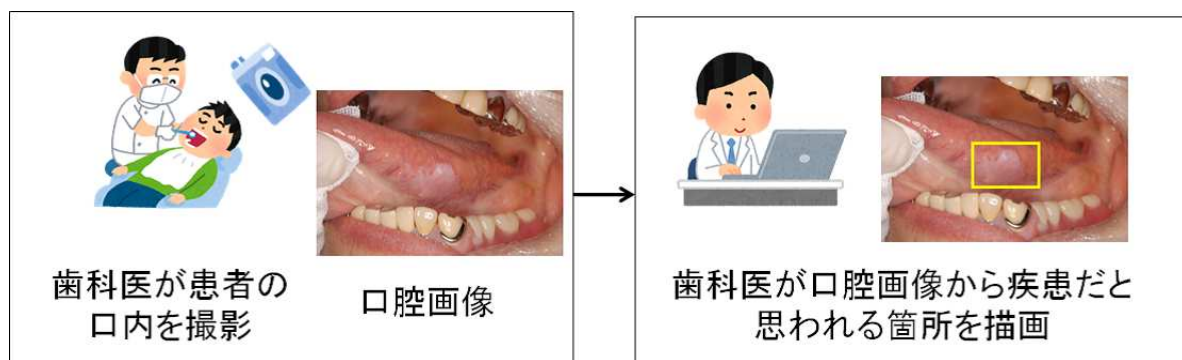


図 2.1: 口腔画像内の疾患切り出し描画方法

だが、どこに疾患があるか明記されていないため、専門医が設定した範囲を正解と仮定する。画像選定では (図 2.2) に示すように

- (a) 専門医と歯科医で全く異なる場所を切り出している画像
 - (b) 専門医と歯科医の切り出した共通範囲が小さい画像
 - (c) 専門医と比べて歯科医の切り出し範囲が極端に大きいもしくは小さい画像
- のいずれかを満たす画像を除外した。画像選定手順を以下に示す。

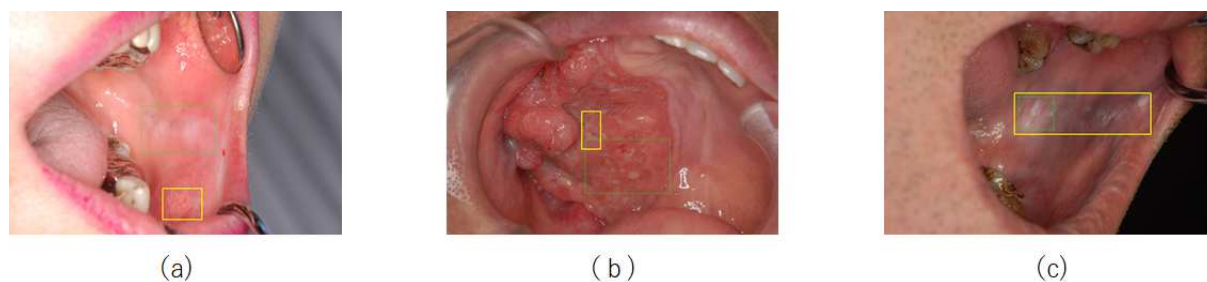


図 2.2: 切り出し傾向を正確に捉えるために除外すべき画像例 黄色:歯科医, 緑:専門医 (a) 専門医と歯科医で全く異なる場所を切り出した例 (b) 歯科医と専門医の共通範囲が小さい例 (c) 歯科医の切り出し範囲が極端に大きい例

- 1: 全く異なる場所を指定した場合 (図 2.2(a))
- 2: 専門医と歯科医が切り出した範囲の midpoint を求める (図 2.3(a)). そして 2つの midpoint を直線で結び midpoint を求める. 直線の midpoint が専門医と歯科医が切り出した範囲の共通範囲に存在しない場合 (図 2.3(b))
- 3: 切り出し面積に関して, 1) 専門医を基準 (専門医/歯科医), 2) 歯科医を基準 (歯科医/専門医) にしてそれぞれ倍率を求める. これを全口腔画像に対して行い, 結果の例を (図 2.4) に示す. 各

口腔画像の面積比における倍率の平均 (μ), 標準偏差 (γ) を求める. 歯科医の切り出し範囲が極端に大きいもしくは小さい例に対して以下の2つの条件どちらかに当てはまる例を除外する.

(1): 専門医/歯科医で $\mu+2\gamma$ 以上かつ歯科医/専門医で $\mu-\gamma$ 以下を満たす

(2): 歯科医/専門医で $\mu+2\gamma$ 以上かつ専門医/歯科医で $\mu-\gamma$ 以下を満たす

$\mu-\gamma$ に設定した理由として $\mu-2\gamma$ に設定すると負の値になる場合が多くあったためである.

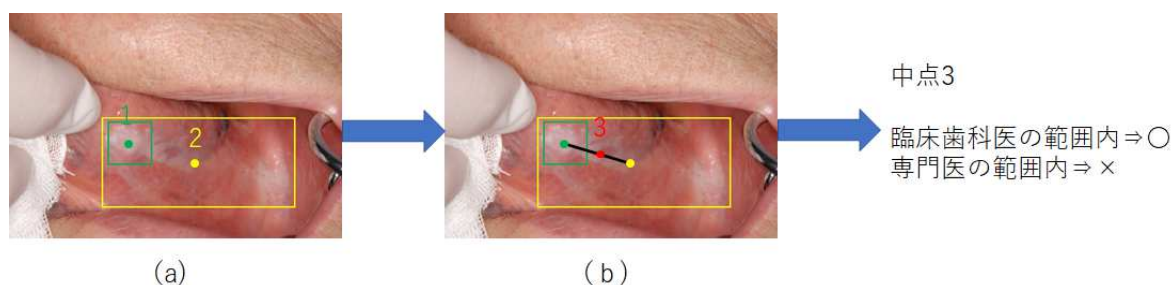


図 2.3: 専門医と歯科医の共通範囲が小さい場合に対する除外方法 黄色: 歯科医, 緑: 専門医 (a) 中点を算出 1: 歯科医 2: 専門医 (b) 1 と 2 の中点を直線で結び, 直線の中点 3 を算出. 中点 3 が歯科医と専門医の切り出した範囲の共通範囲かどうか確認

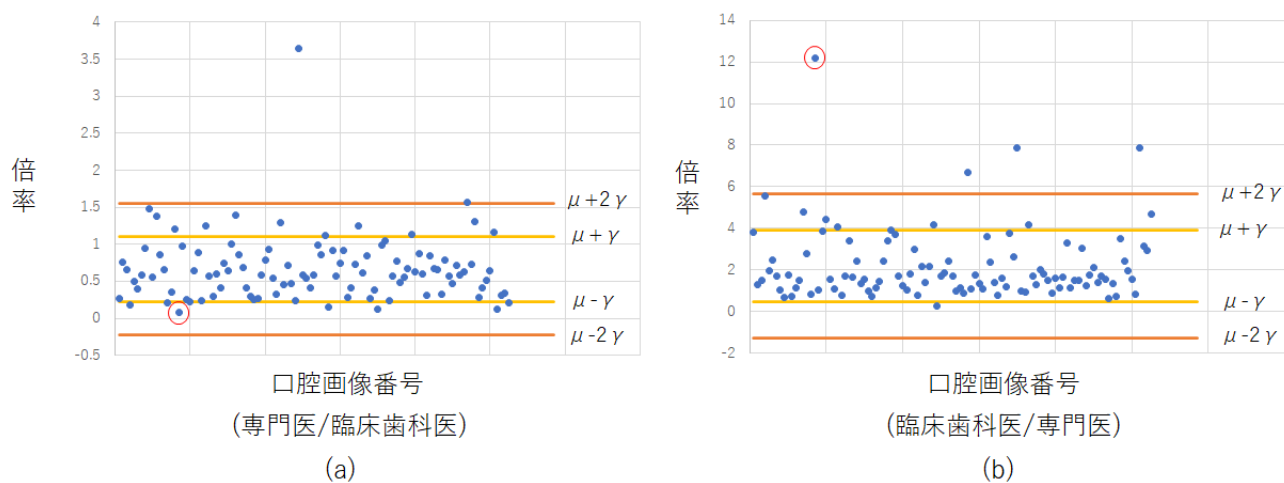


図 2.4: 切り出し面積における倍率 (a) 専門医/歯科医 (b) 歯科医/専門医

例として赤い丸の中にある点に注目すると (a) は $\mu-\gamma$ 以下, (b) は $\mu+2\gamma$ 以上を満たすため除外される

2.4 実データを用いた画像選定

2.4.1 概要

使用したデータは口腔画像枚数が 117 枚であり、切り出しを行った人は専門医 1 名，歯科医 2 年目 1 名，歯科医 5 年目 1 名，歯科医 15 年目 2 名，歯科医 30 年目 4 名と工学部の学生 9 名の計 18 名である。これらを用いて 2.3 項で述べた画像選定を行った。

2.4.2 選定結果

画像選定 1（全く異なる場所）と 2（中点が共通範囲にない）を行った結果を表 3.1 に示す。表 3.1 内の欠損率は専門医の枚数を基準に算出している。欠損率が高い人は学生 5 の 27 枚，学生 7 の 19 枚，学生 9 の 41 枚である。次に画像選定 3 の切り出し範囲が極端に大きいもしくは小さい例を除外した結果を表 2.2，画像選定 3 を実施する前後の専門医/歯科医 (学生) 切り出し範囲倍率の変化を図 2.5 に示す。図 2.5 内の青線は画像選定 3 実行前を，橙色の線は画像選定 3 実行後を示している。多くの人は画像選定 3 を行うことで極端な切り出し例は除外されたが，画像選定 1 と 2 の結果より欠損率が高かった学生 5, 7, 9 の 3 名は変化がない結果となった。以上より，3 名に関しては口腔画像内の疾患位置がどこかわかっていない可能性があると考えられるため，個人の傾向を捉えることは難しいと判断し，以降の実験では 3 名のデータを使用しない。

2.5 切り出し傾向の特定

画像選定 1-3 を行い，除外されなかった口腔画像に対して，専門医/歯科医 (学生) を基に面積，縦の長さ，横の長さの倍率を求め，各倍率より切り出し傾向について調査を行った。切り出し傾向については専門医/歯科医 (学生)=1 を基準にし，1 以上は「大きい」，1 未満は「小さい」とし，歯科医 8 名，学生 6 名のデータを用いる。その結果を表 2.3 に示す。表 2.3 に示す面積，縦の長さ，横の長さの「大」と「小」は専門医と比較し，どのような切り出し傾向であるかを示している。「大」となっている人は専門医より大きく切り出す傾向があるということになる。また括弧内の数字については切り出す傾向の可能性を使用した画像枚数から算出しているため，数字が高い程，その切り出し傾向である可能性が高いといえる。

面積倍率において切り出し傾向が大きい（専門医より大きく切り出す）人は縦と横の長さも長い傾向が見られる場合が多かった。また切り出し傾向が小さい人は縦と横の長さが短いといった同様の傾向が見られた。特に学生 1 に関しては面積倍率が大きい結果であったが縦と横の長さに注目すると短い傾向である。しかし各括弧内の数字に注目すると面積，縦と横の長さ，全て 50[%] 前後であるため口腔画像によって大きい，もしくは小さく切り出す可能性がある。そのため，切り

表 2.1: 全く異なる場所 (画像選定 1) と中点が共通範囲にない例 (画像選定 2) を除外した結果

	画像枚数 [枚]	欠損率 [%]
専門医	117	0
歯科医 2 年目	115	1.7
歯科医 5 年目	111	5.1
歯科医 15 年目 1	112	4.3
歯科医 15 年目 2	112	4.3
歯科医 30 年目 1	105	10.3
歯科医 30 年目 2	110	6
歯科医 30 年目 3	111	5.1
歯科医 30 年目 4	111	5.1
学生 1	103	12
学生 2	105	10.3
学生 3	112	4.3
学生 4	105	10.3
学生 5	90	23.1
学生 6	108	7.7
学生 7	98	16.2
学生 8	107	8.6
学生 9	76	35

出し傾向の特定が難しいと考えられる。また、学生 1 と同様、歯科医 30 年目 1、学生 4 も面積、縦横の長さが 50[%] 前後であり、切り出し傾向の特定が難しいと考えられるため、以降の実験では 3 名のデータは使用しない。

一方で専門医より大きく切り出す傾向が見られた人は、歯科医 2 年目、歯科医 5 年目、歯科医 15 年目 1、歯科医 30 年目 2、歯科医 30 年目 4、学生 3、学生 6、学生 8 の 8 名である。小さく切り出す傾向が見られた人が歯科医 15 年目 2、歯科医 30 年目 3、学生 2 の 3 名である。これらをまとめたものを表 2.4、表 2.5 に示す。

2.6 画像選定で除外した共通画像の調査

画像選定を行った結果、除外された口腔画像番号が共通していることが判明した。5 章で示すパラメータ推定を行うための入力データに切り出し範囲の座標データを用いる。また、入力データを作成する前準備として各口腔画像に対して正規化を行うため、人数が少ない場合、個人の切り出し傾向が変化する可能性が考えられる。そこで除外された画像を調査し、全体の口腔画像枚数

表 2.2: 全く異なる場所 (画像選定 1) と中点が共通範囲にない例 (画像選定 2) と極端な切り出し例 (画像選定 3) を除外した結果

	画像枚数 [枚]	欠損率 [%]
専門医	117	0
歯科医 2 年目	110	5.98
歯科医 5 年目	109	6.83
歯科医 15 年目 1	107	8.55
歯科医 15 年目 2	103	12
歯科医 30 年目 1	101	13.7
歯科医 30 年目 2	101	13.7
歯科医 30 年目 3	105	10.3
歯科医 30 年目 4	99	15.4
学生 1	92	21.4
学生 2	95	18.8
学生 3	100	14.5
学生 4	100	14.5
学生 5	90	23.1
学生 6	105	10.3
学生 7	96	17.9
学生 8	99	15.4
学生 9	76	35.0

を調整する必要がある。

調査を行った結果を図 2.6 に示す。図 2.6 では 11 人中、何人が除外されたかに関する画像枚数を記載している。口腔画像 117 枚中、11 人誰も除外されていない口腔画像が 51 枚、11 人中 1 人以下が除外されている口腔画像 76 枚、11 人中 2 人以下が除外されている口腔画像は 91 枚であった。今回閾値として 11 人の 80% を指定したため 11 人中 2 人以下の 91 枚を採用した。総口腔画像枚数 117 枚から 91 枚に減らした場合の推移を表 2.6 に示す。口腔画像枚数の調整を行うことで欠損率が大幅に減少していることから、口腔画像によっては疾患の場所を見つけることが難しかったと考えられるため、撮影でも様々な角度で複数枚撮ったり動画で撮る必要があるなど検討を行う必要がある。

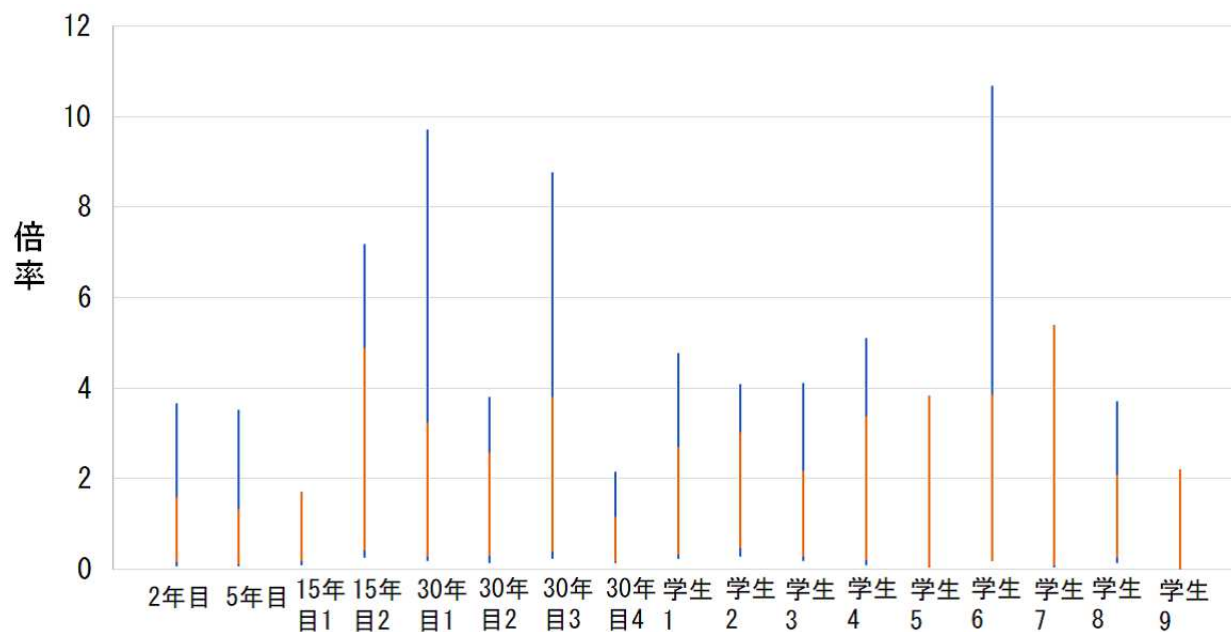


図 2.5: 面積最大最小倍数 (専門医/歯科医 (学生)) 青色:画像選定 3 実行前 橙色:画像選定 3 実行後

表 2.3: 専門医/歯科医 (学生) の面積と縦横の長さに関する切り出し傾向

	面積倍率	面積	縦の長さ	横の長さ
歯科医 2 年目	0.17-1.56	大 (84[%])	大 (80[%])	大 (73[%])
歯科医 5 年目	0.11-1.3	大 (96[%])	大 (91[%])	大 (92[%])
歯科医 15 年目 1	0.19-1.68	大 (90[%])	大 (78[%])	大 (84[%])
歯科医 15 年目 2	0.29-4.88	小 (70[%])	小 (73[%])	小 (71[%])
歯科医 30 年目 1	0.29-3.23	小 (55[%])	小 (51[%])	小 (53[%])
歯科医 30 年目 2	0.3-2.57	大 (62[%])	大 (63[%])	大 (58[%])
歯科医 30 年目 3	0.4-3.79	小 (62[%])	小 (55[%])	小 (68[%])
歯科医 30 年目 4	0.17-1.15	大 (95[%])	大 (77[%])	大 (76[%])
学生 1	0.34-2.68	大 (52[%])	小 (52[%])	小 (51[%])
学生 2	0.48-3	小 (64[%])	小 (58[%])	小 (66[%])
学生 3	0.29-2.16	大 (64[%])	大 (62[%])	大 (59[%])
学生 4	0.22-3.36	大 (54[%])	大 (73[%])	大 (51[%])
学生 6	0.2-3.83	大 (61[%])	大 (56[%])	大 (56[%])
学生 8	0.26-2.06	大 (60[%])	大 (53[%])	大 (62[%])

表 2.4: 切り出し傾向特定が困難な人と理由

疾患範囲とは違う箇所が多い	切り出し傾向が曖昧
学生 5	歯科医 30 年目 1
学生 7	学生 1
学生 9	学生 4

表 2.5: 切り出し傾向が見られた人

大きく切りだす傾向	小さく切り出す傾向	切り出し傾向が曖昧
歯科医 2 年目 歯科医 5 年目 歯科医 15 年目 1 歯科医 30 年目 2 歯科医 30 年目 4 学生 3 学生 6 学生 8	歯科医 15 年目 2 歯科医 30 年目 3 学生 2	歯科医 30 年目 1 学生 1 学生 4

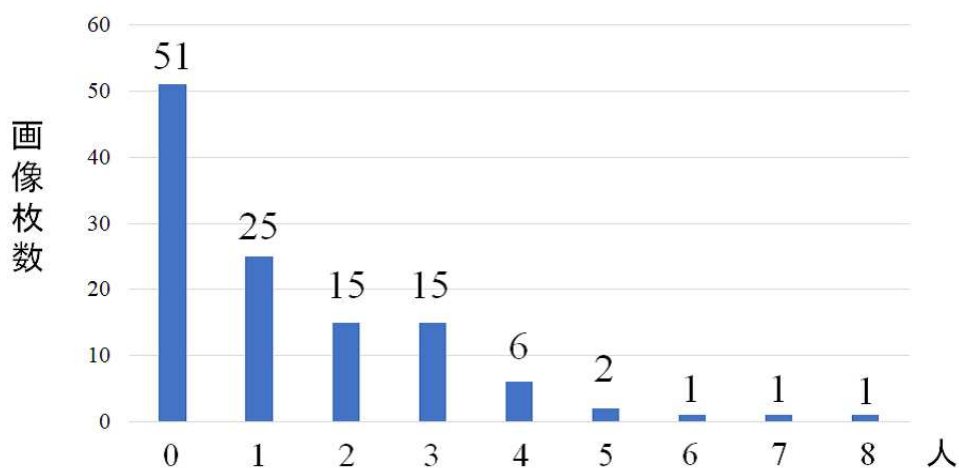


図 2.6: 除外された共通枚数 (対象 11 人)

表 2.6: 口腔画像 117 枚から 91 枚に調整した結果, 個人の画像枚数と欠損率の推移

	画像枚数 [枚]		欠損率 [%]	
	117 枚	91 枚	117 枚	91 枚
専門医				
歯科医 2 年目	110	88	5.98	3.3
歯科医 5 年目	109	87	6.83	4.4
歯科医 15 年目 1	107	90	8.55	1.1
歯科医 15 年目 2	103	85	12	6.59
歯科医 30 年目 2	101	85	13.7	6.59
歯科医 30 年目 3	105	88	10.3	3.3
歯科医 30 年目 4	99	84	15.4	7.69
学生 2	95	82	18.8	9.89
学生 3	100	86	14.53	5.49
学生 6	105	88	10.26	3.3
学生 8	99	83	15.38	8.79

2.7 おわりに

本章では口腔画像 117 枚を専門医 1 名, 歯科医 8 名, 学生 9 名が切り出し範囲指定したデータを用いて実験を行った. 2.2 節では実験に使用する口腔画像について述べた. 2.3 節では描画データは手動によるものであるため, 疾患範囲と異なる場所や極端な切り出し範囲も含まれる. それらは個人の切り出し傾向を正確に捉えられない可能性があるため, 画像選定を行う必要があることから具体的な方法について述べた. 2.4 節では実際のデータを用いて画像選定を行った結果, 学生 3 名が疾患範囲と異なる場所を切り出している例が多く見られた. 2.5 節では切り出し範囲の面積, 縦, 横の長さを参照した結果, 歯科医 5 名, 学生 3 名が大きく切り出す傾向がみられ, 歯科医 2 名, 学生 1 名が小さく切り出す傾向がみられた. また, 口腔画像によって大きく切り出したり小さく切り出す曖昧な切り出し傾向は歯科医 1 名, 学生 2 名の結果であり, これは口腔画像に依存している可能性が高いため, 3 章以降の実験ではデータを用いない. 2.6 節では画像選定において除外された口腔画像が共通していたことから総口腔画像枚数の調整を行った結果, 117 枚から 91 枚に調整した.

第3章 口腔粘膜疾患診断支援システム

3.1 はじめに

本章では、筆者の研究室で開発している診断支援システムの概要ならびに切り出し傾向の結果より改良した診断支援システムに関して説明する。3.2節では、診断支援システムの概要について説明する。3.3節では、診断支援システムの識別で使用する特徴量について説明する。3.4節では、切り出し傾向が見られた人を対象に診断支援システムを用いた識別に関する実験について述べる。3.5節では、切り出し傾向の結果を踏まえ、局所画像を生成する診断支援システムに改良した概要について述べる。3.6節では、切り出し傾向が見られた人を対象に、改良した診断支援システムを用いた識別に関する実験について述べる。3.7節では、本章のまとめである。

3.2 診断支援システム

図 3.1 に筆者の研究室でこれまで開発されてきた診断支援システム（以後、従来システムと呼ぶ）を示す。入力画像は口腔画像から疾患範囲だと思われる範囲を手動で切り出した画像を用いている。切り出し画像に対して解像度低減を行っており、元木らは解像度を 0.2 倍に設定している。0.2 倍に設定した理由として、後述する特徴量抽出が専門医が患者を診断するときの状況に基づき決定された。低解像度化により不必要な特徴量情報が削除され、識別率が向上したことが確認されている。特徴量抽出は白斑特徴量、白斑形状特徴量、発赤特徴量、隆起特徴量、顆粒特徴量の 5 つが識別に有効な特徴量であり、詳細は 3.3 節にて説明する。識別は Support Vector Machine (SVM) を用いている。SVM は、汎化性能が高い学習モデルであり、データ数が少ない場合でも高い識別率が期待できる [26]。SVM の学習で使用したデータは口腔画像内の専門医の範囲より 5 つの特徴量を参照して学習している。扁平上皮癌、白板症、扁平苔癬、正常の各 50 例、計 200 例のデータを用いている。そして、SVM による出力は「扁平上皮癌」、「白板症」、「扁平苔癬」、「正常」のいずれかが出力され、それを識別結果とする。

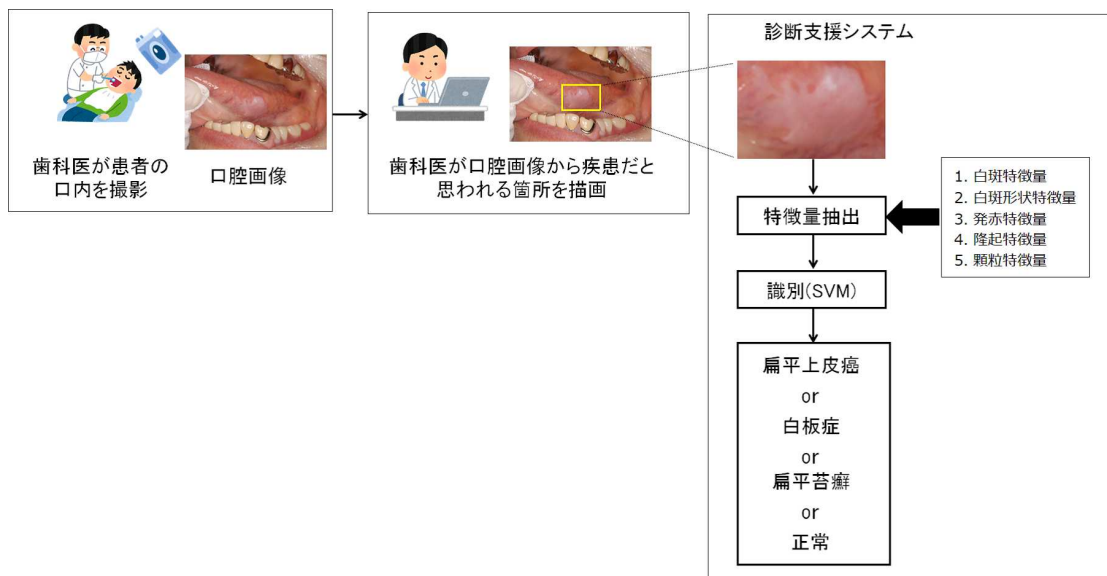


図 3.1: 診断支援システム (従来システム)

3.3 特徴量抽出

従来システムで用いている特徴量に関して述べる。表 3.1 に各口腔粘膜疾患と正常の特徴を示す。癌の特徴として白斑形状が複雑であり、顆粒が存在している。白板症は白斑形状が板状であり、発赤がない。扁平苔癬は白斑形状が線状であり、発赤がある。このように各疾患に対して有効な特徴量を設定することで識別を実現している [27]-[29]。2.3.1 項では白斑特徴量について、2.3.2 項では白斑形状特徴量について、2.3.3 項では発赤特徴量について、2.3.4 項では隆起特徴量について、2.3.5 項では顆粒特徴量について述べる。

表 3.1: 各口腔粘膜疾患および正常の特徴

	白斑	白斑形状	発赤	隆起	顆粒
癌	○	複雑	-	○	○
白板症	○	板状	×	○	×
扁平苔癬	○	線状	○	×	×
正常	×	-	-	×	×

3.3.1 白斑特徴量

白斑形状を基に識別を行う手法を松本らが提案しており、白斑抽出を適切に行うことは口腔粘膜疾患の識別に重要である [30]. 口腔画像から教師付き確率的弛緩法を用いて白斑抽出を行っているが、癌にみられる複雑形状、扁平苔癬にみられる線形状といった白斑の形状を抽出するのは困難である. 白斑抽出を行うため、切り出し画像内に白斑である領域と白斑でない領域があると仮定して、2 値化により白斑抽出を行う. 西らの研究で色の鮮やかさを表す HSV 色空間の彩度 (S) を用いることで白斑領域が分離できることが確認されている. まず、RGB 色空間から HSV 色空間へ変換を行い、彩度画像を生成する. HSV 色空間とは色相、彩度、明度の 3 要素で構成されている. RGB 色空間から HSV 色空間に変換する式を以下に示す. なお $m_{max} = \max(R; G; B)$, $m_{min} = \min(R; G; B)$ とする [31].

$$H = \begin{cases} 0 & (m_{max} = m_{min}) \\ 60 \frac{G-B}{m_{max}-m_{min}} & (R = m_{max}) \\ 60(2 + \frac{B-R}{m_{max}-m_{min}}) & (G = m_{max}) \\ 60(4 + \frac{R-G}{m_{max}-m_{min}}) & (B = m_{max}) \end{cases} \quad (3.1)$$

$$S = \begin{cases} 0 & (m_{max} = m_{min}) \\ \frac{m_{max}-m_{min}}{m_{max}} & (m_{max} \neq m_{min}) \end{cases} \quad (3.2)$$

$$V = m_{max} \quad (3.3)$$

彩度画像のヒストグラムに基づき、ノイズ除去を行う. ヒストグラム中 0.1%未満の画素値はノイズとみなし除外する. 各疾患のヒストグラムを図 3.2 に示す. 図 3.2(d) の正常は白斑がないため彩度が低い値がない. よって、白斑をもつ疾患に比べて彩度のヒストグラムの幅は狭くなることから、彩度のヒストグラム幅を特徴量 S_{broad} とする. S_{broad} を次式に示す.

$$S_{broad} = S_{max} - S_{min} \quad (3.4)$$

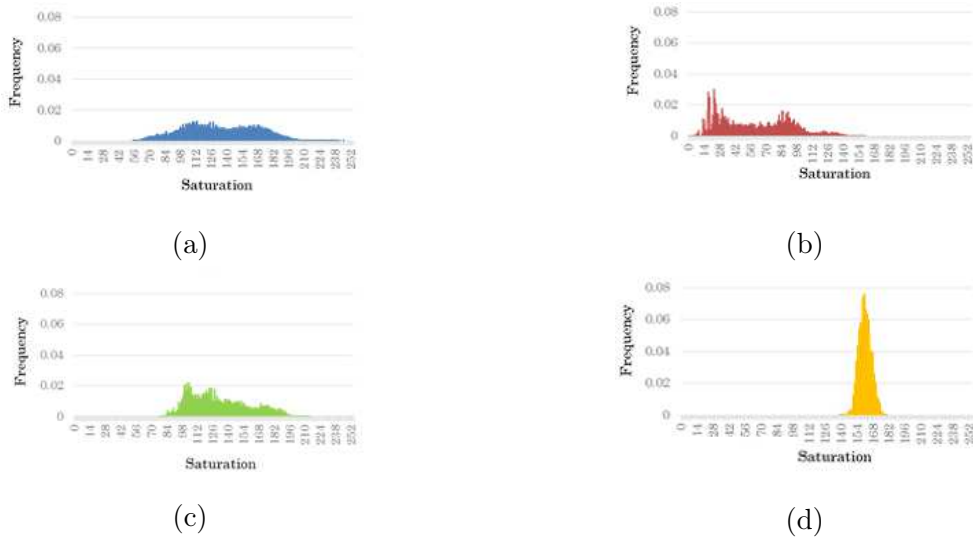


図 3.2: 彩度ヒストグラム (a) 癌 (b) 白板症 (c) 扁平苔癬, (d) 正常

3.3.2 白斑形状特徴量

白斑特徴量で用いた彩度画像のヒストグラムの S 値の最大値 S_{max} , 最小値 S_{min} を用いて閾値 T を決定する.

$$T = \frac{S_{max} - S_{min}}{2} \quad (3.5)$$

画素値が式 3.5 の閾値未満の画素を白斑領域とする. 図 3.3 に各疾患及び正常部位の白斑領域抽出例を示す.

白板症の白斑形状は板状であり, 他の疾患と比べて円形に近いことから, 形状がどれくらい円に近いかを表す円形度を白斑形状の特徴量として設定する. 白斑領域の画素数を S_w , 白斑領域の端の画素数を L とすると円形度 F は,

$$F = 4\pi \frac{S_w}{L^2} \quad (3.6)$$

となる.

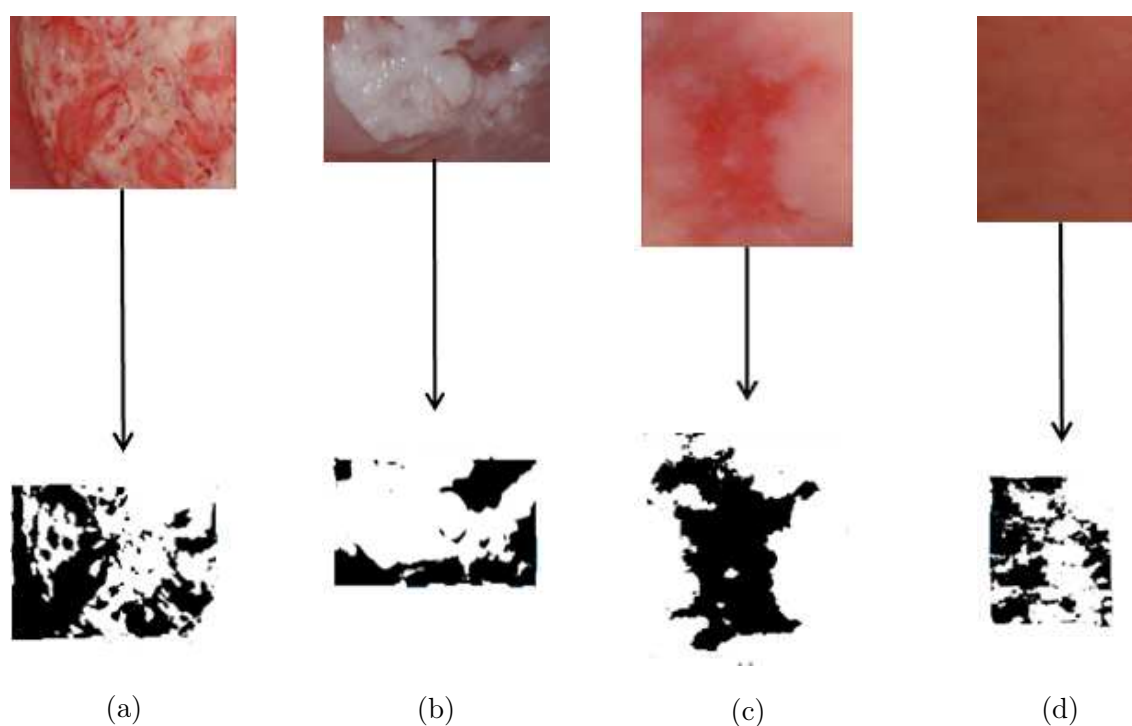


図 3.3: 白斑領域抽出例 (a) 癌 (b) 白板症 (c) 扁平苔癬, (d) 正常

3.3.3 発赤特徴量

発赤特徴量の抽出方法を図 3.4 に示す。HSV 色空間における彩度は色の鮮やかさを表すため、口腔画像において発赤領域は赤いことから彩度値は高いと考えられる。また、扁平苔癬の特徴である発赤は白斑周囲が赤くなることが判明している。そこで切り出し画像と白斑特徴量で得られた白斑領域を用いて、切り出し画像から白斑領域を除外する。そして HSV 変換を行い、彩度画像から彩度値の平均値を算出する。算出して得られた結果を発赤特徴量とする。

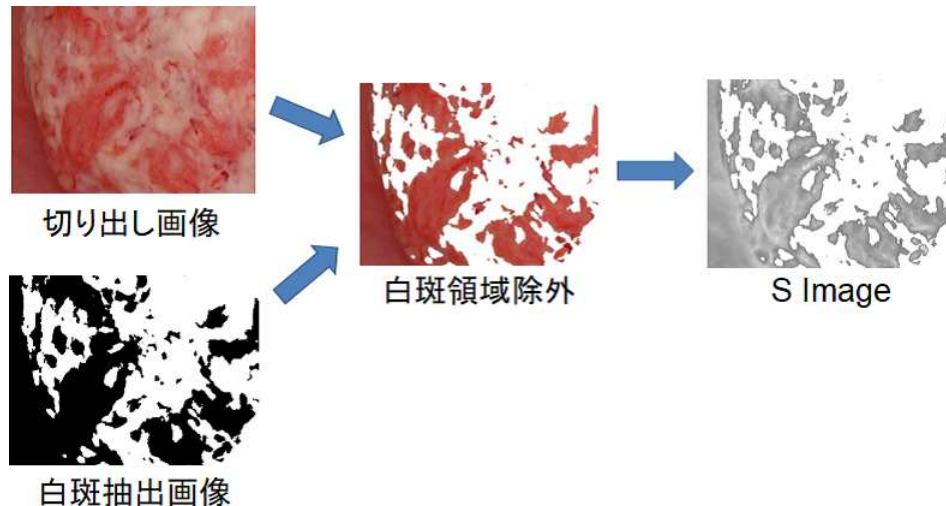


図 3.4: 発赤領域抽出例

3.3.4 隆起特徴量

口腔画像における口腔粘膜疾患の隆起は、病変の隆起による陰影として表せる。つまり、陰影のない領域と比べ陰影のある領域は輝度が低いと考えられる。したがって、輝度の低い領域の有無によって隆起の有無が判断できる。しかし、口腔粘膜疾患は白斑領域と口腔粘膜の赤領域で構成されているので、輝度値の低い領域が口腔粘膜の赤領域となり隆起の有無を判断できない。

そこで、HSV 色空間の明度に注目して隆起を求める。明度画像は光の強弱を表した画像であり、光の強さにより画像の陰影を判断できる。また、光が同じ強さの場合、色に関係なく同じ数値を示す。明度画像を用いることで白色部分と赤色部分の領域にかかわらず病変の隆起を保つことができる。

そのため明度値の勾配強度の平均値を隆起特徴量とする。勾配強度 $v(x, y)$ は以下の式で表せる。

$$v(x, y) = \sqrt{f_x^2(x, y) + f_y^2(x, y)} \quad (3.7)$$

$$f_x(x, y) = V(x + 1, y) - V(x - 1, y) \quad (3.8)$$

$$f_y(x, y) = V(x, y + 1) - V(x, y - 1) \quad (3.9)$$

ここで $f(x, y)$ は勾配ベクトルであり, $V(x, y)$ は明度関数である.

3.3.5 顆粒特徴量

扁平上皮癌の特徴として粒々としたものが広がっている状態を表す顆粒が多くみられる. 口腔画像において, 顆粒状構造は円形の小さな図形が画像中に多く含まれていることを表している. 顆粒特徴量抽出法において, 病変の色でなく構造に着目するため, HSV 色空間の明度画像を用いる. 明度画像の勾配強度を求め, 閾値に基づき 2 値化処理を行う. 図 3.4(b) は閾値を 20 に設定した 2 値化処理後の画像である. 2 値化処理後の流れを図 3.5 に示す. 2 値化画像に基づきモルフォロジー演算におけるオープニング処理を行うことで, 小物体や線などのノイズ除去を行う [32][33]. 円形構造要素の相似形を用いてサイズ密度分布を作成し, 画像における円形小図形の個数をカウントし, 総数を顆粒状特徴量とする [34].



(a)



(b)

図 3.4: 勾配強度画像の 2 値化

(a) 対象部位画像 (b) 2 値化された勾配強度画像

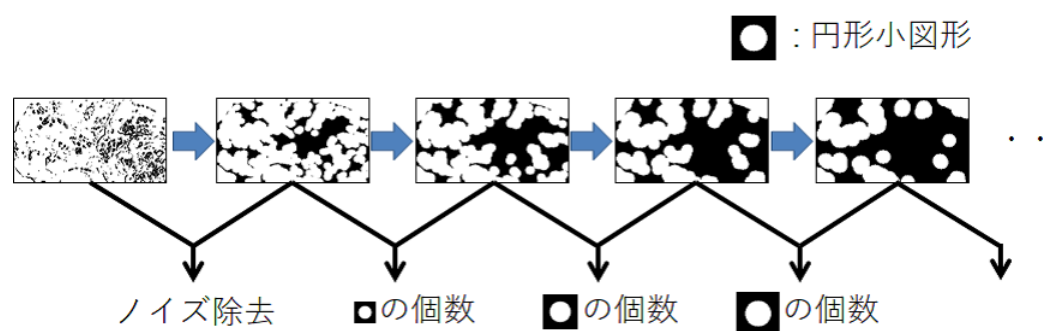


図 3.5: 顆粒特徴量の算出

3.4 従来システムを用いた識別率

3.4.1 実験概要

図 3.1 の従来システムを用いて識別率を算出した。使用したデータは 2.5, 2.6 節で述べたデータを用いる。専門医 1 名，歯科医 7 名，学生 4 名のデータを入力とするが，人毎に使用する総口腔画像枚数は異なる。使用する口腔画像枚数を表 3.2 に示す。なお本実験以降の実験で使用する口腔画像枚数は表 3.2 となる。

本実験では 4 分類の識別率を求めるとともに，偽陽性，偽陰性の識別率も算出する。これは取り扱う疾患が重度の癌と白板症，重度ではない扁平苔癬，正常と分けられるからである。偽陽性と偽陰性の定義を表 3.4 に示す。陽性は病理診断結果と識別結果が癌もしくは白板症とし，偽陽性は病理診断結果が癌もしくは白板症で識別結果は扁平苔癬もしくは正常とする。陰性は病理診断結果と識別結果が扁平苔癬もしくは正常となり，偽陰性は病理診断結果が扁平苔癬もしくは正常だが識別結果は癌もしくは白板症とする。

3.4.2 実験結果

従来システムによる識別率を表 3.4，陽性・陰性の識別率を表 3.5 に示す。表 3.4 より平均は 78.2% という結果となり，表 3.5 より，陽性率は高いが陰性率が低いことが原因で，識別率に影響していると考えられることから癌や白板症に誤識別していると考えられる。

表 3.2: 実験で使用するデータ (口腔画像枚数)

	画像枚数 [枚]
専門医	91
歯科医 2 年目	88
歯科医 5 年目	87
歯科医 15 年目 1	90
歯科医 15 年目 2	85
歯科医 30 年目 2	85
歯科医 30 年目 3	88
歯科医 30 年目 4	84
学生 2	82
学生 3	86
学生 6	88
学生 8	83

表 3.3: 偽陽性, 偽陰性の定義

識別結果 \ 病理診断結果	癌 or 白板症	扁平苔癬 or 正常
	癌 or 白板症	陽性
扁平苔癬 or 正常	偽陽性	陰性

3.5 切り出し傾向結果を考慮した従来システムの改良

表 2.5 より専門医より大きく切り出す傾向が見られた人は 11 人中 8 人であったことから, 専門医が切り出した範囲に近づけるためには切り出し画像を小さくする必要がある. そこで切り出し画像の内側に注目し, 切り出し画像とは異なる局所画像を生成できるシステム (以後, 提案システムと呼ぶ) に改良する. 提案システムを図 3.6 に示す. 局所画像を生成するシステムに改良するにあたり 2 つパラメータを設定した. 1 つは局所画像サイズ倍率パラメータ δ を設定することで, 様々な画像サイズの画像生成が可能になる. 切り出し範囲の幅 Δx , 高さ Δy に対し設定する. δ は切り出し画像の内側に注目することから $0 < \delta < 1$ の範囲内になる.

局所画像サイズ倍率パラメータ δ を設定することで局所画像の画像サイズは $\delta\Delta x$, $\delta\Delta y$ となる. また局所画像の始点は切り出し画像内でランダムに決定することから, 様々な場所に注目した局所画像を生成することができる. そしてランダムで局所画像場所を決定していることから局

表 3.4: 従来システムによる識別率

	識別率 [%]
専門医	90
歯科医 2 年目	76
歯科医 5 年目	67
歯科医 15 年目 1	72
歯科医 15 年目 2	82
歯科医 30 年目 2	84
歯科医 30 年目 3	88
歯科医 30 年目 4	68
学生 2	77
学生 3	85
学生 6	73
学生 8	76
平均	78.2

表 3.5: 従来システムによる陽性，陰性識別率

識別率 [%]	陽性	偽陽性	陰性	偽陰性
専門医	100	0	95	5
歯科医 2 年目	94	6	79	21
歯科医 5 年目	92	8	73	27
歯科医 15 年目 1	94	6	79	21
歯科医 15 年目 2	98	2	78	22
歯科医 30 年目 2	96	4	88	12
歯科医 30 年目 3	98	2	87	13
歯科医 30 年目 4	96	4	69	31
学生 2	94	6	82	18
学生 3	98	2	86	14
学生 6	96	4	71	29
学生 8	94	6	81	19
平均	95.3	4.17	80.7	19.3

所画像によって識別結果が異なる可能性がある．そのため1枚の局所画像だと，提案システムを利用する度に識別結果が変化する可能性が危惧される．そこで2つ目のパラメータとして画像枚数パラメータ N を設定することで頑健なシステムを実現する．画像枚数パラメータ $N = 1$ の場合は切り出し画像のみ用いる．それぞれの画像に対して特徴量抽出を行い，SVMを用いて識別結果を出力する．特徴量抽出および識別器に関しては従来システムと同一である．識別結果が N 個出力されるため，多数決で1番得票数が多かった疾患を出力結果とする．得票数が同数の場合は疾患の重度を考慮して癌 > 白板症 > 扁平苔癬 > 正常の条件を設定した [35]．

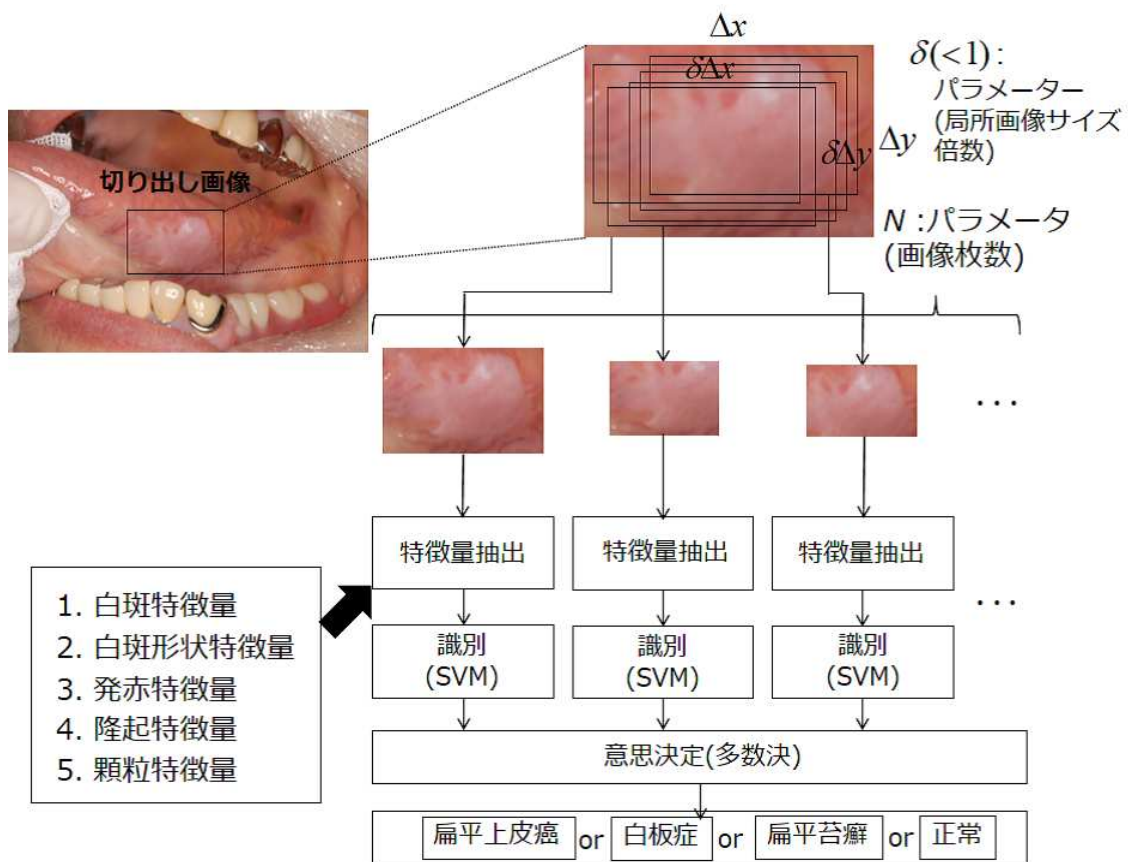


図 3.6: 局所画像生成を追加した診断支援システム (提案システム)

3.6 提案システムを用いた識別率

3.6.1 実験概要

図 3.6 の提案システムを用いて従来システムより識別率が向上するか実験を行う。使用データは 3.4 節の実験で使用したデータと同じ表 3.2 を用いる。3.4 節と同様、識別率だけでなく、偽陽性、偽陰性の識別率も調査した。

図 3.6 の診断支援システムのパラメータは従来システムの切り出し画像のみの $\delta = 1, N = 1$ および局所画像を含む ($\delta = 0.9, N = 10$), ($\delta = 0.9, N = 30$), ($\delta = 0.8, N = 10$), ($\delta = 0.8, N = 30$), ($\delta = 0.7, N = 10$), ($\delta = 0.7, N = 30$) の 7 パターンで行った。

3.6.2 実験結果

各パラメータにおける識別率を表 3.6 に、偽陽性を表 3.7 に、偽陰性を表 3.8 に示す。表 3.6 に示される値は各パラメータを設定した場合の識別率 [%] を示しており、平均識別率も示している。表 2.5 より、切り出し傾向を踏まえると、局所画像を含まない切り出し画像 ($\delta = 1, N = 1$) が一番識別率が高い人は歯科医 15 年目 2, 歯科医 30 年目 3 であった、この 2 人は切り出し傾向が専門医より小さく切り出す傾向の人である。一方で専門医より大きく切り出す傾向が見られた人は全員、切り出し画像 ($\delta = 1, N = 1$) の識別率と比べ、局所画像を含めた場合の方が識別率が高い結果となった。この結果より切り出し傾向と識別率の推移は関係しているのではないかと考えられる。また、平均識別率に注目すると一番識別率が高いパラメータは $\delta = 0.7, N = 10$ の 81.9% であり、歯科医 2 年目, 歯科医 15 年目 1, 歯科医 30 年目 2, 歯科医 30 年目 4 の 4 名の最も高い識別率を示したパラメータと同じである。しかし、それ以外の人々は識別率が一番高いパラメータが $\delta = 0.7, N = 10$ ではない。このため切り出し傾向を考慮し、影響を軽減する診断支援システムを実現するためには切り出し範囲と表 3.6 の識別率のデータを参照し、システム利用者毎に高識別率を示すパラメータ δ, N を推定する必要がある。

次に偽陽性の表 3.7 に注目すると、局所画像枚数 $N = 30$ に設定することで偽陽性率が上昇した人が多く見られた結果より、癌や白板症に対して誤識別が増加することは望ましくない。また、表 3.8 に注目すると、局所画像を含めた方が切り出し画像 ($\delta = 1, N = 1$) のみと比べて偽陰性率は減少している結果が得られた。そのため、局所画像を含んだ場合に識別率が向上した理由として、偽陰性の減少に関係していると考えられる。

また、5 章で局所画像サイズ・枚数パラメータの推定実験を行うが、評価方法の 1 つとしてパラメータ推定精度を取り扱う。そこで表 3.6 の各パラメータによる識別率を参照し、順位付けを行った。順位は 1-7 まで設定し、低いほど識別率が高いパラメータである。しかし、人によって同率の

識別率となっていることから、該当箇所は表 3.7 に示す偽陽性率の識別率を参照し、仮に偽陽性率も同じ識別率の場合は表 3.8 に示す偽陰性率を基に順位付けを行っている。ただし、偽陰性率でも同じ識別率を示している場合は同一順位とする。その結果を表 3.9 に示す。

表 3.6: 提案システムによる識別率 [%]

δ	1.0	0.9		0.8		0.7	
N	1	10	30	10	30	10	30
専門医	90	91	90	90	90	85	82
歯科医 2 年目	76	80	78	80	80	85	82
歯科医 5 年目	67	69	67	75	75	79	81
歯科医 15 年目 1	72	82	72	82	73	87	71
歯科医 15 年目 2	82	82	76	79	75	76	70
歯科医 30 年目 2	84	84	72	84	72	87	71
歯科医 30 年目 3	88	86	74	84	74	82	72
歯科医 30 年目 4	68	76	69	81	72	83	69
学生 2	77	79	78	80	78	79	77
学生 3	85	87	87	85	86	85	86
学生 6	73	74	74	76	76	74	75
学生 8	76	81	81	82	81	81	80
平均値	78.2	80.9	76.5	81.5	77.67	81.9	76.3

表 3.7: 提案システムによる偽陽性率 [%]

δ	1.0	0.9		0.8		0.7	
N	1	10	30	10	30	10	30
専門医	0	2	0	4	4	4	9
歯科医 2 年目	6	6	8	6	8	6	8
歯科医 5 年目	8	10	4	10	12	6	6
歯科医 15 年目 1	6	6	30	6	33	2	36
歯科医 15 年目 2	2	0	26	0	24	2	22
歯科医 30 年目 2	4	4	33	4	33	4	35
歯科医 30 年目 3	2	2	27	4	25	4	25
歯科医 30 年目 4	4	4	26	4	24	6	28
学生 2	6	6	8	0	4	2	2
学生 3	2	2	2	4	3	4	4
学生 6	4	4	6	4	4	4	4
学生 8	6	4	4	4	4	4	6
平均	4.17	4.17	14.5	4.17	14.83	4	15.41

表 3.8: 提案システムによる偽陰性率 [%]

δ	1.0		0.9		0.8		0.7	
N	1	10	30	10	30	10	30	
専門医	5	3	5	3	3	5	5	
歯科医 2 年目	21	16	13	16	16	11	11	
歯科医 5 年目	27	24	49	16	16	16	14	
歯科医 15 年目 1	21	11	5	13	3	8	3	
歯科医 15 年目 2	22	14	0	22	0	19	0	
歯科医 30 年目 2	12	12	0	12	0	12	0	
歯科医 30 年目 3	13	13	5	13	5	13	3	
歯科医 30 年目 4	31	17	8	14	5	14	3	
学生 2	18	18	18	21	21	21	18	
学生 3	14	9	9	9	9	6	6	
学生 6	29	26	23	20	17	26	23	
学生 8	19	14	14	14	14	19	19	
平均	19.33	14.75	12.42	14.42	9.08	14.17	8.75	

表 3.9: 提案システムによる識別率を参照した局所画像サイズ・枚数パラメータ順位

δ	1.0		0.9		0.8		0.7	
N	1	10	30	10	30	10	30	
専門医	2	1	2	4	4	6	7	
歯科医 2 年目	7	3	6	3	5	1	2	
歯科医 5 年目	7	5	6	3	4	2	1	
歯科医 15 年目 1	5	2	6	3	4	1	7	
歯科医 15 年目 2	2	1	5	3	6	4	7	
歯科医 30 年目 2	2	2	5	2	5	1	7	
歯科医 30 年目 3	1	2	6	3	5	4	7	
歯科医 30 年目 4	7	3	5	2	4	1	6	
学生 2	6	3	5	1	4	2	6	
学生 3	5	1	1	7	4	6	3	
学生 6	7	4	6	2	1	4	3	
学生 8	7	2	2	1	2	5	6	

3.7 おわりに

本章では本研究で使用する口腔粘膜疾患診断支援システムについて説明を行った。3.2 章ではこれまで研究室で開発していた従来システムについて、切り出し画像から特徴量抽出を行い SVM を用いた識別について説明を述べた。3.3 節では従来システムで用いる特徴量に関して説明を行った。使用した特徴量は白斑特徴量，白斑抽出特徴量，発赤特徴量，隆起特徴量，顆粒特徴量の 5 つである。3.4 節では，従来システムを用いた識別を行い，識別率や偽陽性率，偽陰性率に関して結果

を述べた。3.5節では従来システムを改良した説明を行った。切り出し画像に加え新たに局所画像を設定し、局所画像は切り出し傾向が専門医より大きい傾向する人が多い調査結果を踏まえ、切り出し画像の内側に注目した。局所画像サイズパラメータ σ 、画像枚数パラメータ N を設定することで局所画像を生成する提案システムに改良した。3.6節では3.5節で述べた提案システムを用いて、3.4節の実験と同様のデータを用いて局所画像サイズパラメータ σ 、局所画像枚数 N を設定し、識別率の推移を確認した。識別結果より、切り出し傾向が局所画像を含めた識別率に関係あるのではないかという結果が得られた。また、12人の平均識別率で最も高い識別率は81.9%の $\sigma = 0.7$ 、 $N = 10$ であるが、全員が $\sigma = 0.7$ 、 $N = 10$ のパラメータが最も高い識別率ではない。そのため切り出し傾向を考慮し、高識別率を示すパラメータを切り出し画像より推定を行う必要がある。また、偽陽性率に関して、 $N = 30$ の場合に急激に偽陽性率が高くなった人も見られた。偽陰性率は局所画像を含めることで低くなることが確認され、識別率が向上した要因になったと考えられる。

第4章 局所画像サイズ，枚数パラメータ δ ， N 推定方法

4.1 はじめに

本章では，局所画像サイズ，枚数パラメータ推定で用いる手法について説明する．4.2 節では，推定入力データとして使用する切り出し場所・範囲のデータに関して述べる．4.3 節では，局所画像サイズ，枚数パラメータ推定方法について述べる．4.4 節では，重回帰分析に関して述べる．4.5 節では，リッジ回帰に関して述べる．4.6 節では，Canonical Correlation Analysis(CCA) に関して述べる．4.7 節では，Regularized Canonical Correlation Analysis(RCCA) に関して述べる．4.8 節では，Canonical Correlation Analysis SOM(CCA-SOM) に関して述べる．4.9 節では，本章のまとめである．

4.2 切り出し場所・範囲におけるデータ化

切り出し範囲から得られる情報は口腔画像における始点終点の x, y 座標である．また使用する口腔画像は複数枚あることから正規化を行う必要がある．また，始点終点のデータでは，どのような切り出し方を行ったか判断が難しい，そこで始点終点より中点と切り出し範囲の縦横の長さを算出し正規化を行うことで局所画像サイズ・枚数パラメータ推定で用いる入力データを作成する．図 4.1 に切り出し範囲におけるデータ作成の過程を示す．人数は専門医 1 名，歯科医 7 名，学生 4 名の計 12 人であるが，口腔画像の選定により口腔画像によって切り出しを行った人数が変化するため，人数を p とする．口腔画像 91 枚に対して切り出し画像の始点および終点の座標を $x_{P,1-91}^S, y_{P,1-91}^S, x_{P,1-91}^E, y_{P,1-91}^E$ とし，中点 $x_{P,1-91}^C, y_{P,1-91}^C$ と縦の長さ $h_{P,1-91}$ と横の長さ $w_{P,1-91}$ を算出する．その後，各要素において切り出した人数に応じて口腔画像毎に正規化を行う．これを 91 枚の口腔画像に対して行うことで $x_{P,1-91}^{C'}, y_{P,1-91}^{C'}, h'_{P,1-91}, w'_{P,1-91}$ を求めることができる．そして，各要素人毎に正規化の平均を行い， X_P^C, Y_P^C, H_P, W_P を算出する．この正規化の平均を局所画像サイズ・枚数パラメータ推定で用いる入力データとする．

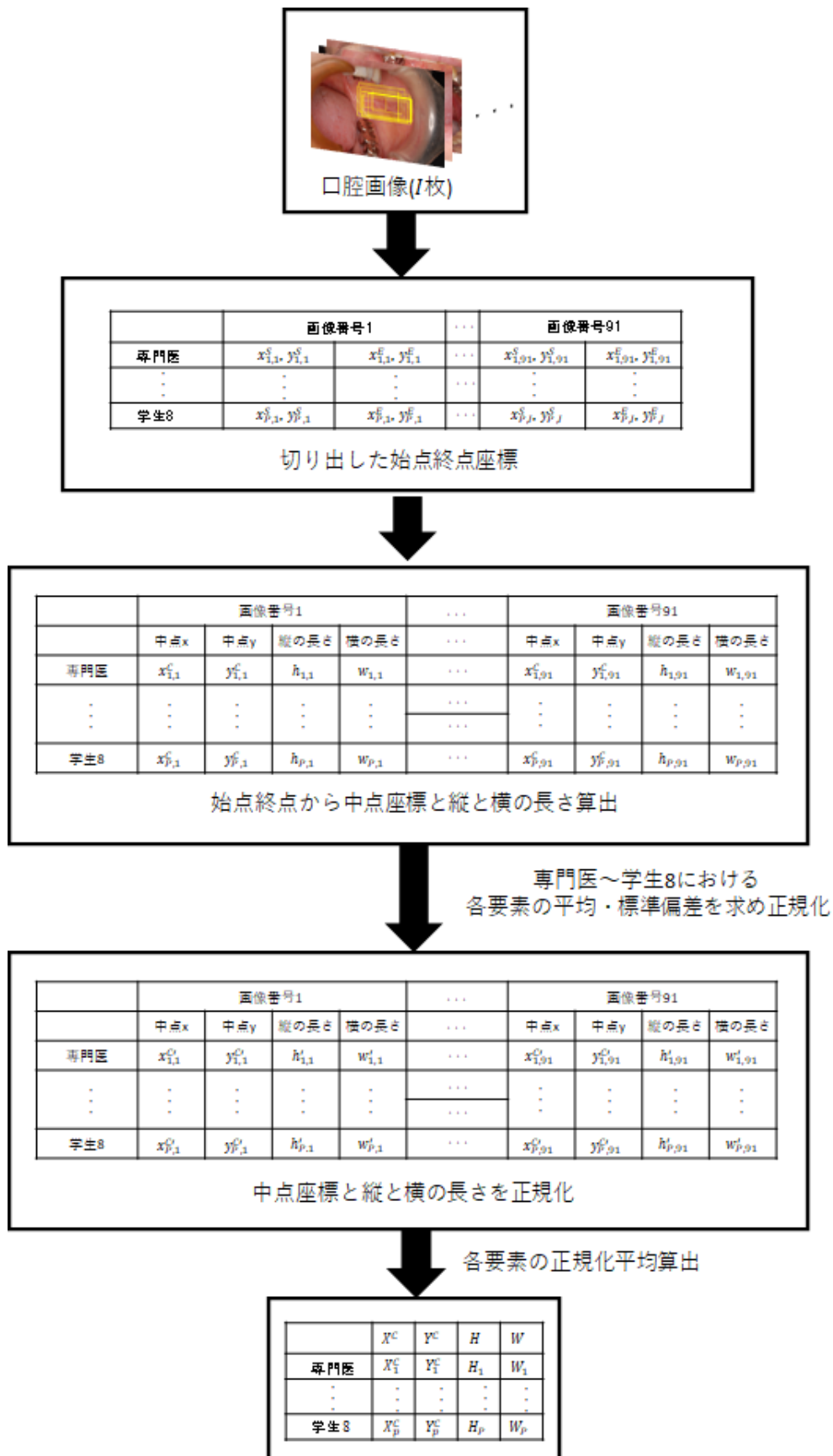


図 4.1: 切り出し場所・範囲におけるデータ化手順

4.3 局所画像サイズ、枚数パラメータ δ , N 推定方法

局所画像サイズ・枚数パラメータ δ , N 推定を行うにあたり、推定に用いる学習モデルを事前に作成する必要がある。学習モデル作成手順を図 4.2 に示す。学習モデルで使用するデータは表

4.1 の切り出し範囲データおよび表 3.6 の提案システムを用いた識別率である。今回切り出しを行った人は 12 名のため、パラメータ推定を行う人のデータを除外した 11 名のデータを用いる。切り出し範囲データおよび識別率データを入力することで学習モデルを作成していくが、局所画像サイズ・枚数パラメータの推定方法が確立されていない。そこで今回、重回帰分析、リッジ回帰、Canonical Correlation Analysis(CCA), Regularized Canonical Correlation Analysis(RCCA), Canonical Correlation Analysis SOM(CCA-SOM) を用いる。いずれかの方法が選択された後、学習モデルが生成される。

学習モデル生成後、局所画像サイズ・枚数パラメータの推定手順を、図 4.3 に示す。入力データは学習モデル作成で使用していない(パラメータ推定を行う)1 人のデータを用いる。ただし口腔画像枚数に関しては、学習モデル作成時に選択した推定方法が頑健であるかどうか検証するために交差検証を行っている。そのため、口腔画像枚数 82-91 枚からランダムで選択された 20 枚を除外した画像による切り出し範囲データを入力データとする。図 4.2 で作成した学習モデルを用いることで各局所画像サイズ・枚数パラメータに対応した識別率が算出される。その後、各識別率に関して大小比較を行い、一番高い識別率に対応する、局所画像サイズ・枚数パラメータを推定結果とする。

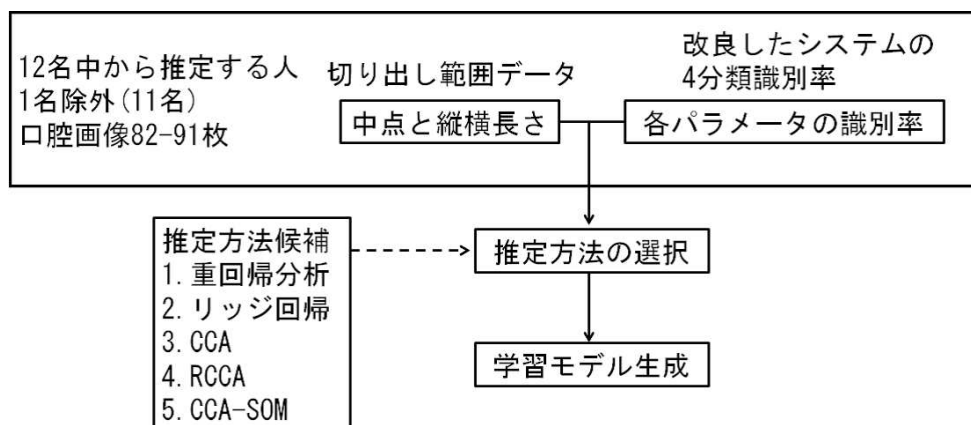


図 4.2: 局所画像サイズ・枚数パラメータ推定に用いる学習モデル作成手順

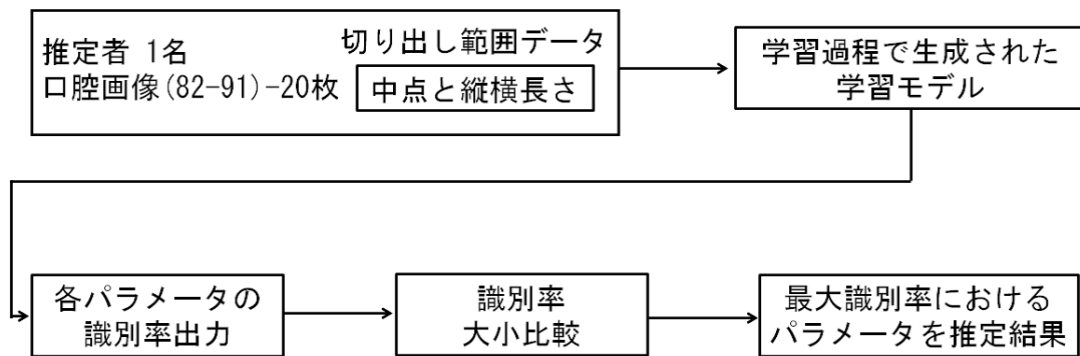


図 4.3: 局所画像サイズ・枚数パラメータ推定手順

4.4 重回帰分析

2つ以上の説明変数 $x_{1,2,\dots,N}$ と目的変数 y を用いて 5.1 式に示す重回帰モデルを想定し, 回帰母数 $\beta_{0,1,\dots,N}$ を最小二乗法により推定する分析を重回帰分析という.

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_N x_N \quad (4.1)$$

ここで残差 e を 5.2 式に, 残差平方和 S_e を 5.3 式に示す.

$$e = y_N - \hat{y} = y_N - (\hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \dots + \hat{\beta}_N x_N) \quad (4.2)$$

$$S_e = \sum_{i=1}^N e^2 = \sum_{i=1}^N y_N - (\hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \dots + \hat{\beta}_N x_N)^2 \quad (4.3)$$

S_e を最小にする $\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_N$ を求める [36].

4.5 リッジ回帰

リッジ回帰は過学習を防ぐために正則化項の L2 ノルムを導入する方法である. 重回帰分析との違いは予測と実際の値の二乗誤差を小さくするとともに複雑なモデルを作らないように β の値を小さくすることである. 5.3 式に L2 ノルムを追加すると,

$$S_e = \sum_{i=1}^N e^2 - \lambda \|\beta\|_2^2 \quad (4.4)$$

になる。 $\|\beta\|_2$ は各回帰母数の二乗和であるため、

$$\|\beta\|_2 = \sqrt{\beta_1^2 + \beta_2^2 + \dots + \beta_N^2} \quad (4.5)$$

とする。またパラメータ λ は設計者が設定するもしくはクロスバリデーションで決定するハイパーパラメータであり、値を大きくすることで正則化を大きくし、モデルの重みを小さくすることができるため、過学習を軽減することができる [37]。

4.6 Canonical Correlation Analysis

Canonical Correlation Analysis(CCA)とはそれぞれの観測データを線形変換した値間の相関係数が最大になるような線形変換のパラメータを決定する手法である。2つの測定方法によって得られた N 個の観測データ $X = x_1, x_2, \dots, x_N$, $Y = y_1, y_2, \dots, y_N$ はそれぞれベクトル \mathbf{X} , \mathbf{Y} とすることができる。また計算を簡単にするため \mathbf{X} , \mathbf{Y} の平均を 0 とする。このとき、 \mathbf{X} , \mathbf{Y} それぞれを線形変換して得られる 1 次元の値 $u(X)$, $v(Y)$ は、

$$u(\mathbf{X}) = \mathbf{a}^T \mathbf{X} \quad v(\mathbf{Y}) = \mathbf{b}^T \mathbf{Y} \quad (4.6)$$

と表すことができる。上式の \mathbf{a} , \mathbf{b} のパラメータを算出する。ただし、 \mathbf{a}^T はベクトル \mathbf{a} の転置である。これらの値間の相関係数 $\rho(\mathbf{a}, \mathbf{b})$ は、共分散を標準偏差で割った値であり次式で表すことができる。

$$\rho(\mathbf{a}, \mathbf{b}) = \frac{\mathbf{E}[u(\mathbf{X})v(\mathbf{Y})]}{\sqrt{\mathbf{E}[u(\mathbf{X})^2]}\sqrt{\mathbf{E}[v(\mathbf{Y})^2]}} \quad (4.7)$$

$\mathbf{E}[f(\mathbf{X})]$ はアンサンブル平均

$$\mathbf{E}[f(\mathbf{X})] = \frac{1}{N} \sum_{i=1}^N f(\mathbf{X}_i) \quad (4.8)$$

を表す。ここで \mathbf{X} , \mathbf{Y} の共分散行列をそれぞれ $V_{XX} = \mathbf{E}[\mathbf{X}\mathbf{X}^T]$, $V_{XY} = \mathbf{E}[\mathbf{X}\mathbf{Y}^T]$, $V_{YY} = \mathbf{E}[\mathbf{Y}\mathbf{Y}^T]$ とすると、式 5.7 は、

$$\rho(\mathbf{a}, \mathbf{b}) = \frac{\mathbf{a}^T V_{XY} \mathbf{b}}{\sqrt{\mathbf{a}^T V_{XX} \mathbf{a}} \sqrt{\mathbf{b}^T V_{YY} \mathbf{b}}} \quad (4.9)$$

になる．CCA では $\rho(\mathbf{a}, \mathbf{b})$ を最大とするような \mathbf{a} ， \mathbf{b} を求める． \mathbf{a} ， \mathbf{b} に正の数を掛けても分母と分子で打ち消すことで， $\rho(\mathbf{a}, \mathbf{b})$ の値が変動しないことから $\rho(\mathbf{a}, \mathbf{b})$ 式の分母の標準偏差を 1 と仮定すると，

$$\arg \max_{\mathbf{a}, \mathbf{b}} \mathbf{a}^T V_{XY} \mathbf{b} \quad s.t. \quad \mathbf{a}^T V_{XX} \mathbf{a} = \mathbf{b}^T V_{YY} \mathbf{b} = 1 \quad (4.10)$$

の 2 次関数の最大化問題になる．最大値問題に対してラグランジュの未定乗数法を用いることで，

$$L(\mathbf{a}, \mathbf{b}, \lambda_X, \lambda_Y) = \mathbf{a}^T V_{XY} \mathbf{b} + \lambda_a (1 - \mathbf{a}^T V_{XX} \mathbf{a}) + \lambda_b (1 - \mathbf{b}^T V_{YY} \mathbf{b}) \quad (4.11)$$

の $\mathbf{a}, \mathbf{b}, \lambda_X, \lambda_Y$ の極値問題になる．関数 L を \mathbf{a}, \mathbf{b} で偏微分を行い，0 とすることで，

$$V_{XY} \mathbf{b} = 2\lambda_a V_{XX} \mathbf{a} \quad (4.12)$$

$$V_{XY}^T \mathbf{a} = 2\lambda_b V_{YY} \mathbf{b} \quad (4.13)$$

となる．また $\lambda_X = \lambda_Y$ であることから $\lambda' = 2\lambda_X = 2\lambda_Y$ と置き換えると，式 5.12，5.13 式は

$$\begin{vmatrix} O & V_{XY} \\ V_{XY}^T & O \end{vmatrix} \begin{vmatrix} \mathbf{a} \\ \mathbf{b} \end{vmatrix} = \lambda' \begin{vmatrix} V_{XX} & O \\ O & V_{YY} \end{vmatrix} \begin{vmatrix} \mathbf{a} \\ \mathbf{b} \end{vmatrix} \quad (4.14)$$

となる．

4.7 Regularized Canonical Correlation Analysis(RCCA)

Regularized Canonical Correlation Analysis(RCCA) は CCA に正則化項 L2 ノルムを追加することである．リッジ回帰では 1 つの正則化パラメータを付与したが，RCCA では β_a と β_b の 2 つのパラメータを用いる．5.9 式の $\rho(\mathbf{a}, \mathbf{b})$ は，

$$\rho(\mathbf{a}, \mathbf{b}) = \frac{\mathbf{a}^T V_{XY} \mathbf{b}}{\sqrt{\mathbf{a}^T V_{XX} \mathbf{a} + \lambda \|\beta_a\|^2} \sqrt{\mathbf{b}^T V_{YY} \mathbf{b} + \lambda \|\beta_b\|^2}} \quad (4.15)$$

となる．また 5.14 式の固有値問題に対しても次式で表現される [39]．

$$\begin{vmatrix} O & V_{XY} \\ V_{XY}^T & O \end{vmatrix} \begin{vmatrix} \mathbf{a} \\ \mathbf{b} \end{vmatrix} = \lambda' \begin{vmatrix} V_{XX} + \lambda \mathbf{I} & O \\ O & V_{YY} + \lambda \mathbf{I} \end{vmatrix} \begin{vmatrix} \mathbf{a} \\ \mathbf{b} \end{vmatrix} \quad (4.16)$$

4.8 Canonical Correlation Analysis SOM

Canonical Correlation Analysis SOM(CCA-SOM) は 2 つの SOM を用いてそれぞれの観測データ集合を解析し，2 種類の測定方法によって得られた観測データ集合の関係性を解析する手法である．2

つの潜在区間が一致した潜在空間 z となるように Metric を変更を行う。データ $X = x_1, x_2, \dots, x_N$, $Y = y_1, y_2, \dots, y_N$ の関係性を解析する場合、推定した潜在変数が一致するデータ X に対応する Metric 行列 A およびデータ Y に対応する B と潜在変数と写像を推定を行う。潜在変数 $\zeta_{l_n^*}$ からデータ x_n への写像を $u_{l_n^*}$, 潜在変数 $\zeta_{k_n^*}$ からデータ y_n への写像を $v_{k_n^*}$ としたときの評価関数は次式で表すことができる。

$$F = \frac{1}{N} \|A u_{l_n^*} - A x_n\| + \frac{1}{N} \|B v_{k_n^*} - B y_n\|, s.t. |det A| = const, |det B| = const \quad (4.17)$$

評価関数が最小になる Metric 行列 A, B を算出することで潜在変数を一致することができる。評価関数には推定を行う行列が2つあるため、最小化を行う場合、 A を推定するため、Metric 行列 B を固定し、評価関数 F が最小になるようにする。また、 B を推定するため、Metric 行列 A を固定し、評価関数 F が最小になるようにする。Metric 行列 A および B の初期値は単位行列であり、評価関数 F が収束するまで交互に推定を行う。

CCA-SOM のアルゴリズムを以下に示す。STEP1 4 を Metric 行列が収束するまで繰り返す。

STEP1

Metric 行列を A 用いてデータ $X = (x_1, x_2, \dots, x_N)$ を SOM で学習して潜在変数を推定する。初期状態では Metric 行列 A は単位行列とする。

勝者ユニットの決定

$$l_n^* = \arg \max_l \|A u_l - A x_n\|^2 \quad (4.18)$$

事後分布の決定

$$r_{nk} = C \exp\left[-\frac{d(z_n^*, \zeta_k)^2}{2\sigma^2}\right] \quad (4.19)$$

$$\sigma = \sigma_{min} + (\sigma_{max} - \sigma_{min}) \exp\left(-\frac{t}{\tau}\right) \quad (4.20)$$

参照ベクトルの更新

$$u_k = \frac{\sum_n r_{nk} x_n}{\sum_n r_{nk}} \quad (4.21)$$

ここで z_n^* は n 番目のデータのベクトル x_n に対する勝者ユニットのインデックス、 r_{nk} は x_n に対する勝者ユニットを中心とした場合の近傍関数から求めた事後分布である。 $d(z_n^*, \zeta_k)^2$ は z_n^* と ζ_k 間のマップ距離を示す。 σ は近傍関数に用いる分散であり、学習回数 t が増えるにつれて、時定数 τ に従って変化する。学習が収束するまで繰り返し行うことで Step2 に移行する。

STEP2

Step1 で推定した潜在変数を教師データとして、データ Y の Metric 行列 B を推定する。評価関数 F の最小化を行うことで、Metric 行列 B を推定するが、Metric 行列 A を固定することで、

評価関数 F の最小化は，

$$F_b = \frac{1}{N} \|\mathbf{B}v_{k_n^*} - \mathbf{B}y_n\| \quad s.t \ |\det \mathbf{B}| = const \quad (4.22)$$

の最小化と考えることができる．ここでデータ \mathbf{Y} と写像 v_k の誤差ベクトル e_{bn} と誤差の共分散行列 \mathbf{S}_b を

$$\mathbf{S}_b = \frac{1}{N} e_{bn} e_{bn}^T \quad (4.23)$$

$$\mathbf{E} = (e_1, e_2, \dots, e_N) \quad (4.24)$$

$$e_n = y'_n - u'_{k_n^*}$$

とすることで式 3.15 は次式に変形することができる．

$$F_b = \frac{1}{N} [\mathbf{B} \mathbf{S}_b \mathbf{B}^T] \quad s.t \ |\det \mathbf{B}| = const \quad (4.25)$$

F_b にラグランジュの未定乗数法を適応することで Metric 行列は，

$$\mathbf{B} = \mathbf{S}_b^{-\frac{1}{2}} \quad (4.26)$$

と求めることができる．

STEP3

STEP1 と同様，Metric 行列を \mathbf{B} 用いてデータ $\mathbf{Y} = (\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_N)$ を SOM で学習して潜在変数を推定する．初期状態では Metric 行列 \mathbf{B} は単位行列とする．

勝者ユニットの決定

$$z_n^* = \arg \max_l \|\mathbf{B}u_l - \mathbf{B}y_n\|^2 \quad (4.27)$$

事後分布の決定

$$r_{nl} = C \exp\left[-\frac{d(z_n^*, \zeta_l)^2}{2\sigma^2}\right] \quad (4.28)$$

$$\sigma = \sigma_{min} + (\sigma_{max} - \sigma_{min}) \exp\left(-\frac{t}{\tau}\right) \quad (4.29)$$

参照ベクトルの更新

$$v_l = \frac{\sum_n r_{nl} y_n}{\sum_n l_{nl}} \quad (4.30)$$

以上を学習が収束するまで繰り返し行う．

STEP4

Step3 で推定した潜在変数を教師データとして，データ \mathbf{X} の Metric 行列 \mathbf{A} を推定する．評価関数 F の最小化を行うことで，Metric 行列 \mathbf{A} を推定するが，Metric 行列 \mathbf{B} を固定することで，評価関数 F の最小化は，

$$F_a = \frac{1}{N} \|\mathbf{A}u_{l_n^*} - \mathbf{A}x_n\| \quad s.t \ |\det \mathbf{A}| = const \quad (4.31)$$

の最小化と考えることができる．ここでデータ \mathbf{A} と写像 u_l の誤差ベクトル e_{an} と誤差の共分散行列 \mathbf{S}_a を

$$\mathbf{S}_a = \frac{1}{N} e_{an} e_{an}^T \quad (4.32)$$

$$e_n = x'_n - u'_{l_n^*}$$

とする．Step2 と同様，ラグランジュの未定乗数法を適応することで Metric 行列は，

$$\mathbf{A} = \mathbf{S}_a^{-\frac{1}{2}} \quad (4.33)$$

と求めることができる [40].

4.9 おわりに

本章では，局所画像サイズ，枚数パラメータ推定で用いる手法について説明を行った．4.2 節では，推定入力データとして使用する切り出し場所・範囲のデータ作成方法として，口腔画像枚数が複数枚あるため正規化を行う必要がある．また，始点終点の座標データを中点と縦横長さデータに変換した．4.3 節では，局所画像サイズ，枚数パラメータ推定方法について学習モデル作成と推定手順について述べた．学習モデル作成ではどの推定方法が適しているか不明であるため 6 章で 5 つの手法を試し検証する．4.4 節では，重回帰分析に関して述べた．4.5 節では，リッジ回帰について述べた．4.6 節では，Canonical Correlation Analysis(CCA) に関して述べた．4.7 節では，Regularized Canonical Correlation Analysis(RCCA) について述べた．4.8 節では，Canonical Correlation Analysis SOM(CCA-SOM) に関して述べた．

第5章 局所画像サイズ，枚数パラメータ δ ， N 推定実験

5.1 はじめに

本章では，提案システムの局所画像サイズ，枚数パラメータ δ ， N の推定方法に関する実験を行った．6.2 節では，実験概要について説明する．6.3 節では，重回帰分析を用いた場合の推定結果を示す．6.4 節では，リッジ回帰を用いた場合の推定結果を示す．6.5 節では，CCA を用いた推定結果を示す．6.6 節では，RCCA を用いた推定結果を示す．6.7 節では，CCA-SOM を用いた推定結果を示す．6.8 節では，パラメータ推定結果のまとめに関して述べる．6.9 節では，本章のまとめである．

5.2 実験概要

局所画像サイズ，枚数パラメータ δ ， N の推定においてどの手法がパラメータ推定に適しているか実験を行った．データは専門医 1 名，歯科医 7 名，学生 4 名の切り出し範囲データおよび表 3.6 の提案システムによる識別率を使用した．図 4.2 に示す学習モデル作成ではパラメータ推定を行う 1 人をあらかじめ除外した 11 名，口腔画像枚数 82-91 枚の切り出し範囲データおよび提案システムによる識別率を入力データとしている．その後，学習モデル作成で用いる推定方法は重回帰分析，リッジ回帰，CCA，RCCA，CCA-SOM の 5 つを候補とした．リッジ回帰および RCCA で使用するパラメータ λ は 10,1,0.1 で，leave-one-out による交差検証により λ を決定している．また学習モデル作成後の入力データとして，学習モデル作成で使用していない 1 人の口腔画像枚数 82-91 枚の中からランダムで選ばれた 20 枚を除外した枚数の切り出し範囲データを入力とした．ランダムで 20 枚減らした理由として，1 人当たり 10 回交差検証を行うことで推定で用いる口腔画像が変化しても切り出し傾向が特定できるか検証するためである．学習モデルより得られた各パラメータの識別率を比較し，最大識別率を示したパラメータを推定結果とする．本実験では 3 つの観点から推定方法の妥当性を検証する．以下に 3 つの観点を示す．

1. 個人の切り出し傾向に合わせた推定方法であるかどうか，人によって推定されるパラメータが異なっているか確認．

2. 推定方法が頑健であるかどうか，1人当たり 10 回試行することによって同一のパラメータを推定できているかどうか確認
3. 推定方法の精度を確認するため，パラメータ推定結果と提案システムによる識別率を順位付け (表 3.9) した結果と比較

5.3 重回帰分析による推定結果

重回帰分析によって推定されたパラメータを表 5.1 に示す。括弧内に示した数字は表 3.9 の識別率の順位付けを示している。括弧内の数字が低い値かつ頻度が多い場合は正しく推定されているといえる。表 5.1 より人によって違うパラメータを推定していたことから，使用者の切り出し傾向に考慮したパラメータが推定された。頑健さに関して重回帰分析では 12 人中 8 名が 10 回とも同じパラメータを推定していることが確認できた。しかし，学生 3 は $\delta = 1$ ， $N = 1$ が 5 回， $\delta = 0.8$ ， $N = 10$ が 5 回と，パラメータが変動したが順位に関して， $\delta = 1$ ， $N = 1$ が 5 位， $\delta = 0.8$ ， $N = 10$ が 7 位であることから正しく推定ができなかったと考えられる。また，推定精度では，歯科医 15 年目 1 の平均順位が 4.3，30 年目 4，学生 2，3 の平均順位が 6 といった識別率が低いパラメータが推定された。

表 5.1: 局所画像サイズ，枚数パラメータ δ ， N 推定結果 (重回帰分析)

δ	1.0	0.9		0.8		0.7		平均順位
N	1	10	30	10	30	10	30	
専門医	0(2)	10(1)	0(2)	0(4)	0(4)	0(6)	0(7)	1
歯科医 2 年目	0(7)	0(3)	0(6)	0(3)	0(5)	10(1)	0(2)	1
歯科医 5 年目	0(7)	0(5)	0(6)	0(3)	0(4)	10(2)	0(1)	2
歯科医 15 年目 1	0(5)	0(2)	0(6)	0(3)	9(4)	0(1)	1(7)	4.3
歯科医 15 年目 2	0(2)	10(1)	0(5)	0(3)	0(6)	0(4)	0(7)	1
歯科医 30 年目 2	3(2)	1(2)	0(5)	6(2)	0(5)	0(1)	0(7)	2
歯科医 30 年目 3	10(1)	0(2)	0(6)	0(3)	0(5)	0(4)	0(7)	1
歯科医 30 年目 4	0(7)	0(3)	0(5)	0(2)	0(4)	0(1)	10(6)	6
学生 2	10(6)	0(3)	0(5)	0(1)	0(4)	0(2)	0(6)	6
学生 3	5(5)	0(1)	0(1)	5(7)	0(4)	0(6)	0(3)	6
学生 6	0(7)	10(4)	0(6)	0(2)	0(1)	0(4)	0(3)	4
学生 8	0(7)	9(2)	0(2)	1(1)	0(2)	0(5)	0(6)	1.9

5.4 リッジ回帰による推定結果

リッジ回帰によって推定されたパラメータを表 5.2 に示す。重回帰分析同様、括弧内に示した数字は表 4.2 で得られた識別率の順位付け結果を示している。重回帰分析と同じ、人毎に異なるパラメータが推定されていることから、使用者の切り出し傾向を考慮したパラメータが推定された。頑健さに関して重回帰分析では 12 人中 10 名が 10 回とも同じパラメータを推定していることが確認できた、また、推定精度では学生 3 の平均順位が 6.1、学生 6 の平均順位が 4、学生 8 の平均順位が 4.2 と識別率が低いパラメータが推定されているが、歯科医に注目すると、重回帰分析と比べ識別率が高いパラメータを推定していることが確認できた。

表 5.2: 局所画像サイズ, 枚数パラメータ δ , N 推定結果 (リッジ回帰)

δ	1.0	0.9		0.8		0.7		平均順位
N	1	10	30	10	30	10	30	
専門医	0(2)	10(1)	0(2)	0(4)	0(4)	0(6)	0(7)	1
歯科医 2 年目	0(7)	0(3)	0(6)	0(3)	0(5)	10(1)	0(2)	1
歯科医 5 年目	0(7)	0(5)	0(6)	0(3)	0(4)	10(2)	0(1)	2
歯科医 15 年目 1	0(5)	0(2)	0(6)	0(3)	0(4)	10(1)	0(7)	1
歯科医 15 年目 2	10(2)	0(1)	0(5)	0(3)	0(6)	0(4)	0(7)	1
歯科医 30 年目 2	0(2)	0(2)	0(5)	10(2)	0(5)	0(1)	0(7)	2
歯科医 30 年目 3	10(1)	0(2)	0(6)	0(3)	0(5)	0(4)	0(7)	1
歯科医 30 年目 4	0(7)	0(3)	0(5)	0(2)	0(4)	10(1)	0(6)	1
学生 2	0(6)	10(3)	0(5)	0(1)	0(4)	0(2)	0(6)	3
学生 3	0(5)	0(1)	0(1)	1(7)	0(4)	9(6)	0(3)	6.1
学生 6	0(7)	0(4)	0(6)	0(2)	0(1)	10(4)	0(3)	4
学生 8	0(7)	2(2)	0(2)	0(1)	0(2)	8(5)	0(6)	4.2

5.5 CCA による推定結果

CCA によって推定されたパラメータを表 5.3 に示す．重回帰分析同様，括弧内に示した数字は表 4.2 で得られた識別率の順位付け結果を示している．使用者の切り出し傾向を考慮しているかに関して，人毎に違うパラメータを推定していることを確認することができた．頑健さに関して重回帰分析では 12 人中 11 名が 10 回とも同じパラメータを推定していることが確認できた，しかし，歯科医 30 年目 4 は $\delta = 0.8$ ， $N = 10$ が 5 回， $\delta = 0.7$ ， $N = 10$ が 5 回であるが，順位に関しては $\delta = 0.8$ ， $N = 10$ が 2 位， $\delta = 0.7$ ， $N = 10$ が 1 位という結果のため，正しく推定できていたと考えられる．また，推定精度に関して歯科医 30 年目 2 の平均順位が 1，歯科医 30 年目 4 の平均順位が 1.5，学生 2 の平均順位が 1 と高識別率を実現するパラメータが推定されているが，3 人以外は順位が 3-6 位の識別率を示すパラメータを推定していた．

表 5.3: 局所画像サイズ，枚数パラメータ δ ， N 推定結果 (CCA)

δ	1.0	0.9		0.8		0.7		平均順位
N	1	10	30	10	30	10	30	
専門医	0(2)	0(1)	0(2)	0(4)	0(4)	10(6)	0(7)	6
歯科医 2 年目	0(7)	0(3)	10(6)	0(3)	0(5)	0(1)	0(2)	6
歯科医 5 年目	0(7)	10(5)	0(6)	0(3)	0(4)	0(2)	0(1)	5
歯科医 15 年目 1	0(5)	0(2)	0(6)	10(3)	0(4)	0(1)	0(7)	3
歯科医 15 年目 2	0(2)	0(1)	0(5)	10(3)	0(6)	0(4)	0(7)	3
歯科医 30 年目 2	0(2)	0(2)	0(5)	0(2)	0(5)	10(1)	0(7)	1
歯科医 30 年目 3	0(1)	0(2)	0(6)	0(3)	0(5)	10(4)	0(7)	4
歯科医 30 年目 4	0(7)	0(3)	0(5)	5(2)	0(4)	5(1)	0(6)	1.5
学生 2	0(6)	0(3)	0(5)	10(1)	0(4)	0(2)	0(6)	1
学生 3	0(5)	0(1)	0(1)	0(7)	0(4)	10(6)	0(3)	6
学生 6	0(7)	0(4)	0(6)	0(2)	0(1)	10(4)	0(3)	4
学生 8	0(7)	0(2)	0(2)	0(1)	0(2)	10(5)	0(6)	5

5.6 RCCA による推定結果

RCCA によって推定されたパラメータを表 5.4 に示す。重回帰分析同様、括弧内に示した数字は表 4.2 で得られた識別率の順位付け結果を示している。RCCA の結果は、12 人全員 $\delta = 0.7$, $N = 10$ のパラメータのみ推定した結果となった。また、 $\delta = 0.7$, $N = 10$ は提案システムの識別率 (表 3.6) において 12 人の平均識別率が一番高かった識別率を示したパラメータと同じである。RCCA は CCA に正則化を導入した手法であるため、学習モデル作成で用いた 2 つの観測データ (切り出し範囲データおよび識別率) のパラメータ \mathbf{a} , \mathbf{b} を比較した。CCA によるパラメータ \mathbf{a} の値を 5.5, RCCA を 5.6 に示す。また、CCA によるパラメータ \mathbf{b} の値を 5.7, RCCA を 5.8 に示す。表中に記載している人は推定を行う人である。切り出し範囲データに関して CCA と RCCA を比較すると、中点 y , 縦横の長さの分散に大きな変化が見られる。また識別率データに関して CCA と RCCA を比較すると、切り出し範囲データ同様、分散に大きな変化が見られる。 \mathbf{a} , \mathbf{b} で共通している点は RCCA の方が分散が少ないということである。分散は切り出し傾向の特徴を示しているため、分散が低いと推定において同じパラメータしか選択されないことになる。この原因として考えられるのは正則化である。正則化を導入することで切り出し傾向の特徴が吸収されることから、切り出し傾向を考慮していない推定方法であると考えられる。

表 5.4: 局所画像サイズ, 枚数パラメータ δ , N 推定結果 (RCCA)

δ	1.0	0.9		0.8		0.7		平均順位
N	1	10	30	10	30	10	30	
専門医	0(2)	0(1)	0(2)	0(4)	0(4)	10(6)	0(7)	6
歯科医 2 年目	0(7)	0(3)	0(6)	0(3)	0(5)	10(1)	0(2)	1
歯科医 5 年目	0(7)	0(5)	0(6)	0(3)	0(4)	10(2)	0(1)	2
歯科医 15 年目 1	0(5)	0(2)	0(6)	0(3)	0(4)	10(1)	0(7)	1
歯科医 15 年目 2	0(2)	0(1)	0(5)	0(3)	0(6)	10(4)	0(7)	4
歯科医 30 年目 2	0(2)	0(2)	0(5)	0(2)	0(5)	10(1)	0(7)	1
歯科医 30 年目 3	0(1)	0(2)	0(6)	0(3)	0(5)	10(4)	0(7)	4
歯科医 30 年目 4	0(7)	0(3)	0(5)	0(2)	0(4)	10(1)	0(6)	1
学生 2	0(6)	0(3)	0(5)	0(1)	0(4)	10(2)	0(6)	2
学生 3	0(5)	0(1)	0(1)	0(7)	0(4)	10(6)	0(3)	6
学生 6	0(7)	0(4)	0(6)	0(2)	0(1)	10(4)	0(3)	4
学生 8	0(7)	0(2)	0(2)	0(1)	0(2)	10(5)	0(6)	5

表 5.5: 学習モデル作成時に用いた切り出し範囲データのパラメータ \mathbf{a} (CCA)

	中点 x	中点 y	縦の長さ	横の長さ
専門医	0.765	0.067	-0.403	-0.497
歯科医 2 年目	-0.098	-0.235	0.851	-0.459
歯科医 5 年目	-0.139	-0.23	0.963	-0.018
歯科医 15 年目 1	-0.254	0.384	0.444	0.769
歯科医 15 年目 2	-0.143	-0.196	0.93	-0.309
歯科医 30 年目 2	-0.12	-0.139	0.8	-0.571
歯科医 30 年目 3	-0.048	-0.115	0.881	-0.456
歯科医 30 年目 4	-0.169	0.139	-0.685	0.695
学生 2	0.061	-0.098	-0.34	0.933
学生 3	0.138	0.096	-0.34	0.933
学生 6	-0.161	-0.124	0.941	0.269
学生 8	0.016	0.034	0.899	-0.436
分散	0.066	0.031	0.41	0.33

表 5.6: 学習モデル作成時に用いた切り出し範囲データのパラメータ \mathbf{a} (RCCA)

	中点 x	中点 y	縦の長さ	横の長さ
専門医	-0.381	-0.565	-0.005	-0.023
歯科医 2 年目	-0.4	-0.564	-0.007	-0.012
歯科医 5 年目	-0.433	-0.523	-0.141	0.091
歯科医 15 年目 1	-0.389	-0.549	-0.014	-0.031
歯科医 15 年目 2	-0.323	-0.351	-0.058	-0.052
歯科医 30 年目 2	-0.417	-0.552	-0.023	-0.043
歯科医 30 年目 3	-0.438	-0.53	0.002	-0.053
歯科医 30 年目 4	-0.469	-0.457	-0.102	0.028
学生 2	-0.045	-0.499	0.002	-0.06
学生 3	-0.436	-0.515	-0.025	-0.039
学生 6	-0.389	-0.566	-0.09	0.051
学生 8	-0.503	-0.449	-0.026	-0.04
分散	0.012	0.004	0.002	0.002

表 5.7: 学習モデル作成時に用いた識別率データのパラメータ \mathbf{b} (CCA)

δ	1.0	0.9		0.8		0.7	
N	1	10	30	10	30	10	30
専門医	-0.271	0.357	0.068	0.062	-0.64	-0.181	0.59
歯科医 2 年目	0.209	-0.33	0.565	-0.238	0.327	0.193	-0.572
歯科医 5 年目	0.163	-0.243	0.816	0.0285	-0.348	0.041	-0.354
歯科医 15 年目 1	-0.132	0.751	0.328	-0.466	-0.235	-0.12	-0.154
歯科医 15 年目 2	0.305	-0.017	0.79	0.125	-0.261	-0.318	-0.312
歯科医 30 年目 2	-0.291	0.547	-0.042	-0.209	0.582	-0.19	-0.443
歯科医 30 年目 3	-0.058	0.487	0.576	-0.165	0.133	-0.354	-0.508
歯科医 30 年目 4	-0.059	-0.014	0.288	0.386	-0.724	-0.274	0.406
学生 2	0.38	-0.481	0.532	0.33	-0.458	-0.143	-0.041
学生 3	0.425	-0.149	0.278	0.241	-0.691	-0.289	0.319
学生 6	0.48	-0.35	0.749	-0.075	-0.163	-0.022	-0.402
学生 8	0.214	0.392	-0.130	-0.562	0.531	0.021	-0.431
分散	0.066	0.154	0.095	0.082	0.193	0.024	0.141

5.7 CCA-SOM による推定結果

5.7.1 推定方法

CCA-SOM による推定方法を図 5.1 に示す。11 名の切り出し範囲のデータと識別率データを入力とし、CCA-SOM を用いることで切り出し範囲、識別率マップを生成される。マップにはユニット毎に色分けがされており、切り出し範囲マップと識別率マップにおける学習データで使用した人の勝者位置は同一場所になる。学習データで使用した全員に対して、一致していることを確認した後、推定したい人の切り出し範囲データを入力することで座標、識別率マップそれぞれに勝者位置の決定する。その後、識別率マップに注目して、パラメータ δ , N の識別率が可視化されている Component Map を参照する。それぞれの Component Map における勝者位置の値を比較することで、最大識別率を示した Component Map における δ , N を推定されたパラメータとする。

5.7.2 推定結果

CCA-SOM によって推定されたパラメータを表 5.9 に示す。重回帰分析同様括弧内に示した数字は表 4.2 で得られた識別率の順位付け結果を示している。CCA-SOM の結果として、どの人も同じ値のパラメータが推定されることはなく、切り出し傾向に考慮したパラメータが推定されていたが、1 人当たり 10 回の試行回数で推定されるパラメータに一貫性がない。例として歯科医 2 年目は $\delta = 1.0$, $N = 1$ が 5 回, $\delta = 0.8$, $N = 10$ が 4 回, $\delta = 0.8$, $N = 30$ が 1 回であることから推定で用いた口腔画像によって推定結果が変化すると考えられる。また順位に関しても

表 5.8: 学習モデル作成時に用いた識別率データのパラメータ b (RCCA)

δ	1.0		0.9		0.8		0.7	
N	1	10	30	10	30	10	30	
専門医	-0.086	0.024	0.062	0.081	0.06	0.104	0.022	
歯科医 2 年目	-0.108	0.026	0.099	0.074	0.085	0.087	0.003	
歯科医 5 年目	-0.094	0.023	0.076	0.09	0.096	0.07	0.022	
歯科医 15 年目 1	-0.069	0.023	0.076	0.066	0.072	0.075	0.026	
歯科医 15 年目 2	0.011	0.029	0.033	0.042	0.044	0.055	0.05	
歯科医 30 年目 2	-0.051	0.034	0.06	0.077	0.055	0.082	0.007	
歯科医 30 年目 3	-0.026	0.041	0.035	0.083	0.039	0.084	0.008	
歯科医 30 年目 4	-0.01	-0.036	0.061	0.003	0.123	-0.013	0.148	
学生 2	-0.049	0.039	0.073	0.063	0.066	0.066	0.008	
学生 3	-0.063	0.02	0.057	0.074	0.066	0.082	0.033	
学生 6	-0.068	0.033	0.076	0.069	0.072	0.077	0.005	
学生 8	-0.08	0.019	0.062	0.087	0.078	0.089	0.011	
分散	0.001	0.0004	0.0003	0.0005	0.0005	0.0008	0.001	

$\delta = 1.0$, $N = 1$ が 7 位 (最下位) と $\delta = 0.8$, $N = 10$ が 3 位ということから入力する口腔画像に依存する可能性が高い. これにより, 推定方法が頑健であるかにおいて, 他の手法 (重回帰 8 人, リッジ回帰 10 人, CCA 11 人, RCCA 12 人) と比較すると CCA-SOM は 6 名と少ない.

表 5.9: 局所画像サイズ, 枚数パラメータ δ , N 推定結果 (CCA-SOM)

δ	1.0		0.9		0.8		0.7		平均順位
N	1	10	30	10	30	10	30		
専門医	0(2)	0(1)	10(2)	0(4)	0(4)	0(6)	0(7)	2	
歯科医 2 年目	5(7)	0(3)	0(6)	4(3)	1(5)	0(1)	0(2)	5.2	
歯科医 5 年目	0(7)	0(5)	0(6)	0(3)	0(4)	10(2)	0(1)	2	
歯科医 15 年目 1	0(5)	0(2)	0(6)	10(3)	0(4)	0(1)	0(7)	3	
歯科医 15 年目 2	8(2)	0(1)	0(5)	0(3)	0(6)	2(4)	0(7)	2.4	
歯科医 30 年目 2	0(2)	6(2)	0(5)	2(2)	0(5)	2(1)	0(7)	1.2	
歯科医 30 年目 3	0(1)	10(2)	0(6)	0(3)	0(5)	0(4)	0(7)	2	
歯科医 30 年目 4	0(7)	0(3)	0(5)	10(2)	0(4)	0(1)	0(6)	2	
学生 2	0(6)	10(3)	0(5)	0(1)	0(4)	0(2)	0(6)	3	
学生 3	6(5)	0(1)	0(1)	1(7)	0(4)	3(6)	0(3)	5.5	
学生 6	0(7)	1(4)	0(6)	0(2)	0(1)	9(4)	0(3)	4	
学生 8	0(7)	0(2)	0(2)	1(1)	0(2)	9(5)	0(6)	4.6	

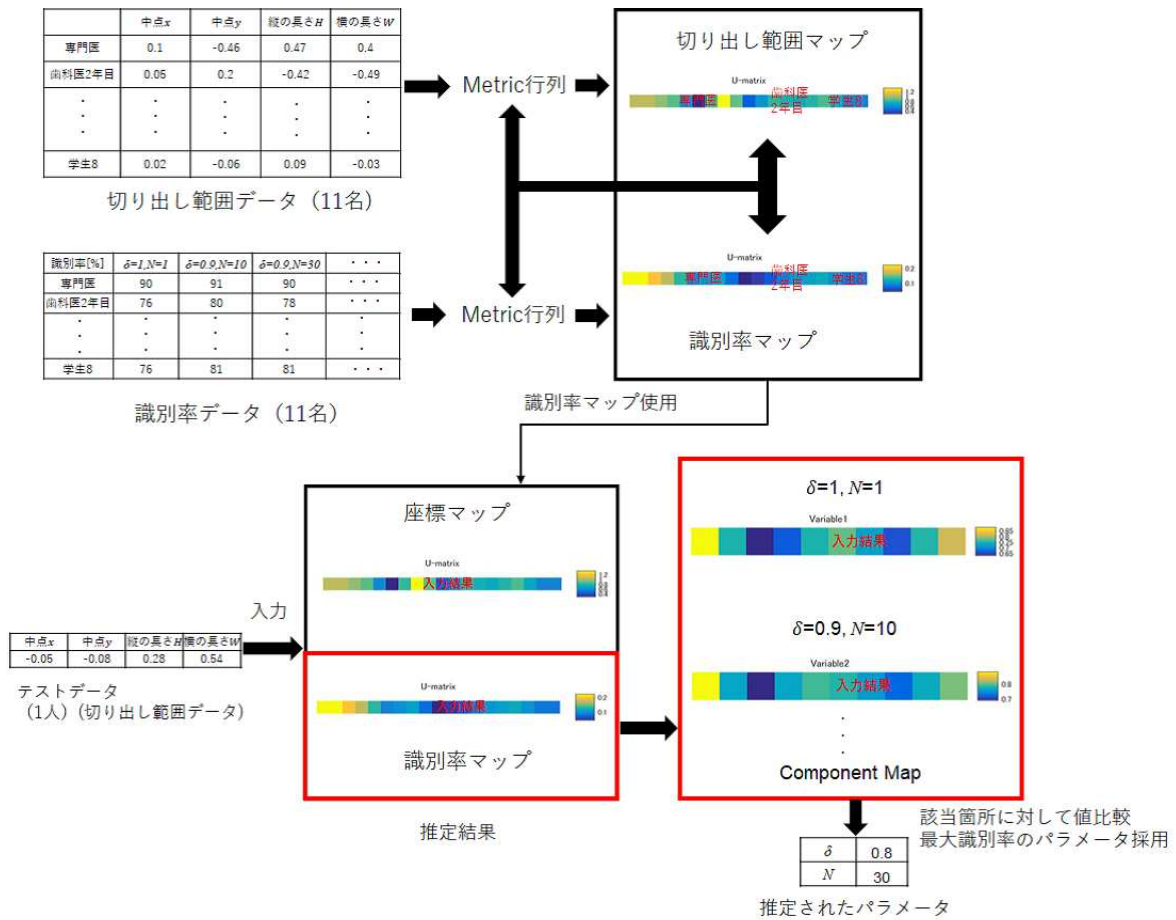


図 5.1: CCA-SOM を用いた局所画像サイズ・枚数パラメータ δ, N 推定方法

5.8 5手法によるパラメータ推定のまとめ

重回帰分析, リッジ回帰, CCA, RCCA, CCA-SOM によるパラメータ推定のまとめについて述べる. 識別率の順位付け結果をまとめた表を表 6.8 に示す. ここでは 10 回試行の平均を示しており, また推定精度を比較するために, 12 人の平均順位を算出している. この中で 1 番推定精度が高いのは平均順位 2.3 のリッジ回帰である. この理由として, 重回帰分析と比べ正則化を行うことで過学習を軽減し, 重回帰分析ではパラメータ推定が難しいと考えられる人に対しても高識別率なパラメータを推定することが可能になったのではないかと考えられる. また, CCA, RCCA についても正則化を行うことで平均順位が 3.8 から 3.1 と向上を確認できたが, 全ての人に対して $\delta = 0.7, N = 10$ のパラメータを推定したことから使用者の切り出し傾向を考慮したパラメータ推定ができていないと考えられる. また, CCA-SOM についても, 表 6.8 では確認できないが推定で用いる口腔画像に依存する可能性が考えられることから, パラメータ推定方法に適してい

ないのではないかと考える。

個人の切り出し傾向に合わせた推定が可能であるか，推定方法が頑健であるか，推定精度の3つの点を踏まえた結果，リッジ回帰が局所画像サイズ・枚数パラメータ δ , N の推定に適していると考えられる。

識別率に関し，個人の切り出し傾向を考慮した局所画像サイズ・枚数パラメータ推定を導入することで従来システムとの比較を表 3.6 の識別率結果を参照して行った。比較対象を以下に示す。

1. 従来システムを用いた 12 人の平均識別率 ($\delta = 1$, $N = 1$)
2. 提案システムを用いた 12 人の平均最大識別率 ($\delta = 0.7$, $N = 10$)
3. リッジ回帰を用いた 1 人当たり 10 回の推定されたパラメータを提案システムに適応した場合の 12 人平均識別率
4. 提案システムを用いた 12 人の最大識別率に注目した平均識別率

なお個人の切り出し傾向を考慮したものは 3 のみであり，1, 2, 4 に関しては個人の切り出し傾向を考慮していない。結果を図 5.2 に示す。従来システムと提案システム ($\delta = 0.7$, $N = 10$) に改良することで 3.7% 向上した。また提案システム ($\delta = 0.7$, $N = 10$) とリッジ回帰を用いて個人の切り出し傾向を考慮した場合の識別率を比較すると 1.3% 向上した。これは従来システムを比較すると 5% 向上した。

表 5.10: 5 手法による識別率の平均順位

	重回帰分析	リッジ回帰	CCA	RCCA	CCA-SOM
専門医	1	1	6	6	2
歯科医 2 年目	1	1	6	1	2
歯科医 5 年目	2	2	5	2	2
歯科医 15 年目 1	4.3	1	3	1	3
歯科医 15 年目 2	1	1	3	4	2.4
歯科医 30 年目 2	2	2	1	1	1.2
歯科医 30 年目 3	1	1	4	4	2
歯科医 30 年目 4	6	1	1.5	1	2
学生 2	6	3	1	2	3
学生 3	6	6.1	6	6	5.5
学生 6	4	4	4	4	4
学生 8	1.9	4.2	5	5	4.6
平均	3	2.3	3.8	3.1	3.1

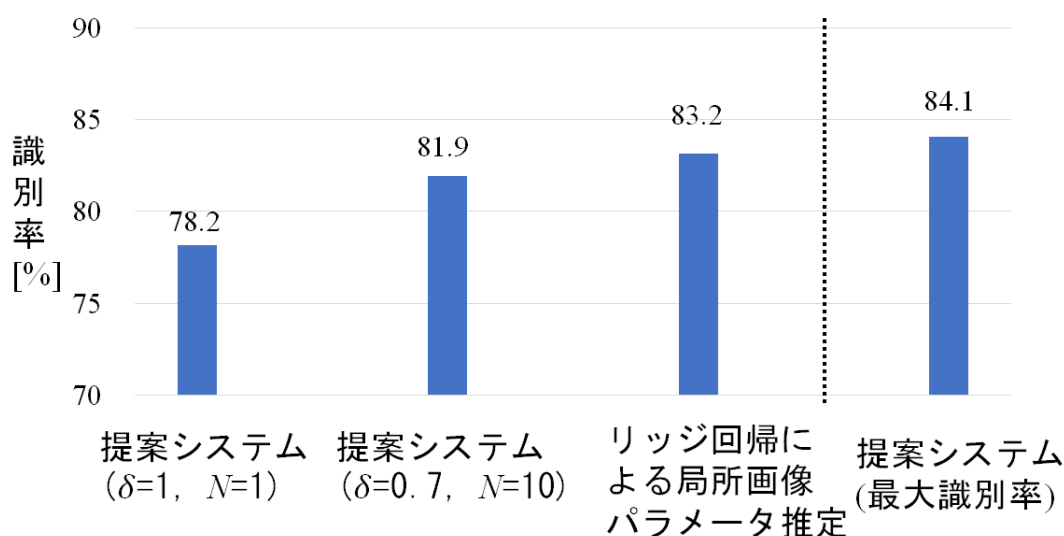


図 5.2: 使用データ 12 人, 口腔画像 81-92 枚における平均識別率の比較

5.9 おわりに

本章では切り出し位置・範囲と識別率から局所画像サイズ, 枚数パラメータ推定実験について述べた. 5.2-5.7 節では重回帰分析, リッジ回帰, CCA, RCCA, CCA-SOM によるパラメータ推定結果について述べた. 5.8 節では 5.2-5.7 節の実験結果のまとめについて説明した. その結果, リッジ回帰が局所画像サイズ, 枚数パラメータ推定に適している手法であると考えられる. この理由として, 正則化を行うことで過学習を軽減し, 推定精度が向上したと考えられる. 重回帰分析とリッジ回帰, CCA と RCCA では正則化を行うことでパラメータ推定精度は向上したが, RCCA はどの人に対しても, $\delta=0.7, N=10$ の 1 種類のパラメータしか推定されないことから, 正則化を導入することで個人の切り出し傾向の特徴が失われてしまったと考えられる. また, CCA-SOM に関しては同じ人が切り出している口腔画像によって推定されるパラメータが異なっていることから, 個人の切り出し傾向よりも推定で用いる口腔画像に依存していると考えられる. また, 識別率に関してリッジ回帰を推定方法に用いることで局所画像サイズ, 枚数パラメータ推定を行うことで従来システムより 5% の向上を確認することができた.

第6章 リッジ回帰を用いた切り出し傾向特定に必要な画像枚数の調査

6.1 はじめに

本章では、4, 5章の局所画像サイズ、枚数パラメータ推定の推定手順 (図 4.3) で、口腔画像 (83-91枚)-20枚のデータを用いてパラメータ推定を行ったが、新規の歯科医が診断支援システムを使用した場合、初回使用前に切り出し傾向を特定するために専門医の切り出し範囲が描画されている本実験で使用した口腔画像に切り出し範囲の指定を行う必要がある。そのため、約 60枚の口腔画像に対して切り出しを行ってもらう過程は手間を要すると考えられる。そこで切り出し傾向が特定できるための口腔画像枚数が最低何枚必要であるか調査を行った。本実験ではリッジ回帰を推定方法とし、結果について述べる。6.2節では、実験条件について説明する。6.3節では、実験結果について述べる。6.4節では、本章のまとめである。

6.2 実験条件

本実験では6.2節で述べた実験条件と同じである。専門医1名、歯科医7名、学生4名を対象に行った。使用データは表 4.1の切り出し範囲データおよび表 3.6の提案システムを用いた識別率である。また、1人あたり10回の試行を行うことで、頑健な推定が実現できているか確認する。図 4.2の学習モデル作成手順に変更はないが推定方法はリッジ回帰を用いる。学習モデル作成後、図 4.3の推定手順に沿うが、入力データに変更がある。推定を行う人1名の口腔画像で5章の実験では(82-91枚)-20枚を入力していたが、口腔画像枚数を70,60,50,40,30,20,10枚に設定する。ただし推定する人の識別率データは用いない。この切り出し範囲データを学習モデルに入力することで、各パラメータに対応した識別率が算出される。各パラメータによる識別率を比較し、最大の識別率を示したパラメータを推定結果とすることで、局所画像サイズ、枚数パラメータの推移を検証した。

6.3 実験結果

口腔画像 70, 60, 50, 40, 30, 20, 10 枚の場合の推定結果を表 6.1-6.6 に示す。括弧内に示した数字は表 4.2 で得られた識別率の順位付け結果を示している。また各口腔画像枚数の平均順位をまとめた表を表 6.7 に示す。表 6.7 の平均に注目すると、口腔画像が 70 枚の時に 2.2 位、10 枚の時に 2.5 位と大きく変化は見られなかった。枚数を減らすことで影響を及ぼした人は専門医と学生 6 の 2 名であることから多くの人は口腔画像枚数を変化しても影響はみられなかった。そこで表 6.1-6.6 に注目し、推定が頑健であるかに注目した。口腔画像枚数が 70 枚の場合歯科医 30 年目 2 以外は 10 回中 10 回同じパラメータを推定した。なお歯科医 30 年目 2 に関して $\delta=0.8, N=10$ で 2 番目に高識別のパラメータが 9 回、 $\delta=0.7, N=10$ で 1 番の高識別パラメータが 1 回と良好な推定精度であると考えられる。

次に口腔画像枚数 60 枚では歯科医 15 年目 2 と歯科医 30 年目 2 の 2 名以外 10 回中 10 回同じパラメータを選択している。また歯科医 15 年目 2 と歯科医 30 年目 2 に関して、70 枚の歯科医 30 年目 2 と同様の傾向が見られた。

口腔画像枚数 50 枚の場合、学生 6、学生 8 の 2 名以外 10 回中 10 回同じパラメータを選択している。学生 6 は $\delta=0.9, N=10$ と $\delta=0.7, N=10$ の 2 パターンのパラメータを推定しているが、どちらも 4 番目に高識別のパラメータを推定しているため問題はないと考えられる。学生 8 に関して、 $\delta=0.9, N=10$ の 2 番目に識別率が高いパラメータを 9 回、 $\delta=0.7, N=10$ の 5 番目に高い識別率パラメータを 1 回推定したことから、口腔枚数 50-60 枚で特定の人の切り出し傾向を間違える可能性が僅かだがあると考えられる。

口腔画像枚数 40 枚の場合では、歯科医 30 年目 2、学生 3、学生 6 の 3 名以外 10 回中 10 回同じパラメータを選択している。歯科医 30 年目 2 と学生 3 に関しては口腔画像枚数 70 枚の歯科医 30 年目 2 と同様の傾向であり、学生 6 に関して口腔画像枚数 50 枚の学生 6 と同じ傾向であった。

口腔画像枚数 30 枚では、専門医、歯科医 15 年目 2、30 年目 2、学生 3、6、8 が 10 回中 10 回同じパラメータを選択していなかった。これまで口腔画像枚数 70-40 枚に関しては 1-3 人以外 10 回とも同じパラメータを推定傾向であったが、口腔画像枚数 30 枚では 6 人しか同じパラメータを推定していない。特に専門医や学生 6、8 といった識別率が低い順位のパラメータを推定してしまうことから、口腔画像枚数が 30-40 枚にかけて推定の頑健さが失われる可能性が増加すると考えられる。口腔画像枚数が 20 枚と 10 枚の場合でも同様の現象が見られた。

また、識別率に関して表 3.6 の識別率結果を参照し、口腔画像枚数 10-70 枚における 1 人当たり 10 回推定されたパラメータを適応した平均識別率の比較を行った。その結果を表 6.9 に示す。平均識別率において口腔画像枚数 10-70 枚での識別率の変化に大きな変化は見られなかった。表

6.7 の識別率の平均順位に口腔画像枚数によって違いが見られなかったように識別率に対しても

83.1-83.2% という結果であった。

以上より，推定精度の観点から大きな変化が見られなかったため，推定の頑健さ（1人当たり10回同一パラメータを推定していたかどうか）に注目すると，口腔画像枚数が40枚から30枚に大きな変化がみられたため，口腔画像枚数40枚が望ましいと考えられる．そのため，診断支援システムを利用する新規の歯科医は本論文で使用した口腔画像92枚の中から40枚切り出しを行うことで個人の切り出し傾向の影響を軽減した識別が可能になる，しかし，2章の画像選定で述べた全く異なる場所を切り出し範囲に設定する例などもあることを考慮して，予備10枚を加えた50枚の口腔画像を対象に切り出しを行う方がより個人の切り出し傾向を正確に捉えることができると考える．

表 6.1: 局所画像サイズ，枚数パラメータ δ ， N 推定結果 (口腔画像 70 枚)

δ	1.0	0.9		0.8		0.7		平均順位
N	1	10	30	10	30	10	30	
専門医	0(2)	10(1)	0(2)	0(4)	0(4)	0(6)	0(7)	1
歯科医 2 年目	0(7)	0(3)	0(6)	0(3)	0(5)	10(1)	0(2)	1
歯科医 5 年目	0(7)	0(5)	0(6)	0(3)	0(4)	10(2)	0(1)	2
歯科医 15 年目 1	0(5)	0(2)	0(6)	0(3)	0(4)	10(1)	0(7)	1
歯科医 15 年目 2	10(2)	0(1)	0(5)	0(3)	0(6)	0(4)	0(7)	2
歯科医 30 年目 2	0(2)	0(2)	0(5)	9(2)	0(5)	1(1)	0(7)	1.9
歯科医 30 年目 3	10(1)	0(2)	0(6)	0(3)	0(5)	0(4)	0(7)	1
歯科医 30 年目 4	0(7)	0(3)	0(5)	0(2)	0(4)	10(1)	0(6)	1
学生 2	0(6)	10(3)	0(5)	0(1)	0(4)	0(2)	0(6)	3
学生 3	0(5)	0(1)	0(1)	10(7)	0(4)	0(6)	0(3)	7
学生 6	0(7)	0(4)	0(6)	0(2)	0(1)	10(4)	0(3)	4
学生 8	0(7)	10(2)	0(2)	0(1)	0(2)	0(5)	0(6)	2

表 6.2: 局所画像サイズ, 枚数パラメータ δ , N 推定結果 (口腔画像 60 枚)

δ	1.0	0.9		0.8		0.7		平均順位
N	1	10	30	10	30	10	30	
専門医	0(2)	10(1)	0(2)	0(4)	0(4)	0(6)	0(7)	1
歯科医 2 年目	0(7)	0(3)	0(6)	0(3)	0(5)	10(1)	0(2)	1
歯科医 5 年目	0(7)	0(5)	0(6)	0(3)	0(4)	10(2)	0(1)	2
歯科医 15 年目 1	0(5)	0(2)	0(6)	0(3)	0(4)	10(1)	0(7)	1
歯科医 15 年目 2	8(2)	2(1)	0(5)	0(3)	0(6)	0(4)	0(7)	1.8
歯科医 30 年目 2	0(2)	0(2)	0(5)	9(2)	0(5)	1(1)	0(7)	1.9
歯科医 30 年目 3	10(1)	0(2)	0(6)	0(3)	0(5)	0(4)	0(7)	1
歯科医 30 年目 4	0(7)	0(3)	0(5)	0(2)	0(4)	10(1)	0(6)	1
学生 2	0(6)	10(3)	0(5)	0(1)	0(4)	0(2)	0(6)	3
学生 3	0(5)	0(1)	0(1)	10(7)	0(4)	0(6)	0(3)	7
学生 6	0(7)	0(4)	0(6)	0(2)	0(1)	10(4)	0(3)	4
学生 8	0(7)	10(2)	0(2)	0(1)	0(2)	0(5)	0(6)	2

表 6.3: 局所画像サイズ, 枚数パラメータ δ , N 推定結果 (口腔画像 50 枚)

δ	1.0	0.9		0.8		0.7		平均順位
N	1	10	30	10	30	10	30	
専門医	0(2)	10(1)	0(2)	0(4)	0(4)	0(6)	0(7)	1
歯科医 2 年目	0(7)	0(3)	0(6)	0(3)	0(5)	10(1)	0(2)	1
歯科医 5 年目	0(7)	0(5)	0(6)	0(3)	0(4)	10(2)	0(1)	2
歯科医 15 年目 1	0(5)	0(2)	0(6)	0(3)	0(4)	10(1)	0(7)	1
歯科医 15 年目 2	10(2)	0(1)	0(5)	0(3)	0(6)	0(4)	0(7)	2
歯科医 30 年目 2	0(2)	0(2)	0(5)	10(2)	0(5)	0(1)	0(7)	2
歯科医 30 年目 3	10(1)	0(2)	0(6)	0(3)	0(5)	0(4)	0(7)	1
歯科医 30 年目 4	0(7)	0(3)	0(5)	0(2)	0(4)	10(1)	0(6)	1
学生 2	0(6)	10(3)	0(5)	0(1)	0(4)	0(2)	0(6)	3
学生 3	0(5)	0(1)	0(1)	10(7)	0(4)	0(6)	0(3)	7
学生 6	0(7)	2(4)	0(6)	0(2)	0(1)	8(4)	0(3)	4
学生 8	0(7)	9(2)	0(2)	0(1)	0(2)	1(5)	0(6)	2.3

6.4 おわりに

本章では新規の歯科医が初めて診断支援システムを利用する前に約 60 枚の口腔画像に対して切り出しを行ってもらう過程は手間を要すると考え、個人の切り出し傾向が特定できるために口腔画像枚数が何枚必要になるか調査を行った。本実験では局所画像サイズ・枚数パラメータ推定による学習モデル生成後の推定で用いる入力データとして用いる口腔画像枚数を 10, 20, 30, 40, 50, 60, 70 枚を使用した。実験の結果、平均順位および識別率に関して、10 枚から 70 枚まで大きな

表 6.4: 局所画像サイズ, 枚数パラメータ δ , N 推定結果 (口腔画像 40 枚)

δ	1.0	0.9		0.8		0.7		平均順位
N	1	10	30	10	30	10	30	
専門医	0(2)	10(1)	0(2)	0(4)	0(4)	0(6)	0(7)	1
歯科医 2 年目	0(7)	0(3)	0(6)	0(3)	0(5)	10(1)	0(2)	1
歯科医 5 年目	0(7)	0(5)	0(6)	0(3)	0(4)	10(2)	0(1)	2
歯科医 15 年目 1	0(5)	0(2)	0(6)	0(3)	0(4)	10(1)	0(7)	1
歯科医 15 年目 2	10(2)	0(1)	0(5)	0(3)	0(6)	0(4)	0(7)	2
歯科医 30 年目 2	0(2)	0(2)	0(5)	9(2)	0(5)	1(1)	0(7)	1.9
歯科医 30 年目 3	10(1)	0(2)	0(6)	0(3)	0(5)	0(4)	0(7)	1
歯科医 30 年目 4	0(7)	0(3)	0(5)	0(2)	0(4)	10(1)	0(6)	1
学生 2	0(6)	10(3)	0(5)	0(1)	0(4)	0(2)	0(6)	3
学生 3	0(5)	0(1)	0(1)	9(7)	0(4)	1(6)	0(3)	6.9
学生 6	0(7)	1(4)	0(6)	0(2)	0(1)	9(4)	0(3)	4
学生 8	0(7)	10(2)	0(2)	0(1)	0(2)	0(5)	0(6)	2

表 6.5: 局所画像サイズ, 枚数パラメータ δ , N 推定結果 (口腔画像 30 枚)

δ	1.0	0.9		0.8		0.7		平均順位
N	1	10	30	10	30	10	30	
専門医	0(2)	9(1)	0(2)	0(4)	0(4)	1(6)	0(7)	1.6
歯科医 2 年目	0(7)	0(3)	0(6)	0(3)	0(5)	10(1)	0(2)	1
歯科医 5 年目	0(7)	0(5)	0(6)	0(3)	0(4)	10(2)	0(1)	2
歯科医 15 年目 1	0(5)	0(2)	0(6)	0(3)	0(4)	10(1)	0(7)	1
歯科医 15 年目 2	7(2)	3(1)	0(5)	0(3)	0(6)	0(4)	0(7)	1.7
歯科医 30 年目 2	0(2)	0(2)	0(5)	8(2)	0(5)	2(1)	0(7)	1.8
歯科医 30 年目 3	10(1)	0(2)	0(6)	0(3)	0(5)	0(4)	0(7)	1
歯科医 30 年目 4	0(7)	0(3)	0(5)	0(2)	0(4)	10(1)	0(6)	1
学生 2	0(6)	10(3)	0(5)	0(1)	0(4)	0(2)	0(6)	3
学生 3	0(5)	0(1)	0(1)	8(7)	0(4)	2(6)	0(3)	6.8
学生 6	2(7)	3(4)	0(6)	0(2)	0(1)	5(4)	0(3)	4.6
学生 8	0(7)	8(2)	0(2)	0(1)	0(2)	2(5)	0(6)	2.6

変化がみられなかった。しかし、パラメータ推定の頑健さに注目した結果、口腔画像枚数 40 枚では 12 人中 9 名 10 回中 10 回同じパラメータを推定したが、30 枚において 12 人中 6 人と減少した。そのため、口腔画像枚数は 40 枚程度が望ましいと考えられる。新規で診断支援システムを使用する利用者は個人の切り出し傾向を捉えるために事前に口腔画像 40 枚切り出す必要があるが、疾患部位と異なる場所を切り出す可能性を考慮して 10 枚程度増やした 50 枚用意した方が好ましい。

表 6.6: 局所画像サイズ, 枚数パラメータ δ , N 推定結果 (口腔画像 20 枚)

δ	1.0	0.9		0.8		0.7		平均順位
N	1	10	30	10	30	10	30	
専門医	0(2)	9(1)	0(2)	0(4)	0(4)	1(6)	0(7)	1.6
歯科医 2 年目	0(7)	0(3)	0(6)	0(3)	0(5)	10(1)	0(2)	1
歯科医 5 年目	0(7)	0(5)	0(6)	0(3)	0(4)	10(2)	0(1)	2
歯科医 15 年目 1	0(5)	0(2)	0(6)	0(3)	0(4)	10(1)	0(7)	1
歯科医 15 年目 2	5(2)	5(1)	0(5)	0(3)	0(6)	0(4)	0(7)	1.5
歯科医 30 年目 2	0(2)	0(2)	0(5)	7(2)	0(5)	3(1)	0(7)	1.7
歯科医 30 年目 3	10(1)	0(2)	0(6)	0(3)	0(5)	0(4)	0(7)	1
歯科医 30 年目 4	0(7)	0(3)	0(5)	0(2)	0(4)	10(1)	0(6)	1
学生 2	0(6)	10(3)	0(5)	0(1)	0(4)	0(2)	0(6)	3
学生 3	0(5)	0(1)	0(1)	10(7)	0(4)	0(6)	0(3)	7
学生 6	5(7)	1(4)	0(6)	0(2)	0(1)	4(4)	0(3)	5.5
学生 8	0(7)	9(2)	0(2)	0(1)	0(2)	1(5)	0(6)	2.3

表 6.7: 局所画像サイズ, 枚数パラメータ δ , N 推定結果 (口腔画像 10 枚)

δ	1.0	0.9		0.8		0.7		平均順位
N	1	10	30	10	30	10	30	
専門医	0(2)	7(1)	0(2)	0(4)	0(4)	3(6)	0(7)	2.5
歯科医 2 年目	0(7)	0(3)	0(6)	0(3)	0(5)	10(1)	0(2)	1
歯科医 5 年目	0(7)	0(5)	0(6)	0(3)	0(4)	10(2)	0(1)	2
歯科医 15 年目 1	0(5)	0(2)	0(6)	0(3)	0(4)	10(1)	0(7)	1
歯科医 15 年目 2	2(2)	8(1)	0(5)	0(3)	0(6)	0(4)	0(7)	1.4
歯科医 30 年目 2	0(2)	0(2)	0(5)	8(2)	0(5)	2(1)	0(7)	1.8
歯科医 30 年目 3	10(1)	0(2)	0(6)	0(3)	0(5)	0(4)	0(7)	1
歯科医 30 年目 4	0(7)	0(3)	0(5)	0(2)	0(4)	10(1)	0(6)	1
学生 2	2(6)	8(3)	0(5)	0(1)	0(4)	0(2)	0(6)	3.6
学生 3	0(5)	0(1)	0(1)	10(7)	0(4)	0(6)	0(3)	7
学生 6	3(7)	2(4)	0(6)	0(2)	0(1)	5(4)	0(3)	4.9
学生 8	0(7)	9(2)	0(2)	0(1)	0(2)	1(5)	0(6)	2.3

表 6.8: 口腔画像枚数 10-70 枚による識別率の平均順位

	70 枚	60 枚	50 枚	40 枚	30 枚	20 枚	10 枚
専門医	1	1	1	1	1.6	1.6	2.5
歯科医 2 年目	1	1	1	1	1	1	1
歯科医 5 年目	2	2	2	2	2	2	2
歯科医 15 年目 1	1	1	1	1	1	1	1
歯科医 15 年目 2	2	1.8	2	2	1.7	1.5	1.4
歯科医 30 年目 2	1.9	1.9	2	1.9	1.8	1.7	1.8
歯科医 30 年目 3	1	1	1	1	1	1	1
歯科医 30 年目 4	1	1	1	1	1	1	1
学生 2	3	3	3	3	3	3	3.6
学生 3	7	7	7	6.9	6.8	7	7
学生 6	4	4	4	4	4.6	5.5	4.9
学生 8	2	2	2.3	2	2.6	2.3	2.3
平均	2.2	2.2	2.3	2.2	2.3	2.4	2.5

表 6.9: 口腔画像 10-70 枚における 1 人当たり 10 回推定されたパラメータを適応した平均識別率

	識別率 [%]
70 枚	83.2
60 枚	83.2
50 枚	83.1
40 枚	83.2
30 枚	83.1
20 枚	83.2
10 枚	83.1

第7章 結論

本論文では、口腔粘膜疾患診断支援システム利用時に生じる口腔画像内の口腔粘膜疾患だと思われる場所に対して切り出す際に生じる個人の特性を考慮し、影響を軽減する手法を提案した。

従来システムでは、口腔画像から疾患だと思われる場所を手動で切り出す必要があり、切り出し画像を入力画像とすることで疾患の識別を行う。従来システムは切り出し範囲のみを入力画像としていた。しかし、切り出し範囲の位置、サイズが利用者によって異なることから識別率に影響を及ぼす問題があった、個人の切り出し傾向をシステムに反映させることで、従来システム利用時より識別率を向上を目的とした。

はじめに個人の切り出し傾向を知るため、口腔画像 117 枚に対して切り出しを行った 18 名のデータに対して行った。しかし、口腔画像によって全く異なる場所を切り出し範囲に設定しているなど、個人の切り出し傾向が捉えられない可能性があったため画像選定を行った。本実験では口腔画像内の疾患位置が不明であることから専門医の切り出し範囲を基準に行った。その結果、画像選定の除外基準を多く満たしていた人、多くの人が間違った切り出し範囲を設定した口腔画像を除外することで 18 名から 12 名に、口腔画像は 112 枚から 83-91 枚に調整した。これらのデータより切り出し範囲の傾向を調査した結果、大きく切り出す傾向が見られた人が 8 名、小さく切り出す傾向が 3 名であった。

切り出し傾向の結果より、専門医と比較して切り出し範囲を大きく切り出す傾向の人が多かったことから、切り出し画像だけでなく、切り出し画像の内側に注目した、様々なサイズの局所画像を生成する提案システムに改良した。提案システムを用いて識別率の変化を確認した結果、大きく切り出す傾向が見られた人全員、局所画像を含んだ提案システムを用いることで識別率は向上したが、人によって一番識別率が高い局所画像サイズ・枚数パラメータ δ 、 N が異なっていた。 δ は局所画像サイズ倍率で N は画像枚数を指す。そこで切り出す人の傾向を考慮し軽減したシステムを実現するためには、局所画像サイズ・枚数パラメータを推定を行う必要があると考えた。

パラメータ推定に用いるデータとして切り出し範囲の midpoint と縦横の長さ と識別率に注目したが適切な推定方法が不明であったため、重回帰分析、リッジ回帰、CCA、RCCA、CCA-SOM を用いた。12 名を対象に交差検証による推定を行った。個人の切り出し傾向を考慮した推定方法であるか、頑健な推定は実現できているか、推定精度の 3 点を考慮した結果、リッジ回帰がパラメー

タ推定に適している手法であると判明した。また、従来システムと提案システム（パラメータ推定あり）の識別率に関して、5%の向上が確認できた。また、提案システムのパラメータ推定ありと各パラメータによる12人の平均識別率が1番高い識別率を比較した結果、パラメータ推定を行うことで1.3%の識別率向上を確認できた。

しかし、診断支援システムの利用にあたり、事前準備として切り出し傾向を特定するために口腔画像枚数を70枚前後切り出してもらう行為は大変であると考えられる。そこでパラメータ推定に必要な最小限の口腔画像枚数は何枚か調査した。推定手法ではリッジ回帰を行い、口腔画像枚数70, 60, 50, 40, 30, 20, 10枚に設定し、アンサンブル学習によって得られた識別率を順位付けすることで口腔画像枚数に推定精度を推移を確認した。その結果口腔画像枚数50枚で1部の人に対して切り出し傾向が見られない可能性が生じること、30枚でさらに切り出し傾向が見られない人が増えたため口腔画像に依存したパラメータ推定になってしまうことが判明した。そのため、40枚の口腔画像枚数が切り出し傾向を特定するのに十分な枚数であると考えられるが、口腔画像によって全く異なる場所を切り出し範囲に設定する可能性なども考慮して50枚が適切であると考ええる。

今後の課題として、切り出し範囲が小さい傾向に対する高識別の実現が挙げられる。切り出し範囲が大きい傾向の場合に用いたアンサンブル学習を実装することで解決できそうではあるが局所範囲の設定として切り出し範囲の外側に注目する必要がある。しかし、切り出し範囲の外側に歯や医療機器などが存在し不要な情報も含まれる可能性が考えられる。そのため、局所画像を生成する前に歯や医療機器の情報を取り除く必要がある。また、多くの人が疾患がある場所を発見できなかった口腔画像が複数枚みられたことから口腔画像の撮影においても何かしらのルールを設定する必要がある。

謝辞

本研究を遂行するにあたり、終始懇切丁寧なご指導を賜りました九州工業大学大学院生命体工学研究科人間知能システム工学専攻 堀尾 恵一 教授に心から感謝致します。本論文をまとめるにあたり、有意義なご助言とご討論を頂いた九州工業大学大学院生命体工学研究科人間知能システム工学専攻、井上 創造 教授、田向 権 教授、九州工業大学工学部電気電子工学科電子システム工学コース、河野 英昭 准教授に謝意を申し上げます。また、これまでの学生生活において、多くのご支援を頂いた九州工業大学大学院生命体工学研究科人間知能システム工学専攻堀尾研究室をはじめとした多くの研究室の方々ならびに大島商船高等専門学校情報工学科の先生方に厚く御礼申し上げます。

参考文献

- [1] 井上孝, 小林隆太郎.『 歯科医院で気づく・見落とさない!色と形からみる口腔粘膜病変 デジタルハイジーン別冊,』 医歯薬出版, p.35, 2011.
- [2] 最新がん統計, https://ganjoho.jp/reg_stat/statistics/stat/cancer/3_oral.html, 2021年9月6日閲覧.
- [3] 柴原孝彦. "口腔外科学を学んで -口腔がん検診活動 28年, 歯科学法," Vol.120, No.3, pp285-303, 2020.
- [4] M.W. Finkelstein. "A Guide to Clinical Differential Diagnosis of Oral Mucosal Lesions," dentalcare.com, 2010.
- [5] M.R. Oczko and M. Mendak. "Differential diagnosis of oral leukoplakia and lichen planus -on the basis of literature and own observations," J Stoma, Vol.64, No.5-6, pp.355-370, 2011.
- [6] S.E. Mofty. "Early detection of oral cancer," J.Oral Maxillofac.Surg, Vol1, pp.25-31, Egypt, 2010.
- [7] M.W. Lingen, J.R. Kalmar, T. Karrison and P.M. Speight. "Critical evaluation of diagnostic aids for the detection of oral cancer," Oral Oncology, Vol.44, Issue 1, pp.10-22, 2007.
- [8] Cesar Rivera. "Essentials of oral cancer," International Journal of Clinical and Experimental Pathology, Vol.8, No.9, 2015.
- [9] H. Bessho, T. Onda, T. Yakushiji, N. Yamamoto, T. Nomura, K. Sato, H. Saito, T. Yamauchi, A. Katakura, N. Takano, T. Shibahara, "Development of oral cancer navigation system by Department of Oral and Maxillofacial Surgery," Oral Science in Japan, pp.99-100, 2014.
- [10] 六反田賢, 柳本惣市, 鳴瀬智史, 川北晃子, 坂本由紀, 梅田正博. "遠隔口腔粘膜疾患患者相談システムの構築," 日本口腔診断学会雑誌, Vol.30, No.1, pp83-88, 2017.

- [11] 平成28年医師・歯科医師・薬剤師調査の概況, <https://www.mhlw.go.jp/toukei/saikin/hw/ishi/16/dl/kekka.pdf>, 2019年10月15日閲覧.
- [12] T. Morikawa, A. Kozakai, A. Kosugi, H. Bessho, T. Shibahara. "Image processing analysis of oral cancer, oral potentially malignant disorders, and other oral diseases using optical instruments," *International Journal of Oral and Maxillofacial Surgery*, Available online 7, pp.1-7, 2019.
- [13] 森川貴迪, 別所央城, 小坂井絢子, 小杉彩歌, 柴原孝彦. "口腔粘膜観察用光学機器 IllumiScan® による口腔扁平上皮癌ならびに口腔扁平苔癬の解析," *歯科学報*, Vol.117, No.5, pp383-392, 2017.
- [14] Eric C. Yang, Imran S. Vohra, Hawraa Badaoui, et al. "Development of an integrated multimodal optical imaging system with real-time image analysis for the evaluation of oral premalignant lesions," *Journal of Biomedical Optics*, Vol.24, Issue2, 2019.
- [15] 片山洋平, 松井智一, 佐藤佑磨, 中岡黎, Chang Xin, Dang Chenyu, 松田裕貴. "口腔内画像撮影支援システムの提案," 第28解マルチメディア通信と分散処理ワークショップ論文集, pp253-256, 2020.
- [16] 土生学, 國領真也, 吉岡泉, 富永和宏. "口腔内写真を用いた口腔粘膜疾患診断支援システムは実用に耐えうるか?," *九州歯科学会雑誌*, Vol.71, p29, 2017.
- [17] Gizem Tanriver, Merva Soluk Tekkesin, and Onur Ergen. "Automated Detection and Classification of Oral Lesions Using Deep Learning to Detect Oral Potentially Malignant Disorders," *cancers*, Vol.13, No.11, 2021.
- [18] Roshan Alex Welikala, Paolo Remagnino, Jian Han Lim, et al. "Automated Detection and Classification of Oral Lesions Using Deep Learning for Early Detection of Oral Cancer," *IEEE Access*, Vol.8, pp.132677-132693, 2021.
- [19] Assaf Gottlieb, Robert Hoehndorf, Michel Dumontier, and Russ B Altman. "Ranking Adverse Drug Reactions with Crowdsourcing," *Journal of Medical Internet Research*, Vol.17, iss.3, e80, pp.1-13, 2015.
- [20] Aditi Ramachandran, Chien-Ming Huang, and Brian Scassellati. "Give Me a Break!: Personalized Timing Strategies to Promote Learning in Robot-Child Tutoring," *Proceedings*

- of the 2017 ACM/IEEE International Conference on Human-Robot Interaction, pp.146-155, 2017.
- [21] Daniel Leyzberg, Aditi Ramachandran, and Brian Scassellati. "The Effect of Personalization in Longer-Term Robot Tutoring," *ACM Transactions on Human-Robot Interaction*, Vol.7, Issue. 3, pp.1-19, 2018.
- [22] 西 勇作, 三澤 秀明, 大谷 泰志, 土生 学, 堀尾 恵一, 富永 和弘. "口腔画像から抽出した構造要素による扁平上皮癌の検出," 電子情報通信学会 SIS 研究会, pp.7-10, 2011.
- [23] 西 勇作, 大谷 泰志, 土生 学, 堀尾 恵一, 富永 和弘. "口腔画像に基づく口腔粘膜疾患診断支援における勾配強度ヒストグラムの有意性の検証," 電子情報通信学会 SIS 研究会, pp.19-22, 2012.
- [24] Y. Nishi et al. "Discrimination of Oral Mucosal Disease Inspired by Diagnostic Process of Specialist," *J. of Medical and Bioengineering*, Vol.2, No.1, pp.57-61, 2013.
- [25] 元木慎也, 斎藤謙太郎, 土生学, 堀尾恵一, 富永和宏. "画像処理による口腔粘膜疾患の識別における解像度の与える影響に関する考察," 電気情報通信学会, SIS2014 Vol.3, pp.13-16, 2014.
- [26] C.J.C.Burges, "A Tutorial on Support Vector Machine for Pattern Recognition", *Data Mining and Knowledge Discovery*, No.2, pp.121-167, 1998.
- [27] Lester WB, "Oral Medicine: Diagnosis and Treatment, Lippincott Company: sixth edition," pp.96-111, 608-612, 1971.
- [28] Irwin WS, "Oral medicine A clinical approach with basic science correlation: second edition," The C.V. Company, pp.109-117, 342-360, 1973.
- [29] Norman KW and Paul WG. "Differential Diagnosis of Oral Regions:third edition," The C.V Mosby Company, pp.60-89, 123-128, 160-162, 1985.
- [30] 堀尾恵一, 松本宗兵, 末竹規哲, 大谷泰志, 土生学, 富永和宏, 山川烈. "確率的弛緩法を用いた口腔粘膜疾患診断支援システム," 26th Fuzzy System Symposium, pp.385-388, 2010.
- [31] 岩崎 洋一郎. "画像処理入門," 培風館, p.149, 2010.
- [32]]小畑 秀文. "モルフォロジー," コロナ社, 1996.

- [33] 浅野 晃, 浅野 千恵, 木森 義隆, 棟安 実治, 延原 肇, 藤尾 光彦. "非線形画像・信号処理 モルフォロジーの基礎と応用," 丸善株式会社, pp.65-67, 2010.
- [34] 浅野 晃. "マセマティカルモルフォロジーの思想," 信学会, Vol.4, No.2, pp.113-122, 2010.
- [35] Nanto Ozaki et al. "Ensemble Classifier which is Robust against Individual Difference in Diagnosis Support System of Oral Mucosal Disease," SISA2016, pp.258-262, 2016.
- [36] 永田靖, 棟近雅彦. "多変量解析入門," サイエンス社, pp.61-86, 2001.
- [37] Trevor Hastie, Robert Tibshirani, Jerome Friedman. "統計的学習の基礎 -データマイニング・推論・予測," 共立出版, pp.76-81, 2014.
- [38] 赤穂昭太郎. "正準相関分析入門-複数種類の観測からの共通情報抽出法-, " 日本神経回路学会誌, Vol.20, No.2, pp62-72, 2013.
- [39] Natalia Y. Bilenko, Jack L. Gallant. "Pyrcca: Regularized Kernel Canonical Correlation Analsis in Python and Its Applications to Neuroimaging," Frontiers in Neuroinformatics, Vol.10, No.49, pp.1-12, 2016.
- [40] 米田 圭佑, 中野 貴理博, 堀尾 恵一, 古川 徹生. "自己組織化マップによるマルチビュー関係データの包括的可視化," 知能と情報, Vol.30, No.2, 2018.