

修士論文

# 癌病理画像における癌巣抽出支援の検討

同志社大学大学院 工学研究科 情報工学専攻  
博士前期課程 2008年度 773番

渡辺 章人

指導教授 三木 光範教授

2010年1月23日

## Abstract

This thesis presents an algorithm that extracts target cancer images from pathological images. New diagnostic approaches and treatment techniques for various diseases have been developed. In these methods, medical images of pathology diagnosis field are analyzed and discussed using clinical examination technique and development of medical imagers. However, these methods still require medical doctors to check the pathological images and make appropriate diagnoses. However, the number of doctors who can diagnose and check these images is decreasing. Therefore, there is a demand for pathological image diagnostic support systems that can aid in checking pathological images. Pathological image diagnostic support systems carry out diagnostic processing to obtaining medical diagnostic information by computer processing of the images to be interpreted. Specifically, these systems derive characteristic values of pathological images and these characteristics are checked and compared with previous cases. In most systems, the obtained characteristics can be derived from pathological images. However, it is very difficult for medical doctors to determine the process for extracting the affected region. In addition, most image processing is dependent on an individual problem. Therefore, several processing methods have been applied to identify appropriate processing techniques. Therefore, the development of an automated method for extracting the affected region would be very useful.

In this thesis, a novel algorithm to derive the targets from images is proposed. The proposed algorithm mimics the process of diagnosis by medical doctors. In this process, the targets are narrowed down in coarse images. The narrowed targets are refined in the fine images. The proposed algorithm, Support Vector Machine (SVM), extracts the affected region using classifiers, and judges whether the extracted area is the affected region using color information.

The proposed algorithm was also applied to extract a cancer nest from mouth cancer images to determine its effectiveness. The results of experiments indicated that the proposed algorithm was effective for extracting a cancer nest from mouth cancer images.

# 目次

<b>1</b>	<b>序論</b>	<b>1</b>
<b>2</b>	<b>病理画像診断支援</b>	<b>2</b>
2.1	病理画像とは . . . . .	2
2.2	病理画像診断支援の概要 . . . . .	2
2.3	病理画像診断支援に期待される効果 . . . . .	3
<b>3</b>	<b>口腔癌</b>	<b>4</b>
3.1	口腔癌の現状 . . . . .	4
3.2	口腔癌の組織学的特徴 . . . . .	5
<b>4</b>	<b>病理診断医の診断方法に基づく癌胞巣抽出戦略</b>	<b>5</b>
4.1	病理診断医の診断方法のモデル化 . . . . .	5
4.2	病理診断医の診断方法に基づく癌胞巣抽出戦略 . . . . .	5
<b>5</b>	<b>サポートベクターマシン</b>	<b>8</b>
5.1	パターン認識 . . . . .	8
5.2	サポートベクターマシンの原理 . . . . .	8
<b>6</b>	<b>色ヒストグラムインタセクション</b>	<b>10</b>
6.1	色ヒストグラム . . . . .	10
6.2	色ヒストグラムインタセクション . . . . .	11
<b>7</b>	<b>提案戦略の実装</b>	<b>12</b>
7.1	SVM を利用した大域領域における病変部位抽出識別器の生成 . . . . .	12
7.2	SVM を利用した局所領域における病変部位抽出識別器の生成 . . . . .	13
7.3	色ヒストグラムインタセクションを利用した病変部位判定 . . . . .	14
<b>8</b>	<b>提案戦略の有効性の検証</b>	<b>15</b>
8.1	特徴量の検討 . . . . .	16
8.2	同一画像における有効性の検証 . . . . .	18
8.3	同一検体における有効性の検証 . . . . .	20
<b>9</b>	<b>結論</b>	<b>22</b>

# 1 序論

近年，医療における病理診断分野において，新しい臨床検査法や画像装置の開発，精度の向上，症例の蓄積により，様々な疾病の新たな診断法や治療法が知られるようになってきた．一方，日本では病理診断を行っている医師が極端に不足している．その為，計算機を利用することで病理診断の支援を行う，病理画像診断支援システムに対するニーズが近年増加している<sup>1)</sup>病理画像診断支援システムとは，病理画像から癌などの構成要素の特徴量を抽出し，抽出した特徴量を計測後，症例と照らし合わせるにより関連性を数値的に算出し，医師に対して過去の症例との関連性や病気の進行具合の提示等を行うものである．病理画像診断支援システムの多くは，各症例に合わせた特殊な画像処理によって病理部位等の目標部位抽出を行う．しかしながら，このような病理画像を解析するための画像処理手法を人手によって考案することは困難であると考えられる．また，元来，画像処理は医用に限らず問題試行な面が強く，それぞれの問題毎に種々の方法を試みて適当な処理方法を探す仮定をとることが多い．しかし，その中でも基本的な常識，考え方，手法などが存在すると考えられている<sup>2)</sup>．従って，本研究では，対象とする画像について，問題毎に特化した特徴量を利用するのではなく，画像情報として広く一般的な情報を特徴量として利用することを考える．そして，病理医の診断過程をモデル化し，既存の学習手法を適用することにより，病理の現場において，顕微鏡画像を対象とした汎用的な識別器の生成が可能と考える．

本論文では，病理医の診断過程のモデル化として，2階層の識別器とテクスチャの色情報を利用した目標病変部位抽出手法を提案する．前者は顕微鏡画像に対し，拡大縮小を行い診察を行う模様を模擬したものであり，後者は構成要素の形状や包含要素の分布状況から診断する模様を模擬したものである．また，精度や特徴を確認するために，提案手法を口腔癌病理画像における癌巣抽出処理に適用した．本論文の構成を以下に示す．まず，第2章で病理画像診断支援の概要について述べ，第3章では，本論文で処理の対象とする口腔癌の説明を行う．そして，第4章で提案する手法の概要について述べ，第5章，第6章で提案手法で利用する既存手法の原理について説明を行う．第7章では，提案手法の具体的な実装方法について触れ，その有効性を第8章で検証する．最後に，第9章で結論を述べる．

## 2 病理画像診断支援

### 2.1 病理画像とは

病理画像とは、人体から疾患の原因とみなされる細胞を直接摘出してパラフィン等で固定した後に、薄片化してヘマトキシリン・エオシン (Hematoxylin-Eosin:HE) 等で染色した病理スライドの顕微鏡画像である。近年では、この病理スライド全体をスキャナで撮影することにより、バーチャルスライドを作成可能となった。このバーチャルスライド画像は、顕微鏡下でスライド観察する時と同様に任意の倍率の病理画像を得ることが可能である。本研究では、バーチャルスライドにより得た病理画像を扱う。Fig. 2.1 に本研究で扱う病理画像の一例を示す。

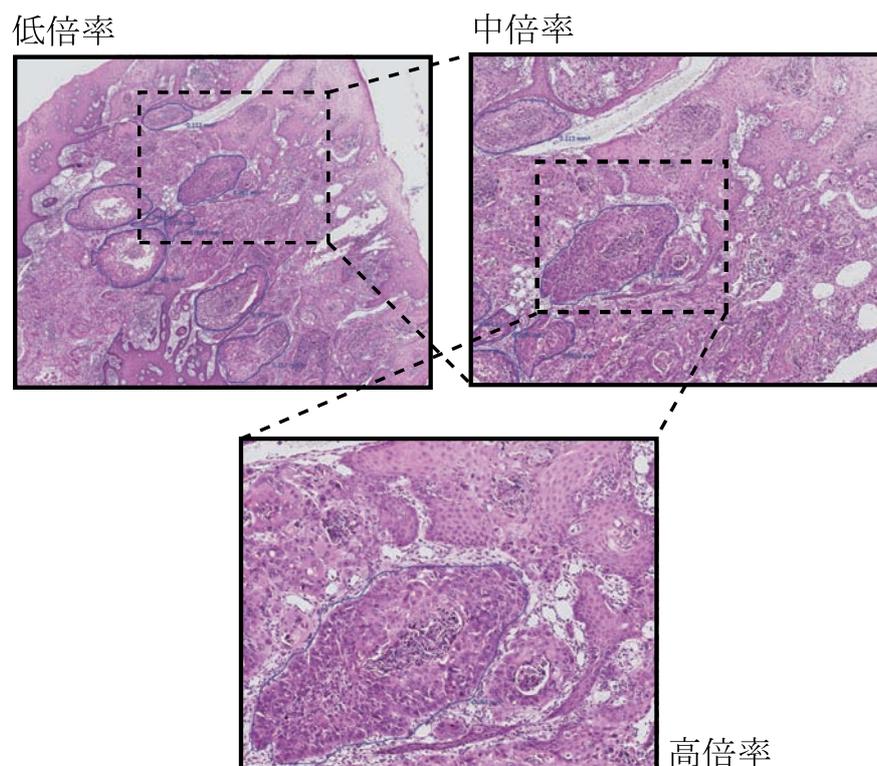


Fig. 2.1 病理画像の一例<sup>1</sup>

### 2.2 病理画像診断支援の概要

一般的に、病理医は病理組織の構成要素を判断することによって病理診断を行う。病理組織の構成要素といっても、その形状は様々である。そのため、病理医が見ても病変部位かどうかの判断に迷う組織の構造や細胞の変化が数多く存在する。この場合、病理医は、組織の構造の変化の程度や細胞の中の核の大きさの比率などを計測する。また、構成要素の判断は病理医の主観によって行われる為、その人の経験、体調などに左右されてしまう可能性がある<sup>3)</sup>。従って、病理診断の解析を自動化することにより、病理医の診断の支援が可能であると考えられる。また近年、日本では癌などの病理診断を行っている医師が極端に不足している。そのため、一人一人の医師に対する負担が増大しているが、診断支援を行うことで、病理医の負担の軽減にもつながると期待している。

<sup>1</sup>画像提供：東北大学加齢医学研究所病態臓器構築研究分野福本学教授

そこで、病理診断を自動化し、医師の病理診断を支援する病理画像診断支援システムが考案されている。病理画像診断支援システムの流れを Fig. 2.2 に示す。

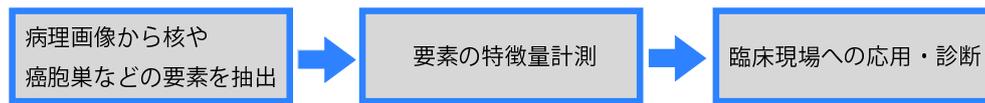


Fig. 2.2 病理画像診断支援システム

Fig. 2.2 に示すように、病理画像診断支援システムは、病理画像から癌などの構成要素を抽出し、要素の特徴量を計測することで癌かどうかを判断し、病理診断を行うという一連の行程を自動化したものである。以下にそれぞれのプロセスについて説明する。

#### Step1 病理画像から核や癌胞巣などの構成要素の抽出

病理画像から、病気の原因である核や癌の構成要素の要素の特徴量を抽出する。

#### Step2 要素の特徴量計測

Step1 で抽出された要素を定量化し、特徴量を計測する。

#### Step3 臨床現場への応用・診断

特徴量計測したデータをもとに、癌などの病気かどうかの判断をし、臨床現場へ応用する。

本研究では、Fig. 2.2 に示した病理画像診断支援システムの中でも、最初のプロセスである病理画像からの構成要素自動抽出について考える。

### 2.3 病理画像診断支援に期待される効果

病理画像診断支援に期待される効果は以下の通りである。

#### (1) 病理医の診断支援

病変組織の構成要素といっても、その形状は様々である。そのため、病理医が見ても癌かどうかの判断に迷う組織の構造や細胞の変化が数多く存在する。この場合、病理医は、組織の構造の変化の程度や細胞の中の核の大きさの比率などを計測するが、病理診断の解析を自動化することにより、病理医の診断の支援が可能であると期待している。また近年、日本では癌などの病理診断を行っている医師が極端に不足している。そのため、一人一人の医師に対する負担が増大しているが、診断支援を行うことで、病理医の負担の軽減にもつながると期待している。

#### (2) 病理診断の迅速化

高齢化社会に伴い、癌などの検診が今後増大していくと考えられるが、病理診断の解析を自動化することにより、迅速で正確な病理診断が可能であると期待している。

#### (3) 癌の病理診断における新たな指標への期待

癌の病理診断において、癌と判断するための計測手段が癌組織の大きさや重さしかなく、十分な判断材料が少ないのが現状である。そこで、癌構成要素を抽出することで、癌構成要素の特徴の数値化が可能になり、癌診断の新たな指標が考案されることを期待する。

### 3 口腔癌

本章では、提案手法の有効性検証の為に、検証データとして利用した口腔癌について、その概要と組織学的特徴を説明し、本論文における抽出対象について言及する。

#### 3.1 口腔癌の現状

口腔癌はヒトの口腔内で発症する癌である。近年、わが国の口腔癌患者は増加の傾向にあり、治療成績向上にも関わらず、口腔癌による死亡数は1950年の約700人から2008年の約6600人と過去50年間に約7.2倍の増加が記録されている<sup>4)</sup>同様の傾向は世界にも見られ、Macfarlenらによると24ヶ国中19ヶ国で口腔癌の増加傾向があり、特に中央、西ヨーロッパ諸国において、過去30年間で男性患者が3倍以上に増加していることが指摘されている<sup>5)</sup>口腔癌の一種である舌癌の一例をFig. 3.1に示す。



Fig. 3.1 口腔癌(舌癌)<sup>1</sup>

癌の治療においては手術による病変の切除が基本であるが、補助的に化学療法や放射線療法が用いられる。たとえば術前補助療法(ネオアジュバント療法)は、手術前に化学療法や放射線療法を行い、癌の縮小と転移の防止を期待するものである。しかし化学療法や放射線療法の効果は患者ごとに異なるため、癌が縮小したり消滅したりせず、場合によれば何の効果も得られないまま癌が拡大する場合もある。また薬の種類も複数あり、どの患者にどの薬を用いれば最大の効果が得られるのかも予測できないのが現状である。このような術前補助療法では、効果がある患者にとっては、手術による切除範囲が狭められまた早い段階で転移を防止することができるため有用な方法であるといえる。一方、効果がない患者にとっては、次の治療の時期を遅らせることになるばかりか無用な副作用(有害事象)を与えることになり、QOL(Quality of Life:生活の質)の低下を招く。これらのことから、患者ごとに化学療法や放射線療法の効果が事前に予測できることが望ましいが、現段階ではその方法が確立されていない。効果予測を行うアプローチとして、病理画像内の構成要素の特徴を基に判断を行うことが検討されており、本研究における病理部位抽出が前述の研究の支援にも繋がると考えられる。

<sup>1</sup>画像提供：東北大学加齢医学研究所病態臓器構築研究分野福本学教授

### 3.2 口腔癌の組織学的特徴

口腔癌は扁平上皮癌と唾液腺癌にわけられ、その多くは扁平上皮癌である<sup>6)</sup>。扁平上皮癌では、粘膜上皮層の扁平上皮細胞が癌化し、そこから拡大していく。病理画像を見ると、上皮層内の正常な細胞の他、癌化した扁平上皮細胞、結合織部分には炎症細胞、リンパ球、間質細胞などが確認できる。病理医は、上皮層内の細胞の並び方、上皮細胞の核の大きさや形や色むら、炎症細胞やリンパ球の数や広がりなどによって癌かどうかの判定を行う。癌細胞は胞巣とよばれる塊を形成することがあり、これも癌の判定の重要な要素となる。癌胞巣はその形状や大きさによって、悪性度や浸潤様式、治療効果の判定に用いられる場合もある<sup>7)</sup>。なお、本論文では口腔癌病理画像における癌胞巣の抽出を対象とする。

## 4 病理診断医の診断方法に基づく癌胞巣抽出戦略

本章では、病理診断医の診断方法の概要を述べ、それに基づいた癌胞巣抽出についての戦略を述べる。なお、本章で掲載する病理画像は、東北大学加齢医学研究所病態臓器構築研究分野福本学教授より提供されたものである。

### 4.1 病理診断医の診断方法のモデル化

病理医は予め、過去の症例についての知識を有しており、それにより個人毎に何らかの識別器を持っていると考えられる。2.1節でも示したように、病理診断医は事前に染色された検体または病理画像について、拡大縮小を行い、倍率を変化させることで診断を行っている。また、人の目で病理部位の判断を行うため、極小範囲(画像ではピクセル単位)での診断を行っているとは考え難い。従って、病理診断医が診断を行う際に以下の点について考慮し、診断を行っているとは仮定できる。

仮定 1 対象画像の倍率を調整することで、病理部位の特定を行う。

仮定 2 極小範囲(ピクセル単位)を見て診断を行うのではなく、注目領域とその近傍領域の特徴量を判断することで、病理部位の特定を行う。

仮定 3 注目領域のテクスチャを見て、病理部位の特定を行う。

仮定 4 病理部位特定は過去の症例より得た知識に依存している。

### 4.2 病理診断医の診断方法に基づく癌胞巣抽出戦略

#### 4.2.1 基本方針

1章で述べたように、本研究の目的は病理医の診断過程のモデル化による識別器の生成にある。従って、本節では4.1節で列挙した仮定を踏まえ、診断過程を模擬するための戦略提案を行う。

病理診断において、倍率が低い画像での診断は大域的な領域での病変部位判別、倍率の高い画像での診断は局所的な領域での病変部位判別と定義することができる。本論文では、一定の大きさを持った領域単位での病理部位判定を行い、また注目する領域の大きさを変更することにより、大域的な領

域での病変部位判別と局所的な領域での病変部位判別を行う．そのため，判別処理を自動構築するための学習画像及び病変部位の判定に利用する入力画像を， $m \times m$  ピクセルと  $n \times n$  ピクセル ( $m > n$ ) の小領域に分割を行う．以後，分割した領域をセル，領域のサイズをセルサイズと呼ぶ．また，注目領域の病変部位判定を行うために，注目しているセルの特徴量だけでその領域の病理判定を行うのではなく，注目しているセルと隣接する 8 個のセル (近傍セル) の特徴量を利用することにより，注目領域の病変部位判定を行うことを考える．病理画像のセル分割について，Fig. 4.1 に示す．

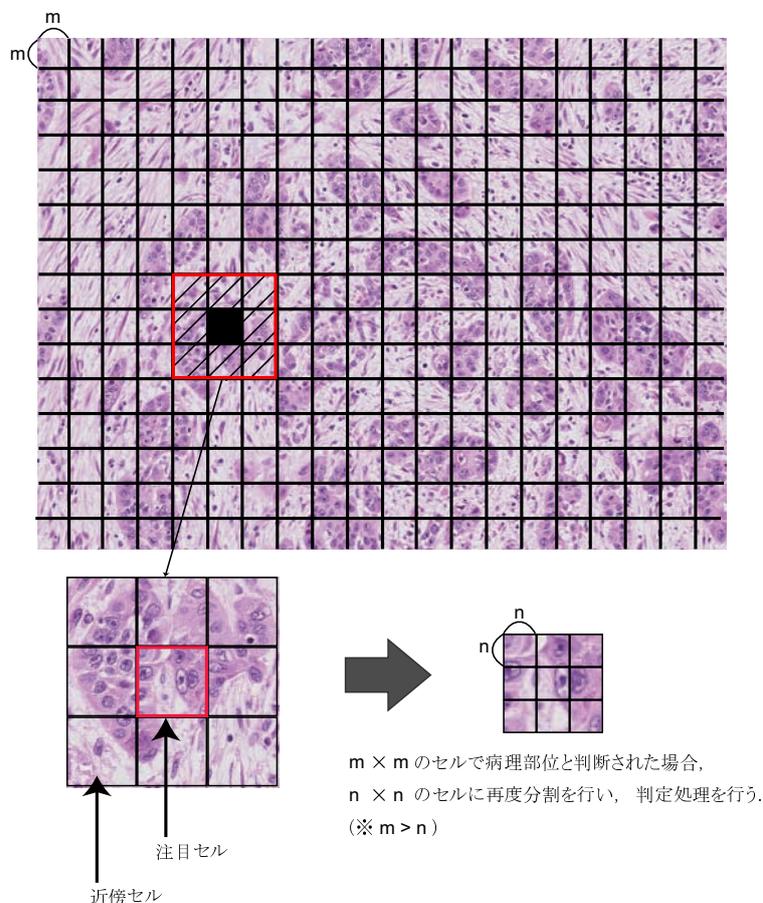


Fig. 4.1 病理画像のセル分割<sup>1</sup>

まず， $m \times m$  ピクセルのセルサイズにおいて，注目領域と近傍領域の特徴より識別器を生成する．この時，生成された識別器は，あくまで病変部位が含まれる領域を大まかに特定する識別器に過ぎない．従って，識別器によって抽出された領域について，さらに詳しく病変部位の特定を行うための識別器を生成する必要がある．そこで，識別器によって抽出された領域を  $n \times n$  ピクセルのセルサイズを利用し，さらに細かい領域単位で再度特徴量の計測を行い，識別器の生成を行うことを考える．

さらに，抽出した領域より，テクスチャによる病変部位判別を行うことを考える．本稿では，上記の戦略の具体的な実装として以下を行い，戦略の検証を行う．

- (1) サポートベクターマシンを利用した大域領域における病変部位抽出識別器の生成
- (2) サポートベクターマシンを利用した局所領域における病変部位抽出識別器の生成

<sup>1</sup>画像提供：東北大学加齢医学研究所病態臓器構築研究分野福本学教授

### (3) 色ヒストグラムインタセクションを利用した病理部位判定

サポートベクターマシン (Support Vector Machine: SVM) とは, V. Vapnik などによって提案されたパターン識別手法である<sup>9,10</sup>). また, 色ヒストグラムインタセクションとは, SWAIN らによって提案された領域の色ヒストグラムを元に類似度を判定する手法である<sup>8</sup>). 両手法の詳細については 5 章と 6 を, 提案戦略の実装の詳細については 7 章を参照されたい.

なお, 本研究では提案戦略の実装として SVM を利用しているため, 予め学習用画像データを必要とする. 学習用画像データは, 病理画像より抽出対象を含んだ領域を抽出した画像と, その画像から識別器が抽出すべき領域をマーキングしたものを一つの学習セットとする. 学習画像データの一例を Fig. 4.2 に示す.

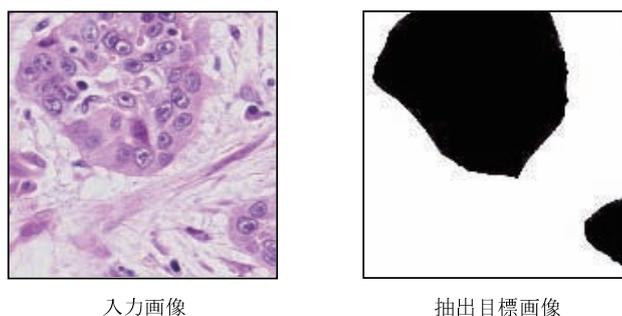


Fig. 4.2 学習用画像データの一例<sup>1</sup>

#### 4.2.2 提案戦略で用いる特徴量

提案アルゴリズムでは, 注目領域と近傍領域が持つ色情報を特徴量として, SVM によって識別器の生成を行う. 色情報としては, RGB モデル及び人間の感性に似た色の表現方法として利用される HSV モデル<sup>11</sup>), RGB モデルより算出したグレースケール値を用いた. 利用した特徴値を以下に示す.

- SVM を利用した識別器生成に利用する特徴量
  - 注目領域と近傍領域の各領域の R,G,B 値の平均値よりなる 27 ベクトルの特徴量
  - 注目領域と近傍領域の各領域の H,S,V 値の平均値よりなる 27 ベクトルの特徴量
  - 注目領域と近傍領域の各領域のグレースケール値の平均値よりなる 9 ベクトルの特徴量
- ヒストグラムインタセクションを利用したテクスチャ判別に利用する特徴量
  - 注目領域の持つ色ヒストグラム

#### 4.2.3 提案手法による抽出領域の評価方法

提案アルゴリズムで抽出した領域についての評価を行うため, 本稿では, 抽出対象を正しく抽出する確率 (true positive fraction:TPF), 非抽出対象を誤って抽出する確率 (false positive fraction:FPF) を用いる. なお, TPF 及び FPF は, 式 (4.2) のように定義した.

$$TPF = \frac{A_C \cap A_R}{A_R} \quad (4.1)$$

$$FPF = \frac{A_C \cap \bar{A}_R}{A_C} \quad (4.2)$$

ここで、 $A_R$  は抽出対象領域、 $A_C$  は提案手法によって抽出された領域を示す。従って、抽出の精度が高くなるほど、 $TPF$  の値が 1.0 に、 $FPF$  の値が 0.0 に近づき、抽出対象のみを正しく抽出できた場合は  $TPF$  が 1.0、 $FPF$  が 0.0 となる。

## 5 サポートベクターマシン

サポートベクターマシン (Support Vector Machine: SVM) は、1960 年代に Vapnik らが考案した Optimal Separating Hyperplane (OSH) を起源とし<sup>9)</sup>、1990 年代にカーネル学習と組み合わせた非線形の識別手法に拡張された。カーネルトリックにより非線形の識別関数が構成できるように拡張された SVM は、現在知られている手法の中で最もパターン認識性能の優秀な学習モデルの一つである<sup>10)</sup>。

本章では、まずパターン認識について説明し、次に SVM の数学的な原理を説明する。

### 5.1 パターン認識

パターン認識とは、認識対象がいくつかの概念に分類できるとき、観測されたパターンをそれらの概念のうちの 1 つに対応させることである<sup>12)</sup>。この概念をクラスと呼ぶ。

$l$  個の学習されたデータが与えられているとする。各学習データは、特徴ベクトル  $x_i \in R^n$ 、 $i = 1, \dots, l$  と、それに割り当てられたクラス  $y_i \in -1, 1$  の組からなる。 $x_i$  の各要素は計測された各特徴量を表す。さらに、関数  $f: R^n \rightarrow R$  が次の条件を満たすものとする。

$$y_i = 1 \text{ のとき } f(x_i) > 0$$

$$y_i = -1 \text{ のとき } f(x_i) < 0$$

$f(x) = 0$  を満たす点の集合は 2 つのクラスの境界面をなす。このような  $f$  を識別関数と呼ぶ。

パターン認識は 2 つの段階からなる。まずクラスが既知の学習データから識別関数  $f$  を求め (学習段階)、次にクラスが未知のデータを  $f$  の正負により 2 つのクラスのどちらかに分類する (識別段階)。新たなデータ  $x$  に対してクラス  $y$  は  $y = \text{sgn}[f(x)]$  で推定される。ここで  $\text{sgn}$  は

$$\text{sgn}[\alpha] = \begin{cases} 1 & \alpha \geq 0 \text{ のとき} \\ -1 & \alpha < 0 \text{ のとき} \end{cases}$$

である。未知のデータを正しく分類した割合 (識別率) が高い識別関数を求めることがパターン認識の目標である。

### 5.2 サポートベクターマシンの原理

SVM は 2 クラスの分類を行う機械学習の一種で、与えられた学習データの中でサポートベクトルと呼ばれるクラス境界近傍に位置する学習データと、識別面との距離であるマージンを最大化するよ

うに分離超平面を構築し，クラス分類を行う．線形で分類が難しい場合，カーネルトリックによって入力空間を有限，あるいは無限次元の特徴空間へ写像し，特徴空間上で線形分離を行うことで非線形の問題にも適用が可能である．これらの工夫から，従来パターン認識の分野で多く用いられていた逆誤差伝播法によって学習を行う多層パーセプトロンと比べ高い汎用性を持ち，2次の凸計画問題として定式化されるため学習の結果として最適解を得ることができるという特徴を持つ．

### 5.2.1 分離超平面の決定

線形分離可能な正負2つのクラス ( $y_i \in \{-1, 1\}$ ) に分けられた  $l$  個の学習データ  $x_i \in \mathbb{R}^n, i = 1, \dots, l$  を超平面で分離する場合を考える．求める超平面の法線ベクトルを  $w$ ，原点との垂直な距離を  $\frac{|b|}{\|w\|}$  とすると， $\mathbb{R}^n$  の超平面は，

$$w^T x + b = 0 \quad (5.1)$$

と表すことができる．式 (5.1) に対する各学習データ  $x_i (i = 1, \dots, l)$  のマージンを

$$\gamma_i = \frac{y_i(w^T x_i + b)}{\|w\|}$$

と定義する．超平面により  $i$  番目の学習データが正しく識別されている場合には  $\gamma_i > 0$  となり，その値はデータを表す点と超平面との距離に等しい．また， $\gamma_i$  のうち最小のものを  $\gamma$  とし，学習データ集合に対する超平面のマージンと呼ぶ．SVM は，学習データを完全に分離する超平面のうち，学習データ集合に対するマージン  $\gamma$  を最大にする超平面を識別面とする．Fig. 5.1 に，SVM の概念図を示す．

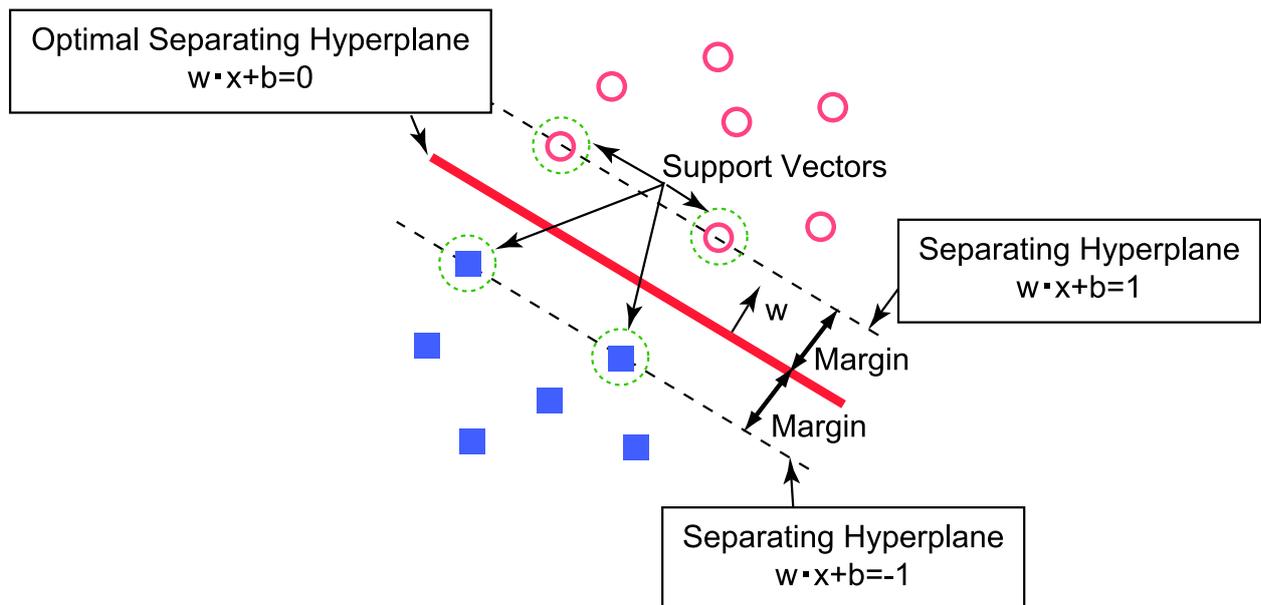


Fig. 5.1 Support vector machine

パラメータ  $w, b$  を定数倍しても表現する超平面は同じであるという冗長性を取り除くために，

$$\min_{i=1, \dots, l} |w^T x_i + b| = 1 \quad (5.2)$$

という制約条件を加えると，マージン  $\gamma$  の最大化は以下のように定式化される．

$$\begin{aligned} & \text{minimize} && \mathbf{w}^T \mathbf{w} \\ & \text{subject to} && y_i(\mathbf{w}^T \mathbf{x}_i + b) \geq 1 \quad (i = 1, \dots, l) \end{aligned}$$

この制約条件付きの最適化問題は，ラグランジュ未定乗数  $\alpha_i (\geq 0)$  を用いて，以下に示すラグランジアン(Lagrangian)の鞍点を求める問題に置き換えることができる．

$$L(\mathbf{w}, b, \alpha) = \frac{1}{2} \mathbf{w}^T \mathbf{w} - \sum_{i=1}^l \alpha_i \{y_i(\mathbf{w}^T \mathbf{x}_i + b) - 1\} \quad (5.3)$$

鞍点において  $L$  の勾配が 0 になることから，

$$\frac{\partial L}{\partial \mathbf{w}} = \mathbf{w} - \sum_{i=1}^l y_i \alpha_i \mathbf{x}_i = 0 \quad (5.4)$$

$$\frac{\partial L}{\partial b} = - \sum_{i=1}^l y_i \alpha_i = 0 \quad (5.5)$$

が得られる．これらを式 (5.3) に代入して整理すると， $L$  が  $\alpha$  に関して鞍点で最大値をとることから，SVM の学習は二次計画問題に帰着する．

$$\begin{aligned} & \text{maximize} && \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i,j=1}^l y_i y_j \alpha_i \alpha_j \mathbf{x}_i^T \mathbf{x}_j, \\ & \text{subject to} && \sum_{i=1}^l y_i \alpha_i = 0, \quad \alpha_i \geq 0, \quad (i = 1, \dots, l) \end{aligned}$$

したがって，局所解の問題は存在せずに，最適化計算によって超平面が決定される．

また，最適解は式 (5.6) の相補性条件を満たさなくてはならない．

$$\alpha_i \{y_i(\mathbf{w}^T \mathbf{x}_i + b) - 1\} = 0, \quad (i = 1, \dots, l) \quad (5.6)$$

これは，式 (5.2) より，超平面に最も近い学習データに対してのみ  $\alpha_i \neq 0$  となることを示している．式 (5.4) より，超平面はサポートベクトルと呼ばれる超平面に最も近い学習データのみで決定される．

SVM では，分類の性能はサポートベクトルの数によって特徴づけられる．その結果，学習データ数と比較して特徴空間の次元数が大きい場合でも，高い汎化性能を持つことが期待できる．

## 6 色ヒストグラムインタセクション

### 6.1 色ヒストグラム

色ヒストグラムとは，ある画像の持つ色彩情報を RGB や HSV といった色空間に変換後，色彩情報の割合を示したものである．たいていの場合，色彩情報の数値を任意の区間 (ヒストグラム・ビン) によって区切り，その区間内に存在する画素をカウントされ，任意の区間毎での割合をヒストグラムとして利用する．色ヒストグラムの例を Fig. 6.1 に示す．

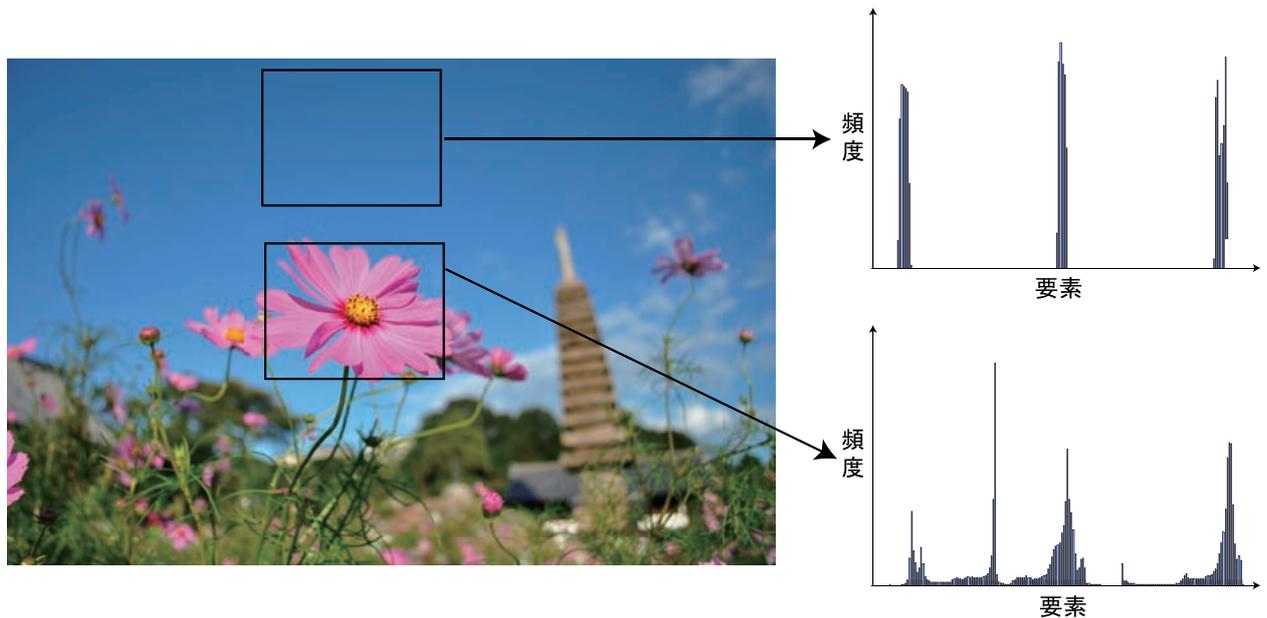


Fig. 6.1 色ヒストグラムの場合

## 6.2 色ヒストグラムインタセクション

色ヒストグラムインタセクションは，SWAINらによって提案した色ヒストグラムを利用した画像照合法である<sup>8)</sup>。照合するためのモデル画像と照合対象となる画像について，それぞれの色ヒストグラムを求めることができる。それぞれについて， $n$ 個のヒストグラム・ビンを持つヒストグラム  $I$  と  $M$  が与えられたとき，2つのヒストグラムの交点 (Intersection)  $S_{IM}$  は式 (6.1) のように定義される。

$$S_{IM} = \sum_{j=1}^n \min(I_j, M_j) \quad (6.1)$$

この時，交点  $S_{IM}$  はヒストグラムを領域の総ピクセル数に伴い，上限が変化するため，式 (6.2) のように従い正規化を行う。

$$H(I, M) = \frac{\sum_{j=1}^n \min(I_j, M_j)}{\sum_{j=1}^n M_j} \quad (6.2)$$

正規化により， $H(I, M)$  は 0~1 の範囲をとり，領域 A と領域 B の類似度を表す値となる。

色ヒストグラムインタセクションは対象とする領域の色の構成割合だけに基づいて類似度計算を行うため，画像を構成する色が一致していてもその構成割合が異なると類似度は低くなる。しかしながら，本研究で利用する病理画像は，症例によって異なるテクスチャを持っているが，同じ症例では数パターンのテクスチャに限定されることが視認されている。従って，色の構成要素によってそれらのテクスチャを識別することは容易と考えられる。

## 7 提案戦略の実装

本章では、提案戦略の実装について述べる。なお、本章で掲載する病理画像は、東北大学加齢医学研究所病態臓器構築研究分野福本学教授より提供されたものである。

### 7.1 SVM を利用した大域領域における病変部位抽出識別器の生成

本節では、SVM を利用した大域領域における病変部位抽出識別器の生成についての説明を行う。以後、本節の手順で生成された識別器を第一識別器と呼ぶ。第一識別器の生成手順を以下に示す。

**STEP.1** 学習用画像データの入力画像及び抽出目標画像を  $m \times m$  ピクセルのセルに切り分け、両画像について同じ座標にあるセルについて対応付けを行う。

**STEP.2** 切り分けた入力画像のセル毎に特徴量の計測を行う。

**STEP.3** 閾値  $T_1 (0 < T_1 \leq 1)$  を基に抽出対象領域についてのラベル付けを行う。抽出目標画像のセル毎に、マーキングされたピクセル数を計測し、セル内でマーキングされたピクセル数の割合が  $T_1$  を超える場合、対応づけられた入力画像のセルを抽出対象領域とする。すなわち、抽出対象領域には 1、非抽出対象領域には  $-1$  のラベルを付ける。

**STEP.4** 各セルの特徴量とラベルを学習データセットとし、SVM にて識別器の学習を行う。

セルサイズを 10px、学習画像データに Fig. 4.2 を用い、上記のフローで構築した第一識別器を原画像に適用した一例を Fig. 7.1 に示す。

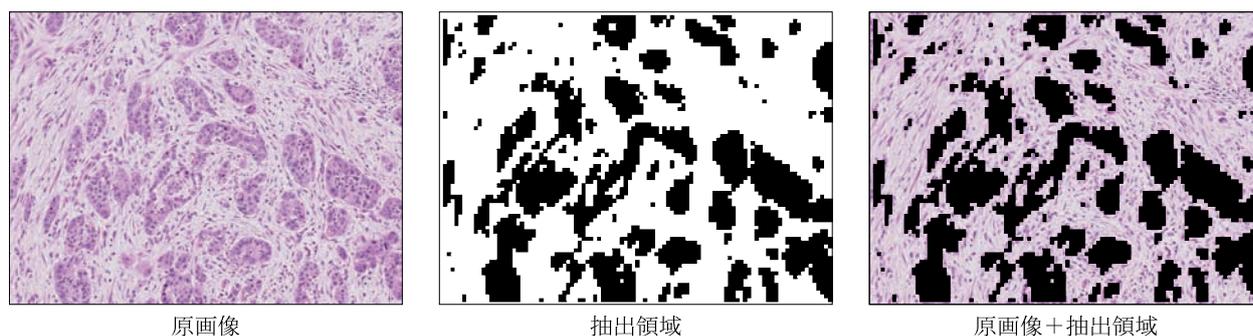


Fig. 7.1 セルサイズ 10px で生成した第一識別器の適用例

Fig. 7.1 よりわかるように、セルサイズ  $m$  を 10px にした場合、病変部位の大まかな特定が可能であることがわかる。また、セルサイズを 2px に変更し同様の手順を用いて識別器を生成し、原画像に適用した例を Fig. 7.2 に示す。

Fig. 7.2 より、セルサイズを小さくすると病変部位の境界付近での抽出について正確に行えていることが分かるが、病変部位以外での誤抽出が多く、また正常に抽出できた病変部位内に過剰な分割が発生している。そのため、病変部位抽出に適当とは言えない。

本論文で利用した口腔癌病理画像の 10 倍率バーチャルスライド画像については、2px のセルサイズより 10px のセルサイズが妥当であった。しかし、セルサイズは対象とする画像に応じて適当なサ

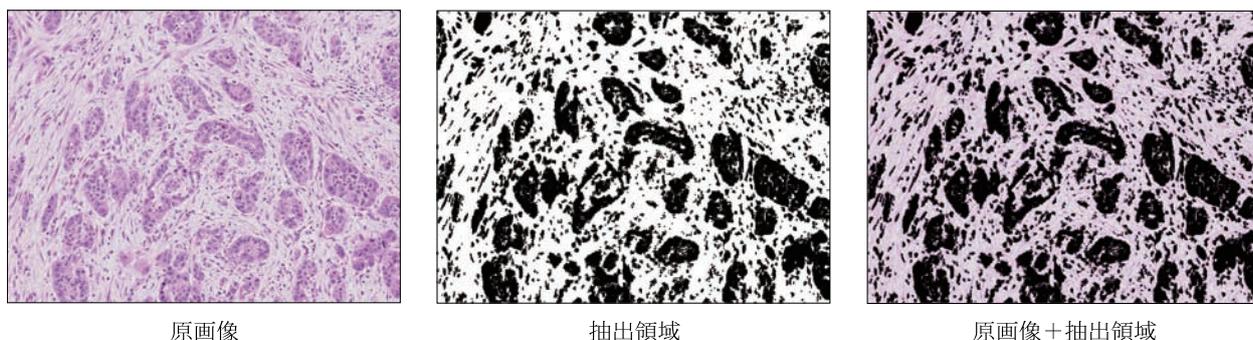


Fig. 7.2 セルサイズ 2px で生成した第一識別器の適用例

イズが異なると考えられる．従って，対象とする病理画像に応じて，セルサイズの検証は必要となると考えられる．

## 7.2 SVM を利用した局所領域における病変部位抽出識別器の生成

本節では，SVM を利用した局所領域における病変部位抽出識別器の生成についての説明を行う．以後，本節の手順で生成された識別器を第二識別器と呼ぶ．7.1 節で説明したように，第一識別器による病変部位の判定では，セルサイズを大きくとることによって，病変部位の大まかな特定が可能であった．よって，その結果で得られた病理対象部位について，再び病理の学習を行い，更に高精度な識別器の生成を考える．

まず，はじめにセルサイズ 10px で作成した第一識別器の精度について検証を行う．第一識別器のを学習画像に適応した結果を Fig. 7.3 に示す．

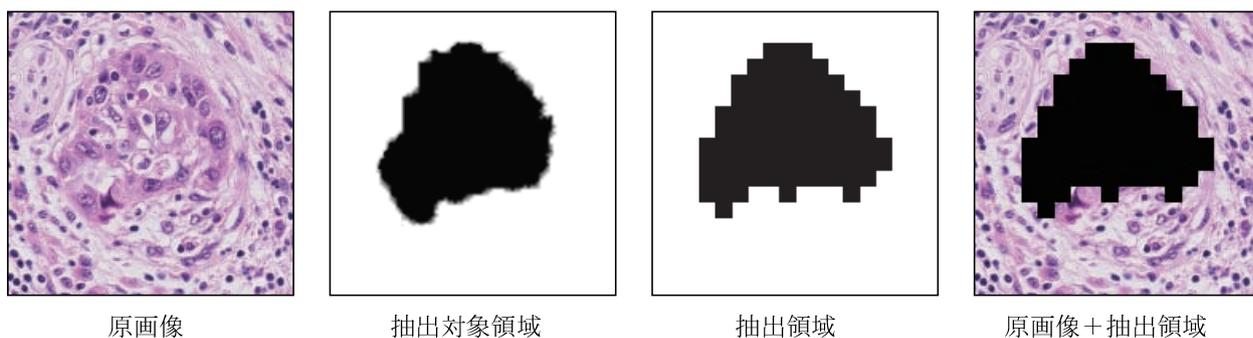


Fig. 7.3 セルサイズ 10px で生成した第一識別器の精度

Fig. 7.3 に示したように，第一識別器では抽出した領域の外郭部分について，わずかながら病変部位の抽出を行えていないことがわかる．そのため，この領域のみを学習データとして利用するのではなく，抽出した領域より一セル分外郭を含めた領域について学習用データとして扱う．

第二識別器の生成手順を以下に示す．なお，下記の手順でのマスク画像とは，7.1 節で作成した識別器により得た結果の領域より一セル分外郭を含めた領域をマーキングした画像である．

**STEP.1** 学習用画像セットの入力画像及び抽出目標画像を  $n \times n$  ピクセルのセルに切り分ける．同じ座標にあるセルについて対応付けを行う．この時，セルサイズは，7.1 節の識別器生成時のセルサイズより小さいものを利用する．

**STEP.2** 切り分けた入力画像のセル毎に特徴量の計測を行う．

**STEP.3** 閾値  $T_2(0 < T_2 \leq 1)$  を基に抽出対象領域についてのラベル付けを行う．抽出目標画像のセル毎に，マーキングされたピクセル数を計測し，セル内でマーキングされたピクセル数の割合が閾値を超える場合，対応づけられた入力画像のセルを抽出対象領域とする．すなわち，抽出対象領域には 1，抽出対象領域には -1 のラベルを付ける．

**STEP.4** 各セルの特徴量とラベルを学習データセットとする．

**STEP.5** マスク画像を基に，学習データセットの選択を行う．STEP.3 と同様に，閾値  $T_3(0 < T_3 \leq 1)$  を設定し，マスク画像のセル毎に，マーキングされたピクセル数を計測し，セル内でマーキングされたピクセル数の割合が  $T_3$  を超える場合，学習データセットとして選択する．超えない場合は，学習データセットより除去する．

**STEP.6** STEP.5 にて選択された学習データセットを利用し，SVM にて識別器の学習を行う．

セルサイズ  $n$  を 2px，学習画像データに Fig. 4.2 を用い，上記のフローで構築した第二識別器を原画像に適用した一例を Fig. 7.4 に示す．

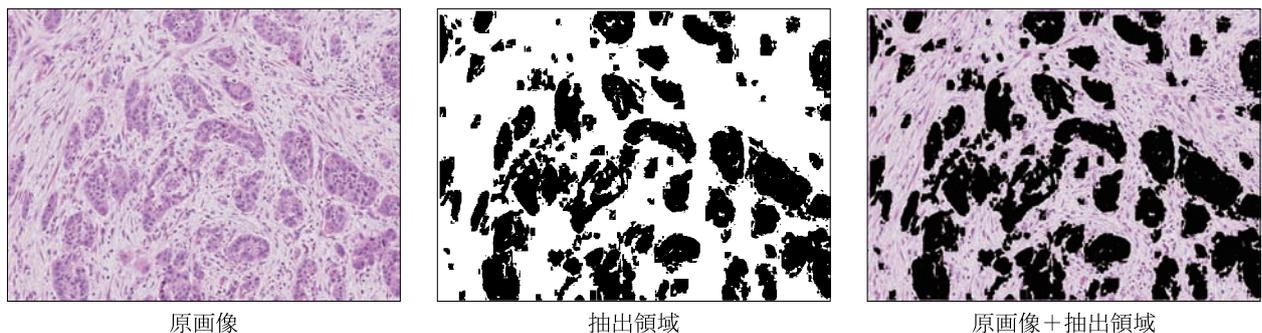


Fig. 7.4 セルサイズ 2px で生成した第二識別器の適用例

Fig. 7.4 に示したように，誤検出が低下し，且つ，病変部位の境界付近の抽出と病変部位内での過剰な分割の防止ができていることがわかる．

### 7.3 色ヒストグラムインタセクションを利用した病変部位判定

第一識別器と第二識別器によって病変部位抽出が行われた出力画像の一例を Fig. 7.5 に示す．

Fig. 7.5 より分かるように，抽出された領域の中には明らかに病変部位でない領域が含まれている．本節では，領域毎に色ヒストグラムを比較することで最終的な病変部位の判定を行うことを提案する．色ヒストグラムによる病変部位判定の手順を以下に示す．

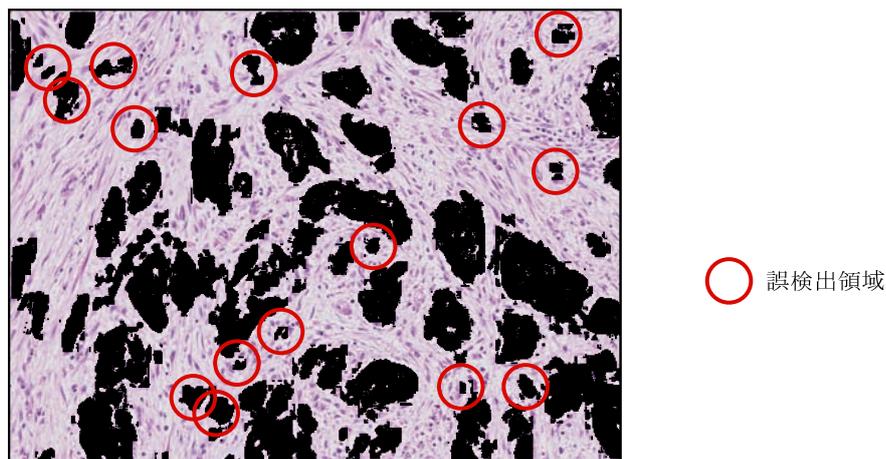


Fig. 7.5 第一識別器と第二識別器による病変部位抽出の例

STEP.1 学習用画像データの入力画像より，病変部位（抽出目標画像にてマーキングされている領域）についての色ヒストグラム  $M$  を計測する．

STEP.2 第一識別器と第二識別器によって抽出を行った病変部位について，連続する領域毎にヒストグラム  $I_j (0 \leq j < \text{連続する領域数})$  を計測する．

STEP.3 色ヒストグラムインタセクションにより  $I_j (0 \leq j < \text{連続する領域数})$  と  $M$  との類似度を計測する．

STEP.4 閾値  $T_4 (0 < T_4 \leq 1)$  を設定し，類似度が  $T_4$  を超える領域について病変部位と判定する．

なお，式 (6.2) では類似度の計測を行う領域を同サイズのものとして正規化を行っているが，本手法では異なる大きさの領域についても類似度の計測を行う為，以下のように式の変換を行った．

$$H(I_j, M) = \sum_{k=1}^n \min\left(\frac{I_{jk}}{\sum_{j=1}^n I_{jk}}, \frac{M_k}{\sum_{j=1}^n M_j}\right) \quad (7.1)$$

学習画像データに Fig. 4.2 を用い，上記のフローにより類似度による病変部位判定を例を Fig. 7.6 に示す．

Fig. 7.6 の場合， $T_4$  の値が 8.0 ~ 8.25 の時，病変部位の判定が正確に行われていると考えられる．

## 8 提案戦略の有効性の検証

本章では，提案戦略の有効性の検証を行う為に，まず特徴量について検証し，戦略に用いるべき特徴量を決める．また，その特徴量をもとに提案戦略を適用し有効性の検証を行う．なお，本論文で取り扱う問題について利用する SVM は C-SVM とし，線形カーネルを用いる．また，線形カーネルのパラメータである  $C$  は 1.0 とする．また，検証に利用した共通のパラメータを，Table 8.1 に示す．

なお，本章で掲載する病理画像は，東北大学加齢医学研究所病態臓器構築研究分野福本学教授より提供されたものである．

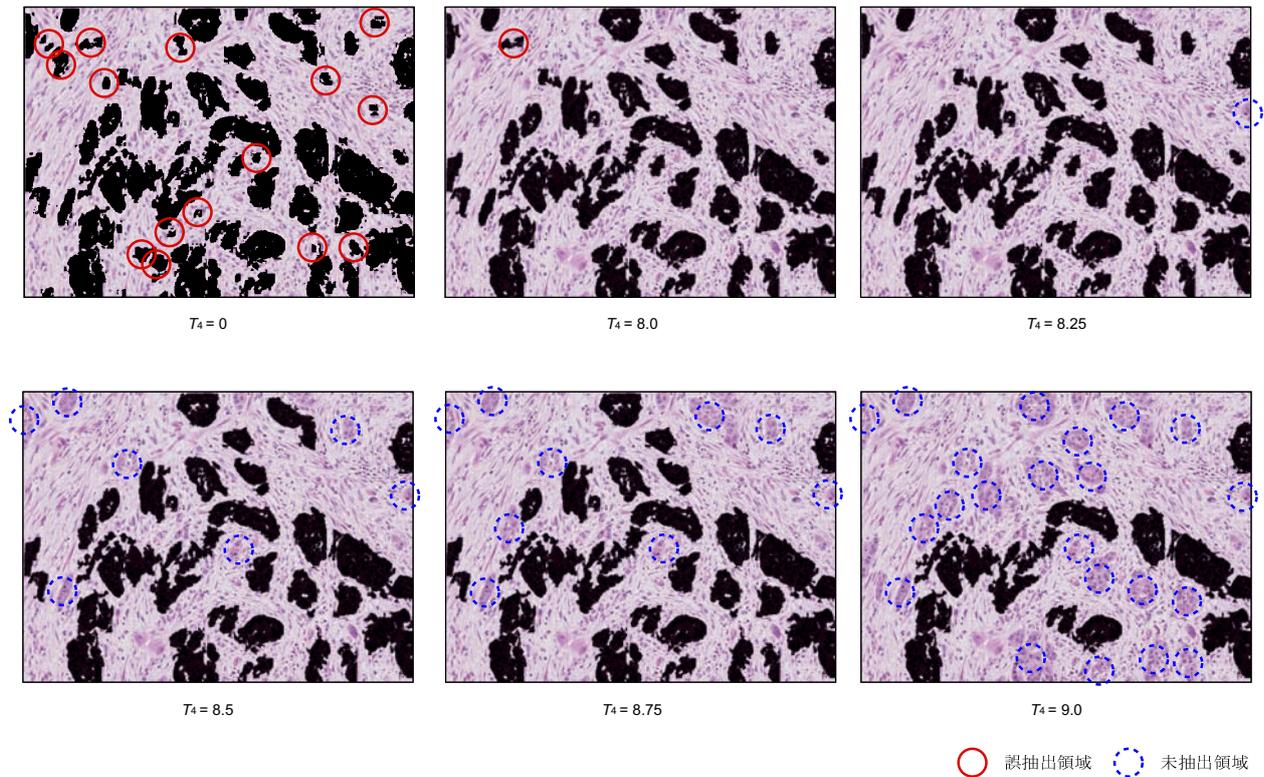


Fig. 7.6 類似度による病変部位判定

Table 8.1 検証で利用する共通パラメータ

$m$	10px
$n$	2px
$T_1$	0.25
$T_2$	0.25
$T_3$	0.25
$T_4$	0.8

## 8.1 特徴量の検討

本節では、第一識別器、第二識別器生成時に利用する特徴量の検証を行う。利用する特徴量の検討として、以下の三つの特徴量について比較を行う。

- 特徴量 1:注目領域と近傍領域の各領域の R,G,B 値の平均値よりなる 27 ベクトルの特徴量
- 特徴量 2:注目領域と近傍領域の各領域の H,S,V 値の平均値よりなる 27 ベクトルの特徴量
- 特徴量 3:注目領域と近傍領域の各領域のグレースケール値の平均値よりなる 9 ベクトルの特徴量

学習用画像として Fig. 8.1 を用い、作成した識別器の精度を検証するために Fig. 8.2 を用いた。各特徴量を基に第一識別器、第二識別器を作成し、両識別器を Fig. 8.2 に適用した結果を Fig. 8.3 に、またそれらについて評価を行った結果を Table 8.2 に示す。

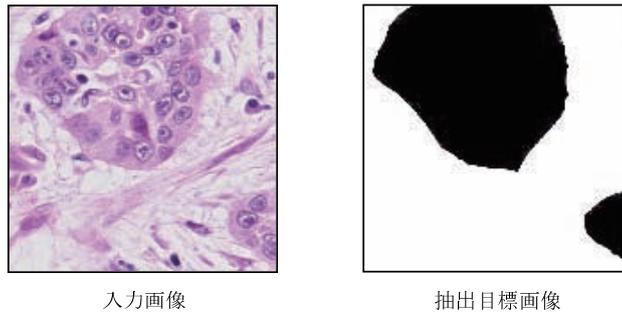


Fig. 8.1 学習用画像

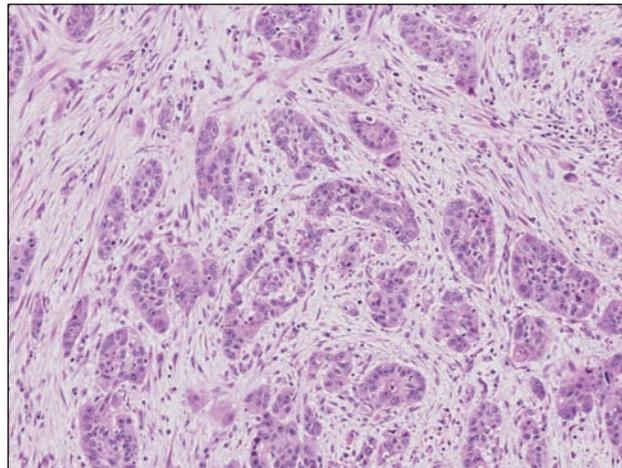


Fig. 8.2 識別器適用画像

Table 8.2 異なる色空間の特徴値を利用した識別器の精度

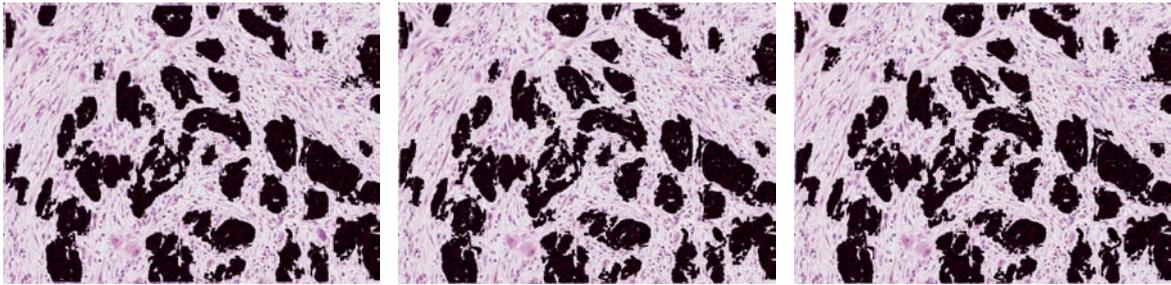
	$TPF$	$FPF$
特徴量 1	0.9310	0.1918
特徴量 2	0.9342	0.2125
特徴量 3	0.9418	0.2356

Fig. 8.3 より、どの特徴量についても病変部位の検出が可能であることが視覚的に確認できる。また、Table 8.2 より、 $TPF$  はどの特徴量もほぼ同等であり、 $FPF$  はわずかながら特徴量 1 が良い。従って、以降の検証では特徴量 1 を用いることとする。

次に、近傍情報利用の有効性の検証を行う為、以下の特徴量について検証を行う。

- 特徴量 1:注目領域と近傍領域の各領域の R,G,B 値の平均値よりなる 27 ベクトルの特徴量
- 特徴量 4:注目領域の R,G,B 値の平均値よりなる 3 ベクトルの特徴量

なお、前述の検証と同様に、学習用画像として Fig. 8.1 を用い、作成した識別器の精度を検証するために Fig. 8.2 を用いた。各特徴量を基に第一識別器、第二識別器を作成し、両識別器を Fig. 8.2 に



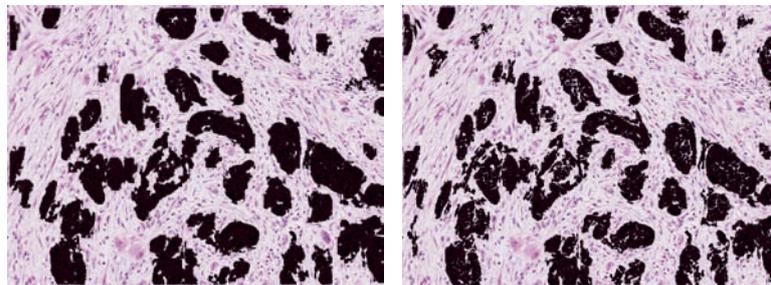
(a) 特徴量 1

(b) 特徴量 2

(c) 特徴量 3

Fig. 8.3 異なる色空間の特徴値を利用した識別器の適用結果

適用した結果を Fig. 8.4 に、またそれらについて評価を行った結果を Table 8.3 に示す。



(a) 特徴量 1

(b) 特徴量 4

Fig. 8.4 近傍情報を利用した識別器の適用結果

Table 8.3 近傍情報利用の有無による識別器の精度

	$TPF$	$FPF$
特徴量 1	0.9310	0.1918
特徴量 4	0.9040	0.2010

Table 8.3 より、近傍情報を利用することにより、 $TPF$  値が高く、そして  $FPF$  値が低くなっていることがわかり、近傍情報利用が有効に働いていると考えられる。

## 8.2 同一画像における有効性の検証

本節では、提案手法の学習性能の検証として、識別器生成に用いた学習用画像の原画像への適用を行う。検証を行うデータとして、同一検体（検体 A）より得た 10 枚の病理画像（付図を参照）を利用した。それぞれの画像における評価値を Table 8.4 に示す。

Table 8.4 同一画像より生成した識別器の精度

学習用画像	$TPF$	$FPF$
A-1	0.9310	0.1918
A-2	0.8949	0.1607
A-3	0.8677	0.1653
A-4	0.9310	0.1918
A-5	0.9197	0.2022
A-6	0.8306	0.1941
A-7	0.9105	0.4008
A-8	0.8532	0.2025
A-9	0.9265	0.3368
A-10	0.9150	0.3708

Table 8.4 より，識別器を学習用画像の原画像へ適用した場合，高い  $TPF$  値を得ることが分かった．しかしながら，A-7 や A-9，A-10 においては  $FPF$  値も高くなっている．A-7，A-9，A-10 での適用結果画像を Fig. 8.5 に示す．

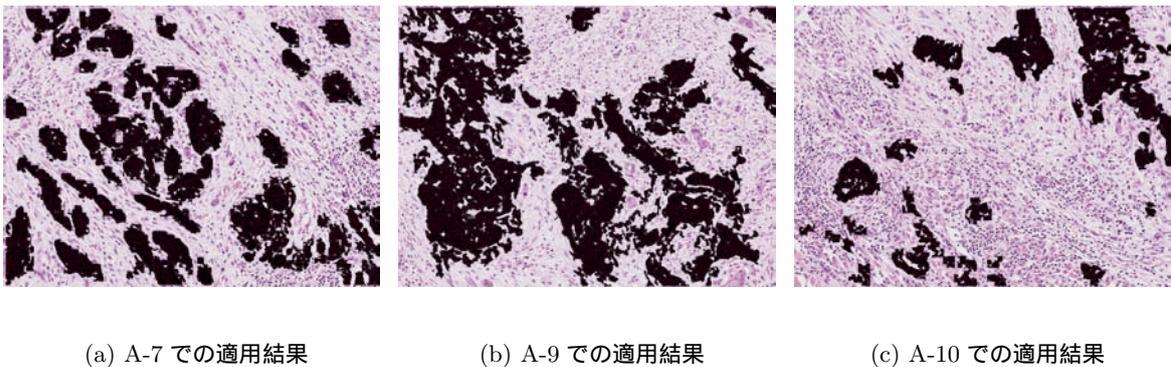
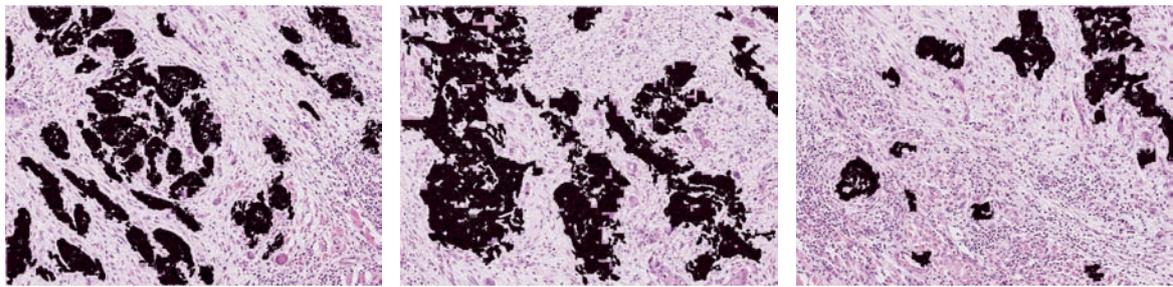


Fig. 8.5 同一画像より生成した識別器の適用結果

Fig. 8.5 より，病変部位の抽出を行えているにも関わらず，他の領域にて誤検出を行っていることが分かる．これは，学習用画像が病変部位抽出の目的に特化したものであり，病変部位ではない部分における病変部位判断の学習には役に立っていなかったことを示す．SVM ではポジティブデータセット (本手法において 1 のラベル付けを行われた学習データ) とネガティブデータセット (本手法において -1 のラベル付けを行われた学習データ) を必要とするが，これまで与えてきた学習用画像では，ポジティブデータセットについては概ね良好に与えることができていたのに対し，ネガティブデータセットについて与えることができていなかったと考えられる．そこで，これまででは病変部位とその周辺についての学習用画像を基に識別器を生成していたが，病変部位でない部分を学習画像に含むことによる精度向上について検討した．A-7 や A-9，A-10 において，病変部位でない部分を学習画像に含

み，識別器を生成し，その識別器をそれぞれの画像に適用した結果を Fig. 8.6 に，それらの評価値を Table 8.5 に示す．



(a) A-7 での適用結果

(b) A-9 での適用結果

(c) A-10 での適用結果

Fig. 8.6 ネガティブデータセットを考慮した識別器の適用結果

Table 8.5 ネガティブデータセットを考慮した識別器の精度

学習用画像	$TPF$	$FPF$
A-7	0.9116	0.2819
A-9	0.8422	0.2480
A-10	0.8362	0.2134

Fig. 8.6 と Fig. 8.5 を比較すると誤検出が低下していることが確認できる．また Table 8.5 より，病変部位でない部分を学習画像に含むことによってより精度の高い識別器を生成できていることがわかる．

### 8.3 同一検体における有効性の検証

本節では，提案手法の学習性能の検証として，8.2 節で作成した識別器を同一検体の画像への適用を行う．検証を行うデータとして，8.2 節と同様に，同一検体 (検体 A) より得た 10 枚の画像を利用した．それぞれの画像より作成した識別器を他の 9 枚の画像に適用した評価値の平均値と中央値を Table 8.6 に示す．

Table 8.6 より，A-1，A-3，A-4，A-6，A-7，A-10 の画像を基に生成した識別器は概ね良好な結果となったが，その他の識別器については有効でないことが分かる．

従って，任意の画像より切り出した学習用画像より得た識別器を同一検体に適用する場合，その結果は学習用画像によって大きく左右されることが予想される．この問題を考慮する為に，複数のポジティブデータ用の画像とネガティブデータ用の画像より生成した識別器を適用した．識別器生成に利用した学習用画像を Fig. 8.7 に，適用結果を Table 8.7 示す．

Table 8.7 に示したように，複数のポジティブデータ用の画像とネガティブデータ用の画像より生成した識別器は，異なる画像に対し，一定の精度での抽出が行えることが分かる．しかし，A-2，A-8

Table 8.6 異なる画像より生成した識別器の精度

学習用画像	$TPF$		$FPF$	
	$TPF_{ave}$	$TPF_{med}$	$FPF_{ave}$	$FPF_{med}$
A-1	0.8135	0.8281	0.3206	0.2865
A-2	0.3583	0.3635	0.1463	0.1334
A-3	0.8107	0.8360	0.3640	0.3076
A-4	0.8976	0.9279	0.3722	0.3281
A-5	0.6407	0.7207	0.1876	0.1292
A-6	0.9066	0.9177	0.3578	0.3394
A-7	0.8895	0.9237	0.4310	0.3860
A-8	0.5311	0.5474	0.1366	0.1391
A-9	0.4467	0.4931	0.4734	0.4539
A-10	0.8418	0.8760	0.2527	0.2072

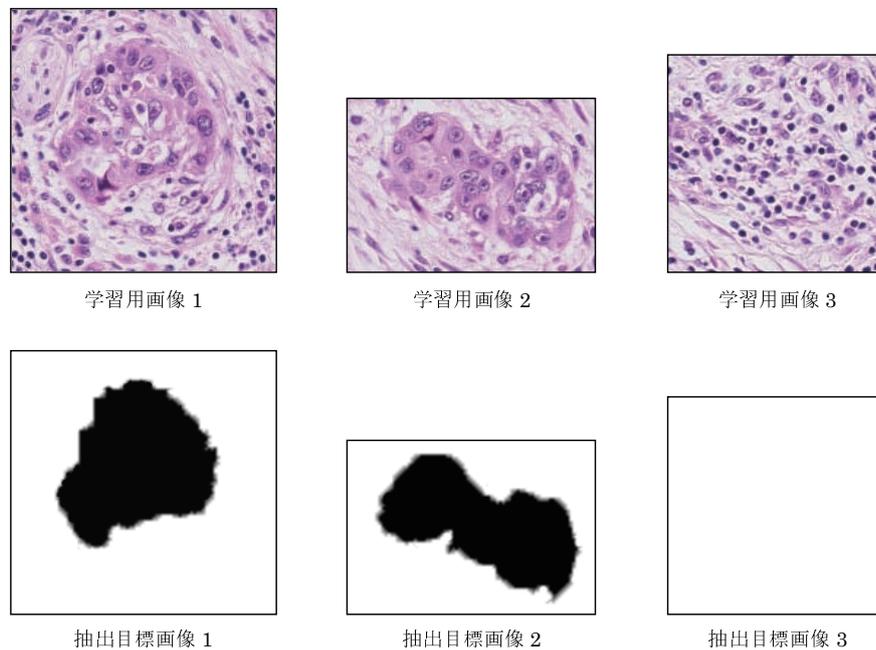


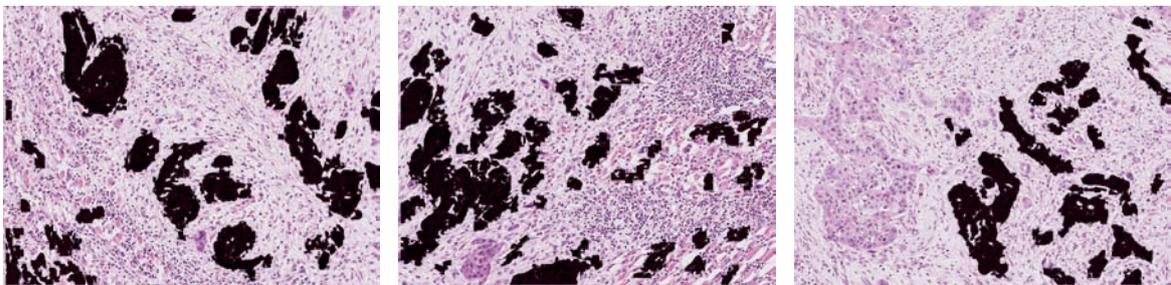
Fig. 8.7 学習用画像

においては  $FPF$  が高く、A-9 においては  $TPF$  が極端に低いことがわかる。それぞれの抽出結果を Fig. 8.8 に示す。

Fig. 8.8 に示したように、A-2、A-8 では学習用画像には含まれないテクスチャを持った部位でご検出が発生し、A-9 では左側の大きな癌巣が抽出できていない。A-2、A-8 については、8.2 節で示したように、ネガティブデータを考慮することによって誤検出を低下できると考えられる。また、A-9 について、検証作業を行った結果、色ヒストグラムインタセクションを利用した病変部位判定にて用

Table 8.7 複数の学習画像より生成した識別器の精度

適用対象画像	$TPF$	$FPF$
A-1	0.9045	0.2018
A-2	0.9574	0.4315
A-3	0.8899	0.2033
A-4	0.9040	0.1718
A-5	0.9654	0.3140
A-6	0.8014	0.1514
A-7	0.9254	0.2262
A-8	0.8898	0.4855
A-9	0.3755	0.1542
A-10	0.9128	0.3433
平均	0.8526	0.2683
中央値	0.9042	0.2147



(a) A-2 での適用結果

(b) A-8 での適用結果

(c) A-9 での適用結果

Fig. 8.8 複数の学習用画像より生成した識別器の適用結果

いた閾値  $T_4$  の値が適当でないことが分かった．同じ検体内であっても微妙な染色具合の異なりがあり，適用する病理画像毎に閾値を変更する必要があると考えられる．

複数の学習用画像を用いることにより精度の高い識別器の生成を行えたが，今後，学習用画像の組合せや病変判定に用いる閾値  $T_4$  の決定方法について検討の必要がある．

## 9 結論

本論文では，病理医の診断過程をモデル化し，顕微鏡画像における病変部位抽出について検討を行った．病理医の顕微鏡画像を用いた診断は，以下の仮定に基づいていると考えた．

仮定 1 対象画像の倍率を調整することで，病理部位の特定を行う．

仮定 2 極小範囲を見て診断を行うのではなく、注目領域とその近傍領域の特徴量を判断することで、病理部位の特定を行う。

仮定 3 注目領域のテクスチャを見て、病理部位の特定を行う。

仮定 4 病理部位特定は過去の症例より得た知識に依存している。

そして、これらの処理を機械的に実装することにより顕微鏡画像における病変部位抽出が可能であると考え、2階層の識別器生成と病変部位の持つテクスチャの色情報を利用した目標病変部位抽出手法を提案した。

提案手法では、サポートベクターマシン (SVM) により2階層の識別器生成を行い、1階層目の識別器で大域的な病変部位の判定を行い、病変部位と判定された領域に対して2階層目の識別器を適用することにより局所的な領域での病変部位判定を行う。また、2つの識別器によって抽出された連続領域について、色ヒストグラムを求めることにより、抽出対象の持つテクスチャとの類似度を比較し、目標病変部位の抽出を行う。

提案手法の精度や特徴を確認するために、提案手法を口腔癌病理画像における癌胞巣抽出処理に適用した結果、1階層目の識別器により、病変部位の大まかな位置特定が行われ、2階層目の識別器により、病変部位の概形を鮮明にし、誤検出の割合が減少した。そして、色ヒストグラムを用いることにより、更に誤検出を低下することが可能であることが分かった。また、提案手法を学習用画像の原画像及び、同検体の画像に適用した結果、概ね良好な結果を得たが、学習用画像により識別器の精度が大きく左右されることが予想される。これらのことより、学習画像に病変部位周辺の画像だけでなく、意図的に病変部位以外の画像を利用することによって、良い識別器が得られることが分かった。また、学習用画像の原画像と異なる画像について識別器を適応する場合、複数の学習用画像から得られた識別器を利用することによって、良い結果を得ることができた。今後、学習用画像にどのような画像を利用すれば有効な識別器の生成を行えるか、またそのような学習用画像の選択についての自動化を検討する必要がある。

なお、本提案で利用した手法は病変部の抽出を行う画像に対して前処理を行っていない。その為、画像毎の染色具合の異なりにより、識別器の複雑化や精度の悪化が予想される。今後、他検体に対する適用も考慮に入れ、染色具合を調整する為の前処理が必要となると考えられる。

## 謝辞

本研究を遂行するにあたり、多大なる御指導そして御協力を頂きました、同志社大学生命医科学部の廣安知之教授に心より感謝致します。また、様々な指摘、助言をして下さいました、同志社大学理工学部の三木光範教授、吉見真聡助教に心より感謝致します。また、日本電気株式会社の小掠真貴様、東北大学加齢医学研究所病態臓器構築研究分野の福本学教授には本研究に対する指摘、助言及び研究に利用する画像の提供を行って頂きました。心より感謝しております。

そして、本論文を校正してくださいました藤田宗佑さん、富島千歳さんに御礼申し上げます。お忙しい中、ご無理をお願いしたにも関わらず丁寧な校正をしていただき感謝しております。藤田宗佑さんには校正だけでなく、医用画像研究班の一員として、研究ミーティングにおいて多いに議論をして頂きました。同じく、研究ミーティングにて多く議論して頂いた医用画像研究班の野田徹さん、宮部洋太さんに御礼申し上げます。

最後に、知的システムデザイン研究室の皆様には私の研究に関して数多くの議論や助言をして頂きました。また、皆様のおかげで、1年間すばらしい研究生活を送ることができました。この場を借りて厚く御礼申し上げます。

## 参考文献

- 1) Nec : 病理画像解析. <http://www.nec.co.jp/solution/bio/rd/>.
- 2) 英保茂. 医用画像処理. 朝倉書店, 1992.
- 3) 小車顕吾, 高橋正信. 初期肝細胞癌診断支援のための特徴量の考案. 平成 19 年度電子情報通信学会東京支部学生会研究発表会論稿集. 電子情報通信学会東京支部学生会, 2007.
- 4) 人口動態統計 (厚生労働省大臣官房統計情報部編). 厚生統計協会, 2008.
- 5) Gary J. Macfarlane. Rising trends of oral cancer mortality among males worldwide: the return of an old public health problem. *Cancer causes & control*, 1994.
- 6) 天笠光雄, 岡田憲彦, 作田正義, 立川哲彦, 道健一. 口腔癌の早期診断アトラス. 医歯薬出版株式会社, 2008.
- 7) 日本口腔腫瘍学会. 舌癌取り扱い指針 ワーキング・グループ案 (第 1 版). <http://www.jsot.org/t-tumor.pdf>.
- 8) M. J. SWAIN. Color indexing. *Int. J. Comput. Vis.*, Vol. 7, No. 1, pp. 11–32, 1991.
- 9) Corinna Cortes and Vladimir Vapnik. Support-Vector Networks. *Machine Learning*, Vol. 20, No. 3, pp. 273–297, 1995.
- 10) N. Cristianini and J. Shawe-Taylor. *An introduction to Support Vector Machines*. Cambridge University Press, 2000.
- 11) 日本色彩研究所編. デジタル色彩マニュアル, 2004.
- 12) 茨木俊秀, 福島雅夫. 最適化の手法. 情報数学講座 14, 共立出版, 1993.

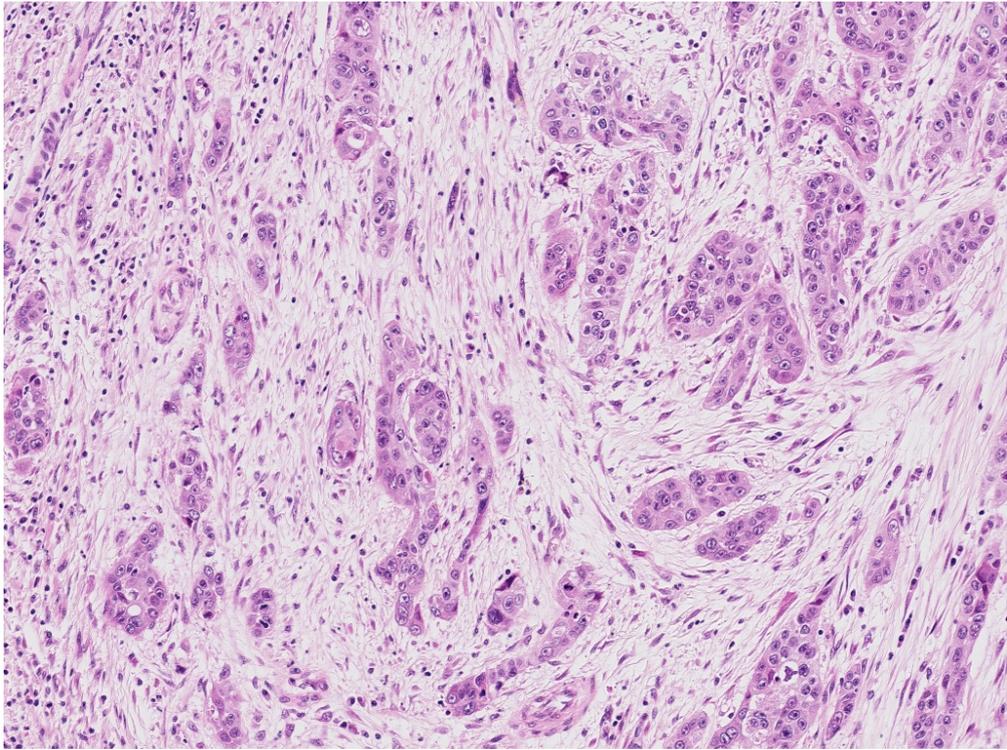
## 付 図

2.1	病理画像の一例	2
2.2	病理画像診断支援システム	3
3.1	口腔癌 (舌癌)	4
4.1	病理画像のセル分割	6
4.2	学習用画像データの一部	7
5.1	Support vector machine	9
6.1	色ヒストグラムの例	11
7.1	セルサイズ 10px で生成した第一識別器の適用例	12
7.2	セルサイズ 2px で生成した第一識別器の適用例	13
7.3	セルサイズ 10px で生成した第一識別器の精度	13
7.4	セルサイズ 2px で生成した第二識別器の適用例	14
7.5	第一識別器と第二識別器による病変部位抽出の例	15
7.6	類似度による病変部位判定	16
8.1	学習用画像	17
8.2	識別器適用画像	17
8.3	異なる色空間の特徴値を利用した識別器の適用結果	18
8.4	近傍情報を利用した識別器の適用結果	18
8.5	同一画像より生成した識別器の適用結果	19
8.6	ネガティブデータセットを考慮した識別器の適用結果	20
8.7	学習用画像	21
8.8	複数の学習用画像より生成した識別器の適用結果	22
1	A-1	1
2	A-2	2
3	A-3	3
4	A-4	4
5	A-5	5
6	A-6	6
7	A-7	7
8	A-8	8
9	A-9	9
10	A-10	10

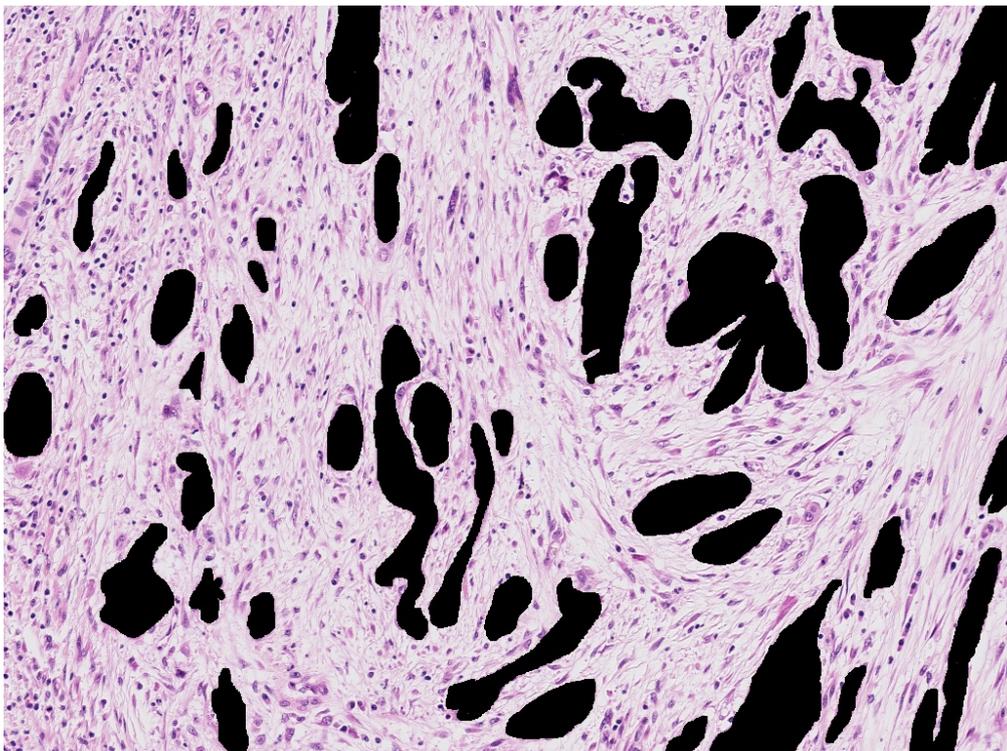
## 付 表

8.1	検証で利用する共通パラメータ	16
8.2	異なる色空間の特徴値を利用した識別器の精度	17

8.3	近傍情報利用の有無による識別器の精度 . . . . .	18
8.4	同一画像より生成した識別器の精度 . . . . .	19
8.5	ネガティブデータセットを考慮した識別器の精度 . . . . .	20
8.6	異なる画像より生成した識別器の精度 . . . . .	21
8.7	複数の学習画像より生成した識別器の精度 . . . . .	22

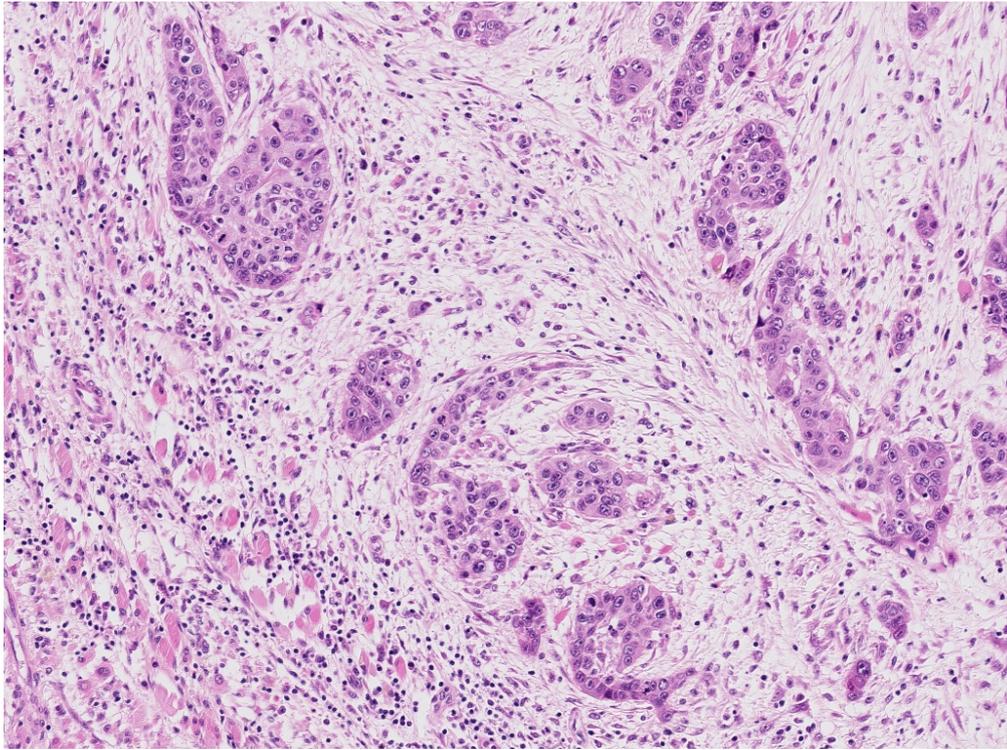


(a) 原画像

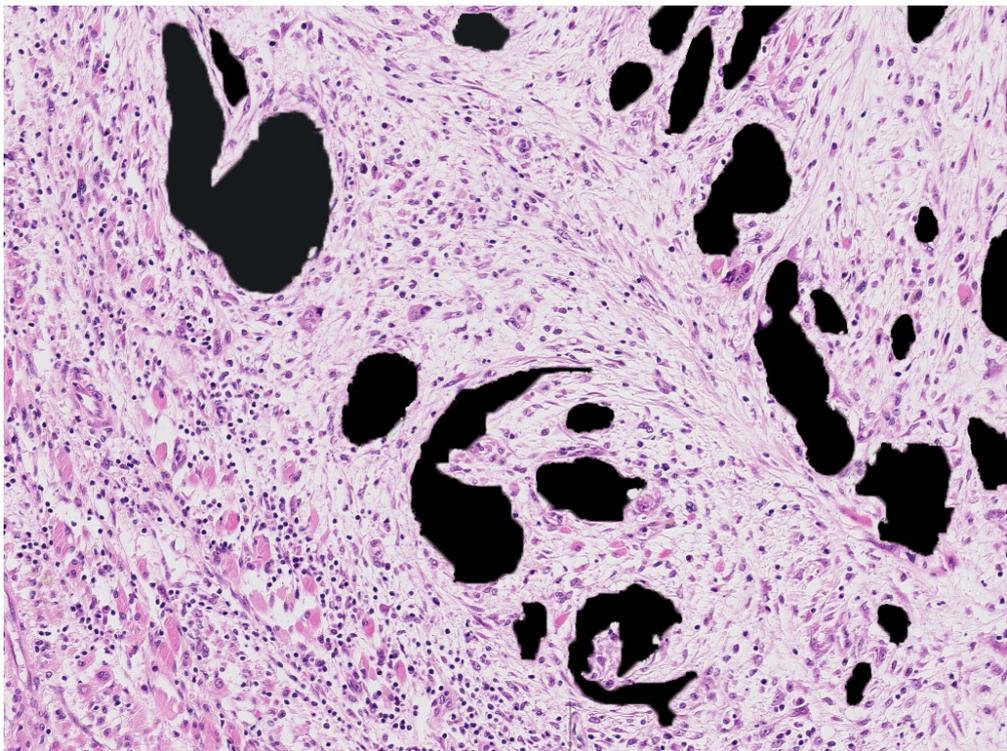


(b) 抽出対象

Fig. 1 A-1

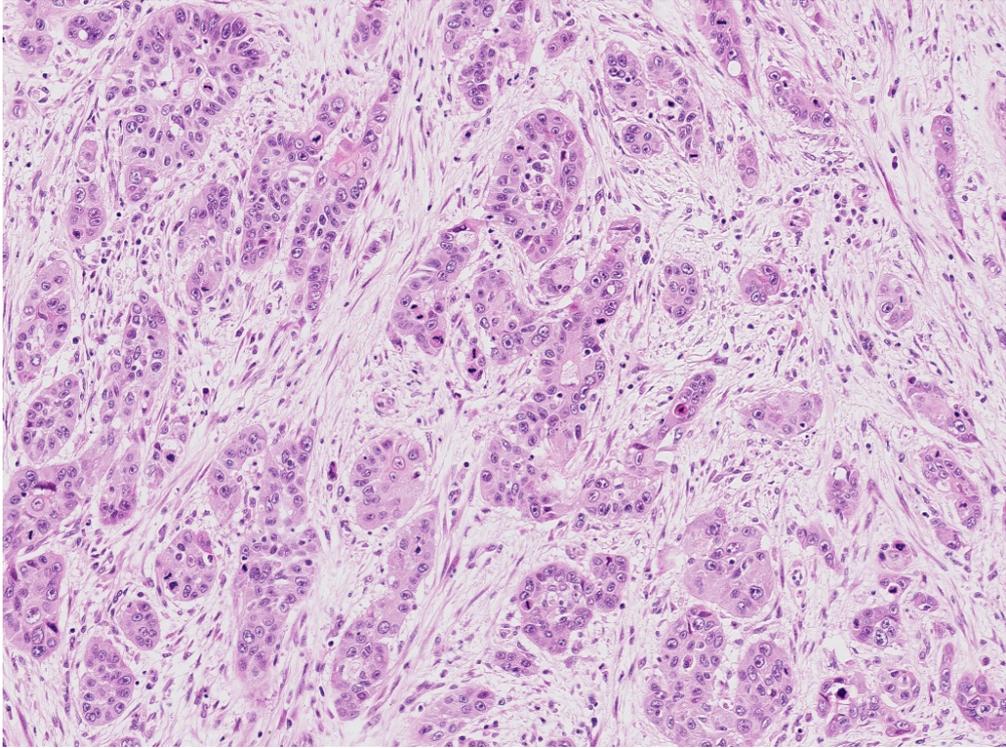


(a) 原画像

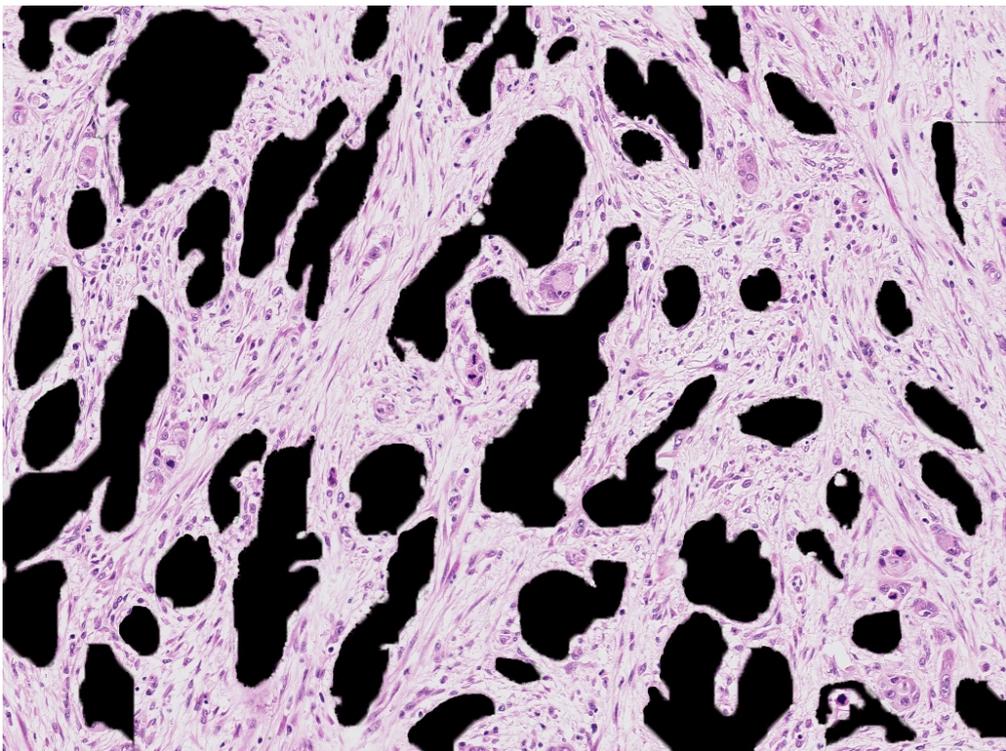


(b) 抽出対象

Fig. 2 A-2

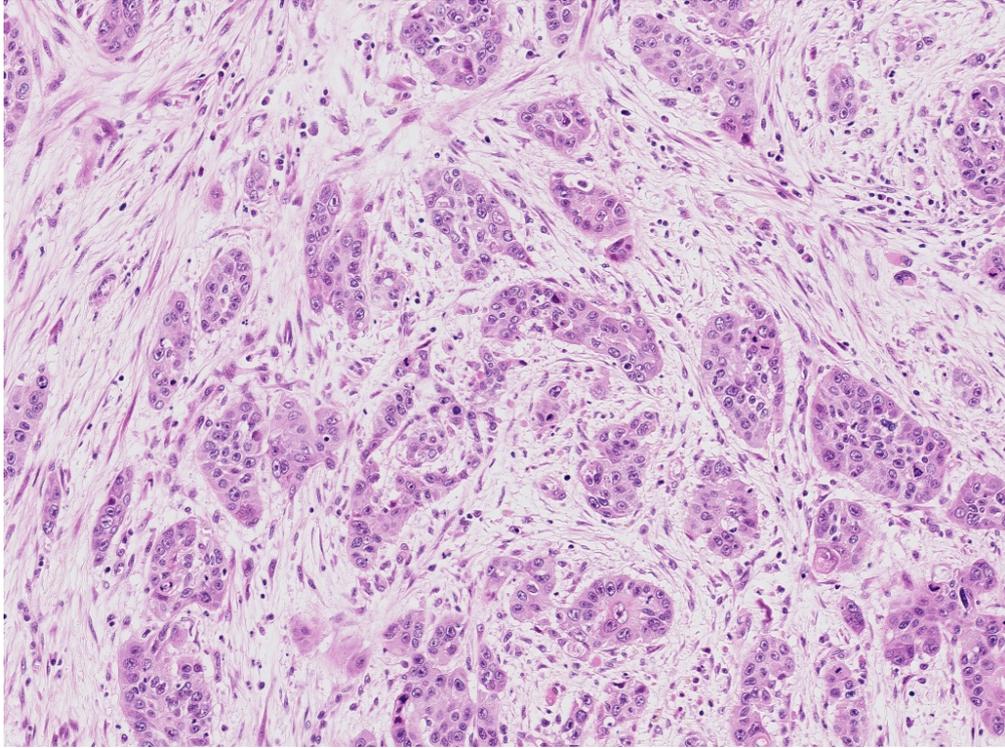


(a) 原画像

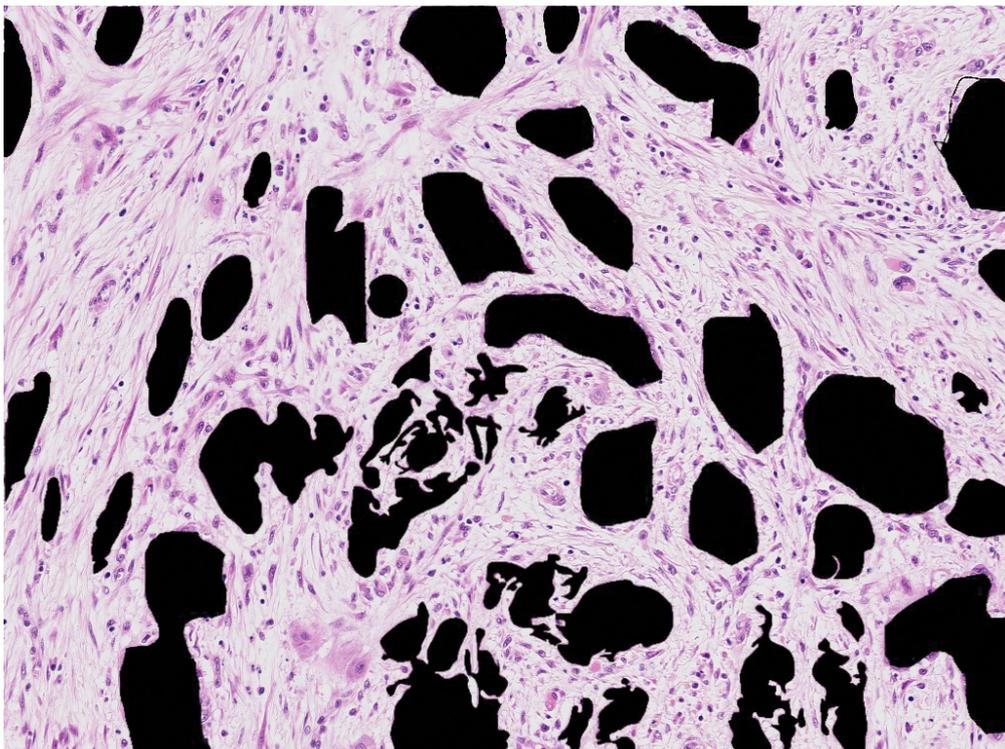


(b) 抽出対象

Fig. 3 A-3

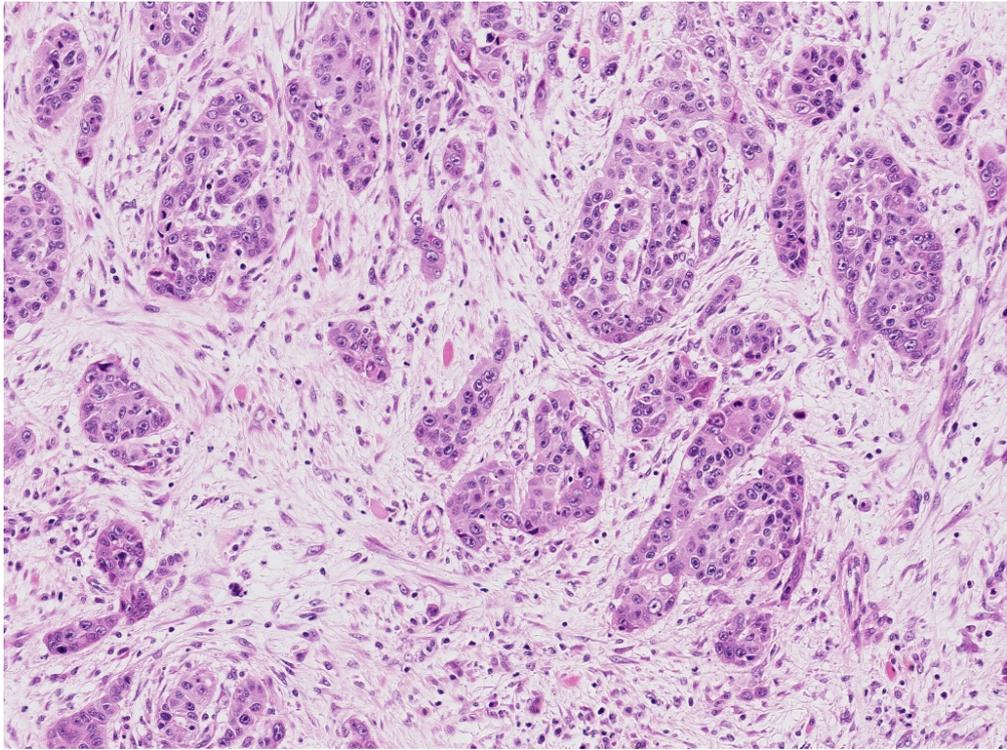


(a) 原画像

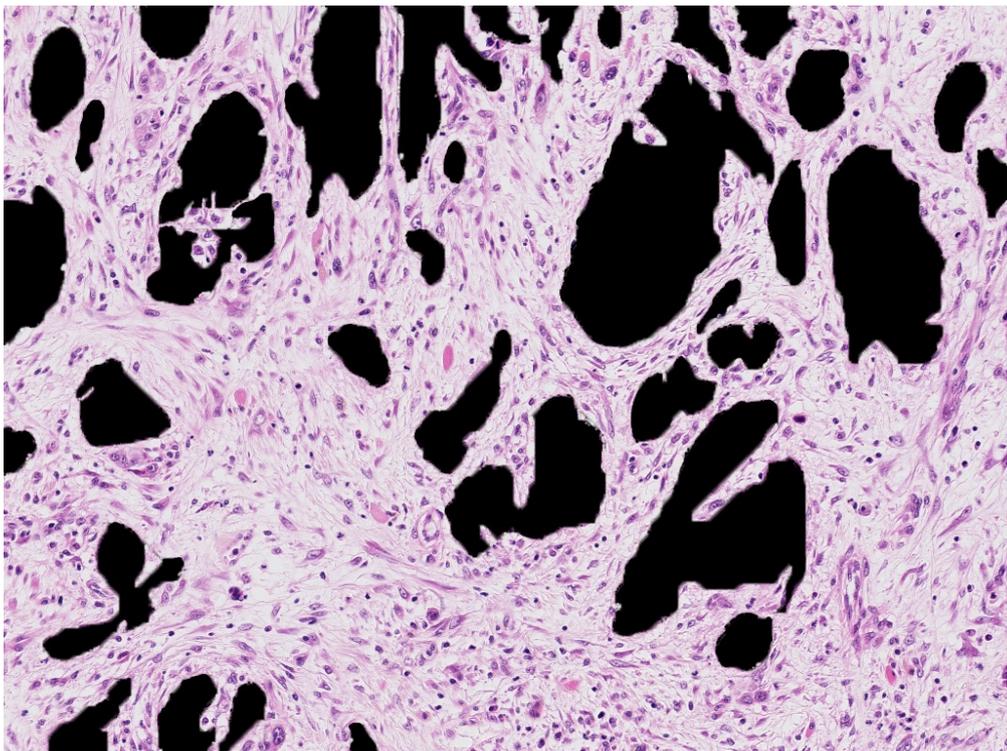


(b) 抽出対象

Fig. 4 A-4

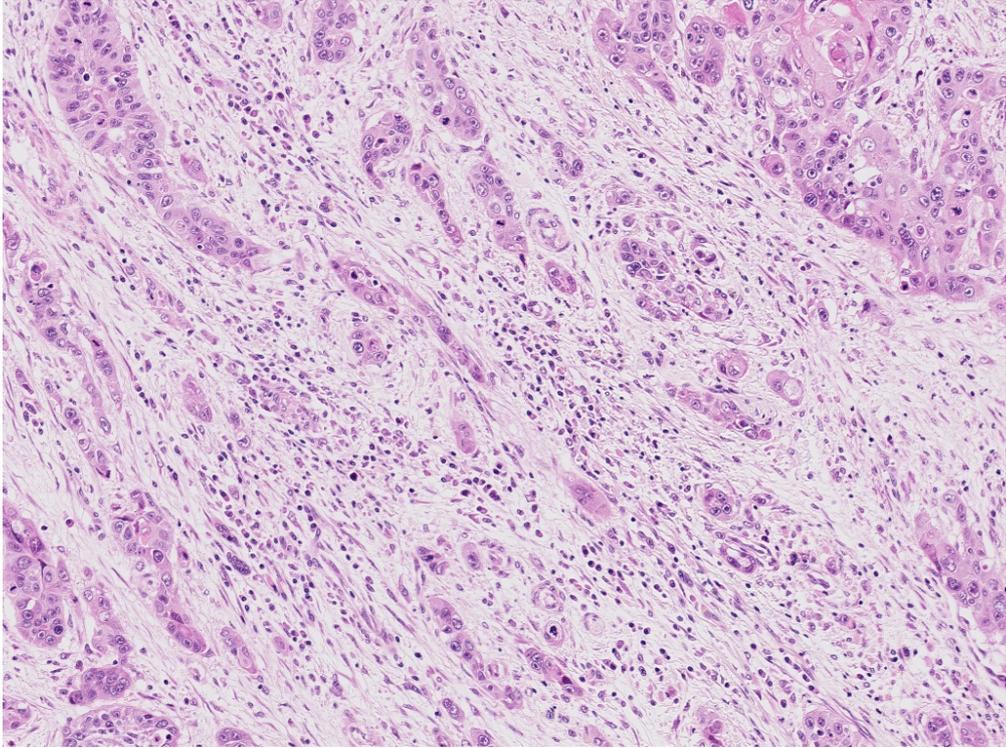


(a) 原画像

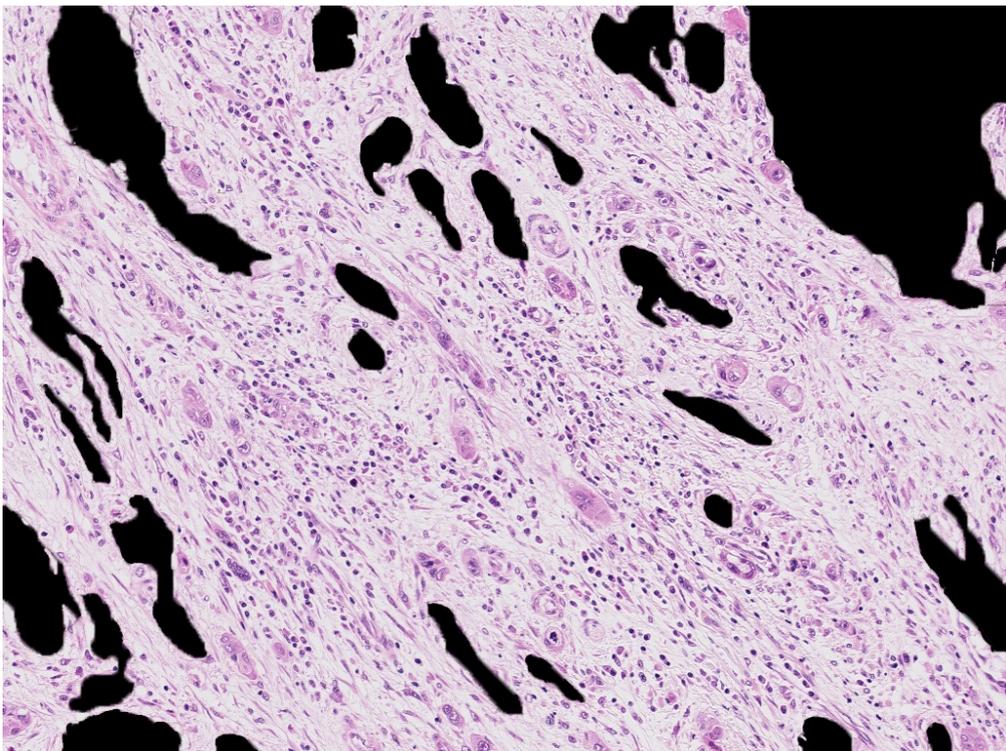


(b) 抽出対象

Fig. 5 A-5

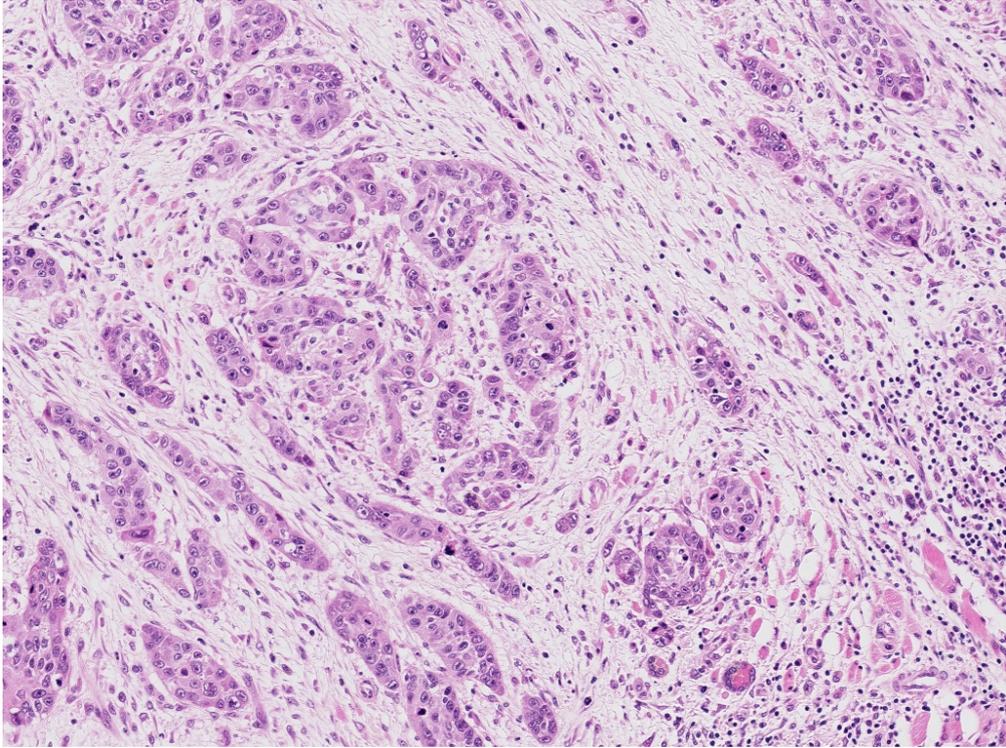


(a) 原画像

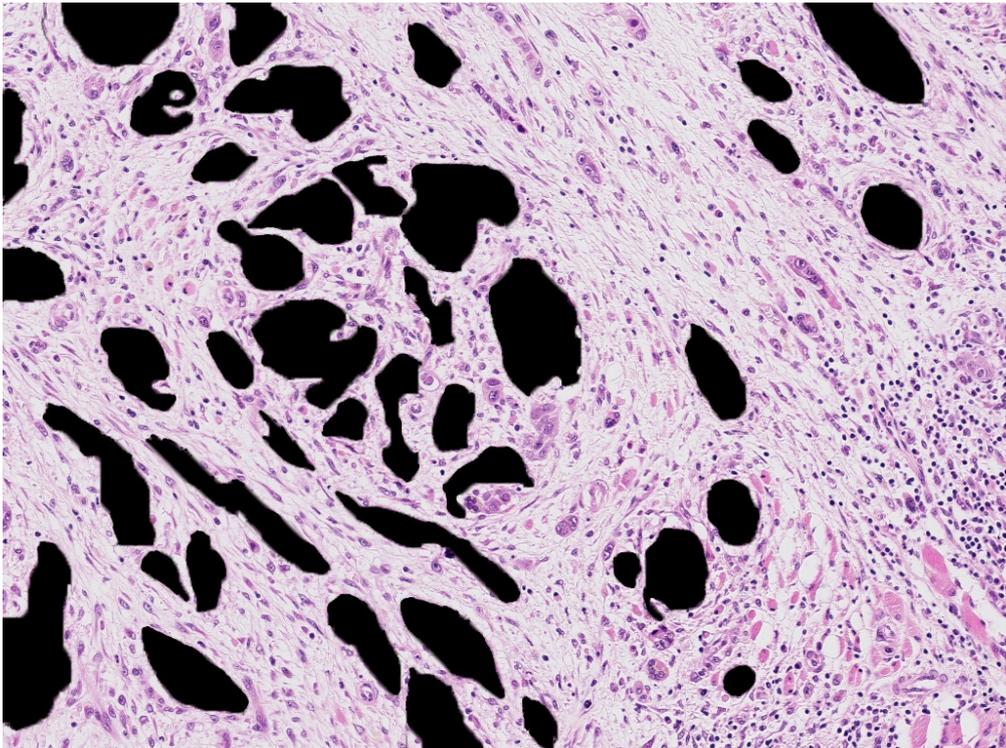


(b) 抽出対象

Fig. 6 A-6

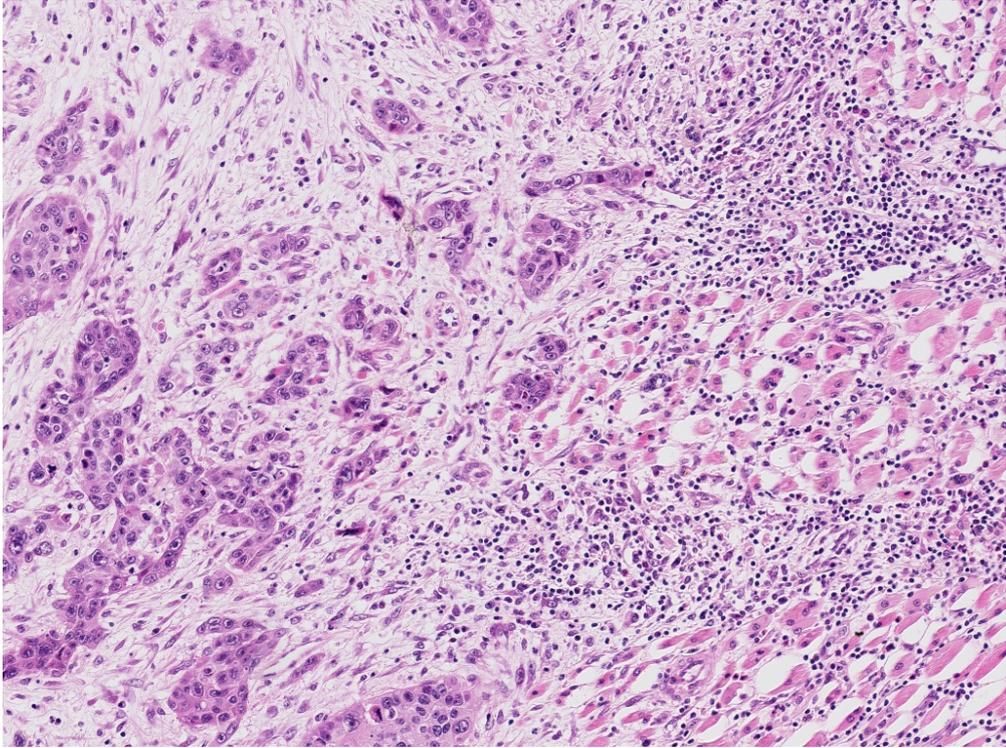


(a) 原画像

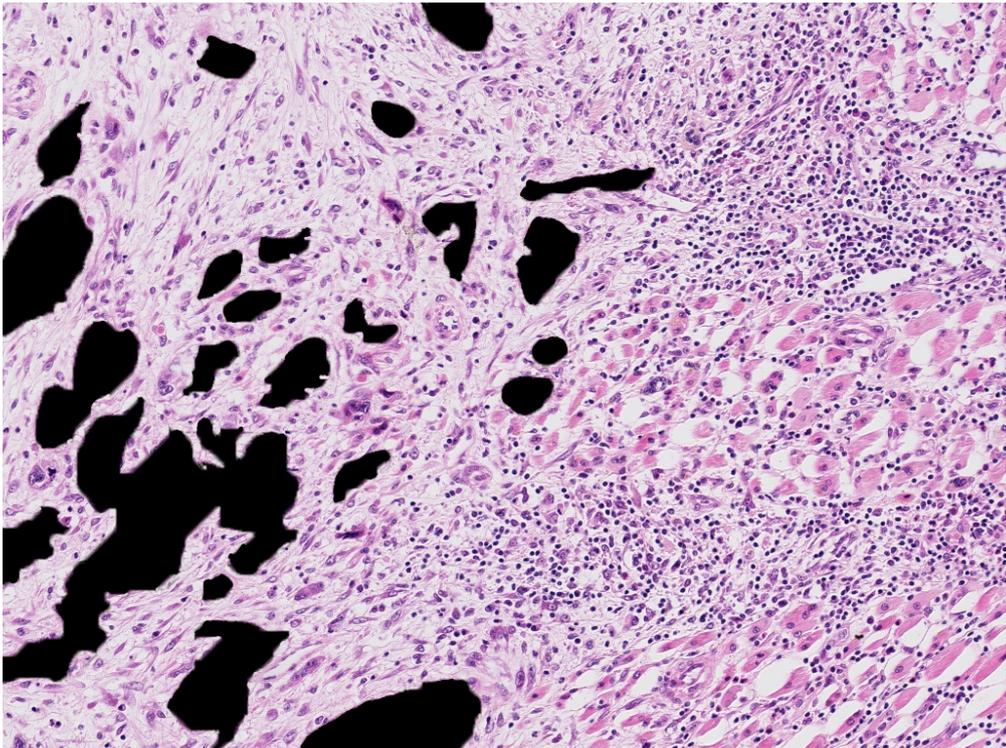


(b) 抽出対象

Fig. 7 A-7

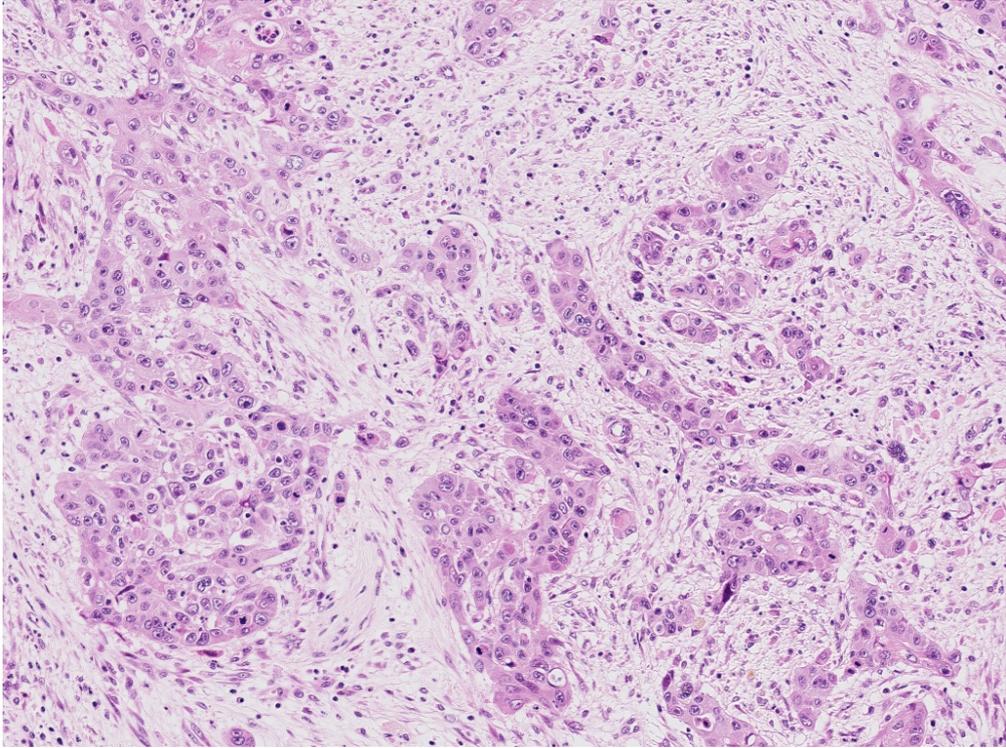


(a) 原画像

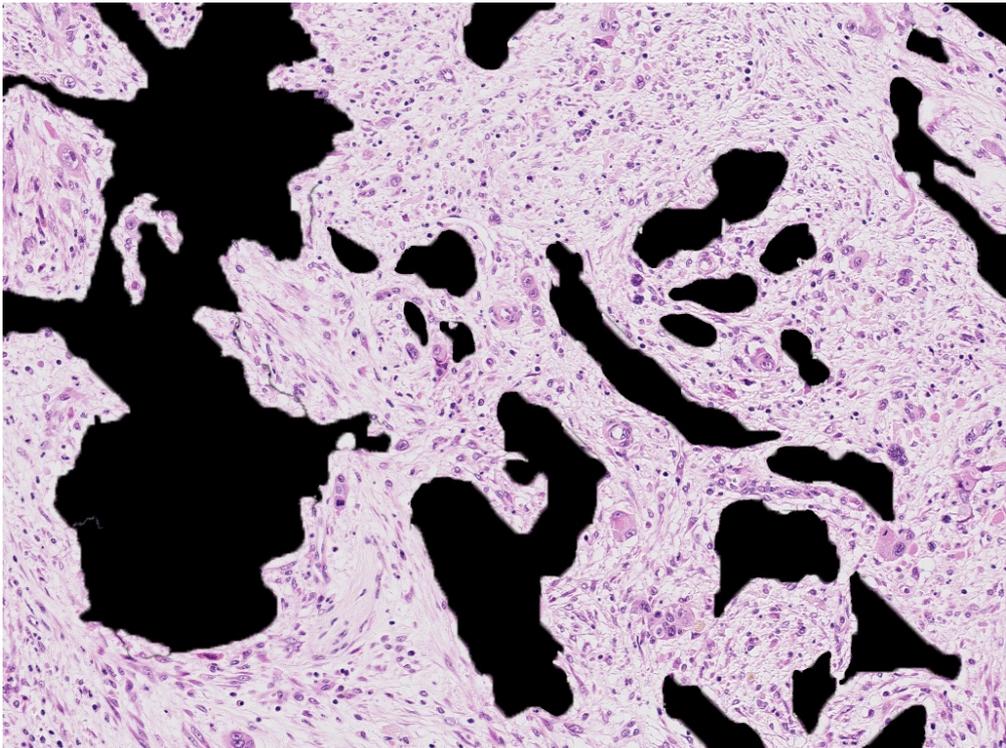


(b) 抽出対象

Fig. 8 A-8

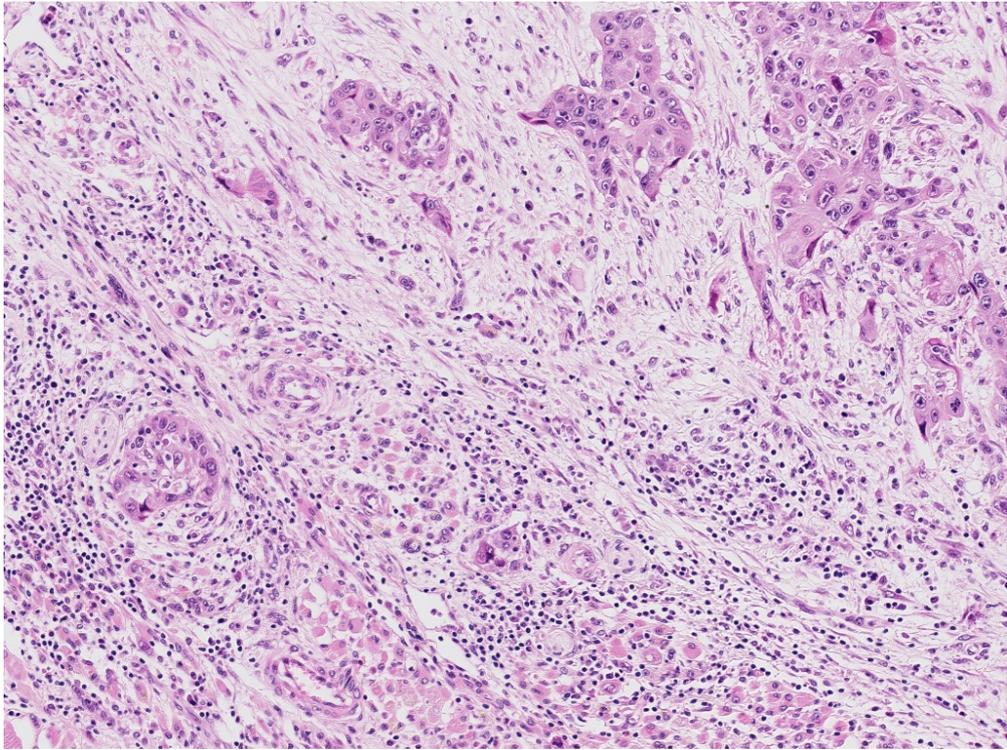


(a) 原画像

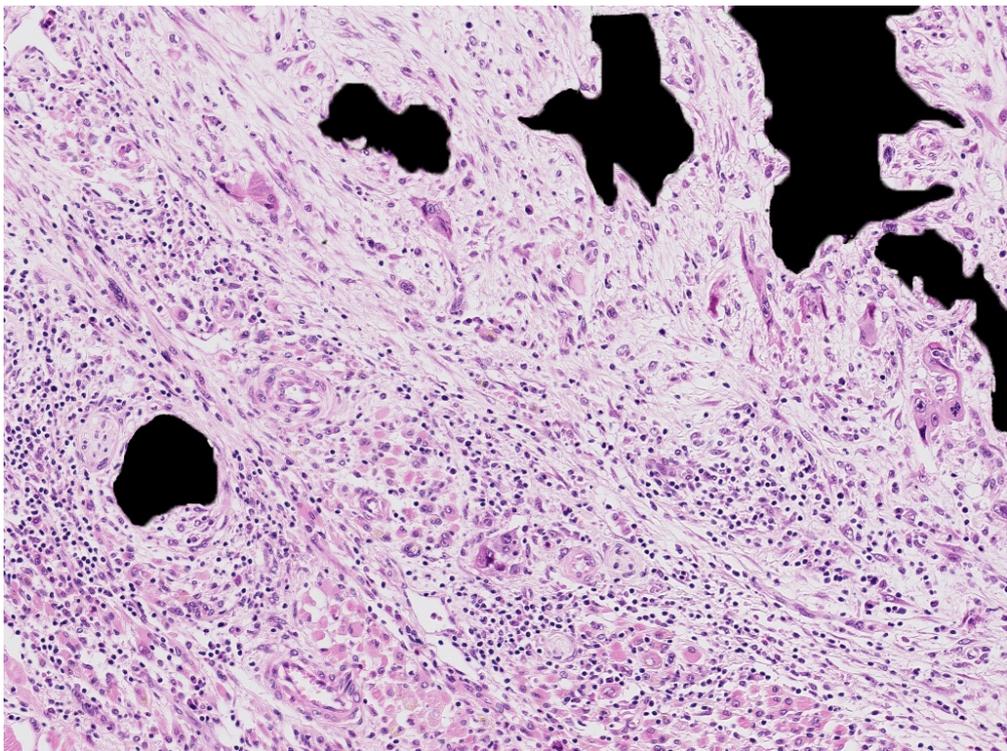


(b) 抽出対象

Fig. 9 A-9



(a) 原画像



(b) 抽出対象

Fig. 10 A-10