

修士論文

癌病理画像における
自動プログラミング手法を用いた
癌胞巣抽出アルゴリズムの検討

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 2009年度 706番

藤田 宗佑

指導教授 三木 光範教授

2011年1月21日

Abstract

In recent years, new diagnostic approaches and treatment techniques for various diseases have been developed. But these methods still require pathologists who check the pathological images on their eyes and make appropriate diagnoses. However, the number of pathologists is not enough. Therefore, there is a demand for pathological image diagnostic support systems that can aid pathologists in checking pathological images. Pathological image diagnostic support system is an system which extract the characteristics of cancer in medical images. And it shows a pathologist the extracted amount of the characteristics and supports diagnosis of a pathologist. However, it is very difficult for pathologists to design the process for extracting the affected region. In addition, most image processing is dependent on an individual problem. Therefore, various processing methods have been applied to find appropriate processing techniques.

In this thesis, automatic constructions of algorithm for image filters that extracts target cancer images from pathological images are discussed. Automatic construction of these filters would be very useful in the medical field. Image processing filters can be expressed as tree topology operations. Genetic Programming (GP) is an evolutionary computation algorithm and one of the methods to design tree topology operations. Simulated Annealing Programming (SAP) is another algorithm to design tree topology operations. These two algorithms, GP and SAP were applied to construct Image Processing Filters and the characteristics of these two algorithms were compared.

The results indicated that SAP requires a lot of iterations to find the optimum solution. Cooling-SAP designed smaller tree topology operations than that of Fixed temperature SAP. On the other hand, GP has strong search capability for finding the optimum solution. GP can also design small tree topology operations by restricting the growth of tree topology operations. As the result, GP shows higher performance than SAP. Then the strategy where the main flow of tree structure is designed by GP and parameters are determined by GA was proposed. Using this strategy, more flexible image filter to extract targeted cancer image can be obtained.

目次

1	序論	1
2	病理診断支援	2
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	病理画像診断支援システムに期待される効果	3
3	口腔癌	4
3.1	口腔癌の現状	4
3.2	口腔癌の組織学的特徴	4
4	自動プログラミング手法による画像処理フィルタの構築	5
4.1	概要	5
4.2	基本フィルタ	6
4.3	木構造状フィルタの評価	7
5	自動プログラミング手法	7
5.1	遺伝的プログラミング (Genetic Programming: GP)	7
5.2	シミュレーテッドアニーリングプログラミング (Simulated Annealing Programming: SAP)	8
6	癌巣抽出アルゴリズムの自動構築	9
6.1	GP におけるプログラムサイズ制限が探索に及ぼす影響	9
6.2	一定温度 SAP における温度パラメータが探索に及ぼす影響	12
6.3	冷却 SAP における温度スケジューリングが探索に及ぼす影響	13
6.4	GP と SAP の比較	15
7	癌巣抽出に対するノイズ除去に関する検討	17
7.1	ノイズ除去処理が探索に及ぼす影響	17
7.2	複数画像探索が探索に及ぼす影響	18
7.3	癌巣抽出のための特徴抽出とパラメータ最適化が探索に及ぼす影響	19
8	結論	23

1 序論

近年、医療の病理診断の分野において、新しい臨床検査法や画像装置の開発、診断精度の向上、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。そして日本では、近年、癌などの増加に伴い、病理医が極端に不足しているのが現状である。そのため、癌などを病理診断するための支援システムに対するニーズが今後増加していくと考えられ、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日本電気株式会社（以下 NEC）で考案されている^{1,2)}。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の特徴量を抽出し、抽出した特徴量を計測後、計測したデータを医師に提示することで、医師が行なう病理診断を支援するシステムである。現在まで、医用画像から病変部位の構成要素を抽出するために、様々なアルゴリズムが考案されている^{1,3)}。

しかし、病理画像診断支援システムに必要とされる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。そこで本研究では、病理画像を用いた画像処理に着目し、病理画像診断支援システムに必要とされる病理部位抽出アルゴリズムの自動構築を行うことを考える。

病理部位抽出のための画像処理アルゴリズムの自動構築を行う手法として、画像処理エキスパートシステム⁴⁻⁶⁾や進化的画像処理^{7,8)}などの手法が提案されている。画像処理エキスパートシステムとは、画像中から抽出したい物体の特徴（線や面など）を直接与えることで過去の履歴や、予め用意された画像処理に関する知識データベースから処理手順やパラメータの決定を行なうものである。一方、進化的画像処理は長尾ら⁹⁾により考案された手法であり、画像処理を組み合わせ最適化問題として捉え、主に遺伝的プログラミング（Genetic Programming: GP）^{10,11)}などの自動プログラミング手法を用いることで画像処理を行うものである。また進化的画像処理は、画像処理エキスパートシステムのような履歴や知識データベースのアプローチとは異なり、最適化を行なうことで画像処理フィルタを構築するため、画像処理の知識やデータベースを必要としない。進化的画像処理はこれまでに、顕微鏡画像から細胞壁抽出⁹⁾や PET 画像から腎臓と肺のセグメンテーション処理¹²⁾など様々な画像処理の課題に適用され、必要な画像処理プロセスが全自動で獲得されることが確認されている。

そこで本研究では、画像処理フィルタの構築を進化的画像処理を用いて行なうこととする。また本研究では、口腔癌病理画像から、癌胞巣を抽出する問題を対象とする。近年、医用画像技術が急速に進展し、各種の病理画像が取得可能となっていると同時に、それらの解析が急務の課題となっている。

進化的画像処理における自動プログラミング手法の解の探索能力は、得られる画像処理フィルタに大きく影響すると考えられる。例えば、プログラムサイズの増加によるブロートは GP において探索効率の低下をもたらす最大の問題である。プログラムサイズを抑制する方法として、藤田らは、金属の焼きなまし過程を模倣したシミュレーテッドアニーリング（Simulated Annealing: SA）¹³⁾を木構造が扱えるように拡張したシミュレーテッドアニーリングプログラミング（Simulated Annealing Programming: SAP）を自動プログラミング手法として提案している¹⁴⁾。しかし、病理画像を対象とした画像処理フィルタの設計において、どのような探索能力を有するのか、GP に対してどのような探索の特性を有するのかについては検討されていない。そのため本稿では、先に述べた病理画像から、病理画像中の癌胞巣を抽出する問題を対象とし、進化的画像処理において GP および SAP の探索の特徴について検討する。また口腔癌における癌胞巣の自動抽出の特性について議論する。

2 病理診断支援

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。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。

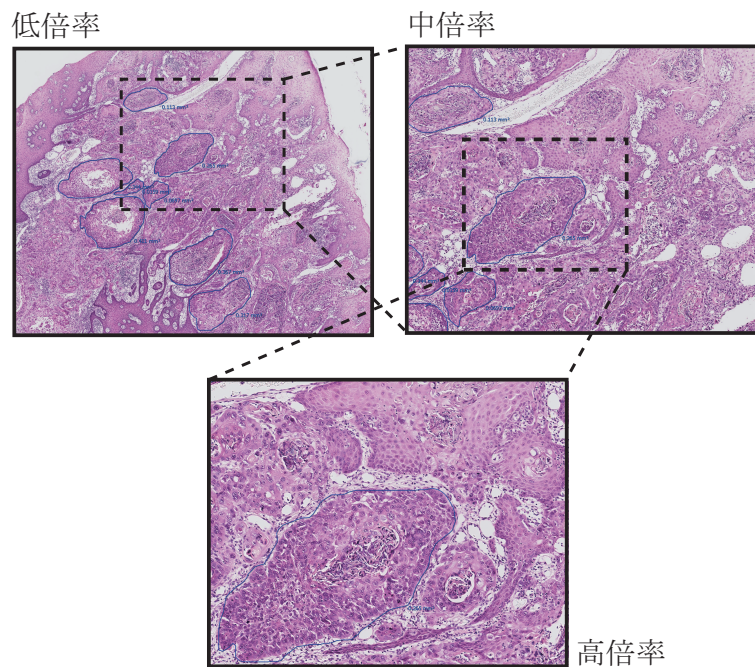


Fig. 2.1 病理画像の例

2.2 病理画像診断支援の概要

一般的に、病理診断を行っている医師（以下、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、組織構造の変化度合や細胞中の核の大きさの比率などを計測する。

しかし、癌の検診が増大している一方で、日本では病理医が極端に不足している。また一般的に、病理診断は病理医によって主観的に行われており、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¹⁵⁾。そのため、経験や体調に影響されない病理診断を支援するシステムに対するニーズが高まってきている。NECでは、医師の病理診断を支援する病理画像診断支援システムを考案している^{1,2)}。

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病変組織の構成要素を抽出し、要素の特徴量を計測することで医師の診断を支援するという一連の行程を自動化したもの

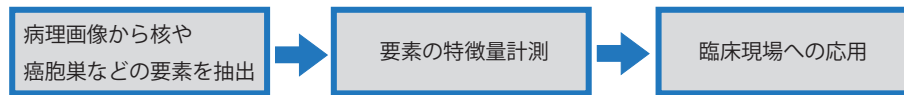


Fig. 2.2 病理画像診断支援システムのフロー

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要素の抽出

病理画像に対して、画像処理を行うことにより、病気の原因である核や癌の構成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

特徴量計測したデータを病理医に提示し、癌などの病気かどうかの判断を促進することで、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を用いた画像処理に着目し、画像処理手法を自動で構築することを考える．

2.3 病理画像診断支援システムに期待される効果

病理画像診断支援システム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．そのため、病理画像の解析を自動化することにより、病理医が行なう診断の支援が可能であると期待されている．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、本システムが病理医の負担の軽減にもつなげると期待されている．

(2) 病理医による病理診断の迅速化

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、病理画像の解析を自動化することにより、病理医による診断を迅速化することが可能であると期待されている．

(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手段が病理医の経験に基づく判断しか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、癌の診断の新たな指標が考案されると期待されている．

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¹⁶⁾。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¹⁷⁾。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)¹

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。例えば術前補助療法（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。

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¹⁸⁾。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

¹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

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¹⁹⁾．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 自動プログラミング手法による画像処理フィルタの構築

4.1 概要

画像処理フィルタの自動構築を行うために，長尾らは進化的画像処理を提案している⁹⁾．進化的画像処理は，一般的に画像処理は既知の単純な画像フィルタ（基本フィルタ）の組み合わせとして表現可能であると仮定して，自動プログラミング手法を用いて最適な基本フィルタの組み合わせを導く手法である．画像処理構築の問題を組み合わせ最適化問題と捉えることで，複雑な処理を行う画像処理でも，画像処理を意識することなく，最適な画像処理フィルタを構築することが可能となる．

本研究では，画像処理フィルタを Fig. 4.1 に示すような木構造状フィルタとして表し，GP および，SAP の自動プログラミング手法を用いて画像処理フィルタの最適化を行なう．木構造状フィルタのように複数系統の組み合わせを構築することで，一系統だけの処理では実現不可能な複雑な処理も可能となる．

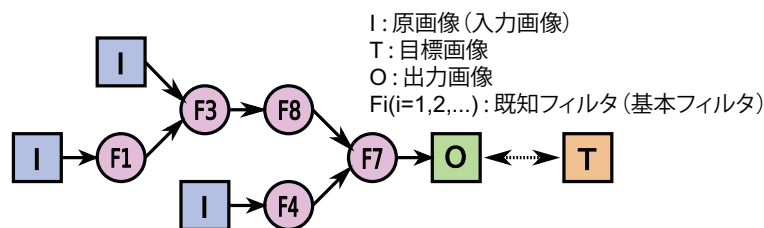


Fig. 4.1 木構造状フィルタによる画像処理の自動化の原理

Fig. 4.1 に示した木構造状フィルタは，原画像（入力画像） I から目標画像 T に至る画像処理アルゴリズムを近似したものである．また Fig. 4.2 に示すような原画像 I と，それに対応する理想的な処理画像である目標画像 T の組である学習用画像セットを事前に与えることで，木構造状フィルタの学習を行なう．なお目標画像 T は原画像 I から人手によって作成したものである．Fig. 4.2 に学習用画像セットの一例を示す．

木構造状フィルタを用いて画像処理を実現するには，まずこの木の終端ノードから原画像 I を入力する．その後，非終端ノードに格納された基本フィルタにより複数系統で順次画像処理を行なう．各非終端ノードに格納された基本フィルタが順次実行された後，最終的に木構造状フィルタのルートノードから 1 つの出力画像 O が出力される．また構築された木構造状フィルタは，出力画像と目標画像を比較することにより評価する．本研究で用いた基本フィルタを 4.2 節，木構造状フィルタの評価方法を 4.3 節でそれぞれ述べる．

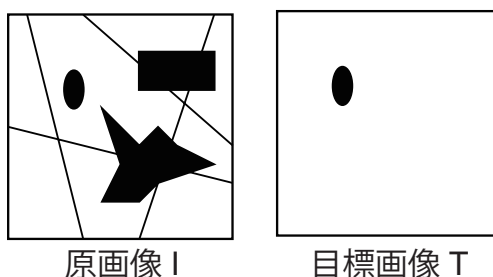


Fig. 4.2 学習用画像セットの一例

4.2 基本フィルタ

本節では、本研究で用いた基本フィルタについて説明する．本研究で用いた一入力の基本フィルタを Table 4.1 に、二入力の基本フィルタを Table 4.2 に示す．これらのフィルタは、長尾らが用いたフィルタ⁹⁾を参考に、階調値を変化させるものを選択している．

Table 4.1 一入力の基本フィルタ

Filter ID	Effect
f101	Median
f102	Inversion
f103	Binarization
f104	Edge emphasis(laplacian)
f105	Edge emphasis(dark edge)
f106	Minimum value of eighborhood
f107	Maximum value of eighborhood
f108	Noise Removal

Table 4.2 二入力の基本フィルタ

Filter ID	Effect
f201	Logical Sum
f202	Logical Product
f203	Algebraic Sum
f204	Algebraic Product
f205	Bounded Sum
f206	Bounded Product

また本研究では、原画像 I が白黒濃淡画像の場合と、カラー画像の場合とで木構造状フィルタの終端ノードを変更する．原画像 I が白黒濃淡画像の場合は、それぞれの終端ノードに対して、原画像 I をそのまま入力する．従って、全ての終端ノードには同じ画像の情報が格納される．また、原画像 I がカラー画像の場合は、Table 4.3 に示すような、原画像 I の各色成分の情報を白黒濃淡画像で表現し、各終端ノードにランダムに格納する²⁰⁾．扱う色成分の情報としては、カラー画像の色の表現法である RGB の R 成分、G 成分、および B 成分と、色空間を扱う HSB の色相 (Hue)、彩度 (Saturation)、および明度 (Brightness) の計 6 つの情報を扱うものとする．なお本論文ではカラー画像の病理画像を扱うため、終端ノードには各色成分の情報を格納した．

Table 4.3 カラー画像を扱う際の終端記号

Filter ID	Component of input image
I01	Red component of input image
I02	Green component of input image
I03	Blue component of input image
I04	Hue component of input image
I05	Saturation component of input image
I06	Brightness component of input image

4.3 木構造状フィルタの評価

各個体の評価は式 (4.1) に示すように、画像処理フィルタからの出力画像 O と目標画像 T について各ピクセル値の差分によって求める。次式は 1.0 を最適解とする最大値問題とする。

$$fitness = \frac{1}{k} \sum_{k=1}^K \left\{ 1 - \frac{\sum_{x=1}^{W_x} \sum_{y=1}^{W_y} |O(x, y) - T(x, y)|}{W_x \cdot W_y \cdot V_{max}} \right\} \quad (4.1)$$

ここで、 $O(x, y)$ は出力画像の各ピクセル、 $T(x, y)$ は目標画像の各ピクセルを示す。また K は学習用画像セット数、 W_x は画像の横サイズ、 W_y は画像の縦サイズ、 V_{max} は最大階調値を示す。複数の学習用画像セットを用いて学習を行なう場合、評価値はそれぞれの学習用画像セットに対する評価値の単純平均として算出する。

5 自動プログラミング手法

本稿では、自動プログラミングの手法²¹⁾として、GP と SAP を用いる。以下に、それぞれの自動プログラミング手法について述べる。

5.1 遺伝的プログラミング (Genetic Programming: GP)

GP は、最適化手法である遺伝的アルゴリズム (Genetic Algorithm: GA)¹⁰⁾ を木構造や、グラフ構造などの構造的な表現が扱えるように拡張した自動プログラミング手法である。GP では、選択、交叉、突然変異といった遺伝的オペレータを繰り返し行うことにより、問題に適した解を生成する。以下に GP のアルゴリズムを示す。

Step 1 初期個体群の生成

複数の個体をランダムに生成し、初期個体群 (母集団) とする。また個体の生成時、各個体の評価を行う。

Step 2 選択

各個体の評価値を判断基準とし、次世代に残す個体を選択する。本論文では、母集団の中からランダムに選択した一定数の個体の中から、評価値が最も高い個体を選択して次世代に残すトーナメント選択を用いる。

Step 3 交叉

交叉対象となる個体から、ランダムに選択した交叉点をルートとする部分木同士を入れ替える (Fig 5.1)。

Step 4 突然変異

突然変異の対象となる個体から、ランダムに選択した突然変異点以下の木を、ランダムに作成した突然変異木に置き換える (Fig 5.2)。

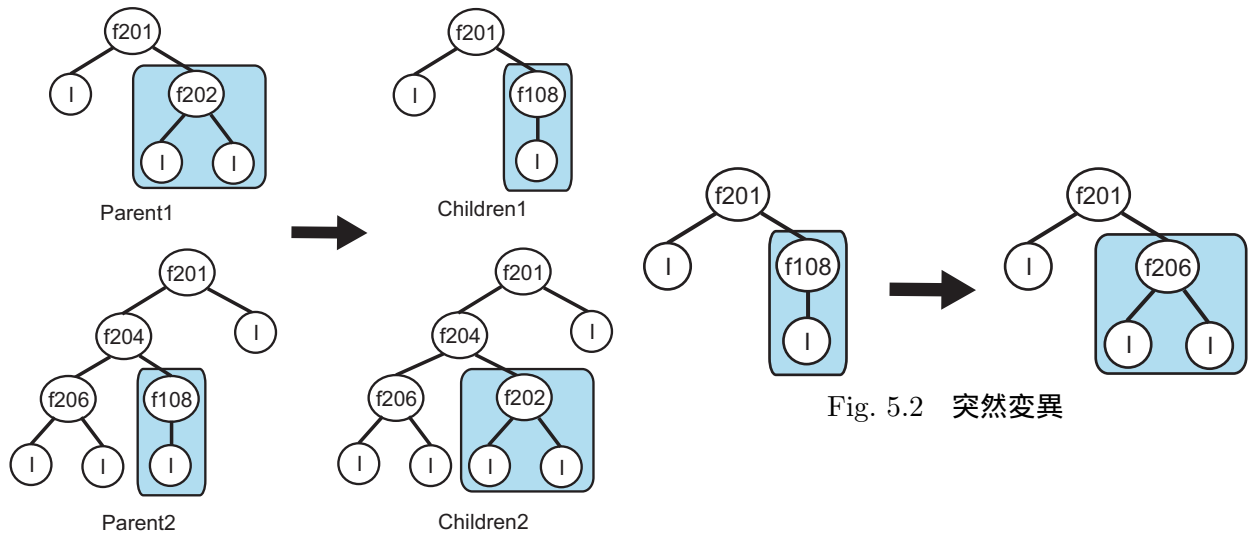


Fig. 5.1 交叉

Fig. 5.2 突然変異

Step 5 終了条件

終了条件に達するまで，Step2～4の処理を繰り返す．本論文では，指定した探索世代数に達すれば終了とする．

また GP では探索が進むにつれて木構造のサイズが増大するブロートという現象が発生すると報告されている^{11,20)}．ブロートが発生する要因は，イントロン (intron) と呼ばれる適合度計算に関与しない部分解の発生や，簡潔に表現されていない冗長な部分解の出現によるとされている．一般に，イントロンはブロートを引き起こす一方で，良好な部分解が交叉によって破壊されることを防ぎ，潜在的に機能的な部分解を含むという理由により，GP の探索性能を向上させる²²⁾．しかしながら，イントロンによる探索性能の向上は，主に明示的に定義されたイントロンによるものであり，探索過程で創発的に出現したイントロンは探索性能の向上よりもブロートを引き起こすことによる弊害の方が顕著である．そのため GP で探索を行なう場合，創発的に出現するイントロンを抑制する必要があるが，その最適な手法は提案されていない．

5.2 シミュレーテッドアニーリングプログラミング (Simulated Annealing Programming: SAP)

SAP¹⁴⁾ は，最適化手法であるシミュレーテッドアニーリング (Simulated Annealing: SA)¹³⁾ を，木構造が扱えるように拡張した自動プログラミング手法である．SAP では，生成処理，受理判定，状態遷移，冷却を繰り返し行うことで，問題に適した解を生成する．以下に SAP のアルゴリズムを示す．

Step 1 初期個体の生成

初期個体をランダムに生成し，その個体の評価を行う．

Step 2 生成処理

現在の個体に対し，GP の突然変異 (Fig 5.2) と同様の処理を行うことにより，新しい個体を生成し，その個体の評価を行う．

Step 3 受理判定, 状態推移

現在の個体の評価値 E と生成処理により生成した新しい個体の評価値 E' との差分 $\Delta E (= E' - E)$, および温度パラメータ T により, 状態を推移するか否かの受理判定を行う. 受理判定には, 式 (5.1) に示す Metropolis 基準¹³⁾を用いる.

$$P_{AC} = \begin{cases} 1 & \text{if } \Delta E \geq 0 \\ \exp(\frac{\Delta E}{T}) & \text{otherwise} \end{cases} \quad (5.1)$$

ここで P_{AC} は受理確率である. Metropolis 基準では, 新しい個体が改良方向 ($\Delta E \geq 0$) の場合は必ず受理し, それ以外の場合は, 温度 T と, 評価値の差分 ΔE によって, 受理する確率が変化する.

Step 4 冷却

Step2,3 の処理を一定回数繰り返した後, 温度パラメータ T を小さくする冷却を行う. ここでは式 (5.2) に示すような指数型アニーリングを用いる.

$$T_{k+1} = \gamma T_k \quad (0.8 \leq \gamma \leq 1) \quad (5.2)$$

ここで, γ は冷却率であり, T_k は現在の温度, T_{k+1} は冷却後の温度を表す. なおこの処理は, 冷却を行なう場合にのみ用いる. このように冷却を用いる SAP を冷却 SAP といい, 冷却を用いず一定温度の温度スケジュールを用いる SAP を一定温度 SAP という.

Step 5 終了条件

終了条件に達するまで, Step2~4 の処理を繰り返す. 本論文では, 指定した探索世代数に達すれば終了とする.

SAP の大きな特徴は, 新しい解候補が改悪方向へ生成された場合でも, その解候補への遷移を確率的に認めることである. この確率は, 式 (5.1) の Metropolis 基準により決定され, 改悪方向への遷移確率は温度パラメータ T によって制御することができる. この温度パラメータ T が大きい (高温) 場合, 改悪方向への遷移を多く受理し, 大域的な探索が行なわれる. 一方, 温度パラメータ T が小さい (低温) 場合, 改悪方向への遷移をほとんど認めず, 局所的な探索が行なわれる. そのため, SAP においてこの温度パラメータをどのように設定するかが非常に重要になる.

6 癌胞巣抽出アルゴリズムの自動構築

本章では, GP および, SAP を用いて画像処理フィルタを構築することで, 癌胞巣抽出アルゴリズムの自動構築を行なう. そして, GP および, SAP の両手法を比較し, 本問題における最適な手法について検討する. 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.

6.1 GP におけるプログラムサイズ制限が探索に及ぼす影響

5.1 節で述べたように, GP では創発的に出現するイントロンを抑制する必要がある. 本稿では, 構築された木構造のフィルタ数 (非終端ノード数) をプログラムサイズと定義し, プログラムサイズを

制限することで、イントロンの出現を防ぐことを考える。本節では、GP におけるプログラムサイズ制限が探索に及ぼす影響について検討する。

6.1.1 実験概要

癌巣抽出アルゴリズムの自動構築問題において、GP のプログラムサイズ制限が探索に及ぼす影響について検討するために、プログラムサイズの制限を変更して数値実験を行なう。プログラムサイズ制限の処理は、交叉、突然変異の処理において、設定されたプログラムサイズよりもサイズの大きい次個体が生成された場合、その処理を無効とし、親個体を子個体とする処理を行なった。本実験に用いたパラメータを Table 6.1 に示す。本実験では単一学習を行ない、学習画像は Fig 6.1 に示すものを用いた。また Table 4.1, Table 4.2 に示す、基本フィルタを用いた。

Parameter	Value
Generations	100
Populations	100
Selection Method	Tournament
Crossover Rate	0.9
Mutation Rate	0.1
Maximum of the number of filters	10, 25, 50, 75

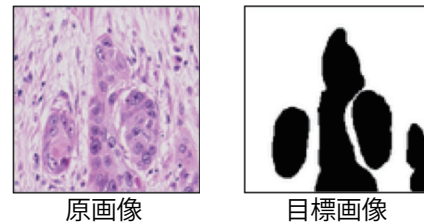


Fig. 6.1 学習用画像セット（単一画像学習）

6.1.2 学習性能

GP の学習性能を Fig 6.2 に示す。Fig 6.2 の横軸は世代数、縦軸は評価値を示す。結果は、各探索時における最適解の 10 試行の中央値である。

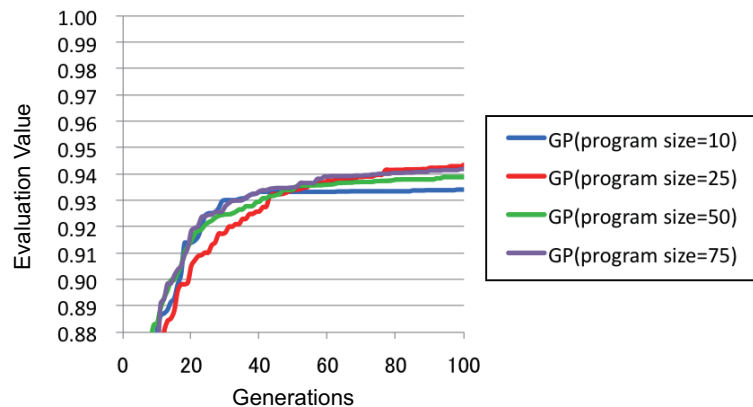


Fig. 6.2 プログラムサイズ制限の探索における学習性能

Fig 6.2 より、プログラムサイズが 10 と制限が厳しい場合、制限が 25 以上の場合と比べて評価値が劣ってしまった。しかしプログラムサイズが 25 以上の場合、プログラムサイズ制限の変更が大きく探索に影響することはないことがわかる。

6.1.3 プログラムサイズ

プログラムサイズの推移を Fig 6.3(a) に示し、計算時間の推移を Fig 6.3(b) に示す。Fig 6.3(a) の横軸は世代数、縦軸はプログラムサイズを示し、Fig 6.3(b) の横軸は世代数、縦軸は計算時間を示す。結果は、各探索時における最適解の 10 試行の中央値である。

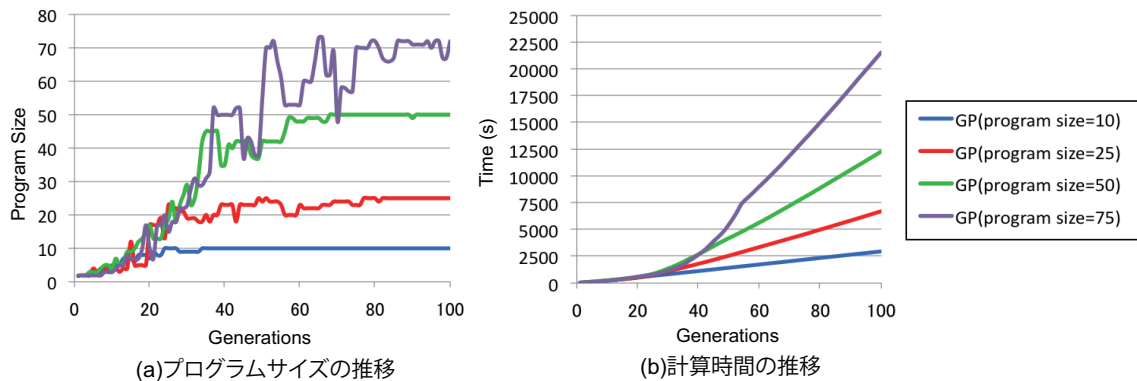


Fig. 6.3 プログラムサイズ制限の探索におけるプログラムサイズの比較

Fig 6.3(a) より、それぞれの GP は設定した制限までプログラムサイズが増大していることがわかる。このことから、GP ではプログラムサイズを制限しないとプログラムサイズが増大し続けることがわかる。このプログラムサイズの差は、Fig 6.3(b) に示したそれぞれの探索での計算時間にも反映し、プログラムサイズが増大するほど、計算時間を要することがわかった。このようにプログラムサイズが増大したのは GP による交叉が原因であると考えられる。Fig 6.4 に、プログラムサイズを 75 に制限した GP と交叉を用いない GP におけるプログラムサイズの比較を示す。なお、交叉を用いない GP にはプログラムサイズの制限を用いていない。

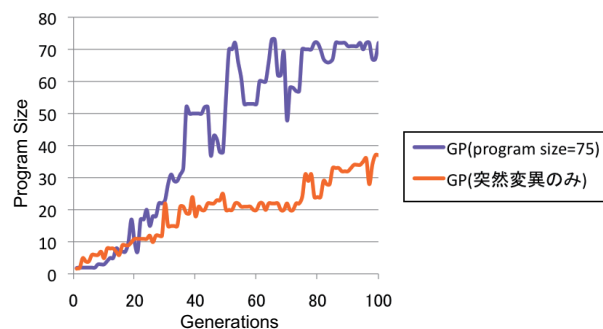


Fig. 6.4 GP(プログラムサイズ制限:25) と GP(交叉なし) におけるプログラムサイズの比較

Fig 6.4 より、交叉を用いないことにより、40 世代以降の爆発的なプログラムサイズの増加が抑制されていることがわかる。突然変異だけの場合、オペレーションを行なうノードがルートノードに近い場合、新しく生成される個体のプログラムサイズは小さくなる。一方交叉では、オペレーションを行なうノードが、ルートノードに近い場合と終端ノードに近い場合に交叉が行なわれると、一方の木構造のプログラムサイズが莫大に増大する。このような交叉によりプログラムサイズの大きな木構造が構築され、それが母集団中に広がり、また交叉を行なうことでプログラムサイズが増大していくために、結果として大幅に計算時間を要したと考えられる。

6.2 一定温度 SAP における温度パラメータが探索に及ぼす影響

5.2 節で述べたように，SAP では温度パラメータの設定が非常に重要である．本節では，一定温度 SAP における温度パラメータが探索に及ぼす影響について検討する．

6.2.1 実験概要

癌巣抽出アルゴリズムの自動構築問題において，一定温度 SAP の温度パラメータが探索に及ぼす影響について検討するために，一定温度の温度スケジュール²³⁾を用いて数値実験を行なう．本実験に用いたパラメータを Table 6.2 に示す．温度パラメータは Table 6.2 に示した 6 温度を用いた．本実験では単一学習を行ない，学習画像は Fig 6.1 に示したものをを用いた．また本実験においても Table 4.1，Table 4.2 に示す，基本フィルタを用いた．

Table 6.2 一定温度 SAP のパラメータ

Parameter	Value
Number of Evaluations	10000
Fix Temperature	1.6, 0.4, 0.1, 0.025, 0.00625, 0.0015625

6.2.2 学習性能

温度毎の一定温度 SAP の学習性能を Fig 6.5 に示す．Fig 6.5 の横軸は評価計算回数，縦軸は評価値を示す．結果は，各探索時における最適解の 10 試行の中央値である．

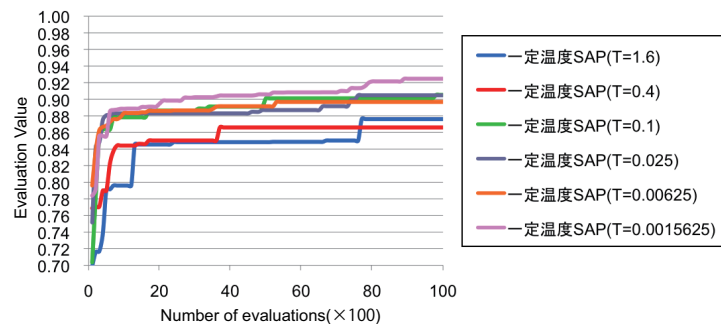


Fig. 6.5 一定温度探索における学習性能

Fig 6.5 より，温度が低いほど探索序盤で高い評価の解が得られており，最終的な評価値も高いことがわかる．一般的に，SAP の問題では重要温度領域が存在すると言われている¹⁴⁾が，本問題では，温度が低くなるにつれ，良い評価が得られていることから，重要温度領域が存在しない問題だと考えられる．温度が高いと次個体が改悪の場合に遷移する確率が高くなり，温度が低い場合と比べ最適な解が得られていないことがわかる．

6.2.3 プログラムサイズ

プログラムサイズの推移を Fig 6.6(a) に示し，Fig 6.6(b) に計算時間の推移を示す．Fig 6.6(a) の横軸は評価計算回数，縦軸はプログラムサイズを示し，Fig 6.6(b) の横軸は評価計算回数，縦軸は計算時間を示す．結果は，各探索時における最適解の 10 試行の中央値である．

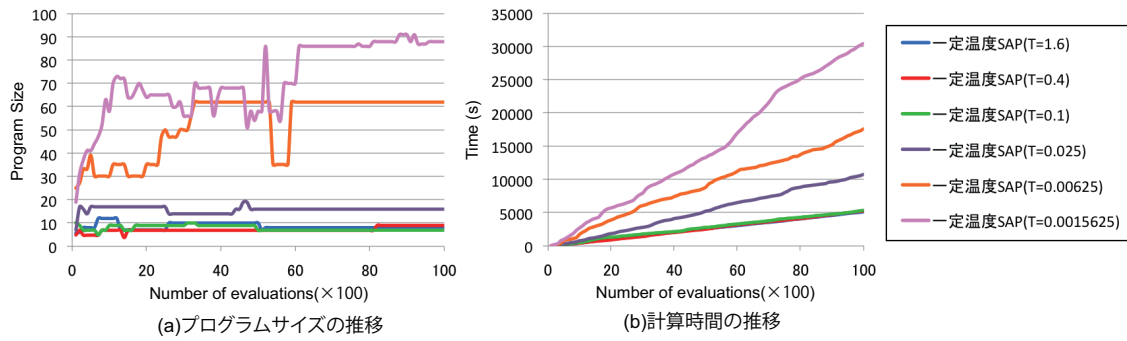


Fig. 6.6 一定温度探索におけるプログラムサイズの比較

Fig 6.6(a) より、温度が低くなるにつれて、プロートが発生していることがわかる。一方、温度が高いと小さなプログラムサイズで収束していることがわかる。これは、Metropolis 基準での改悪の場合に遷移する確率が影響している。オペレーションを行なうノードがルートノードに近い場合、新しく生成される解候補のプログラムサイズは比較的小さくなる。また、これまでに評価をした多くのノードが削除される代わりに、評価をしていないランダムに生成したノードが挿入されるため、評価が悪くなる可能性が高い。よって、温度が高いほどそのような改悪となる解候補を受け取りやすいため、高温ではプログラムサイズが小さくなったと考えられる。一方、温度が低いほどそのような解候補を受け取りにくいために、低温ではプログラムサイズが大きくなったと考えられる。また Fig 6.6(b) より、プログラムサイズの増加は、計算時間の増加に繋がることがわかる。ただし、高い温度ではプロートが抑制できるが、6.2.2 項より、温度を高くしてしまうと、より最適な解が得られる探索ができないといった問題が生じてしまう。

6.3 冷却 SAP における温度スケジューリングが探索に及ぼす影響

本節では、6.2 節で述べた一定温度ではなく、冷却 SAP を用いることで温度スケジューリングが探索に及ぼす影響について述べる。

6.3.1 実験概要

癌細胞抽出アルゴリズムの自動構築問題において、冷却 SAP の温度スケジューリングが探索に及ぼす影響について検討するために、一定温度 SAP と冷却 SAP の比較を行なう。本実験に用いたパラメータを Table 6.3 に示す。なお、最高温度は最大の改悪を 50% の確率で受理する温度を、最低温度は最小の改悪をあるクーリング周期内で 1 回受理する温度を、予備実験によって求めたものである²⁴⁾。また本実験では単一学習を行ない、学習画像は同様に Fig 6.1 に示したものをを用いた。また本実験においても Table 4.1, Table 4.2 に示す、基本フィルタを用いた。

Table 6.3 冷却 SAP のパラメータ

Parameter	Value
Number of Evaluations	10000
Max Temperature	1.44
Min Temperature	0.00015

6.3.2 学習性能

冷却 SAP と一定温度 SAP の学習性能の比較を Fig 6.7 に示す．Fig 6.7 の横軸は評価計算回数，縦軸は評価値を示す．結果は，各探索時における最適解の 10 試行の中央値である．

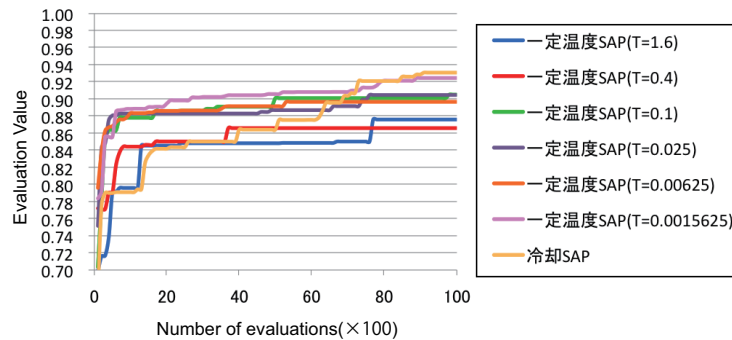


Fig. 6.7 冷却 SAP と一定温度 SAP の学習性能の比較

Fig 6.7 より，探索終了時，冷却 SAP は温度の低い一定温度 SAP と同等，またはそれ以上の性能を示していることがわかる．温度を効率よく冷却させることで，より最適な解が得られることがわかる．一方，温度の低い一定温度 SAP は探索序盤で高い評価の解が得られていることがわかる．

6.3.3 プログラムサイズ

冷却 SAP と一定温度 SAP のプログラムサイズの推移の比較を Fig 6.8(a) に示し，計算時間の推移の結果を Fig 6.8(b) に示す．Fig 6.8(a) の横軸は評価計算回数，縦軸はプログラムサイズを示し，Fig 6.8(b) の横軸は評価計算回数，横軸は計算時間を表す．結果は，各探索時における最適解の 10 試行の中央値である．

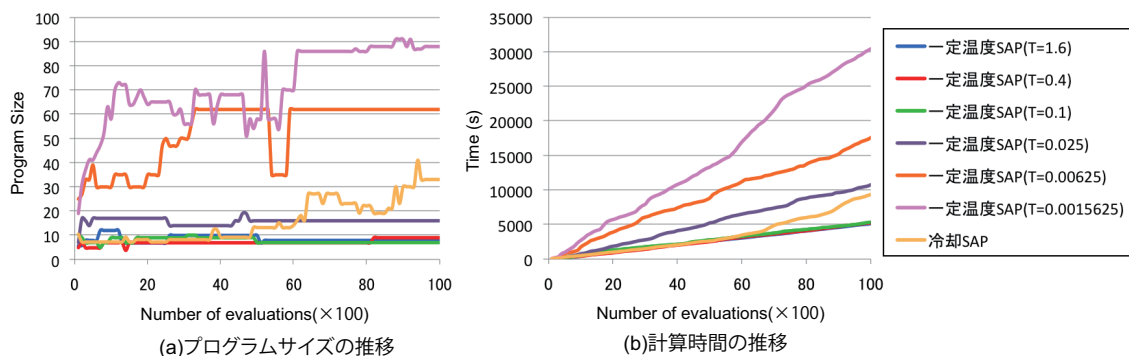


Fig. 6.8 冷却 SAP と一定温度 SAP のプログラムサイズの比較

Fig 6.8(a) より，冷却 SAP では，探索終盤わずかながらプログラムサイズが増加したが，探索序盤の急激なプロットを抑制できることがわかる．また Fig 6.8(b) より，冷却 SAP の探索終了時までの計算時間は，一定温度 SAP(T=0.0015625) での探索のおおよそ 3500step に相当する．SAP(T=0.0015625) での探索の 3500step の評価値と冷却 SAP の探索終了時の評価値を比べた場合，冷却 SAP の方が高い評価が得られていることから，冷却 SAP では探索序盤のプロットを抑制し，一定温度 SAP よりも効率の良い探索が可能であることがわかる．

6.4 GP と SAP の比較

本節では、GP と SAP の性能比較を行ない、それぞれの手法の探索における特徴について検証する。

6.4.1 実験概要

本実験では、口腔癌病理画像から癌胞巣抽出アルゴリズムの自動構築をする問題において、GP、SAP を適用し、それぞれの性能比較を行なう。GP のパラメータは、Table 6.1 と同様のものを用い、最大プログラムサイズは 25 に制限した。また SAP は冷却 SAP を用い、Table 6.3 と同様のパラメータを用いた。本実験でも、単一画像による学習を行い、Fig 6.1 と同様の画像を用い、Table 4.1、Table 4.2 に示す、基本フィルタを用いた。

6.4.2 学習性能

本節では、学習性能について GP と SAP の比較を行なう。Fig 6.9 に、今回用いた GP と SAP の学習性能の比較を示す。Fig 6.9 の横軸は評価計算回数、縦軸は評価値を示す。なお GP は 1 世代で 100 回評価計算を行なうことから、1 世代を評価計算回数 100 とし、GP と SAP の横軸を調整した。結果は、各探索時における最適解の 10 試行の中央値である。また Fig 6.10 にそれぞれの手法の出力結果を示す。なお、Fig 6.10 の上段はそれぞれの手法で構築された画像処理フィルタから出力された画像であり、下段は出力された画像と原画像との論理和処理を行ない、原画像から癌胞巣を抽出した画像である。

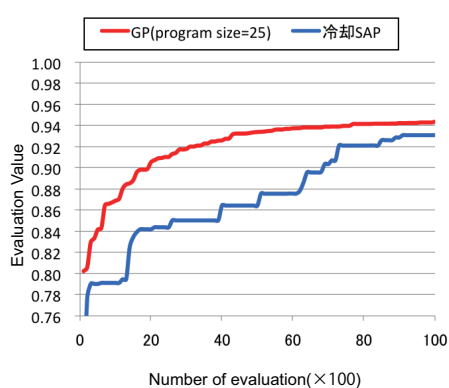


Fig. 6.9 GP と SAP の学習性能の比較

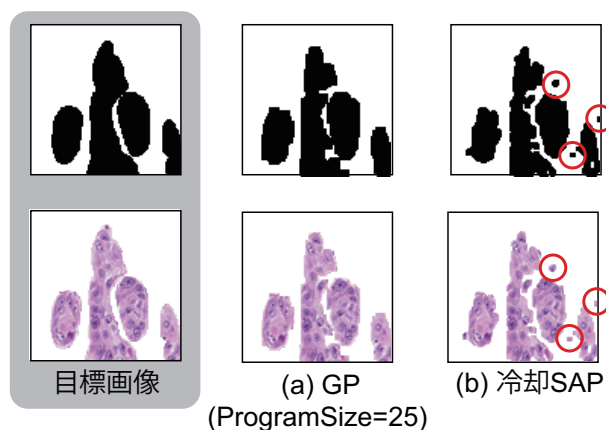


Fig. 6.10 出力結果

Fig 6.9 に示したように、冷却 SAP よりも GP の評価の方が、探索の早い段階で収束しており、探索終了時での評価値も高いことがわかる。このように、GP でプログラムサイズ制限を厳しくした場合でも、SAP よりも学習性能に関して、性能が高いことがわかった。また Fig 6.10 に示したように、冷却 SAP は抽出対象の輪郭部が滑らかではなく、赤丸で示したノイズも発生していることから、GP の方がより目標画像に近いアルゴリズムを構築できることが視覚的にも確認できる。このように GP が少ない評価回数で収束したのは、各世代ごとに母集団を形成し、母集団全体で解更新を行なっていくためだと考える。そのため GP では、評価計算を行なうにつれ、母集団中の各個体が解更新され、比較的评价の高い個体同士を交叉することで、さらに評価の高い解が爆発的に生成されたと考える。

6.4.3 プログラムサイズ

本節では、プログラムサイズについて GP と SAP の比較を行なう。Fig 6.11(a) にプログラムサイズの比較、Fig 6.11(b) に計算時間の比較を示す。Fig 6.11(a) の横軸は評価計算回数、縦軸はプログラムサイズを示し、Fig 6.11(b) の横軸は評価計算回数、横軸は計算時間を表す。結果は、各探索時における最適解の 10 試行の中央値である。

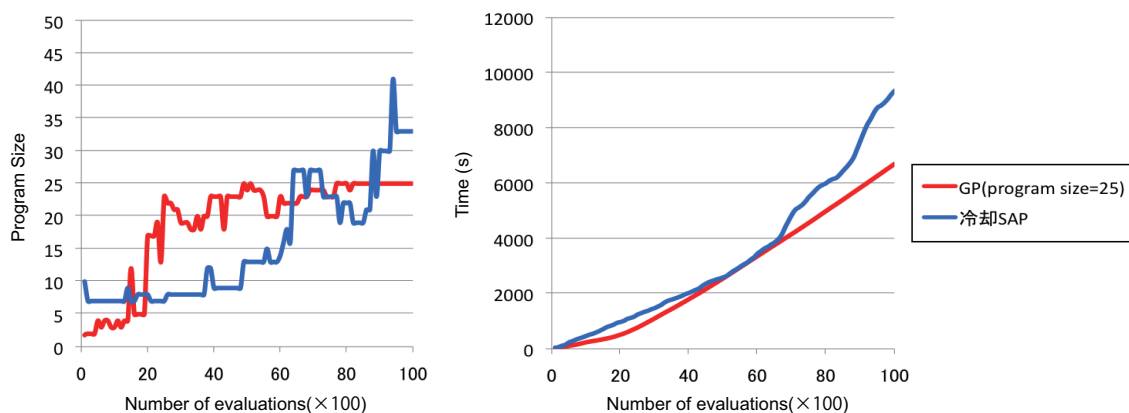


Fig. 6.11 GP と SAP のプログラムサイズの比較

Fig 6.11(a) より、GP の制限を 25 と厳しく設定することで、冷却 SAP よりもプログラムサイズを抑制して探索できることがわかる。また Fig 6.11(b) に示したように、プログラムサイズを抑制することで、計算時間も抑えることが可能である。そのため、制限を厳しくしたとしても、GP を用いた方が学習性能、プログラムサイズの両性能において良い結果が得られることがわかる。

6.4.4 ロバスト性

本節では GP および、SAP でそれぞれ構築した木構造を未知画像に適用することで、両手法におけるロバスト性について検証する。また本実験に用いる未知画像を Fig 6.12 に示し、また未知画像に対する抽出対象部位（癌胞巣）を Fig 6.13 に示す。

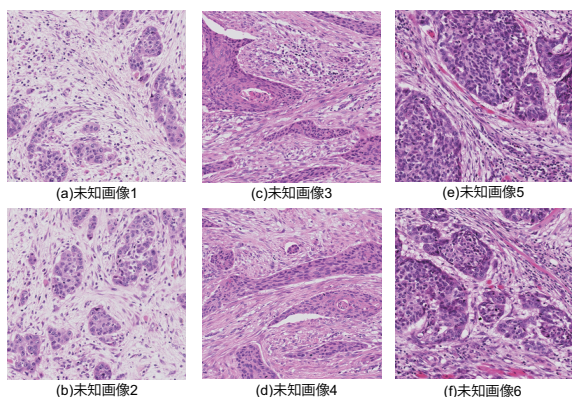


Fig. 6.12 未知画像

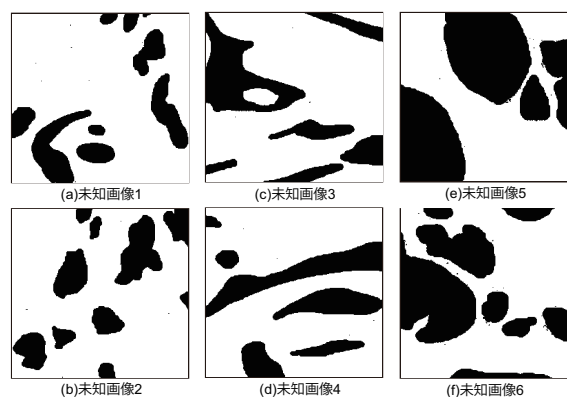


Fig. 6.13 未知画像に対する抽出対象部位

本実験では、未知画像として、口腔癌患者の 3 症例分の画像を用いた。Fig 6.12 に示した画像は口腔癌患者 3 症例分の画像であり、(a) と (b)、(c) と (d)、(e) と (f) がそれぞれ同じ検体の画像の他の領域を切り抜いた画像である。なお (a) と (b) は Fig 6.1 に示した原画像と同じ検体の他の領域を切

り抜いた画像である。GP で構築したアルゴリズムを適用した結果を Fig 6.14 , SAP で構築したアルゴリズムを適用した結果を Fig 6.15 に示す。

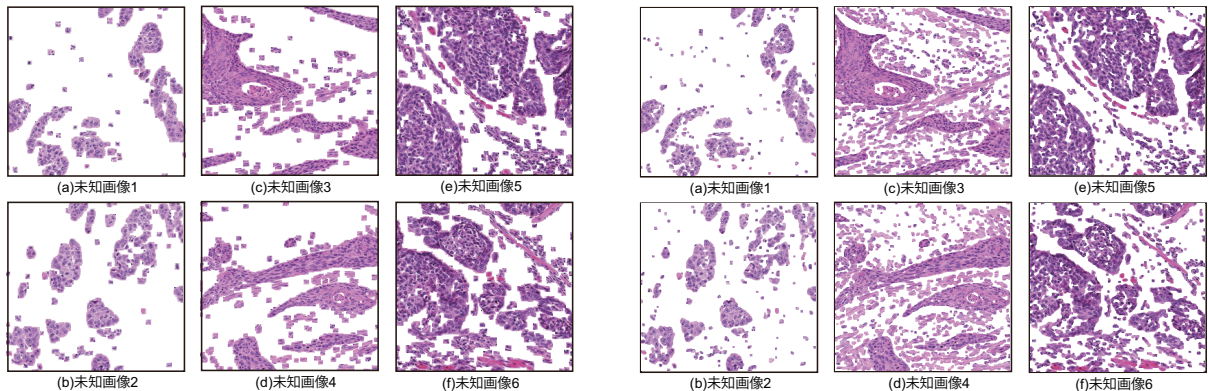


Fig. 6.14 GP(制限:25)を未知画像に適用した結果 Fig. 6.15 冷却SAPを未知画像に適用した結果

Fig 6.14 , Fig 6.15 に示したように , GP , SAP においても , 目標とする部位が抽出できていることが確認できる . しかし , SAP では , 学習に用いていない検体 ((c) ~ (f)) において , 多くのノイズが発生している . 一方 GP ではノイズは多少発生しているが , SAP よりも少ないことがわかる . このように単一画像の学習を行なうだけでは , 未知画像に対してノイズが発生し , 最適な手法であるとはいい難い . そのため次章では , 未知画像においてノイズをいかに除去するか検討する .

7 癌胞巣抽出に対するノイズ除去に関する検討

本章では , 口腔癌病理画像における癌胞巣抽出問題において発生するノイズの除去に関して検討する .

7.1 ノイズ除去処理が探索に及ぼす影響

本節では , ノイズ除去処理が探索に及ぼす影響について検証する .

7.1.1 概要

ノイズ除去処理が探索に及ぼす影響について検証するために , Table 4.1 に示した基本フィルタにノイズ除去処理 (f108) を加えて探索を行なう . このノイズ除去の処理は , ある閾値以下の面積をノイズとみなし , そのノイズを除去する画像処理である . また本実験では , ノイズ除去の処理を自動化するために , ノイズ除去の閾値を抽出されている領域の平均面積とした . 探索には , プログラムサイズを 25 に制限した GP を用い , パラメータは Table 6.1 と同様のものを用いた . また本実験でも , 単一画像による学習を行い , Fig 6.1 と同様の画像を用いた .

7.1.2 実験結果

GP で探索を行ない , 構築されたフィルタを Fig 6.12 に示した未知画像に適用した結果を Fig 7.1 に示す .

Fig 7.1 より , Fig 6.14 に示したノイズ除去処理がない場合と比べると , 特に学習画像と同じ検体である未知画像 (a) , (b) に関してはノイズ (赤丸) が少なくなっていることがわかる . しかし , 学習

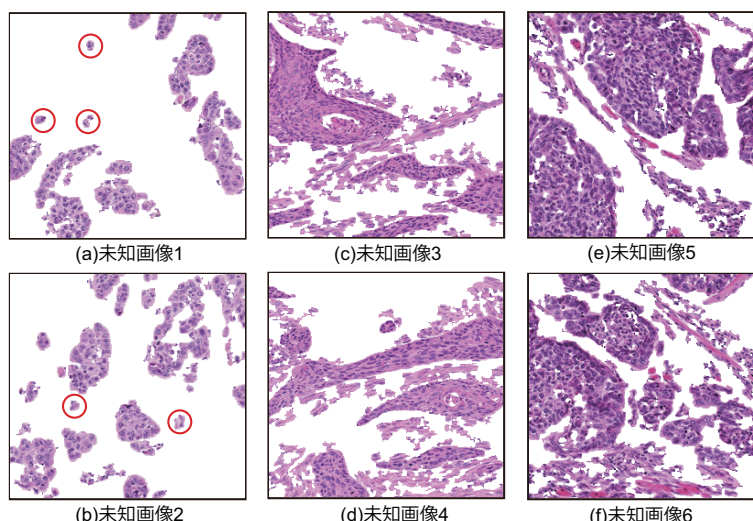


Fig. 7.1 ノイズ除去を加えた GP (Program Size=25) を未知画像に適用した結果

画像に用いなかった未知画像 (c) ~ (f) の検体に関してはノイズが十分には除去できないという結果となった。これは、単一画像探索しか行なっておらず、ノイズ除去処理の閾値が未知画像でうまく作用しなかったためと考えられる。また本節で構築した癌胞巣抽出アルゴリズムを Fig 7.2 に示す。図中に示す画像は、未知画像 (b) を適用した場合の各処理での抽出結果である。

Fig 7.2 より、癌胞巣抽出アルゴリズムは平滑化処理 2 値化処理 ノイズ除去処理 膨張処理の 4 段階の処理に集約されることがわかった。本稿では、この 4 段階の処理を癌胞巣抽出のための特徴とする。

7.2 複数画像探索が探索に及ぼす影響

本節では、複数画像探索が探索に及ぼす影響について検証する。

7.2.1 概要

複数画像探索が探索に及ぼす影響について検証するために、本節では単一画像探索ではなく、複数画像を用いて探索を行なう。本実験でも、探索にはプログラムサイズを 25 に制限した GP を用い、パラメータは Table 6.1 と同様のものを用いた。7.1 節同様、本実験においてもノイズ除去処理 (f108) を用いた。本実験では複数画像による学習を行い、Fig 7.3 に示す学習用画像セットを用いる。学習画像 1 は Fig 6.1 と同様の画像であり、学習画像 2 は学習画像 1 と同検体で、他の領域の癌胞巣を含まない領域を切り抜いた画像である。また学習画像 3 は他検体の癌胞巣を含む画像である。

7.2.2 実験結果

GP で探索を行ない、構築されたフィルタを Fig 6.12 に示した未知画像に適用した結果を Fig 7.4 に示す。

Fig 7.4 に示したように、複数画像で学習することでノイズを除去することができたが、抽出したい部位も除去されていることがわかる。これは、学習画像 2 が癌胞巣を含まない画像で学習したため、学習画像において何も抽出しない画像の方が評価が高くなり、ノイズ除去が過剰に行なわれたため

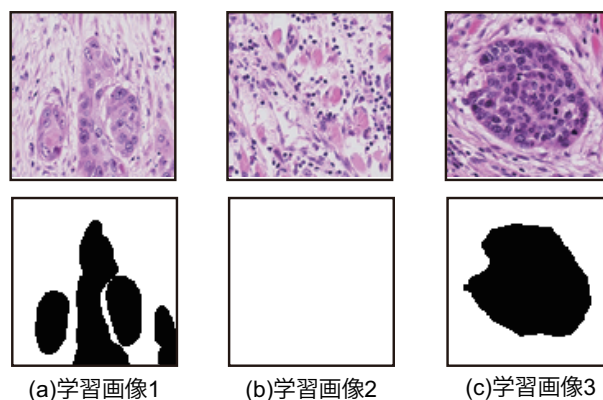


Fig. 7.3 学習画像セット（複数画像学習）

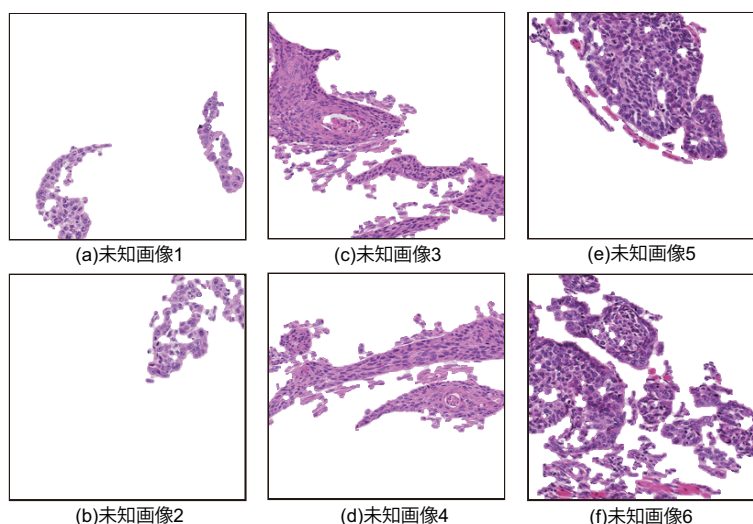


Fig. 7.4 複数画像で学習した GP を未知画像に適用した結果

像から癌巣を抽出するための画像処理をこの4段階の処理とみなし、各々の画像処理のパラメータをGAにより最適化する。最適化を行なうパラメータは平滑化処理の回数、ノイズ除去処理の閾値、膨張処理の回数である。なお、2値化処理の閾値決定に関しては、本アルゴリズムに汎用性を持たせるために大津の閾値判別法^{25,26)}を用いた。また本実験では複数画像探索を行ない、Fig 7.3 に示した学習用画像セットを用いて学習を行なう。

7.3.2 実験結果

パラメータ最適化によって構築したアルゴリズムは平滑化処理（0回） 2値化処理（大津の閾値判別法） ノイズ除去処理（ノイズ面積120） 膨張処理（1回）となった。結果として、本実験では平滑化処理は行なわれなかった。このアルゴリズムを Fig 6.12 に示した未知画像に適用した結果を Fig 7.5 に示す。また未知画像4における本アルゴリズムの画像処理経過を Fig 7.6 に示す。

Fig 7.5 より、パラメータ最適化を用いることで、抽出したい癌巣の部位を抽出したまま、ノイズを抑えることができた。複数画像で学習することで、(c)~(f)での未知画像でもノイズを抑えることができ、汎用性の高いアルゴリズムが構築できていることがわかる。また Fig 7.6 より、ノイズ除去処理を行なう閾値を120と設定することで、抽出したい部位と、ノイズを分けることができ、ノイズ除去が行なえていることがわかる。

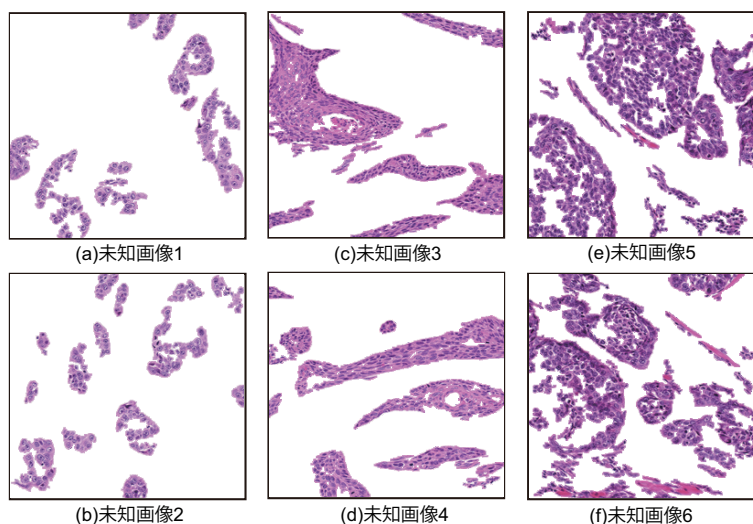


Fig. 7.5 パラメータ最適化したフィルタを未知画像に適用した結果

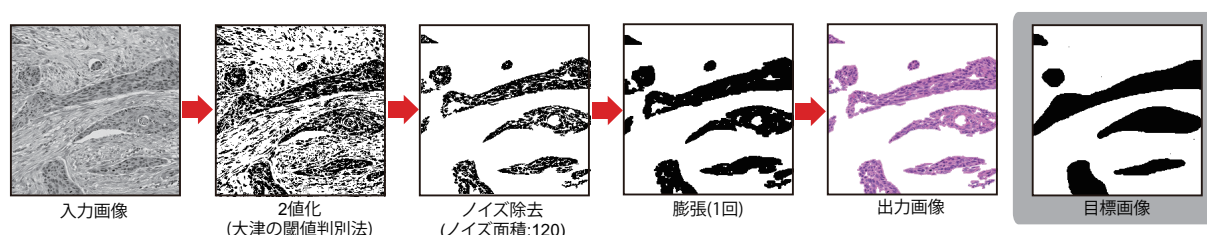


Fig. 7.6 特徴抽出アルゴリズムにおける画像処理経過 (未知画像 4)

ここで、7.1 節及び、7.2 節のアルゴリズムと比較することで、本アルゴリズムの性能を計る。比較には Receiver Operating Characteristic (ROC) 解析³⁾を用いた。True Positive Fraction (TPF) 及び False Positive Fraction (FPF) は以下のように定義した。

$$TPF = \frac{A_C \cap A_R}{A_R} \quad (7.1)$$

$$FPF = \frac{A_C \cap \bar{A}_R}{A_C} \quad (7.2)$$

A_C は本アルゴリズムにより抽出した領域、 A_R は抽出したい領域を示す。TPF は抽出したい領域のうち、実際に抽出できている割合を示し、FPF は抽出した領域のうち、誤抽出である割合を示す。つまり本解析では、 $(TPF, FPF) = (1.0, 0.0)$ となるほど良い手法であるといえる。Fig 6.12 に示した未知画像に対し、両アルゴリズムから出力された画像の TPF, FPF の値をプロットした図を Fig 7.7 に示す。なおこの結果は、各画像において 5 試行分のデータの中央値をプロットしたものである。

○ GP(単一画像探索, ノイズ除去処理) △ GP(複数画像探索, ノイズ除去処理) ▽ パラメータ最適化(複数画像探索)

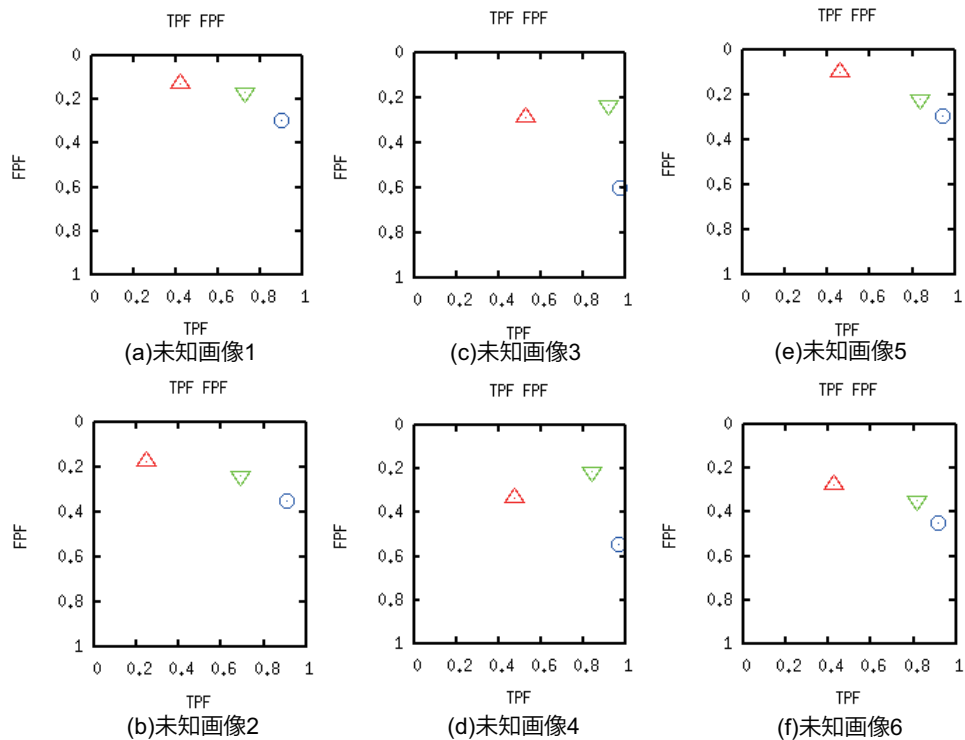


Fig. 7.7 TPF, FPF グラフによる比較

Fig 7.7 より, GP(単一画像探索, ノイズ除去処理) では TPF は一番高い値を示しているが, FPF が一番低く, ノイズが多いことがわかる。これは GP(単一画像探索, ノイズ除去処理) では, 学習に用いた画像が多くのノイズを含んでいないため, ノイズ除去の処理が過剰に行なわれなかったためだと考える。癌巣部位と背景領域の差は, 単一画像の学習でも明確に表現できたが, ノイズの多い未知画像ではノイズが除去できず, 多くのノイズを含む結果となった。また GP(複数画像探索, ノイズ除去処理) は FPF を上げることは出来たが, TPF が低くなり, 抽出対象が抽出できていないことがわかる。これは GP(複数画像探索, ノイズ除去処理) では, 癌巣部位を含まない複数画像で探索することで, 構築する画像処理フィルタにノイズ除去に対する意識を持たせることができたが, 未知画像では, ノイズ除去処理の閾値が上手く作用せず, 癌巣部位も抽出してしまったためだと考える。つまり, ノイズ除去の閾値設定が非常に重要になる。一方パラメータ最適化(複数画像探索) では, 全ての画像において TPF と FPF のバランスが良く, 特に, (c), (d) の未知画像において, TPF を高い値で維持したまま FPF を下げることができている。これは, ノイズ除去処理の閾値を一つに決定し, 癌巣部位とノイズを明確に区別できているためだと考える。ただし, 本アルゴリズムにおいて TPF が 1.0 に近づいていない理由は, 今回再構築したアルゴリズムが非常にシンプルなものであり, 高度な平滑化処理できておらず, 抽出対象である癌巣の中の部位が一部抽出できていなかったためである。今後, 特徴抽出をいかに行なうか検討し, いかに抽出対象の領域を塗りつぶしたまま癌巣抽出が可能なアルゴリズムを検討する必要がある。

8 結論

本論文では、口腔癌病理画像から癌巣を抽出するアルゴリズムの自動構築について述べ、ノイズを除去するアルゴリズムの構築について述べた。

まず自動で癌巣抽出を行なうアルゴリズムを構築するために GP と SAP を用い、両手法について比較を行なった。SAP は最適解を見つけるのに多くの評価計算が必要であるが、冷却 SAP を用いることで、プログラムサイズを抑えて探索することが可能であった。一方、GP は強い探索能力を示したが、プログラムサイズが増大するプロートが発生した。しかし GP において、プログラムサイズを制限することで、プロートを抑制したまま、冷却 SAP よりも高い能力を示すことがわかった。

しかし、単一画像探索を行なった GP だけでは、未知画像に適用した場合、多くのノイズが発生した。そこで GP で構築したアルゴリズムから特徴的なアルゴリズムを抽出し、各処理のパラメータを最適化することで、本問題に適した汎用性のあるアルゴリズムを構築した。本アルゴリズムを検証するために ROC 解析によって、TPF、FPF の値を比較したところ、抽出対象を残したままノイズが除去できていることがわかる。これらのことから、何も知識のないところからある程度のアルゴリズムは GP によって構築でき、GP により構築したアルゴリズムから癌巣抽出のための特徴的なアルゴリズムを抽出し、再構築したアルゴリズムの各処理のパラメータを最適化することで、よりシンプルで汎用的なアルゴリズムが構築できることがわかった。

しかし本稿で再構築したアルゴリズムは単純なものであり、複雑な平滑化処理が行なえず、抽出対象であった癌巣の中の部分が一部抽出できていないため、今後、GP からのアルゴリズム再構築をどのように行なうか検討する必要がある。アルゴリズムの再構築ができれば、全て自動で、汎用性のある癌巣抽出アルゴリズムを構築できることがわかった。また本稿では、病変部位の抽出を行なう病理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宮地正大さん、横田山都さん、宮部洋太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。宮部洋太さんには校正だけでなく、私のコーチとして1年間、多いに議論して頂きました。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、山口浩明さん、南谷祥之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、2年間すばらしい研究生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) 小掠真貴, 齋藤彰. 癌診断支援のための病理画像解析システム. 病理と臨床, Vol. 24, No. 4, pp. 411–415, 2006.
- 3) 榎本潤, 堀雅敏, 佐藤, 嘉伸, 村上卓道, 上甲剛, 中村仁信, 田村進一. マルチスライス ct 画像からの肝臓領域自動抽出. 電子情報通信学会論文誌, No. 9, pp. 2150–2161, 2001.
- 4) 松山隆司, 尾崎正治. トップダウン・セグメンテーションのための画像処理エキスパートシステム. 情報処理, Vol. 27, No. 2, pp. 191–204, 1986.
- 5) 田村秀行, 佐藤宏明, 坂上勝彦他. Dia-expert システムとその知識表現方法. 情報処理, Vol. 29, No. 2, pp. 199–208, 1988.
- 6) 周向栄, 清水昭伸, 長谷川純一, 他. 3次元画像処理手順の自動構築のためのビジョン・エキスパートシステム 3d-impress とその性能評価. 電子情報通信学会論文誌 D-II, No. 11, pp. 1949–1959, 1999.
- 7) 青木紳也, 長尾智晴. 木構造状画像変換の自動構築法 actit. 映像情報メディア学会誌, Vol. 53, No. 6, pp. 888–894, 1999.
- 8) T. Nagao S. Aoki. Automatic construction of tree-structural image transformation using genetic programming, international conference of image processing. pp. 529–533, 1999.
- 9) 長尾智晴. 進化的画像処理. 昭晃堂, 2002.
- 10) 伊庭斉志. 遺伝的プログラミング. 東京電機大学出版局, 1996.
- 11) J.Koza. *Genetic programming, on the programming of computers by means of natural selection*. MIT Press, 1992.
- 12) 中野雄太, 長尾智晴. 3次元画像処理自動構築システム 3d-actit の提案と pet 画像への応用. 医用画像情報学会, Vol. 24, No. 4, pp. 119–125, 2007.
- 13) Rosenbluth A. Rosenbluth M. Teller A. Teller E Metropolis, N. Equation of state calculation by fast computing machines. *Journ. of Chemical Physics*, Vol. 21, pp. 1087–1092, 1953.
- 14) 藤田佳久, 三木光範, 橋本雅文, 廣安知之. シミュレーテッドアニーリングを用いた自動プログラミング. 情報処理, Vol. 19, pp. 89–92, 2007.
- 15) 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. *IEICE*, p. 244, 2008.
- 16) 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.

- 17) Gary J. Macfarlane. Rising trends of oral cancer mortality among males worldwide: the return of an old public health problem. *Cancer causes & control*, 1994.
- 18) 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.
- 19) 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>.
- 20) 前園正宣, 小野智司, 中山茂. 遺伝的プログラミングを用いた画像フィルタ設計におけるパラメータ調整とプロット抑制. 日本計算工学会, Vol. 2006, No. 20060021, 2006.
- 21) 原田実. 自動プログラミングハンドブック. オーム社, 1989.
- 22) Frank Francone Peter Nordin and Wolfgang Banzhaf. Explicitly defined introns and destructive crossover in genetic programming. *Proc. of the Workshop on Genetic Programming: From Theory to Real-World Applications*, pp. 6–22, 1995.
- 23) M Fielding. Simulated annealing with an optimal fixed temperature. *j-SIAM-J-OPT*, Vol. 11, No. 2, pp. 289–307, 2000.
- 24) 三木光範, 上田祐一郎, 廣安知之. 受理率に依存する適応的近傍を持つシミュレーテッドアニーリングの実最適化問題における有効性 - 光通信用利得等化フィルタへの応用. 情報処理学会論文誌, Vol. 49, No. 10, pp. 3567–3575, 2008.
- 25) 昌達慶仁. 詳解 画像処理プログラミング. ソフトバンク クリエイティブ, 2008.
- 26) 井上誠喜, 八木伸行, 林正樹, 中須英輔, 三谷公二. C 言語で学ぶ実践画像処理. オーム社, 2008.

付 図

2.1	病理画像の例	2
2.2	病理画像診断支援システムのフロー	3
3.1	口腔癌 (舌癌)	4
4.1	木構造状フィルタによる画像処理の自動化の原理	5
4.2	学習用画像セットの一例	6
5.1	交叉	8
5.2	突然変異	8
6.1	学習用画像セット (単一画像学習)	10
6.2	プログラムサイズ制限の探索における学習性能	10
6.3	プログラムサイズ制限の探索におけるプログラムサイズの比較	11
6.4	GP(プログラムサイズ制限:25) と GP(交叉なし) におけるプログラムサイズの比較	11
6.5	一定温度探索における学習性能	12
6.6	一定温度探索におけるプログラムサイズの比較	13
6.7	冷却 SAP と一定温度 SAP の学習性能の比較	14
6.8	冷却 SAP と一定温度 SAP のプログラムサイズの比較	14
6.9	GP と SAP の学習性能の比較	15
6.10	出力結果	15
6.11	GP と SAP のプログラムサイズの比較	16
6.12	未知画像	16
6.13	未知画像に対する抽出対象部位	16
6.14	GP(制限:25) を未知画像に適用した結果	17
6.15	冷却 SAP を未知画像に適用した結果	17
7.1	ノイズ除去を加えた GP (Program Size=25) を未知画像に適用した結果	18
7.2	構築されたアルゴリズムにおける画像処理過程	19
7.3	学習画像セット (複数画像学習)	20
7.4	複数画像で学習した GP を未知画像に適用した結果	20
7.5	パラメータ最適化したフィルタを未知画像に適用した結果	21
7.6	特徴抽出アルゴリズムにおける画像処理経過 (未知画像 4)	21
7.7	TPF, FPF グラフによる比較	22

付 表

4.1	一入力の基本フィルタ	6
4.2	二入力の基本フィルタ	6
4.3	カラー画像を扱う際の終端記号	6
6.1	GP のパラメータ	10

6.2	一定温度 SAP のパラメータ	12
6.3	冷却 SAP のパラメータ	13