

高次局所自己相関特徴を用いた病理組織診断支援技術の研究

野里博和[†] 坂無英徳[†] 村川正宏[†]
樋口哲也[†] 大津展之[†] 寺井謙介^{††}
蛭田啓之^{††} 亀田典章^{††}

本論文では、病理組織診断におけるがん組織の自動判別を可能にする、高次局所自己相関特徴 (HLAC) を用いた病理組織診断支援技術を提案する。現在の病院における病理組織診断では、専門資格を有する病理医による顕微鏡観察により組織診断が行われている。しかし、病理医不足による負担の増加は、社会問題となりつつある。そこで、本論文では、病理組織診断における病理医の負担を軽減する、高次局所自己相関特徴を用いた病理診断支援技術を提案する。提案手法では、非癌の顕微鏡画像から抽出した高次局所自己相関特徴の主成分分析により学習し、癌/非癌の判別を行う。本提案手法により、明らかに非癌である組織の自動スクリーニングが可能となり、病理医は癌および癌と疑わしい組織の診断に集中することが可能となる。

Histopathological Diagnostic Support Technology using Higher-order Local Autocorrelation Features

HIROKAZU NOSATO,[†] HIDENORI SAKANASHI,[†]
MASAHIRO MURAKAWA,[†] TETSUYA HIGUCHI,[†] NOBUYUKI OTSU,[†]
KENSUKE TERAI,^{††} NOBUYUKI HIRUTA^{††} and NORIAKI KAMEDA^{††}

This paper proposes a histopathological diagnostic support technology by using correlation-based features of a pathological specimen. Recently, the shortage of clinical pathologists is posing increasing burden on them, and this has caused a serious social problem. In order to overcome this problem, we propose a histopathological diagnostic support technology using higher-order local autocorrelation (HLAC) features. In our method, we use principal component analysis (PCA) and find the eigenvectors to form the subspace of normal features, and we discriminate between tumor and non-tumor by using abnormal value calculated based on the subspace of normal features. Therefore, because this proposed method can automatically screen out non-tumor tissue, clinical pathologists can concentrate on the diagnosis of cancer and their burden can be reduced.

1. はじめに

現在、がんの確定診断は、専門の知識と経験を有する病理専門医 (以下、病理医とする) によって行われており、特に患者から採取される組織を診断する病理組織診断では、病理標本を作製し、顕微鏡観察によって、癌/非癌* などの振り分けおよび確定診断が行われている。近年の医療分野への IT 技術の導入によっ

て、病理標本作製の自動化、顕微鏡の性能向上、特殊・免疫染色技術の向上により、がんの診断精度は向上してきた。

しかしながら、現在、日本における病理医は絶対的に不足しており、病理医不足による負担の増加は、社会問題となりつつある。1998 年以降日本の死因順位の第 1 位は、悪性新生物 (がん) であり、現在の全死亡者に占める割合は、図 1 に示すように、他の主要死因に比べ格段に多い。そのため、病理医への負担は日に日に増している。しかし、根本的な病理医不足が解消するめどは立っておらず、病理医の負担を軽減する新しい技術の開発が求められている。

そこで、この問題を解決するために、我々は、病理組織診断において、病理医の負担を減らすため、明らかな非癌組織を非癌として自動判別する、高次局所

[†] 産業技術総合研究所
National Institute of Advanced Industrial Science and
Technology (AIST)

^{††} 東邦大学医療センター 佐倉病院
Toho University Sakura Medical Center

* 実際は、非癌組織として、良性腫瘍や偽・疑陽性病変などが分類されるが、本論文では、正常に近い組織のみを非癌とし、それ以外の癌あるいは癌と疑わしい組織を癌とする。

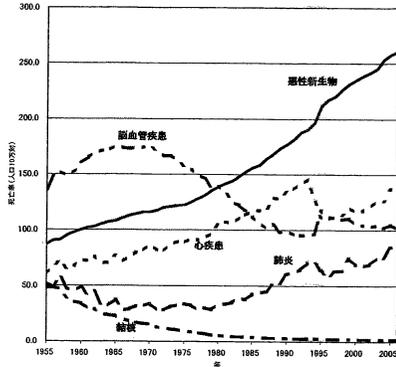


図1 主要死因別粗死亡率年次推移 (1955年～2006年) (資料: 国立がんセンターがん対策情報センター 冊子「がんの統計 2008年」²⁾)

Fig. 1 Trends in Crude Mortality Rate for Leading Causes of Death (1955～2006).

自己相関特徴 (Higher-order Local AutoCorrelation: HLAC)³⁾ を用いた病理診断支援技術を提案する。提案手法では、胃生検における非癌の病理組織顕微鏡画像から高次局所自己相関特徴を抽出し、その特徴の主成分分析により学習することで、癌/非癌の判別を行う。この提案手法により、これまで病理医が診断していた非癌病理標本を自動スクリーニングすることが可能となり、病理医が癌あるいは癌と疑わしい標本の診断に集中することで、診断効率を高めるとともに、見落としなどの発生確率を減らすことが可能となる。

2. 病理組織診断と自動診断の現状

2.1 病理組織診断とその問題点

現在の医療の中で、患者から採取された組織を病理形態学的方法により診断を行う病理組織診断は、臨床医が、より適切な治療を行うために重要な役割を担っている。この病理組織診断には、臓器・組織の一部を針やメス、内視鏡で採取して診断する生検組織診断や、手術によって摘出された臓器・組織を診断する外科組織診断、手術中に切除された病変から凍結標本を作製して良・悪性や手術の適・不適を診断する術中迅速病理診断などがある。これら病理組織診断では、患者の生死にかかわる“がん”の診断が行われており、その診断の精度・迅速性が最も重要視されている。

しかしながら、現在、日本における病理医は絶対的に不足しており、今後さらに需要が増加することを考えると、病理医不足は2015年までに大きな社会問題となる恐れがある¹⁾。そのため、この病理医不足の問題は、近年のがん患者数増加に伴う病理組織診断件数

の増加により、病理医の負担が増え、それが原因で見落としなどの重大な問題に発展する可能性がある。現在、このような状況において、根本的な病理医不足が解消するめどは立っておらず、病理医の負担を軽減する新しい技術開発が求められている。

2.2 自動診断の現状

病理医不足を解消するための技術として注目されているのが、診断の自動化技術である。病理組織診断と同じ、病理診断のひとつである細胞診の分野では、1967年から開発が進められている、田中昇らによる子宮がん早期発見のための細胞診スクリーニング自動化装置 CYBEST⁴⁾⁵⁾ や、1995年に米国FDAの承認を修得した子宮頸部スメアの自動診断装置 AutoPap⁶⁾ などが開発されており、数多くの知見が報告されている。

一方、病理組織診断においては、顕微鏡のIT化による飛躍的な進歩により研究が活発になっており、最新の顕微鏡システムである、バーチャルスライドシステムの拡張技術として、癌自動診断装置の開発が行われている⁷⁾。また、がんの病理診断に関して、産業技術総合研究所と神奈川県立がんセンターが、正常な組織・細胞からの逸脱度 (異常値) をコンピュータが計測することによって、正常な組織・細胞か、がん組織・細胞かを認識する手法の共同研究を進めている⁸⁾。

3. 提案手法

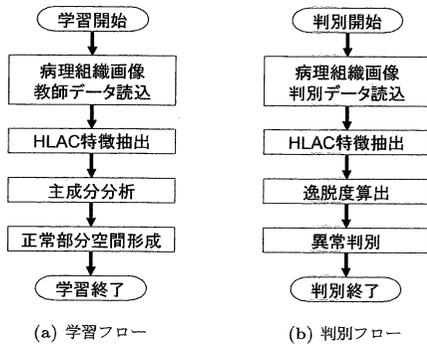
本論文では、問題を解決するアプローチとして、高次局所自己相関特徴 (HLAC)³⁾ を用いた病理組織診断支援技術を提案する。提案手法では、胃生検における非癌の病理組織顕微鏡画像を、高次局所自己相関特徴を用いた主成分分析により学習し、癌/非癌の判別を行う。この高次局所自己相関特徴は、位置に対する不変性と画面に関する加法性の特徴を持つ特徴抽出の1つで、主成分分析と組み合わせることで、有効な特徴が効率良く抽出される。この高次局所自己相関特徴は、多くの実用化を目指した研究がされており、監視カメラによる異常検知技術や、医療情報処理応用に関して高次局所自己相関の応用の可能性が示されている⁹⁾。

本提案手法は、図2(a)に示す学習過程と図2(b)に示す判別過程の2段階に分けることができる。以下では、主な処理手順について具体的に述べていく。

3.1 高次局所自己相関特徴

高次局所自己相関特徴は、対象画像を $f(r)$ とすると、 N 次自己相関関数は、変位方向 (a_1, \dots, a_N) に対して、

$$x_N = \int f(r)f(r+a_1)\cdots f(r+a_N)dr \quad (1)$$



(a) 学習フロー (b) 判別フロー

図2 提案手法フロー

Fig. 2 Flowchart of proposed method.

で定義される。

本論文では、1つの対象画像に対し、高次自己相関関数係数の次数を $N = 2$ とし、次変位方向 a を参照点 r の周囲の 3×3 画素の領域に限定することで、図3に示すような25個の局所パターンから算出した25次元のベクトルを特徴ベクトルとする。

3.2 主成分分析

特徴ベクトル x に主成分分析¹⁰⁾を適用することで、次節後述する正常部分空間¹¹⁾を形成する規底ベクトルを算出することができる。これは、HLAC特徴で構成された特徴ベクトルの自己相関行列 R_x の固有ベクトルにより求められる。

$$R_x U = U \Lambda \quad (2)$$

ここで、 U は固有ベクトルを列とする行列であり、 Λ は固有値を対角要素とする対角行列である。

3.3 正常部分空間と異常判別

正常部分空間を形成する次元数 K は、累積寄与率

$$\eta_K = \frac{\sum_{i=1}^K \lambda_i}{\sum_{i=1}^M \lambda_i} \quad (3)$$

によって決定される。このとき、正常部分空間は、正常累積寄与率 $\eta_K \geq C$ を満たす次元 K までの固有ベクトル $U_K = \{u_1, \dots, u_K\}$ を規定ベクトルにより張られる空間とする。

本論文では、このようにして得られた正常部分空間からの逸脱度を異常判別の指標とする。この逸脱度は、正常部分空間の直交補空間への射影成分の距離として次のように算出することができる。正常部分空間への射影子は、 $P = U_K U_K^T$ として表され、それに対する直交補空間への射影子は、 $P_\perp = I_M - P$ となる。このとき、直交補空間への射影成分の距離 d_\perp は、

$$d_\perp^2 = \|P_\perp x\|^2 = x^T (I_M - U_K U_K^T) x \quad (4)$$

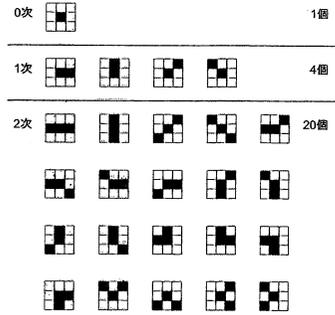


図3 2次までの高次局所自己相関 (HLAC) の変位パターン
Fig. 3 HLAC Patterns up to 2nd order correlation.

と表すことができる。

4. 検証実験

本章では、提案手法の有効性を検証するために行った検証実験について報告する。実験では、実際の胃生検サンプルに対して、提案手法を用いて学習・判別の検証実験を行った。本論文では、提案手法により非癌サンプルの学習により、癌データがきちんと判別できるかどうかの確認を行うために、予め病理医により診断されている明らかな非癌データと明らかな癌データを用いた実験を行った。

4.1 実験データ

学習データとして、病理医によって非癌と診断された50サンプルを用い、判別データには、学習データも含め、図4に示すような非癌、癌、非癌・癌混合のデータを62サンプル用いた。なお、各画像データには、顕微鏡倍率20倍で撮影された2560x1920のjpegカラー画像を反転した後、大津の判別二値化手法¹²⁾にて二値化処理を行った。

4.2 検証実験結果

図5に検証実験結果の逸脱度分布を示す。グラフの縦軸がサンプル数、横軸が逸脱度 d_\perp ($\times 10^5$) である。白抜きが非癌50サンプル、黒が癌6サンプル、灰色が癌・非癌混合6サンプルの度数を示す。図の分布より、癌と混合サンプルの分布は、非癌サンプルの分布と区別可能なことがわかる。

図に示す逸脱度 ($d_\perp = 4.8 \times 10^5$) に閾値を設定した場合、癌と混合サンプルは、異常値を示すことになり、判別することが可能となる。しかし、学習した非癌サンプルの内、3サンプルも閾値を超え、結果としてこれらの非癌サンプルも異常値を示した。この結果は、本論文における学習サンプル数が少ないためと考えられ、さらに多くの非癌サンプルを学習することで、

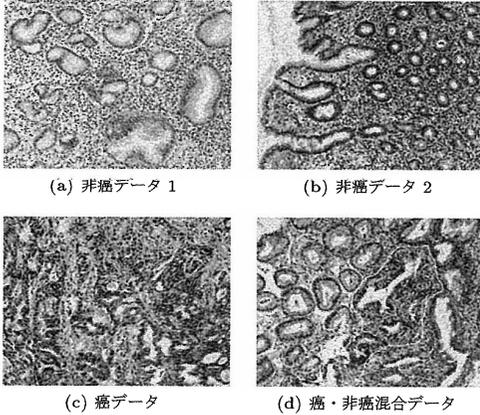


図 4 検証実験データ
Fig. 4 Experimental pathological specimen.

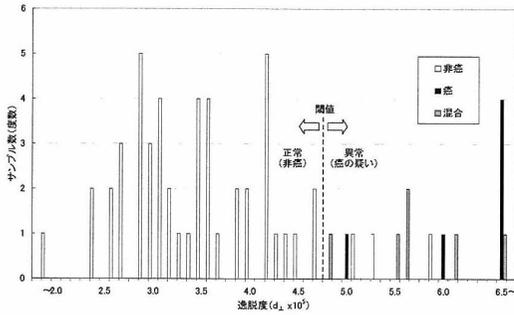


図 5 逸脱度分布
Fig. 5 Frequency deviance.

解決できると考える。

5. おわりに

本論文では、病理医不足の問題を解決するアプローチとして、高次局所自己相関特徴 (HLAC) を用いた病理組織診断支援技術を提案した。提案手法では、胃生検における非癌の病理組織顕微鏡画像を高次局所自己相関特徴を用いた主成分分析により学習し、癌/非癌の判別を行った。検証実験の結果、提案手法により非癌サンプルの学習により、癌データがきちんと判別できることを確認した。

本論文では、予め病理医により診断されている明らかな非癌と明らかな癌の二極化したデータを用いたが、実際の胃生検の病理診断は、その細胞形態や組織構築の異型性などによって、5段階にグループ分類されている。実用化を目指す上で、本提案手法による自動判別に求められる役割は、明らかに非癌組織である正常に近い組織だけを選別することであるので、今後は、

明らかな癌データに加え、非癌組織であるが、病理医の判断を必要とする境界病変などのデータも含めた検証と手法の改良が必要である。

本提案手法が、病理医の負担を軽減し、経験と知識を必要とする腫瘍かどうかの診断に専念させることにより、その診断効率を高めるとともに、見落としなどの発生確率を減らし、より質の高い医療の一端を担う新しい医療技術となることを期待する。

参考文献

- 1) 濃沼信夫: “病理医をめぐる課題と医療制度改革の展望”, 病理と臨床, 23(9), pp. 1025-1030 (2005)
- 2) がんの統計 2008 年, 国立がんセンター がん対策情報センター, http://ganjoho.ncc.go.jp/public/statistics/backnumber/2008_jp.html
- 3) N. Otsu and T. Kurita: “A new scheme for practical flexible and intelligent vision systems,” Proc. IAPR Workshop on Computer Vision, pp. 431-435 (1988)
- 4) 田中昇: “細胞診の自動化について”, 癌の臨床, 42(9), pp. 893-895 (1996)
- 5) 田中昇: “Overview 細胞診自動化: 歴史と実用化にむけて”, 日本臨床細胞学会雑誌, 39(Suppl.1), pp 61 (2000)
- 6) 森川政夫, 清水章, 岡本吉明, 植田政嗣, 植木實, 布引治, 野田定, 伊藤博俊: “細胞診自動診断装置 AutoPap の評価: (第 1 報)”, 日本臨床細胞学会雑誌, 37(Suppl.1), pp 126 (1998)
- 7) 佐藤達賢, 三浦富智, 野坂大喜, 方山揚誠, 森内昭, 中村真一, 鄭衆喜, 高松輝賢: “病理検査の進歩: バーチャルスライド作成システムの開発とその活用”, 臨床病理, 55(4), pp. 344-350 (2007)
- 8) “コンピュータによるがんの病理診断を支援する世界初の画期的なシステムを開発します!”, 神奈川県記者発表資料 (知事会見), 2008.6.17, <http://www.pref.kanagawa.jp/press/0806/051/index.html>
- 9) 医療と情報技術の連携イノベーションフォーラム, 主催東京大学医学部附属病院・産総研, 東京大学本郷キャンパス, 2008.12.02, http://www.jita.or.jp/data/iryojoho_forum/2008poster.pdf
- 10) Richard O. Duda, Peter E. Hart and David G. Stork: Pattern Classification, pp. 114-117, Wiley-Interscience (2000)
- 11) Takuya Nanri and Nobuyuki Otsu: “Unsupervised Abnormality Detection in Video Surveillance,” Proc. IAPR Conf. on Machine Vision Application, pp. 574-577 (2005)
- 12) 大津展之: “判別および最小 2 乗基準に基づく自動しきい値選定法”, 電子通信学会論文誌 D, J63-D-4, pp. 349-356 (1980)