

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第6588444号
(P6588444)

(45) 発行日 令和1年10月9日(2019.10.9)

(24) 登録日 令和1年9月20日(2019.9.20)

(51) Int. Cl.	F I	
GO 1 N 33/48 (2006.01)	GO 1 N 33/48	M
GO 6 T 1/00 (2006.01)	GO 6 T 1/00	2 9 5
GO 6 T 7/00 (2017.01)	GO 6 T 7/00	2 5 0
GO 1 N 33/483 (2006.01)	GO 1 N 33/483	C
GO 1 N 21/27 (2006.01)	GO 1 N 33/48	Z
請求項の数 40 (全 32 頁) 最終頁に続く		

(21) 出願番号	特願2016-546837 (P2016-546837)	(73) 特許権者	507179346
(86) (22) 出願日	平成27年1月23日(2015.1.23)		ベンタナ メディカル システムズ, イ
(65) 公表番号	特表2017-510792 (P2017-510792A)		ンコーポレイテッド
(43) 公表日	平成29年4月13日(2017.4.13)		アメリカ合衆国アリゾナ州85755, ト
(86) 国際出願番号	PCT/EP2015/051302		ゥーソン, イースト・イノベーション・
(87) 国際公開番号	W02015/113895		パーク・ドライブ 1910
(87) 国際公開日	平成27年8月6日(2015.8.6)	(74) 代理人	100140109
審査請求日	平成30年1月9日(2018.1.9)		弁理士 小野 新次郎
(31) 優先権主張番号	61/932, 671	(74) 代理人	100075270
(32) 優先日	平成26年1月28日(2014.1.28)		弁理士 小林 泰
(33) 優先権主張国・地域又は機関	米国 (US)	(74) 代理人	100101373
			弁理士 竹内 茂雄
(31) 優先権主張番号	62/033, 261	(74) 代理人	100118902
(32) 優先日	平成26年8月5日(2014.8.5)		弁理士 山本 修
(33) 優先権主張国・地域又は機関	米国 (US)		
			最終頁に続く

(54) 【発明の名称】 全スライド組織セグメント化のための適応分類

(57) 【特許請求の範囲】

【請求項1】

組織標本の少なくとも二次元の画像を表す画像データを受け取り、前記画像データの複数のサブセットを識別する画像領域データを出力する画像領域識別子モジュール(112、710)であって、各画像データサブセットが、前記画像の固有の空間的に接触する領域を表す、画像領域識別子モジュール(112、710)と、

複数の前記画像領域に対して、組織特性のデータベース(116、217)に問い合わせを行うために、前記それぞれの画像データサブセットを使用して、前記それぞれの画像領域を複数の組織の種類のうちの一つとして分類し、前記それぞれの画像領域を前記分類するステップの信頼を示す信頼スコアを決定する画像領域分類子モジュール(114、720)と、

第1の範囲内に含まれる信頼スコアを有する複数の前記画像領域に対して、結果として得られた変更されたデータベース(218)が、前記それぞれの画像データサブセットから取得した前記それぞれの画像領域および組織特性データに分類された前記組織の種類を表すデータを含むように、前記データベースの変更を実施するデータベース変更子モジュール(730)であって、前記第1の範囲が、前記分類が正しいという確実性を表す信頼スコアを含む信頼スコアの範囲である、データベース変更子モジュール(730)と、

前記第1の範囲とははっきりと異なる第2の範囲内に含まれる信頼スコアを有する複数の前記画像領域に対して、前記変更されたデータベースに問い合わせを行うために、前記それぞれの画像データサブセットを使用して、前記それぞれの画像領域を前記複数の組織

の種類の中の1つとして再分類し、前記複数の組織の種類の中の前記1つを解析結果として出力する画像領域再分類子モジュール(115、740)とを備える、組織解析システム(700)。

【請求項2】

生の画像データを取得するために組織標本を画像化する組織イメージングデバイス(102、750)を備え、

前記受け取った画像データが、前記生の画像データから取得される、
請求項1に記載の組織解析システム。

【請求項3】

染色された組織標本を取得するために前記組織標本を染色する組織染色デバイス(770)と、

前記画像データを取得するために前記染色された組織標本を画像化する組織イメージングデバイス(102、750)とを備え、

前記受け取った画像データが、前記生の画像データから取得される
請求項2に記載の組織解析システム。

【請求項4】

前記データベースを記憶するデータ記憶システム(760)を備え、

前記データベースが、複数の組織画像領域の各々に対して、組織の少なくとも二次元の画像を表すデータと、少なくとも1つの組織特徴を表すデータと、組織の種類を表すデータと、信頼スコアを表すデータとを含み、前記データベースが、前記画像領域分類子モジュールによって前記画像分類および信頼スコアの決定に使用される、事前構築された第1の訓練データベース(217)であり、前記データベース変更子モジュール(730)が、第1の範囲内に含まれる信頼スコアを有する複数の前記画像領域の前記画像データサブセットから取得した前記画像領域および前記組織特性データに分類された前記組織の種類を表す前記データによって構成される第2の訓練データベース(219)を生成し、前記第1および前記第2の訓練データベースを組み合わせて前記変更されたデータベース(218)を提供するように構成される、請求項1から3のいずれか一項に記載の組織解析システム。

【請求項5】

前記画像領域分類子モジュールが、サポートベクタマシン(722)を備え、前記信頼スコアを決定するために前記サポートベクタマシンの出力を使用する、請求項1から4のいずれか一項に記載の組織解析システム。

【請求項6】

前記それぞれの画像領域を前記分類するステップが、前記それぞれの画像データサブセットおよび前記データベースから取得した前記データを使用して少なくとも1つの特徴を前記それぞれの画像領域から抽出するステップを含み、前記特徴がテクスチャ特徴と、生物学的特徴と、強度特徴と、傾斜特徴と、ガボール特徴と、共起特徴と、核特徴とからなる群に属する、請求項1から5のいずれか一項に記載の組織解析システム。

【請求項7】

前記それぞれの画像領域を前記再分類するステップが、空間的近接値と、信頼スコアと、特徴類似値とのうちの少なくとも1つを使用して、前記それぞれの画像データサブセットのデータと前記変更されたデータベースから取得した前記データとを重み付けするステップを含む、請求項1から6のいずれか一項に記載の組織解析システム。

【請求項8】

前記画像領域分類子モジュールが、画像チャンネル抽出器(111、724)を備え、

前記それぞれの画像領域を前記分類するステップが、前記画像チャンネル抽出器を使用して、前記画像の少なくとも前記それぞれの領域を1つまたは複数の構成要素チャンネルに分離するステップと、前記データベースから取得した前記それぞれの画像領域および前記データの前記構成要素チャンネルのうちのいずれかを使用して、少なくとも1つの特徴を前記それぞれの画像領域から抽出するステップとを備え、

10

20

30

40

50

前記特徴が、テクスチャ特徴と、生物学的特徴と、強度特徴と、傾斜特徴と、ガボール特徴と、共起特徴と、核特徴とからなる群に属し、

前記構成要素チャンネルが、ヘマトキシリンチャンネルと、エオシンチャンネルと、輝度チャンネルとからなる群に属する、請求項 1 から 6 のいずれか一項に記載の組織解析システム。

【請求項 9】

前記画像領域分類子モジュールが、前記データベースのデータを使用して、分類子論理を構築し、前記画像領域の各々に対して、前記それぞれの組織の種類および前記それぞれの信頼スコアを決定するために前記分類子論理を前記それぞれの画像領域の前記画像データサブセットに適用し、

前記データベース変更子モジュールが、前記第 1 の範囲内に含まれる信頼スコアを有する画像領域と前記第 2 の範囲内に含まれる信頼スコアを有する画像領域とを見つけ、前記変更されたデータベースを取得するために前記データベースと前記それぞれの組織の種類と前記第 1 の範囲内に含まれる信頼スコアを有することが見つけられた前記画像領域の前記それぞれの画像データサブセットとを組み合わせ、

前記画像領域再分類子モジュールが、前記変更されたデータベースのデータを使用して機械学習により前記分類子論理を変更し、前記第 2 の範囲内に含まれる信頼スコアを有することが見つけられた前記画像領域の各々に対して、前記それぞれの再分類された組織の種類を決定するために、前記変更された分類子論理を前記それぞれの画像領域の前記画像データサブセットに適用し、

前記システムが、前記第 1 の範囲内に含まれる信頼スコアを有することが見つけられた前記画像領域の各々の前記それぞれの組織の種類と、前記第 2 の範囲内に含まれる信頼スコアを有することが見つけられた前記画像領域の各々の前記それぞれの再分類された組織の種類とを分類結果として出力する、請求項 1 から 8 のいずれか一項に記載の組織解析システム。

【請求項 10】

前記データベース変更子モジュールが、前記複数の組織の種類の各々に対して、および前記第 1 の範囲内に含まれる信頼スコアを有する画像領域だけに対して、前記それぞれの組織の種類を有する画像領域の総数を決定し、前記それぞれの組織の種類を有する画像領域の総数が前記それぞれの組織の種類のそれぞれの閾値数を超える組織の種類だけに対して、前記データベースの前記変更を実施する、請求項 1 から 9 のいずれか一項に記載の組織解析システム。

【請求項 11】

前記それぞれの組織の種類を有する画像領域の総数が前記それぞれの組織の種類のそれぞれの閾値数を超える組織の種類に対して、前記データベース変更子モジュールが、前記変更されたデータベースが前記それぞれの画像データサブセットから取得した組織特性データだけを含むように、前記データベースの前記変更を実施する、請求項 10 に記載の組織解析システム。

【請求項 12】

組織標本の少なくとも二次元の画像を表す画像データを受け取るステップ(830)と、

前記画像データの複数のサブセットを識別する画像領域データを生成するステップ(840)であって、各画像データサブセットが、前記画像の固有の空間的に接触する領域を表す、生成するステップ(840)と、

複数の前記画像領域に対して、前記それぞれの画像データサブセットを使用して、および組織特性のデータベースから取得したデータを使用して、前記それぞれの画像領域を複数の組織の種類のうちの 1 つとして分類するステップ(850)と、

前記それぞれの画像領域を前記分類するステップの信頼を示す信頼スコアを決定するステップ(860)と、

第 1 の範囲内に含まれる信頼スコアを有する複数の前記画像領域に対して、結果として得られた変更されたデータベース(218)が、前記それぞれの画像データサブセットが

10

20

30

40

50

ら取得した前記それぞれの画像領域および組織特性データに分類された前記組織の種類を表すデータを含むように、前記データベースの変更を実施するステップ(870)であって、前記第1の範囲が、前記分類が正しいという確実性を表す信頼スコアを含む信頼スコアの範囲である、実施するステップ(870)と、

前記第1の範囲とははっきりと異なる第2の範囲内に含まれる信頼スコアを有する複数の前記画像領域に対して、前記変更されたデータベースに問い合わせを行うために、前記それぞれの画像データサブセットを使用して、前記それぞれの画像領域を前記複数の組織の種類の中の1つとして再分類するステップ(880)と、

前記第1の範囲とははっきりと異なる第2の範囲内に含まれる信頼スコアを有する前記画像領域の中の少なくとも1つに対して、前記それぞれの画像領域の前記再分類された組織の種類を出力するステップ(890)とを含む、組織解析方法。

10

【請求項13】

生の画像データを取得するために組織標本を画像化するステップ(820)であって、前記受け取った画像データが、前記生の画像データから取得される、画像化するステップ(820)を含む、請求項12に記載の組織解析方法。

【請求項14】

染色された組織標本を取得するために前記組織標本を染色するステップ(810)と、前記生の画像データを取得するために前記染色された組織標本を画像化するステップ(820)とを含み、

前記受け取った画像データが、前記生の画像データから取得される、請求項13に記載の組織解析方法。

20

【請求項15】

前記データベースが、複数の組織画像領域の各々に対して、組織の少なくとも二次元の画像を表すデータと、少なくとも1つの組織特徴を表すデータと、組織の種類を表すデータと、信頼スコアを表すデータとを含み、前記データベースが、画像領域分類子モジュールによる前記画像分類および信頼スコアの決定に使用される事前構築された第1の訓練データベース(217)であり、データベース変更子モジュール(730)が、前記第1の範囲内に含まれる信頼スコアを有する複数の前記画像領域の前記画像データサブセットから取得した前記画像領域および前記組織特性データに分類された前記組織の種類を表す前記データによって構成される第2の訓練データベース(219)を生成し、前記第2の訓練データベースの内容の一部または全部を前記第1のデータベースにコピーすることによってなど、前記変更されたデータベース(218)を提供するために前記第1の訓練データベースと前記第2の訓練データベースとを組み合わせる、請求項12から14のいずれか一項に記載の組織解析方法。

30

【請求項16】

信頼スコアを前記決定するステップが、サポートベクタマシンの動作(862)を含む、請求項12から15のいずれか一項に記載の組織解析方法。

【請求項17】

前記それぞれの画像領域を前記分類するステップが、前記それぞれの画像データサブセットと前記データベースから取得した前記データとを使用して、前記それぞれの画像領域から少なくとも1つの特徴を抽出するステップ(854)を含み、前記特徴が、テクスチャ特徴と、生物学的特徴と、強度特徴と、傾斜特徴と、ガボール特徴と、共起特徴と、核特徴とからなる群に属する、請求項12から16のいずれか一項に記載の組織解析方法。

40

【請求項18】

前記それぞれの画像領域を前記再分類するステップが、空間的近接値と、信頼スコアと、特徴類似値とのうちの少なくとも1つを使用して、前記それぞれの画像データサブセットのデータと前記変更されたデータベースから取得した前記データとを重み付けするステップ(882)を含む、請求項12から17のいずれか一項に記載の組織解析方法。

【請求項19】

前記それぞれの画像領域を前記分類するステップが、前記画像の少なくとも前記それぞ

50

れの領域を1つまたは複数の構成要素チャネルに分離するステップ(852)と、前記それぞれの画像領域の前記構成要素チャネルおよび前記データベースから取得した前記データを使用して、少なくとも1つの特徴を前記それぞれの画像領域から抽出するステップ(854)とを含み、

前記特徴が、テクスチャ特徴と、生物学的特徴と、強度特徴と、傾斜特徴と、ガボール特徴と、共起特徴と、核特徴とからなる群に属し、

前記構成要素チャネルが、ヘマトキシリンチャネルと、エオシンチャネルと、輝度チャネルとからなる群に属する、請求項12から17のいずれか一項に記載の組織解析方法。

【請求項20】

前記データベースのデータを使用して分類子論理を構築するステップと、

前記画像領域の各々に対して、前記それぞれの組織の種類と前記それぞれの信頼スコアとを決定するために、前記分類子論理を前記それぞれの画像領域の前記画像データサブセットに適用するステップと、

前記第1の範囲内に含まれる信頼スコアを有する画像領域と、前記第2の範囲内に含まれる信頼スコアを有する画像領域とを見つけるステップと、

前記変更されたデータベースを取得するために前記データベースと、前記それぞれの組織の種類と、前記第1の範囲内に含まれる信頼スコアを有することが見つけられた前記画像領域の前記それぞれの画像データサブセットとを組み合わせるステップと、

前記変更されたデータベースのデータを使用して、機械学習により、前記分類子論理を変更するステップと、

前記第2の範囲内に含まれる信頼スコアを有することが見つけられた前記画像領域の各々に対して、前記それぞれの再分類された組織の種類を決定するために前記変更された分類子論理を前記それぞれの画像領域の前記画像データサブセットに適用するステップと、

前記第1の範囲内に含まれる信頼スコアを有することが見つけられた前記画像領域の各々の前記それぞれの組織の種類と、前記第2の範囲内に含まれる信頼スコアを有することが見つけられた前記画像領域の各々の前記それぞれの再分類された組織の種類とを分類結果として出力するステップとを含む、請求項12から19のいずれか一項に記載の組織解析方法。

【請求項21】

前記複数の組織の種類の各々に対して、および前記第1の範囲内に含まれる信頼スコアを有する画像領域だけに対して、前記それぞれの組織の種類を有する画像領域の総数を決定するステップを備え、

前記データベースの前記変更が、前記それぞれの組織の種類を有する画像領域の総数が前記それぞれの組織の種類のそれぞれの閾値数を超える組織の種類に対してのみ実施される、請求項12から20のいずれか一項に記載の組織解析方法。

【請求項22】

前記それぞれの組織の種類を有する画像領域の総数が前記それぞれの組織の種類のそれぞれの閾値数を超える組織の種類に対して、前記データベースの前記変更が、前記変更されたデータベースが前記それぞれの画像データサブセットから取得した組織特性データだけを含むように実施される、請求項21に記載の組織解析方法。

【請求項23】

電子デバイスによって実装される組織画像のセグメント化のための方法であって、

前記組織画像内の格子点を識別するステップと、

複数の組織の種類のうちの1つとして前記格子点を分類するステップ、および組織の種類の既知の特性のデータベースに基づいて分類された格子点を生成するステップと、

前記分類された格子点に高信頼スコアおよび低信頼スコアのうちの少なくとも1つを割り当てるステップと、

高信頼スコアが割り当てられた前記格子点に基づいて組織の種類の既知の特性の前記データベースを変更するステップ、および変更されたデータベースを生成するステップと、

前記変更されたデータベースに基づいて低信頼スコアが割り当てられた前記格子点を再

10

20

30

40

50

分類するステップとを含む、方法。

【請求項 24】

前記格子点が六角形構成で配列される、請求項 23 に記載の方法。

【請求項 25】

前記複数の組織の種類が、正常組織、腫瘍、壊死、間質、またはリンパ球凝集体のうちの 1 つまたは複数を含む、請求項 23 に記載の方法。

【請求項 26】

前記高信頼スコアと低信頼スコアとを割り当てるステップが、サポートベクタマシン (SVM) の動作を使用するステップをさらに含む、請求項 23 に記載の方法。

【請求項 27】

前記格子点が、前記組織画像に対して一定の間隔でサンプリングされる、請求項 23 に記載の方法。

【請求項 28】

前記格子点を分類するステップが、各格子点の近傍内から特徴を抽出するステップをさらに含む、請求項 23 に記載の方法。

【請求項 29】

前記特徴が、1 つまたは複数のテクスチャ特徴または生物学的特徴をさらに含む、請求項 28 に記載の方法。

【請求項 30】

特徴を抽出する前に、前記組織画像を 1 つまたは複数の構成要素チャンネルに分離するステップをさらに含む、請求項 28 に記載の方法。

【請求項 31】

前記 1 つまたは複数の構成要素チャンネルが、ヘマトキシリンチャンネル、エオシンチャンネル、または輝度チャンネルの 1 つまたは複数を含む、請求項 30 に記載の方法。

【請求項 32】

前記特徴が、強度特徴と、傾斜特徴と、ガボール特徴と、共起特徴と、核特徴とを含む、請求項 28 に記載の方法。

【請求項 33】

前記データベースを変更するステップが、低信頼スコアを有する前記格子点を高信頼スコアを有する前記格子点と比較するステップと、前記割り当てられた高信頼格子点に基づく前記割り当てられた低信頼格子点を変更するステップとをさらに含む、請求項 23 に記載の方法。

【請求項 34】

前記比較が、ランダムフォレストの動作に基づく、請求項 33 に記載の方法。

【請求項 35】

前記比較が、前記低信頼格子点にごく接近した前記高信頼格子点のラベルのパーセンテージに基づく請求項 33 に記載の方法。

【請求項 36】

前記低信頼格子点にごく近接した前記高信頼格子点が、各低信頼格子点の前記ラベルを改良する方により大きな寄与をする、請求項 35 に記載の方法。

【請求項 37】

前記近接が 80 画素以内である、請求項 35 に記載の方法。

【請求項 38】

前記変更されたデータベースが、前記組織画像に関連付けられたものとして記憶される、請求項 33 に記載の方法。

【請求項 39】

各格子点と各格子点の隣接した画素とが画像パッチを表す、請求項 23 に記載の方法。

【請求項 40】

前記組織画像によって構成された 1 つまたは複数の画像パッチの信頼マップを生成するステップをさらに含む、請求項 39 に記載の方法。

10

20

30

40

50

【発明の詳細な説明】

【技術分野】

【0001】

[1]本出願は、参照によりその内容がその全体において本開示に組み込まれている、2014年1月28日出願の米国仮特許出願第61/932,671号および2014年8月5日出願の米国仮特許出願第62/033,261号の優先権を主張するものである。

【0002】

[2]本開示は組織構造を識別することに関する。詳しくは、本開示は、組織の画像をセグメント化するためのシステムおよび方法を対象にする。

【背景技術】

10

【0003】

[3]病理組織画像からの組織セグメント化は、デジタルパソロジー (digital pathology) において重要な問題である。全スライド (whole slide) (WS) 組織画像を所与として、多くの適用例は、正常組織、腫瘍、壊死、リンパ球、および間質など、画像内の組織領域の異なる種類の識別を必要とする。これらの領域の正確な識別は、貴重な診断情報を提供するのに役立つことがある。例えば、標本内のそのような部分の存在の定量的評価は、化学療法などの療法の影響を決定するのに有益であり得る。組織画像セグメント化は、これまで、様々な方法によって対処されてきた。一般に、自動組織画像セグメント化は、特徴抽出および分類を含む機械学習法によって達成される。例えば、小さなパッチを画像の各画素の周りから抽出することができ、様々な方法を使用して、パッチから特徴を抽出することができる。Masahiro, I.らは、画像からスーパー画素をセグメント化し、間質に対応するものとしてリンパ球濃度およびファイバ確率を識別することによって、肝組織画像における間質をセグメント化することに関する。Ozseven, T.らは、サポートベクタマシン (SVM) アルゴリズムおよびガボールフィルタを使用して肝組織上の壊死部分を定量化することに関する。Doyle, S.らは、一次統計、共起、およびガボール特徴などのテクスチャ特徴の豊富な組を計算し、続いてアダブーストを使用して特徴選択をし、それによって異なる解像度で画素分類を実施することを含む前立腺組織画像のがん検出方法を論じる。Sertel, O.らは、ヘッシアン行列から計算された共起特徴および構造的特徴を使用して、組織画像を間質に富んだ、分化する、分化の乏しいおよび未分化の領域に区分化することによって神経芽腫組織スライドを解析することに関する。Nayak, N.らは、辞書学習方法を組織セグメント化問題に適用することに関する。Xu, Y.らは、異なるがんの種類領域への結腸組織画像のセグメント化のために弱教師あり学習とも呼ばれる多事例学習方法を採用することに関する。

20

30

【発明の概要】

【発明が解決しようとする課題】

【0004】

[4]しかし、これらの方法は、画像間に大きな変動があり、訓練標本が限定されることにより不十分である。さらに、訓練画像の手動注釈は、高倍率でWS画像のサイズが大きく、処理されるデータが多量であることにより、骨の折れることである。したがって、従来技術の方法の限定されたセグメント化の精度では、要求が満たされないままとなる。

40

【課題を解決するための手段】

【0005】

本発明は、それぞれの独立請求項において特許請求される、組織解析システムおよび方法、ならびに組織画像のセグメントのための方法ならびに組織画像の適応分類のためのシステムを提供する。本発明の実施形態は、従属請求項に示される。

【0006】

[5]本明細書に開示されるのは、2段階分類方法を使用して上記識別された問題に対処するシステムおよび方法である。本明細書に開示される動作は、WS画像を複数のパッチに分割するステップと、SVMなどの「ソフト」分類を使用して各パッチを第1の分類す

50

るステップと、各パッチの信頼スコアおよびラベルを生成するステップとを含む。分類結果として取得された各パッチの位置、その特徴、その組織の種類、およびその信頼スコアは、データベースに記憶することができる。第2の分類ステップは、低信頼パッチをデータベース内の高信頼パッチと比較するステップと、データベース内のパッチの空間コヒーレンスを増大させるために同様のパッチを使用するステップとを含む。言い換えれば、各低信頼パッチでは、隣接した高信頼パッチが各パッチのラベルを改良する方により大きな寄与をし、それによって、低信頼パッチにおけるセグメント化精度が改善される。訓練データベースを成長させるための既存の適応/能動学習技法と対照的に、開示される動作は単一の訓練データベースを成長させることにあまり関係せず、その代わりに、各テスト画像を独立に処置することに焦点を置きながら、解析中の画像の信頼情報のラベル付けに基づいて分類精度を適応的に改善する。言い換えれば、信頼ラベルパッチデータベースが各画像に対して生成され、類似検索動作が低信頼パッチの分類結果を改良するために画像内で実施される。

10

【0007】

[6]1つの例示的な実施形態において、本開示は、組織画像内の格子点を識別するステップと、複数の組織の種類の中の1つとして格子点を分類するステップ、および組織の種類の見知の特性のデータベースに基づいて分類された格子点を生成するステップと、分類された格子点に高信頼スコアおよび低信頼スコアの中の1つを割り当てるステップと、高信頼スコアが割り当てられた格子点に基づいて組織の種類の見知の特性のデータベースを変更するステップ、および変更されたデータベースを生成するステップと、変更されたデータベースに基づいて低信頼スコアが割り当てられた格子点を再分類するステップとを含む、組織画像のセグメント化のための方法である。方法はコンピュータ実装方法でよい。

20

【0008】

[7]別の例示的な実施形態において、本開示は、前記組織標本に関連付けられた見知の特徴のデータベースとの比較に基づいて組織の種類および信頼スコアを組織標本の画像パッチに割り当てるステップと、同じ組織標本からの1つまたは複数の高信頼画像パッチと画像パッチとの比較に基づいて画像パッチの組織の種類および信頼スコアを改良するステップとを含む動作を実施するために電子デバイスのプロセッサによって実行可能なデジタルで符号化された命令を記憶するためのデジタル記憶媒体であって、高信頼画像パッチが組織標本に関連付けられた高信頼画像パッチのデータベースに記憶される、デジタル記憶媒体である。電子デバイスは、並列処理を支援することができるイメージングシステムなどのシングルまたはマルチプロセッサデータ処理システムを備えることができる。

30

【0009】

[8]さらに別の例示的な実施形態において、本開示は、組織画像の適応分類のためのシステムであって、プロセッサと、プロセッサに通信可能に結合されたメモリであって、ソフト分類に基づいて複数の組織の種類の中の1つとして組織画像内の画素を分類するステップと、画素を画素の分類を改良するために高信頼スコアを有する1つまたは複数の隣接した画素と比較するステップとを含む動作を実施するためにプロセッサによって実行可能であるデジタルで符号化された命令を記憶するメモリとを含むシステムであって、1つまたは複数の隣接した画素の高信頼スコアがソフト分類に基づく、システムである。

40

【図面の簡単な説明】**【0010】**

【図1】[9]本開示の例示的な実施形態による、全スライド画像の適応分類のためのシステムを例示する図である。

【図2A】[10]本開示の例示的な実施形態による、全スライド画像の適応分類のための方法を例示する図である。

【図2B】本開示の例示的な実施形態による、全スライド画像の適応分類のための方法を例示する図である。

【図2C】本開示の例示的な実施形態による、全スライド画像の適応分類のための方法を

50

例示する図である。

【図3】[11]図3 Aは本開示の例示的な実施形態による、5つの異なる組織の種類画像チャンネルのうちの1つを例示する図である。図3 Bは本開示の例示的な実施形態による、5つの異なる組織の種類画像チャンネルのうちの1つを例示する図である。図3 Cは本開示の例示的な実施形態による、5つの異なる組織の種類画像チャンネルのうちの1つを例示する図である。図3 Dは本開示の例示的な実施形態による、5つの異なる組織の種類画像チャンネルのうちの1つを例示する図である。図3 Eは本開示の例示的な実施形態による、5つの異なる組織の種類画像チャンネルのうちの1つを例示する図である。図3 Fは本開示の例示的な実施形態による、5つの異なる組織の種類画像チャンネルのうちの1つを例示する図である。図3 Gは本開示の例示的な実施形態による、5つの異なる組織の種類画像チャンネルのうちの1つを例示する図である。図3 Hは本開示の例示的な実施形態による、5つの異なる組織の種類画像チャンネルのうちの1つを例示する図である。図3 Iは本開示の例示的な実施形態による、5つの異なる組織の種類画像チャンネルのうちの1つを例示する図である。図3 Jは本開示の例示的な実施形態による、5つの異なる組織の種類画像チャンネルのうちの1つを例示する図である。図3 Kは本開示の例示的な実施形態による、5つの異なる組織の種類画像チャンネルのうちの1つを例示する図である。図3 Lは本開示の例示的な実施形態による、5つの異なる組織の種類画像チャンネルのうちの1つを例示する図である。図3 Mは本開示の例示的な実施形態による、5つの異なる組織の種類画像チャンネルのうちの1つを例示する図である。図3 Nは本開示の例示的な実施形態による、5つの異なる組織の種類画像チャンネルのうちの1つを例示する図である。図3 Oは本開示の例示的な実施形態による、5つの異なる組織の種類画像チャンネルのうちの1つを例示する図である。図3 Pは本開示の例示的な実施形態による、5つの異なる組織の種類画像チャンネルのうちの1つを例示する図である。図3 Qは本開示の例示的な実施形態による、5つの異なる組織の種類画像チャンネルのうちの1つを例示する図である。図3 Rは本開示の例示的な実施形態による、5つの異なる組織の種類画像チャンネルのうちの1つを例示する図である。図3 Sは本開示の例示的な実施形態による、5つの異なる組織の種類画像チャンネルのうちの1つを例示する図である。図3 Tは本開示の例示的な実施形態による、5つの異なる組織の種類画像チャンネルのうちの1つを例示する図である。

10

20

【図4】[12]本開示の例示的な実施形態による、多クラスG Pセグメント化の階層的戦略を例示する図である。

30

【図5】[13]図5 Aは本開示の例示的な実施形態による、全スライドセグメント化の結果を例示する図である。図5 Bは本開示の例示的な実施形態による、全スライドセグメント化の結果を例示する図である。図5 Cは本開示の例示的な実施形態による、全スライドセグメント化の結果を例示する図である。図5 Dは本開示の例示的な実施形態による、全スライドセグメント化の結果を例示する図である。

【図6】[14]本開示の例示的な実施形態による、24の全スライド画像の各々の計算された低信頼領域内の従来技術(A₂値)対開示された方法(A₃値)の分類精度を例示するグラフである。

【図7】[15]組織解析システムの実施形態を概略的に示す図である。

40

【図7 a】[16]図7に示す画像領域分類子モジュールの詳細を概略的に示す図である。

【図8】[17]組織解析方法の実施形態の流れ図を概略的に示す図である。

【図8 a】[18]図8に示すステップ860の詳細を概略的に示す図である。

【図8 b】[19]図8に示すステップ880の詳細を概略的に示す図である。

【図8 c】[20]図8に示すステップ850の実施形態を概略的に示す図である。

【図8 d】[21]図8に示すステップ850の実施形態を概略的に示す図である。

【発明を実施するための形態】

【0011】

[22]図に示す実施形態を明らかにする前に、まず、本開示の様々な実施形態を概括的に説明する。

50

[23]本開示は、とりわけ、解析システム、例えば、組織解析システムに関する。システムは、スライド上に用意された生物組織を解析するのに適切であり得る。

【0012】

[24]解析システムは、画像領域識別子モジュール、例えば、解析される組織標本の画像の領域を選択しおよび/または識別する画像領域識別子モジュールを備えることができる。画像領域を選択するステップ/識別するステップは、例えば、空間位置および/または画像コンテンツを含む複数の基準のいずれかの機能として実施することができる。画像領域を画定するステップは、例えば、個々の画像領域のコンテンツおよび/または境界を指定することによって、個々の画像領域を画定する画像領域データを出力するステップを含むことができる。画像領域を選択するステップは、(受け取った)画像データの複数のサブセットを画定する画像領域データを生成するステップを含むことができ、画像領域を画定するステップは、そのような画像領域データを出力するステップを含むことができる。画像領域識別子モジュールは、以下に説明するように、格子点生成モジュールでよい、または格子点生成モジュールを含むことができる。

10

【0013】

[25]画像領域識別子モジュールは、組織標本の画像を表す画像データを受け取ることができる。画像データは、少なくとも二次元の画像、例えば、組織標本の少なくとも二次元の画像を、例えば、100万~10億画素の程度で表すことができる。画像データは、当技術分野で知られているように、複数の画素を含むことができる。画像データは、画像をグレースケール画像、カラー画像(例えば、RGBもしくはCMYK)またはマルチチャンネル画像として表すことができる。マルチチャンネル画像は、例えば、マルチチャンネル画像の明白なチャンネルとして、非可視電磁放射(例えば、UV光)または他のイメージング技法を使用して捕捉された画像情報を含むことができる。

20

【0014】

[26]画像領域識別子モジュールは、画像データを(組織)解析システムの要素でなくてよい画像源から直接または間接的に受け取ることができる。この点において、(組織)解析システムは、(組織)イメージングデバイス、例えば、マルチチャンネル画像、例えば、各チャンネル画像が狭スペクトル帯からの画像捕捉に対応する8または16ビットのグレースケール画像であるいくつかの(例えば、10から16の間などの)チャンネルを有するマルチチャンネル蛍光画像または明視野像または各チャンネルが特定のカラー捕捉に対応する3つのカラーチャンネルを有するRGBカラー画像などの、画像データを取得するために組織標本を画像化する(組織)イメージングデバイスを備えることができる。例えば、画像源は、蛍光画像を生成するイメージングシステムの蛍光顕微鏡、カメラ、光学スキャナ、CCD、もしくは他の光学構成要素、またはRGB画像を生成するイメージングシステムの明視野顕微鏡、カメラ、光学スキャナでよい。イメージングシステムの例は、例えば、分光フィルタホイールまたは全スライドスキャナを有する任意の蛍光または明視野顕微鏡であり得る。

30

【0015】

[27]イメージングデバイスは、非可視電磁放射(例えば、UV光)または他のイメージング技法を利用して画像を捕捉することができる。(組織)イメージングデバイスは、顕微鏡と、顕微鏡によって拡大される画像(組織の)を捕捉するように配列されたカメラとを備えることができる。画像領域識別子モジュールによって受け取られる画像データは、(組織)イメージングデバイスによって捕捉された生の画像データと同一であり得る、および/または(組織)イメージングデバイスによって捕捉された生の画像データから導き出され得る。

40

【0016】

[28]画像領域識別子モジュールは、(受け取った)画像データの複数のサブセットを識別する画像領域データを生成しおよび/または出力することができる。画像データサブセットの任意の個々のサブセットは、画像のそれぞれの領域を表すことができる。画像領域データは、画像データを、例えば、それぞれのサブセットを表すデータ構造にグループ化

50

することによってそれぞれのサブセットを識別することができる。例えば、画像領域データは、複数の（サブセット）データ構造を含むことができ、各（サブセット）データ構造が、単一の（それぞれの）サブセットの画像データを含む。したがって、画像領域識別子モジュールは、少なくとも1つのそのような（サブセット）データ構造を生成することができる。同様に、画像領域データは、それぞれのサブセットにどの画像データ（例えば、画像データの画素）が属するかを画定する境界を指定することによって、それぞれのサブセットを識別することができる。したがって、画像領域識別子モジュールは、そのような境界を指定する画像領域データを生成することができる。例えば、画像領域識別子モジュールは、画像データの複数の画素を格子点として識別するデータおよび幾何形状を表すデータを生成することができ、幾何形状が個々の領域、すなわち、それぞれの格子点に対する画像データのサブセットを画定する。したがって、「格子点（grid point）」および「データ画像のサブセット（data image subset）」という用語の各々は、画像の領域、すなわち、画像内の点／画素およびその点の周りの近傍を指定するものと理解することができる。したがって、「格子点」および「データ画像のサブセット」という用語の各々は、画像の画素の組、例えば、組織標本の領域を表す画素の組を指定することができる。

10

【0017】

[29]領域のいずれかは、空間的に接触する領域、例えば、画像内の点／画素およびその点の周りの空間的に接触する近傍であり得る。したがって、「領域（region）」という用語は、空間的に接触する画像の画素の組、例えば、組織標本の空間的に接触する領域を表す画素の組を指定することができる。

20

【0018】

[30]個々の領域（それぞれの画像データサブセットで表される）は、様々なサイズまたは形状でよい。例えば、領域は、正方形、長方形、六角形、または円形でよい。同様に、領域は単一の画素と同じくらいに小さくてよく、または数十／数百の画素の直径を有することができる。例えば、個々の領域は、100×100画素程度の正方形であり得る。したがって、格子点は、少なくとも1つの寸法において、一定の間隔をおいて配置され得る。例えば、格子点は、正方形または長方形の（二次元の）格子の交差点に配置され得る。同様に、領域は、ハニカム状の配列で配列され得る。したがって、格子点は、一般的なアレイの形態で配列することができ、アレイの交互に並ぶ行の格子点が、行の方向に格子点の間隔の半分だけ、他方の交互に並ぶ行において格子点から行の方向にオフセットされる。画像領域識別子モジュールは、ユーザ定義の領域サイズ、格子点間隔、領域形状／幾何形状、格子点アレイ、格子点／領域配列、領域重複限界など（選択パラメータとして）を使用してそれぞれの画像領域を選択することができる。そのようなユーザ定義のパラメータの基礎となるユーザ相互作用は、解析システムまたは別のシステムによって実施することができる。したがって、ユーザ定義のパラメータは、ネットワークを介してまたはデータ記憶装置から解析システムによって受け取ることができる。

30

【0019】

[31]個々の領域（それぞれの画像データサブセットで表される）は、固有である、すなわち、別の領域と同一でない可能性がある。個々の領域は、重複してもよいし、重複なしでもよい。例えば、個々の領域は、それぞれの個々の領域の面積の30%以下、20%以下、10%以下または0%が他の領域と重複するように配列し／ような形状にすることができる。したがって、画像データのサブセットは、相互排他的でなくてよい。言い換えれば、画像データの複数のサブセットのうちいずれかの1つは、画像データの複数のサブセットのうち別のサブセットに属する画像データを含むことができる。

40

【0020】

[32]解析システムは、画像領域分類子モジュール、例えば、画像領域のうち任意の画像領域を複数の組織の種類のうち1つとして分類する画像領域分類子モジュールを備えることができる。例えば、画像領域分類子モジュールは、画像領域のうち任意の個々の画像領域を複数の組織の種類のうちそれぞれの1つとして個別に分類することができる

50

。画像領域分類子モジュールは、画像領域のうちの各々の個々の画像領域を個別に分類することができる。画像領域分類子モジュールは、複数の組織の種類（画像領域の可能な分類として利用可能）を記憶するメモリを備えることができる。複数の組織の種類は、正常組織と、腫瘍と、壊死と、間質と、リンパ球凝集体とのうちのいずれかを含むことができる。画像領域分類子モジュールは、数千または数万の画像領域、例えば、少なくとも5,000、少なくとも1万または少なくとも2万の画像領域を分類することができる。

【0021】

[33] 画像領域分類子モジュールは、それぞれの画像領域を表す画像データサブセットを使用してそれぞれの画像領域を分類することができる。例えば、画像領域分類子モジュールは、それぞれの画像データサブセットに属する画素に対して画像処理を実施することによって、それぞれの画像領域を分類することができる。画像領域分類子モジュールは、データベース、例えば、組織特性のデータベースに問い合わせを行うためのそれぞれの画像データサブセットを使用して、それぞれの画像領域を分類することができる。例えば、画像領域分類子モジュールは、それぞれの画像領域の特徴をそれぞれの画像データサブセットから導出し、導出した特徴を使用して、データベースに問い合わせを行うことができる。さらに、画像領域分類子モジュールは、データベース、例えば、組織特性のデータベースから取得したデータを使用して、それぞれの画像領域を分類することができる。例えば、画像領域分類子モジュールは、データベースから取得したデータを使用して、機械学習アルゴリズム（個々の画像領域を分類するための）を訓練し、データベースから取得したデータを使用して（それぞれの画像領域の組織の種類を分類するために）、訓練された機械学習アルゴリズムによりそれぞれの画像データサブセットを処理することができる。同様に、画像領域分類子モジュールは、データベースから取得したデータをそれぞれの画像データサブセットに属する画素の画素情報および/またはそれぞれの画像データサブセットに属する画素に対する画像処理の結果と比較することによって、それぞれの画像領域を分類することができる。データベースから取得したデータは、画像、画像特徴、特定の画像情報に帰する分類および/または画像特徴の特定の組に帰する分類を表すことができる。したがって、データベースから取得したデータは、分類情報と画像情報との対および/または分類情報と少なくとも1つの画像特徴を表す情報との対を含むことができる。画像領域分類子モジュールは、以下に説明するように、ソフト分類モジュールでよい、またはソフト分類モジュールを含むことができる。

【0022】

[34] 画像領域分類子モジュールは、信頼スコア、例えば、それぞれの画像領域を分類するステップの信頼を示す信頼スコアを決定しおよび/または出力することができる。したがって、個々の画像領域を任意の分類するステップは、それぞれの信頼スコアを有することができる。任意の信頼スコアは、それぞれの個々の画像領域を分類するステップに関連することができる。信頼スコアは、それぞれの画像領域を分類するステップが正しい、すなわち、信頼スコア「1」である確率を表すことができる。画像領域分類子モジュールは、それぞれの画像データサブセットに属する画素とデータベースから取得した画像情報との類似性の程度を決定することによっておよび/またはそれぞれの画像データサブセットに属する画素に対して実施された画像処理の結果とデータベースから取得した画像特徴情報との類似性の程度を決定することによって、信頼スコアを決定することができる。信頼スコアを出力するステップは、信頼スコアを表すデータを出力するステップを含むことができる。

【0023】

[35] 解析システムは、データベース変更子モジュールを備えることができる。データベース変更子モジュールは、データベースの変更を、例えば、結果としてデータベースを変更する（データベース）動作の実行となる命令を直接または間接的にデータベースに発行することによって、実施することができる。例えば、データベース変更子モジュールは、結果としてデータベースへの/における/からデータの追加/変更/削除となる命令をデータベースに発行することができる。

【 0 0 2 4 】

[36]データベース変更子モジュールは、画像領域のうちのいずれかの、例えば、画像領域分類子モジュールによって分類された画像領域のうちのいずれかのデータベースの変更を実施することができる。例えば、データベース変更子モジュールは、第1の範囲内に含まれる信頼スコアを有する任意の画像領域のデータベースの変更を実施することができる。言い換えれば、データベース変更子モジュールは、画像領域分類子モジュールによる分類するステップが第1の範囲内に含まれる信頼スコアを有する任意の画像領域のデータベースの変更を実施することができる。第1の範囲は、分類が正しいという確実性を表す信頼スコアを含む信頼スコアの範囲であり得る。したがって、データベースは、分類するステップが正しいことが十分に確からしい、すなわち、第1の範囲内に含まれる信頼スコアを有する、画像領域を分類するステップに回答して、データベースの変更を実施することができる。画像領域分類子モジュールは、数百または数千の画像領域の（第1の範囲内に含まれる信頼スコアを有する）、例えば、少なくとも500、少なくとも1,000、少なくとも5,000または少なくとも10,000の画像領域（第1の範囲内に含まれる信頼スコアを有する）のデータベースの変更を実施することができる。

10

【 0 0 2 5 】

[37]データベース変更子モジュールは、それぞれの画像領域に分類された組織の種類を使用して変更を実施することができる。同様に、データベース変更子モジュールは、それぞれの画像領域を分類するステップに関連する信頼スコアを使用して変更を実施することができる。さらに、データベース変更子モジュールは、それぞれの画像データサブセットを使用して、例えば、それぞれの画像データサブセットに属する画素、他の画像領域に対するそれぞれの画像領域の位置に関する情報、それぞれの画像データサブセットに属する画素に対して実施された画像処理の結果、および/またはそれぞれの画像データサブセットから取得した（他の）組織特性データを使用して、変更を実施することができる。したがって、データベース変更子モジュールは、結果として得られる変更されたデータベースがそれぞれの画像領域に分類された組織の種類およびそれぞれの画像データサブセットから取得した組織特性データを表すデータを含むように変更を実施することができる。

20

【 0 0 2 6 】

[38]解析システムは、画像領域再分類子モジュール、例えば、画像領域のうちの任意の画像領域を複数の組織の種類の中の1つとして再分類する画像領域再分類子モジュールを備えることができる。例えば、画像領域再分類子モジュールは、画像領域のうちの任意の個々の画像領域を複数の組織の種類の中のそれぞれの1つとして個別に再分類することができる。画像領域再分類子モジュールは、複数の組織の種類（画像領域の可能な再分類として利用可能）を記憶するメモリを備えることができる。上記のように、複数の組織の種類は、正常組織と、腫瘍と、壊死と、間質と、リンパ球凝集体とのうちのいずれかを含むことができる。画像領域再分類子は、画像領域のうちのいずれかに対して、それぞれの画像領域を再分類するステップによって決定された組織の種類を出力することができる。組織の種類を出力するステップは、組織の種類を表すデータを出力するステップおよび/または組織の種類および/または組織の種類を表すデータ、例えば、画像データなど、それぞれの画像領域に関する他のデータと連動して、再分類が正しいという確実性および/または（組織の）特徴を表す信頼スコアを含むために、変更されたデータベースの変更をさらに実施する命令を出力するステップを含むことができる。

30

40

【 0 0 2 7 】

[39]画像領域再分類子モジュールは、第2の範囲内に含まれる信頼スコアを有する任意の画像領域を再分類することができる。例えば、画像領域再分類子モジュールは、第2の範囲内に含まれる信頼スコアを有する各画像領域を個別に再分類することができる。第2の範囲は、分類が誤っている、すなわち、信頼スコアが「0」以上であるという確実性を表す信頼スコアを含む信頼スコアの範囲であり得る。したがって、画像領域再分類子モジュールは、分類するステップが誤っていることが十分に確からしい、すなわち、第2の範囲内に含まれる信頼スコアを有する画像領域を再分類するステップに回答して、画像領域

50

を再分類することができる。画像領域再分類子モジュールは、以下に説明するように、適応分類モジュールでよい、または適応分類モジュールを含むことができる。

【 0 0 2 8 】

[40] 画像領域再分類子モジュールは、それぞれの画像領域を表す画像データサブセットを使用して、それぞれの画像領域を再分類することができる。例えば、画像領域再分類子モジュールは、それぞれの画像データサブセットに属する画像に対して画像処理を実施することによって、それぞれの画像領域を再分類することができる。画像領域再分類子モジュールは、変更されたデータベース（組織特性の）に問い合わせを行うためのそれぞれの画像データサブセットを使用して、それぞれの画像領域を再分類することができる。例えば、画像領域再分類子モジュールは、それぞれの画像領域の特徴をそれぞれの画像データサブセットから導出し、導出した特徴を使用して、変更されたデータベースに問い合わせを行うことができる。さらに、画像領域再分類子モジュールは、変更されたデータベース、例えば、上記のように変更された組織特性のデータベースから取得したデータを使用して、それぞれの画像領域を再分類することができる。例えば、画像領域再分類子モジュールは、変更されたデータベースから取得したデータを使用して、機械学習アルゴリズム（個々の画像領域を再分類するための）を（再）訓練することができ、変更されたデータベースから取得したデータを使用して（それぞれの画像領域の組織の種類を再分類するために）（再）訓練された機械学習アルゴリズムによりそれぞれの画像データサブセットを処理することができる。同様に、画像領域再分類子モジュールは、変更されたデータベースから取得したデータをそれぞれの画像データサブセットに属する画素の画素情報および/またはそれぞれの画像データサブセットに属する画素に対する画像処理の結果と比較することによって、それぞれの画像領域を再分類することができる。変更されたデータベースから取得したデータは、画像、画像特徴、特定の画像情報に帰する分類および/または画像特徴の特定の組に帰する分類を表すことができる。したがって、変更されたデータベースから取得したデータは、分類情報と画像情報との対および/または分類情報と少なくとも1つの画像特徴を表す情報との対を含むことができる。

【 0 0 2 9 】

[41] 解析システムは、データベースを記憶するデータ記憶システムを備えることができる。データベースは、複数の組織画像領域の各々に対して、組織の少なくとも二次元の画像を表すデータ、少なくとも1つの組織特徴を表すデータ、組織の種類を表すデータ、および信頼スコアを表すデータのうちのいずれかを含むことができる。任意のそれぞれの画像領域のために記憶された少なくとも1つの組織特徴を表すデータは、それぞれの画像領域のために記憶された組織画像から導出されたデータであり得る。同様に、任意のそれぞれの画像領域のために記憶されたデータで表される信頼スコアは、それぞれの画像領域のために記憶されたデータで表される組織の種類が決定された分類するステップの信頼スコアであり得る。さらに、任意のそれぞれの画像領域のために記憶されたデータで表される組織画像は、分類するステップがそれぞれの画像領域のために記憶されたデータで表される組織の種類を生じさせた、それぞれの組織画像領域を分類するステップに使用される組織画像であり得る。

【 0 0 3 0 】

[42] 解析システムは、サポートベクタマシン、例えば、以下に説明するサポートベクタマシンを備えることができる。サポートベクタマシンは、画像領域（再）分類子モジュールの要素であり得る。解析システム/画像領域（再）分類子モジュールは、サポートベクタマシンを使用して、信頼スコア（それぞれの画像領域を（再）分類するステップの）を決定することができる。言い換えれば、信頼スコアを決定するステップは、1つまたは複数のサポートベクタマシンの動作を実行するステップを含むことができる。

【 0 0 3 1 】

[43] 任意のそれぞれの画像領域を（再）分類するステップは、例えば、以下に説明するように、特徴抽出モジュールを用いて、少なくとも1つの特徴をそれぞれの画像領域から抽出するステップを含むことができる。抽出するステップは、それぞれの画像データサブ

10

20

30

40

50

セットを使用して、例えば、それぞれの画像領域の画素情報を使用して、実施することができる。同様に、抽出するステップは、データベース（またはその変更されたバージョン）から取得したデータを使用して、例えば、上記のように、他の画像領域に関するデータベースに記憶されたデータを使用して、実施することができる。抽出するステップは、それぞれの画像領域の画素情報および/またはそのような画素情報から導出されたデータをデータベースから取得したデータと、例えば、データベースから取得したデータのそれぞれに対応する種類と比較することによって実施することができる。抽出するステップは、テクスチャ特徴と、生物学的特徴と、強度特徴と、傾斜特徴と、ガボール特徴と、共起特徴と、核特徴とからなる群に属する特徴を抽出することができる。

【 0 0 3 2 】

[44] 任意のそれぞれの画像領域を再分類するステップは、それぞれの画像データサブセットのデータおよび/または変更されたデータベースから取得したデータを重み付けするステップを含むことができる。重み付けするステップは、空間的近接値、信頼スコアおよび特徴類似値のうちの少なくとも1つを使用して実施することができる。例えば、重み付けするステップは、それぞれの分類が関係しそれぞれの画像領域が再分類されるデータベース内の画像領域の空間的近接（標本上の/画像内の）の関数としてデータベースから取得した分類を重み付けするステップを含むことができる。同様に、重み付けするステップは、それぞれの画像特徴が関係する画像領域の組織の種類分類に対してデータベースに記憶された信頼スコアの関数としてデータベースから取得した画像特徴を重み付けするステップを含むことができる。さらに、重み付けするステップは、データベースから取得した画像特徴の組を、再分類されるそれぞれの画像領域内の画像特徴の組とのそれらのそれぞれの類似性の関数として重み付けするステップを含むことができる。画像特徴の1つの組と画像特徴の別の組との類似性を示す特徴類似値は、一方の組内の個々の特徴の空間的關係と他方の組内の個々の特徴の空間的關係との類似性の関数としておよび/または一方の組内のある種類の個々の特徴の数と他方の組内のある種類の個々の特徴の数との類似性の関数として決定され得る。

【 0 0 3 3 】

[45] 解析システムは、画像チャンネル抽出器、例えば、以下に説明するように、画像チャンネル抽出モジュールを備えることができる。画像チャンネル抽出器は、画像領域（再）分類子モジュールの要素であり得る。

【 0 0 3 4 】

[46] それぞれの画像領域を分類するステップは、例えば、画像チャンネル抽出器を使用して、画像の少なくともそれぞれの領域を、1つまたは複数の構成要素チャンネルに、例えば、ヘマトキシリンチャンネルと、エオシンチャンネルと、輝度チャンネルとからなる群に属する1つまたは複数の構成要素チャンネルに分離するステップを含むことができる。同様に、分離するステップは、任意の画像領域、例えば、画像全体を1つまたは複数の（前述の）構成要素チャンネルに分離するステップを含むことができる。分離するステップは、前述の抽出するステップ（特徴を）の前に実施され得る。抽出するステップ（少なくとも1つの特徴をそれぞれの画像領域から）は、それぞれの画像領域の構成要素チャンネルのいずれかを使用して実施され得る。

【 0 0 3 5 】

[47] 本開示は、とりわけ、解析方法に、例えば、組織解析方法に関する。方法は、スライド上に用意された生物組織を解析するのに適切であり得る。したがって、解析システムの前述の説明は、変更すべきところは変更して、上記の技法を採用した解析方法に適用する。

【 0 0 3 6 】

[48] 本開示の様々な実施形態を上記に概括的に説明してきたので、次に、図に示す実施形態を明らかにする。

[49] 図1Aは、本開示の例示的な実施形態による、適応分類のためのシステム100を示す。システム100は、電子処理デバイス101に結合されたプロセッサ105によ

10

20

30

40

50

て実行される複数の処理モジュールまたは論理命令を記憶するメモリ110を備える。本明細書では、「モジュール(module)」とは、それぞれの機能を提供するソフトウェアもしくはハードウェアモジュールまたはソフトウェアとハードウェアとの組合せを包含する。プロセッサ105とメモリ110に加えて、電子処理デバイス101は、キーボード、マウス、タッチペン、およびディスプレイ/タッチスクリーンなどのユーザ入力および出力デバイスも含む。以下の議論において説明するように、プロセッサ105は、メモリ110に記憶された論理命令を実行し、画像解析、および結果として電子処理デバイス101を動作させるユーザに対してまたはネットワークを介して結果の出力となる他の定量的動作を実施する。

【0037】

[50]例えば、イメージングシステム102は、画像データを1つまたは複数のスキャンされたスライドからメモリ110に提供することができる。画像データは、画像、ならびに画像が生成されたイメージングプラットフォームに関連した任意の情報を含むことができる。例えば、組織切片は、明視野画像の発色ステインまたは蛍光画像の蛍光体に関連付けられた1つまたは複数の異なる生体指標を含む染色アッセイの適用により染色される必要があり得る。染色アッセイは、明視野画像に発色ステイン、蛍光画像に有機蛍光体、量子ドット、または量子ドットと一緒に有機蛍光体、またはステイン、生体指標、およびビューイングもしくはイメージングデバイスの任意の他の組合せを使用することができる。さらに、典型的な切片は、染色アッセイを切片に適用し、結果として染色された標本となる自動染色/アッセイプラットフォームで処理される。市場には染色/アッセイプラットフォームとして使用するのに適切な様々な商品があり、一例は、譲受人Ventana Medical Systems, Inc.のSYMPHONY(商標)製品である。染色された組織はイメージングシステム、例えば、顕微鏡上または顕微鏡を有する全スライドスキャナおよび/またはイメージング構成要素に供給することができ、一例は譲受人Ventana Medical Systems, Inc.のISCAN COREO(商標)製品である。多重組織スライドは、同等の多重化スライドスキャナシステムでスキャンされ得る。イメージングシステム102によって提供される追加の情報は、染色に使用される化学薬品の濃度、染色中の組織に加えらるる化学薬品の反応時間、および/または組織年齢、固定法、持続時間、切片の埋め込まれ方、切開などの組織の事前解析的条件を含む染色プラットフォームに関連した任意の情報を含むことができる。

【0038】

[51]さらに、本明細書に説明する実施形態は、RGB画像を作り出す明視野全スライド(W S)スキャナで画像化された肝臓における結腸直腸がん転移がんからのヘマトキシリンおよびエオシン(H & E)染色切片に言及するが、本開示は、任意の生物試料または組織の画像の任意の種類に適用可能である。スライドなどの、またはスライドでない画像は、基板上に位置する生物試料の全部または一部から生成され得る。本開示は、RGB、明視野、暗視野、および蛍光画像を含む任意の画像の種類にさらに適用可能である。

【0039】

[52]画像チャンネル抽出モジュール111は、特徴抽出および分類を容易にするために、入力画像を異なる画像チャンネルに分離することによって、実行され得る。例えば、ヘマトキシリンの局所量、エオシンの局所量、および輝度を表す別々のチャンネルを画像チャンネル生成モジュール111によって生成することができる。例えば、Ruifrok, A. and Johnston, D., 「Quantification of histochemical staining by color de-convolution (カラーデコンボリューションによる組織化学的染色の定量化)」Analyt. Quant. Cytol. Histol. 23, 291-299 (2001)に説明されている方法などのカラーデコンボリューションまたは非混合法は、元のRGB画像をヘマトキシリン(HTX)およびエオシンチャンネルに分解するのに適用される。さらに、画像の輝度チャンネル(Lab色空間のL構成要素)も識別され得る。これらのチャンネルは組織画像内の異なる組織構造を強調し、したがって、それらは構造画像チャンネルとして表され得る。

10

20

30

40

50

より正確には、H T Xチャンネルは核領域を強調し（図2 Aの灰色領域参照）、エオシンチャンネルは好酸球構造を強調し（図2 Aの暗い領域）、一方、輝度チャンネルは、脂肪性構造、内腔および空間を強調する（図2 Aの明るい領域）。したがって、これらのチャンネルから抽出された特徴は、組織構造を説明する上で有用である。構造画像チャンネルの選択は、各セグメント化問題に対して調整することができる。例えば、I H C染色画像の場合、構造画像チャンネルは、さらに図3 A ~ Tに示すように、対比染色チャンネル、1つまたは複数の免疫組織化学染色チャンネル、色相、および輝度を含むことができる。

【0040】

[53] 格子点生成モジュール112は、画像内のシードポイントの均一な格子をサンプリングし、各シードポイントの間隔または近傍を指定することによって、WS画像を複数のパッチに分割するために実行することができる。例えば、 $d = 80$ 画素の間隔を有する格子点(GP)をWS画像上に重ね合わせることができ、特徴抽出モジュール113が特徴をこれらのGPの近傍から抽出することが可能になり、分類モジュール114および115が特徴を、したがって、GPを異なる組織の種類に分類することが可能になる。間隔のサイズは、80画素に限定されず、変えることができる。さらに、格子は、正方形、長方形、六角形など、任意の形状でよい。

【0041】

[54] 特徴抽出モジュール113は、画像チャンネルのうちの1つまたは複数に対して特徴抽出を実施する。各画像チャンネルに関連付けられた各GPに対して、特徴抽出モジュール113は、これらのポイントの近傍における画像特徴を抽出し、テクスチャ特徴および生物学的特徴を含む異なる種類の画像特徴が抽出される。例えば、隣接サイズ s および画像チャンネル c を所与として、 s, c が特徴が抽出されるチャンネル c におけるサイズ $s \times s$ の近傍を示すとす。すべての s, c, S, C に対して計算された特徴（ここで S, C がそれぞれ選択された隣接サイズおよび選択されたチャンネルの組を示す）がGPを表す豊富な情報を含む特徴ベクトルを生成するために連結される。1つの実験的な実施形態において、例えば、 $S = [100; 200]$ 画素および $C = \{HTX, \text{エオシン}, \text{輝度}\}$ 。さらに、テクスチャ特徴がすべての画像チャンネルに対して計算されるが、生物学的特徴は生物学的構造が存在する画像チャンネルに対して計算される。例えば、細胞核の特徴は、核領域が顕著であるヘマトキシリンチャンネルから抽出される。特徴選択方法は、分類用に良好な特徴のサブセットを選択するために訓練特徴のプールに適用される。例えば、核の豊富な部分における構造、例えば、腫瘍およびリンパ球凝集体(LA)は、ほとんどの兆候をH T Xチャンネルに有するが、これに対して、正常肝臓、壊死、および間質はほとんどの兆候をエオシンチャンネルに有する。これらの構造に関するさらなる詳細については図3 A ~ Tを参照されたい。この相違を捕捉するために、10瓶のヒストグラムを含む強度に基づいた特徴を計算することができ、各 s, c における画素強度の平均および分散とともに特徴として使用することができる。他の適用例の場合、10瓶ヒストグラムに加えて、または10瓶ヒストグラムの代わりに、平均および分散、多かれ少なかれ瓶、平均、標準偏差、尖度、異なる百分率などを有するヒストグラムのような他の記述統計値を計算することができる。瓶のサイズおよび瓶の種類は異なり得る。本明細書に開示される1つの実験的な実施形態において、特徴の総数は $12 \times 2 \times 3 = 72$ である。エオシン（好酸球組織とも呼ばれる）で強力に染色される組織の中で、正常肝臓は、通常大きな均質の細胞群を含み、エオシンおよび輝度チャンネルにおいて同様に向いた縁部を有し、壊死には不規則に向いた縁部、間質には稜線状構造、および図3 A ~ Tにさらに詳細に示すように他の変形を有する、強力な輝度の変動ならびに無秩序な構造を有する。これらのテクスチャの相違を活用して、特徴抽出モジュール113は、3つの画像チャンネルの各々に対して、傾斜、ガボール、共起、および核特徴を抽出することができる。

【0042】

[55] 様々な種類の特徴抽出が本明細書に列挙されている。傾斜抽出では、特徴抽出モジュール113が、まず、画像の傾斜の大きさおよび傾斜の向きを計算することができる。傾斜特徴は、傾斜の大きさの10瓶ヒストグラムと、傾斜ベクトルの向きの10瓶ヒスト

10

20

30

40

50

グラムとを含む。これらの特徴は、均質領域と不均質領域とを区別し、同様に向いた縁部を有する領域と不規則に向いた縁部を有する領域とを区別する。この場合も、ヒストグラムに加えて、平均、標準偏差、尖度、百分率などの異なる記述統計を傾斜の大きさおよび向きの特徴として使用することができる。実験的な例において、特徴の総数は $20 \times 2 \times 3 = 120$ である。ガボール特徴では、特徴抽出モジュール 113 が 3 つの異なる波長および 6 つの異なる向きを使用して 18 のガボールフィルタを生成することができる [Jain, A. K., Farrokhnia, F.: Unsupervised texture segmentation using Gabor filters (ガボールフィルタを使用する教師なしテクスチャセグメント化). In: IEEE Int. Conf. Sys., Man., Cyber., pp. 14 - 19 (1990) 参照]。フィルタ応答の平均および分散が特徴として使用される。特徴として使用される波長の数、向き、および記述統計を各適用例に選択することができる。実験的な例において、特徴の総数は $36 \times 2 \times 3 = 216$ である。共起特徴では、特徴抽出モジュール 113 が画素強度の共起行列 (CM) を計算することができ、この CM からエネルギー、相関、慣性、エントロピー、逆差分モーメント、総和平均、総和分散、総和エントロピー、差分平均、差分分散、差分エントロピー、および 2 つの情報の相関の程度を含む、13 のハラリック特徴を計算することができる [Haralick, R.ら.: Textural Features for Image Classification (画像分類のためのテクスチャ特徴). IEEE Trans. Sys., Man., Cyber. 3 (6), 610 - 621 (1973) 参照]。各チャンネルに対して個別に計算され得る従来のグレーレベルの CM (GLCM) に加えて、チャンネル間またはカラー共起行列 (CCM) をさらに使用することができる。CCM は 2 つの異なる画像チャンネルにおける画素強度の共起から作り出されて、すなわち、変異ベクトル $d = [dx; dy]$ を使用して 2 つのチャンネル $C_i; C_j$ から CCM を計算し、画素強度の共起が C_i における位置 $(x; y)$ で計算され、画素強度が C_j における位置 $(x + dx; y + dy)$ で計算される。CCM の利点は、異なる組織構造 (異なるチャンネルにおいて強調される) 間の空間関係を、それらを明瞭にセグメント化する必要なく、捕捉することである。さらに、ハラリック特徴は、すべての 3 つのチャンネルの GLCM から計算することができ、ハラリック特徴は、すべてのチャンネルの対 (HTX - エオシン、HTX - 輝度およびエオシン - 輝度) の CCM から計算することができる。実験的な実施形態において、特徴の総数は $13 \times 2 \times (3 + 3) = 156$ であり得る。さらに、核特徴を、濃度、形状、サイズ、および細胞核の外観を使用して抽出して、強力な特徴を提供し、それによって、例えば、Masahiro, I.ら.: Automatic segmentation of hepatocellular structure from HE-stained liver tissue (HE 染色肝臓組織からの幹細胞の自動セグメント化). In: Proc. SPIE, pp. 867611 - 867611 - 7 (2013)] に説明される方法を使用して、組織の種類を区別することができる。HTX チャンネルから計算されたテクスチャ特徴は、ある量の核情報を捕捉するが、明確な核関連特徴をさらに計算することができる。例えば、システムは、まず、放射対称ベースの方法 [Parvin, B.ら.: Iterative voting for inference of structural saliency and characterization of subcellular events (構造的性質の推論および細胞内事象の特性化のための反復投票). IEEE Trans. Image Processing 16 (3), 615 - 623 (2007)] を使用して HTX チャンネル (核が最も顕著である) から中心核を検出することができ、続いて、Otsu の方法による核領域のセグメント化 [Otsu, N.: A threshold selection method from gray-level histograms (グレーレベルヒストグラムからの閾値選択方法). IEEE Trans. Sys., Man., Cyber. 9 (1), 62 - 66 (1979)] がある。核領域における画素強度が変動するので、Otsu の方法は各検出された中心核の局所的近傍に適用され得る。セグメント化の結果に基づいて、システムは、(i) 核濃度

10

20

30

40

50

(検出された核の数)、(i_i)核サイズ(核面積の平均)、および(i_{ii})核領域における平均強度値を計算することができる。要するに、合計 $72 + 120 + 216 + 156 + 3 = 567$ の特徴を、各 GP に対する特徴ベクトルを形成するために作り出すことができる。これらの核関連の特徴は、分類のために特徴を作り出すために検出される組織内の核、細胞、腺などの生体の発生、濃度、および特性を捕捉する生物学的特徴の一例である。

【 0043 】

[56] 特徴抽出に続いて、2段分類が組織の外観の変動を効率的におよび確実に処理するために実施される。まず、ソフト分類モジュール 114 が SVM などの「ソフト」分類を使用し、各パッチの信頼スコアおよびラベルを生成して各パッチを分類するために実行され得る。このソフト分類は、既知の特徴を含む外部(事前構築された)訓練データベースを使用し、各 GP のラベルおよび信頼スコアを生成して、WS 画像 W_i 内のすべての GP を分類するステップを含む。例えば、腫瘍領域などの特定の領域の種類の SVM の出力ラベルは、0 から 1 までのスカラー値であり得、その場合、0 は領域が腫瘍である可能性が何もないことを示し、1 は GP が腫瘍領域に属する高い可能性があることを示す。信頼マップが各 GP の信頼決定を使用して、画像内のパッチに対して生成され得る。 W_i からの最高信頼の GP を、外部データベースと組み合わせて、 W_i の適応訓練 DB を生成する内部訓練データベースに追加することができる。例えば、 > 0.8 の信頼スコアは、高信頼 GP とみなすことができ、データベースに追加することができる。訓練データベース 116 は、組み合わせられたデータベースを含むことができる。他の実施形態において、ソフト分類の外部訓練データベースは、訓練データベース 116 内に組み込まれ得る。データベース 116 は、各画像に対するパッチの信頼およびラベルを記憶することもできる。

【 0044 】

[57] 適応分類モジュール 115 は、低信頼パッチを訓練データベース 116 内の高信頼パッチと比較するステップと、データベース内のパッチの空間コヒーレンスを増大させるために同様のパッチを使用するステップとを含む第 2 の分類ステップを実施するために実行される。低信頼パッチの組織特徴に基づいて、低信頼パッチの分類結果を改良するために、類似検索動作が画像内で実施される。言い換えれば、各低信頼パッチに対して、隣接した高信頼パッチが各パッチのラベルを改良する方により大きな寄与をし、それによって、低信頼パッチ内のセグメント化精度が改善される。例えば、上位 10 の同様のパッチを取得することができ、それらからの大半のラベルを低信頼パッチまたは画素の新たなラベルとして使用することができる。したがって、データベース 116 に記憶された適応データベースにより、 W_i 内の低信頼パッチを再分類するステップが可能となる。低信頼パッチの周りの空間的拘束により、より多くの重みを高信頼パッチに与えるステップが可能となり、低信頼パッチからさらに離れている同様のパッチに低い重みを与えることが可能になる。

【 0045 】

[58] 各画像内の高解像度および大画素数により、結果として得られる画像ごとのデータベースは、極めて包括的であり得る。異なる画像間の大きな変動により、開示されるシステムおよび方法が、各画像内のパターンに基づくセグメント化の結果を適応的に改善することが可能になる。本開示の例示的な実施形態において、画像データに関する生体情報、例えば、本明細書に開示される方法により収集したまたは取得したデータは、具体的な画像のデータベース 116 を訓練するために具体的な特徴を設計するのに利用される。類似検索は、同じ画像内の特徴にうまく働き、低信頼領域内のセグメント化精度の改善が可能になる。さらに、改良された信頼マップを、各 GP の 2 番目の段階の信頼決定を使用して、画像内のパッチに対して生成することができる。信頼マップおよび組織の種類のマップは、ユーザ操作端末 101 に出力する、またはネットワークを介して遠隔端末に伝送することができる。信頼マップおよび組織の種類のマップは、電子的に解析して、画像の品質を決定し、または患者の治療または予後のための診断を取得することができる。

【 0046 】

[59]本明細書に説明するように、様々な異なる分類方法を検出された特徴に適用することができる。例示的な実施形態において、これらの異なる方法を評価することができ、ランダムフォレスト分類方法を優れた性能により選択することができる。本明細書に開示される実験的な実施形態において、性能は、肝臓、CRC転移がん、リンパ球、壊死および間質（グラウンドトリスは病理学者によって提供された）の5つの異なる組織の種類の84,000を超えるシードを含むデータベースにより評価された。これらの5つの組織の種類は、組織の種類のカテゴリの例であり、開示されるシステムおよび方法は、これらの5つの組織の種類に限定されない。組織の種類は、異なる種類の組織には、例えば、組織画像が肝臓組織画像でないとき、異なることがある。取得したシード分類精度は89%であった。さらに、画像セグメント化の結果が27の全スライド組織画像に対しても取得された。実験結果は機械支援診断システムの有用性を示している。実験的な実施形態において、セグメント化は、Ozseven, T.: Quantifying the necrotic areas on liver tissues using support vector machine (SVM) algorithm and Gabor filters (サポートベクタマシン (SVM) アルゴリズムおよびガボールフィルタを用いた肝臓組織上の壊死部分の定量化)の研究と同様の従来の教師ありの枠組みを使用して実施することができる。

10

【0047】

[60]図2A~2Cは、本開示の例示的な実施形態による、適応分類のための方法を示す。大きなサイズのWS画像（すなわち、スライドごとに生成される大量のGP）を活用または利用するために、2段階分類手順が、事前構築された訓練用の第1のデータベース(DB) 217が画像内のすべてのGP（ステップ(1)、(2)、(3)）を分類するのに使用される第1の段階を含む。次に、高分類信頼を有するGPが新たな第2の訓練DB(219)とみなされ、それは適応訓練用の変更されたDB * 218（ステップ(4)）を作り出すために217と組み合わせられる。訓練データが分類されなければならないデータと同じWS画像に属するとき分類精度がより高いという仮定に基づいて、* 218は適切なデータを提供して、を使用するとき低信頼で分類されたGPを再分類する（ステップ(5)）。* 218が各WS画像に対して適応的に構築されるので、方法は適応分類と称される。実装形態により、変更されたDB * 218は、事前構築された訓練用の第1のデータベースが次第に改善されるように分析される必要がある後続の画像（例えば、同じ患者から取得した組織標本の同じスライドまたは別のスライドから取られた画像）として、事前構築された訓練用の第1のデータベース(DB) 217と置き換えることができる。この方法では、 $C = \{ \text{正常肝臓、腫瘍、壊死、LA、間質} \}$ は、対象のすべての組織の種類のリストと定義することができる。アルゴリズム1（図2B）におけるテスト標本の信頼スコアは、SVMにおける決定境界までの距離、ランダムフォレストによって生成された投票スコア、またはk最近傍分類子における最近傍のラベルのパーセンテージを使用して取得することができる。アルゴリズム1は、アルゴリズム2（図2Cに示す）を参照する。図2Bおよび2Cの実施形態において、「信頼閾値 τ_c

20

30

[61]」は、0から1の間の信頼範囲をそれぞれ τ_c から1の間および0から τ_c の間の第1および第2の範囲に分割することに留意されたい。「テストデータ」は分類されるべき画像データサブセットである。

40

【0048】

[62]図3は、本開示の例示的な実施形態による、5つの組織の種類を例示する。図3A、3B、3C、3D、および3Eは、それぞれ正常肝臓、壊死、間質（腫瘍周辺間質）、腫瘍、およびリンパ球凝集体(LA)切片からの結腸直腸がん転移がんのH&E染色のスキンを示す。図3F~3Jは、これらの組織の種類の高々に対応するHTX構造画像チャンネルを示し、図3K~Oはエオシン構造画像チャンネルを示し、3P~3Tは輝度構造画像チャンネルを示す。これらのチャンネルの高々は、組織画像内の異なる組織構造を強調し、したがって、それらは構造画像チャンネルと称される。腫瘍組織は腫瘍内間質（図3Dにおける。図3Nにおいて顕著である）を含むことがあるが、しかし、組織は依然として固形

50

腫瘍とみなされ得る。

【 0 0 4 9 】

[63]本明細書に述べるように、核の豊富な部分における構造（例えば、腫瘍およびLA）はHTXチャンネル（図3F～3J）内にほとんどの兆候を有することがあるが、正常肝臓、壊死、および間質はエオシンチャンネル（図3K～3O）内にほとんどの兆候を有する。この相違を捕捉するために、10瓶ヒストグラムを含む、強度に基づいた特徴を計算し、各s、cにおける画素強度の平均および分散とともに特徴として使用することができる。他の適用例では、10瓶ヒストグラム、平均、および分散に加えてまたはそれらの代わりに、多かれ少なかれ瓶、平均、標準偏差、尖度、異なる百分率などを有するヒストグラムのような他の記述統計値を計算することができる。瓶のサイズおよび瓶の種類は異なり得る。本明細書に開示される1つの実験的な実施形態において、特徴の総数は $12 \times 2 \times 3 = 72$ である。好酸球組織の中で、正常肝臓は、通常、大きな均質の細胞群を含み、エオシンおよび輝度チャンネル内に同様に向いた縁部を有する（図3Kおよび3P）。対照的に、壊死の場合、これらのチャンネルは、不規則に向いた縁部を有する、強力な輝度の変動および無秩序な構造を含む（図3Lおよび3Q）。最後に、間質では、これらのチャンネルは、より多くの稜線状構造（図3Mおよび3R）を含む。好塩基性組織の場合、腫瘍は、典型的には、より大きな核を含み、画素強度がLAよりもHTXチャンネル内の核領域においてより低い（図3Iおよび3J）。

10

【 0 0 5 0 】

[64]実験的な実施形態において、提案した方法を評価するのに使用されるデータセットは、 $26,600 \times 22,800$ 画素の平均サイズを用いてVentana iScan HT全スライドスキャナ（ $0.465 \mu\text{m}$ /画素）上で20倍の倍率でデジタル化された結腸直腸がんからの転移がんによる27のスライドの肝臓標本を含んでいた。27の画像の各々において、専門観察者によっていくつかのGPが選択され、5つの組織の種類に割り当てられ、結果として合計84,000を超えるラベル付きGPのGPデータセットとなっている。実験の第1の部分において、従来の訓練および分類手順が適応分類手順なしでGPデータセットに対して実施された。目的は抽出された特徴の弁別力を検証することである。GPデータセットは、3つの群に分割され、2つが訓練用であり、1つが検証用である。過剰適合を回避するために、データは、同じ画像からのGPが訓練およびテストデータの両方に同時に存在しないように分割された。プロセスは各々異なる訓練および検証の群に対して3回繰り返され、平均分類精度が報告されている。異なる分類子の性能が比較され、すなわち、k最近傍（kNN）、サポートベクタマシン（SVM）、およびランダムフォレスト（RF）である。さらに、特徴空間の高次元性により、主成分解析（PCA）および最小冗長 - 最大関連（mRMR）[Peng, H.: Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy（相互情報に基づく特徴選択：最大依存、最大関連、および最小冗長の基準）。IEEE Trans. Pattern Analysis and Machine Intelligence 27(8), 1226-1238 (2005)]が全特徴の組と競争で次元縮退について検討されている。多クラス分類問題が2つの戦略、すなわち、1対1および階層的戦略を使用して多値2値分類問題を組み合わせることによって解決される。本開示の例示的な実施形態による、多クラスGPセグメント化の階層的戦略の例示の図4を参照されたい。

20

30

40

【 0 0 5 1 】

[65]表1は、分類子の異なる選択、次元縮退方法、および多クラス分類の戦略の標準偏差によるすべての分類精度（%）をまとめたものである。

【 0 0 5 2 】

【表 1】

分類子	1対1戦略			階層的戦略		
	全特徴	mRMR	PCA	全特徴	mRMR	PCA
SVM	87.7 (2.5)	87.4 (1.8)	87.8 (2.7)	87.3 (4.2)	88.0 (3.0)	88.3 (6.2)
RF	89.8 (3.3)	89.6 (4.4)	85.1 (7.6)	89.9 (3.5)	89.4 (2.8)	81.9 (5.6)
kNN	85.3 (3.7)	89.3 (3.4)	85.5 (3.2)	85.0 (3.1)	89.0 (4.1)	75.9 (6.1)

【 0 0 5 3 】

[66]適応分類方法を図2に示すようにWSセグメント化にさらに適用することができる。GPデータセットは、一個抜き交差検証手順の訓練データとして使用される。WS画像 W_i におけるセグメント化が、 W_i 以外のスライドのラベル付きGPを使用して、アルゴリズム1(図2B)における事前構築された訓練DBとして実施される。プロセスは、DB内のすべての27のWS画像に対して繰り返すことができる。

10

【 0 0 5 4 】

[67]表1のGP分類結果に基づいて、WSセグメント化のmRMRおよび階層的戦略を、低い計算コストで優位性のある分類精度を提供するので、使用することができる。使用される分類子はRF(第1の分類段に対して)およびkNN(改良分類段に対して)であり、それらは表1のものと同様の優位性のある検証の後選択することができるが、適応ステップ後の結果を比較した後である。外部DBが大きく、RFとしてのアンサンブル分類子がより適切である大きな特徴変動を含むが、内部DBがより小さく、kNNとしてのより単純な分類子がより適切であるより低い特徴分散を含むと仮説を立てることができる。

20

$$J I (S_t, G_t) = | S_t \cap G_t | / | S_t \cup G_t |$$

J I [0 , 1] により、J I のより大きな値がより優れたセグメント化の結果に対応する。下記表2において、従来の方(アルゴリズム1における第1の分類段だけを実施する)によって取得した各組織の種類の7つのWS画像の平均J I 値は、提案した適応分類方法と比較される。さらに、 W_i の全セグメント化精度(すべての組織の種類を考慮することによる)が次式として計算される。

30

【 0 0 5 5 】

【数 1】

$$S A = \frac{|\{l(p)|l(p) = g(p)\}|}{|W_i|}$$

【 0 0 5 6 】

ここで、 $l(p)$ および $g(p)$ は画素 p の割り当てられたラベルおよびグラウンドトゥースラベルを表し、 $|\cdot|$ はある組の濃度を表す。従来のおよび提案した方法によって取得した7つのWS画像の平均SA値は、それぞれ、72%および74%である。これらの評価は、従来の方(アルゴリズム1)に比べて提案した方法の改善された性能を示す。

40

【 0 0 5 7 】

【表 2】

方法	肝臓	腫瘍	壊死	LAs	間質
従来の分類	0.50	0.31	0.54	0.44	0.44
適応分類	0.54	0.33	0.58	0.45	0.44

表 2 : 対象の5つの組織の種類の従来の分類方法および適応分類方法の平均 J I 値

【 0 0 5 8 】

[68]実験結果から、GPデータセット(表1)に対して取得したGP分類精度は、GPデータセットではなくWS画像が過渡的組織領域(混乱領域)を含むので、セグメント化

50

精度（SA値）よりも高いことが観察された。これらの過渡的領域内のGPの近傍は、1つより多くの組織の種類を含み、それによってそれらを分類するのがより困難となる。SA値はJI値よりも高く、それは各クラスがすべての他のクラスの偽陰性部分に寄与する5つのクラスセグメント化問題に対して予測される。さらに、第2段分類子が実験的にk最近傍として選択された。GPデータセットは、一個抜き交差検証手順における事前構築されたDB F^* （図2B参照）として使用された。WS画像 W_i におけるセグメント化は、 W_i 以外のスライドのラベル付きGPを使用して実施された。

【0059】

[69]図5A～5Dは、本開示の例示的な実施形態による、全スライド（WS）のセグメント化の結果を例示する。図5Aは、サイズ18,500×17,200画素の入力WS画像を示す。図5Bは、異なる影付きの領域がそれぞれ腫瘍531、肝臓532、LA533、壊死534、および間質535領域をそれぞれ示すセグメント化グラウンドトールースを示す。図5Cは、従来の方法を使用したセグメント化結果を示す。図5Dは、提案の適応分類方法を使用したセグメント化結果を示す。図5Cにおいて誤って分類された領域540の一部が図5Dでは訂正して示される。

【0060】

[70]セグメント化グラウンドトールース（専門観察者によって提供される）を使用して、分類精度を、再分類段を適用する前および後に、高信頼領域（ x^h ）および低信頼領域における各組織の種類 t_j （ $j \in [1, 5]$ ）に対して計算することができる（それぞれの分類されたラベルは $l(x_i)$ および $l^*(x_i)$ であり、ここで、 $x_i \in x^h$ ）。 A_1 、 A_2 、および A_3 として表されるこれらの精度は、次式として計算される。すなわち、

【0061】

【数2】

$$A_k = \frac{|\{x_i \in S_k | l_k(x_i) = g(x_i) = t_j\}|}{|S_k|} \quad (k = [1, 3])$$

【0062】

ここで、 $g(x_i)$ はWS画像 $S_1 = x^h$ 、 $S_2 = S_3 = x^l$ 、 $l_1(x_i) = l_2(x_i) = l(x_i)$ 、 $l_3(x_i) = l^*(x_i)$ 内の画素 x_i のグラウンドトールースラベルを表す。5つの組織の種類の各々に対するすべてのWS画像にわたる A_1 、 A_2 、および A_3 の平均値が、表3に示される。

【0063】

【表3】

精度	正常肝臓	LAs	腫瘍	壊死	間質
A_1	0.76	0.55	0.81	0.76	0.74
A_2	0.52	0.33	0.42	0.55	0.58
A_3	0.56	0.34	0.44	0.60	0.66

表3：再分類段の前（ A_2 ）および後の（ A_3 ）高信頼領域（ A_1 ）および低信頼領域における分類精度

【0064】

[71]さらに、平均 A_2 および A_3 の値は、図6に示すように、24のWS画像の各々のすべての組織の種類にわたってプロットされる。この実験的実施形態において、以下の観察が取得されている。すなわち、(i) A_1 値は、一貫して A_2 値よりも高く、RF分類子によって選択された高信頼領域がWS画像における信頼できる領域であり、適応DBに適切であること示している。(ii) A_3 値は A_2 値よりも高く、スライド間組織の変動

の存在下で分類結果を改善する上で2段階適応分類方法の有用性を示している。(i i i) 図6に示すように、2段階適応分類方法は、従来技術の方法の結果を大抵改善している(24の画像のうち23の画像に改善が得られている)。

【0065】

[72]したがって、包括的な枠組みが、一般的には組織画像のセグメント化問題に、具体的には肝臓のH&E染色切片における組織セグメント化に対処するために提供される。異なる種類の特徴が異なる構造の画像チャネルから抽出され(カラーデコンボリューション手順およびLab色空間への変換を使用して取得され)、組織構造を説明するのに使用される。セグメント化を実施するために、適応分類方法は、事前構築された訓練データベースを使用し、事前構築された訓練データベースを改良するために高信頼スコアを有する分類されたGPを使用し、それによって、低信頼GPを再分類するのにより適切である適応訓練データベースを生成して、第1のGP分類を実施するステップを含む。そのような適応訓練データベースは、各々の新たなスライドに対して個別に生成され、入力WS画像の大きなサイズにより、数多くの高信頼GPが第1の分類段からの各スライドに予測され、それによって、訓練用の組の改良がより信頼できるものになる。

10

【0066】

[73]本開示の例示的な実施形態の前述の開示は、例示および説明の目的で提示されてきた。網羅的であることが意図されないし、新規の特徴を正確な形態に限定することが意図されない。本明細書に説明する実施形態の多くの変形および変更は、上記の開示を考慮すると当業者には明らかであろう。本開示の範囲は、本明細書に添付される特許請求の範囲によってのみ、およびそれらの同等物によってのみ、定義されるものとする。

20

【0067】

[74]電子処理デバイスは、典型的には、プロセッサ、オペレーティングシステム、システムメモリ、メモリ記憶デバイス、入出力コントローラ、入出力デバイス、およびディスプレイデバイスなどの周知の構成要素を含む。電子処理デバイスの多くの可能な構成および構成要素があり、キャッシュメモリ、データバックアップユニット、および多くの他のデバイスを含むこともできることを当該技術分野の当業者にはやはり理解されよう。入力デバイスの例には、キーボード、カーソル制御デバイス(例えば、マウス)、マイクロホン、スキャナなどが含まれる。出力デバイスの例には、ディスプレイデバイス(例えば、モニタまたはプロジェクタ)、スピーカ、プリンタ、ネットワークカードなどが含まれる。ディスプレイデバイスは、視覚情報を与えるディスプレイデバイスを含むことができ、この情報は、典型的には、論理的におよび/または物理的に画素のレイとして編成され得る。入力および出力インターフェースを提供するための様々な周知のまたは将来のソフトウェアプログラムのいずれかを備えることができるインターフェースコントローラを含むこともできる。例えば、インターフェースは、1つまたは複数のグラフィカル表現をユーザに提供する、一般に「グラフィカルユーザインターフェース」と称される(しばしばGUIと称される)ものを含むことができる。インターフェースは、典型的には、関連技術分野の当業者に周知の選択または入力の手段を使用してユーザ入力を受けることが有効とされる。インターフェースは、タッチスクリーンデバイスでもよい。同じまたは代替実施形態において、電子処理デバイスの適用例は、「コマンドラインインターフェース」と称される(しばしばCLIと称される)ものを含むインターフェースを採用することができる。CLIは、典型的には、アプリケーションとユーザとの間のテキストベースの相互作用を提供する。典型的には、コマンドラインインターフェースは、ディスプレイデバイスを通じてテキストのラインとして出力を提供し、入力を受け取る。例えば、いくつかの実装形態は、関連技術分野の当業者には周知のUnix Shell、またはMicrosoft .NET frameworkなどのオブジェクト指向の種類のプログラミングアーキテクチャを採用したMicrosoft Windows Powershellなどの「シェル」と称されるものを含むことができる。

30

40

【0068】

[75]インターフェースは1つまたは複数のGUI、CLIまたはそれらの組合せを含む

50

ことができることを関連技術分野の当業者は理解するであろう。

[76]プロセッサは、Intel Corporationによって製作されたCeleron、Core、もしくはPentiumプロセッサ、Sun Microsystemsによって製作されたSPARCプロセッサ、AMD Corporationによって製作されたAthlon、Sempron、Phenom、もしくはOpteronプロセッサなどの市販のプロセッサを含むことができ、またはプロセッサは、入手可能である、もしくは入手可能となる他のプロセッサの1つであり得る。プロセッサのいくつかの実施形態は、マルチコアプロセッサと称されるものを含むことができ、および/または単一コアもしくはマルチコア構成の並列処理技術を採用することが有効にされ得る。例えば、マルチコアアーキテクチャは、典型的には、2つ以上のプロセッサ「実行コア」を備える。本例では、各実行コアは、マルチスレッドの並列実行を可能にする独立プロセッサとして実施することができる。さらに、プロセッサは一般に32ビットまたは64ビットアーキテクチャと称されるもので、または現在周知のもしくは将来開発される他のアーキテクチャ構成で構成することができることを関連技術分野の当業者は理解するであろう。

10

【0069】

[77]プロセッサは、典型的には、オペレーティングシステムを実行するが、オペレーティングシステムは、例えば、Microsoft CorporationのWindowsの種類オペレーティングシステム、Apple Computer Corp.のMac OS Xオペレーティングシステム、多くのベンダーから入手可能なUnixまたはLinux（登録商標）の種類オペレーティングシステムまたはオープンソースと称されるもの、別のもしくは将来のオペレーティングシステム、またはそれらのいくつかの組合せであり得る。オペレーティングシステムは、よく知られたやり方でファームウェアおよびハードウェアとインターフェースし、プロセッサが様々なプログラミング言語で書くことができる様々なプログラムの機能を調整し実行するのを容易にする。オペレーティングシステムは、典型的には、プロセッサと協働して、電子処理デバイスの他の構成要素の機能を調整し実行する。オペレーティングシステムは、スケジューリング、入出力制御、ファイルおよびデータ管理、メモリ管理、ならびに通信制御および関連サービスも、すべて周知の技法により、提供する。

20

【0070】

[78]システムメモリは、所望の情報を記憶するのに使用することができ、電子処理デバイスによってアクセスすることができる、様々な周知のまたは将来のメモリ記憶デバイスのいずれかを含むことができる。デジタル記憶媒体は、デジタルで符号化された命令、データ構造、プログラムモジュール、または他のデータなど、情報の記憶のための任意の方法または技術で実装される揮発性および不揮発性、取り外し可能および非取り外し可能媒体を含むことができる。例には、任意の一般に入手可能なランダムアクセスメモリ（RAM）、読み出し専用メモリ（ROM）、電子的に消去可能プログラム可能読み出し専用メモリ（EEPROM）、デジタル多用途ディスク（DVD）、常駐ハードディスクまたはテープなどの磁気媒体、読み出し書き込みコンパクトディスクなどの光媒体、または他のメモリ記憶デバイスが含まれる。メモリ記憶デバイスは、コンパクトディスクドライブ、テープドライブ、取り外し可能ハードディスクドライブ、USBもしくはフラッシュドライブ、またはディスクドライブを含む、様々な周知のまたは将来のデバイスのいずれかを含むことができる。そのような種類のメモリ記憶デバイスは、典型的には、それぞれコンパクトディスク、磁気テープ、取り外し可能ハードディスク、USBもしくはフラッシュドライブ、またはフロッピーディスクなどのプログラム記憶媒体から読み出しおよび/またはプログラム記憶媒体に書き込む。これらのプログラム記憶媒体または現在使用されているもしくは後で開発され得る他のものうちのいずれかをプログラム製品とみなすことができる。理解されるように、これらのプログラム記憶媒体は、典型的には、ソフトウェアプログラムおよび/またはデータを記憶する。ソフトウェアプログラムは、制御論理とも呼ばれるが、典型的には、メモリ記憶デバイスと連動して使用されるシステムメモリおよび/またはプログラム記憶デバイスに記憶される。いくつかの実施形態におい

30

40

50

て、中に記憶された制御論理（プログラムコードを含むソフトウェアプログラム）を有するデジタル記憶媒体を備えるプログラム製品が説明される。制御論理は、プロセッサによって実行されたとき、本明細書に説明する機能をプロセッサに実施させる。他の実施形態において、いくつかの機能は、例えば、ハードウェア状態マシンを使用して主にハードウェアで実装される。本明細書に説明する機能を実施するためのハードウェア状態マシンの実装形態は、当該技術分野の当業者には明らかであろう。入出力コントローラは、人間であるかマシンであるかにかかわらず、局部であるか遠隔であるかにかかわらず、ユーザからの情報を受けて処理するための様々な周知のデバイスのいずれかを含み得る。そのようなデバイスは、例えば、モデムカード、ワイヤレスカード、ネットワークインターフェースカード、サウンドカード、または様々な周知の入力デバイスのためのいずれかの他の種類のコントローラを含む。出力コントローラは、人間であるかマシンであるかにかかわらず、局部であるか遠隔であるかにかかわらず、情報をユーザに提示するための様々な周知のディスプレイデバイスのいずれかのコントローラを含み得る。現在説明する実施形態において、電子処理デバイスの機能要素は、システムバスを介して互いに通信する。電子処理デバイスのいくつかの実施形態は、ネットワークまたは他の種類の遠隔通信を使用していくつかの機能要素と通信することができる。当該技術分野の当業者には明らかであるように、測定器制御および/またはデータ処理アプリケーションは、ソフトウェアで実装された場合、システムメモリおよび/またはメモリ記憶デバイスに読み込み、システムメモリおよび/またはメモリ記憶デバイスから実行することができる。測定器制御および/またはデータ処理アプリケーションの全部または部分は、読み出し専用メモリまたはメモリ記憶デバイスの同様のデバイスに常駐することもできるが、そのようなデバイスは、測定器制御および/またはデータ処理アプリケーションがまず入出力コントローラを通じて読み込まれることを必要としないものである。測定器制御および/またはデータ処理アプリケーションまたはその部分は、周知のやり方で実行に有利なようにシステムメモリまたはキャッシュメモリまたは両方に、プロセッサによって読み込むことができることを当該技術分野の当業者には理解されよう。また、電子処理デバイスは、システムメモリに記憶された1つまたは複数のライブラリファイル、実験データファイル、およびインターネットクライアントを含むことができる。例えば、実験データは、検出された信号値または1つまたは複数の合成によるシーケンシング（SSB）実験またはプロセスに関連付けられた他の値など、1つまたは複数の実験またはアッセイに関連したデータを含み得る。さらに、インターネットクライアントは、ネットワークを使用して、別の電子処理デバイス上の遠隔サービスにアクセスすることが有効なアプリケーションを含むことができ、例えば、一般に「ウェブブラウザ」と称されるものを備えることができる。本例では、いくつかの一般に採用されるウェブブラウザは、Microsoft Corporationから入手可能なMicrosoft Internet Explorer、Mozilla CorporationからのMozilla Firefox、Apple Computer Corp.からのSafari、Google CorporationからのGoogle Chrome、または当技術分野で現在周知のまたは将来開発される他の種類のウェブブラウザを含む。また、同じまたは他の実施形態において、インターネットクライアントは、生物学的用途のデータ処理アプリケーションなど、ネットワークを介して遠隔情報にアクセスすることが有効な専用ソフトウェアアプリケーションを含むことができる、または専用ソフトウェアアプリケーションの要素であり得る。

【0071】

[79]ネットワークは、当技術分野の当業者にはよく知られた、多くの様々な種類のネットワークのうちの1つまたは複数を含むことができる。例えば、ネットワークは、通信するために、一般にTCP/IPプロトコル群と称されるものを採用することができる、ローカルまたはワイドエリアネットワークを含むことができる。ネットワークは、一般にインターネットと称される、相互接続されたネットワークの全世界システムを備えたネットワークを含むことができ、または様々なイントラネットアーキテクチャも含み得る。ネットワーク環境におけるユーザの中には、ハードウェアおよび/またはソフトウェアシステ

10

20

30

40

50

ムとの情報トラフィックを制御するために、一般に「ファイアウォール」と称される（パケットフィルタまたは境界保護デバイスと称されることもある）ものを採用することを選ぶことができるものがあることを関連技術分野の当業者はやはり理解するであろう。例えば、ファイアウォールは、ハードウェアもしくはソフトウェア要素またはそれらのいくつかの組合せを備えることができ、典型的には、例えば、ネットワーク管理者などのユーザによって導入されるセキュリティポリシーを実施するように設計される。

【 0 0 7 2 】

[80] 図 7 は、例えば、上記のように、本開示による、組織解析システム 7 0 0 の実施形態を概略的に示す。

[81] 例示する実施形態において、組織解析システム 7 0 0 は、画像領域識別子モジュール 7 1 0 と、画像領域分類子モジュール 7 2 0 と、データベース変更子モジュール 7 3 0 と、画像領域再分類子モジュールと、任意選択の組織イメージングデバイス 7 5 0 と、任意選択のデータ記憶システム 7 6 0 と、任意選択の組織染色デバイス 7 7 0 と、複数の通信リンク 7 8 1（読みやすさのために、通信リンクのうちの 1 つだけが参照符号を有する）を備えた通信バス 7 8 0 とを備える。通信バス 7 8 0 および通信リンク 7 8 1 は、前述の構成要素 7 1 0 ~ 7 7 0 を通信可能に相互接続する。

【 0 0 7 3 】

[82] 図 7 a は、図 7 に示す画像領域分類子モジュール 7 2 0 の詳細を概略的に示す。

[83] 例示する実施形態において、画像領域分類子モジュール 7 2 0 は、任意選択のサポートベクタマシン 7 2 2 ならびに任意選択の画像チャネル抽出器 7 2 4 を備える。サポートベクタマシン 7 2 2 および画像チャネル抽出器 7 2 4 のいずれかは、互いに、および通信バス 7 8 0 および通信リンク 7 8 1 を介して前述の構成要素 7 1 0 ~ 7 7 0 のいずれかに通信可能に相互接続することができる。

【 0 0 7 4 】

[84] 図 8 は、例えば、上記のように、本開示による、組織解析方法の実施形態の流れ図 8 0 0 を概略的に示す。

[85] 例示する実施形態において、流れ図 8 0 0 は、組織標本を染色する任意選択のステップ 8 1 0 と、（染色された）組織標本を画像化する任意選択のステップ 8 2 0 と、画像データを受け取るステップ 8 3 0 と、画像領域データを生成するステップ 8 4 0 と、画像領域を分類するステップ 8 5 0 と、信頼スコアを決定するステップ 8 6 0 と、データベースの変更を実施するステップ 8 7 0 と、画像領域を再分類するステップ 8 8 0 と、再分類された組織の種類を出力するステップ 8 9 0 とを含む。

【 0 0 7 5 】

[86] 図 8 a は図 8 に示すステップ 8 6 0 の詳細を概略的に示す。

[87] 例示する実施形態において、ステップ 8 6 0 は、サポートベクタマシンの動作を実施する任意選択のステップ 8 6 2 を含む。

【 0 0 7 6 】

[88] 図 8 b は、図 8 に示すステップ 8 8 0 の詳細を概略的に示す。

[89] 例示する実施形態において、ステップ 8 8 0 は、データを重み付けする任意選択のステップ 8 8 2 を含む。

【 0 0 7 7 】

[90] 図 8 c は、図 8 に示すステップ 8 5 0 の実施形態を概略的に示す。

[91] 例示する実施形態において、ステップ 8 5 0 は、少なくとも 1 つの特徴を画像領域から抽出する任意選択のステップ 8 5 4 を含む。

【 0 0 7 8 】

[92] 図 8 d は、図 8 に示すステップ 8 5 0 の実施形態を概略的に示す。

[93] 例示する実施形態において、ステップ 8 5 0 は、画像領域を構成要素チャネルに分離する任意選択のステップ 8 5 2 と、少なくとも 1 つの特徴を画像領域から抽出する任意選択のステップ 8 5 4 とを含む。

【 0 0 7 9 】

10

20

30

40

50

[94]本開示において、「することができる(may)」という動詞は、任意選択性/非強制を指定するのに使用される。言い換えれば、「することができる(may)」はすることができる、する必要がないものである。本開示において、「備える(comprise)」という動詞は、含むの意味で理解することができる。したがって、「備える(comprise)」という動詞は、他の要素/動作の存在を除外しない。本開示において、「第1の(first)」、「第2の(second)」、「最上位の(top)」、「最下位の(bottom)」などの関係語は、必ずしも任意の実際のそのような関係または順序をそのような実体または動作間に必要とすることなくまたは暗示することなく1つの実体または動作を別の実体または動作と区別することだけに使用することができる。

【0080】

10

[95]本開示において、「任意の(any)」という用語は、それぞれの要素のうちの任意の数を指定するものとして、例えば、それぞれの要素のうちの1つ、少なくとも1つ、少なくとも2つ、各々または全部を指定するものとして理解することができる。同様に、「任意の(any)」という用語は、それぞれの要素の任意の集合を指定するものとして、例えば、それぞれの要素の1つまたは複数の集合、それぞれの要素のうちの1つ、少なくとも1つ、少なくとも2つ、各々もしくは全部を含む集合を指定するものとして理解することができる。それぞれの集合は、要素のうちの同じ数を含む必要はない。

【0081】

[96]本開示において、「少なくとも1つ(at least one)」という表現は、任意の(整数)数または(整数)数の範囲(所与の文脈において技術的に妥当である)を指定するのに使用される。したがって、「少なくとも1つ(at least one)」という表現は、とりわけ、1つ、2つ、3つ、4つ、5つ、10、15、20または100として理解することができる。同様に、「少なくとも1つ(at least one)」という表現は、とりわけ、「1つまたは複数(one or more)」、「2つ以上(two or more)」または「5つ以上(five or more)」として理解することができる。

20

【0082】

[97]本開示において、カッコ内の表現は、任意選択として理解することができる。本開示では、引用符は引用符内の表現を比喩的な意味で理解することもできることを強調することができる。本開示では、引用符は論議中の特定の表現を識別することができる。

30

【0083】

[98]本開示において、例えば、「できる(may)」という動詞の使用またはカッコの使用を通じて、多くの特徴が任意選択であるものとして説明される。簡潔さおよび読みやすさのために、本開示は、明示的に、任意選択の特徴の組から選択することによって取得することができる各々のおよびあらゆる置き換えを列挙しない。しかし、本開示は、すべてのそのような置き換えを明示的に開示すると解釈するものとする。例えば、3つの任意選択の特徴を有するものとして説明されたシステムは、7つの異なるやり方で、すなわち、3つの可能な特徴のうちの1つだけにより、3つの可能な特徴のうちの任意の2つにより、または3つの可能な特徴のすべての3つにより、具現化することができる。

【0084】

40

さらに、本開示の代表的な実施形態を説明する際に、本明細書は、本発明の方法および/またはプロセスを特定のステップの配列として提示してきた可能性がある。しかし、方法またはプロセスが本明細書に記載した特定のステップの順序を利用しない限りにおいて、方法またはプロセスは、説明したステップの特定の配列に限定されてはならない。当技術分野の当業者が理解するように、ステップの他の配列が可能であり得る。したがって、本明細書に記載された特定のステップの順序は、特許請求の範囲に対する限定とみなしてはならない。さらに、本開示の方法および/またはプロセスを対象にした特許請求の範囲は、記載された順序でそれらのステップの性能に限定されてはならず、配列は変えることができ、それでも本開示の精神および範囲内にとどまることを当技術分野の当業者は容易に理解することができる。

50

【図1】

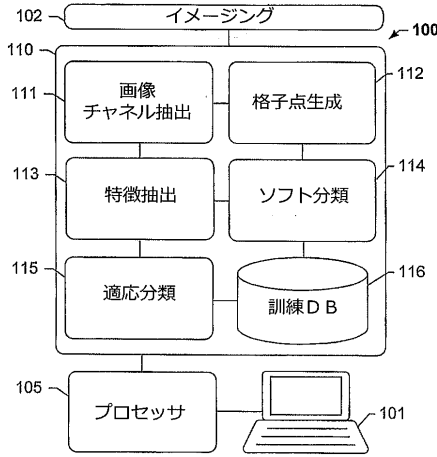


FIG. 1

【図2A】

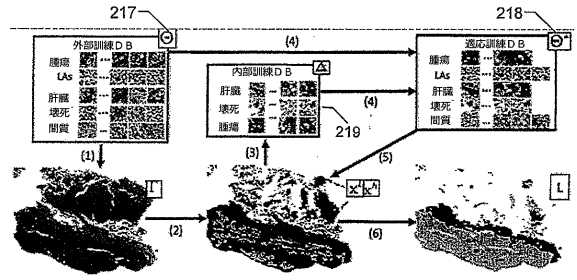


FIG. 2A

【図2B】

アルゴリズム1 適応分類
 入力：事前構築された訓練DB Θ 、テストデータ Γ 、信頼閾値 δ_c
 出力：分類結果 $L = \{l(x_i) \in T \mid x_i \in \Gamma\}$
 1: Θ を使用して分類子 F を構築する。
 2: 各 x_i に F を適用し、ラベル $l(x_i)$ 、および信頼スコア $c(x_i) \in [0, 1]$ を取得する。
 3: 高信頼試料 $x^h = \{x_i \mid c(x_i) \geq \delta_c\}$ および低信頼試料 $x^l = \{x_i \mid c(x_i) < \delta_c\}$ を見つける。
 $\Delta = \{(x_i, l(x_i)) \mid x_i \in x^h\}$ とする。
 4: アルゴリズム2を使用して、 Θ と Δ とを組み合わせることによって、適応訓練DB Θ^* を生成する。アルゴリズム2では、われわれは Θ を外部訓練DBとみなし、 Δ を内部訓練DBとみなす。
 5: Θ^* を使用して分類子 F^* を学習する。
 6: F^* を使用して、 x^l の試料を再分類し、それによって $\{l^*(x_i) \mid x_i \in x^l\}$ を取得する。
 $T: L = \{(x_i) \mid x_i \in x^h\} \cup \{(x_i) \mid x_i \in x^l\}$ を生成する。

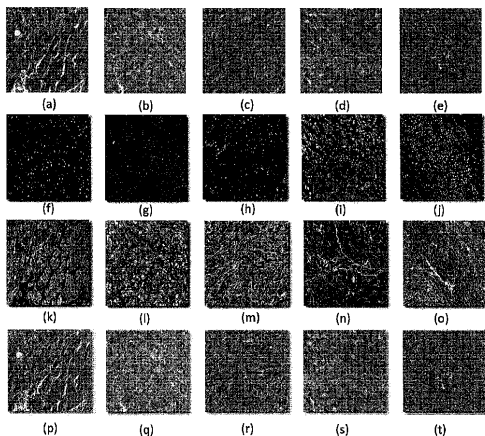
Fig. 2B

【図2C】

アルゴリズム2 適応訓練データベース生成
 入力：外部訓練DB Θ 、内部訓練DB Δ 、組織の種類閾値 δ_t
 適応訓練データベース Θ^*
 1: $T(D) = (T(D) \subset T)$ がDBD、すなわち、 $\forall t_i \in T(D)$ 、カード $\{x_i \in D \mid l(x_i) = t_i\} > \delta_t$ (カード：組内の要素の数)で表される組織の種類リストを表すとす。これは、D内に十分な t_i の試料がある場合、組織の種類 t_i がDで表される。
 2: $T(\Theta) = T$ とする、すなわち、外部DBが完全なDBであるとする(これはそれがすべての組織の種類を含むことを意味する)が、 Δ が完全なDBでない可能性がある。この仮定は、事前構築された訓練DBが完全なDBでないアルゴリズム1からの動機付けによるが、第1の分類ステップから取得した高信頼試料がすべての組織の種類を含まない可能性がある。
 3: $\Theta^* = \{(x_i, l(x_i)) \mid (x_i, l(x_i)) \in \Delta \wedge l(x_i) \in T(\Delta)\} \cup \{(x_i, l(x_i)) \mid (x_i, l(x_i)) \in l(x_i) \in T(\Theta) \wedge l(x_i) \notin T(\Delta)\}$ を生成する(組織の種類 t_i が Δ で表される場合、われわれは Θ^* を生成するために Θ ではなく t_i の試料を Δ に使用する)。

Fig. 2C

【図3A-T】



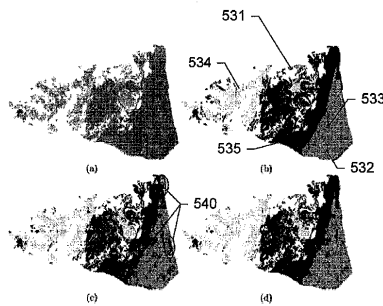
FIGS. 3A-T

【図4】



FIG. 4

【図5A-D】



FIGS. 5A-D

【 図 6 】

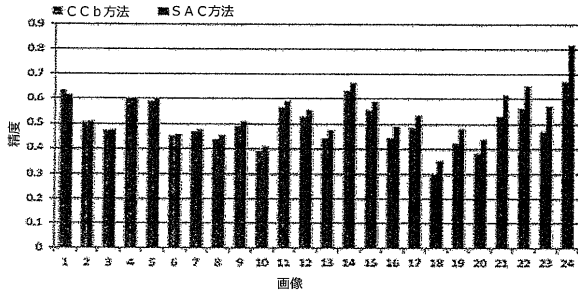


FIG. 6

【 図 7 】

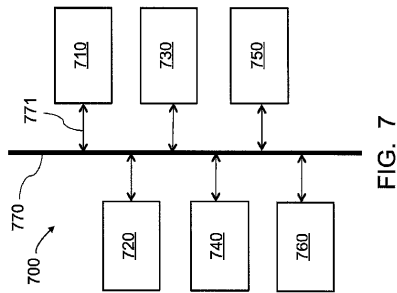


FIG. 7

【 図 7 a 】

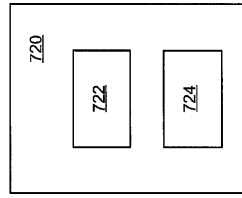


FIG. 7a

【 図 8 】

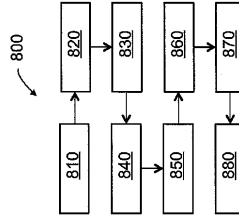


FIG. 8

【 図 8 a 】

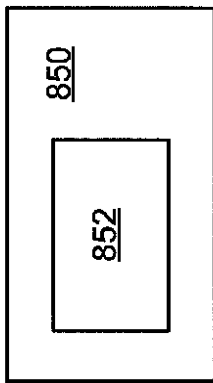


FIG. 8a

【 図 8 b 】

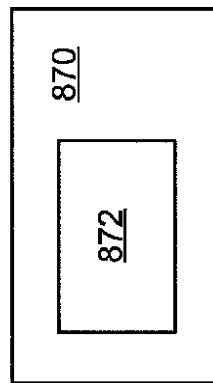



FIG. 8b

【 8 c】

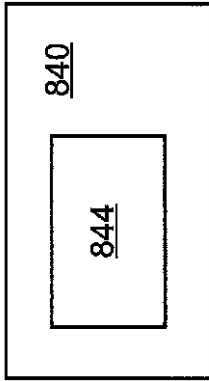



FIG. 8c

【 8 d】

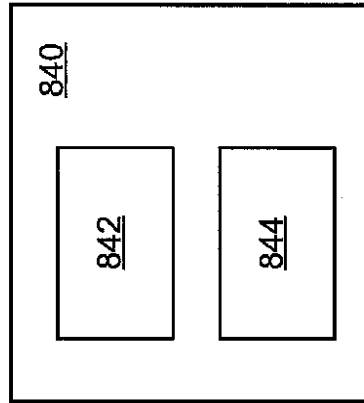


FIG. 8d

フロントページの続き

(51)Int.Cl. F I
G 0 1 N 21/27 A

(74)代理人 100147991

弁理士 鳥居 健一

(72)発明者 ブレドノ, イェルク

アメリカ合衆国カリフォルニア州 9 4 1 3 1, サンフランシスコ, トゥエンティセブンス・ストリート 3 9 7

(72)発明者 チュッカ, スリニヴァス

アメリカ合衆国カリフォルニア州 9 5 1 2 9, サンノゼ, オーク・ツリー・ドライブ 1 0 4 5

(72)発明者 チェン, ティーン

アメリカ合衆国カリフォルニア州 9 5 1 2 4, サンノゼ, ラフトン・ドライブ 4 9 3 2

(72)発明者 シェフドテル, クリストフ

アメリカ合衆国カリフォルニア州 9 4 0 8 5, サニーヴェール, ノース・ベイビュー・アベニュー 6 0 4

(72)発明者 グエン, キエン

アメリカ合衆国カリフォルニア州 9 4 0 8 6, サニーヴェール, サウス・メアリー・アベニュー 4 1 5, アパートメント 1 5

審査官 赤坂 祐樹

(56)参考文献 特開 2 0 1 0 - 2 0 3 9 4 9 (J P , A)

特開 2 0 1 1 - 0 4 7 7 1 4 (J P , A)

特開 2 0 0 1 - 3 1 1 7 3 0 (J P , A)

特表 2 0 0 9 - 5 1 6 5 5 1 (J P , A)

McKenna, SJ et al., Immunohistochemical analysis of breast tissue microarray images using contextual classifiers, J Pathol Inform., 2 0 1 3 年 3 月 3 0 日, 4(Suppl), S13

(58)調査した分野(Int.Cl., DB名)

G 0 1 N 3 3 / 4 8 - 3 3 / 9 8

G 0 6 T 1 / 0 0

G 0 6 T 7 / 0 0