

(19) 日本国特許庁(JP)

(12) 公表特許公報(A)

(11) 特許出願公表番号

特表2018-502279

(P2018-502279A)

(43) 公表日 平成30年1月25日(2018.1.25)

(51) Int.Cl.	F I	テーマコード (参考)
GO 1 N 33/48 (2006.01)	GO 1 N 33/48 M	2 G O 4 5
GO 1 N 33/53 (2006.01)	GO 1 N 33/53 Y	4 B O 2 9
GO 6 T 7/00 (2017.01)	GO 6 T 7/00 6 3 0	4 B O 6 3
GO 1 N 33/574 (2006.01)	GO 1 N 33/574 A	5 L O 9 6
C 1 2 M 1/34 (2006.01)	GO 1 N 33/48 P	

審査請求 未請求 予備審査請求 未請求 (全 53 頁) 最終頁に続く

(21) 出願番号 特願2017-525014 (P2017-525014)
 (86) (22) 出願日 平成27年11月9日 (2015.11.9)
 (85) 翻訳文提出日 平成29年6月27日 (2017.6.27)
 (86) 国際出願番号 PCT/EP2015/076105
 (87) 国際公開番号 W02016/075096
 (87) 国際公開日 平成28年5月19日 (2016.5.19)
 (31) 優先権主張番号 62/077,536
 (32) 優先日 平成26年11月10日 (2014.11.10)
 (33) 優先権主張国 米国 (US)
 (31) 優先権主張番号 62/144,364
 (32) 優先日 平成27年4月8日 (2015.4.8)
 (33) 優先権主張国 米国 (US)

(71) 出願人 507179346
 ベンタナ メディカル システムズ, インコーポレイテッド
 アメリカ合衆国アリゾナ州85755, トゥーソン, イースト・イノベーション・パーク・ドライブ 1910
 (74) 代理人 100140109
 弁理士 小野 新次郎
 (74) 代理人 100118902
 弁理士 山本 修
 (74) 代理人 100106208
 弁理士 宮前 徹
 (74) 代理人 100120112
 弁理士 中西 基晴

最終頁に続く

(54) 【発明の名称】 組織学画像中の核の分類

(57) 【要約】

(1) 組織標本の画像を入力として提供するステップと、(2) 画像内の核の特徴から核特徴メトリクスを計算するステップ(111)と、(3) 画像を使用して対象の核に基づいて文脈情報メトリクスを計算するステップ(112)と、(4) 核特徴メトリクスと文脈情報メトリクスの組合せを使用して画像中の細胞を分類するステップ(113)とを含む、組織標本の画像内の細胞を分類するコンピュータデバイス(14)及びコンピュータ実現方法が開示される。

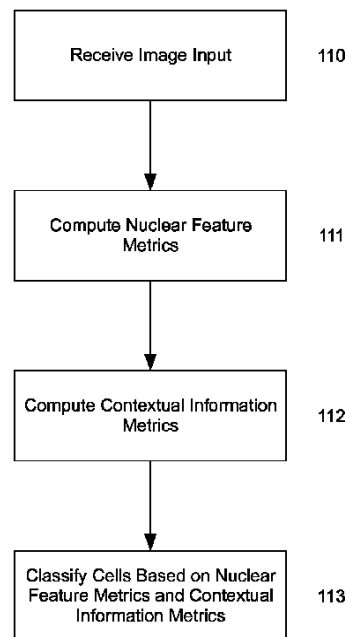


FIG. 1C

【特許請求の範囲】**【請求項 1】**

1つ又は複数のプロセッサ及び少なくとも1つのメモリを備える、PD-L1バイオマーカーの存在に関するIHCアッセイで染色された組織標本の画像(110)内の細胞を分類するためのコンピュータシステム(14)であって、前記少なくとも1つのメモリが、前記1つ又は複数のプロセッサによる実行のための非一時的コンピュータ可読命令であって、前記1つ又は複数のプロセッサに、

前記組織標本の前記画像中の細胞核内の特徴から導出される核特徴メトリクスを計算させ(111)、

前記組織標本の前記画像内の対象の核の文脈情報メトリクスを計算させ(112)、

前記核特徴メトリクス及び文脈情報メトリクスに基づいて、前記組織標本の前記画像内の細胞を、陽性免疫細胞、陽性腫瘍細胞、陰性免疫細胞及び陰性腫瘍細胞又は他の細胞のうち少なくとも1つとして分類させる

ための非一時的コンピュータ可読命令を記憶するコンピュータシステム(14)。

【請求項 2】

前記核特徴が、形態学特徴、外観特徴及び背景特徴からなるグループから選択される、請求項1に記載のコンピュータシステム。

【請求項 3】

前記核特徴が、局所総合染色強度を表す第1の画像チャンネル、PD-L1バイオマーカーの存在を示すIHCラベルの強度を表す第2の画像チャンネル、及び/又は細胞核の存在を示す局所対比染色強度を表す第3の画像チャンネル上で計算される、請求項1又は2のいずれかに記載のコンピュータシステム。

【請求項 4】

前記文脈情報メトリクスが、(i)隣接する核を記述しているデータ、及び(ii)前記対象の核を取り囲んでいる画像テクスチャ特徴のうち少なくとも1つから導出される、請求項1から3のいずれか一項に記載のコンピュータシステム。

【請求項 5】

前記画像テクスチャ特徴が、前記組織標本の前記画像中の対象の核を取り囲んでいる画像パッチから導出される、請求項4に記載のコンピュータシステム。

【請求項 6】

前記導出される画像テクスチャ特徴が、テクストンヒストグラム特徴、Garbor特徴、Haralick特徴、強度特徴のヒストグラム、及び傾斜量及び傾斜配向特徴のヒストグラムからなるグループから選択される、請求項5に記載のコンピュータシステム。

【請求項 7】

前記導出される画像テクスチャ特徴がテクストンヒストグラム特徴であり、例えば前記テクストンヒストグラムが、前記テクストンの各々にそれぞれ割り当てられる前記画像パッチに含まれているピクセルの数を示す、請求項6に記載のコンピュータシステム。

【請求項 8】

前記テクストンヒストグラム特徴が、(1)フィルタ応答画像であって、個々のフィルタ応答画像が1つ又は複数のフィルタ応答を含むフィルタ応答画像のリストを得るために、前記組織標本の前記画像に対して最大応答フィルタのバンクを適用すること、(2)前記フィルタ応答を前記フィルタ応答画像からテクストンにクラスタ化すること、(3)前記組織標本の前記画像中の個々のピクセルを前記テクストンのうちの1つに割り当てること、及び(4)前記対象の核の各々を取り囲んでいる前記画像パッチ中のすべてのピクセルから前記テクストンヒストグラムを計算することによって導出される、請求項7に記載のコンピュータシステム。

【請求項 9】

前記導出される画像テクスチャ特徴が強度特徴のヒストグラムであり、強度特徴の前記ヒストグラムが、一次染色チャンネル、対比染色チャンネル、IHC染色チャンネル及び輝度チャンネルからなるグループから選択される画像チャンネルから導出される、請求項6から8の

10

20

30

40

50

いずれか一項に記載のコンピュータシステム。

【請求項 10】

前記対象の核を取り囲んでいる前記画像パッチ内の強度ベース特徴を計算するために、異なる画像チャンネルからの信号の差が解析され、前記強度ベース特徴が、例えば前記異なる画像チャンネルの前記画像パッチ内のピクセル強度値から導出されるメトリック値を含む、請求項 9 に記載のコンピュータシステム。

【請求項 11】

前記導出される画像テクスチャ特徴が Haralick 特徴であり、前記 Haralick 特徴が、前記画像パッチ中のピクセルとその指定された隣接ピクセルとの間の角関係に基づいて共起行列を計算するための命令を提供することによって導出される、請求項 6 から 10 のいずれか一項に記載のコンピュータシステム。

10

【請求項 12】

隣接する核を記述している前記データが、クラスタ割当てのヒストグラム (630) から導出される、請求項 4 から 11 のいずれか一項に記載のコンピュータシステム。

【請求項 13】

クラスタ割当ての前記ヒストグラム (630) が、(1) クラスタ中心を得るために、核特徴ベクトルに対して K 平均アルゴリズムを適用すること、(2) 対象の特定の核の個々の隣接する核を最も近いクラスタ中心に割り当てること、及び (3) 前記割当てに基づいてクラスタ割当ての前記ヒストグラムを計算することによって導出される、請求項 12 に記載のコンピュータシステム。

20

【請求項 14】

(2a) 個々の隣接する核の前記核特徴ベクトルから個々のクラスタの中心までの Euclidean 距離を測定するステップ、及び (2b) 前記個々の隣接する核を、その中心がその核の核特徴ベクトルに最も近いクラスタに割り当てるステップをさらに含む、請求項 13 に記載のコンピュータシステム。

【請求項 15】

前記組織標本の前記画像の前記細胞が、サポート・ベクター・マシンを使用して分類される、請求項 1 から 14 のいずれか一項に記載のコンピュータシステム。

【請求項 16】

前記標本が肺組織標本である、請求項 1 から 15 のいずれか一項に記載のコンピュータシステム。

30

【請求項 17】

前記組織標本の前記画像内の前記細胞の前記分類が訓練された分類器によって実施され、前記命令が前記 1 つ又は複数のプロセッサに、

(a) 個々の訓練核から 1 つ又は複数の核特徴メトリクスを抽出することによって個々の訓練画像の個々の訓練核に対する核特徴ベクトルを生成すること、(b) K 平均アルゴリズムを使用してクラスタ化手順を前記核特徴ベクトルに対して実施することによって、複数の (C 個の) 事前訓練済みクラスタを獲得すること、(c) 前記訓練画像のうちの 1 つの中の前記訓練核に隣接する個々の核を、(c1) 個々の個別の隣接する核の前記核特徴ベクトルから個々のクラスタの中心までの Euclidean 距離を測定し、かつ、(c2) 個々の訓練画像の個々の核の個々の隣接する核を、その核の中心を表す前記核特徴ベクトルまでの最も短い Euclidean 距離を有するクラスタに割り当てることによって、前記複数の C 個のクラスタのうちの 1 つに割り当てること、(d) 前記訓練画像中の前記訓練核の各々に対して、前記 1 つの訓練核の隣接する核であるすべての訓練核の前記クラスタ割当てのヒストグラムを計算することによって前記訓練核の文脈特徴メトリクスを決定すること、及び (e) 前記訓練画像の各々における個々の訓練核に対して、前記核に対して決定された前記核特徴及び前記文脈特徴の核特徴を前記訓練核のための単一の完全な特徴ベクトルに結合すること、(f) すべての訓練核の前記完全な特徴ベクトルを入力として使用して分類モデルを訓練することによって、前記分類器の計算訓練を複数の訓練画像のうちの 1 つの中で識別された核である複数の訓練核に対して自動的に実施する

40

50

ことによって前記分類器を生成させる、請求項 1 から 16 のいずれか一項に記載のコンピュータシステム。

【請求項 18】

前記命令が前記 1 つ又は複数のプロセッサに、前記組織標本の前記画像中の前記対象の核の各々に対して、

(a) 前記訓練された分類器によって、前記核特徴メトリクスの前記計算を実施することによって前記対象の核に対する核特徴メトリクスベクトルを生成するステップと、(b) 前記組織標本の前記画像内の前記対象の核の隣接する核である前記組織標本画像中の個々の核を、(c 1) 個々の個別の隣接する核の前記核特徴ベクトルから前記訓練された識別器の前記 (C 個の) クラスタ中心までの Euclidean 距離を測定し、かつ、(c 2) 前記個々の隣接する核を、その中心がその核の前記核特徴ベクトルに最も近いクラスタに割り当てることによって、前記 (C 個の) 事前訓練済みクラスタのうちの 1 つに割り当てるステップと、(d) 前記対象の核の前記隣接する核の前記クラスタ割当てのヒストグラムを計算することによって、前記組織標本の前記画像中の個々の対象の核の前記文脈情報メトリクスの前記計算を実施するステップと、(e) 前記組織標本の前記画像中の前記対象の核の前記核特徴ベクトルを前記文脈情報メトリクスと結合して、前記対象の核の完全な特徴ベクトルにするステップと、(f) 前記訓練された分類器を前記対象の核の前記完全な特徴ベクトルに対して適用してそれを分類することによって、前記対象の核を含む前記組織標本の前記画像内の前記細胞のうちの 1 つの分類を実施するステップと

10

を含む方法を実施させる、請求項 17 に記載のコンピュータシステム。

20

【請求項 19】

PD-L1 バイオマーカの存在に関する IHC アッセイで染色された組織標本の画像内の細胞を分類するコンピュータ実現方法であって、前記組織標本の前記画像内の核の特徴から核特徴メトリクスを計算するステップと、前記組織標本の前記画像内の対象の核に基づいて文脈情報メトリクスを計算するステップと、前記核特徴メトリクスと文脈情報メトリクスの組合せを使用して、前記組織標本の前記画像内の前記細胞を分類するステップとを含み、前記細胞が、陽性免疫細胞、陽性腫瘍細胞、陰性免疫細胞及び陰性腫瘍細胞又は他の細胞のうちの少なくとも 1 つとして分類されるコンピュータ実現方法。

【請求項 20】

前記方法が、前記細胞内の個々の核を識別するための前景分割マスクを生成するステップをさらに含む、請求項 19 に記載のコンピュータ実現方法。

30

【請求項 21】

前記核特徴が、形態学特徴、外観特徴及び背景特徴からなるグループから選択される、請求項 19 又は 20 に記載のコンピュータ実現方法。

【請求項 22】

前記文脈情報メトリクスが、(i) 隣接する核を記述しているデータ、及び (i i) 対象の核を取り囲んでいる画像テクスチャ特徴のうちの少なくとも 1 つから導出される、請求項 19 から 21 のいずれか一項に記載のコンピュータ実現方法。

【請求項 23】

前記方法が、対象の特定の核を取り囲んでいる画像パッチを生成するステップをさらに含み、前記画像テクスチャ特徴が、前記生成された画像パッチから導出される、請求項 22 に記載のコンピュータ実現方法。

40

【請求項 24】

前記導出される、対象の核を取り囲んでいる画像テクスチャ特徴が、テクストンヒストグラム特徴、Garbor 特徴、Haralick 特徴、強度特徴のヒストグラム、及び傾斜量及び傾斜配向特徴のヒストグラムからなるグループから選択される、請求項 23 に記載のコンピュータ実現方法。

【請求項 25】

前記導出される画像テクスチャ特徴がテクストンヒストグラム特徴である、請求項 24 に記載のコンピュータ実現方法。

50

【請求項 26】

前記テキストヒストグラム特徴が、(1)フィルタ応答画像のリストを得るために、前記組織標本の前記画像に対して最大応答フィルタのバンクを適用すること、(2)前記フィルタ応答を前記フィルタ応答画像からテキストにクラスタ化すること、(3)前記組織標本の前記画像中の個々のピクセルを前記テキストのうちの1つに割り当てること、及び(4)対象の核を取り囲んでいる画像パッチ中のすべてのピクセルからテキストヒストグラムを計算するための命令を提供することによって導出される、請求項25に記載のコンピュータ実現方法。

【請求項 27】

前記導出される画像テクスチャ特徴が強度特徴のヒストグラムであり、強度特徴の前記ヒストグラムが、一次染色チャンネル、対比染色チャンネル、IHC染色チャンネル及び輝度チャンネルからなるグループから選択される画像チャンネルから計算される、請求項24から26のいずれか一項に記載のコンピュータ実現方法。

10

【請求項 28】

前記対象の核を取り囲んでいる前記画像パッチ内の強度ベース特徴を計算するために、異なる画像チャンネルからの信号の差が捕獲される、請求項27に記載のコンピュータ実現方法。

【請求項 29】

前記導出される画像テクスチャ特徴がHaralick特徴であり、前記Haralick特徴が、前記画像パッチ中のピクセルとその指定された隣接ピクセルとの間の角関係に基づいて共起行列を計算することによって導出される、請求項24から28のいずれか一項に記載のコンピュータ実現方法。

20

【請求項 30】

隣接する核を記述している前記データが、クラスタ割当てのヒストグラム(630)から導出される、請求項22から29のいずれか一項に記載のコンピュータ実現方法。

【請求項 31】

クラスタ割当ての前記ヒストグラムが、(1)クラスタ中心を得るために、核特徴ベクトルに対してK平均アルゴリズムを適用すること、(2)対象の特定の核の個々の隣接する核を最も近いクラスタ中心に割り当てること、及び(3)前記割当てに基づいてクラスタ割当ての前記ヒストグラムを計算することによって導出される、請求項30に記載のコンピュータ実現方法。

30

【請求項 32】

(2a)個々の隣接する核の前記核特徴ベクトルから個々のクラスタの中心までのEuclidean距離を測定するステップ、及び(2b)前記個々の隣接する核を、その中心がその核の核特徴ベクトルに最も近いクラスタに割り当てるステップをさらに含む、請求項31に記載のコンピュータ実現方法。

【請求項 33】

前記組織標本の前記画像の前記細胞が、サポート・ベクター・マシンを使用して分類される、請求項19から32のいずれか一項に記載のコンピュータ実現方法。

【請求項 34】

前記組織標本の前記画像内の前記細胞の前記分類が訓練された分類器によって実施され、前記方法が、

40

(a)個々の訓練核から1つ又は複数の核特徴メトリクスを抽出することによって個々の訓練画像の個々の訓練核に対する核特徴ベクトルを生成するステップと、(b)K平均アルゴリズムを使用してクラスタ化手順を前記核特徴ベクトルに対して実施することによって、複数の(C個の)事前訓練済みクラスタを獲得するステップと、(c)前記訓練画像のうちの1つの中の前記訓練核に隣接する個々の核を、(c1)個々の個別の隣接する核の前記核特徴ベクトルから個々のクラスタの中心までのEuclidean距離を測定し、かつ、(c2)個々の訓練画像の個々の核の個々の隣接する核を、その核の中心を表す前記核特徴ベクトルまでの最も短いEuclidean距離を有するクラスタに割り当

50

ることによって、前記複数のC個のクラスタのうちの一つに割り当てるステップと、(d)前記訓練画像中の前記訓練核の各々に対して、前記一つの訓練核の隣接する核であるすべての訓練核の前記クラスタ割当てのヒストグラムを計算することによって前記訓練核の文脈特徴メトリクスを決定するステップと、(e)前記訓練画像の各々における個々の訓練核に対して、前記核に対して決定された前記核特徴及び前記文脈特徴の核特徴を前記訓練核のための単一の完全な特徴ベクトルに結合するステップと、(f)すべての訓練核の前記完全な特徴ベクトルを入力として使用して分類モデルを訓練するステップとによって、前記分類器の計算訓練を複数の訓練画像のうちの一つの中で識別された核である複数の訓練核に対して自動的に実施するステップと

をさらに含む、前記請求項19から33のいずれか一項に記載のコンピュータ実現方法

10

【請求項35】

前記組織標本の前記画像中の前記対象の核の各々に対して、

(a)前記訓練された分類器によって、前記核特徴メトリクスの前記計算を実施することによって前記対象の核に対する核特徴メトリクスベクトルを生成するステップと、(b)前記組織標本の前記画像内の前記対象の核の隣接する核である前記組織標本画像中の個々の核を、(c1)個々の個別の隣接する核の前記核特徴ベクトルから前記訓練された識別器の前記(C個の)クラスタ中心までのEuclidean距離を測定し、かつ、(c2)前記個々の隣接する核を、その中心がその核の前記核特徴ベクトルに最も近いクラスタに割り当てることによって、前記(C個の)事前訓練済みクラスタのうちの一つに割り当てるステップと、(d)前記対象の核の前記隣接する核の前記クラスタ割当てのヒストグラムを計算することによって、前記組織標本の前記画像中の個々の対象の核の前記文脈情報メトリクスの前記計算を実施するステップと、(e)前記組織標本の前記画像中の前記対象の核の前記核特徴ベクトルを前記文脈情報メトリクスと結合して、前記対象の核の完全な特徴ベクトルにするステップと、(f)前記訓練された分類器を前記対象の核の前記完全な特徴ベクトルに対して適用してそれを分類することによって、前記対象の核を含む前記組織標本の前記画像内の前記細胞のうちの一つの分類を実施するステップと

20

を実施するステップをさらに含む、請求項34に記載のコンピュータ実現方法。

【請求項36】

請求項1から18までのいずれか一項に記載のコンピュータシステム(14)及び画像化装置(12)を備える細胞分析器。

30

【請求項37】

PD-L1発現のための腫瘍標本をスコアリングする方法であって、

(a)前記腫瘍標本中の腫瘍細胞及び免疫細胞を識別するステップであって、

(a1)請求項1から18までのいずれかに記載のシステム

(a2)請求項19から35までのいずれかに記載の方法、又は

(a3)請求項36に記載の細胞分析器

のうち少なくとも一つを使用して識別するステップと、

(b)PD-L1を発現する腫瘍細胞及び免疫細胞の数及び/又は前記細胞中のPD-L1発現の相対強度を決定するステップと、

40

(c)ステップ(b)で決定された前記PD-L1発現に従って腫瘍を類別するステップと

を含む方法。

【請求項38】

PD-L1の前記発現が、前記腫瘍中のPD-L1タンパク質及び/又はPD-L1 mRNAを特異的に検出することによって決定される、請求項37に記載の方法。

【請求項39】

前記細胞が免疫組織化学的に検出されたPD-L1タンパク質の少なくとも部分膜染色を有している場合、前記細胞がPD-L1を発現していると思なされる、請求項37又は38のいずれか一項に記載の方法。

50

【請求項 40】

前記腫瘍が、それぞれステップ (b) から計算される修正 H - スコア (MHS)、修正比率スコア (MPS)、又は MHS と MPS の両方によって類別される、請求項 37 から 39 のいずれか一項に記載の方法。

【請求項 41】

1 つ又は複数のプロセッサ及び少なくとも 1 つのメモリを備える、PD - L1 バイオマーカーの存在に関する IHC アッセイで染色された組織標本の画像内の細胞を分類するためのコンピュータシステムであって、前記少なくとも 1 つのメモリが、前記 1 つ又は複数のプロセッサによる実行のための非一時的コンピュータ可読命令であって、前記 1 つ又は複数のプロセッサに、前記組織標本の前記画像から核特徴メトリクス (111) 及び文脈情報メトリクス (112) を導出するための特徴抽出モジュール (101) を走らせ、また、前記組織標本の前記画像内の前記細胞を分類するための分類モジュール (102) であって、前記細胞が、陽性免疫細胞、陽性腫瘍細胞、陰性免疫細胞及び陰性腫瘍細胞又は他の細胞のうちの少なくとも 1 つとして分類される分類モジュール (102) を走らせるための非一時的コンピュータ可読命令を記憶するコンピュータシステム。

10

【請求項 42】

前記文脈情報メトリクスが、(i) 隣接する核を記述しているデータ、及び (ii) 前記対象の核を取り囲んでいる画像テクスチャ特徴のうちの少なくとも 1 つから導出される、請求項 41 に記載のコンピュータシステム。

【請求項 43】

前記導出される画像テクスチャ特徴が、前記対象の核を取り囲んでいる画像パッチから計算される、請求項 42 に記載のコンピュータシステム。

20

【請求項 44】

前記導出される画像テクスチャ特徴が、文脈 - テクスチャ方法又は文脈 - テクストン方法のうちの 1 つの適用を介して計算される、請求項 43 に記載のコンピュータシステム。

【請求項 45】

隣接する核を記述している前記データが、文脈条件付き確率場方法又は文脈語の多重集合方法のうちの 1 つの適用を介して計算される、請求項 42 から 44 のいずれか一項に記載のコンピュータシステム。

【請求項 46】

前記システムが、前記組織標本の組織試料の画像を得るために画像収集モジュールを走らせるステップをさらに含む、請求項 41 から 45 のいずれか一項に記載のコンピュータシステム。

30

【発明の詳細な説明】

【技術分野】

【0001】

関連出願の相互参照

[001] 本出願は、2015年4月8日に提出した米国仮特許出願第62/144364号の出願日の利益、及び2014年11月10日に提出した米国仮特許出願第62/077536号の利益を主張するものであり、これらの開示は、参照により本明細書に組み込まれている。

40

【背景技術】

【0002】

[002] 組織病理学画像から生物学的構造を自動で識別することは、デジタル病理学における中心的な課題である。多くの用途では、スライド組織の全体画像が与えられると、正常組織、腫瘍組織、壊死、リンパ管領域及び間質の中に出現する画像中の異なるタイプの細胞又は他の構造を識別することが必要になる。実際、標本中のこのような細胞又は構造の存在の定量評価は、癌を治療するために特定の化学療法薬品を選択するなど、特定の療法の影響を決定するために、しばしば必要とされる。例えば PD - L1 (Programmed death - ligand 1: プログラム死リガンド1) を発現する腫瘍細胞

50

は、PD-1/PD-L1経路の活性化を介して免疫応答を抑制するとされており、また、データによると、PD-L1腫瘍状態がPD-1誘導療法(directed therapies)及びPD-L1誘導療法に対する応答の前兆であり得ることが示されている。したがってPD-L1核分類及び定量化は、デジタル病理学における重要な仕事である。

【0003】

[003]定量PD-L1組織分析には、タイプ(腫瘍、免疫、間質、等々)に応じた、又はPD-L1染色に対する応答に応じた細胞又は核の検出及びラベル付けがしばしば必要である。PD-L1バイオマーカは、腫瘍細胞及び免疫細胞の膜の上に発現され得る。画像データの生物学的に意味のあるあらゆる自動分析は、第1にすべての細胞及び染色パターンを検出し、かつ、(1)PD-L1陽性免疫細胞、(2)PD-L1陽性腫瘍細胞、(3)PD-L1陰性免疫細胞(免疫組織化学(IHC)染色を使用することなく、その核によって視認できる細胞)、(4)PD-L1陰性腫瘍細胞(PD-L1染色を使用することなく、その核の出現によって識別される細胞)、(5)間質細胞、正常組織細胞、等々を含むあらゆる他の細胞、及び/又は(6)アーチファクト、背景染色、等々を含む、細胞を表さない染色のうちの一つとしてそれらを識別しなければならない。

10

【0004】

[004]PD-L1の文脈においては、分析は、細胞及びそれらのIHC染色を検出するだけでなく、追加として、IHC染色が存在する理由を決定し、かつ、分類しなければならない。例えば局所染色の吸収は、PD-L1陽性腫瘍細胞、PD-L1陽性免疫細胞又は非標的人工染色によって引き起こされ得る。さらに、免疫細胞及び腫瘍細胞は、閉空間近辺で、互いに接触しているPD-L1陽性及び陰性細胞と共に存在し得る。実際、単一の細胞を適切に識別するためには、細胞の核及び可能な膜の染色の外観が、複数の細胞、それらの外観、及びそれらの局所近辺における染色パターンと共に評価されなければならない。

20

【発明の概要】

【発明が解決しようとする課題】

【0005】

[005]高倍率におけるスライド画像全体のサイズが大きく、また、処理されるべきデータが大量であるため、病理学者による画像の評価は問題である。実際、スライド組織画像全体に存在する細胞又は細胞核の数は、典型的には 10^4 程度になり、実行不可能ではないにしても、病理学者によるこのような仕事の手動での実施を困難にしている。したがってそれ自体の出現及びその局所組織文脈における細胞の出現に基づいて個々の細胞又は細胞核を識別することができる自動定量アッセイを開発することが望ましい。

30

【課題を解決するための手段】

【0006】

[006]本開示の一態様では、1つ又は複数のプロセッサ及び少なくとも1つのメモリを備える、PD-L1バイオマーカの存在に関するIHCアッセイで染色された組織標本の画像内の細胞を分類するためのコンピュータシステムに関し、少なくとも1つのメモリは、1つ又は複数のプロセッサによる実行のための非一時的コンピュータ可読命令であって、その1つ又は複数のプロセッサに、組織標本の画像中の細胞核内の特徴から導出される核特徴メトリクスを計算させ、組織標本の画像内の対象(of interest)核の文脈情報メトリクスを計算させ、また、核特徴メトリクス及び文脈情報メトリクスに基づいて、組織標本の画像内の細胞を、陽性免疫細胞、陽性腫瘍細胞、陰性免疫細胞及び陰性腫瘍細胞又は他の細胞のうちの一つとして分類させるための非一時的コンピュータ可読命令を記憶する。

40

【0007】

[007]いくつかの実施形態では、核特徴は、形態学特徴、外観特徴及び背景特徴からなるグループから選択される。いくつかの実施形態では、核特徴は、局所総合染色強度、PD-L1バイオマーカの存在を示すIHCラベルの強度、及び細胞核の存在を示す局所対

50

比染色強度を表す画像チャンネル上で計算される。

【0008】

[008]いくつかの実施形態では、文脈情報メトリクスは、(i)隣接する核を記述しているデータ、及び(ii)対象の核を取り囲んでいる画像テクスチャ特徴のうち少なくとも1つから導出される(対象の核のために)。

【0009】

[009]いくつかの実施形態では、画像テクスチャ特徴は、組織標本の画像中の対象の核を取り囲んでいる画像パッチから導出される。いくつかの実施形態では、導出される画像テクスチャ特徴は、テクストン(texton)ヒストグラム特徴、Garbor特徴、Haralick特徴、強度特徴のヒストグラム、及び傾斜量(gradient magnitude)及び傾斜配向特徴のヒストグラムからなるグループから選択される。

10

【0010】

本明細書において使用されている「隣接する核」は、例えば画像内の対象の核から定義済みの最大距離以内に存在する核である。

「画像テクスチャ特徴」は、例えば画像解析機能によって計算される特性値又はメトリクスであり、画像領域のテクスチャを定量化する。画像テクスチャ特徴は、画像中又は画像の選択された領域の色又は強度の空間配置に関する情報を提供する。

【0011】

本明細書において使用されている「強度特徴のヒストグラム」は、例えば画像又は画像チャンネルの強度値の発生頻度の分布である。例えばヒストグラムは、その画像中に見出される個々の異なる強度値における画像中のピクセルの数を示すグラフとして表され得る。

20

【0012】

「テクストン」は、例えばピクセルプロブ(blob)、又は基準ピクセルから最大距離未満の離れた位置に存在する一組のピクセルの一組の1つ又は複数の属性であり、それにより前記属性が観察されているか、又は画像内で反復性であることが期待される。例えばテクストンは、配向された線形フィルタ出力の頻繁に共起(co-occur)する組合せであってもよい。ピクセルプロブは、例えば核を含む細胞内に存在するものとして識別される核プロブ又はピクセル領域であってもよい。基準ピクセルは、例えば核中心又は細胞中心あるいは細胞膜であってもよい。したがって「テクストン」は、「視覚語」、例えば特定のサイズ又は寸法の楕円、特定の平均強度値の円、特定の強度分布又はパターンを有するピクセルプロブ、等々で見なされ得る。

30

【0013】

「テクストンヒストグラム特徴」は、例えば画像中又は特定の画像チャンネル中で識別されるテクストン(「視覚語」)の発生頻度の分布である。例えばテクストンヒストグラムは、特定のタイプのテクストンの数を示すグラフとして表され得る。例えば軸a1.1及びa1.2を有する「楕円体1」、軸a2.1及びa2.2を有する「楕円体2」、及び直径d1及び強度値範囲=[09-110]を有する「円1」の3つのタイプのテクストンは、画像から抽出され得る。テクストンヒストグラム特徴は、テクストン「楕円体1」が79回見出され、テクストン「楕円体2」が画像中で1124回見出されたこと、また、「円1」が画像中で34回見出されたことを示すヒストグラムであり得る。

40

【0014】

[0010]本明細書において使用されている「傾斜」は、例えば、前記特定のピクセルを取り囲んでいる一組のピクセルの強度値傾斜を考慮することによって特定のピクセルに対して計算されたピクセルの強度傾斜である。個々の傾斜は、x軸及びy軸がデジタル画像の2つの直交縁によって画定される座標系に対する特定の「配向」を有し得る。「傾斜配向特徴」は、前記座標系内の傾斜の配向を示すデータ値であり得る。

【0015】

[0011]「陽性腫瘍細胞」又は「PD-L1陽性腫瘍細胞」は、腫瘍細胞として識別された細胞であり、支配的にPD-L1バイオマーカを発現する。「陰性腫瘍細胞」又は「PD-L1陰性腫瘍細胞」は、腫瘍細胞として識別された細胞であり、PD-L1バイオマ

50

ーカをごく弱く発現するか、又は全く発現しない。「支配的に」は、例えば染色されたPD-L1バイオマーカから放出される光信号を識別するために使用される色チャンネルのピクセル強度が所与の閾値強度より強いことを意味することができ、また、「弱い発現」は、例えば染色されたPD-L1バイオマーカから放出される光信号を識別するために使用される色チャンネルのピクセル強度が所与の閾値強度より弱いことを意味することができる。PD-L1染色によって生じる信号の強度は、腫瘍細胞又は非腫瘍細胞としての細胞の分類に対する影響を有する1つの特徴であり得るが、追加特徴も同じく評価され、したがって分類は、腫瘍細胞として分類されたPD-L1陽性細胞、並びに非腫瘍細胞として分類されたPD-L1陽性細胞をもたらし得る。「陰性リンパ球」又は「PD-L1陰性リンパ球」は、PD-L1バイオマーカをごく弱く発現するか、又は全く表さないリンパ球細胞として分類された細胞である。「陽性リンパ球」又は「PD-L1陽性リンパ球」は、腫瘍細胞として識別されたリンパ球であり、支配的にPD-L1バイオマーカを発現する。

10

【0016】

[0012]「前景分割マスク」は、例えば、他のピクセル（「背景」を構成している）からの1つ又は複数のピクセルプロブ（「前景ピクセル」として使用される）の分割を可能にする分割アルゴリズムによって生成される画像マスクである。例えば前景分割マスクは、核分割アルゴリズムによって生成されることが可能であり、また、組織切片を描写している画像上への前景分割マスクの適用は、画像中の核プロブの識別を可能にし得る。

【0017】

[0013]いくつかの実施形態では、導出される画像テクスチャ特徴は、テクストンヒストグラム特徴である。いくつかの実施形態では、テクストンヒストグラム特徴は、(1)フィルタ応答画像のリストを得るために、組織標本の画像に対して最大応答フィルタのバンクを適用すること、(2)フィルタ応答をフィルタ応答画像からテクストンにクラスタ化すること、(3)組織標本の画像中の個々のピクセルをテクストンのうちの1つに割り当てること、及び(4)対象の核を取り囲んでいる画像パッチ中のすべてのピクセルからテクストンヒストグラムを計算することによって導出される。いくつかの実施形態では、導出される画像テクスチャ特徴は強度特徴のヒストグラムであり、強度特徴のヒストグラムは、一次染色チャンネル、対比染色チャンネル、IHC染色チャンネル（例えばPD-L1の存在を示すチャンネル）及び輝度チャンネルからなるグループから選択される画像チャンネルから導出される。いくつかの実施形態では、対象の核を取り囲んでいる画像パッチ内の強度ベース特徴を計算するために、異なる画像チャンネルからの信号の差が解析される。

20

30

【0018】

[0014]「フィルタバンク」は、2つ以上のフィルタの集合である。例えばフィルタバンクは、エッジフィルタ、パーフィルタ及びスポットフィルタの混合を複数のスケール及び配向、並びに異なる位相で備えることができ、例えばGaussフィルタのLaplace演算子及びGaussフィルタを備えることができる。回転不変フィルタセット並びに回転変化フィルタセットが存在している。48個のフィルタを有する多重スケール多重配向フィルタバンクの一例は、Leung-Malik (LM) Filter Bankである。画像は、しばしば、強度の無作為変化によって、又は乏しいコントラストを有することによって害われる。1つ又は複数のフィルタリングされた画像を得るための画像上へのフィルタの適用は、ピクセル強度値を微分データ値に変換するために実施されることが可能であり、微分データ値は、関連する画像特性、例えば特定のテクスチャを表す画像特性を明らかにし、及び/又はより良好なコントラストを有する。

40

【0019】

[0015]本発明の実施形態には、「最大応答」(MR)フィルタを備えるフィルタバンクが優先的に使用される。最大応答フィルタバンクは、特定のタイプの少なくとも1つのフィルタの複数のコピーを備えるフィルタバンクであり、個々のコピーは別様に配向される。前記特定のフィルタの複数のコピーを画像又は画像領域のピクセル強度に対して異なる回転角度で適用することにより、フィルタの適用が最大の信号を返した、前記異なる配向

50

のうちの1つが識別される。したがってMRフィルタセットは、対象の特定の核に対する、特定のフィルタタイプが最大信号を返した角度の識別、及びさらなる解析のために前記特定のフィルタによって提供されるフィルタリングされた画像のみの考察を可能にし得る。別法としては、角度は、フィルタリング結果の回転不変性を達成する方法で、フィルタリングされた画像又は任意の微分データ値を画像から抽出された他の特徴に整列させるために使用され得る。したがってMRフィルタバンクの使用は、テクスチャ特徴メトリクスとして使用され得る配向に対するより高次の共起統計量の計算を可能にする最大応答の角度を記録することができるMRフィルタを含み得る。最大応答の角度を記録することができるMRフィルタバンクを使用することにより、回転した特徴を同じテクスチャにマッピングすることによってより重要なテクスチャを生成することができ、延いてはクラスタ化の間に考慮すべき特徴空間の寸法を低減することができる。

10

【0020】

[0016]いくつかの実施形態では、導出される画像テクスチャ特徴はHaralick特徴であり、Haralick特徴は、画像パッチ中のピクセルとその指定された隣接ピクセルとの間の角関係に基づいて共起行列を計算することによって導出される。いくつかの実施形態では、隣接する核を記述しているデータは、クラスタ割当てのヒストグラムから導出される。いくつかの実施形態では、クラスタ割当てのヒストグラムは、(1)クラスタ中心を得るために、核特徴ベクトルに対してK平均アルゴリズムを適用することであって、例えばクラスタ中心は、クラスタ中心のクラスタに割り当てられたすべての核特徴ベクトルから導出される核特徴ベクトルであってもよい、核特徴ベクトルに対してK平均アルゴリズムを適用すること、(2)対象の特定の核の個々の隣接する核を最も近いクラスタ中心に割り当てること、及び(3)割当てに基づいてクラスタ割当てのヒストグラムを計算することによって導出される。クラスタ割当てのヒストグラムは、特定の画像又は画像領域に存在する核が特定のクラスタに割り当てられる数の多さを示すヒストグラムである。いくつかの実施形態では、(2a)個々の隣接する核の核特徴ベクトルから個々のクラスタの中心までのEuclidean距離を測定し、かつ、(2b)個々の隣接する核を、その中心がその核の核特徴ベクトルに最も近いクラスタに割り当てるための命令が実行される。いくつかの実施形態では、対象の特定の核の個々の隣接する核の最も近いクラスタ中心への割当ては、反復して繰り返され、それにより洗練された(refined)クラスタ中心を提供するためにクラスタ中心が再計算される2つのサブステップ2a、2bで実施され得る。いくつかの実施形態では、組織標本の画像の細胞は、サポート・ベクター・マシンを使用して分類される。いくつかの実施形態では、標本は肺組織標本である。

20

30

【0021】

[0017]本開示の別の態様では、PD-L1バイオマーカの存在に関するIHCアッセイで染色された組織標本の画像内の細胞を分類するコンピュータ実現方法であって、組織標本の画像内の核の特徴から核特徴メトリクスを計算するステップと、組織標本の画像内の対象の核に基づいて文脈情報メトリクスを計算するステップと、核特徴メトリクスと文脈情報メトリクスの組合せを使用して(分類器(classifier)の入力として)、組織標本の画像内の細胞を分類するステップとを含み、細胞は、陽性免疫細胞、陽性腫瘍細胞、陰性免疫細胞及び陰性腫瘍細胞又は他の細胞のうち少なくとも1つとして分類される。いくつかの実施形態では、方法は、細胞内の個々の核を識別するための前景分割マスクを生成するステップをさらに含む。いくつかの実施形態では、核特徴は、形態学特徴、外観特徴及び背景特徴からなるグループから選択される。

40

【0022】

[0018]いくつかの実施形態では、文脈情報メトリクスは、(i)隣接する核を記述しているデータ、及び(ii)対象の核を取り囲んでいる画像テクスチャ特徴のうち少なくとも1つから導出される。いくつかの実施形態では、方法は、対象の特定の核を取り囲んでいる画像パッチを生成するステップをさらに含む、画像テクスチャ特徴は、生成された画像パッチから導出される。いくつかの実施形態では、導出される、対象の核を取り囲ん

50

でいる画像テクスチャ特徴は、テクストンヒストグラム特徴、Garbor特徴、Haralick特徴、強度特徴のヒストグラム、及び傾斜量及び傾斜配向特徴のヒストグラムからなるグループから選択される。いくつかの実施形態では、導出される画像テクスチャ特徴は、テクストンヒストグラム特徴である。いくつかの実施形態では、テクストンヒストグラム特徴は、(1)フィルタ応答画像のリストを得るために、組織標本の画像に対して最大応答フィルタのバンクを適用することによって、例えばフィルタバンクは、少なくとも1つのフィルタタイプの複数のコピーを備えることができ、個々のコピーは異なる配向を有し、個々のフィルタ応答画像は複数のフィルタ応答を含むことができ、それにより「フィルタ応答」は、特定の配向を有するフィルタのうちの1つを画像の特定のサブ領域に対して適用する結果であり、前記サブ領域のサイズは、適用されるフィルタによって決定することができ、フィルタ応答は、前記サブ領域内で識別されるエッジ又はプロブなどの前記サブ領域内のいくつかの特徴を表し、最大応答フィルタの「フィルタ応答」は、それらの配向のみが互いに異なる同じフィルタの複数のコピーを特定のサブ領域に対して適用する際に得られる最大フィルタ応答であり、最大フィルタ応答のみが得られ、かつ、さらに処理されるため、最大応答フィルタバンクを適用することによって得られるフィルタ応答は回転不変であり、フィルタ応答は、例えば画像又は画像サブ領域で識別されるエッジ又はプロブなどの1つ又は複数の特徴を表すか、又は示すことができ、フィルタ応答は、例えば畳込み及び/又は相関フィルタであってもよく、あるいはフィルタが適用された複数の画像ピクセルから導出された、重み付けされたデータ値であってもよい、組織標本の画像に対して最大応答フィルタのバンクを適用すること、(2)フィルタ応答をフィルタ応答画像からテクストンにクラスタ化することによって、例えば「フィルタ応答」は、特定のフィルタを特定の画像領域のピクセル強度値に対して適用することによって導出される一組のピクセルデータ値であってもよく、クラスタ化の間に識別されたクラスタ中心、あるいは前記クラスタ又はクラスタ中心から導出されたデータ値は、テクストンとして使用され得る、フィルタ応答をフィルタ応答画像からテクストンにクラスタ化すること、(3)組織標本の画像中の個々のピクセルをテクストンのうちの1つに割り当てること、及び(4)対象の核を取り囲んでいる画像パッチ中のすべてのピクセルからテクストンヒストグラムを計算するための命令を提供することによって導出される。

10

20

30

40

50

【0023】

[0019]いくつかの実施形態では、導出される画像テクスチャ特徴は強度特徴のヒストグラムであり、強度特徴のヒストグラムは、一次染色チャンネル、対比染色チャンネル、IHC染色チャンネル(例えばPDL1の存在を示すチャンネル)及び輝度チャンネルからなるグループから選択される画像チャンネルから計算される。例えば「輝度画像」は、白色光チャンネルを介して受け取られる光スペクトル中のすべての波長からのデータを含むグレースケール画像である。いくつかの実施形態では、対象の核を取り囲んでいる画像パッチ内の強度ベース特徴を計算するために、異なる画像チャンネルからの信号の差が捕獲される。いくつかの実施形態では、導出される画像テクスチャ特徴はHaralick特徴であり、Haralick特徴は、画像パッチ中のピクセルとその指定された隣接ピクセルとの間の角関係に基づいて共起行列を計算する命令を提供することによって導出される。

【0024】

[0020]いくつかの実施形態では、隣接する核を記述しているデータは、クラスタ割当てのヒストグラムから導出される。いくつかの実施形態では、クラスタ割当てのヒストグラムは、(1)クラスタ中心を得るために、核特徴ベクトルに対してK平均アルゴリズムを適用することによって、例えば画像中で識別された個々の核(又は核プロブ)がそれぞれの核特徴ベクトルによって表され、核特徴ベクトルが前記核の核特徴メトリクス及び文脈情報メトリクスを含み、K平均クラスタ化アルゴリズムの適用は、同様の核特徴ベクトルのクラスタ及びそれぞれのクラスタ中心を反復して識別することを含むことができ、クラスタ中心は、反復して洗練されることが可能である、核特徴ベクトルに対してK平均アルゴリズムを適用すること、(2)対象の特定の核の個々の隣接する核(それぞれの核特徴ベクトルによって表される)を最も近いクラスタ中心に割り当てること、及び(3)割当

てに基づいてクラスタ割当てのヒストグラムを計算することによって導出される。したがって例えば「クラスタ割当てのヒストグラム」は、特定のクラスタに割り当てられる、対象の核の隣接する細胞の数を示す。

【0025】

[0021]いくつかの実施形態では、方法は、(2a)個々の隣接する核の核特徴ベクトルから個々のクラスタの中心までのEuclidean距離を測定するステップと、(2b)個々の隣接する核を、その中心がその核の核特徴ベクトルに最も近いクラスタに割り当てるステップとをさらに含む。いくつかの実施形態では、組織標本の画像の細胞は、サポート・ベクター・マシンを使用して分類される。Euclidean距離は、Euclidean空間における2点間の「通常の」(すなわち直線)距離である。核特徴が非メトリックデータ値に関連している場合、このデータ値は、Euclidean距離の計算を可能にするためにメトリック空間で表され得るメトリック値に変換され得る。追加又は別法として、追加距離関数を使用して、核特徴ベクトルの類似性(したがって個々に表された核の類似性)を決定するための基本である核特徴の距離が計算され得る。異なる核の核特徴メトリクスが類似していればいるほど、また、同様の核特徴の数が多ければ多いほど、2つの核特徴ベクトル及びそれぞれの核が同じクラスタ中心に割り当てられる可能性が高くなる。

10

【0026】

[0022]本開示の別の態様は、本明細書において説明されるコンピュータシステム、及び画像化装置を備える細胞分析器である。

20

[0023]本開示の別の態様では、PD-L1発現のための腫瘍標本をスコアリング(scoring)する方法であって、方法は、(a)コンピュータデバイス、細胞分析器、又は本明細書において説明される方法のうちの任意を使用して、腫瘍標本中の腫瘍細胞及び免疫細胞を識別するステップと、(b)PD-L1を発現する腫瘍細胞及び免疫細胞の数、及び/又は前記細胞中のPD-L1発現の相対強度を決定するステップと、(c)(b)で決定されたPD-L1発現に従って腫瘍を類別するステップとを含む。いくつかの実施形態では、PD-L1の発現は、腫瘍中のPD-L1タンパク質及び/又はPD-L1 mRNAを特異的に検出することによって決定される。いくつかの実施形態では、細胞は、IHCによって検出されたPD-L1タンパク質の少なくとも部分膜染色を細胞が有している場合、PD-L1を発現していると思なされる。いくつかの実施形態では、腫瘍は、修正H-スコア(modified H-score)(MHS)又は修正比率スコア(modified proportion score)(MPS)のうちの一方又は両方に従って類別され、これらはいずれもステップ(b)から計算される。

30

【0027】

[0024]H-スコアは、例えば核免疫反応力の程度を評価する方法である。バイオマーカに依じて、H-スコア計算のための異なる手法が使用され得る。実例による例を与えるために、ステロイド受容体核に対するH-スコアは、0から300の範囲を与える式、 $3 \times$ 強く染色する核の百分率 + $2 \times$ 穏やかに染色する核の百分率 + 弱く染色する核の百分率によって獲得され得る。

【0028】

40

[0025]いくつかの実施形態では、MHSの割当ては、(i)検査されたすべての腫瘍巢中の生存能力のある腫瘍細胞及び染色された単核炎症性細胞のすべてにわたる、染色を有さない細胞、弱い染色を有する細胞(+1)、穏やかな染色を有する細胞(+2)及び強い染色を有する細胞(+3)に対する4つの個別の百分率を予測するステップであって、細胞は、弱い染色百分率、穏やかな染色百分率又は強い染色百分率に含まれる少なくとも部分膜染色を有していなければならず、また、すべての4つの百分率の合計が100に等しいステップと、(ii)予測された百分率を $1 \times$ (弱い染色細胞のパーセント) + $2 \times$ (穏やかな染色細胞のパーセント) + $3 \times$ (強い染色細胞のパーセント)の式に入力し、かつ、式の結果をMHSとして組織切片に割り当てるステップとを含み、MPSの割当ては、検査されたすべての腫瘍巢中の生存能力のある腫瘍細胞及び単核炎症性細胞のすべて

50

にわたる、任意の強度の少なくとも部分膜染色を有する細胞の百分率を予測するステップと、結果として得られた百分率をMPSとして組織切片に割り当てるステップとを含み、MHS及びMPSの両方が割り当てられる場合、その割当ては、いずれかの順序で、又は同時になされ得る。

【0029】

[0026]例えば「なし」、「弱い」、「穏やか」及び「強い」の4つのカテゴリが例えば非重畳強度閾値範囲として定義されることが可能であり、例えば細胞ピクセル領域は、平均強度値が5%未満である場合は「染色しない」細胞と見なされ、平均強度値が5%を超え、かつ25%未満である場合は「弱い染色」を有する細胞と見なされ、平均強度値が25%に等しいか、又はそれより大きく、かつ、75%未満である場合は「穏やかな染色」を有する細胞と見なされ、また、平均強度値が75%に等しいか、又はそれより大きい場合は「強い染色」を有する細胞と見なされ得る。

10

【0030】

[0027]本開示のさらに別の態様では、IHCアッセイにおける抗PD-L1抗体で染色された腫瘍組織切片中のPD-L1発現をスコアリングする方法である。いくつかの実施形態では、これらのスコアリングプロセスのスコアリング結果を使用して、例えばPD-1アンタゴニストに対する被検体の応答を予測するための臨床試験における、また、癌患者を治療する方法における登録基準(enrollment criteria)として、PD-1アンタゴニストで治療するための患者が選択され得る。

【0031】

[0028]本開示の別の態様では、組織標本の画像内の細胞を分類し、及び/又は定量化するためのコンピュータデバイス又はシステムであって、1つ又は複数のプロセッサ及び少なくとも1つのメモリを備え、少なくとも1つのメモリは、1つ又は複数のプロセッサによる実行のための非一時的コンピュータ可読命令であって、その1つ又は複数のプロセッサに、組織標本の画像中の細胞又は細胞核を検出させ、組織標本の画像中の細胞核内の特徴から導出される核特徴メトリクスを計算させ、組織標本の画像内の対象の核の文脈情報メトリクスを計算させ、また、核特徴メトリクス及び文脈情報メトリクスに基づいて、画像内の細胞を分類させるための非一時的コンピュータ可読命令を記憶する。いくつかの実施形態では、組織標本は、PD-L1バイオマーカの存在のために染色されている。いくつかの実施形態では、組織標本は、PD-L1バイオマーカの存在のためにIHCアッセイで染色されており、アッセイは、色素ラベル、化学発光ラベル又は蛍光ラベル、及び対比染色を含む。他の実施形態では、IHCアッセイは、PD-L1バイオマーカの存在を3,3'-ジアミノベンジジンで染色し、また、対比染色としてヘマトキシリンを使用する。いくつかの実施形態では、細胞は、PD-L1陽性免疫細胞、PD-L1陽性腫瘍細胞、PD-L1陰性免疫細胞、PD-L1陰性腫瘍細胞又は他の細胞(PD-L1画像)として分類される。

20

30

【0032】

[0029]本開示の別の態様では、組織標本の画像内の細胞を分類するコンピュータ実現方法であって、組織標本画像中の対象の細胞又は細胞核を検出するステップと、それらの出現及び染色アッセイに対する応答に基づいて、画像内の核から核特徴メトリクスを計算するステップと、画像内の核から文脈情報メトリクスを計算するステップと、核特徴メトリクスと文脈情報メトリクスの組合せを使用して画像内の核を分類するステップとを含む。「それらの出現及び染色アッセイに対する応答に基づいて」という表現は、前記核特徴に関する情報は、メトリック特徴値を計算するための入力として使用されることを暗に示し得る。いくつかの実施形態では、核特徴メトリクス及び文脈情報メトリクスは、特徴抽出モジュールを使用して計算される。いくつかの実施形態では、核は分類モジュールを使用して分類され、分類モジュールは、特徴抽出モジュールによって計算されたメトリクスを使用する。いくつかの実施形態では、組織標本は、PD-L1バイオマーカの存在のために染色されている。いくつかの実施形態では、組織標本は、PD-L1バイオマーカの存在のためにIHCアッセイで染色されており、アッセイは、色素ラベル、化学発光ラベル

40

50

又は蛍光ラベル、及び対比染色を含む。他の実施形態では、IHCアッセイは、PD-L1バイオマーカの存在を3,3'-ジアミノベンジジンで染色し、また、対比染色としてヘマトキシリンを使用する。いくつかの実施形態では、細胞は、PD-L1陽性免疫細胞、PD-L1陽性腫瘍細胞、PD-L1陰性免疫細胞、PD-L1陰性腫瘍細胞又は他の細胞(PD-L1画像)として分類される。

【0033】

[0030]本出願人は、現在開示されている方法は、従来技術による方法と比較すると優れた結果を提供することを示した。本出願人は、一組の従来核特徴と組み合わせると、PD-L1染色画像中の核を分類するためには文脈情報が有用であることを示すために、広範囲に渡る実験評価を実施した。実際、文脈情報メトリクスを核メトリクスと共に使用すると、細胞が核メトリック特徴を使用して排他的に分類される従来技術と比較して核分類精度が改善される。また、本出願人は、核特徴メトリクスと組み合わせると、従来技術と比較してかなり優れた分類結果を提供する文脈情報メトリクスを導出するいくつかの方法を立証した。したがって本出願人は、PD-L1染色組織画像中の核の自動分類を可能にする方法であって、それ自体の出現(核特徴メトリクス)及びその局所組織文脈における細胞の出現(文脈情報メトリクス)に基づいて個々の核を識別する方法を開発した。PD-L1染色組織画像に対する精度は、PD-L1が細胞膜で同じく発現されるため、とりわけ高いことが分かっており、したがって核を取り囲んでいる領域に関する強度情報は、PD-L1陽性細胞及びPD-L1陰性細胞への細胞の正確な分類を促進し得る。

10

【0034】

[0031]特許ファイル又は出願ファイルは、カラーで実行された少なくとも1枚の図面を含む。カラー図面を有する本特許又は特許出願公開のコピーは、要求及び必要な手数料の支払いに応じて事務所に提供される。

20

【図面の簡単な説明】

【0035】

【図1A】[0032]開示される技術の実施形態による、標本を分析するためのコンピュータベースシステムを示す図である。

【図1B】[0033]コンピュータベースシステム及び方法で使用されるモジュールの概説を示すフローチャートである。

【図1C】[0034]核メトリクス及び文脈情報メトリクスに基づいて細胞を分類するステップの概説を示すフローチャートである。

30

【図2A】[0035]PD-L1で染色された肺癌標本からの視野を示す図である。

【図2B】PD-L1で染色された肺癌標本からの視野を示す図である。

【図3】[0036]図3Aは、四角の点で示された対象の異なる核の近辺領域を示す図である。図3Bは、四角の点で示された対象の異なる核の隣接する領域を示す図である。

【図4A】[0037]画像パッチの例であって、円が重畳するパッチがPD-L1に対して陽性の領域に対してのみ生成される画像パッチの例を示す図である。

【図4B】画像パッチの例であって、円が重畳するパッチがPD-L1に対して陽性の領域に対してのみ生成される画像パッチの例を示す図である。

【図5A】[0038]パッチ分類結果であって、赤が主として免疫細胞(IC)からの染色を示し、緑が主として腫瘍細胞(TC)からの染色を示し、また、黄が主として非標的からの染色を示すパッチ分類結果を示す図である。

40

【図5B】[0039]パッチ分類結果であって、円が主として免疫細胞(IC)からの染色を示し、正方形が主として腫瘍細胞(TC)からの染色を示し、また、黄が主として非標的からの染色を示すパッチ分類結果を示す図である。

【図6A】[0040]文脈「語の多重集合(Bag of Words)」方法を使用してテキスト特徴を計算するステップ方式による方法を示す図である。

【図6B】文脈「語の多重集合」方法を使用してテキスト特徴を計算するステップ方式による方法を示す図である。

【図6C】[0041]文脈-テキスト方法を使用して文脈情報を導出するステップを示すフ

50

ローチャートである。

【図 6 D】[0042]文脈 - テクストン方法を使用して文脈情報を導出するステップを示すフローチャートである。

【図 7 A】[0043] P D - L 1 染色肺組織画像中の核の 5 つの等級を示す図である。

【図 7 B】 P D - L 1 染色肺組織画像中の核の 5 つの等級を示す図である。

【図 8】[0044]図 8 a は、パラメータ選択に応じた文脈「語の多重集合」の精度を示す図である。図 8 b は、パラメータ選択に応じた文脈 - C R F 方法の精度を示す図である。

【図 9 - 1】[0045]分類結果の例を示す図である。

【図 9 - 2】分類結果の例を示す図である。

【図 1 0】分類結果の例を示す図である。

10

【図 1 1 A】[0046]本明細書において説明される、提案されるスコアリング方法を使用して得られる例示的結果を示す図である。

【図 1 1 B】本明細書において説明される、提案されるスコアリング方法を使用して得られる例示的結果を示す図である。

【図 1 1 C】本明細書において説明される、提案されるスコアリング方法を使用して得られる例示的結果を示す図である。

【図 1 1 D】本明細書において説明される、提案されるスコアリング方法を使用して得られる例示的結果を示す図である。

【発明を実施するための形態】

【 0 0 3 6】

20

[0047]本明細書において使用されているように、単数形の表現には、複数ではないことを文脈が明確に示していない限り、複数指示を含む。同様に、「又は」という語には、「又は」ではないことを文脈が明確に示していない限り、「及び」を含むことが意図されている。

【 0 0 3 7】

[0048]「備えている」、「含んでいる」、「有している」、等々という用語は交換可能に使用されており、同じ意味を有している。同様に、「備える」、「含む」、「有する」、等々も交換可能に使用されており、同じ意味を有している。詳細には、用語の各々は、「備えている」の共通の米国特許法定義と無矛盾に定義されており、したがって「少なくとも以下の」を意味する公開用語として解釈され、また、追加特徴、制限、態様、等々を排他しないものとして同じく解釈される。したがって例えば「構成要素 a、b 及び c を有しているデバイス」は、デバイスは少なくとも構成要素 a、b 及び c を含むことを意味している。同様に、「ステップ a、b 及び c を含んでいる方法」という語句は、方法は少なくともステップ a、b 及び c を含むことを意味している。さらに、ステップ及びプロセスは、本明細書においては特定の順序で概説されることがあるが、当業者は、ステップ及びプロセスの順序は変化し得ることを認識するであろう。

30

【 0 0 3 8】

[0049]本開示は、組織画像中の細胞及び細胞核、例えば P D - L 1 バイオマーカの存在のために染色された組織中の細胞及び細胞核の分類を補助するための文脈情報メトリクスの導出及び後続する使用を対象としている。本開示は、(1)核特徴メトリクスの頂部への条件付き確率場 (c o n d i t i o n a l r a n d o m f i e l d) (「 C R F 」) モデルの適用 (「 文脈 - C R F 方法」)、(2)核を中心とする画像パッチ中の追加テキスト特徴の抽出 (及びテキスト特徴と核特徴メトリクスの結合) (「 文脈 - テクストン方法」)、(3)核を中心とする画像パッチ中のテキストヒストグラムの計算 (及びテキストヒストグラム特徴と核特徴メトリクスの結合) (「 文脈 - テクストン方法」)、及び (4)隣接する核の外観を捕獲するための「語の多重集合」モデルの使用 (「 語の多重集合」モデルは、訓練核特徴を多数のクラスタ中心にクラスタ化し、かつ、対象の個々の核の隣接する核をこれらのクラスタに割り当てる) (「 文脈 - B o W 方法」) を含む、文脈情報メトリクスを導出する 4 つの異なる方法を説明する。

40

【 0 0 3 9】

50

[0050]何らかの特定の理論に縛られるつもりはないが、対象の核の文脈情報（すなわち対象の核を中心とする領域の隣接する核又は画像テクスチャを記述している情報）は、細胞核の分類には有用であるとされている。一例として、細胞核は、その近辺の細胞及び他の生物学的構造を考慮することにより、より確信をもってラベル付けされ得る（例えば腫瘍細胞中の核として）とされている。実際、本出願人は、文脈情報メトリクスを核特徴メトリクスと共に考察することにより、核特徴のみに基づく細胞の分類と比較すると、分類精度が体系的に改善されることを示した（例えば本明細書における例1及び表1を参照されたい）。

【0040】

[0051]本明細書において開示される技術の少なくともいくつかの実施形態は、PD-L1バイオマーカのための免疫組織化学（IHC）染色で前処理された組織標本から捕獲されたデジタル画像を解析するためのコンピュータシステム及び方法に関する。いくつかの実施形態では、本開示において評価中の組織標本は、PD-L1バイオマーカが存在しているところでDABで染色されており、また、そこではヘマトキシリンが対比染色として使用された（例えば、PD-L1で染色された肺癌標本からの視野を提供する図2A及び2Bを参照されたい）。いくつかの実施形態では、PD-L1（SP263）アッセイが最新染色器機のVENTANA Benchmarkシリーズと共に使用される。本明細書において説明される例示的实施形態は、PD-L1バイオマーカのためのIHC染色の適用を開示しているが、技術は、組織標本中の異なるタイプの細胞及び/又は画像領域を検出するために、他のプローブ及び/又はアッセイで処理された組織標本の画像を解析するために使用され得ることは認識されよう。

10

20

【0041】

[0052]IHCは、組織の標本を対象のペプチドに特異的に結合する実体と接触させることによって組織中の対象のペプチドを検出するための技法である（「特異的結合実体」）。いくつかの例では、特異的結合実体は、抗体、抗体の抗原結合断片、単鎖抗体、Affimers（商標）、及びDARPIN、アフーマ、及びDARPINからなるグループから選択される。本明細書において使用されているように、「ペプチド」という用語は、オリゴペプチド、ポリペプチド及びタンパク質を含む、ペプチド結合によって互いに連結されたアミノ酸の鎖を含む任意の分子を包含するものとする。例えばいくつかの実施形態では、IHCは、その対応する抗原に対しそれを検出するための特異的結合実体と、染料を処理して視覚化用の染色を生成する酵素ステップとに依拠している。他の方法及びアッセイが当業者によって代替され得る。観察された組織中に特異的ペプチドが存在すること、又は存在しないことを明らかにすることにより、IHCは、腫瘍の起点に存在する細胞タイプの決定を支援する。IHCによって明らかにされたタンパク質に応じて、検出された癌のタイプに対する特異的治療処置が同じく選択また適合され得る。

30

【0042】

[0053]本明細書において使用されているように、「抗体」という用語は、所望の生物学的活性又は結合活性を示す任意の形態の抗体を意味している。したがって「抗体」という用語は、最も広範囲の意味で使用され、また、それらに限定されないが、とりわけ、単クローン抗体（全長（full length）単クローン抗体を含む）、ポリクローン抗体、多重特異性抗体（例えば2重特異性抗体）、ヒト化抗体、完全ヒト抗体、キメラ抗体及びカメライズ化（camelized）単領域抗体を包含する。

40

【0043】

[0054]本明細書において使用されているように、特に指示されていない限り、「抗体断片」又は「抗原結合断片」は、抗体の抗原結合断片、すなわち完全長抗体によって拘束された抗原に特異的に結合する能力を保持する抗体断片、例えば1つ又は複数のCDR領域を保持する断片を意味している。抗体結合断片の例には、それらに限定されないが、Fab断片、Fab'断片、F(ab')₂断片及びFv断片、二重特異性抗体、線形抗体、単鎖抗体分子、例えばsc-Fv、ナノボディ及び抗体断片から形成された多重特異性抗体がある。

50

【0044】

[0055] *Affimers* (商標) は、抗体の特異性及び結合親和力を模倣する改変タンパク質であるが、はるかに小さく、また、約14kDaの分子量を有している。それらは、高度に安定であり、また、特異的な標的タンパク質のための高親和力結合表面を提供するペプチドループを示すように改変されるとされている。

【0045】

[0056] *DARPs* (*designed ankyrin repeat proteins*) は、典型的には高度に特異的な、高親和力標的タンパク質結合を示す、遺伝学的に改変された抗体模倣性タンパク質である。それらは天然アンキリンタンパク質から導出され、また、3つのタンパク質の少なくとも3つ、通常は4つ又は5つの反復モチーフからなっている。*Programmed cell death 1 ligand 1* (プログラム細胞死1リガンド1) (*PD-L1*) は、細胞の調節及び液素性免疫応答に含まれる1型膜貫通タンパク質である。*PD-L1* は、主として、抗原提示細胞、胎盤、及び黒色腫、びまん性大細胞型B細胞性リンパ腫、及び肺、結腸、直腸、腎臓並びに他の臓器の癌腫などのいくつかの腫瘍の中で発現される。正常組織及び腫瘍組織における免疫組織化学(IHC)用途のためには、2つのウサギ単クローン抗ヒト*PD-L1*抗体、すなわち*SP142* (California州PleasantonのSpring Bioscience) 及びクローン*E1L3N* (Massachusetts州DanversのCell Signaling Technology) が商用的に入手可能である。参照によりその開示全体が本明細書に組み込まれているWO2014165422 (PCT/US2014/032305) に、腫瘍組織中の*PD-L1*を検出するための他のIHCアッセイが開示されている。

10

20

【0046】

[0057] 本明細書において使用されているように、「ペプチド」という用語は、オリゴペプチド、ポリペプチド及びタンパク質を含む、ペプチド結合によって互いに連結されたアミノ酸の鎖を含む任意の分子を包含するものとする。

【0047】

[0058] 「標本」又は「組織標本」は、それらに限定されないが、とりわけバクテリア、酵母菌、原生動物及びアメーバなどの単細胞生体、組織化学的分析又は細胞化学的分析のために適している、分析されるべき細胞及び/又は組織の形態学的特性を保持している標本などの多細胞生体(健康な、又は外見上健康なヒト被検体からの標本、又は癌などの診断又は調査されるべき状態又は疾病による影響を受けたヒト患者からの標本を含む植物又は動物など)を含む、任意の生きた生体から得られた、任意の生きた生体によって排泄された、又は任意の生きた生体によって分泌された任意の固体標本又は流体標本であってもよい。例えば生物学的標本は、例えば血液、血漿、血清、尿、胆汁、腹水、唾液、脳脊髄液、房水又は硝子体液あるいは任意の身体分泌物、漏出液、滲出液(例えば膿瘍又は感染あるいは炎症の任意の他の部位から得られる流体)、又は関節(例えば正常関節又は疾病による影響を受けた関節)から得られる流体から得られる生物学的流体であってもよい。また、生物学的標本は、任意の臓器又は組織(腫瘍生検などの生検標本又は剖検標本を含む)から得られる標本であってもよく、あるいは細胞(一次細胞であれ、あるいは培養された細胞であれ)又は任意の細胞、組織又は臓器によって条件付けられた媒質を含むことも可能である。いくつかの例では、生物学的標本は核抽出物である。特定の例では、標本は、開示された細胞小球切片標本のうちの1つなどの品質管理標本である。他の例では、標本は試験標本である。例えば試験標本は、被検体から得られた生物学的標本から準備される細胞、組織又は細胞小球切片である。一例では、被検体は、危険な状態にある被検体又は獲得された被検体である。標本は、当業者に知られている任意の方法を使用して準備され得る。標本は、日常的なスクリーニングのために被検体から獲得されるか、又は遺伝的異常、感染又は異常増殖などの障害を有していることが疑われる被検体から獲得され得る。開示される方法の説明した実施形態は、「正常」標本と呼ばれる、遺伝的異常、疾病、障害、等々を有していない標本にも同じく適用され得る。標本は、複数の標的を含むこ

30

40

50

とができる。

【0048】

[0059] 標本を分析するためのコンピュータベース標本分析器は、図1Aに示されている。当業者は、他のコンピュータシステムが利用され得ること、また、本明細書において説明されているコンピュータシステムは、追加構成要素、例えば分析器、スキャナ、等々に通信結合され得ることを認識するであろう。これらの追加構成要素及び利用され得る様々なコンピュータのうちのいくつかは、本明細書においてさらに説明される。一般に、画像化装置12は、それには限定されないが、1つ又は複数の画像捕獲デバイスを含むことができる。画像捕獲デバイスは、それらに限定されないが、カメラ（例えばアナログカメラ、デジタルカメラ、等々）、光学（例えば1つ又は複数のレンズ、センサ焦点レンズグループ、顕微鏡対物レンズ、等々）、画像化センサ（例えば電荷結合デバイス（CCD）、相補性金属酸化膜半導体（CMOS）画像センサ、等々）、写真フィルム、等々を含むことができる。デジタル実施形態では、画像捕獲デバイスは、共同してオンザフライ集束を証明する複数のレンズを含むことができる。CCDセンサは、標本のデジタル画像を捕獲することができる。デジタル画像を生成する1つの方法は、標本の少なくとも一部を含む顕微鏡スライドの領域を含む走査領域を決定するステップを含む。走査領域は、複数の「スナップショット」に分割され得る。画像は、個々の「スナップショット」を結合することによって生成され得る。いくつかの実施形態では、画像化装置12は標本全体の高解像度画像を生成し、このような装置のための一例は、Ventana Medical Systems社（Arizona州Tucson）のVENTANA iScan HTスライドスキャナである。

10

20

【0049】

[0060] コンピュータデバイス14は、デスクトップコンピュータ、ラップトップコンピュータ、タブレット、等々を含むことができ、また、デジタル電子回路機構、ファームウェア、ハードウェア、メモリ、コンピュータ記憶媒体、コンピュータプログラム、プロセッサ（プログラムされたプロセッサを含む）、等々を含むことができる。図1の図解されている計算システム14は、スクリーン16及びタワー18を有するデスクトップコンピュータである。タワー18は、デジタル画像を2進形態で記憶することができる。画像は、ピクセルの行列に同じく分割され得る。ピクセルは、ビット深度によって画定される1つ又は複数のビットのデジタル値を含むことができる。ネットワーク20又は直接接続は、画像化装置12及びコンピュータシステム14を相互接続する。コンピュータシステムは、一連のコンピュータ実行可能命令を使用してプログラムされる1つ又は複数のプロセッサを含み、命令はメモリに記憶される。

30

【0050】

[0061] 図1B及び1Cを参照すると、命令が実行されると、コンピュータシステムのプロセッサのうちの少なくとも1つが、組織標本のデジタル画像などの入力100及び110を受け取る。いくつかの実施形態では、デジタル画像は、PD-L1バイオマーカの存在のために染色された組織標本のデジタル画像である。必要な入力提供されると、特徴抽出モジュール101が実行されて、核特徴メトリクス111及び文脈情報メトリクス112を導出する。核特徴メトリクス及び文脈情報メトリクスは、導出されると、分類モジュール102に提供されて細胞核を分類113し、かつ、出力103をユーザに提供する。出力は、ディスプレイ、メモリ、又は当分野における適切な任意の他の手段に提供され得る。

40

【0051】

[0062] 特徴抽出モジュール

[0063] 一般に、特徴抽出モジュールは、画像データを受け取り、入力された画像データに基づいて特定のメトリクスを導出し（例えばメトリクスのベクトル）、また、これらのメトリクスを分類モジュールに出力する。本明細書においてさらに詳細に説明されるように、特徴抽出モジュールは、入力画像内の細胞の核特徴から核特徴メトリクスを導出し（図1C、ステップ111）、また、本明細書において説明される4つの方法のうちの1つ

50

を利用して、入力画像から対象の核 (NoI) の文脈情報を同じく導出する (図 1 C、ステップ 112)。次に、細胞核が分類され得るよう、様々なメトリクスが分類モジュールに供給される。分類のために核特徴メトリクスのみが使用される従来技術と比較すると、核特徴メトリクスと文脈情報メトリクスの組合せは、優れた分類精度を提供する (例 1 を参照されたい)。

【0052】

[0064]「特徴メトリクス」は、例えば 1 つ又は複数の特徴から導出されたデータ値である。例えば特徴メトリクスは、特定の特徴の定量特性を示す数値データ値、ヒストグラム、分布、等々であってもよい。

【0053】

[0066]特徴抽出モジュールは、1 つ又は複数の画像チャンネルに基づいて核メトリクス及び文脈情報メトリクスを計算することができ、また、画像チャンネルは、当業者に知られている任意の手段によって導出され得る。例えばいくつかの実施形態では、画像チャンネルは、色逆たたみこみ (deconvolution) 又は非混合を介して導出され得る (すなわち色逆たたみこみスキームを使用して、個々の画像が RGB 色空間から、利用される染色のスペクトル特性によってモデル化される新しい空間に変換される)。非混合は、例えば Zimmermann 「Spectral Imaging and Linear Unmixing in Light Microscopy (光学顕微鏡法におけるスペクトル画像化及び線形非混合)」Adv Biochem Engin/Biotechnol (2005) 95:245-265、及び C. L. Lawson and R. J. Hanson、「Solving least squares Problems (最小 2 乗問題の解決)」, Prentice Hall, 1974、Chapter 23、p. 161 に記載されており、これらの開示は、参照によりその全体が本明細書に組み込まれている。実際、入力画像は、組織標本内の染色及びラベルのための局所染色強度を表す画像チャンネルを提供するために非混合され得る。異なるチャンネルは、組織画像中の異なる組織構造を強調し、これらの個々のチャンネルに目を向けることにより、細胞の核及び膜染色の出現が、これらの細胞の局所近辺における複数の細胞、これらの細胞の外観及び染色パターンと共に評価され得る。

【0054】

[0067]いくつかの実施形態では、選択される画像チャンネルは、利用されるアッセイに基づかれる (例えば一次染色チャンネル、IHC 染色チャンネル又は対比染色チャンネル、PD-L1 を示す染色チャンネル)。当然、異なるアッセイのために異なるチャンネルが使用され得る。PD-L1 バイオマーカー染色の文脈においては、いくつかの実施形態では、様々な特徴が導出され得るチャンネルは、ヘマトキシリンチャンネル、輝度チャンネル及び DAB チャンネルであり、この場合も使用される特定のアッセイで決まる。いくつかの実施形態では、組織標本の画像は、スライド上の染色及びラベルのための局所染色強度を表す画像チャンネルの提供などのために非混合されることが可能であり、一例は、ヘマトキシリンチャンネル及び DAB チャンネルである。いくつかの実施形態では、輝度チャンネルは、 L^*a^*b 色空間の L 成分に基づかれる (L^*a^*b 色空間では、「L」チャンネルはピクセルの輝度を表し、「A」チャンネルは、ピクセルの赤成分及び緑成分を反映し、また、「B」チャンネルは、ピクセルの青成分及び黄成分を表す)。

【0055】

[0068]核特徴

[0069]「核特徴」は、例えば細胞核の特徴又は前記核を含む細胞の特徴であってもよく、核又は細胞は組織標本の画像中で識別されている。

【0056】

[0070]核特徴メトリクスは、最初に、細胞又は細胞核毎に、それらの視覚特性及び記述子 (descriptors)、例えばそれぞれ以下で説明される形態学特徴、外観特徴及び背景特徴に基づいて計算される。

【0057】

10

20

30

40

50

[0071]本明細書において使用されている「形態学特徴」は、例えば核又は核を含む細胞の形状及び寸法を示す特徴である。例えば形態学特徴は、核プロブに含まれている、又は核プロブを取り囲んでいるピクセルに対して様々な画像解析アルゴリズムを適用することによって計算され得る。

【0058】

[0072]本明細書において使用されている「外観特徴」は、例えば、核を識別するために使用される核プロブに含まれている、又は核プロブを取り囲んでいるピクセルのピクセル強度値を比較することによって特定の核に対して計算された特徴であり、比較されるピクセル強度は、異なる画像チャンネルから導出される（例えば背景チャンネル、PD-L1バイオマーカの染色のためのチャンネル、等々）。

10

【0059】

[0073]「背景特徴」は、例えば背景特徴が画像から抽出された核を含む細胞の細胞形質特徴及び細胞膜特徴に存在する外観及び/又は染色を示す特徴である。背景特徴及び対応するメトリクスは、例えば、核を表している核プロブを識別し、識別されたセットの細胞のすぐ隣のピクセル領域（例えば核プロブ境界の周りの20個のピクセルのリボン - 約9ミクロン - の厚さ）を解析し、したがってこの核を有する細胞の細胞形質及び膜における外観及び染色の存在を細胞のすぐ隣の領域と共に捕獲することにより、デジタル画像中に描写されている核及び対応する細胞に対して計算され得る。

【0060】

[0074]当然、当業者に知られている他の特徴が考慮され、また、核特徴メトリクスを計算するための基本として使用され得る。核特徴は、組織標本中の核、細胞、等々を含む生物学的対象の発生、密度及び他の同様の特性を捕獲することができ、また、これらの対象を検出することにより、分類に使用するためのメトリクスを導出することができる（従来技術の場合と同様に単独で、又はここでは文脈情報メトリクスとの組合せで）。これらの核特徴から計算される様々な核特徴メトリクスは、核特徴メトリクスのベクトルとして提供され、また、文脈情報メトリクスと共に分類モジュールに供給される。

20

【0061】

[0075]いくつかの実施形態では、入力として受け取られたデジタル画像は、核中心の検出及び/又は核の分割などのために予備処理される。例えば、当業者に広く知られている技法を使用した放射対称票決 (radial-symmetry voting) に基づいて、核中心を検出する命令が提供され得る（参照によりその開示全体が本明細書に組み込まれている、Parvin、Bahramらの「Iterative voting for inference of structural saliency and characterization of subcellular events. (構造的特徴の推定に対する反復票決及び細胞レベル下の事象の特性化)」Image Processing、IEEE Transactions on 16.3 (2007): 615 - 623を参照されたい)。

30

【0062】

[0076]核は、次に、引き続いて、核毎に個別に計算された閾値を使用して分割される。例えば、核領域のピクセル強度は変化するとされているため、Otsuの方法を使用して核の周りの領域が分割され得る。当業者には認識されるように、Otsuの方法を使用して、等級内変化を最小化することによって最適閾値が決定され、これは当業者に知られている。より詳細には、Otsuの方法を使用して、クラスタ化ベース画像閾値化、又は2進画像に対するグレーレベル画像の低減が自動的に実施される。アルゴリズムは、画像は、2モードヒストグラムに従って2つの等級のピクセルを含むことを仮定している（前景ピクセル及び背景ピクセル）。アルゴリズムは、次に、2つの等級の結合散布量 (combined spread) (クラス内分散) が最小になるか、又は等価になるように (ペアワイズ2乗距離 (pairwise squared distances) の合計は一定であるため) それらを分割する最適閾値を計算し、したがってそれらの等級間変化は最大である。

40

50

【0063】

[0077]次に、核特徴メトリクスが、組織標本中の細胞の核から抽出された特徴から導出される。核特徴メトリクスの計算は当分野でよく知られており、知られている任意の核特徴が本開示の文脈に使用され得る。計算され得るメトリクスの非制限例は、以下を含む。

【0064】

[0078]形態学特徴から導出されるメトリクス：（面積、短軸長及び長軸長、周囲、半径及び固体性）[面積＝核領域におけるピクセルの総数；Minor/Major Axis Length：領域と同じ正規化第2中心モーメントを有する楕円の短軸/長軸の長さ（ピクセル内の）を規定するスカラー；周囲：核領域の境界上のピクセルの数；半径：核の中心から核の境界ピクセルまでの平均距離；固体性：同じく領域に存在している凸閉包中のピクセルの割合を規定するスカラー（Area/Convex Areaとして計算される）]

[0079]外観特徴から導出されるメトリクス：異なる画像チャンネルから計算されたピクセル強度及び傾斜量の百分位数値（例えば10番目、50番目及び95番目の百分位数値）。例えば、最初に対象の核を表している核プロブ内の画像チャンネル（例えばHTX、DAB、輝度の3つのチャンネル）の複数のICの各々のピクセル値のX - 百分位数値（X = 10、50、95）の数Pが識別される。したがってP × IC個の特徴値が前記核プロブによって表される対象の特定の核に対して計算され、Pは、調査された百分位数値の数を示し、また、ICは、特徴値が計算される画像チャンネルの数を表す。この例では、「核特徴メトリクス」として作用する3 × 3個の特徴値がNoIに対して計算される。

【0065】

[0080]さらに、NoIを表している核プロブを取り囲んでいるピクセルのリボン内の同じ数及びタイプの画像チャンネルのピクセル値の同じX - 百分位数値が識別され、「文脈情報メトリクス」として作用するP × IC個の特徴値（ここでは3 × 3 = 9個の特徴値）がNoIに対して計算される。最後に、P × IC個（ここでは9個）の核特徴メトリクス及びP × IC個（ここでは9個）の文脈情報メトリクスが、例えば2 × P × IC個（ここでは9 + 9 = 18個）の特徴メトリクスを含む結合特徴ベクトルに結合され、これらは対象の核の外観特徴として使用される。前記メトリクスは核領域の特性（例えば暗褐色の核、暗い青色の核、明るい青色の核、等々）を記述することができ、また、例えば膜染色が明るい茶色、暗褐色又は全く染色がない、等々の場合は、核の周りの膜領域（リボン領域）を記述するため、外観特徴メトリクスの計算は有利であり得る。

【0066】

[0081]PD - L1バイオマーカのためのIHC染色の文脈では、画像チャンネルは、分割された核中のヘマトキシリン、輝度（LAB色空間）及びDABである。外観特徴を決定するために必要な局所画像強度は、参照によりその開示全体が組み込まれている、Rufrok, Arnout C及びDennis A. Johnston「Quantification of histochemical staining by color deconvolution.（色逆たたみこみによる組織化学的染色の定量化）」Analytical and quantitative cytology and histology/the International Academy of Cytology [及び] American Society of Cytology 23.4（2001）：291 - 299によって開示されている方法を含む、当業者に知られている任意の方法によって計算され得る。

【0067】

[0082]背景特徴から導出されるメトリクス：これらのメトリクスは核外観特徴に類似しているが、個々の核境界の周りの20個のピクセルのリボン（約9ミクロン）の厚さで計算され、したがってこの核を有する細胞の細胞形質及び膜に存在する外観及び染色を細胞のすぐ隣の領域と共に捕獲する。このサイズは、核を区別するための有用な情報を提供するために使用され得る核の周りの十分な量の背景組織領域を捕獲するように選択される。これらの特徴は、参照によりその開示全体が本明細書に組み込まれている、J. Kon

10

20

30

40

50

gらの「A comprehensive framework for classification of nuclei in digital microscopy imaging: An application to diffuse gliomas, (デジタル顕微鏡画像化における核の分類のための包括的フレームワーク:びまん性グリオーマに対する適用)」in ISBI、2011年、頁2128-2131によって開示されている特徴に類似している。これらの特徴を使用して、周囲の組織が間質であるか、又は上皮であるかが決定され得るとされている。何らかの特定の理論に縛られるつもりはないが、PD-L1バイオマーカは、主として細胞膜を染色し、かつ、核境界の外側に沿って茶色のリボン、茶色のスポット又は両方を生成するため、これらの背景特徴は、膜染色パターンを同じく捕獲するとされている。

10

【0068】

[0083]文脈情報

[0084]核特徴メトリクスが計算されると、文脈情報メトリクスが対象の核(NoI)毎に導出される。NoIの文脈情報、すなわちNoIを中心とする領域における隣接する核又は画像テクスチャを記述している情報は、そのラベルを予測するための有用な証拠を提供するとされている。例えば核は、隣接する核が同じく腫瘍組織に属していることを条件として、腫瘍としてより確信的にラベル付けされ得るとされている。同様に、核は、核を取り囲んでいる領域のテクスチャパターンが腫瘍領域で見出されたテクスチャパターンに類似していることを条件として、腫瘍としてより確信的にラベル付けされ得るとされている。これは、異なるNOIの近辺領域を示す図3A及び3Bに示されている(長方形の点で示されている)。赤(中間の幅の線)、緑(細い幅の線)及び青(太い幅の線)の長方形は、それぞれPD-L1陽性免疫細胞核、PD-L1陽性腫瘍細胞核及びPD-L1陰性腫瘍細胞核の近辺領域を示している。図3A及び3Bに提供されている少なくとも例から、示されている近辺領域におけるテクスチャ情報及び核情報は互いに異なっており、核特徴メトリクスと共に考察されると、優れた分類結果の獲得を可能にするのは、このタイプの文脈情報であることは明らかである。

20

【0069】

[0085]文脈情報は、当業者に知られている任意の方法によって導出され得る。いくつかの実施形態では、文脈情報は、(1)文脈-テクスチャ方法、(2)文脈-テクストン方法、(3)文脈-CRF方法、及び(4)文脈-語の多重集合(BoW)方法のうち少なくとも1つから導出される。いくつかの実施形態では、文脈情報には、個々の核の近辺におけるすべての画像ピクセルからの追加特徴の計算が必要である(例えば本明細書における文脈-テクスチャ方法及び文脈-テクストン方法を参照されたい)。他の実施形態では、文脈情報メトリクスは事前計算され、隣接する特定の核の文脈情報を評価するための追加特徴を計算する必要はない。その代わりに、事前計算済み核特徴及びラベルが隣接する核から利用される。例えば核特徴及び文脈情報メトリクスの事前計算は、画像内又は前記画像中の対象の領域内で識別されたすべての核に対して一度実施され得る。

30

【0070】

[0086]文脈-テクスチャ方法

[0087]文脈-テクスチャ方法は、個々のNoIを中心とする画像パッチから一組のテクスチャ特徴を計算するために使用される。より詳細には、文脈-テクスチャ方法は、個々のNoIの周りの領域におけるテクスチャパターンの捕獲を可能にし、この情報を使用して、NoIが存在し得る組織の局所タイプの識別が補助される(例えば任意のNoIの周りの領域は、固化した腫瘍、リンパ球(免疫細胞)の集合体、間質及び/又は総合染色応答を含み得る)。例えば間質は、線維様テクスチャによって特性化され、一方、可変サイズの複数の「プロブ」の存在は腫瘍領域の特性である。線維様テクスチャ又はプロブを取り囲んでいる領域の画像パッチ中のテクスチャ特徴を計算することにより、情報は、領域における任意の細胞又は細胞核の、間質に属するものとしての分類を補助することができ、これは腫瘍組織とは対照的でありまた、あるいはその逆も真である。

40

【0071】

50

[0088] PD-L1染色組織の文脈では、PD-L1バイオマーカを発現しないリンパ球（「陰性リンパ球」）を有する領域は、小さい青のプロブによって特性化され、PD-L1バイオマーカを発現するリンパ球（「陽性リンパ球」）を有する領域は、小さい青のプロブ及び茶色のプロブによって特性化され、PD-L1バイオマーカを支配的に発現する細胞（「陽性腫瘍細胞」）を有する腫瘍領域は、大きい青のプロブ及び茶色のリングによって特性化され、また、PD-L1バイオマーカを発現しない細胞（「陰性腫瘍細胞」）の腫瘍領域は、大きい青のプロブによってのみ特性化される。

【0072】

[0089]一般に、また、図6Cを参照すると、文脈-テクスチャ方法は、個々のNoIを中心とする画像パッチを捕獲することによって実施される（ステップ320）。いくつかの実施形態では、核に関する豊富な文脈情報を提供する適度に大きい組織領域を捕獲するサイズ $S \times S$ を有するパッチサイズが選択される。他の実施形態では、パッチサイズは、任意の $S \times S$ 寸法における約50個のピクセルと約200個のピクセルの間の範囲である。さらに他の実施形態では、約150個のピクセルのパッチサイズ（約70ミクロン）が使用される。

10

【0073】

[0090]画像パッチが捕獲されると（ステップ320）、テクスチャ特徴が個々のパッチ内で計算される（ステップ321）。いくつかの実施形態では、計算されるテクスチャ特徴は、強度のヒストグラム、傾斜量及び傾斜配向のヒストグラム、Gabor特徴及びHaralick特徴などの特徴を含み、これらの各々は、本明細書においてさらに説明される。計算されるテクスチャ特徴は、文脈情報メトリクスとして分類モジュールに出力される（ステップ322）。

20

【0074】

[0091]いくつかの実施形態では、画像はパッチの格子に仕切られる。パッチの分配は多くの異なる方法で実現され得る。例えばパッチは、円形又は長方形であってもよい。いくつかの実施形態ではパッチが重畳する。他の実施形態では、パッチの中心間の距離（半径）は、同じであっても、あるいは異なってもよい。例えばパッチの格子の個々のパッチは、パッチ内の任意の位置に配置され得る1つ又は複数の核を含むことができる。前記核のうちの1つがNoIとして使用される場合、前記核を含むパッチが、前記パッチに含まれているすべての他の核がNoIの隣接する核として見なされる領域を画定する。したがって核の近辺を画定しているパッチは、いくつかの実施形態ではNoIの周囲に中心が置かれ得るが、例えばパッチの格子の周囲の格子点に同じく中心が置かれ得る。当然、当業者は、提供される任意のこれらの変数を使用して、パッチの任意の格子がアセンブルされ得ることを認識するであろう。そのために、点の格子が画像中に一様に分配される（点と点の間隔距離は一定である）。個々の点では、パッチと呼ばれる、点を中心とする長方形（又は円形）の領域が生成される。パッチのサイズ及び形状の効果は、実験によって評価されなければならない。一般に、何らかの特定の理論に縛られるつもりはないが、極めて小さいパッチは、生物学的に意味のあるラベルを割り当てるための十分な情報を提供しないとされている。一方、何らかの特定の理論に縛られるつもりはないが、大きいパッチには長い計算時間が必要であり、その上、複数の組織タイプを含むことがあり、それは、方法にとっては好ましくないとされている。したがってパッチサイズは、差し迫った組織特性を反映するために、分析問題毎に選択される。

30

40

【0075】

[0092]いくつかの実施形態では、生成されるパッチは、「スーパーピクセル」の形態である。「スーパーピクセル」は、複数の隣接するピクセルを網羅している画像の副領域である。「スーパーピクセル」は、画像を自由形態形状を使用して非交差画像パッチに分割する。いくつかの実施形態では、形状は、個々のスーパーピクセルが目標サイズ範囲に合致し、かつ、1つのタイプの組織又は細胞を支配的に含むように選択され得る。スーパーピクセルは、「グラフベースアルゴリズム」、「傾斜-上昇ベースアルゴリズム」、「SLICアルゴリズム」、平均変位及び正規化切断を含む多くの方法によって生成され得る。したがって実

50

施形態によれば、スーパーピクセル生成手順は、個々のパッチがスーパーピクセルであるパッチを生成するための画像に対して適用され得る。

【0076】

[0093]スーパーピクセル生成に対するグラフベース手法は、個々のピクセルをグラフ中のノードとして取り扱う。いくつかの実施形態では、2つのノード間の縁重量 (edge weights) は、隣接するピクセル間の類似性に比例する。一方、傾斜上昇ベースアルゴリズムは、グラフ上で定義される費用関数を最小化することによってスーパーピクセルを生成する。ピクセルの大きな初期クラスタ化から開始して、傾斜上昇方法は、スーパーピクセルを形成するための何らかの集束基準に合致するまで反復してクラスタを洗練する。

10

【0077】

[0094]実施形態によれば、「パッチ」(すなわちスーパーピクセル)として使用される隣接するピクセルセットを識別するために、単純な線形反復クラスタ化が使用される。単純な線形反復クラスタ化 (SLIC) は、スーパーピクセルを生成するための k 平均の適合であり、(i) 探索空間をスーパーピクセルサイズに比例する領域に制限することにより、最適化における距離計算の数が劇的に低減される(これは、ピクセルの数 N における線形化の複雑性を低減し、また、スーパーピクセルの数 k に無関係であるとされている)、及び(ii) 重み付き距離測度が色と空間近接性を結合し、その一方で、それと同時にスーパーピクセルのサイズ及びコンパクト性に対する制御を提供する、という2つの重要な区別がある。(参照によりその開示全体が本明細書に組み込まれている、Achantaraの「SLIC Superpixels Compared to State-of-the-Art Superpixel Methods」(最新のスーパーピクセル方法と比較したSLICスーパーピクセル) IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 34, No. 11, 2012年11月を参照されたい。)例えばスーパーピクセルサイズに比例する領域は、スーパーピクセルを識別するために使用されるスーパーピクセル領域の定義済み上部リミットと全く同じであってもよい。スーパーピクセルの上部サイズリミットは、例えば10000個のピクセルであってもよい。

20

【0078】

[0095]色と空間近接性の間の結合の重みは、例えば0.2に設定され得る。これらのパラメータは、訓練フェーズの間、とりわけ高い検証精度を提供することが観察されている。

30

【0079】

[0096]個々の核に割り当てられる文脈特徴(又は「文脈情報メトリクス」)は、単一の組織タイプ等級又は周囲の組織が可能組織タイプのリストに属する確率のいずれかを含むことができるとされている。いくつかの実施形態では、パッチは、特異的な組織タイプのためにのみ生成される。例えばパッチは、図4A及び4Bに示されている領域などの、PD-L1などのバイオマーカに対して陽性の領域のためにのみ生成され得る。さらに、図5A及び5Bは、パッチ分類結果を示したもので、主として免疫細胞によるPD-L1染色を有する領域は、赤(又は円)で示されており、主として腫瘍細胞による染色PD-L1染色を有する領域は、緑(又は正方形)で示されており、また、非標的PD-L1染色は、黄(又は三角形)でそれぞれ示されている。核 n_i がパッチ p_i に属すると仮定し、また、パッチ p_i が主として特定の細胞タイプの細胞を含む組織の均質な領域からのパッチ、例えば免疫細胞領域からのパッチとして分類されると仮定すると、核 n_i の特徴は、[n_i の核特徴 + パッチラベル情報(前記特定の細胞タイプからのパッチ - 例えば免疫細胞 - 領域)] を含む。したがってパッチ(すなわち「スーパーピクセル」)には、一組の定義済み又は学習済み細胞タイプのうちの1つだけが割り当てられ得る。しかしながらこれは、パッチは異なる細胞タイプの細胞の核を含み得ることを排除しない。しかしながら特定の細胞タイプの細胞に含まれている核の確率は、核を含むパッチが異なる細胞タイプを割り当てたところでは低減され得る。

40

50

【0080】

[0097]パッチ生成方法には無関係に、文脈情報メトリクスは、パッチ領域すなわちスーパーピクセルから導出される文脈テクスチャ特徴である。いくつかの実施形態では、テクスチャ特徴は、異なる画像チャンネル（ヘマトキシリン、輝度及びDAB）から導出され、また、強度のヒストグラム、傾斜量及び傾斜配向のヒストグラム、Gabor特徴及びHaralick特徴を含む。

【0081】

[0098]いくつかの実施形態では、テクスチャ特徴は、異なる画像チャンネルから計算される。例えば異なる画像チャンネルは、組織標本を準備するために使用される染色又は対比染色に基づかれ得る（例えばヘマトキシリン、輝度、IHCチャンネル、PDL1染色チャンネル）。いくつかの実施形態では、異なる画像チャンネルからの信号の差を捕獲して、組織構造を記述するのに役に立ち得る強度ベース特徴が計算される。これは、値の範囲を「ビンニング（binning）」することによって達成され、すなわち値（強度）の全範囲が一連の小さいインターバル（interval）に分割され、次に、どれだけ多くの値が個々のインターバルに入るかが計数される。したがって「強度ベース特徴」は、ピクセル又は一組のピクセルのビンニングされた強度値であり得る。これらの特徴は、分類モジュールに供給され得る。他の実施形態では、画像の傾斜量及び傾斜配向を計算することによって傾斜特徴が決定される。いくつかの実施形態では、傾斜特徴は、傾斜量のヒストグラム及び/又は傾斜ベクトル配向のヒストグラムを含む。例えば傾斜特徴は、傾斜量の10-ビンヒストグラム（bin histogram）及び傾斜ベクトル配向の10-ビンヒストグラムを含み得る。これらの特徴は、例えばパッチ内のピクセルのために選択的に計算され、パッチは、例えばスーパーピクセル生成アルゴリズムによって識別され得る。これらの特徴は、均質領域を非均質領域から区別することができ、また、同様に配向された縁を有する領域を無作為に配向された縁を有する領域から区別することができる。ヒストグラムの計算は、値の範囲の「ビンニング」に関する上記説明と同様である。ヒストグラムに加えて、いくつかの実施形態では、平均、標準偏差、尖度（curtosis）、百分位数、等々のような異なる記述統計量が傾斜量及び傾斜配向の特徴として導出され得る。これらの特徴は、分類モジュールに供給され得る。

【0082】

[0099]他の実施形態では、メトリクスは、Gaborフィルタを適用することによって抽出されるテクスチャ特徴に基づいて計算される。

[00100]「Gabor特徴」は、例えば、1つ又は複数のGaborフィルタをデジタル画像に対して適用することによってデジタル画像から抽出されたデジタル画像の特徴である。1つ又は複数のGaborフィルタは、異なる周波数及び/又は配向を有し得る。Gaborフィルタは、例えば、画像中のパターンを検出するため、例えば縁を検出するために使用され得る線形フィルタである。Gaborフィルタの周波数及び配向表現は、ヒト視覚システムのそれらと同様であり、それらは、テクスチャ表現及び区別にとりわけ適していることが分かっている。Gaborフィルタは、画像解析、例えば縁検出のためにしばしば使用される線形フィルタである。例えばGaborフィルタは、正弦平面波によって変調されたGaussカーネル関数であってもよい。

【0083】

[00101]Gaborフィルタは、ヒト視覚システムの周波数及び配向感度特性をモデル化する能力を有しているとされている。Gaborフィルタは、画像を複数の異なる配向及び異なるスケールのログ-Gaborフィルタで巻き込み、次に、同じスケールの異なる配向の応答を平均して回転不変特徴を得る。Gaborフィルタの応答は、Gaborフィルタを一組の画像ピクセルの強度値に対して適用した結果である。NoIを含む画像パッチのピクセルに対して計算された応答は、NoIの文脈情報メトリクスとして使用され得る。いくつかの実施形態では、Gaborフィルタを使用して、文脈情報メトリクスとして使用され得る平均、標準偏差、平均応答に対する最小-最大比率が計算される。Gaborフィルタ及びそれらの適用に関するさらなる情報は、参照によりその開示全体が

10

20

30

40

50

本明細書に組み込まれている、Jain, A. K., Farrokhnia, F. 「Unsupervised texture segmentation using Gabor filters. (Gaborフィルタを使用した非監視テクスチャ分割)」IEEE Int. Conf. System, Man., Cyber., 頁14 - 19 (1990)に見出され得る。この場合も、これらの特徴は、分類モジュールに供給され得る。

【0084】

[00102]さらに他の実施形態では、文脈情報メトリクスはHaralick特徴を含む

[00103]Haralick特徴は、テクスチャのパターン中に出現するパターンに関する情報を捕獲するとされている。Haralickテクスチャ値は、共起行列を使用して計算される。この行列は、2つのピクセル（これらは何らかの距離によって互いに分離され得る）間の角関係及び距離の両方の関数であり、2つの指定されたピクセル間の関係の発生数を示す。「Haralickテクスチャ特徴」又は「Haralick特徴」は、例えば共起行列から抽出されたデジタル画像の特徴であり、互いに対する関係で特定の位置を有するデジタル画像のピクセル中の画像強度が共に生じる様子に関する情報を含む。Haralick特徴を計算するために、共起行列は、例えば、行列の強度レベルをデジタル画像中で識別された個々の対象内で観察された最大及び最小強度に基づくことによって正規化され得る。

【0085】

[00104]Haralick、Shanmugan及びDinstein(1973)は、これを「グレートン空間依存行列」と呼んでいる。本発明の実施形態に使用されるこれらの実施態様は、何らかの距離dによって分離されるピクセル間の4つの方向(0°、45°、90°及び135°)を考察している。(参照によりその開示全体が本明細書に組み込まれている、Haralick, R., Shanmugan, K., 及びDinstein, I. 「Textural Features for Image Classification. (画像分類のためのテクスチャ特徴)」IEEE Transactions on Systems, Man, and Cybernetics 3, no. 6 (1973): 610 - 621を参照されたい。)

[00105]実施形態によれば、共起行列(すなわち空間依存性共起行列)は、NoIを中心とするパッチ中のピクセルに対して計算される。実施形態によれば、共起行列は、複数の定義済み方向(すなわち「角度」)の各々に対して計算され、例えば4つの方向0°、45°、90°及び135°に対して計算される。

【0086】

[00106]生成された1つ又は複数の共起行列から、自己相関、コントラスト、相関、非類似性、エネルギー、エントロピー、均質性、最大確率、分散、和平均、和分散、和エントロピー、差分散、差エントロピー、相関の2つの情報測度、逆差、正規化逆差及び逆モーメントを含む複数の特徴が計算され得る。パラメータの各々は、共起行列中の異なるデータエントリー間の関係、例えば特徴「茶色の色チャネル中の高強度値」と、グレー値傾斜サイズに対する特定のビン値との相関を表すことができる。これらの値を考察中の個々のチャンネルから抽出し、かつ、平均、標準偏差及び個々の特徴画像のモードを取ることにより、極めて多くの共起特徴が得られる。これらの特徴のうちの任意の特徴が文脈情報メトリクスとして使用され得る。

【0087】

[00107]共起行列は、特定の細胞タイプ又は組織タイプの指標であり得る生物学的情報を示し得るため、パッチ中のピクセルに対する共起行列を計算することは有利であり得る。例えば共起行列及びその共起行列から導出される文脈情報メトリクスは、青いピクセル(核内のピクセル)が茶のピクセル(膜染色のピクセル)にしばしば接近する(距離d内で)様子を記述し得る。

【0088】

10

20

30

40

50

[00108]いくつかの実施形態によれば、グレーレベル共起行列(「GLCM」)は、画像チャンネル毎に個々に計算され、また、それぞれのHaralickテクスチャ値が個々の画像チャンネルから個別に導出される。

【0089】

[00109]チャンネル毎に個々に計算される従来のグレーレベル共起行列(「GLCM」)に加えて、チャンネル間すなわち色共起行列(「CCM」)が使用され得る。CCMは、2つの異なる画像チャンネルにおけるピクセル強度の共起から生成され、すなわち、変位ベクトル(例えば $d = [d_x; d_y]$)を使用して、2つのチャンネル(例えば $C_i; C_j$)からCCMを計算する。共起は、 C_i の位置($x; y$)におけるピクセル強度及び C_j の位置($x + d_x; y + d_y$)におけるピクセル強度から計算される。CCMは、異なる組織構造間の空間関係を、それらを明確に分割する必要なく捕獲する利点を提供するとされている(異なるチャンネルで強調される)。例えば第1のバイオマーカは、細胞膜の外部表面で発現されることが知られており、また、第2のバイオマーカは、細胞膜の内部表面で発現されることが知られており、第1及び第2のバイオマーカは、その信号が2つの異なる画像チャンネルで捕獲される異なる染料によって染色される場合、内部膜タンパク質及び外部膜タンパク質は、常に、又は支配的に、互いに空間的に極めて近接して信号を生成することになるため、2つの異なるチャンネルの信号の強度値は相関することになる(定義済みオフセットで)。前記空間近接は、異なるチャンネルにおけるピクセル強度相関の形態のCCM行列で捕獲され得る。

10

【0090】

20

[00110]いくつかの実施形態では、Haralick特徴は、考察中のすべてのチャンネルのGLCMから計算される。この場合も、CCMから計算された特徴のうちの任意の特徴が文脈情報メトリクスとして使用され得る。チャンネル間行列は、実施形態によれば、GLCM行列と同じ又は同様の方法で計算される。複数の異なる角度及び距離が考察され得る。唯一の相違は、ピクセルの対が2つの異なるチャンネルから拾い上げられ、例えばピクセル p_1 が画像チャンネル1に属し、また、 p_2 が画像チャンネル2に属し、その一方でこれらの2つのピクセルは、同じ座標系に存在すると見なされることである(したがってこれらの間の距離及び角度が計算され得る)。これらの特徴は、同様に分類モジュールに供給され得る。

【0091】

30

[00111]文脈 - テクストン方法

[00112]文脈 - テクストン方法は、個々のNoIを中心とする画像パッチからテクストンマップのヒストグラムを計算する。(参照によりその全体が組み込まれている、Malik、Jitendraらの「Textons, Contours and Regions: Cue Integration in Image Segmentation. (テクストン、輪郭及び領域: 画像分割におけるキュー統合)」s.i.: IEEE Computer Society、1999年、Proceedings of the International Conference on Computer Vision, Corfu, Greece.)テクストンマップは、以下のように計算され得る。

40

【0092】

[00113]文脈 - テクスチャ方法と同様、この方法の目的も、個々のNoIの周りの領域のテクスチャパターンを同じく捕獲することである。しかしながら、上で説明したようにテクスチャ特徴から文脈情報メトリクスを導出する代わりに、図6Dを参照すると、フィルタ応答画像のリストを得るために、最大応答フィルタのバンクが組織標本の画像に対して(又はそのチャンネル画像に対して)適用される(ステップ330)。(参照によりその全体が組み込まれている、Varma及びZisserman「Classifying images of materials: Achieving viewpoint and illumination independence, (材料の画像の分類: 視点及び照明独立性の達成)」in Computer Vision ECCV 2

50

002、2002年、vol. 2352、頁255 - 271を参照されたい。) 個々のフィルタ応答画像は、1つ又は複数のフィルタ応答を含むデジタル画像である。「フィルタ応答」は、[0018]で定義されるフィルタ応答であってもよい。訓練画像から導出されたフィルタ応答画像、及びその中に含まれているフィルタ応答が収集され、かつ、「テクストン」と呼ばれる複数のK個のクラスタにクラスタ化される(ステップ331)。

【0093】

[00114]例えば複数の最大応答フィルタを組織標本の画像に対して適用することによって得られるフィルタ応答の各々は、直径、強度、等々などのいくつかの特性値を有するベクトルであり得る。前記フィルタ応答のクラスタ化は、一組のK個のクラスタを提供することができ、それによりクラスタ中心が前記クラスタに属しているすべてのフィルタ応答の平均値のベクトルとしてクラスタ毎に反復して計算される。したがって個々のクラスタ中心は、「平均」フィルタ応答ベクトル(その値は、前記クラスタ中心に割り当てられたフィルタ応答のそれぞれの特徴ベクトルの平均値である)、又は前記クラスタ中心に割り当てられたフィルタ応答から導出された他の形態の「プロトタイプ」フィルタ応答ベクトルであり得る。クラスタのクラスタ中心を表す前記「微分」フィルタ応答ベクトルは、「テクストン」として使用される。例えば個々のクラスタ中心は、特定の画像パッチ上への個々のフィルタの一組の投影を表す。反復して洗練され得る前記K個の「クラスタ中心テクストン」は、K平均クラスタ化アルゴリズムの出力として提供され得る。クラスタ化アルゴリズムのための基準は、個々のフィルタ応答ベクトルを最も近くの中心に割り当てた後に、中心からの2乗された距離の合計が最小化されるようなK個の「中心」を見出すことである。したがって組織標本画像に含まれている情報を処理することにより、K個のテクストンのテクストン語彙が自動的に抽出される。次に、テクストンマップがクラスタ中心を構成しているテクストンから計算され得る。

10

20

【0094】

[00115]複数の訓練されたクラスタ中心に基づいて、組織標本の画像の個々のピクセルがK個のテクストンのうちの1つに割り当てられる(ステップ332)。

[00116]例えば割当ては、画像中の個々のピクセル(又はNoIの周りを中心とする画像パッチ中の少なくとも個々のピクセル)が、マップされる前記ピクセルを含む一組のピクセルのために生成されたフィルタ出力の特性であるテクストンのうちの1つに割り当てられるように実施され得る。個々のピクセルは、テクストンのうちの厳密に1つにマップされるため、画像は、異なるテクストンに割り当てられた領域に仕切られる。前記「仕切られた画像」は、「テクストンマップ」と呼ばれ得る。

30

【0095】

[00117]次に、NoIを中心とするサイズ $S \times S$ を有する画像パッチ中のすべてのピクセルからテクストンヒストグラムが計算される(ステップ333)。いくつかの実施形態では、パッチサイズは、任意の $S \times S$ 寸法における約50個のピクセルと約200個のピクセルの間の範囲である。他の実施形態では、約150個のピクセルのパッチサイズ(約70ミクロン)が使用される。前記パッチサイズ範囲は、テクスチャ関連文脈情報が予測パラメータである細胞タイプを正確に識別するためにとりわけ適していることが観察されている。

40

【0096】

[00118]この方法に従って出力として分類器に供給される文脈情報メトリクス(ステップ334)は、テクストンヒストグラムである。テクストンヒストグラムは、NoIを取り囲んでいるピクセルの前記パッチ中のテクストンの発生の頻度を示す。テクストンヒストグラムは、分類器による組織スライド画像中の細胞の細胞タイプの識別を可能にするための分類器への入力として提供される。テクストンの分配は、組織の個々のタイプのための区別サインを提供するとされており、計算された核特徴メトリクスと共に分類器への入力として使用される。

【0097】

[00119]文脈 - C R F 方法

50

[00120]文脈 - C R F方法は、条件付き確率場 (C R F) モデルを使用して分類結果の均質性を強化する (参照によりその開示全体が本明細書に組み込まれている、J . L a f f e r t y らの C o n d i t i o n a l R a n d o m F i e l d s : P r o b a b i l i s t i c M o d e l s f o r S e g m e n t i n g a n d L a b e l i n g S e q u e n c e D a t a . (条件付き確率場 : シーケンスデータを分割し、かつ、ラベル付けするための確率的モデル) I C M L 、 頁 2 8 2 - 2 8 9 、 2 0 0 1 年を参照されたい) 。本明細書における B o W モデルのような C R F モデルは、隣接する核からの事前計算済み核特徴 / メトリクス及びラベルを文脈情報として利用し、それにより追加特徴抽出を必要とすることなく文脈情報の組み込みを可能にする (文脈テキストチャ方法及び文脈テキストトン方法と比較して) 。 C R F モデルは、ペアワイズ制約を組み込むための自然な方法を提供し、同じ等級に属する隣接領域を補強するとされている。

10

【 0 0 9 8 】

[00121]文脈情報 (C I) メトリクスの計算のために、C R F モデルを使用してラベル付け結果の均質性が促進され、すなわち C R F モデルは、近くに配置された核が同じラベルを有するように促進する。均質な組織領域 (組織、間質、等々) は、通常、同じタイプの核を含むとされている。それぞれすべての核ラベル及び核特徴ベクトルのセットとして表される $y = \{ y_1, y_2, \dots, y_n \}$ 及び $x = \{ x_1, x_2, \dots, x_n \}$ である場合、ラベル付け問題は、最適化問題 $y^* = \arg \min_y p(y | x)$ として形式化される。このような問題は、通常、有効な推定のための確率分布を要因化するためにグラフ表現を使用して解かれる。ここではグラフ $G = (V, E)$ を使用して核間の関係がモデル化され、個々の頂点 $v_i \in V$ は核 n_i に対応する。縁 $e_{ij} \in E$ は、2つの核 n_i と n_j の間の距離が d 未満である場合、 v_i と v_j の間に生成される。いくつかの実施形態では、 $d = S / 2 = 75$ 個のピクセルである。 y^* は、G i b b s エネルギー関数

20

【 0 0 9 9 】

【 数 1 】

$$E(y) = \sum_i \varphi_u(y_i | x_i) + \alpha \sum_{i,j} \varphi_p(y_i | y_j)$$

【 0 1 0 0 】

を最小化することによって得られる。

30

[00122]単項ポテンシャル (u n a r y p o t e n t i a l) $\varphi_u(y_i | x_i)$ は、S V M 分類器などの分類器の負ログ確率出力として計算される。いくつかの実施形態では、 v_i と v_j の間に縁が存在する場合、 G を使用して $\varphi_p(y_i | y_j)$ が $[y_i, y_j]$ に設定され、さもなければ $\varphi_p(y_i | y_j) = 0$ が使用される。エネルギー出力関数は、グラフ切断手法を使用して最小化される。(参照によりその開示全体が組み込まれている、Y . B o y k o v らの「Fast approximate energy minimization via graph cuts, (グラフ切断を介した高速近似エネルギー最小化) 」I E E E P A M I 、 v o l . 2 3 、 頁 1 2 2 2 - 1 2 3 9 、 2 0 1 1 年を参照されたい。) 規則化パラメータ α は重要なパラメータとされている。実際、より大きい値の α は、ラベル付けのより高い均質性をもたらすとされている。規則化パラメータの選択は、図 8 a 及び 8 b に示されている (最良の規則化パラメータを見出すために異なるパラメータが試験される) 。この方法は、実際の出力を生成しないため、他の方法とは異なっている。その代わりにこの方法では、核分類は、最初に核特徴のみを使用して実施され、次に、初期核ラベルが得られる。引き続いて C R F 方法は、初期核ラベルの洗練を支援して核均質性を強化し、すなわち 2 つの近くに配置された核が同じラベルを有する可能性がより高くなる。

40

【 0 1 0 1 】

[00123]文脈 - B o W 方法

[00124]語の多重集合モデルは、基本ブロック (語) の頻度に基づく単純で、しかも強力な表現技法である。広く使用されている特徴符号化方法である語の多重集合 (B o W)

50

は、画像から抽出される局所特徴は互いに独立していると仮定しており、個々の画像中に出現する個々の視覚「語」の頻度を計数するだけである。本明細書において使用されているように、「語」は核特徴ベクトルである。

【0102】

[00125]この方法は、NOIの文脈情報は、その隣接する核の外観（例えば数及び/又は細胞タイプ分布）を介して記述され得る、という観察を使用している。いくつかの実施形態では、隣接する核は、距離 $d = 75$ 個のピクセル内の核として画定される。他の実施形態では、距離の範囲は、約 25 個のピクセルと約 100 個のピクセルの間である。

【0103】

[00126]隣接する核の外観を文脈情報として捕獲するために、「語の多重集合」を使用して、以下で概説される、また、図3A及び3Bに示されているステップを使用してそれらの外観が定量化される。

【0104】

[00127]最初に、一組の訓練画像が提供される。個々の訓練画像は、PD-L1陽性及び陰性腫瘍細胞並びに非腫瘍細胞、例えばリンパ球細胞を含む、異なるタイプの複数の細胞を含む。訓練画像は、核識別アルゴリズムのための入力として使用される。

【0105】

[00128]核が検出され（ステップ600）、次に、当分野で知られているように、また、上で説明した方法に従って、訓練画像中のすべての核に対する核特徴が抽出される（ステップ610を参照されたい）。次に、「C個の」クラスタ中心を得るために、すべての訓練核特徴に対するK平均アルゴリズムを使用してクラスタ化手順が実施される（ステップ620を参照されたい）。いくつかの実施形態では、K平均アルゴリズムは候補「語」をクラスタ化し、また、得られたクラスタ中心（C個）は、「視覚語彙」を構築するために獲得される。例えば訓練画像中の核の各々に対して抽出された核特徴メトリクスは、核特徴ベクトルとして記憶される。個々の識別された核は、前記核の核特徴及び/又は文脈情報メトリクスを含むそのそれぞれの核特徴ベクトルによって表される。次に、K平均アルゴリズムが訓練画像の核特徴ベクトルに対して反復して適用され、それにより個々のクラスタの中心が、前記クラスタ中心に割り当てられたすべての核特徴メトリクスのそれぞれの核特徴メトリクスの平均値のベクトルとして導出された計算核特徴ベクトルとして表される。クラスタ中心を表している核特徴ベクトルは、反復毎に再計算され、かつ、洗練され、訓練画像の核の核特徴と、クラスタ中心を表しているベクトル中のそれぞれの特徴値との間の相違が最小化されるよう、反復毎に個々の核特徴ベクトルがそれぞれのクラスタ中心に再割り当てされる。終了基準に到達すると、複数のクラスタ中心は、訓練フェーズの結果として、「視覚語彙」として提供される。それにより訓練された分類器、及び例えばC個の核特徴ベクトルの形態の一組のC個の事前訓練済みクラスタ中心が訓練フェーズで得られる。

【0106】

[00129]次に、組織標本の画像中で、例えば本発明の他の実施形態に対して既に説明した画像分割アルゴリズムによって核が識別される。組織標本の画像は、訓練画像を使用して訓練された、又は訓練ステップを実施した機械学習アプリケーションから「視覚ライブラリ」（すなわちC個のクラスタ中心を有する訓練された分類モデル）を受け取った分類器のための入力として使用される。

【0107】

[00130]組織標本の前記画像中の所与のNOIに対して、次に、C個の中心までの個々の隣接する核の核特徴ベクトルのEuclidean距離を使用して、その隣接する核が最も近くのクラスタ中心に割り当てられ、また、クラスタ割当てのヒストグラムが計算される（ステップ630を参照されたい）。クラスタ割当てのヒストグラムは、特定のクラスタに属するNOIの前記環境に含まれている核の数を示す。近辺における核のクラスタ割当ては、NOIの周りの環境を記述するための手段である。このヒストグラムは、NOIの文脈情報メトリックである。最後に、核特徴及び文脈特徴が訓練及び分類のために単

10

20

30

40

50

一のベクトルに結合される。

【0108】

[00131]いくつかの実施形態では、文脈 - B o W方法は、(a) 個々の核から抽出された1つ又は複数の核特徴に基づいて、訓練画像の個々の核に対する核特徴ベクトルを生成すること、(b) K平均アルゴリズムを使用してクラスタ化手順を核特徴ベクトルに対して実施することによって、複数のC個の事前訓練済みクラスタを獲得すること、(c) 訓練核に隣接する個々の核を、(c 1) 個々の個別の隣接する核の核特徴ベクトルから個々のC個のクラスタの中心までのE u c l i d e a n距離を測定し、かつ、(c 2) 個々の隣接する核を、その中心がその核の核特徴ベクトルに最も近いクラスタに割り当てることによって、複数のC個のクラスタのうち1つに割り当てること、(d) 隣接する核のクラスタ割当てのヒストグラムを計算することによって、個々の訓練核の文脈特徴を決定すること、及び(e) 核特徴及び文脈特徴を訓練核のための単一の完全な特徴ベクトルに結合すること、並びに(f) すべての訓練核の完全な特徴ベクトルを使用して分類モデルを訓練することによって、複数の訓練画像からの複数の訓練核上で訓練される。

10

【0109】

[00132]いくつかの実施形態では、文脈 - B o W方法は、(a) 核から抽出された1つ又は複数の核特徴に基づいて、試験画像の個々の核に対する核特徴ベクトルを生成すること、(b) 対象の核の個々の個別の隣接する核を、(c 1) 個々の個別の隣接する核の核特徴ベクトルから複数のクラスタの中心までのE u c l i d e a n距離を測定し、かつ、(c 2) 個々の隣接する核を、その中心がその核の核特徴ベクトルに最も近いクラスタに割り当てることによって、複数の事前訓練済みクラスタのうち1つに割り当てること、(d) 隣接する核のクラスタ割当てのヒストグラムを計算することによって、対象の核の文脈特徴を決定すること、及び(e) 対象の核の核特徴ベクトルを文脈特徴と結合して、対象の核のための完全な特徴ベクトルにすること、(f) 訓練された分類モデルを対象の核の完全な特徴ベクトルに対して適用してそれを分類することによって試験画像中の対象の核を分類する。

20

【0110】

[00133]いくつかの実施形態では、B o W方法は文脈 - C R F方法より安定であり、及び/又は必要な計算コストが文脈 - テクスチャ方法及び文脈 - テクストン方法より安価であるとされている(追加特徴を計算する必要がないため)。

30

【0111】

[00134]分類モジュール

[00135]核メトリクス及び文脈情報メトリクスが特徴抽出モジュールによって導出されると、メトリクスは、タイプ(例えば腫瘍、免疫、間質、等々)、又は特定の染色(例えばP D L 1の存在を示す染色)に対する応答に従って細胞核を検出し、かつ、ラベル付けするために分類モジュールに提供される。いくつかの実施形態では、分類器が訓練され、次に、陽性腫瘍、陰性腫瘍、陽性リンパ球、非標的染色、及びその他を含むP D - L 1染色組織中の核の5つの等級を区別するために使用される(P D - L 1染色肺組織画像中の核の5つの等級を示し、陽性腫瘍、陰性腫瘍、陽性リンパ球、非標的染色及びその他がそれぞれ緑矢印(「E」)、青矢印(「A」)、赤矢印(「B」)、黄矢印(「C」)及び青緑矢印(「D」)で示されている図7 A及び7 Bを参照されたい)。訓練中、専門家観察者によって、当業者に知られている手順に従って提供されるグラウンドトゥース識別と共に例示的細胞が提供される。

40

【0112】

[00136]いくつかの実施形態では、分類モジュールは、Support Vector Machine(「SVM」)である。一般に、SVMは、統計的学習理論に基づかれる分類技法であり、非線形ケースの場合、非線形入力データセットがカーネルを介して高次元線形特徴空間に変換される。何らかの特定の理論に縛られるつもりはないが、サポート・ベクター・マシンは、2つの異なる等級を表す一組の訓練データEをカーネル関数Kによって高次元空間に投入するとされている。この変換されたデータ空間では、非線形デ

50

ータが変換され、したがって等級を分離して等級分離を最大化するための平らな線が生成され得る（識別超平面）。次に、試験データがKを介して高次元空間に投入され、それらが超平面に関して降下する場所を基本として試験データが分類される。カーネル関数Kは、データが高次元空間に投入される方法を定義する。

【0113】

[00137]以下の例1で詳細に説明されるように、本出願人は、PD-L1核の分類には文脈情報が有用であることを示すために広範囲にわたる実験評価を実施した。さらに、本出願人は、提案される文脈-Bow方法は、典型的には数万の細胞核を含む組織標本の分析に関連する良好な分類精度を低計算コストで提供するため、魅力的であることを示した。

10

【0114】

[00138]例1

[00139]PD-L1染色肺標本からの腫瘍のある組織から検出された核を有する核データベースを使用して包括的評価が実施された。スライドがVENTANA iScan HTスキャナ上で走査され、0.465 μ mピクセルサイズのRGB全スライド画像が得られた。

【0115】

[00140]これらのPD-L1染色組織画像（図7A及び7Bを参照されたい）の場合、5つの異なるタイプの核を分類しなければならなかった。PD-L1データベースは、PD-L1染色肺標本から得られた600 \times 700個のピクセルの256個の画像を含んでいた。スライドは、再度VENTANA iScan HTスキャナ上で走査され、0.465 μ mピクセルサイズのRGB全スライド画像が得られた。これらの画像から、PD-L1陽性腫瘍細胞（2,414個）からの核、PD-L1陰性腫瘍細胞（1,620個）からの核、PD-L1陽性リンパ球（2,851個）からの核、非標的染色細胞（1,753個）からの核及び残りのPD-L1陰性細胞（1,834個）からの核を含む5つのタイプの多数の核が病理学者によって選択された（図7A及び7B）。これらは、PD-L1陽性腫瘍核、PD-L1陰性腫瘍核、PD-L1陽性リンパ球、非標的染色応答（非標的染色）、及び他の細胞の残りのPD-L1陰性核である。PD-L1マーカは、細胞の膜を染色して、核境界に沿って茶色のプロブ又は茶色のリングを生成する。表1は、3つのデータベースを要約したものである。図9及び10は、分類結果の追加例を提供したものである。図9及び10では、緑の点（ひし形）、青い点（開いた円）、赤い点（開いた正方形）、黄色い点（三角形）及び青緑の点（アスタリスク）は、それぞれ陽性腫瘍、陰性腫瘍、陽性リンパ球、非標的染色及びその他の核を表している。図9E及び9Fは、組織の特定の切片を示しており、また、同様に陽性腫瘍、陰性腫瘍、陽性リンパ球、非標的染色及びその他の核を示している。

20

30

【0116】

[00141]10倍画像ベース交差検証がこのデータベースに対して実施され、平均精度及び標準偏差が報告される。個々の画像中の核が訓練又は試験のいずれかのために使用され、すなわち核は、同じ画像から取り上げられた核を使用して訓練されたモデルは決して試験されなかった。文字通り他の研究と同様、SVM分類器が使用された。核特徴のみ（文脈情報なし）が使用された場合に得られる分類精度、並びに4つの文脈-認識方法（context-aware methods）によって得られる精度が表1に報告されている。文脈-CRF方法及び文脈-Bow方法が、図8A及び8Bにおけるパラメータ選択に関して、それらの精度をプロットすることによってさらに比較された。異なるパラメータ値（図8A及び8Bに示されている）を使用した文脈-Bow方法及び文脈CRF方法の性能の評価は、文脈-Bow方法は、パラメータ選択に対してより安定であることを立証した。

40

【0117】

[00142]上で言及したように、文脈-CRF方法の重要なパラメータは規則化パラメータであり、一方、文脈-Bow方法の重要なパラメータはクラスタの数Cである。「規則

50

化パラメータ」は、例えば所与のクラスタの核に割り当てられるラベルの均質性を制御するパラメータである。このパラメータの値が大きいほど、同じ等級に割り当てられた核に同じラベルが割り当てられるようになる確率が高くなる。

【0118】

[00143] 3つのすべてのデータベースに対して、文脈 - C R F方法の試験は、2 [0 : 0 0 1 ; 0 : 0 1 ; 0 : 0 5 ; 0 : 1 ; 0 : 5 ; 1] を含み、また、文脈 - B o W方法は、C 2 [1 0 ; 3 0 ; 5 0 ; 7 0 ; 1 0 0 ; 1 5 0 ; 2 0 0] を使用している。結果は、最良の精度が得られるパラメータ選択に対して報告されている。文脈 - テクスチャ方法の場合、重要なパラメータはパッチサイズSのみである。このパラメータは、文脈 - C R F方法及び文脈 - B o W方法におけるパラメータdと整合するよう、150個のピクセルとして固定されている。パラメータdは、例えば2つの核 n_i と n_j の間の距離閾値であってもよく、2つの核は、それらの距離がd未満である場合、隣接する核と見なされ、dは、例えば75pxであってもよい。公平な比較のために同じサイズの局所組織近辺が強制されている。すべての分類精度が表2に報告されている。

10

【0119】

[00144] これらの結果は、(i) 核メトリクスと相俟った文脈情報メトリクスの使用は、核メトリクスのみに依拠している従来技術と比較すると、核分類精度の改善を提供する、(ii) 提案される文脈 - B o W方法は、文脈 - C R F方法より良好に実施すると思われる、(iii) また、文脈 - テクスチャ方法及び文脈 - テクストン方法は、文脈 - B o W方法より幾分か良好に実施するが、それらは、計算的に高価である、個々の核を中心とする画像パッチに対するテクスチャ特徴又はフィルタ応答の余計な計算を必要とすると思われる、という主要な結論をもたらした。それとは対照的に、文脈 - B o W方法及び文脈 - C R F方法は、隣接する核からの事前計算済み核特徴及びラベルをC Iとして利用し、したがってそれらの方法は、追加特徴抽出を必要とすることなくC Iを組み込むことができる。

20

【0120】

【表1】

表1：核メトリクスのみを使用する場合、および同じ核メトリクスを文脈情報メトリクスと組み合わせて使用する場合に得られる分類精度 (s. d.) (C Iメトリクスは、4つの異なる方法から得られるものとして示されている)。すべての実例において、C Iメトリクスと核メトリクスの組合せは、比較的優れた結果をもたらした。

30

データベース	核特徴のみ	文脈 テクスチャ	文脈 テクストン	文脈 - C R F	文脈 - B o W
PD-L1	76.3 (9.3)	83.0 (10.1)	83.5 (8.2)	80.8 (8.1)	82.0 (7.2)

【0121】

[00145] 例2

40

[00146] P D - L 1 染色スライドをスコアリングするための例が図11A ~ Dに示されている。P D - L 1 (茶色のD A Bでラベルが付けられている) 及び対比染色ヘマトキシリンで免疫組織化学的に染色された組織標本のデジタル画像が図11A及び11Bに示されている。開示されている方法を実現している自動分析は、細胞を個々に検出し (この解像度では見られない)、かつ、それらをP D - L 1 陽性腫瘍細胞、P D - L 1 陰性腫瘍細胞、P D - L 1 陽性リンパ球又は任意の他の細胞のうちの一つとしてラベル付けした。これらの自動生成読取りから、免疫及び免疫細胞に関して、このスライド上の組織がそのP D - L 1 状態に対して個々にスコアされた。腫瘍細胞をスコアするために、P D - L 1 陽性腫瘍細胞の微小部分が腫瘍細胞の総数 (すなわちP D - L 1 陽性腫瘍細胞及びP D - L 1 陰性腫瘍細胞) で割られる。スライド上の組織は、腫瘍細胞の約90%がP D - L 1 に

50

対して陽性としてスコアされた。免疫細胞をスコアするために、これらの細胞を含む画像中の面積が測定される。PD-L1陽性リンパ球（免疫細胞）を含む腫瘍面積の微小部分がスコアされる。図11B及び11Dに示されている組織の場合、腫瘍の面積は 43.9 mm^2 であり（緑又は軽く影が付けられたオーバーレイとして示されている）、また、PD-L1陽性免疫細胞を有する面積は 3.2 mm^2 である（赤又は暗く影が付けられたオーバーレイとして示されている）。したがってこの組織スライドに対するPD-L1免疫スコアは、 $3.2 / (3.2 + 43.9) = 6.8\%$ である。

【0122】

[00147]本開示の実施形態を實踐するための他の構成要素

[00148]本開示の別の態様では、組織学画像中の核を分類する方法であって、前記方法は、NoIの抽出された特徴及びNoIを取り囲んでいる領域に関する文脈情報に基づいて対象の核（NoI）を分類するようにプログラムされたコンピュータプロセッサを備えたコンピュータ装置上で組織学画像を解析するステップを含む。いくつかの実施形態では、文脈情報は、NoIを取り囲んでいる核に関する文脈情報を含む。いくつかの実施形態では、文脈情報は、文脈-CRF方法又は文脈-Bow方法を使用して計算される。いくつかの実施形態では、文脈-Bow方法は、（a）個々の核から抽出された1つ又は複数の核特徴に基づいて、訓練画像の個々の核に対する核特徴ベクトルを生成すること、（b）K平均アルゴリズムを使用してクラスタ化手順を核特徴ベクトルに対して実施することによって、複数のC個の事前訓練済みクラスタを獲得すること、（c）訓練核に隣接する個々の核を、（c1）個々の個別の隣接する核の核特徴ベクトルから個々のC個のクラスタの中心までのEuclidean距離を測定し、かつ、（c2）個々の隣接する核を、その中心がその核の核特徴ベクトルに最も近いクラスタに割り当てることによって、複数のC個のクラスタのうち1つに割り当てること、（d）隣接する核のクラスタ割当てのヒストグラムを計算することによって、個々の訓練核の文脈特徴を決定すること（「隣接する核」は、例えば前記訓練核のコアからの最大距離以内に存在する核、及び/又は前記訓練核の周りを中心とするパッチであって、例えば150個のピクセル（約70ミクロン）の寸法を有するパッチ内に存在する核である）、及び（e）核特徴及び文脈特徴を訓練核のための単一の完全な特徴ベクトルに結合すること、（f）すべての訓練核の完全な特徴ベクトルを使用して分類モデルを訓練することによって、複数の訓練画像からの複数の訓練核上で訓練されている。いくつかの実施形態では、文脈-Bow方法は、（a）核から抽出された1つ又は複数の核特徴に基づいて、試験画像の個々の核に対する核特徴ベクトルを生成すること、（b）NoIの個々の個別の隣接する核を、（c1）個々の個別の隣接する核の核特徴ベクトルからC個のクラスタの中心までのEuclidean距離を測定し、かつ、（c2）個々の隣接する核を、その中心がその核の核特徴ベクトルに最も近いクラスタに割り当てることによって、C個の事前訓練済みクラスタのうち1つに割り当てること、（d）隣接する核のクラスタ割当てのヒストグラムを計算することによって、NoIの文脈特徴を決定すること、及び（e）NoIの核特徴ベクトルを文脈特徴と結合して、NoIのための完全な特徴ベクトルにすること、（f）訓練された分類モデルをNoIの完全な特徴ベクトルに対して適用してそれを分類することによって試験画像中のNoIを分類する。

【0123】

[00149]訓練並びに試験フェーズでは、クラスタ割当てのヒストグラムは、特定のクラスタ（又はクラスタ中心）に割り当てられる特定のNoIに隣接する核である核の数を示す。「割り当てられる」は、隣接する核の核特徴ベクトルの、その割り当てられたクラスタの中心までのEuclidean距離が、他のすべてのクラスタの中心までのEuclidean距離よりも短いことを暗に意味し得る。したがって個々のクラスタ中心は、核特徴ベクトルとして同じく表されることが可能であり、また、k平均クラスタ化を実施している間、反復して計算され、かつ、洗練され得る。クラスタ中心核特徴ベクトルは、微分、例えば平均核特徴メトリクスのベクトル、又は特定の反復で前記クラスタに割り当てられるすべての隣接する核の核特徴ベクトルとして計算され得る。

【 0 1 2 4 】

[00150]本開示のコンピュータシステムは、組織標本に対する1つ又は複数の準備プロセスを実施することができる標本処理装置に結合され得る。準備プロセスは、それらに限定されないが、標本の脱パラフィン化、標本の条件付け（例えば細胞条件付け）、標本の染色、抗原回収の実施、免疫組織化学染色（ラベル付けを含む）又は他の反応の実施、及び/又はインサイチューハイブリダイゼーション（例えばS I S H、F I S H、等々）染色（ラベル付けを含む）又は他の反応の実施、並びに顕微鏡法、微量分析、質量分析方法又は他の分析方法のための標本を準備するための他のプロセスを含むことができる。

【 0 1 2 5 】

[00151]標本は組織標本を含むことができる。組織の標本は、その中又は上に標的が存在し得る任意の液体、半固体又は固体物質（又は材料）であってもよい。詳細には、組織標本は、生物学的標本又は生物学的組織から得られる組織標本であってもよい。組織は、生体内の同様の機能を実施する相互接続された細胞の集合であってもよい。いくつかの例では、生物学的標本は、ヒト被検体などの動物被検体から得られる。生物学的標本は、それらに限定されないが、とりわけバクテリア、酵母菌、原生動物及びアメーバなどの単細胞生体、多細胞生体（健康な、又は外見上健康なヒト被検体からの標本、又は癌などの診断又は調査されるべき状態又は疾病による影響を受けたヒト患者からの標本を含む植物又は動物など）を含む、任意の生きた生体から得られた、任意の生きた生体によって排泄された、又は任意の生きた生体によって分泌された任意の固体標本又は流体標本であってもよい。例えば生物学的標本は、例えば血液、血漿、血清、尿、胆汁、腹水、唾液、脳脊髄液、房水又は硝子体液あるいは任意の身体分泌物、漏出液、滲出液（例えば膿瘍又は感染あるいは炎症の任意の他の部位から得られる流体）、又は関節（例えば正常関節又は疾病による影響を受けた関節）から得られる流体から得られる生物学的流体であってもよい。また、生物学的標本は、任意の臓器又は組織（腫瘍生検などの生検標本又は剖検標本を含む）から得られる標本であってもよく、あるいは細胞（一次細胞であれ、あるいは培養された細胞であれ）又は任意の細胞、組織又は臓器によって条件付けられた媒質を含むことも可能である。いくつかの例では、生物学的標本は核抽出物である。特定の例では、標本は、開示された細胞小球切片標本のうちの1つなどの品質管理標本である。他の例では、標本は試験標本である。例えば試験標本は、被検体から得られた生物学的標本から準備される細胞、組織又は細胞小球切片である。一例では、被検体は、危険な状態にある被検体、又は特定の条件又は疾病にかかった被検体である。いくつかの実施形態では、標本は乳房組織である。

【 0 1 2 6 】

[00152]処理装置は、固定剤を標本に加えることができる。固定剤は、架橋剤（アルデヒド、例えばホルムアルデヒド、パラホルムアルデヒド及びグルタルアルデヒド、並びに非アルデヒド架橋剤など）、酸化剤（例えば、無水オスミン酸及びクロム酸などの、金属イオン及び錯体）、タンパク質変性剤（例えば酢酸、メタノール及びエタノール）、未知の機構の固定剤（例えば塩化第二水銀、アセトン及びピクリン酸）、組合せ試薬（例えばカルノア（C a r n o y）の固定剤、メタカルン（m e t h a c a r n）、ブアン（B o u i n）液、B 5固定剤、ロスマン（R o s s m a n）液及びジェンドル（G e n d r e）液、マイクロ波及び多様な固定剤（例えば排除体積定着及び蒸気定着）を含むことができる。

【 0 1 2 7 】

[00153]標本がパラフィン中に埋め込まれた標本である場合、標本は、適切な脱パラフィン化流体を使用して脱パラフィン化され得る。廃棄物除去剤が脱パラフィン化流体を除去すると、任意の数の物質が連続的に標本に加えられ得る。物質は、前処理（例えばタンパク質の架橋、核酸の露出、等々）、変性化、ハイブリダイゼーション、洗浄（例えば嚴重洗浄）、検出（例えば視覚又はマーカ分子をプローブに連結する）、増幅（例えばタンパク質、遺伝子、等々の増幅）、対比染色、カバーガラス化、等々のためのものであってもよい。

10

20

30

40

50

【0128】

[00154] 標本処理装置は、広範囲にわたる物質を標本に加えることができる。物質は、それらに限定されないが、染料、プローブ、試薬、リンス及び/又はコンディショナを含む。物質は、流体（例えば気体、液体又は気体/液体混合物）、等々であってもよい。流体は、溶媒（例えば極性溶媒、非極性溶媒、等々）、溶液（例えば水溶液又は他のタイプの溶液）、等々であってもよい。試薬は、それらに限定されないが、染料、湿潤剤、抗体（例えば単クローン抗体、ポリクローン抗体、等々）、抗原回収流体（例えば水溶性又は非水溶性ベース抗原回収溶液、抗原回収緩衝剤、等々）、等々を含むことができる。プローブは、検出可能ラベル又はリポータ分子に取り付けられた、隔離された核酸又は隔離された合成オリゴ核酸塩であってもよい。ラベルは、放射性アイソトープ、酵素基質、補因子、リガンド、化学発光又は蛍光薬品、ハプテン及び酵素を含むことができる。

10

【0129】

[00155] 標本処理装置は、Ventana Medical Systems社が販売しているBENCHMARK XT器機及びSYMPHONY器機などの自動装置であってもよい。Ventana Medical Systems社は、自動分析を実施するためのシステム及び方法を開示している、米国特許第5650327号、第5654200号、第6296809号、第6352861号、第6827901号及び第6943029号、並びに米国公開特許出願第20030211630号及び第20040052685号を含む多くの米国特許の譲受人であり、これらの各々は、参照によりその全体が本明細書に組み込まれている。別法としては、標本は手動で処理され得る。

20

【0130】

[00156] 標本が処理されると、ユーザは、標本を担ったスライドを画像化装置に輸送することができる。ここで使用される画像化装置は、明視野イメージスライドスキャナである。明視野イメージの1つは、Ventana Medical Systems社が販売しているiScan Coreo（商標）明視野スキャナである。自動化実施形態では、画像化装置は、IMAGING SYSTEMS AND TECHNIQUES（画像化システム及び技法）という名称の国際特許出願第PCT/US2010/002772号（特許公開第WO/2011/049608号）に開示されているデジタル病理学デバイス、又は2011年9月9日に出願した、IMAGING SYSTEMS, CASSETTES, AND METHODS OF USING THE SAME（画像化システム、カセット及びその使用方法）という名称の米国特許出願第61/533114号に開示されているデジタル病理学デバイスである。国際特許出願第PCT/US2010/002772号及び米国特許出願第61/533114号は、参照によりその全体が組み込まれている。他の実施形態では、画像化装置は、顕微鏡に結合されたデジタルカメラを含む。

30

【0131】

[00157] 画像化システム又は装置は、多重スペクトル画像化(MSI)システム又は蛍光顕微鏡法システムであってもよい。ここで使用される画像化システムはMSIである。MSIは、通常、病理学標本の分析には、ピクセルレベルにおける画像のスペクトル分布へのアクセスを提供することによるコンピュータ化された顕微鏡ベース画像化システムが必要である。様々な多重スペクトル画像化システムが存在しているが、これらのシステムのすべてに共通している動作態様は、多重スペクトル画像を形成する能力である。多重スペクトル画像は、特定の波長又は電磁スペクトルにわたる特定のスペクトル帯域幅で画像データを捕獲する画像である。これらの波長は、光フィルタによって、又は例えば赤外線(IR)などの可視光範囲の範囲を超える波長における電磁放射を含む所定のスペクトル成分を選択することができる他の器機を使用することによって選出され得る。

40

【0132】

[00158] MSIは、光学画像化システムを含むことができ、その一部は、所定の数Nの離散光帯域を画定するべく同調可能なスペクトル選択システムを含む。光学システムは、広帯域光源を使用した伝送で光検出器上に照明された組織標本を画像化するように適合さ

50

れ得る。一実施形態では例えば顕微鏡などの拡大システムを含むことができる光学画像化システムは、通常は光学システムの単一の光出力と空間的に整列される単一の光軸を有する。システムは、画像が異なる離散スペクトル帯域で獲得されることを補償するなどのためにスペクトル選択システムが調整すなわち同調されると（例えばコンピュータプロセッサを使用して）、組織の一連の画像を形成する。装置は、追加で、獲得された一連の画像から、組織の少なくとも1つの視覚的に知覚可能な画像が出現するディスプレイを含むことも可能である。スペクトル選択システムは、回折格子、薄膜干渉フィルタなどの光フィルタの集合、又はユーザ入力又は事前プログラム済みプロセッサのコマンドのいずれかに応答して、光源から標本を介して検出へ向けて伝送される光のスペクトルから特定の通過帯域を選択するように適合された任意の他のシステムなどの光学分散要素を含むことができる。

10

【0133】

[00159]代替実施態様では、スペクトル選択システムは、N個の離散スペクトル帯域に対応するいくつかの光出力を画定する。このタイプのシステムは、光学システムからの伝送光出力を取り込み、かつ、識別されたスペクトル帯域中の標本をこの識別されたスペクトル帯域に対応する光路に沿った検出器システム上に画像化する方法で、この光出力の少なくとも一部をN個の空間的に異なる光路に沿って方向を向けなおす。

【0134】

[00160]本明細書において説明されている本主題の実施形態及び動作は、デジタル電子回路機構の中、又は本明細書において開示されている構造及びそれらの構造的等価物を含むコンピュータソフトウェア、ファームウェア又はハードウェアの中、あるいはそれらのうちの1つ又は複数の組合せの中で実現され得る。本明細書において説明されている本主題の実施形態は、1つ又は複数のコンピュータプログラム、すなわちデータ処理装置による実行のため、又はデータ処理装置の動作を制御するためにコンピュータ記憶媒体上に符号化されたコンピュータプログラム命令の1つ又は複数のモジュールとして実現され得る。

20

【0135】

[00161]コンピュータ記憶媒体は、コンピュータ可読記憶デバイス、コンピュータ可読記憶基板、ランダム又はシリアルアクセスメモリアレイ又はデバイス、あるいはそれらのうちの1つ又は複数の組合せであっても、あるいはそれらの中に含まれ得る。さらに、コンピュータ記憶媒体は伝搬信号ではないが、コンピュータ記憶媒体は、人工的に生成された伝搬信号中に符号化されたコンピュータプログラム命令のソース又は行先であってもよい。また、コンピュータ記憶媒体は、1つ又は複数の個別の物理的構成要素又は媒体（例えば複数のCD、ディスク又は他の記憶デバイス）であっても、あるいはそれらの中に含まれ得る。本明細書において説明されている動作は、1つ又は複数のコンピュータ可読記憶デバイス上に記憶された、又は他のソースから受け取られるデータ上でデータ処理装置によって実施される動作として実現され得る。

30

【0136】

[00162]「プログラムされたプロセッサ」という用語は、一例としてプログラマブルマイクロプロセッサ、コンピュータ、チップ上システム又は複数のチップ上システム、又はそれらの組合せを含む、データを処理するためのあらゆる種類の装置、デバイス及び計算機を包含する。装置は、専用論理回路機構、例えばFPGA（書替え可能ゲートアレイ）又はASIC（特定用途向け集積回路）を含むことができる。また、装置は、ハードウェアに加えて、当該コンピュータプログラムのための実行環境を生成するコード、例えばプロセッサファームウェア、プロトコルスタック、データベースマネジメントシステム、オペレーティングシステム、クロスプラットフォームランタイム環境、仮想計算機、又はそれらのうちの1つ又は複数の組合せを構成するコードを含むことも可能である。装置及び実行環境は、ウェブサービス、分散計算及び格子計算インフラストラクチャなどの様々な異なる計算モデルインフラストラクチャを実現することができる。

40

【0137】

50

【00163】コンピュータプログラム（プログラム、ソフトウェア、ソフトウェアアプリケーション、スクリプト又はコードとしても知られている）は、コンパイル済み又は翻訳済み言語、宣言又は手続き形言語を含む任意の形態のプログラミング言語で作成されることが可能であり、また、コンピュータプログラムは、独立型プログラムとして、あるいはモジュールとして、コンポーネント、サブルーチン、オブジェクト又は計算環境に使用するために適した他のユニットを含む任意の形態で展開され得る。コンピュータプログラムは、必ずしもその必要はないが、ファイルシステム内のファイルに対応し得る。プログラムは、他のプログラム又はデータ（例えばマーク付け言語文書に記憶されている1つ又は複数のスクリプト）を保持しているファイルの一部、当該プログラム専用の単一のファイル又は複数の調整済みファイル（例えば1つ又は複数のモジュール、サブプログラム又はコードの一部を記憶するファイル）に記憶され得る。コンピュータプログラムは、1つのコンピュータ上、又は1つのサイトに配置されている、あるいは複数のサイトにわたって分散され、通信ネットワークによって相互接続された複数のコンピュータ上で実行されるように展開され得る。

10

20

30

40

50

【0138】

【00164】本明細書において説明されているプロセスフロー及び論理フローは、入力データ上で動作し、かつ、出力を生成することによってアクションを実施するために、1つ又は複数のコンピュータプログラムを実行する1つ又は複数のプログラマブルプロセッサによって実施され得る。また、プロセスフロー及び論理フローは、専用論理回路、例えばFPGA（書替え可能ゲートアレイ）又はASIC（特定用途向け集積回路）によって実施されることも可能であり、また、装置は、専用論理回路、例えばFPGA（書替え可能ゲートアレイ）又はASIC（特定用途向け集積回路）として同じく実現され得る。

【0139】

【00165】コンピュータプログラムを実行するために適したプロセッサには、一例として、汎用及び専用の両方のマイクロプロセッサ、及び任意の種類デジタルコンピュータの任意の1つ又は複数のプロセッサがある。通常、プロセッサは、リードオンリメモリ又はランダムアクセスメモリあるいはそれらの両方から命令及びデータを受け取ることになる。コンピュータの本質的な要素は、命令に従ってアクションを実施するためのプロセッサ、及び命令及びデータを記憶するための1つ又は複数のメモリデバイスである。通常、コンピュータは、データを記憶するための1つ又は複数の大容量記憶装置、例えば磁気ディスク、光磁気ディスク又は光ディスクを含むことも可能であり、あるいはデータを受け取り、もしくはデータを転送し、又はそれらの両方のためにこれらの大容量記憶装置に動作結合され得る。しかしながら、コンピュータはこのようなデバイスを有する必要はない。さらに、コンピュータは別のデバイスの中に埋め込まれることも可能であり、例えばほんのいくつかを挙げると、携帯電話、パーソナルデジタルアシスタント（PDA）、移動オーディオ又はビデオプレーヤ、ゲームコンソール、広域位置決めシステム（GPS）受信機又は携帯型記憶装置（例えばユニバーサルシリアルバス（USB）フラッシュドライブ）の中に埋め込まれ得る。コンピュータプログラム命令及びデータを記憶するために適したデバイスには、一例として半導体メモリデバイス、例えばEPROM、EEPROM、及びフラッシュメモリデバイス、磁気ディスク、例えば内部ハードディスク又は取外し可能ディスク、光磁気ディスク、及びCD-ROM並びにDVD-ROMディスクを始めとする、あらゆる形態の不揮発性メモリ、媒体及びメモリデバイスがある。プロセッサ及びメモリは、専用論理回路機構によって補足され、あるいは専用論理回路機構の中に組み込まれ得る。

【0140】

【00166】ユーザとの対話を提供するために、本明細書において説明されている主題の実施形態は、表示デバイス、例えばユーザに情報を表示するためのLCD（液晶ディスプレイ）、LED（発光ダイオード）ディスプレイ又はOLED（有機発光ダイオード）ディスプレイ、及びユーザがコンピュータに入力を提供することができるキーボード及び位置決め装置、例えばマウス又はトラックボールを有するコンピュータ上で実現され得る。い

くつかの実施態様では、タッチスクリーンを使用して情報が表示され、かつ、ユーザから入力が受け取られ得る。他の種類のデバイスを使用してユーザとの対話が提供されることも可能であり、例えばユーザに提供されるフィードバックは、任意の形態の知覚フィードバック、例えば視覚フィードバック、聴覚フィードバック又は触覚フィードバックであってもよく、また、ユーザからの入力は、音響、音声又は触覚入力を含む任意の形態で受け取られ得る。さらに、コンピュータは、ユーザが使用しているデバイスに文書を送り、また、ユーザが使用しているデバイスから文書を受け取ることによってユーザと対話することができ、例えばウェブブラウザから受け取った要求に応答してユーザのクライアントデバイス上のウェブブラウザにウェブページを送ることによって対話することができる。

【0141】

[00167]本明細書において説明されている主題の実施形態は、例えばデータサーバとしてのバックエンド構成要素を含む計算システム、又はミドルウェア構成要素、例えばアプリケーション・サーバを含む計算システム、あるいはフロントエンド構成要素、例えばユーザが本明細書において説明されている主題の実施態様と対話することができるグラフィカル・ユーザ・インタフェース又はウェブブラウザを有するクライアント・コンピュータを含む計算システム、あるいは1つ又は複数のこのようなバックエンド構成要素、ミドルウェア構成要素又はフロントエンド構成要素の任意の組合せを含む計算システムの中で実現され得る。システムの構成要素は、デジタルデータ通信の任意の形態の媒介物、例えば通信ネットワークによって相互接続され得る。通信ネットワークの例には、ローカル・エリア・ネットワーク(「LAN」)及び広域ネットワーク(「WAN」)、インターネット(例えばInternet)及びピア・トゥ・ピアネットワーク(例えばアドホック・ピア・トゥ・ピアネットワーク)がある。例えば図1のネットワーク20は、1つ又は複数のローカル・エリア・ネットワークを含むことができる。

【0142】

[00168]計算システムは、任意の数のクライアント及びサーバを含むことができる。クライアント及びサーバは、通常、互いに遠隔に位置しており、典型的には通信ネットワークを介して対話する。クライアントとサーバの関係は、それぞれのコンピュータ上で走るコンピュータプログラムによって生じ、互いにクライアント・サーバ関係を有している。いくつかの実施形態では、サーバは、データ(例えばHTMLページ)をクライアントデバイスに伝送する(例えばクライアントデバイスと対話しているユーザにデータを表示し、また、ユーザからユーザ入力を受け取るために)。クライアントデバイスで生成されたデータ(例えばユーザ対話の結果)は、サーバでクライアントデバイスから受け取られ得る。

【0143】

[00169]以上、本明細書における開示について、特定の実施形態を参照して説明したが、これらの実施形態は、本開示の原理及び用途についての単なる例証にすぎないことを理解されたい。したがって添付の特許請求の範囲で定義されている本開示の精神及び範囲を逸脱することなく、多くの修正が実例実施形態に加えられ得ること、また、他の構造が工夫され得ることを理解されたい。書面にされた上記明細書は、当業者による本開示の実践を可能にするのに十分であると見なされる。

10

20

30

40

【図 1 A】

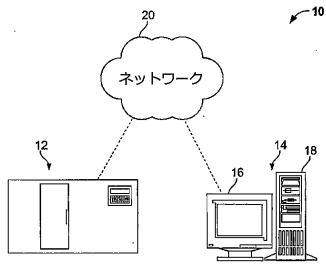


FIG. 1A

【図 1 B】

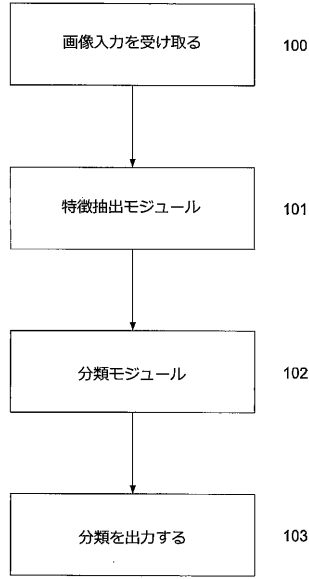


FIG. 1B

【図 1 C】

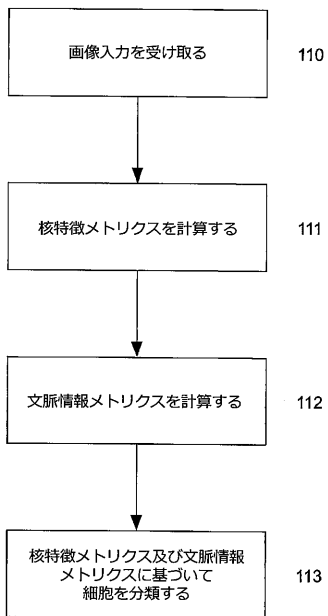
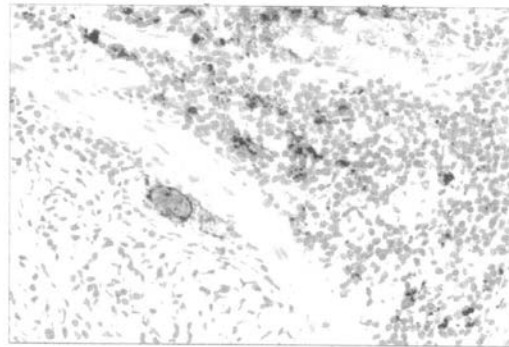


FIG. 1C

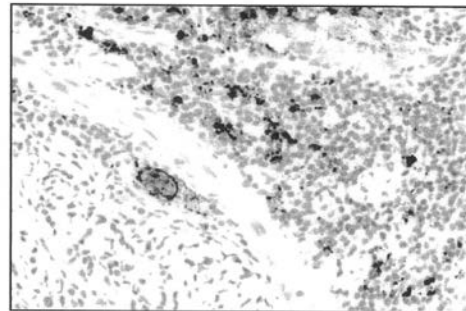
【図 2 A】

FIG. 2A



【図 2 B】

FIG. 2B

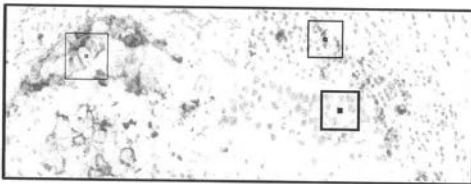


【 図 3 】

FIG. 3A

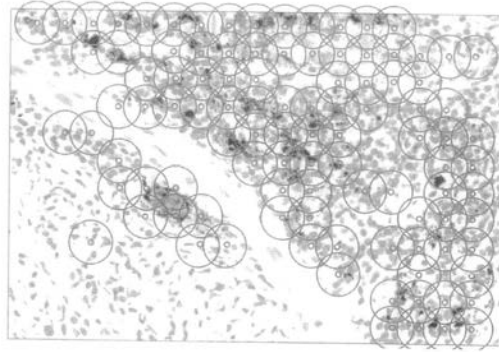


FIG. 3B



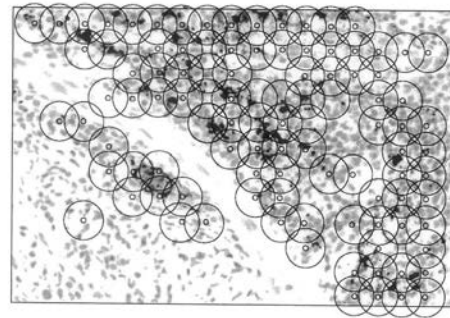
【 図 4 A 】

FIG. 4A



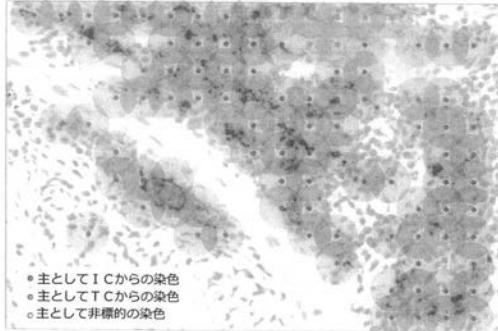
【 図 4 B 】

FIG. 4B



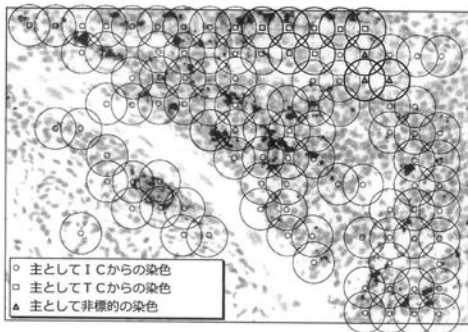
【 図 5 A 】

FIG. 5A



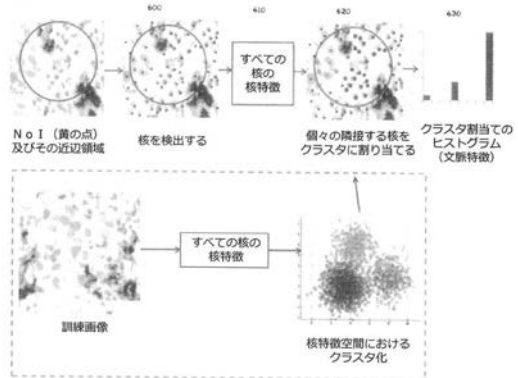
【 図 5 B 】

FIG. 5B



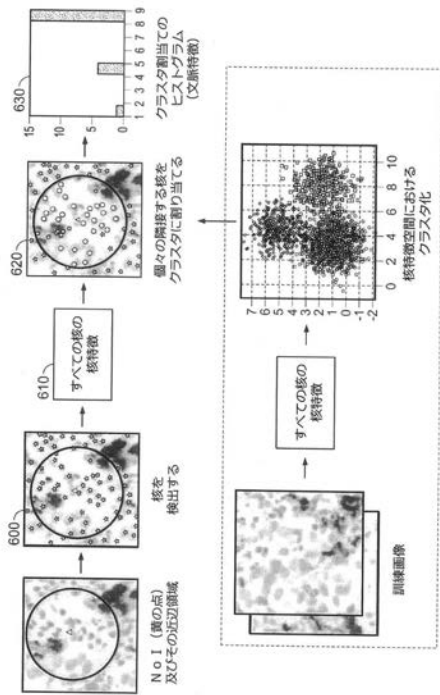
【 図 6 A 】

FIG. 6A



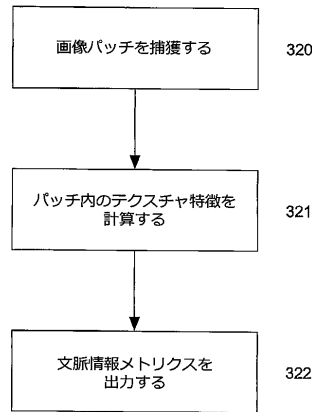
【 図 6 B 】

FIG. 6B



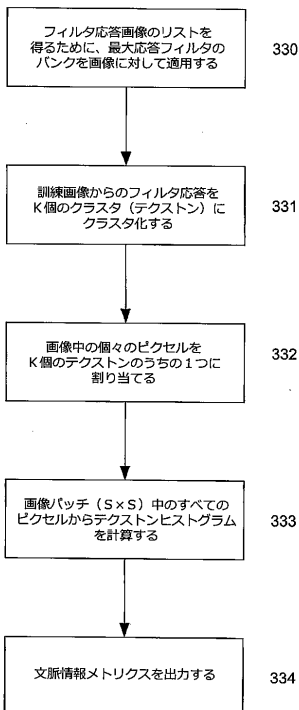
【 図 6 C 】

FIG. 6C



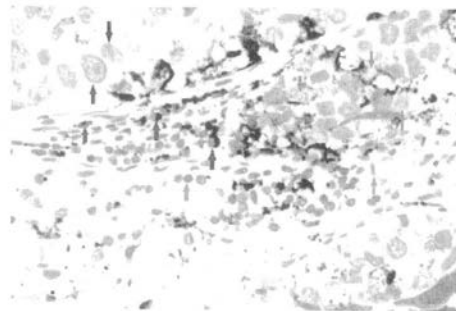
【 図 6 D 】

FIG. 6D



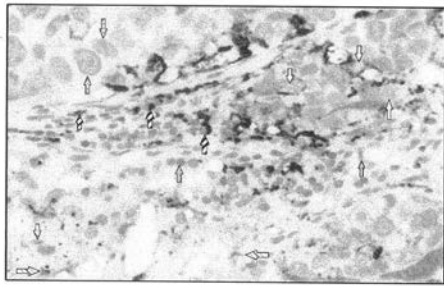
【 図 7 A 】

FIG. 7A



【 図 7 B 】

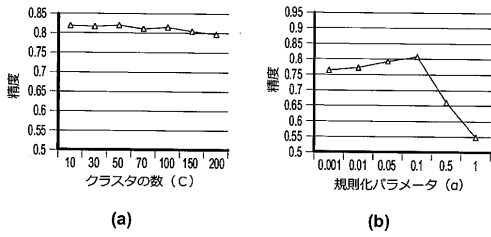
FIG. 7B



- ↓ A
- ↓ B
- ↓ C
- ↓ D
- ↓ E

【 図 8 】

FIG. 8



【 図 9 - 2 】

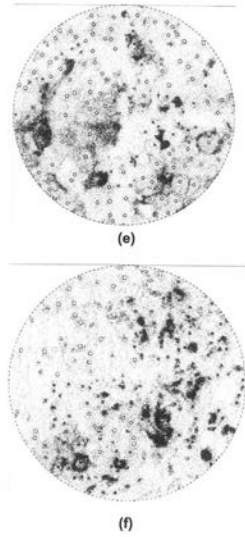
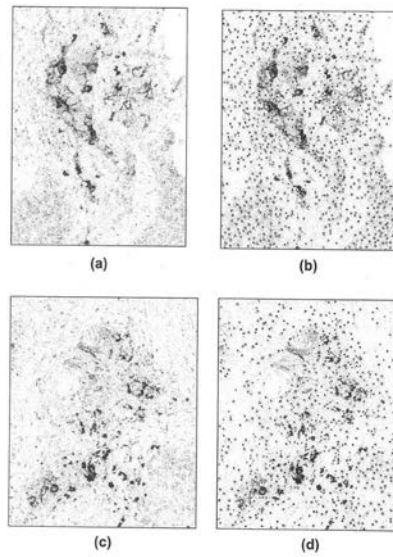


図9 (続き)

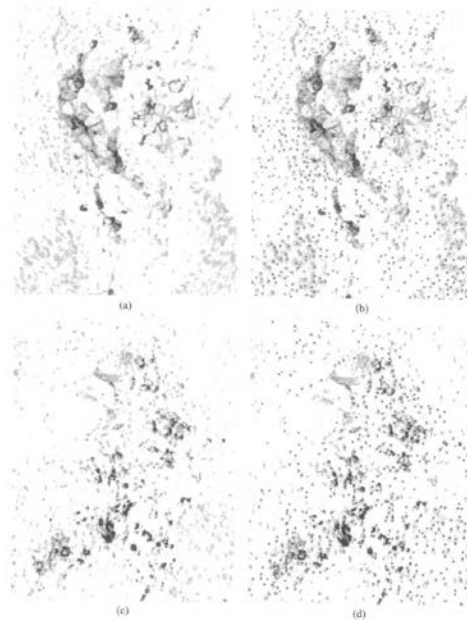
【 図 9 - 1 】

FIG. 9



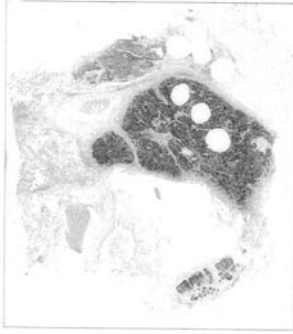
【 図 10 】

FIG. 10



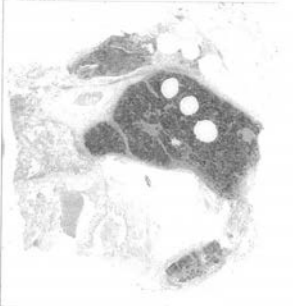
【 図 1 1 A 】

FIG. 11A



【 図 1 1 B 】

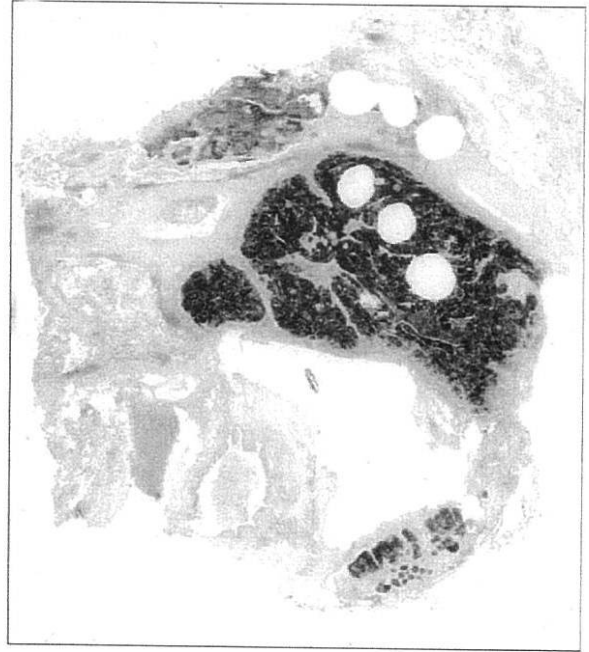
FIG. 11B



• PD-L1 陽性免疫細胞を有する面積 = 3.2 mm²
• 腫瘍面積 = 43.9 mm²

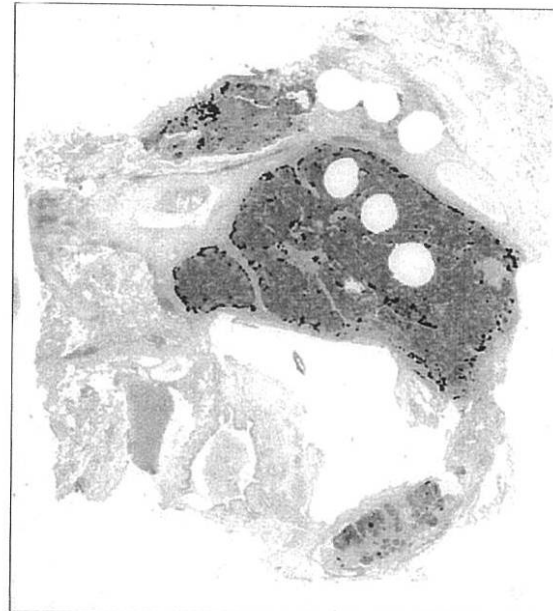
【 図 1 1 C 】

FIG. 11C



【 図 1 1 D 】

FIG. 11D



• PD-L1 陽性免疫細胞を有する面積 = 3.2 mm²
• 腫瘍面積 = 43.9 mm²

【 国際調査報告 】

INTERNATIONAL SEARCH REPORT

International application No PCT/EP2015/076105

A. CLASSIFICATION OF SUBJECT MATTER		
INV. G06K9/00 G06K9/46 ADD. G06K9/62		
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED		
Minimum documentation searched (classification system followed by classification symbols) G06K		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) EPO-Internal		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2014/165422 A1 (MERCK SHARP & DOHME [US]; DOLLED-FILHART MARISA [US]; EMANCIPATOR KENN) 9 October 2014 (2014-10-09) cited in the application	1-5, 16, 19-23, 36-44, 46
Y	the whole document ----- -/--	6-11, 15, 24-29, 33
<input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/> See patent family annex.		
* Special categories of cited documents :		
"A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier application or patent but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed		"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family
Date of the actual completion of the international search 8 February 2016		Date of mailing of the international search report 26/04/2016
Name and mailing address of the ISA/ European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016		Authorized officer Hermes, Lothar

INTERNATIONAL SEARCH REPORT

International application No

PCT/EP2015/076105

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	MARCIAL GARCI-A ROJO ET AL: "Review of imaging solutions for integrated quantitative immunohistochemistry in the Pathology daily practice", FOLIA HISTOCHEMICA ET CYTOBIOLOGICA, vol. 47, no. 3, 19 February 2010 (2010-02-19), pages 349-354, XP055247918, PL ISSN: 0239-8508, DOI: 10.2478/v10042-008-0114-4	1-5,16, 19-23, 36-44,46
Y	abstract page 352, column 1, lines 15-24 page 351, column 2, lines 7-9	6-11,15, 24-29,33
Y	----- ULAS, AYDIN ET AL.: "Hybrid Generative-Discriminative Nucleus Classification of Renal Cell Carcinoma", SIMILARITY-BASED PATTERN RECOGNITION, SPRINGER BERLIN HEIDELBERG, BERLIN, HEIDELBERG, vol. 7005, 28 September 2011 (2011-09-28), pages 77-89, XP019166550, ISBN: 978-3-642-24470-4 the whole document	6,9,10, 24,27,28
Y	----- ONCEL TUZEL ET AL: "Classification of hematologic malignancies using texton signatures", PATTERN ANALYSIS AND APPLICATIONS, SPRINGER-VERLAG, LO, vol. 10, no. 4, 30 March 2007 (2007-03-30), pages 277-290, XP019545403, ISSN: 1433-755X, DOI: 10.1007/S10044-007-0066-X the whole document	7,8,11, 15, 24-26, 29,33
A	----- THOMPSON R HOUSTON ET AL: "Costimulatory B7-H1 in renal cell carcinoma patients: Indicator of tumor aggressiveness and potential therapeutic target", PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES, NATIONAL ACADEMY OF SCIENCES, US, vol. 101, no. 49, 7 December 2004 (2004-12-07), pages 17174-17179, XP002520405, ISSN: 0027-8424, DOI: 10.1073/PNAS.0406351101 the whole document	1,3,9, 19,27, 38,39, 41,46
	----- -/--	

INTERNATIONAL SEARCH REPORT

International application No

PCT/EP2015/076105

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	<p>CHUAN-YONG MU ET AL: "High expression of PD-L1 in lung cancer may contribute to poor prognosis and tumor cells immune escape through suppressing tumor infiltrating dendritic cells maturation", MEDICAL ONCOLOGY, SPRINGER-VERLAG, NEW YORK, vol. 28, no. 3, 6 April 2010 (2010-04-06), pages 682-688, XP019937695, ISSN: 1559-131X, DOI: 10.1007/S12032-010-9515-2 the whole document</p> <p>-----</p>	1,3,9, 19,27, 38,39, 41,46
A	<p>D. S. CHEN ET AL: "Molecular Pathways: Next-Generation Immunotherapy--Inhibiting Programmed Death-Ligand 1 and Programmed Death-1", CLINICAL CANCER RESEARCH, vol. 18, no. 24, 19 October 2012 (2012-10-19), pages 6580-6587, XP055124329, ISSN: 1078-0432, DOI: 10.1158/1078-0432.CCR-12-1362 page 6580 - page 6581 page 6584 - page 6585</p> <p>-----</p>	1,3,9, 19,27, 38,39, 41,46
A	<p>MARY E. KEIR ET AL: "PD-1 and Its Ligands in Tolerance and Immunity", ANNUAL REVIEW OF IMMUNOLOGY., vol. 26, no. 1, 2 January 2008 (2008-01-02), pages 677-704, XP055099324, US ISSN: 0732-0582, DOI: 10.1146/annurev.immunol.26.021607.090331 pages 692-695</p> <p>-----</p>	1,3,9, 19,27, 38,39, 41,46
A	<p>ATUPELAGE CHAMIDU ET AL: "Computational hepatocellular carcinoma tumor grading based on cell nuclei classification", JOURNAL OF MEDICAL IMAGING, SOCIETY OF PHOTO-OPTICAL INSTRUMENTATION ENGINEERS, 1000 20TH ST. BELLINGHAM WA 98225-6705 USA, vol. 1, no. 3, October 2014 (2014-10), page 34501, XP060047793, ISSN: 2329-4302, DOI: 10.1117/1.JMI.1.3.034501 [retrieved on 2014-10-09] the whole document</p> <p>-----</p>	1-5,19, 21-23, 41-44,46
	----- -/--	

INTERNATIONAL SEARCH REPORT

International application No

PCT/EP2015/076105

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	<p>JUN KONG ET AL: "A comprehensive framework for classification of nuclei in digital microscopy imaging: An application to diffuse gliomas", 2011 8TH IEEE INTERNATIONAL SYMPOSIUM ON BIOMEDICAL IMAGING: FROM NANO TO MACRO (ISBI 2011), IEEE, UNITED STATES, 30 March 2011 (2011-03-30), pages 2128-2131, XP031944961, DOI: 10.1109/ISBI.2011.5872833 ISBN: 978-1-4244-4127-3 cited in the application Sect. 2.2 figure 2</p>	1-5,19, 21-23, 41-43,46
A	<p>----- NGUYEN KIEN ET AL: "Prostate Cancer Grading: Use of Graph Cut and Spatial Arrangement of Nuclei", IEEE TRANSACTIONS ON MEDICAL IMAGING, IEEE SERVICE CENTER, PISCATAWAY, NJ, US, vol. 33, no. 12, 10 July 2014 (2014-07-10), pages 2254-2270, XP011565182, ISSN: 0278-0062, DOI: 10.1109/TMI.2014.2336883 [retrieved on 2014-11-25] Sect. IV.A</p>	1-6,11, 15, 19-24, 29,33, 41-44
A,P	<p>----- KIEN NGUYEN ET AL: "Using contextual information to classify nuclei in histology images", 2015 IEEE 12TH INTERNATIONAL SYMPOSIUM ON BIOMEDICAL IMAGING (ISBI), 16 April 2015 (2015-04-16), pages 995-998, XP055243377, DOI: 10.1109/ISBI.2015.7164038 ISBN: 978-1-4799-2374-8 the whole document</p> <p>-----</p>	1-11,15, 16,19, 29,33, 36,37, 41,44,46

INTERNATIONAL SEARCH REPORT

International application No.
PCT/EP2015/076105**Box No. II Observations where certain claims were found unsearchable (Continuation of Item 2 of first sheet)**

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
2. Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of Item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. As all searchable claims could be searched without effort justifying an additional fees, this Authority did not invite payment of additional fees.
3. As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

1-3, 5-11, 15, 16, 19-21, 23-29, 33, 36-41, 43, 44, 46(completely); 4, 22
42(partially)**Remark on Protest**

- The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- No protest accompanied the payment of additional search fees.

International Application No. PCT/ EP2015/ 076105

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. claims: 1-3, 5-11, 15, 16, 19-21, 23-29, 33, 36-41, 43, 44, 46(completely); 4, 22, 42(partially)

Computer systems and computer-implemented methods for cell classification according to W02014/165422

- 1.1. claims: 6-11, 24-29

Computer systems and computer-implemented methods for cell classification, wherein texture features are used, which are not mentioned in W02014/165422

- 1.2. claims: 15, 33

Computer system and computer-implemented method for cell classification, wherein cells are classified with a support vector machine

2. claims: 12-14, 17, 18, 30-32, 34, 35, 45(completely); 4, 22, 42(partially)

Computer systems and computer-implemented methods for cell classification, wherein cells are classified based on data describing neighboring nuclei

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/EP2015/076105

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2014165422 A1	09-10-2014	EP 2981821 A1	10-02-2016
		US 2016084839 A1	24-03-2016
		WO 2014165422 A1	09-10-2014

フロントページの続き

(51) Int. Cl.		F I		テーマコード (参考)	
C 1 2 Q	1/06	(2006.01)	C 1 2 M	1/34	A
G 0 6 F	19/24	(2011.01)	C 1 2 Q	1/06	
			G 0 6 F	19/24	

(81) 指定国 AP(BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), EA(AM, AZ, BY, KG, KZ, RU, TJ, TM), EP(AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OA(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG), AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US

(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 シェフトテル, 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

ベトナム国ホーチミン・シティ, グエン・フー・ソー・ストリート 9, プー・ホアン・アン・アパートメント

F ターム (参考) 2G045 AA26 CB01 DA13 FA19 JA01

4B029 AA07 BB11 BB17 CC03 CC08 FA02 FA03

4B063 QA05 QA19 QQ08 QQ79 QR48 QR77 QS33 QS36 QS39 QX01

5L096 AA02 AA06 BA06 BA13 DA01 EA43 FA14 FA15 FA33 FA34

FA35 FA54 FA59 FA65 FA66 FA67 FA72 GA02 GA19 GA30

GA34 GA40 GA41 HA09 JA11 KA04 MA07

专利名称(译)	组织学图像中细胞核的分类		
公开(公告)号	JP2018502279A	公开(公告)日	2018-01-25
申请号	JP2017525014	申请日	2015-11-09
[标]申请(专利权)人(译)	文塔纳医疗系统公司		
申请(专利权)人(译)	本塔纳医疗系统公司		
[标]发明人	ブレドノイエルク シェフドテルクリストフ グエンキエン		
发明人	ブレドノ,イエルク シェフドテル,クリストフ グエン,キエン		
IPC分类号	G01N33/48 G01N33/53 G06T7/00 G01N33/574 C12M1/34 C12Q1/06 G06F19/24		
CPC分类号	G06K9/0014 G06K9/00147 G06K9/4642 G06K9/6223 G06K9/6269 G06K9/20 G06K9/4604 G06T7/00		
FI分类号	G01N33/48.M G01N33/53.Y G06T7/00.630 G01N33/574.A G01N33/48.P C12M1/34.A C12Q1/06 G06F19/24		
F-TERM分类号	2G045/AA26 2G045/CB01 2G045/DA13 2G045/FA19 2G045/JA01 4B029/AA07 4B029/BB11 4B029/BB17 4B029/CC03 4B029/CC08 4B029/FA02 4B029/FA03 4B063/QA05 4B063/QA19 4B063/QQ08 4B063/QQ79 4B063/QR48 4B063/QR77 4B063/QS33 4B063/QS36 4B063/QS39 4B063/QX01 5L096/AA02 5L096/AA06 5L096/BA06 5L096/BA13 5L096/DA01 5L096/EA43 5L096/FA14 5L096/FA15 5L096/FA33 5L096/FA34 5L096/FA35 5L096/FA54 5L096/FA59 5L096/FA65 5L096/FA66 5L096/FA67 5L096/FA72 5L096/GA02 5L096/GA19 5L096/GA30 5L096/GA34 5L096/GA40 5L096/GA41 5L096/HA09 5L096/JA11 5L096/KA04 5L096/MA07		
代理人(译)	山本修 宮前徹 中西 基晴 健一牌坊		
优先权	62/077536 2014-11-10 US 62/144364 2015-04-08 US		
外部链接	Espacenet		

摘要(译)

(1) 提供组织样本的图像作为输入，(2) 根据图像中核的特征计算核特征度量 (111)，并且 (3) 使用图像将核建立在目标上。用 (112) 和 (4) 使用核特征量度和上下文信息量度的组合来对图像中的细胞进行分类来计算上下文信息量度 (113)。公开了用于对细胞进行分类的计算机设备 (14) 和计算机实现的方法。

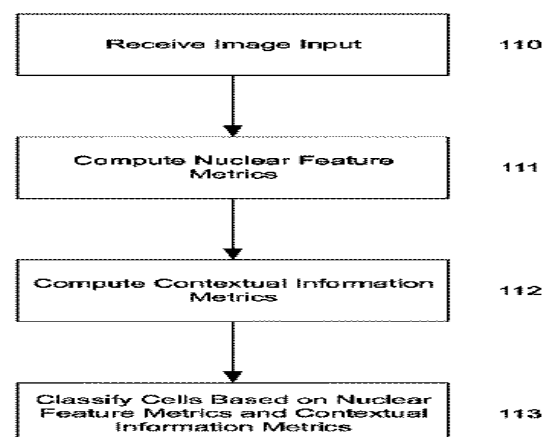


FIG. 1C

