

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7599552号
(P7599552)

(45)発行日 令和6年12月13日(2024.12.13)

(24)登録日 令和6年12月5日(2024.12.5)

(51)国際特許分類

F I

G 0 6 T	7/00	(2017.01)	G 0 6 T	7/00	6 3 0
G 0 6 V	10/82	(2022.01)	G 0 6 T	7/00	3 5 0 C
G 0 6 V	20/69	(2022.01)	G 0 6 V	10/82	
G 0 1 N	33/48	(2006.01)	G 0 6 V	20/69	
			G 0 1 N	33/48	M

請求項の数 20 (全29頁) 最終頁に続く

(21)出願番号	特願2023-509767(P2023-509767)
(86)(22)出願日	令和3年7月29日(2021.7.29)
(65)公表番号	特表2023-537978(P2023-537978 A)
(43)公表日	令和5年9月6日(2023.9.6)
(86)国際出願番号	PCT/US2021/043717
(87)国際公開番号	WO2022/035609
(87)国際公開日	令和4年2月17日(2022.2.17)
審査請求日	令和5年2月13日(2023.2.13)
(31)優先権主張番号	63/065,404
(32)優先日	令和2年8月13日(2020.8.13)
(33)優先権主張国・地域又は機関	米国(US)

(73)特許権者	507179346 ペンタナ 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 8 5 7 5 5 , トゥーソン, イースト・イノベーショ ン・パーク・ドライブ 1 9 1 0
(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	100196508 弁理士 松尾 淳一
(74)代理人	100173565 弁理士 末松 亮太
(72)発明者	フェローズ, ハドリー

最終頁に続く

(54)【発明の名称】 デジタル病理学のためのアクティブラーニングシステム

(57)【特許請求の範囲】

【請求項 1】

デジタル病理システムにおけるアクティブラーニングのための方法であって、
 1つ以上のパラメータを使用して画像を処理してラベルを生成するように構成された機械学習モデルにアクセスすることと、
 前記機械学習モデルを実行することであって、
 少なくとも1つのデジタル病理画像のそれぞれの少なくとも一部を複数の予測ラベルに変換し、
 前記複数の予測ラベルのそれぞれについて信頼性メトリックを生成する
 ように、前記機械学習モデルを実行することと、
 前記少なくとも1つのデジタル病理画像のそれぞれについて、前記デジタル病理画像のための確実性マップを生成することであって、
 前記確実性マップが値のセットを含み、前記値のセットの各値が、(i)前記デジタル病理画像の部分に対応し、(ii)前記部分内の点の位置または領域に対応する1つ以上の信頼性メトリックの統計値を識別する、
 生成することと、
 前記少なくとも1つのデジタル病理画像のそれぞれの前記少なくとも一部を描写して、対応する信頼性メトリックに基づいて予測ラベルを差別的に表すインターフェースを利用することであって、
 前記値のセットに基づく前記確実性マップの表現を利用することを含む、

10

20

前記インターフェースを利用することと、

前記インターフェースの利用に応じて、前記複数の予測ラベルのうちの少なくとも1つを確認、拒否、または置換するラベル入力を受け取ることと、

前記ラベル入力に基づいて前記機械学習モデルの前記1つ以上のパラメータを更新することと

を含む、方法。

【請求項2】

前記複数の予測ラベルのうちの予測ラベルが、前記少なくとも1つのデジタル病理画像の前記少なくとも一部内の細胞の分類および/または特定のタイプの細胞の推定量を識別する、請求項1に記載の方法。

10

【請求項3】

前記インターフェースが、

1つの拡大レベルにおいて信頼性メトリックに関する統計値を表すためのユーザ入力を受け取り、

別の拡大レベルにおいて信頼性メトリックを表し、

前記拡大レベルを変更するためのユーザ入力を受け取る、

ように構成される、請求項1または請求項2に記載の方法。

【請求項4】

新たなデジタル病理画像を受信することと、

前記新たなデジタル病理画像を、前記新たなデジタル病理画像または前記新たなデジタル病理画像の一部を特徴付ける1つ以上の新たなラベルに変換することと、

20

前記1つ以上の新たなラベルを出力することと

をさらに含む、請求項1～3のいずれか一項に記載の方法。

【請求項5】

前記デジタル病理画像の初期バージョンを前処理して、各関心のある特徴の位置を検出することをさらに含み、

前記デジタル病理画像の前記少なくとも一部を変換することが、前記デジタル病理画像の前記少なくとも一部内の各検出された関心のある位置に対応するラベルを生成することを含む、請求項1～4のいずれか一項に記載の方法。

【請求項6】

30

前記複数の予測ラベルのそれぞれについて、前記予測ラベルの前記信頼性メトリックに基づいてマーカーのサイズ、色、または形状を選択することをさらに含み、

前記インターフェースが、前記選択されたサイズ、色、または形状を有するマーカーを使用して、前記複数の予測ラベルの、マーカーを使用した各予測ラベルを表すことにより、対応する信頼性メトリックに基づいて予測ラベルを差別的に表す、請求項1～5のいずれか一項に記載の方法。

【請求項7】

前記機械学習モデルが深層ニューラルネットワークを含む、請求項1～6のいずれか一項に記載の方法。

【請求項8】

40

システムであって、

1つ以上のデータプロセッサと、

命令を含む非一時的コンピュータ可読記憶媒体であって、前記命令が、前記1つ以上のデータプロセッサ上で実行されると、前記1つ以上のデータプロセッサに、動作のセットを実行させる、非一時的コンピュータ可読記憶媒体と、を備え、

前記動作のセットが、

1つ以上のパラメータを使用して画像を処理してラベルを生成するように構成された機械学習モデルにアクセスすることと、

前記機械学習モデルを実行することであって、

少なくとも1つのデジタル病理画像のそれぞれの少なくとも一部を複数の予測ラベ

50

ルに変換し、

前記複数の予測ラベルのそれぞれについて信頼性メトリックを生成する

ように、前記機械学習モデルを実行することと、

前記少なくとも1つのデジタル病理画像のそれぞれについて、前記デジタル病理画像のための確実性マップを生成することであって、

前記確実性マップが値のセットを含み、前記値のセットの各値が、(i)前記デジタル病理画像の部分に対応し、(ii)前記部分内の点の位置または領域に対応する1つ以上の信頼性メトリックの統計値を識別する、

生成することと、

前記少なくとも1つのデジタル病理画像のそれぞれの前記少なくとも一部を描写して、対応する信頼性メトリックに基づいて予測ラベルを差別的に表すインターフェースを利用することであって、

前記値のセットに基づく前記確実性マップの表現を利用することを含む、

前記インターフェースを利用することと、

前記インターフェースの利用に応じて、前記複数の予測ラベルのうちの少なくとも1つを確認、拒否、または置換するラベル入力を受け取ることと、

前記ラベル入力に基づいて前記機械学習モデルの前記1つ以上のパラメータを更新すること

を含む、システム。

【請求項9】

前記複数の予測ラベルのうちの予測ラベルが、前記少なくとも1つのデジタル病理画像の前記少なくとも一部内の細胞の分類および/または特定のタイプの細胞の推定量を識別する、請求項8に記載のシステム。

【請求項10】

前記インターフェースが、

1つの拡大レベルにおいて信頼性メトリックに関する統計値を表すためのユーザ入力を受け取り、

別の拡大レベルにおいて信頼性メトリックを表し、

前記拡大レベルを変更するためのユーザ入力を受け取る、

ように構成される、請求項8または請求項9に記載のシステム。

【請求項11】

前記動作のセットが、

新たなデジタル病理画像を受け取ることと、

前記新たなデジタル病理画像を、前記新たなデジタル病理画像または前記新たなデジタル病理画像の一部を特徴付ける1つ以上の新たなラベルに変換することと、

前記1つ以上の新たなラベルを出力することと

をさらに含む、請求項8～10のいずれか一項に記載のシステム。

【請求項12】

前記動作のセットが、

前記デジタル病理画像の初期バージョンを前処理して、各関心のある特徴の位置を検出すること

をさらに含み、前記デジタル病理画像の前記少なくとも一部を変換することが、前記デジタル病理画像の前記少なくとも一部内の各検出された関心のある位置に対応するラベルを生成することを含む、請求項8～11のいずれか一項に記載のシステム。

【請求項13】

前記動作のセットが、

前記複数の予測ラベルのそれぞれについて、前記予測ラベルの前記信頼性メトリックに基づいてマーカーのサイズ、色、または形状を選択すること

をさらに含み、前記インターフェースが、前記選択されたサイズ、色、または形状を有するマーカーを使用して、前記複数の予測ラベルの、マーカーを使用した各予測ラベルを

10

20

30

40

50

表すことにより、対応する信頼性メトリックに基づいて予測ラベルを差別的に表す、請求項 8 ~ 11 のいずれか一項に記載のシステム。

【請求項 14】

前記機械学習モデルが深層ニューラルネットワークを含む、請求項 8 ~ 13 のいずれか一項に記載のシステム。

【請求項 15】

1 つ以上のデータプロセッサに、動作のセットを実行させるように構成された命令を記録した非一時的機械可読記憶媒体であって、

前記動作のセットが、

1 つ以上のパラメータを使用して画像を処理してラベルを生成するように構成された機械学習モデルにアクセスすることと、

前記機械学習モデルを実行することであって、

少なくとも 1 つのデジタル病理画像のそれぞれの少なくとも一部を複数の予測ラベルに変換し、

前記複数の予測ラベルのそれぞれについて信頼性メトリックを生成する

ように、前記機械学習モデルを実行することと、

前記少なくとも 1 つのデジタル病理画像のそれぞれについて、前記デジタル病理画像のための確実性マップを生成することであって、

前記確実性マップが値のセットを含み、

前記値のセットの各値が、(i) 前記デジタル病理画像の部分に対応し、(i i) 前記部分内の点の位置または領域に対応する 1 つ以上の信頼性メトリックの統計値を識別する、

生成することと、

前記少なくとも 1 つのデジタル病理画像のそれぞれの前記少なくとも一部を描写して、対応する信頼性メトリックに基づいて予測ラベルを差別的に表すインターフェースを利用することであって、

前記値のセットに基づく前記確実性マップの表現を利用することを含む、

前記インターフェースを利用することと、

前記インターフェースの利用に応じて、前記複数の予測ラベルのうちの少なくとも 1 つを確認、拒否、または置換するラベル入力を受け取ることと、

前記ラベル入力に基づいて前記機械学習モデルの前記 1 つ以上のパラメータを更新することと

を含む、非一時的機械可読記憶媒体。

【請求項 16】

前記複数の予測ラベルのうちの予測ラベルが、前記少なくとも 1 つのデジタル病理画像の前記少なくとも一部内の細胞の分類および / または特定のタイプの細胞の推定量を識別する、請求項 15 に記載の非一時的機械可読記憶媒体。

【請求項 17】

前記インターフェースが、

1 つの拡大レベルにおいて信頼性メトリックに関する統計値を表すためのユーザ入力を受け取り、

別の拡大レベルにおいて信頼性メトリックを表し、

前記拡大レベルを変更するためのユーザ入力を受け取る、

ように構成される、請求項 15 または請求項 16 に記載の非一時的機械可読記憶媒体。

【請求項 18】

前記動作のセットが、

新たなデジタル病理画像を受け取ることと、

前記新たなデジタル病理画像を、前記新たなデジタル病理画像または前記新たなデジタル病理画像の一部を特徴付ける 1 つ以上の新たなラベルに変換することと、

前記 1 つ以上の新たなラベルを出力することと

10

20

30

40

50

をさらに含む、請求項 15 ~ 17 のいずれか一項に記載の非一時的機械可読記憶媒体。

【請求項 19】

前記動作のセットが、

前記デジタル病理画像の初期バージョンを前処理して、各関心のある特徴の位置を検出することをさらに含み、

前記デジタル病理画像の前記少なくとも一部を変換することが、前記デジタル病理画像の前記少なくとも一部内の各検出された関心のある位置に対応するラベルを生成することを含む、請求項 15 ~ 18 のいずれか一項に記載の非一時的機械可読記憶媒体。

【請求項 20】

前記動作のセットが、

前記複数の予測ラベルのそれぞれについて、前記予測ラベルの前記信頼性メトリックに基づいてマーカーのサイズ、色、または形状を選択することをさらに含み、

前記インターフェースが、前記選択されたサイズ、色、または形状を有するマーカーを使用して、前記複数の予測ラベルの、マーカーを使用した各予測ラベルを表すことにより、対応する信頼性メトリックに基づいて予測ラベルを差別的に表す、請求項 15 ~ 19 のいずれか一項に記載の非一時的機械可読記憶媒体。

【発明の詳細な説明】

【技術分野】

【0001】

関連出願の相互参照

本出願は、2020年8月13日に提出された米国仮特許出願第63/065,404号の利益および優先権を主張するものであり、その全体があらゆる目的のために参照により本明細書に組み込まれる。

【0002】

分野

本開示は、デジタル病理、特に、デジタル病理画像の一部または全部を自動的に検出、特徴付けおよび/または分類するために機械学習モデルを効率的に訓練するためにアクティブラーニングを使用するための技術に関する。

【背景技術】

【0003】

背景

デジタル病理は、スライド（例えば、組織病理または細胞病理ガラススライド）をデジタル画像にスキャンすることを含む。デジタル画像内の組織および/または細胞は、その後、疾患の診断、治療に対する応答の評価、および疾患と戦うための薬剤の開発を含む様々な理由のために、デジタル病理画像分析を使用して検査され、および/または病理学者によって解釈されることができる。デジタル画像（実質的に透明である）内の組織および/または細胞を検査するために、病理スライドは、特定の組織および/または細胞成分に選択的に結合する様々な染色アッセイ（例えば、免疫染色）を使用して調製されることができる。

【0004】

デジタル病理画像を分析するための機械学習モデルを訓練することは、グラウンドトゥールズで手動でラベル付けされた画像の大きなセットを必要とする。例えば、手動ラベリングは、画像内の特定の領域内の所与のタイプ（例えば、腫瘍細胞）の各細胞の位置（例えば、点の位置および/または境界）を識別することを含むことができる。これらのラベル付き画像を生成することは、収集するのに面倒で時間がかかる。さらにまた、訓練画像を取得することは、プライバシー上の懸念のために困難であり得る。

【0005】

現在、訓練データセットを準備するためにグラウンドトゥールズを収集するために、ラベル付けされるべき画像の選択または画像の選択部分がデータのプール（例えば、利用可能な画像）からランダムに選択される。しかしながら、ラベル付けされるべき画像または

10

20

30

40

50

部分または画像をランダムに選ぶことは、効率的な手法ではない。ランダムに選択されたサンプルである画像または画像の部分は、機械学習モデルの訓練において最も有益なものではない可能性があり、関心のある描写を欠く可能性がある。したがって、ランダムに選択された画像または画像のランダムに選択された部分をラベル付けすることは、訓練プロセスにいかなる値も追加することなく、リソース（例えば、モデル訓練に關与する病理学者の時間および計算リソース）を浪費する可能性がある。

【発明の概要】

【0006】

概要

様々な実施形態の態様および特徴は、添付の図面を参照して例を説明することによってより明らかになるであろう。

10

【0007】

いくつかの実施形態では、デジタル病理システムにおけるアクティブラーニングのための方法が提供される。1つ以上のパラメータを使用して画像を処理してラベルを生成するように構成された機械学習モデルがアクセスされる。機械学習モデルは、少なくとも1つのデジタル病理画像のそれぞれの少なくとも一部を複数の予測ラベルに変換し、複数の予測ラベルのそれぞれについて信頼性メトリックを生成するように実行される。少なくとも1つのデジタル病理画像の少なくとも一部を描写し、対応する信頼性メトリックに基づいて予測ラベルを差別的に表すインターフェースが利用される。インターフェースの利用にตอบสนองして、複数の予測ラベルのうちの少なくとも1つを確認、拒否、または置換するラベル

20

【0008】

本方法は、1つ以上の追加の動作をさらに含むことができる。

【0009】

1つ以上の追加の動作は、デジタル病理画像の確信マップを生成することを含むことができ、デジタル病理画像の各部分の値は、部分内の点位置または領域に対応する1つ以上の信頼性メトリックの統計値に対応し、インターフェースを利用することは、確実性マップの表現を利用することを含む。

【0010】

インターフェースは、1つの拡大レベルで領域信頼性メトリックに関する統計値を表すための入力を受信し、別の拡大レベルで信頼性メトリックを表し、拡大レベルを変更するための入力を受信するように構成されることができる。

30

【0011】

1つ以上の追加の動作は、新たなデジタル病理画像を受信することと、新たなデジタル病理画像を、新たなデジタル病理画像または新たなデジタル病理画像の一部を特徴付ける1つ以上の新たなラベルに変換することと、1つ以上の新たなラベルを出力することと、を含むことができる。

【0012】

1つ以上の追加の動作は、デジタル病理画像の初期バージョンを前処理して各関心のある特徴の位置を検出することを含むことができ、デジタル病理画像の少なくとも一部を変換することは、デジタル病理画像の少なくとも一部内の各検出された関心のある位置に対応するラベルを生成することを含む。

40

【0013】

1つ以上の追加の動作は、複数の予測ラベルのそれぞれについて、予測ラベルの信頼性メトリックに基づいてマーカーのサイズ、色、または形状を選択することを含むことができ、インターフェースは、選択されたサイズ、色、または形状を有するマーカーを使用してマーカーを使用して複数の予測ラベルの各予測ラベルを表すことにより、対応する信頼性メトリックに基づいて予測ラベルを差別的に表す。

【0014】

50

機械学習モデルは、深層ニューラルネットワークを含むことができる。

【0015】

いくつかの実施形態では、本明細書に開示される1つ以上の技術の一部または全部を使用して訓練された機械学習モデルによって生成された結果に基づいて、ユーザによって被験者の診断を決定することと、診断に基づいて被験者に対して特定の治療を潜在的に選択、推奨および/または投与することと、を含む方法が提供される。

【0016】

いくつかの実施形態では、本明細書に開示される1つ以上の技術の一部または全部を使用して訓練された機械学習モデルによって生成された結果に基づいて、被験者を選択、推奨および/または投与するための治療をユーザによって決定することを含む方法が提供される。

10

【0017】

いくつかの実施形態では、本明細書に開示される1つ以上の技術の一部または全部を使用して訓練された機械学習モデルによって生成された結果に基づいて、被験者が臨床試験に参加する資格があるか、または臨床試験において被験者を特定の cohorts に割り当てる資格があるかどうかをユーザによって決定することを含む方法が提供される。

【0018】

いくつかの実施形態では、1つ以上のデータプロセッサと、命令を含む非一時的コンピュータ可読記憶媒体であって、命令が、1つ以上のデータプロセッサ上で実行されると、1つ以上のデータプロセッサに、本明細書に開示された1つ以上の方法の一部または全部を実行させる、非一時的コンピュータ可読記憶媒体と、を含む、システムが提供される。

20

【0019】

いくつかの実施形態では、非一時的機械可読記憶媒体に有形に具現化され、1つ以上のデータプロセッサに、本明細書に開示された1つ以上の方法の一部または全部を実行させるように構成された命令を含む、コンピュータプログラム製品が提供される。

【0020】

採用された用語および表現は、限定ではなく説明の用語として使用され、そのような用語および表現の使用において、示され、説明された特徴またはその一部の任意の均等物を除外する意図はないが、特許請求される本発明の範囲内で様々な修正が可能であることが認識される。したがって、特許請求の範囲に記載された本発明は、実施形態および任意の特徴によって具体的に開示されているが、本明細書に開示された概念の変更および変形は、当業者によってあてにされてもよく、そのような変更および変形は、添付の特許請求の範囲によって定義される本発明の範囲内にあると見なされることを理解されたい。

30

【0021】

特許または出願ファイルは、カラーで作成された少なくとも1つの図面を含む。カラー図面を含む本特許または特許出願公開の写しは、請求に応じて、必要な手数料を支払うことにより、特許庁によって提供されることになる。

【0022】

様々な実施形態の態様および特徴は、添付の図面を参照して例を説明することによってより明らかになるであろう。

40

【図面の簡単な説明】

【0023】

【図1】組織染色プロセスの例を示している。

【図2】機械学習モデルを使用してデジタル病理画像を処理するための(1つ以上のコンピューティングシステムを使用して実行される)プロセスフローを示している。

【図3】本開示のいくつかの態様にかかるアクティブラーニングワークフローを示している。

【図4】デジタル病理プロセスフローの一部または全部を使用するための例示的なワークフローを示している。

【図5】この点で使用されることができ例示的なシステムおよびアプリケーションを有

50

するネットワーク500を開示している。

【図6】図6A～図6Dは、機械学習モデルを訓練するためのラベルの受信を容易にするためのスライド画像全体の一部または全部の例示的な画像を示している。

【発明を実施するための形態】

【0024】

詳細な説明

特定の実施形態が説明されているが、これらの実施形態は例としてのみ提示されており、保護の範囲を限定することを意図するものではない。本明細書に記載された装置、方法、およびシステムは、様々な他の形態で具現化されることができる。さらにまた、保護の範囲から逸脱することなく、本明細書に記載の例示的な方法およびシステムの形態の様々な省略、置換、および変更を行うことができる。

10

【0025】

I. 概要

様々な実施形態では、機械学習モデル（例えば、深層学習または従来の機械学習モデル）を訓練するためにラベル付けされるデータセット内の特定の要素を動的に選択する、デジタル病理学のためのアクティブラーニングシステムが提供される。特定の要素は、所与のスライド、画像内の所与のバッチ、細胞の所与の描写などを含むことができる。選択された特定の要素は、（例えば、細胞描写の一部のみが標識のために選択されるように）データセット内の要素の不完全なサブセットを含むことができる。特定の要素の選択は、機械学習モデルを効率的に訓練して正確な予測を生成するのを支援することができる。例えば、特定の要素の選択は、データ要素がラベル付けのためにランダムに選択された場合に閾値精度を達成するために必要とされるラベル付きデータの一部（例えば、75%未満、50%未満または25%未満）のみで閾値精度を達成するようにモデルを訓練することを可能にすることができる。アクティブラーニングシステムの使用は、ランダムに選択された訓練データを使用する場合と比較して、より少ないラベル付きデータを使用してモデルを訓練することを可能にする。さらにまた、ラベル付けされるべき特定の要素の選択および機械学習モデルの訓練が反復的に実行されることができ、モデルの弱点が繰り返し検出されて改善される。

20

【0026】

デジタル病理アクティブラーニングシステムは、1人以上の病理学者または1人以上の画像サイエンティストと相互作用することができ、および/または1人以上の病理学者または1人以上の画像サイエンティストによって使用されることができる。例えば、1つ以上の時点のそれぞれにおいて、アクティブラーニングシステムは、ラベル付けのために選択された特定の要素を病理学者の装置に送信、提示、または識別することができる。次いで、病理学者は、機械学習モデルを訓練するために使用されることができるラベルを含む入力を提供することができる。次いで、訓練された機械学習モデルが使用されて、入力データ（例えば、病理学者とすることができるエンティティまたは異なるエンティティによって提供または識別される）を処理し、入力データに関連する装置に出力されるラベル予測を生成することができる。

30

【0027】

様々な実施形態では、アクティブラーニングシステムは、高解像度画像の記憶および取得、データの視覚化、対話型データ補正、ならびにアルゴリズム訓練および推論のための画像分析を含む。初期ラベル付きデータセットは、機械学習モデル（パラメータが初期化されていないモデル、または事前訓練もしくはランダム初期化プロセスを使用してパラメータが初期化されているモデル）を最初に訓練するために使用される。初期ラベル付きデータセットは、ランダムまたは擬似ランダムに選択された要素のラベルを含むことができる。次いで、モデルが使用されて、他の要素に対応する予測のセットを生成する。

40

【0028】

各予測は、予測ラベルおよびラベルに関する対応する信頼性を含むことができる。ラベル付けされる（またはラベルレビューのための）選択要素を識別し、および/または様々

50

な要素のラベル付け（またはラベルレビュー）が他の要素と比較して訓練を容易にすると予測される程度を表す（例えば、スケーリングされた視覚的表現またはテキストを使用する）インターフェースが、ユーザ（例えば、病理学者または画像サイエンティスト）にとって利用されることができる。例えば、スライド画像全体に対応するヒートマップが提示されることができ、色は、描写されたスライス内の様々な領域に関連する予測の信頼性を表す。インターフェースは、ユーザが所与の視野にズームして特定の領域を拡大し（例えば、ラベルを確認または拒否するため、ラベルを肯定的に識別するためなど）、予測ラベルをレビューすることを可能にすることができる。これらのユーザ入力を使用して、モデルの再訓練および/または訓練を継続することができる（例えば、モデルの1つ以上のパラメータを更新するために）。場合によっては、インターフェースは、ユーザ識別またはユーザ確認ラベルが高い（または最大信頼性）に関連付けられるように、および/または最も低い信頼性メトリック（例えば、視野内、スライド全体の画像内、またはデータセット内の他のものと比較して）に関連付けられた要素が顕著に識別される（例えば、所与の色、マーカータイプ、線タイプ、線幅などを介して）ように、それ自体を動的に更新する。

10

【0029】

場合によっては、モデルは、ユーザ入力のバッチに基づいて定期的にさらに訓練される（例えば、要素ラベルを識別するか、または予測ラベルにフィードバックを提供する）。この場合、モデルの最新バージョンからの出力に基づいて生成された予測ラベルを識別するために、インターフェースが定期的に更新または定期的に提供されてもよい。

20

【0030】

予測ラベルの生成と、最も不確かな予測に対応するユーザフィードバックの提供と、モデルの再訓練との間の反復は、所定の期間、所定の反復回数、またはモデルが所望の性能に到達するまで（その時点で、モデルが展開されることができる）継続することができる。分類器をさらに微調整するために、反復が後に継続されることができる。

【0031】

高性能画像サーバは、モデルを構築、訓練、および/または使用するための1つ以上のアプリケーションをサポートすることができる。モデルの訓練および推論からの結果が生成され、データストアに大規模に記憶されることができる。APIは、モデルのコントローラ（モデルを訓練および/または実行する）とデータストアとの間の相互作用を可能にすることができる。統合プラットフォームは、これらのコンポーネントの全てを統合するために使用されることができ、ユーザがウェブブラウザ内でモデルを訓練することを可能にする。

30

【0032】

II. 定義

本明細書に使用される場合、動作が何かに「基づく」場合、これは動作が、何かに少なくともいくらか、または少なくとも部分的に基づいていることを意味する。

【0033】

本明細書で使用される場合、「実質的に (substantially)」、「およそ (approximately)」、「および「約 (about)」という用語は、当業者によって理解されるように、大部分が指定されるものであるが、必ずしも完全には指定されないもの（および完全に指定されるものを含む）として定義される。任意の開示された実施形態では、「実質的に」、「およそ」、または「約」という用語は、指定されたものの「[割合]以内」で置き換えられることができ、割合は、0.1、2、5、および20%を含む。

40

【0034】

本明細書で使用される場合、「サンプル」、「生物学的サンプル」または「組織サンプル」という用語は、ウイルスを含む任意の生物学的材料および/または生物から得られる生体分子（例えば、タンパク質、ペプチド、核酸、脂質、炭水化物、またはそれらの組み合わせ）を含む任意のサンプルを指す。生物の他の例は、哺乳類（ヒト、ネコ、イヌ、ウ

50

マ、ウシ、およびブタなどの獣医動物、ならびにマウス、ラット、霊長類などの実験動物など)、昆虫、環形動物、クモ形類動物、有袋類、爬虫類、両生類、細菌、および菌類などを含む。生物学的サンプルは、組織サンプル(組織切片や組織の針生検など)、細胞サンプル(Pap塗抹検体もしくは血液塗抹検体などの細胞学的塗抹検体、またはマイクロダイセクションによって得られた細胞のサンプルなど)、または細胞分画、断片または細胞小器官(細胞を溶解し、遠心分離などによってそれらの成分を分離することによって得られる)を含む。生物学的サンプルの他の例は、血液、血清、尿、精液、糞便、脳脊髄液、間質液、粘膜、涙、汗、膿、生検組織(例えば、外科的生検または針生検によって得られる)、乳頭吸引物、耳垢、乳、腔液、唾液、ぬぐい液(頬スワブなど)、または最初の生物学的サンプルに由来する生体分子を含む任意の材料を含む。特定の実施形態では、本明細書で使用される「生物学的サンプル」という用語は、被験者から得られた腫瘍またはその一部から調製されたサンプル(均質化または液化されたサンプルなど)を指す。

10

【0035】

本明細書で使用される場合、「生物学的物質または構造体」という用語は、生物学的構造(例えば、細胞核、細胞膜、細胞質、染色体、DNA、細胞、細胞塊など)の全体または一部を含む天然材料または構造体を指す。

【0036】

本明細書で使用される場合、「デジタル病理画像」は、染色されたサンプルのデジタル画像を指す。

【0037】

本明細書で使用される場合、「アーチファクト」という用語は、外因性因子の結果としての調製された顕微鏡スライドにおける人工構造または組織変化を指す。アーチファクトは、例えば、外科的除去、固定、組織処理、包埋、セクションニング、染色、および取り付け手順中に発生する可能性がある。組織病理学におけるアーチファクトは、例えば、固定前アーチファクト、固定アーチファクト、骨組織に関連するアーチファクト、組織処理アーチファクト、切開に関連するアーチファクト、浮遊および取り付けに関連するアーチファクト、染色アーチファクト、取り付けアーチファクト、生物学的アーチファクト、撮像アーチファクトなどを含むことができる。例えば、固定前アーチファクトは、注入アーチファクト、スクイーズアーチファクト(例えば、組織の裂け目または組織のひだ)、硫化アーチファクト、デンプンアーチファクト、自己分解アーチファクトなどを含むことができる。固定アーチファクトは、例えば、ホルマリン顔料、水銀顔料、氷晶アーチファクト、凍結アーチファクト、ストリーミングアーチファクトなどを含むことができる。浮遊および取り付けに関連するアーチファクトは、例えば、断面の折り目およびしわ、汚染、気泡などを含むことができる。染色アーチファクトは、例えば、残留ワックス、エオシンへの酢酸の添加に関連するアーチファクト、ヘマトキシリンの媒染によるアーチファクト、ヘマトキシリンの蛍光光沢によるアーチファクトなどを含むことができる。取り付けアーチファクトは、例えば、残留水および気泡を含むことができる。

20

30

III. デジタル病理画像の生成

【0038】

組織学的染色は、関心のある特徴を強調し、生体サンプルの切片化された組織または細胞のコントラストを高めるために広く使用されている。例えば、染色は、特定のタイプの細胞をマークするために、および/または特定のタイプの核酸および/またはタンパク質にフラグを立てて顕微鏡検査を支援するために使用されることができる。次いで、染色されたサンプルが評価されて、サンプル中の目的の特徴の量(例えば、カウント、密度または発現レベルを含むことができる)および/または関心のある特徴の1つ以上の特徴(例えば、互いに対する、または他の特徴、形状特性などに対する関心のある特徴の位置)を決定または推定することができる。組織学的染色のプロセスは、固定、処理、包埋、切片化、染色および画像化などのいくつかの段階を含むことができる。

40

【0039】

いくつかの実施形態では、組織切片の免疫組織化学染色は、生物学的サンプル中の特定

50

のタンパク質の存在を同定するために使用される組織学的染色の一種である。例えば、特定のタンパク質（例えば、抗原）の発現レベルは、以下によって決定される：（a）特定の抗体タイプを有するサンプルの免疫組織化学分析を実施すること；（b）サンプル中のタンパク質の存在および/または発現レベルを決定すること。いくつかの実施形態では、免疫組織化学染色強度が、参照サンプル（例えば、対照細胞株染色サンプル、非癌性被験者からの組織サンプル、所定のレベルのタンパク質発現を有することが知られている参照サンプル）から求められる参照に対して求められる。

【0040】

図1は、組織学的染色プロセス100の例を示している。組織学的染色プロセス100の段階110は、サンプルを保存し、サンプルの分解を遅らせるために使用されることが
10
できるサンプル固定を含む。組織学において、固定とは、一般に、化学組成を保持し、天然サンプル構造を保存し、細胞構造を分解から維持するために化学物質を使用する不可逆的プロセスを指す。固定はまた、切片化のために細胞または組織を硬化させることができる。固定剤は、架橋タンパク質を使用してサンプルおよび細胞の保存を強化することができる。固定剤は、いくつかのタンパク質に結合して架橋し、脱水によって他のタンパク質を変性させることができ、これは組織を硬化させ、そうでなければサンプルを分解する可能性がある酵素を不活性化することができる。固定剤はまた、細菌を死滅させることができる。

【0041】

固定剤は、例えば、調製されたサンプルの灌流および浸漬によって投与されることが
20
できる。メタノール、ブイン固定剤および/またはホルムアルデヒド固定剤、例えば中性緩衝ホルマリン（NBF）またはパラフィン-ホルマリン（パラホルムアルデヒド-PFA）を含む様々な固定剤が使用されることができ、サンプルが液体サンプル（例えば、血液サンプル）である場合、サンプルがスライド上に塗抹され、固定前に乾燥されてもよい。

【0042】

固定プロセスは、組織学的試験の目的のためにサンプルおよび細胞の構造を保存するの
30
に役立つことができるが、固定は、組織抗原の秘匿化をもたらし、それによって抗原検出を減少させることができる。したがって、ホルマリンは、抗原を架橋し、エピトープをマスクすることができるため、固定は、一般に免疫組織化学の制限因子と考えられる。場合によっては、固定されたサンプルを無水シトラコン酸（可逆的タンパク質架橋剤）によって処理することおよび加熱することを含む、架橋の効果を逆転させるための追加のプロセスが行われる。

【0043】

組織学的染色プロセス100の段階120は、サンプル処理および包埋を含む。サンプル
40
処理は、固定されたサンプル（例えば、固定された組織サンプル）に、パラフィンワックスなどの適切な組織学的ワックスを浸透させることを含むことができる。組織学的ワックスは、水またはアルコールに不溶性であってもよいが、キシレンなどのパラフィン溶媒に可溶性であってもよい。したがって、組織内の水がキシレンによって置換される必要があり得る。そうするために、サンプル中の水をアルコールで徐々に置換することによってサンプルが最初に脱水されることができ、これは組織を増加する濃度のエチルアルコール（例えば、0から約100%）に通過させることによって達成されることができ、水がアルコールによって置換された後、アルコールは、アルコールと混和性であるキシレンによって置換されてもよい。包埋は、サンプルを温パラフィンワックスに包埋することを含むことができる。パラフィンワックスは、キシレンに可溶であり得るため、溶融ワックスは、キシレンによって充填され、以前に水によって充填されていた空間を充填することができる。ワックス充填サンプルが冷却されて、切片切断のためにマイクロームにクランプすることができる硬化ブロックを形成することができる。場合によっては、上記の例示的な手順からの逸脱は、抗体、化学物質、または他の固定剤の浸透の障害をもたらすパラフィンワックスの浸透をもたらす。

【0044】

10

20

30

40

50

組織学的染色プロセス100の段階130は、サンプルの切片化を含む。切片化は、検査のために顕微鏡スライド上にサンプル（例えば、包埋されて固定された組織サンプル）を取り付ける目的で、包埋ブロックからサンプルの薄いスライスを切断するプロセスである。切片化は、ミクロトームを使用して行うことができる。場合によっては、組織は、ドライアイスまたはイソペンタン中で急速に凍結されることができ、次いで冷蔵キャビネット（例えば、クライオスタット）中でコールドナイフによって切断されることができ、液体窒素などの他のタイプの冷却剤が使用されて組織を凍結させることができる。光学顕微鏡によって使用するための切片は、一般に4～10μm程度の厚さである。場合によっては、切片は、エポキシ樹脂またはアクリル樹脂に包埋されることができ、これにより、より薄い切片（例えば、< 2μm）が切断されることを可能にすることができる。切片は、ガラススライド上に置かれることができる。

10

【0045】

組織学的染色プロセス100の段階140は、（組織サンプルまたは固定液体サンプルの切片の）染色を含む。染色の目的は、呈色反応によって異なるサンプル成分を同定することである。ほとんどの細胞は無色透明である。したがって、細胞を可視化するために組織学的切片を染色する必要がある。染色プロセスは、一般に、特定の化合物、構造、分子、または特徴（例えば、細胞内特徴）の存在を確認または定量するために、色素または染色剤をサンプルに添加することを含む。例えば、染色は、組織切片から特定のバイオマーカーを同定または強調するのに役立つことができる。他の例では、染色が使用されて、生物学的組織（例えば、筋線維または結合組織）、細胞集団（例えば、異なる血球）、または個々の細胞内の細胞小器官を同定または強調することができる。

20

【0046】

多くの染色溶液は水性である。したがって、組織切片を染色するためには、染色溶液を切片に適用する前に、ワックスが溶解され、水によって置換される（再水和）必要がある。例えば、切片は、キシレン、エチルアルコールの濃度を減少させること（約100%から0%）、および水を順次通過させることができる。染色されると、切片が再度脱水され、キシレンに入れられてもよい。次いで、切片がキシレンに溶解した封入剤中の顕微鏡スライド上に封入されることができる。サンプル切片を保護するために、カバースリッブが上部に配置されることができる。カバースリッブの縁部の周りのキシレンの蒸発は、装着媒体を乾燥させ、カバースリッブをスライドにしっかりと接合することができる。

30

【0047】

様々なタイプの染色プロトコルが使用されて染色を行うことができる。例えば、例示的な免疫組織化学染色プロトコルは、インキュベーション中にスライドからの試薬の漏出を防ぐためにサンプル（例えば、組織切片）の周りに疎水性バリアラインを使用すること、非特異的染色の内因性源（例えば、酵素、遊離アルデヒド基、免疫グロブリン、特異的染色を模倣し得る他の無関係な分子）を遮断するために組織切片を試薬で処理すること、組織への抗体および他の染色試薬の浸透を促進するために透過化緩衝液とともにサンプルをインキュベートすること、特定の温度（例えば、室温、6～8℃）で一定期間（例えば、1～24時間）一次抗体とともに組織切片をインキュベートすること、洗浄緩衝液を使用してサンプルをすすぐこと、別の特定の温度（例えば、室温）で別の期間二次抗体とともにサンプル（組織切片）をインキュベートすること、水緩衝液を使用してサンプルを再びすすぎ、すすいだサンプルを色素原（例えば、DAB）とともにインキュベートすること、および反応を停止させるために色素原を洗い流すことを含む。場合によっては、その後、対比染色が使用されてサンプルの全体「ランドスケープ」を識別し、組織標的の検出に使用される主要な色の基準として機能する。対比染色剤の例は、ヘマトキシリン（青色から紫色の染色剤）、メチレンブルー（染色剤青色）、トルイジンブルー（核深青色、多糖類ピンク色から赤色の染色剤）、核ファストレッド（Kernectrot色素とも呼ばれ、赤色に染色する）、メチルグリーン（染色剤緑色）、非核発色性染色剤、例えばエオシン（ピンク色の染色剤）などを含むことができる。当業者は、他の免疫組織化学染色技術を実施して染色を行うことができることを認識するであろう。

40

50

【 0 0 4 8 】

別の例では、組織切片染色のためにH & E染色プロトコルが実施されることができる。H & E染色プロトコルは、金属塩または媒染剤と混合したヘマトキシリン染色剤をサンプルに適用することを含む。次いで、サンプルを弱酸溶液ですすいで過剰な染色（分化）を除去し、続いて弱アルカリ水中で青みを付けることができる。ヘマトキシリンの適用後、サンプルがエオシンによって対比染色されることができる。他のH & E染色技術が実施されることができることが理解されよう。

【 0 0 4 9 】

いくつかの実施形態では、目的の特徴がどれであるかに応じて、様々なタイプの染色を使用して染色を行うことができる。例えば、D A Bは、I H C染色のための様々な組織切片に使用されることができ、D A Bは、染色画像において関心のある特徴を示す茶色をもたらす。別の例では、D A B色がメラニン色素によってマスクされることができ、アルカリホスファターゼ（A P）がI H C染色のための皮膚組織切片に使用されることができる。一次染色技術に関して、適用可能な染色は、例えば、好塩基性および好酸性染色、ヘマチンおよびヘマトキシリン、硝酸銀、三色染色などを含むことができる。酸性染料は、組織または細胞中のカチオン性または塩基性成分、例えばタンパク質および細胞質中の他の成分と反応することができる。塩基性色素は、組織または細胞中のアニオン性または酸性成分、例えば核酸と反応することができる。上記のように、染色系の一例はH & Eである。エオシンは、負に帯電したピンク色の酸性染料とすることができ、ヘマトキシリンは、ヘマテインおよびアルミニウムイオンを含む紫色または青色の塩基性染料とすることができ、

【 0 0 5 0 】

組織学的染色プロセス100の段階150は、医用撮像を含む。顕微鏡（例えば、電子顕微鏡または光学顕微鏡）が使用されて、染色されたサンプルを拡大することができる。例えば、光学顕微鏡は、約数百ナノメートルなど、1 μ m未満の分解能を有することができる。ナノメートルまたはサブナノメートルの範囲でより細かい詳細を観察するために、電子顕微鏡が使用されることができる。拡大されたサンプルのデジタル画像は、デジタルカメラ（例えば、顕微鏡と一体化されているか、または顕微鏡に接続されている）を使用して取り込まれることができる。デジタル画像は、スライド全体および/または染色された組織スライス全体を描写するスライド全体画像を含むことができる。

【 0 0 5 1 】

段階160において、染色された切片の画像が記憶される。画像は、ローカル、リモート、および/またはクラウドサーバに記憶されることができる。各画像は、被験者の識別子および日付（例えば、サンプルが収集された日付および/または画像が取り込まれた日付）と関連付けて記憶されてもよい。画像は、さらに、別のシステム（例えば、病理学者に関連付けられたシステム、または自動化もしくは半自動化された画像分析システム）に送信されてもよい。

【 0 0 5 2 】

プロセス100に対する変更が企図されることが理解されよう。例えば、サンプルが液体サンプルである場合、段階120（処理および包埋）および/または段階130（切片化）は、プロセスから省略されてもよい。

I V . デジタル病理画像変換の例示的なプロセスフロー

【 0 0 5 3 】

図2は、機械学習モデルを使用してデジタル病理画像を処理するための（1つ以上のコンピューティングシステムを使用して実行される）プロセスフロー200を示している。本明細書でさらに説明するように、デジタル病理画像を処理することは、デジタル病理画像を使用して機械学習モデルを訓練すること、またはデジタル病理画像の一部または全部

10

20

30

40

50

を、機械学習モデルの訓練された（または部分的に訓練された）バージョンを使用して1つ以上の結果に変換することを含むことができる。

【0054】

図2に示すように、プロセスフロー200は、いくつかの段階、すなわち、画像記憶段階205、前処理段階210、ラベリング段階215、（アクティブ）モデル訓練段階220、および結果生成段階225を含む。

【0055】

画像記憶段階205は、生体サンプルスライド（例えば、組織スライド）からまたは生体サンプルスライドの全体から予め選択された領域のデジタル画像245のセットを提供するために（例えば、前処理段階210によって）アクセスされる1つ以上の画像データストア260を含むことができる。各画像データストア260に記憶され、画像記憶段階210においてアクセスされる各デジタル画像245は、図1に示すプロセス100の一部または全部にしたがって生成されたデジタル病理画像を含むことができる。

10

【0056】

前処理段階210において、デジタル画像245のセットのうちのそれぞれ、複数、または全ては、対応する前処理済み画像250を生成するために、1つ以上の技術を使用して前処理される。前処理は、画像を切り取ることを含んでもよい。場合によっては、前処理は、全ての特徴を同じスケール（例えば、同じサイズスケールまたは同じカラースケールまたは彩度スケール）にするための標準化または再スケーリング（例えば、正規化）をさらに含むことができる。特定の例では、画像は、所定の画素（例えば、2500画素）の最小サイズ（幅または高さ）または所定の画素（例えば、3000画素）の最大サイズ（幅または高さ）でサイズ変更され、元のアスペクト比を維持する。

20

【0057】

場合によっては、前処理は、デジタル画像内の関心のある各特徴を位置特定および/またはセグメント化することを含む。関心のある特徴を位置特定することは、画像の点位置および/または境界を予測することを含むことができる。例えば、位置特定は、描写された各核に対応する点位置を予測すること、または描写された各細胞膜に対応する閉鎖形状を予測することを含むことができる。特徴位置特定は、1つ以上の自動検出技術を使用して自動的に実行されてもよい。

【0058】

場合によっては、位置特定は、半自動的にまたは自動的に（特徴位置特定コントローラ252によって）実行される。例えば、画像にわたる強度（例えば、前処理済み画像）は、正規化または正則化されることができ、強度は、閾値化またはフィルタリングされることができ、および/または（例えば、物体、線および/または形状を検出するように構成される）アルゴリズムが適用されることができ、次いで、各境界出力および/または点位置が関心のある特徴として識別されることができ、場合によっては、メトリックは、各関心のある特徴（例えば、関心のある特徴を特定する信頼性を示すメトリック）に関連付けられ、関心レベルは、メトリックに基づいてスケーリングすることができる。

30

【0059】

場合によっては、位置特定は、前処理済み画像250に対応する別の画像を使用して、または同じベース画像245に対応する異なる前処理済み画像を使用して実行される。例えば、所与のデジタル病理画像について、所与の前処理動作が実行されて、核または細胞膜によって吸収される色素に対応する色を強調することができ、前処理済み画像を使用して細胞を位置特定することができる。一方、示された各細胞が所与の細胞型として分類（例えば、ラベリングを介して、または機械学習モデルを使用して）されるべきである場合、他の細胞小器官が確実に示されるようにするために、異なる前処理が使用されることができ、別の例として、サンプルの連続したスライスは、異なる染色を使用して染色されてもよく、その結果、スライスのいくつかは、（例えば、細胞の検出を容易にするためのH&E染色を使用して）関心のある特徴の位置特定を促進するために染色され、隣接するスライスは、異なるバイオマーカーの染色による細胞分類を促進するために染色される。

40

50

この場合、関心のある特徴の位置特定は、位置特定を識別するために使用される画像から、ラベル付けに使用される画像にマッピング、変換、および/またはマッピングされることができ。

【0060】

いくつかの前処理済み画像250は、ラベリング段階215において部分的または完全にラベル付けされる。ラベリング段階215において割り当てられるラベルは、人間のユーザ（例えば、病理学者または画像サイエンティスト）からの入力に基づいて識別される「グランドトゥルース」ラベルである。したがって、ラベリング段階は、1つ以上の前処理済み画像の一部または全部をユーザによって操作される装置に送信および/または提示することを含むことができる。場合によっては、ラベリング段階215は、ユーザによって操作される装置において提示される（例えば、APIを使用する）インターフェースを利用することを含み、インターフェースは、ラベルを識別する入力を受け入れるための入力コンポーネントを含む。例えば、入力コンポーネントは、ドロップダウンメニュー、ラジオボタンメニュー、テキストボックス（例えば、1つ以上の文字を受信するように構成される）、鉛筆ツール（例えば、点位置、曲線、境界などを識別するように構成される）などを含むことができる。

10

【0061】

場合によっては、インターフェースは、（例えば）テキスト命令および/または視覚化を介して伝達されることができ特定のラベルがどのラベルおよび/またはどの程度要求されているかを識別することができる。例えば、特定の色、サイズ、および/または記号は、ラベルが他の描写に対して画像内の特定の描写（例えば、特定の細胞または領域）に対して要求されていることを表すことができる。複数の描写に対応するラベルが要求される場合、インターフェースは、（1つの識別された描写にラベルを提供すると、ラベル付けのための次の描写の識別がトリガされるように）各描写を同時に識別してもよく、または各描写を順次に識別してもよい。

20

【0062】

場合によっては、ユーザが特定の数のラベル（例えば、特定のタイプの）を識別するまで、各画像が提示される。例えば、所与の全スライド画像または全スライド画像の所与のパッチは、ユーザがそれぞれマクロファージの描写に対応する15点の位置を識別するまで提示されることができ、その時点で、インターフェースは、異なる全スライド画像または異なるパッチの画像を（例えば、閾値数の画像またはパッチがラベル付けされるまで）提示することができる。したがって、場合によっては、インターフェースは、関心のある特徴の不完全なサブセットのラベルを要求および/または受け入れるように構成され、ユーザは、潜在的に多くの描写のうちのどれがラベル付けされるかを決定することができる。

30

【0063】

インターフェースは、（例えば、様々な視野を拡大するため、またはスライド全体のスケールに戻すために）画像のズームインおよびズームアウトおよび/または画像の移動をサポートすることができる。場合によっては、低倍率では、ラベルが要求される画像の部分は、1つのタイプの視覚的識別子（例えば、カラーマップ）を使用して識別されてもよく、高倍率では、ラベルが要求される画像の部分は、代替的または追加的に、別のタイプの視覚的識別子（例えば、マーカー記号、マーカーサイズ、および/またはマーカー形状）を使用して識別されてもよい。場合によっては、ラベルが要求される画像の部分の視覚的インジケータの特性は、倍率とともに変化する。例えば、点位置のマーカーのサイズは、拡大レベルと関連させることができる。

40

【0064】

ラベリング段階215において識別されるラベル255は、（例えば）細胞の分類、所与の細胞が特定のタイプの細胞であるかどうかに関するバイナリ表示、前処理済み画像250（または前処理済み画像250を有する特定の領域）が特定のタイプの描写（例えば、壊死またはアーチファクト）を含むかどうかに関するバイナリ表示、（例えば、特定のタイプのアーチファクトを識別する）スライドレベルまたは領域固有の描写のカテゴリ特

50

性、（例えば、領域内の特定のタイプの細胞の量、描写されたアーチファクトの量、または壊死領域の量を識別する）数字などを識別することができる。場合によっては、ラベル 255 は位置を含む。例えば、ラベル 255 は、特定のタイプの細胞の核の点位置または特定のタイプの細胞の点位置を識別することができる。別の例として、ラベル 250 は、描写された腫瘍、血管、壊死領域などの縁または境界を含むことができる。関心のある特徴に応じて、所与のラベル化された前処理済み画像 250 は、単一のラベル 255 または複数のラベル 255 に関連付けられることができる。後者の場合、各ラベル 255 は、（例えば）そのラベルが前処理済み画像 250 内のどの位置または部分に対応するかに関する指示に関連付けられることができる。

【0065】

ラベリング段階 215 において、ラベル 255 および対応する前処理済み画像 250 が使用されて、機械学習モデル 260 を訓練することができる。機械学習モデル 260 は、（例えば）畳み込みニューラルネットワーク（「CNN」）、深層ニューラルネットワーク、初期ニューラルネットワーク、残差ニューラルネットワーク（「Resnet」）、U-Net、V-Net、シングルショットマルチボックス検出器（「SSD」）ネットワーク、リカレントニューラルネットワーク（「RNN」）、正規化線形ユニット（「ReLU」）、ロングショートタームメモリ（「LSTM」）モデル、ゲートド回帰ユニット（「GRU」）モデルなど、またはそれらの任意の組み合わせを含む。機械学習モデルは、決定木モデル、ランダムフォレストモデル、サポートベクターマシン、および/または回帰モデルを含むことができる。

【0066】

機械学習モデルは、固定された 1 つ以上のハイパーパラメータ（例えば、プログラムによって定義される）を含むおよび/またはそれを使用して構成されることができる。例えば、ハイパーパラメータは、ニューラルネットワーク内の層の数、層内のノードの数、学習率などを定義することができる。

【0067】

機械学習モデル 260 を訓練することは、（例えば、損失を最小化するため、または目的を最大化するために）損失関数または目的関数を使用してパラメータのセット（例えば、1 つ以上の係数および/または重み）を学習することを含むことができる。各パラメータは、パラメータの値が訓練中に調整されるように、調整可能な変数であってもよい。例えば、損失関数または目的関数は、描写された表現の正確な分類を最適化し、（例えば、形状、サイズ、均一性などを特徴付ける）所与のタイプの特徴の特徴付けを最適化し、所与のタイプの特徴の検出を最適化し、および/または所与のタイプの特徴の正確な位置特定を最適化するように構成されることができる。

【0068】

任意の段階における訓練は、人間のユーザが利用可能なデータの不完全なサブセットをラベル付けすることを要求することを含むことができる。場合によっては、最初の訓練反復中に、ランダムもしくは擬似ランダム計算アルゴリズムを使用して、またはラベル化されるサブセットをユーザが選択するように要求することによって、サブセットが識別されることができる。

【0069】

最初の訓練反復に関して、機械学習モデル 260 は、最初の訓練の前にランダムまたは擬似ランダムに設定されたパラメータ値によって初期化されてもよい（ただし、そうである必要はない）。場合によっては、パラメータが別のタスクのために学習された転移学習を使用して初期パラメータ値が定義される。場合によっては、パラメータは、ラベル化される複数の特徴のサブセットのランダム、擬似ランダム、または完全に差別的なユーザ選択を含む初期訓練段階の代わりに、転移学習を使用して定義される。

【0070】

後続の反復に関して、パラメータは、前の反復からのパラメータ定義にしたがって設定されることができ、（アクティブ）モデル訓練段階 220 において実行される訓練は、ラ

10

20

30

40

50

ベル付け段階 2 1 5 においてラベル付けされるデータを選択または優先順位付けする際によりアクティブになることができる。より具体的には、機械学習モデル 2 6 0 のパラメータの少なくとも一時的な定義が最初の（または後続の）反復によって識別されるため、結果生成段階 2 2 5 においてモデル 2 6 0 が使用されて、他の前処理済み画像 2 5 0 を予測ラベル 2 6 5 に変換することができる。各予測ラベル 2 6 5 は、予測ラベル 2 6 5 を生成するために使用される画像および/または画像の一部、ならびに信頼性メトリック 2 7 0 に関連付けられることができる。

【 0 0 7 1 】

ユーザの装置（例えば、病理学者に）に利用されるインターフェースは、低信頼性メトリック（例えば、所定の絶対閾値または相対閾値を下回る）に関連付けられた画像の部分を識別または提示することができる。例えば、ラベルが画像全体に対応する場合、アクティブモデル訓練は、所与の訓練反復からの最低信頼性メトリック 2 7 0 に関連する所定数の画像を提示すること、および/または所定の閾値を下回る所与の訓練反復からの信頼性メトリック 2 7 0 に関連する全ての画像を提示することを含むことができる。別の例として、所定の閾値を下回る所与の訓練反復からの信頼性メトリック 2 7 0 と関連付けられた画像の初期グループが識別されることができ、画像の初期グループの（所定の数の画像を有する）サブセットがランダムまたは擬似ランダムに選択されて提示されることができる。

【 0 0 7 2 】

インターフェースは、ラベリングを容易にするために、本明細書に記載の 1 つ以上の特徴を含むことができる。例えば、インターフェースは、ラベルおよび/またはズーム機能を識別する入力を受け入れるための 1 つ以上の入力コンポーネントを有することができる。別の例として、インターフェースは、ラベルまたはラベルレビューが画像のどの部分に対して要求されているかを示すために、色（例えば、ヒートマップを介して）または他の視覚的インジケータを使用することができる。インターフェースは、対応する予測ラベル 2 6 5 の指示に関連して各画像または各画像部分を提示することができる。例えば、画像部分の点位置を識別するマーカーの色は、バイナリまたはカテゴリ予測ラベルの値を識別することができる。別の例として、テキストの見出しまたはオーバーレイは、画像の予測ラベルを識別することができる。

【 0 0 7 3 】

場合によっては、ラベリングまたはラベルレビューのために画像または画像の部分を選択するのではなく、インターフェースは、所与の訓練画像内のモデルを訓練するために使用される全ての画像または 1 つ以上の画像の全ての部分に対するラベリングまたはラベルレビューをサポートするように構成される。例えば、インターフェースは、多くの予測ラベル 2 6 5 のそれぞれの視覚的識別子を同時にまたは順次に提示することができ、予測ラベル 2 6 5 を確認または拒否するそのような予測ラベル 2 6 5 の入力を受け入れることができる。この場合、それにもかかわらず、インターフェースは、低いまたは比較的低い信頼性に関連する予測ラベルを強調してもよく、および/または低いまたは比較的低い信頼性に関連する予測ラベルのラベルレビューを明示的に要求してもよい。例えば、視野内の画像の各部分にマーカーがオーバーレイされることができる。マーカーの色は、予測ラベル 2 6 5 を表すことができ、マーカーのサイズは、対応する信頼性メトリック 2 7 0 と逆

【 0 0 7 4 】

ユーザがインターフェースと相互作用して画像部分または画像の新たなラベルを識別すると、新たなラベルは、画像部分または画像のグランドトゥルースラベル 2 5 5 であると見なされる。ユーザが画像部分または画像の所与の予測ラベルを確認するためにインターフェースと相互作用すると、画像部分または画像のグランドトゥルースラベル 2 5 5 が予測ラベルであると定義される。ユーザが予測ラベルを拒否するためにインターフェースと相互作用すると、2 つの潜在的なラベルのみが利用可能である場合（ラベル付けがバイナリであることを意味する）、画像部分または画像のグランドトゥルースラベル 2 5 5 は、2 つの潜在的なラベルの他方であると定義される。

10

20

30

40

50

【 0 0 7 5 】

次いで、新たなグラントゥールスラベル 2 5 5 が使用されて、次の訓練反復中に機械学習モデル 2 6 0 を訓練することができる。例えば、機械学習モデル 2 6 0 は、前の反復中に学習されたパラメータ値によって初期化されてもよく、その後、モデルは、新たなグラントゥールスラベル 2 5 5 を使用して訓練されてもよい。別の例として、機械学習モデル 2 6 0 は、複数のラベリングセッションおよび/または複数の訓練反復に関連して識別されたラベル 2 5 5 を使用して再訓練されることができ、最新のラベリングセッションに関連するラベル 2 5 5 は、他のラベルよりも高度に重み付けされることができ。

【 0 0 7 6 】

予測ラベル 2 6 5 の生成と信頼性メトリック 2 7 0 (結果生成段階 2 2 5) との間の反復は、低信頼性予測のラベル付けまたはラベルレビューを容易にするインターフェースを提供し、新たなラベルを使用して機械学習モデル 2 6 0 を訓練することは、停止条件が満たされるまで継続することができる。訓練完了条件は、(例えば) 所定の反復回数が完了し、信頼性メトリック 2 7 0 に基づいて生成された統計値(例えば、平均もしくは中央信頼性メトリック、または特定の値を上回る信頼性メトリックの割合)が所定の信頼性閾値を超え、直近のラベルレビューにおいてレビューされて拒否されたラベルの割合が所定のラベル拒否閾値を下回った場合、および/またはラベルレビューに関与していたユーザ装置がラベルレビューアプリケーションを閉じる場合に満たされるように構成されることができる。場合によっては、ユーザ装置からの対応する要求の受信に 응답して、新たな訓練反復が開始されることができ。

【 0 0 7 7 】

次いで、(結果生成段階 2 2 5 において) 訓練された機械学習モデル 2 6 0 が使用されて、新たな前処理済み画像 2 5 0 を処理して、新たな予測ラベル 2 6 5 および潜在的に新たな信頼性メトリック 2 7 0 を生成することができる。新たな予測ラベル(および潜在的に新たな信頼性メトリック 2 7 0) は、初期ラベリングまたはラベルレビューに使用されたものと同じまたは類似のタイプのインターフェースを使用して提示されることができる。様々な場合では、新たな予測ラベルを提示するインターフェースは、予測ラベル 2 6 5 を拒否または置換するための入力コンポーネントを含むか、または欠いていてもよい。

【 0 0 7 8 】

場合によっては、分析コントローラ 2 8 0 は、基礎となる画像の処理を要求したエンティティに利用される 1 つ以上の分析結果 2 8 5 を生成する。予測ラベル 2 6 5 の代わりに、またはそれに加えて、分析結果 2 8 5 が提供されてもよい。分析結果 2 8 5 は、予測ラベルに基づくことができる。例えば、分析結果 2 8 5 は、特定のラベルに関連する示された細胞の数または割合を識別することができる。

【 0 0 7 9 】

プロセスフロー 2 0 0 は例示的なものであり、異なる段階および/または構成要素を使用するプロセスフロー 2 0 0 が企図されることが理解されよう。例えば、場合によっては、ネットワークは、モデルを訓練するために使用される画像および/またはモデルによって処理された画像が生画像(例えば、画像データストアから)であるように、前処理段階 2 1 0 を省略することができる。別の例として、前処理段階 2 1 0 および(アクティブ)モデル訓練段階 2 2 0 のそれぞれは、本明細書に記載の 1 つ以上の動作を実行するためのコントローラを含むことができることが理解されよう。同様に、ラベリング段階 2 1 5 は、特徴位置特定コントローラ 2 5 2 に関連して示されており、結果生成段階 2 2 5 は、分析コントローラ 2 8 0 に関連して示されているが、各段階に関連するコントローラは、特徴位置特定および/または分析結果の生成以外の本明細書に記載の他の動作をさらにまたは代替的に促進することができる。さらに別の例として、図 2 に示すプロセスフロー 2 0 0 の描写は、(例えば、様々なインターフェースがどのように機能するかなどを定義した機械学習モデル 2 6 0 のアーキテクチャを選択した) プログラマに関連付けられた装置、(例えば、ラベリング段階 2 1 5 において) 初期ラベルまたはラベルレビューを提供するユーザに関連付けられた装置、および(初期ラベルまたはラベルレビューを提供したユー

10

20

30

40

50

ザと同じユーザまたは異なるユーザとすることができる) 所与の画像のモデル処理を要求するユーザに関連付けられた装置の描写された表現を欠いている。これらの装置の描写がないにもかかわらず、プロセスフロー 200 は、装置のうちの 1 つ、複数、または全ての使用を含むことができ、実際には、初期ラベルまたはラベルレビューを提供する対応する複数のユーザに関連付けられた複数の装置、および/または様々な画像のモデル処理を要求する対応する複数のユーザに関連付けられた複数の装置の使用を含むことができる。

【0080】

IV. A. 全スライド画像の部分をラベル付けするためのモデルの例示的な訓練

図 3 は、本開示のいくつかの態様にかかるアクティブラーニングワークフロー 300 を示している。アクティブラーニングワークフロー 300 に表される動作は、1 つ以上のコンピュータシステム (例えば、モデル訓練コンピュータシステムおよび/またはラベルもしくはラベルレビューを提供する 1 人以上のユーザの 1 つ以上の装置) を使用して実行されることができる。アクティブラーニングワークフロー 300 において実行される動作の一部または全部は、プロセスフロー 200 (図 2 に示す) のラベリング段階 215 および/または (アクティブ) モデル訓練段階 220 の間に実行されてもよい。

10

【0081】

アクティブラーニングワークフローは、初期化動作 305 およびアクティブラーニング動作 310 の双方を含む。この例示的な例では、ブロック 315 において、前処理動作が実行されて、(例えば、特定の順序でコマンドまたはスクリプトのセットを実行することをサポートすることができる VDP SDK として表されるようなフレームワークを使用して) スライド全体の画像内の特徴を検出する。特徴は、(例えば) 細胞、アーチファクト、核などを含むことができる。

20

【0082】

ブロック 320 において、1 つ以上のスライド画像全体が開かれる。スライド画像全体を開くことは、ローカルまたはリモート記憶装置からユーザ装置にスライド画像全体をロードすることを含むことができる。スライド画像全体は、ユーザ (例えば、病理学者、画像サイエンティスト、または他の適切なユーザ) が閲覧ソフトウェア内で特定のタイプの入力 (例えば、セッションの開始および/またはスライド画像全体の識別する) を提供したことに応答して開かれることができる。スライド全体画像を開くと、スライド全体画像がユーザ装置に提示されるようにトリガすることができる。

30

【0083】

ブロック 325 において、初期アルゴリズムからの結果がアクセスされ、結果はスライド画像全体に対応する。初期アルゴリズムは、機械学習モデルを含むことができ、機械学習モデルは、ランダムに選択されたパラメータ値または擬似ランダムに選択されたパラメータ値によって初期化されたモデルを含むことができる。初期アルゴリズムは、転移学習 (すなわち、異なるタスクに対応するコンテキストにおいて学習されている) を使用して定義されたパラメータ値を有する機械学習モデルを含むことができる。結果は、1 つ以上の予測ラベルを含むことができる。

【0084】

ブロック 330 において、1 つ以上の視野 (FOV) が選択される。各視野は、スライド画像全体の一部を識別することができる。視野は、(例えば、ランダムまたは擬似ランダム技術を使用して、または所定の順序を使用して) 自動的に選択されてもよく、または (例えば、スライド画像全体の提示の一部をクリックして、またはスライド画像全体の領域を選択して) ユーザ入力に基づいて選択されてもよい。視野を選択することは、(例えば、スライド全体の画像の代わりに、またはそれに加えて) 視野をユーザ装置に提示させることができる。

40

【0085】

ブロック 335 において、視野内に存在する (ブロック 315 において識別された) 関心のある特徴のうちの 1 つ、複数、または全てがラベル付けされる。ラベルのそれぞれは、ユーザから受信した入力に基づいて定義されることができる。これらのラベルのそれぞれ

50

これは、グランドトゥールースラベル340であるように定義されることができ、これはブロック345において対応する特徴に関するデータと関連付けて保存される。

【0086】

ブロック350において、訓練に使用される機械学習モデルおよび特徴（例えば、1つ以上の核の特徴、1つ以上の細胞の特徴、1つ以上の形態学的特徴、および/または1つ以上の構造的特徴）が選択される。例えば、選択は、リスト、プルダウンメニューなどに提示された複数のタイプのモデルの中から特定のタイプのモデルを識別するユーザ装置において受信された入力に基づいて行われてもよい。特徴は、自動的にまたはユーザ入力に基づいて選択されてもよい。複数のタイプのモデルのそれぞれは、モデルアーキテクチャ、パラメータ初期化（例えば、異なるコンテキストにおける訓練に基づいて決定される）などに関して互いに異なっていてもよい。選択は、（例えば、ユーザ装置で受信した入力に基づいて）モデルの1つ以上のハイパーパラメータを定義することをさらに含むことができる。別の例として、モデルは、ランダム、擬似ランダム、または所定の順序を使用して、複数のモデルの中から選択されてもよい。

10

【0087】

ブロック355において、選択された機械学習モデルは、グランドトゥールースラベル340および視野の対応する部分を使用して訓練される。場合によっては、グランドトゥールースラベルは、（ブロック355における）モデルを訓練するための第1のサブセットと検証のための第2のサブセットとの2つのサブセットに分割される。グランドトゥールースラベルの第2のサブセットを使用して1つ以上の検証メトリックが計算されることができ、ブロック360において、検証メトリックが少なくともベースライン性能品質を示すかどうか決定されることができ、そうでない場合、ワークフロー300は、ブロック350に戻ることができ、別のモデルが選択されることができ。

20

【0088】

検証メトリックが少なくともベースライン性能品質を示すと決定された場合、ワークフロー300は、アクティブラーニング動作310に遷移する。

【0089】

ブロック365において、ブロック350において選択され、ブロック355において学習されたパラメータを有する機械学習モデルは、同じまたは新たな全スライド画像上で実行される。機械学習モデルは、画像内の複数の点位置のそれぞれまたは領域内の複数の部分のそれぞれに対応する予測ラベルおよび対応する信頼性メトリックを出力することができる。信頼性メトリックが使用されて、（ブロック365においても）スライド全体画像またはスライド全体画像の一部に対応するヒートマップを生成することができ、ヒートマップは、描写された領域の異なる部分にわたる予測ラベルの確からしさ（例えば、信頼性メトリック）を表す。例えば、ヒートマップは、平滑化技術を使用することによって、または複数の領域のそれぞれについて統計値を計算することによって生成されてもよい。

30

【0090】

ブロック370において、機械学習モデルが十分に訓練されているかどうか決定される。そうである場合、ブロック372において、モデルは、モデルを訓練するために使用される全てのグランドトゥールースラベルとともに保存される。

40

【0091】

そうでない場合、（ブロック365からの）機械学習モデルによって生成された375における結果（予測ラベル）のように、（例えば、ユーザ装置において）確実性ヒートマップが表示される。場合によっては、ヒートマップおよび結果は同時に表示される。例えば、各予測の視覚的インジケータは、関心のある対応する特徴に関連付けられた位置においてヒートマップ上にオーバーレイされてもよい。別の例として、（例えば、全スライド画像の、または所与の視野についての）ヒートマップは、結果と重複しない方法（例えば、2つのフレームにおいて）で同時に提示されてもよい。場合によっては、ヒートマップは、第1の拡大レベルセットにわたって提示され、結果は、第2の拡大レベルセットにわたって提示される（第1の拡大レベルセットは、第2の拡大レベルセットと重複してもよ

50

く、または重複しなくてもよい)。

【0092】

ブロック380において、(例えば、ヒートマップが表示されていたユーザ装置における対応する入力を検出に回答して)新たな視野が選択される。ブロック385において、関心のある1つ以上の特徴がラベル付けされ、それによって、1つ以上の追加のグラウンドトゥールースラベル(gt)が作成される。したがって、ブロック390に示されるように、グラウンドトゥールースラベル(G)の全セットは、初期化動作305からのグラウンドトゥールースラベルと、アクティブラーニング動作310を介して識別された各追加のグラウンドトゥールースラベルとを含むように定義されることができる。

【0093】

ブロック395において、機械学習モデルは、グラウンドトゥールースラベルの全セットを使用して再訓練される。ブロック397において、グラウンドトゥールースラベルの全セットが記憶され(例えば、グラウンドトゥールースラベルの任意の以前に記憶された全セットを置き換える)、および/または(ブロック385からの)新たに取得されたラベルが、以前に記憶されたグラウンドトゥールースラベルを補うために記憶される。モデルは、(ブロック370において)停止基準が満たされたと決定されるまで訓練され続けることができる。

【0094】

IV. B. デジタル病理プロセスフローを使用するための例示的なワークフロー

図4は、デジタル病理プロセスフロー(例えば、図2に示すプロセスフロー200)の一部または全部を使用するための例示的なワークフローを示している。図示されたワークフローは、1つ以上のコンピューティングシステム(例えば、人間のアナテータに関連するコンピューティングシステム、ラベル予測を要求するユーザに関連するコンピューティングシステム、および/またはリモートコンピューティングシステム(例えば、クラウドで動作する))上で実行されることができる。

【0095】

ブロック405において、画像の一部または全部(例えば、全スライド画像の一部または全部)に注釈が付けられる。注釈付けは、(例えば)画像の一部にラベル(例えば、バイナリラベル、分類ラベル、数値ラベルなど)を割り当てること、および/または所与の特徴を描写すると予測される画像の特定の部分を識別することを含むことができる。注釈付けは、アナテータによって操作される装置において受信される入力にしたがって実行されてもよい。注釈付けは、異なる画像および/または異なる関心部分にわたって繰り返し実行されてもよい。注釈付けは、追加的または代替的に、(例えば、個々の細胞を最初に位置特定させ、個々の細胞を別々に分類するために)異なる特徴をラベル付けするために実行されてもよい。

【0096】

場合によっては、注釈付けは、プロセスフロー200のブロック215において実行される。場合によっては、注釈付き画像は、各画像が注釈を介して生成された対応するラベルデータに関連して記憶されるように、画像データストア240に記憶されてもよい。

【0097】

ブロック410において、データ収集が実行される。データ収集は、特定の注釈付き画像および/または注釈付き画像の特定の部分を識別することを含むことができる。特定の画像および/または特定の部分は、それらが任意の特定のタイプの関心のある特徴を識別するように注釈付けされたこと、それらが特定のタイプの関心のある特徴の少なくとも一部を示したこと、それらが少なくとも1つの特定のタイプの関心のある特徴を示したこと、および/またはそれらが少なくとも特定のタイプの関心のある特徴の閾値量を示したことを示すように注釈付けされていてもよい。例えば、データ収集は、少なくとも10個の細胞が示され、少なくとも2個のマクロファージが示されていることを示すように注釈が付けられた各画像パッチを識別するために実行されることができる。別の例として、データ収集は、(マクロファージが検出されたかどうかにかかわらず)各マクロファージ細胞を識別するように注釈付けされた各画像パッチを識別するために実行されることができる。

10

20

30

40

50

【0098】

場合によっては、データ収集は、画像データストア240にクエリを送信し、対応する画像および/または画像部分を識別する結果を受信することによって、前処理段階210において実行されてもよい。場合によっては、データ収集は、ラベリング段階215の一部として実行される。

【0099】

ブロック415において、データ準備が実行される。データ準備は、(例えば)本明細書に開示される1つ以上の前処理技術による前処理を含むことができる。データ準備は、1つ以上のタイルまたはパッチを定義することを含むことができる(例えば、それぞれが画像の一部に対応する)。データ準備は、組織スライスの少なくとも一部を描写する画像の一部を識別することができるグラウンドトゥルースマスクを定義することを含むことができる。データ準備は、前処理段階210において実行されることことができる。

10

【0100】

ブロック420において、準備されたデータ(複数の画像および各画像に対応する少なくとも1つのラベルを含むことができる)が使用されて、機械学習モデルを訓練および/または微調整することができる。訓練は、ラベル(例えば、および対応する信頼性メトリック)の予測を生成するために機械学習モデルの現在のバージョンを使用すること、(例えば、損失関数を使用して)予測ラベルを真のラベルと比較すること、および比較に基づいてモデルのパラメータを調整することを含むことができる。訓練は、モデルおよび/または1つ以上の静的変数(例えば、画像の特徴)を使用して現在のモデルを適合させることをさらにまたは追加的に含むことができる。訓練または微調整は、モデル訓練段階220において実行されることことができる。

20

【0101】

ブロック425において、訓練されたモデルが評価されることことができる。評価は、(例えば)予測ラベルの感度、特異性および/または精度を特徴付けることことができる。評価は、モデル訓練段階220において実行されることことができる。評価は、所与のメトリックが所定の閾値を超えるかどうかを決定することを含むことことができる。メトリックが予め定義された閾値を超えない場合、予め定義された閾値を超えるまで、または別の条件が満たされるまで(例えば、少なくとも所定数の訓練反復が完了する)、訓練(または微調整)および評価が繰り返されることことができる。

30

【0102】

ブロック430において、スライド全体の画像推論が実行される。スライド全体の画像推論は、段階225において実行されてもよく、部分内の画像の部分に関連する信頼性メトリックに基づいて、スライド全体の画像の複数の部分のそれぞれについての統計値を生成することを含んでもよい。場合によっては、推論は、部分内の画像の部分に関連付けられた予測ラベルに基づいて、スライド全体画像の複数の部分のそれぞれについての統計値をさらにまたは代替的に生成する。したがって、統計値は、画像部分内のラベルの信頼性を表すことができ、および/または画像部分の全体的な予測ラベルを反映することことができる。ブロック430は、結果生成段階225または(アクティブ)モデル訓練段階220によって実行されることことができる。

40

【0103】

ブロック435において、スライド全体の画像内の複数の部分のそれぞれについての信頼性メトリック統計値を示す確実性ヒートマップが生成される(例えば、結果生成段階225または(アクティブ)モデル訓練段階220において)。この画像は、画像レビューまたはラベルのどの部分が要求されるかに関して人間のアノテータを導くのに有用であり得る。例えば、赤色領域は、予測ラベルのレビューが青色領域と比較して非常に優先されることを示すことことができる。

【0104】

ブロック440において、(例えば、アノテータの装置で受信された入力に基づいて)新たなラベルが識別される。新たなラベルは、信頼性の低い予測が生成された画像部分に

50

関連して提供されたラベルに対応することができる。

【0105】

次いで、ワークフローは、ブロック420に戻り、新たなラベルに基づいて訓練を継続することができる。反復は、停止条件が満たされるまで（例えば、少なくとも所定数のラベルが注釈付けによって提供され、少なくとも所定数の訓練反復が実行され、少なくとも所定のモデル精度が達成され、人間のオペレータからの入力モデルの受け入れに対応しているなど）継続することができる。

V. デジタル病理処理のための例示的なフレームワーク

【0106】

本明細書で示されるように、1つ以上のシステムのそれぞれは、画像を処理するために機械学習モデルを訓練または使用するために開示された様々な動作および/または処理を実行することに関与することができる。図5は、この点で使用されることができる例示的なシステムおよびアプリケーションを有するネットワーク500を開示している。ネットワーク500は、プロセスフロー200の一部もしくは全部、図3に示すアクティブラーニングワークフローの一部もしくは全部、および/または図4に示す例示的なワークフローの一部もしくは全部をサポートするコンポーネントを含むことができる。

10

【0107】

システムおよび/またはアプリケーションは、高解像度および/または高速画像の記憶および取得、画像の視覚化、データの視覚化、対話型データ補正、および/または画像分析（例えば、画像訓練および推論）を個別にまたは集成的にサポートすることができる。

20

【0108】

ネットワーク500は、相互作用支援コンポーネント510のセットによってサポートされるフロントエンドアプリケーションサブネットワーク505を含む。ユーザ装置512は、コンテンツ管理システム515と通信する（例えば、通信を送信し、通信を受信する）ことができ、コンテンツ管理システムは、ユーザ装置512とのアクティブラーニングセッションを1つ以上管理することができる。アクティブラーニングセッションを管理することは、ネットワーク500内の1つ以上の他のコンポーネントによって実行される動作を調整することを含むことができる。

【0109】

ビューアコンポーネント520は、ユーザ装置512に提供されるインターフェースを定義および提供することができる。インターフェースは、（例えば）画像、画像にラベルを付けるための1つ以上の入力コンポーネント、1つ以上の予測ラベル、または本明細書で説明される他の情報を含むことができる。

30

【0110】

ビューアコンポーネント520は、要求を画像管理コンポーネント530に送信することによって画像にアクセスすることができる。要求は、（例えば）要求されている画像のタイプの指示（例えば、器官、染色タイプ、被験者の疾患タイプ、被験者の年齢などを識別すること）、特定の画像の識別子（例えば、特定の被験者を識別すること）、認可情報（例えば、画像管理コンポーネント530がアクセス特権を評価することを可能にするように、ユーザ装置512に関連付けられたユーザを識別すること）、特定の臨床試験の識別、および/または（特定の画像タイプおよび/または特定の画像アクセス制限に関連付けられることができる）訓練または使用されている特定の機械学習モデルの識別を含むことができる。場合によっては、要求は、単一の画像に対するものである。場合によっては、要求は、（ビューアコンポーネント520が部分的または全体的に、同時にまたは別々に提示することができる）複数の画像に対するものである。

40

【0111】

画像管理コンポーネント530は、要求に関連する許可を評価することができ、要求が許可されていると決定された場合、画像についてリモートまたはローカル画像ファイルシステム530に問い合わせることができる。場合によっては、画像管理コンポーネント530は、画像をビューアコンポーネント520に戻す前に画像を前処理する。例えば、画

50

像管理コンポーネント 5 3 0 は、画像をトリミングまたはスケーリングすることができる。

【 0 1 1 2 】

追加的または代替的に、ビューアコンポーネント 5 2 0 は、画像前処理を調整するために分析プラットフォームコンポーネント 5 3 5 と通信することができる。例えば、前処理は、画像に描写された 1 つ以上の細胞のそれぞれの点位置を識別することを含むことができる。分析プラットフォームコンポーネント 5 3 5 は、機械学習モデルの訓練および/または使用をさらに容易にすることができる。例えば、(例えば、細胞が特定の細胞タイプであるかどうか、画像がアーチファクトを描写しているかどうか、腫瘍境界がどこにあるかなどを予測する) ユーザ装置 5 1 2 からの入力に基づいて識別されたラベルが使用されて、機械学習モデルを訓練することができる。続いて、機械学習モデルが使用されて、そのような予測および潜在的に対応する信頼性メトリックを生成することができる。機械学習モデルの訓練および使用は、(例えば) 高速プロセッサおよび大規模メモリを含む高性能コンポーネント 5 4 0 を使用して実行されることができる。高性能コンポーネント 5 4 0 は、それぞれが 1 つ以上のコーディングプラットフォームをサポートすることができる複数のノード 5 4 5 を含むことができる。場合によっては、各ノード 5 4 5 は、並列処理をサポートするように実質的に独立して動作する。

10

【 0 1 1 3 】

分析プラットフォーム 5 3 5 は、低信頼性メトリックに関連付けられた予測ラベルを検出ことができ、(例えば、コンテンツ管理システム 5 1 5 および/またはビューアコンポーネント 5 2 0 を介して) これらの予測ラベルのラベルレビューの要求を容易にすることができる。本明細書でさらに説明するように、各ラベルの確認、拒否、または交換を示すユーザ装置 5 1 2 からのフィードバックが使用されて、モデルを再訓練または微調整することができる。

20

【 0 1 1 4 】

VI . 実施例

機械学習モデルを、デジタル病理画像におけるマクロファージの描写を検出するように構成した。各画像について、前処理ステップを実行して、示された各核の位置を検出した。

【 0 1 1 5 】

図 6 A は、サンプルの染色されたスライスの例示的な全体スライド画像を描写している。各白色ボックスは、ユーザがボックスをクリックすると拡大することができる異なる視野に対応する。図 6 B ~ 図 6 D の描写のそれぞれは、同じ特定の視野に対応する。ドットのそれぞれは、(前処理において識別された) 描写された核の予測位置を表す。

30

【 0 1 1 6 】

図 6 B では、核の各表現は互いに同じである。図 6 B に示すインターフェースは、機械学習モデルの初期訓練中にユーザに提示され、ユーザは、マクロファージであるか否かを標識する核のサブセットを選択することができる。これらのラベル(および同様の技術を使用して得られた他のラベル)が使用されて、モデルを最初に訓練した。

【 0 1 1 7 】

次いで、モデルは、各識別された核について、核がマクロファージまたは他の細胞型に対応するかどうかを予測した。各核について、モデルは、予測の信頼性を示す信頼性メトリックをさらに出力する。図 6 C は、同じ視野および核の識別を示すが、核検出に関連するドットのサイズおよび輝度は、信頼性メトリックに基づいてスケーリングされる。(具体的には、閾値を超える信頼性メトリックに対応する核は十字によって表され、他は黒丸によって表される。) これらの視覚的特性は、(例えば、明るいドットをクリックし、さらに入力を提供することによって) 不確実な細胞をラベル付けし、および/または不確実な細胞のラベルをレビューするようにユーザを導くことができる。

40

【 0 1 1 8 】

様々なラベル予測を検証または修正した第 1 回目の入力を受信すると、追加のラベルを使用してモデルをさらに訓練した。次いで、モデルは、様々な細胞がマクロファージであるかどうかに関する予測を再生成し、予測メトリックおよび信頼性メトリックを再生成し

50

た。

【0119】

図6Dは、同じ視野および核の識別を示すが、核検出に関連するドットのサイズおよび輝度は、新たな信頼性メトリックに基づいてスケーリングされる。これらの更新された視覚的特性は、更新されたモデルが最も確かでない画像の部分ラベル付けまたはレビューするようにユーザを導くことができる。

【0120】

VII.さらなる考察

本開示のいくつかの実施形態は、1つ以上のデータプロセッサを含むシステムを含む。いくつかの実施形態では、システムは、命令を含む非一時的コンピュータ可読記憶媒体であって、命令が、1つ以上のデータプロセッサ上で実行されると、1つ以上のデータプロセッサに、本明細書に開示された1つ以上の方法の一部または全部および/または1つ以上のプロセスの一部または全部を実行させる、非一時的コンピュータ可読記憶媒体を含む。本開示のいくつかの実施形態は、1つ以上のデータプロセッサに、本明細書に開示された1つ以上の方法の一部または全部および/または1つ以上のプロセスの一部または全部を実行させるように構成された命令を含む、非一時的機械可読記憶媒体において有形に具現化されたコンピュータプログラム製品を含む。

10

【0121】

使用された用語および表現は、限定ではなく説明の用語として使用され、そのような用語および表現の使用において、示されて説明された任意の特徴の均等物またはその一部を除外する意図はないが、特許請求の範囲に記載された発明の範囲内で様々な変更が可能であることが認識される。したがって、特許請求の範囲に記載された本発明は、実施形態および任意の特徴によって具体的に開示されているが、本明細書に開示された概念の変更および変形は、当業者によってあてにされてもよく、そのような変更および変形は、添付の特許請求の範囲によって定義される本発明の範囲内にあると見なされることを理解されたい。

20

【0122】

本説明は、好ましい例示的な実施形態のみを提供し、本開示の範囲、適用可能性または構成を限定することを意図しない。むしろ、例示的な実施形態の本説明は、様々な実施形態を実装するための可能な説明を当業者に提供する。添付の特許請求の範囲に記載の趣旨および範囲から逸脱することなく、要素の機能および配置に様々な変更を加えることができることが理解される。

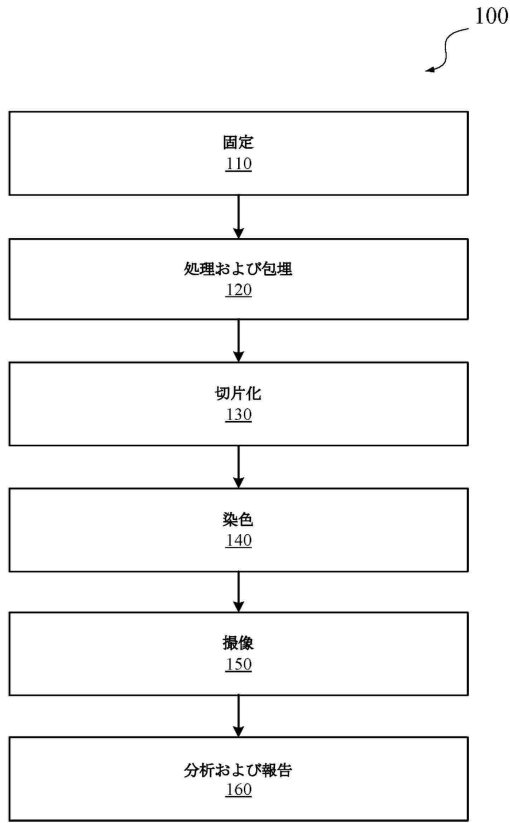
30

【0123】

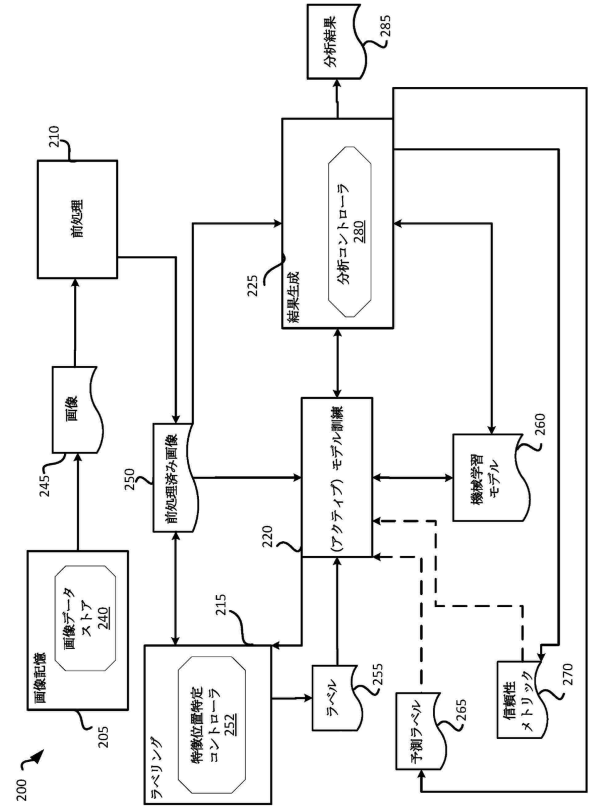
実施形態の完全な理解を提供するために、本明細書では具体的な詳細が与えられる。しかしながら、これらの具体的な詳細なしで実施形態が実施されることが理解されよう。例えば、回路、システム、ネットワーク、プロセス、および他の構成要素は、実施形態を不必要に詳細に不明瞭にしないために、ブロック図形式の構成要素として示されてもよい。他の例では、実施形態を不明瞭にすることを避けるために、周知の回路、プロセス、アルゴリズム、構造、および技術が不必要な詳細なしに示されてもよい。

40

【図面】
【図 1】

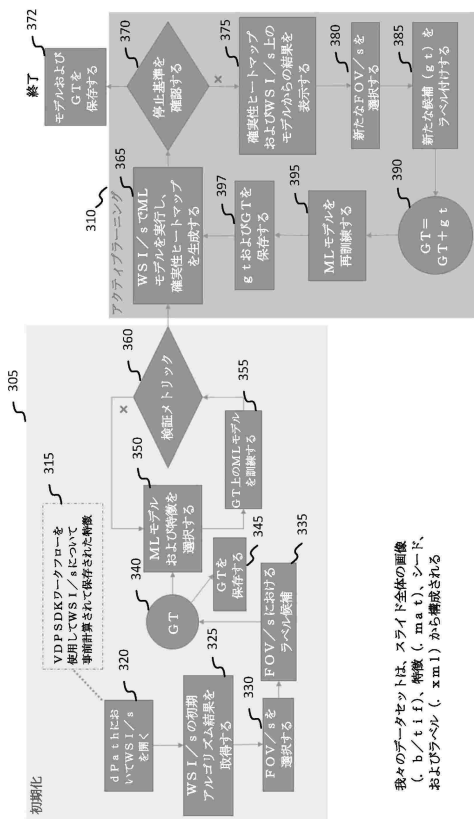


【図 2】

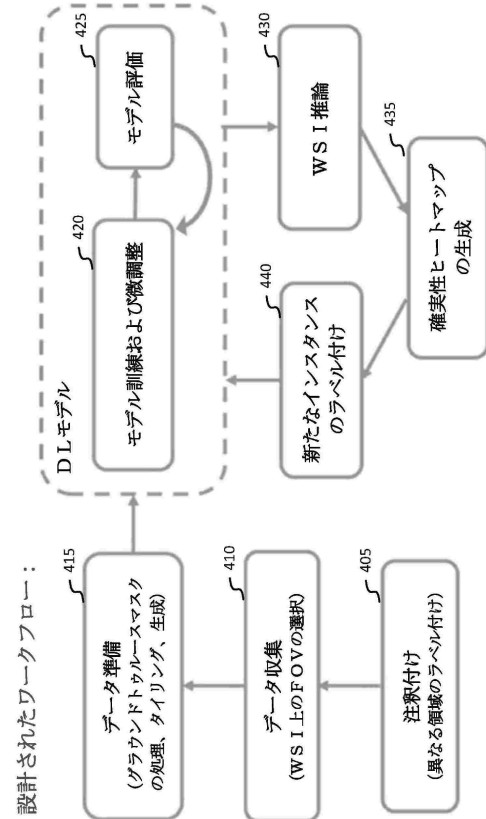


10
20

【図 3】

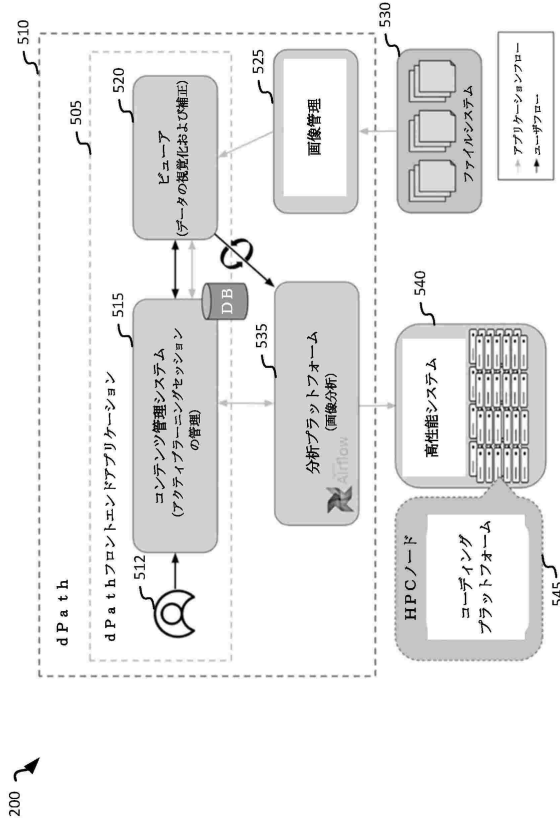


【図 4】



30
40
50

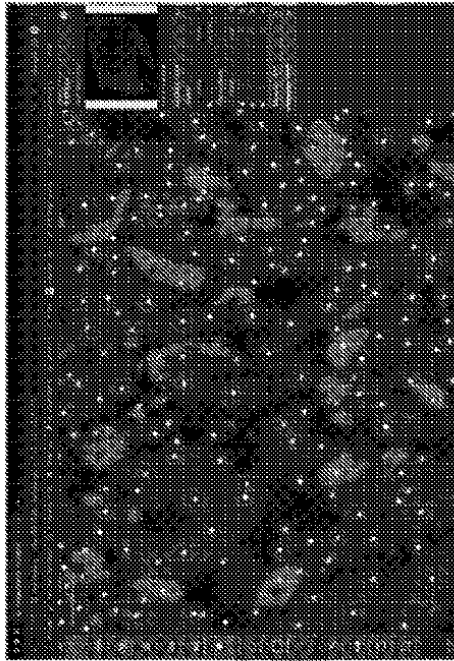
【図 5】



【図 6 A】



【図 6 B】



【図 6 C】




10

20

30

40

50

【 6 D】

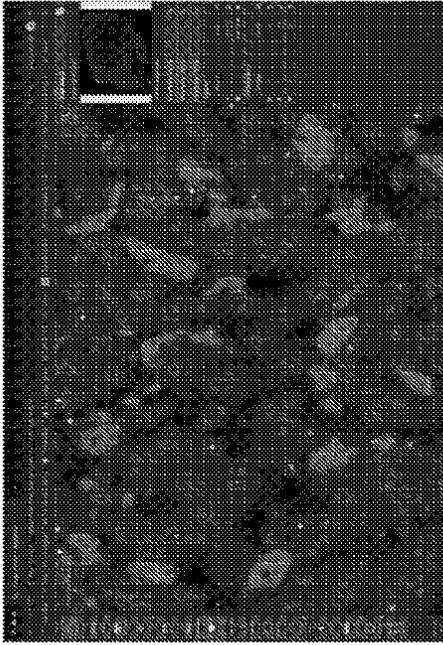


FIG. 6D

10

20

30

40

50

フロントページの続き

(51)国際特許分類

F I
G 0 1 N 33/48 Z

アメリカ合衆国アリゾナ州 8 5 7 5 5 , トゥーソン , イースト・イノベーション・パーク・ドライブ 1 9 1 0 , ベンタナ メディカル システムズ , インコーポレイテッド

(72)発明者 ホジャステ , メールヌーシュ

アメリカ合衆国アリゾナ州 8 5 7 5 5 , トゥーソン , イースト・イノベーション・パーク・ドライブ 1 9 1 0 , ベンタナ メディカル システムズ , インコーポレイテッド

(72)発明者 ラーセン , ジャスティン

アメリカ合衆国アリゾナ州 8 5 7 5 5 , トゥーソン , イースト・イノベーション・パーク・ドライブ 1 9 1 0 , ベンタナ メディカル システムズ , インコーポレイテッド

(72)発明者 マーティン , ジム・エフ

アメリカ合衆国アリゾナ州 8 5 7 5 5 , トゥーソン , イースト・イノベーション・パーク・ドライブ 1 9 1 0 , ベンタナ メディカル システムズ , インコーポレイテッド

(72)発明者 ムラリ , ニディン

アメリカ合衆国アリゾナ州 8 5 7 5 5 , トゥーソン , イースト・イノベーション・パーク・ドライブ 1 9 1 0 , ベンタナ メディカル システムズ , インコーポレイテッド

(72)発明者 シェイクザデー , ファヒーマ

アメリカ合衆国アリゾナ州 8 5 7 5 5 , トゥーソン , イースト・イノベーション・パーク・ドライブ 1 9 1 0 , ベンタナ メディカル システムズ , インコーポレイテッド

審査官 菊池 伸郎

(56)参考文献 国際公開第 2 0 1 9 / 1 0 8 8 8 8 (W O , A 1)

国際公開第 2 0 1 9 / 1 1 0 5 8 3 (W O , A 1)

特開 2 0 1 9 - 0 2 3 3 9 2 (J P , A)

特開 2 0 2 0 - 0 1 3 4 2 7 (J P , A)

(58)調査した分野 (Int.Cl. , D B 名)

G 0 6 T 7 / 0 0

G 0 6 V 1 0 / 8 2

G 0 6 V 2 0 / 6 9

G 0 1 N 3 3 / 4 8