(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局

(43) 国際公開日
2024年12月5日(05.12.2024)

WIPO | PCT

(10) 国際公開番号
WO 2024/247990 A1

(51) 国際特許分類:
G06V 10/774 (2022.01) *G06T 7/00* (2017.01)
A61B 5/00 (2006.01) *G16H 30/20* (2018.01)

(21) 国際出願番号： PCT/JP2024/019506

(22) 国際出願日： 2024年5月28日(28.05.2024)

(25) 国際出願の言語： 日本語

(26) 国際公開の言語： 日本語

(30) 優先権データ：
特願 2023-090399 2023年5月31日(31.05.2023) JP

(71) 出願人: デンカ株式会社(**DENKA COMPANY LIMITED**) [JP/JP]; 〒1038338 東京都中央区日本橋室町二丁目１番１号 Tokyo (JP).

(72) 発明者: 小鮒 弘幸 (**KOBUNA, Hiroyuki**); 〒1038338 東京都中央区日本橋室町二丁目１番１号 デンカ株式会社内 Tokyo (JP). 福田 皓一(**FUKUDA, Koichi**); 〒1038338 東京都中央区日本橋室町二丁目１番１号 デンカ株式会社内 Tokyo (JP). 板倉 聡志(**ITAKURA, Satoshi**); 〒1038338 東京都中央区日本橋室町二丁目１番１号 デンカ株式会社内 Tokyo (JP). 奥田 修二郎(**OKUDA, Shujiro**); 〒9500917 新潟県新潟市中央区天神１丁目１番地 プラーカ３ ２Ｆ ＨＵＢ ＳＴＡＴＩＯＮ ＫＥＮＴＯ内 合同会社ＳＯＡＲＯ内 Niigata (JP).

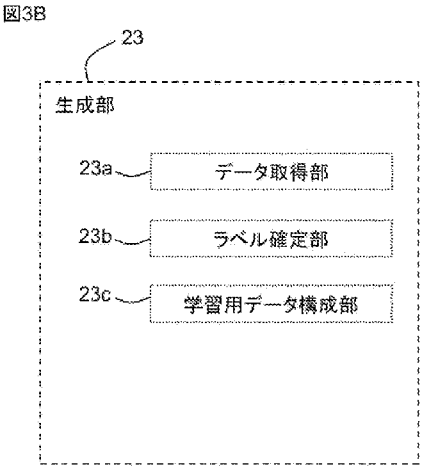
(74) 代理人: ＳＫ弁理士法人, 外(**SK INTELLECTUAL PROPERTY LAW FIRM** et al.); 〒1500021 東京都渋谷区恵比寿西二丁目２０番１７号 ヒューリック代官山ビル３階３０１号室 Tokyo (JP).

(81) 指定国(表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG,

(54) **Title:** PROGRAM, GENERATION METHOD, AND SYSTEM

(54) 発明の名称： プログラム、生成方法及びシステム

図3B



23 Generation unit
23a Data acquisition unit
23b Label finalizing unit
23c Learning data configuration unit

(57) **Abstract:** Provided are a program, a generation method, and a system for providing a novel learning data generation technique that is highly convenient. A program for generating learning data from a medical image is provided, the program causing a computer to execute generation processing wherein a plurality of pieces of information are associated with a region image that is an image of at least one region in the medical image. In the generation processing, the computer is caused, by means of a manual operation received via an operation input unit or automatic processing by the computer, to perform label association with the region image using the plurality of pieces of information, to generate the learning data from the medical image.

WO 2024/247990 A1

［続葉有］

KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, 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, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) 指定国(表示のない限り、全ての種類の広域保護が可能):ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, RU, TJ, TM), ヨーロッパ (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

添付公開書類:

一　国際調査報告(条約第21条(3))

(57) 要約:利便性の高い新規な学習用データ生成技術を提供するプログラム、生成方法及びシステムを提供する。医用画像から学習用データを生成するためのプログラムであって、前記プログラムは、生成処理を、コンピュータに実行させ、前記医用画像における少なくとも１つの領域の画像である領域画像に複数の情報が対応付けられ、前記生成処理は、操作入力部を介して受け付けた手動操作又はコンピュータの自動処理によって、コンピュータに、前記複数の情報により前記領域画像へのラベル対応付けを行わせて、前記医用画像から前記学習用データを生成させる、プログラムが提供される。

明　細　書

発明の名称 ： プログラム、生成方法及びシステム

技術分野

[0001]　本発明は、機械学習アルゴリズムの学習用データを生成するプログラム、生成方法及びシステムに関する。

背景技術

[0002]　特許文献１のデータ生成装置は、機械学習アルゴリズムの学習用データを生成する。学習用データは、複数の画像と、各画像に対応付けされたラベルとで構成される。

先行技術文献

特許文献

[0003] 特許文献1：国際公開第２０２１／２０６０５３号

発明の概要

発明が解決しようとする課題

[0004]　医用画像の学習用データ生成のために、専門知識を有する者が手作業で多数の画像に一つずつラベル付けを行うという実情がある。このような実情は、医用画像の判断に高度の専門性、正確性等が要求されることに起因する。しかし、多数の画像に手作業でラベル付けをする作業負担等が課題となっている。より利便性の高い学習用データ生成技術が望まれていた。

[0005]　本発明は、利便性の高い新規な学習用データ生成技術を提供するプログラム、生成方法及びシステムを提供する。

課題を解決するための手段

[0006]　本開示によれば、少なくとも以下の発明が提供される。

［１］医用画像から学習用データを生成するためのプログラムであって、前記プログラムは、生成処理を、コンピュータに実行させ、前記医用画像における少なくとも１つの領域の画像である領域画像に複数の情報が対応付けられ、前記生成処理は、操作入力部を介して受け付けた手動操作又はコンピュ

ータの自動処理によって、コンピュータに、前記複数の情報により前記領域画像へのラベル対応付けを行わせて、前記医用画像から前記学習用データを生成させる、プログラム。

［２］［１］に記載のプログラムであって、前記複数の情報は第１及び第２情報を含み、前記第１情報が第１ラベルであり、前記第２情報が第２ラベルであり、第１アノテーション及び第２アノテーションにより同一の前記領域画像に対応付けられた第１ラベル及び第２ラベルを取得する処理を、コンピュータに実行させる、プログラム。

［３］［１］又は［２］に記載のプログラムであって、前記複数の情報の各々がラベルであり、前記生成処理は、操作入力部を介して受け付けた手動操作又はコンピュータの自動処理によって、コンピュータに、同一の前記領域画像に対応付けられた複数のラベルから前記領域画像の確定ラベルを決定させて、前記医用画像から生成した学習データ画像と前記確定ラベルとを対応付けさせることで、前記学習用データを生成させる、プログラム。

［４］［１］に記載のプログラムであって、前記複数の情報は、第１及び第２情報を含み、前記領域画像は、第１領域画像および第２領域画像を含み、前記生成処理は、コンピュータに、前記第１情報により前記第１領域画像について決定したラベルを前記第２情報による選別を行った前記第２領域画像に対応付けさせて、前記学習用データを生成させる、プログラム。

［５］［４］に記載のプログラムであって、複数の前記領域画像を含む入力データを取得する処理を、コンピュータに実行させ、前記複数の前記領域画像に含まれる複数の前記第１領域画像は、第１及び第２医用画像の各々から生成され、前記第１及び第２医用画像は前記第１情報と対応付けられ、前記第１医用画像は、前記第１情報が予め定めた基準を満たす画像であり、前記第２医用画像は、前記第１情報が前記基準を満たさない画像であり、前記入力データの前記複数の前記領域画像に含まれる複数の前記第２領域画像を分類する機械推論を学習済判別モデルに行わせる判別処理を、コンピュータに実行させ、前記機械推論による前記領域画像の分類結果が、前記第２情報で

あり、前記生成処理は、コンピュータに、前記第１医用画像から生成された前記第１領域画像と同一の前記領域画像に含まれ、前記分類結果として第１分類グループに分類された前記第２領域画像に前記第１情報が前記基準を満たすことを表すラベルを対応付けさせ、且つ、前記第２医用画像から生成された前記第１領域画像と同一の前記領域画像に含まれ、前記分類結果として前記第１分類グループに分類された前記第２領域画像に前記第１情報が前記基準を満たさないことを表すラベルを対応付けさせて、前記学習用データを生成させる、プログラム。

［６］［４］又は［５］に記載のプログラムであって、前記医用画像は病理画像を含む、プログラム。

［７］［６］に記載のプログラムであって、前記第１情報は、遺伝子変異タイプ、ＴＭＢ－Ｈ（高腫瘍遺伝子変異量）タイプ、生化学的解析の結果、又は患者の年齢、性別、病歴、免疫組織化学的解析の結果、生活習慣もしくは予後予測の少なくとも１つを含む、プログラム。

［８］［１］～［７］のいずれか１つに記載のプログラムであって、前記複数の情報の少なくとも１つは、指標値を備えるか又は指標値であり、前記指標値は情報の内容の確からしさを表す、プログラム。

［９］［２］又は［３］を引用する、［８］に記載のプログラムであって、前記領域画像と対応付けられるラベルが、指標値を備えるか又は指標値であり、前記指標値はラベルの分類結果がどの程度確からしいかを表す、プログラム。

［１０］［５］を引用する［８］に記載のプログラムであって、前記機械推論により前記学習済判別モデルが前記領域画像の分類結果の確からしさを表す確率値を出力し、前記確率値を前記指標値に使用する、プログラム。

［１１］医用画像から学習用データを生成する生成方法であって、生成ステップを、コンピュータに実行させ、前記医用画像における少なくとも１つの領域の画像である領域画像に複数の情報が対応付けられ、前記生成ステップは、操作入力部を介して受け付けた手動操作又はコンピュータの自動処理に

よって、コンピュータに、前記複数の情報により前記領域画像へのラベル対応付けを行わせて、前記医用画像から前記学習用データを生成させる、方法。

［１２］医用画像から学習用データを生成するシステムであって、生成部を備え、前記医用画像における少なくとも１つの領域の画像である領域画像に複数の情報が対応付けられ、前記生成部は、操作入力部を介して受け付けた手動操作又はコンピュータの自動処理によって、前記複数の情報により前記領域画像へのラベル対応付けを行うことで、前記医用画像から前記学習用データを生成する、システム。

[0007] 本発明によれば、領域画像に対応付けられた複数の情報により領域画像へのラベル対応付けを行うという新規な技術が提供される。これにより高い利便性で学習用データを生成可能である。

図面の簡単な説明

[0008] [図1]実施形態に係るシステム１（サーバ１）、各ユーザ端末３、及び診断装置４の一例を示す図である。

[図2]図２Ａは、サーバ１のハードウェア構成の一例を示すブロック図である。図２Ｂは、ユーザ端末３のハードウェア構成の一例を示すブロック図である。

[図3]図３Ａは、第１実施形態のサーバ１の機能構成を示すブロック図である。図３Ｂは、第１実施形態の生成部２３の機能ブロック図である。

[図4]第１実施形態の制御動作を表す処理フローの第１の例（グリッド方式、手動確定処理）である。

[図5]第１実施形態の制御動作を表す処理フローの第２の例（グリッド方式、自動確定処理）である。

[図6]第１実施形態の制御動作を表す処理フローの第３の例（指定枠方式、手動確定処理）である。

[図7]図７Ａは、図６（指定枠方式）の選択画面２７でグリッド表示をオンにした状態を例示している。図７Ｂは、図７Ａのグリッドに従って領域画像変

換を行った結果の一例である。

[図8]第１実施形態の制御動作を表す処理フローの第４の例（指定枠方式、自動確定処理）である。

[図9]図９Ａは、図７Ａで指定枠が３個の場合を例示している。図９Ｂは、図９Ａのグリッドに従って領域画像変換を行った結果の一例である。

[図10]図１０Ａはラベルが指標値を備える画像の例であり、図１０Ｂはラベルが指標値である画像の例である。

[図11]図１１Ａは、第２実施形態のサーバ１の機能構成を示すブロック図である。図１１Ｂは、第２実施形態の生成部２３の機能ブロック図である。

[図12]第２実施形態の制御動作を表す処理フローの第１の例である。

[図13]第２実施形態の制御動作を表す処理フローの第２の例（指標値を用いる処理）である。

発明を実施するための形態

[0009] 以下、図面を用いて本発明のいくつかの実施形態について説明する。以下に示す実施形態中で示した各種特徴事項は、互いに組み合わせ可能である。また、各特徴事項について独立して発明が成立する。

[0010] ＜１．用語等＞

（１－１．アノテーション）

機械学習分野におけるアノテーションという用語は、文脈に応じて、画像へのラベル付けの注釈行為を指したり、付与されたラベルの群を指したりする。注釈行為としてのアノテーションは、「機械学習用の対象情報（電子情報）に正解データとしての注釈情報を対応付ける行為」を意味している。画像、動画、音声などの各種データがアノテーションの対象となりうるが、例えば画像アノテーションには物体検出、領域画像抽出及び画像分類が含まれる。画像アノテーションは、アノテーション対象に対してラベルを対応付ける。アノテーション対象は、各々の画像（画像の全体）でもよく、当該画像中の領域画像（エリア、当該画像の全体も含む）でもよい。ラベルは、タグ又はメタデータとも称される。ラベルが分類内容（分類結果）を表して教師

あり学習時における正解データとなる。実施形態での「アノテーション」は、異なる定義を明記した場合や文脈上明らかな矛盾が生ずる場合を除いて、「画像において少なくとも１つの領域が指定される場合に、この領域の画像である領域画像にラベルを対応付けること」をいうものとする。

[0011]（１－２．アノテーション対象画像）

実施形態でアノテーションの対象とされる画像は、少なくとも、「医用画像（Ｍｅｄｉｃａｌ　ｉｍａｇｉｎｇ）」を含むものとする。医用画像は、「主に病気の診断及び治療のために、人体内部の構造や機能を画像として可視化した画像」を含む。医用画像には、具体的には、単純Ｘ線撮影、コンピュータ断層撮影（ＣＴ）、核磁気共鳴画像法（ＭＲＩ）、超音波断層画像（ＵＳ）、核医学検査、及び血管造影などが含まれてもよい。本開示に係る医用画像は、上記各種画像に加えて「内視鏡画像」及び「病理画像」を更に含み得る。内視鏡画像は、例えば食道、胃、腸、その他の消化器系画像を含んでもよい。病理画像は、具体的には、「内視鏡や手術等で採取した組織を、（１）劣化しないように固定し、（２）観察が容易になるように染色し、（３）数ミクロン程度に薄切した標本を拡大撮影した（顕微鏡）画像」であってもよい。

[0012] 本発明は任意の医用画像に対して同様に適用可能であり、任意のラベル分類種別、分類内容を設定可能である。ラベル分類種別は、例えば病理の種別に関する情報である。一例として、病理の種別は、「腫瘍」又は「超変異」のいずれか一方であってもよい。例えば、医用画像の各々が腫瘍の存在が疑われる画像である場合には病理の種別として「腫瘍」が設定されてもよく、その場合、ラベルの分類内容は「腫瘍」又は「非腫瘍」とすることができる。複数の医用画像の各々が超変異の存在が疑われる画像である場合には病理の種別として「超変異」が設定されてもよく、その場合、ラベルの分類内容は「超変異」又は「非超変異」とすることができる。以下の第１及び第２実施形態では、医用画像の一例として病理画像を採用し、ラベルの分類種別は一例として腫瘍かどうかの分類であり、ラベルの分類内容は「腫瘍」又は「

非腫瘍」であるものとする。

[0013] 実施形態では、医用画像の一例として病理画像を採用し、ラベルの分類種別は腫瘍かどうかの分類であり、ラベルの分類内容は「腫瘍」又は「非腫瘍」である。

[0014] （１－３．領域及び領域画像）

医用画像に領域（エリア）を指定する方法は、様々である。実施形態では、幾つかの例でグリッド方式を採用し、他の幾つかの例で指定枠方式を採用する。

[0015] 「グリッド方式」は、グリッド（つまり格子状の境界線）により、画像における所望の位置及び範囲に、領域を指定する。グリッドが複数のマスを有する場合、グリッドのマスの１つが１つの領域を規定可能である。グリッドは均一間隔でもよく、各領域が均一サイズの矩形に指定されてもよく、このような領域は「タイル」と称してもよい。グリッド間隔は、規定であるか又はユーザが任意指定可能である。なおここで言う「グリッド」は、広義の意味であり、矩形格子に限定されず、平面敷き詰め可能な任意多角形を周期配列した格子を採用可能である。グリッドは、例えばハニカムグリッド等でもよい。ここでいう「ハニカムグリッド」は広義の意味であり、例えば正六角形ハニカムでもよいが、例えば正三角形又は平行四辺形のハニカム構造でもよく、平行六角形（３組の対辺が平行だが各組の間で辺の長さが不一致の六角形）のハニカム構造でもよい。実施形態では、図４及び図５の例でグリッド方式のアノテーションを採用している。

[0016] 指定枠方式は、ユーザが任意に指定する指定枠により、画像における所望の位置及び範囲に、領域を指定する。指定枠はバウンディングボックスと称されることがある。指定枠は、任意のサイズ及び任意の形状にユーザが設定可能である。指定枠は、例えば矩形でもよいがこれに限られず、任意の多角形又は閉曲線でもよい。実施形態では、図６及び図８の例で指定枠方式のアノテーションを採用している。

[0017] 以下の説明では、便宜上、グリッドで指定された領域を、理解を容易にす

るために「グリッド領域」と称することがある。また、便宜上、指定枠で指定された領域を、理解を容易にするために「指定枠領域」と称することがある。其の一方で、単に「領域」と記載した場合には、グリッド領域と指定枠領域のうち片方を指すことが文脈上明らかな場合等を除いて、グリッド領域と指定枠領域との両方を含むことを意図することがある。

[0018] 画像に領域が指定された場合、その領域に対応する画像を「領域画像」とも称す。領域画像がグリッド方式で均一矩形サイズに指定された画像である場合、この領域画像はタイル画像と称されてもよい。領域及び領域画像のサイズ（縦方向及び横方向のピクセル数）は、学習モデルに入力可能な学習用データと同一サイズであってもよい。画像に複数の指定枠が設定された場合、２つの指定枠が完全一致していれば指定枠領域が互いに一致するが、各指定枠が相違していれば指定枠領域も相違する。指定枠同士の重なりがある場合、複数の指定枠領域が重複した重複領域が発生する。

[0019] 実施形態のプログラム、方法及びシステムはグリッド方式と指定枠方式とのうち一方又は両方のアノテーションを採用可能である。もし両方を採用した場合はユーザの設定操作により及び／又はコンピュータ自動処理により今回のアノテーションでどの方式を使用するかが設定されてもよい。

[0020] （１－４．学習用データ）

通常、アノテーションは学習用データセットを構成するために行われる。医用画像にアノテーションを行う場合、学習用データセットは、医用画像を切り出した「学習データ画像」と、学習データ画像に対応付けられたラベルとで構成される。学習データ画像のサイズ（縦方向及び横方向のピクセル数）は、学習モデルに入力可能なサイズに設定される。「学習データ画像」は、グリッドに沿って医用画像を切出したものでもよく、指定枠（バウンディング枠）で医用画像の一部を切出したものでもよい。学習データ画像は、アノテーション時の領域画像に従って医用画像を切出したものでもよく、つまり、アノテーション時の領域画像と学習データ画像とが一致してもよい。しかしこれに限られず、学習データ画像は、アノテーション時の領域画像とは

異なってもよい。一例として、指定枠により医用画像にアノテーションがされた場合に、これをグリッド（例えば均一間隔の矩形グリッド）で切り出すことで、学習データ画像を生成してもよい。他の例として、ハニカムグリッド等でアノテーションがされた場合に、これを他のグリッド（例えば均一間隔の矩形グリッド）で切り出すことで、学習データ画像を生成してもよい。

[0021] 注意されたい点として、アノテーション作業中の段階又は学習データ生成処理の前段階では、医用画像の切り出しは必要ではない。例えば後述の作業画面２６及び選択画面２７により医用画像の各領域画像が表示されることがあるが、この表示段階では医用画像の切り出しは必要ではない。アノテーションが完了し、その後に学習用データセット作成指示が出された段階で、医用画像の切り出しを実施してもよい。

[0022] ＜２．第１実施形態＞

（２－１．システム１及び診断装置４の概要）

図１に示すように、第１実施形態のシステム１、ユーザ端末３及び診断装置４は、無線又は有線の通信ネットワーク２を介して互いに通信可能である。図１のシステム１は、医用画像から学習用データを生成可能である。システム１は、コンピュータであり、具体的にはサーバである。サーバがプログラムを実行することでシステム１が提供される。便宜上、システム１をサーバ１と表記することがある。システム１を利用するユーザＵ００１、Ｕ００２は例えば病理医である。

[0023] システム１は、学習用データを生成する機能を持つ生成部２３（図３Ａ参照）を備える。生成部２３は、入力された画像（具体的には医用画像）を用いて、学習用データを生成可能である。学習用データは、診断装置４が備える学習モデルを学習させるための教師データとして使用可能である。診断装置４は、生成された学習用データを用いて学習した学習済モデルを用いて、医用画像に病態的な異常が存在するか否かの判定を行う。

[0024] システム１は、記憶部２０に記憶された医用画像（実施形態では病理画像）と、グリッド（複数の領域の境界線）とを重ねて画面に表示可能である。

画面の一例は、ユーザ端末３の表示デバイス３４（図２Ｂ参照）上の画面である。システム１は、ユーザによる領域画像へのアノテーション操作を受け付ける。つまりシステム１は、画面に表示された領域画像１つ１つについて、ユーザＵ００１、Ｕ００２からラベルの入力を受け付ける。ラベル入力操作用ＧＵＩとして、システム１は後述の図４における作業画面２６を生成し、ユーザ端末３に表示させる。

[0025] システム１は、１枚の病理画像から学習用データを生成してもよいし、複数の病理画像から学習用データを生成してもよい。複数の病理画像から学習用データを生成する場合、システム１は、病理画像を複数の領域画像に分割してラベルの入力を受け付ける処理を、病理画像ごとに繰り返す。全ての病理画像についてユーザによるラベルの付与が完了すると、学習用データを生成する。例えばシステム１は、グリッドに従って医用画像を切り出すことで学習データ画像を生成し、この学習データ画像の画像データと各領域画像に付与されたラベルとを対応づけることで学習用データを生成する。生成された学習用データは、診断装置４に送られる。

[0026] 診断装置４は、システム１から送られてきた学習用データを用いて学習モデルを学習（訓練）させる。学習モデルの具体的構成に限定はなく、例えば、機械学習により所定の能力を持たせることが可能なニューラルネットワークでもよい。診断装置４は、機械学習により生成された学習済みモデルに診断対象の病理画像を入力し、学習済みモデルからの出力結果に基づいて、当該病理画像に病態的な異常が存在するか否かを判断する。

[0027] システム１は、アノテーションを支援する支援機能として、選択画面２７を提供可能である。当該支援機能は後ほど詳細に説明する。

[0028] （２－２．システム１のハードウェア構成）

図２Ａは、システム１のハードウェア構成例を示す図である。システム１は、ＣＰＵ（Central Processing Unit）、ＧＰＵ（Graphical processing unit）等のプロセッサ１１、メモリ、ＨＤＤ（Hard Disk Drive）及び／又はＳＳＤ（Solid State Drive）等の記憶装置１２、有線又は無線通信を行う通

信ＩＦ（Interface）１３を備える。システム１は、通信ＩＦ１３を介して、ユーザ端末３からのユーザ入力操作を受け付けたり、ユーザ端末３の表示デバイス３４（図２Ｂ参照）に各種画面を表示させたりする。

[0029]（１－３．ユーザ端末３のハードウェア構成）

図２Ｂに示すように、ユーザ端末３は、一例として、プロセッサ等を含む制御部３１と、不揮発性メモリ等を含む記憶部３２と、無線通信又は／及び有線通信を行うための通信部３３と、表示デバイス３４と、スピーカ３５と、マイク３６と、カメラ３７と、操作デバイス３８とを備える。操作デバイス３８は、操作入力部の一例であり、キーボードやマウス等の任意の入力デバイスである。ユーザ端末３の具体的構成に限定はなく、各種の携帯端末、スマートフォン、タブレット端末、あるいはノートＰＣなどでもよい。ユーザ端末３がタッチパネルディスプレイを持つ場合には、表示デバイス３４と操作デバイス３８とがタッチパネルディスプレイに一体化されてもよい。操作デバイス３８とマイク３６は、ユーザ端末３でユーザが各種の入力操作を行うための入力インターフェイス（あるいは入力手段）となる。

[0030]（２－４．システム１および生成部２３の機能ブロック構成）

図３Ａは、システム１の機能ブロック構成例を示す図である。システム１は、記憶部２０と、表示制御部２１と、入力部２２と、生成部２３とを備える。記憶部２０は、システム１が備える記憶装置１２を用いて実現することができる。システム１のプロセッサ１１が、記憶装置１２に記憶されたプログラムを実行することにより、表示制御部２１と入力部２２と生成部２３が提供される。当該プログラムは、ネットワーク上で提供可能でもよいし、記憶媒体に格納されてもよい。当該プログラムを格納した記憶媒体は、コンピュータ読み取り可能な非一時的な記憶媒体（Non-transitory computer readable medium）であってもよい。非一時的な記憶媒体は特に限定されないが、例えば、ＵＳＢメモリ又はＣＤ－ＲＯＭ等の記憶媒体であってもよい。

[0031]（２－４－１．記憶部２０の構成）

記憶部２０は各種のデータベース（ＤＢ）を備える。記憶部２０は、一例

として、アカウントＤＢと、画像ＤＢと、領域ＤＢと、アノテーションＤＢと、ラベルＤＢと、学習用データＤＢとを記憶している。各ＤＢは格納データに識別ＩＤが付与されていて、互いに関連するデータは相互参照可能である。

[0032] アカウントＤＢは、各ユーザのユーザＩＤ、ユーザ詳細情報、及びログイン履歴などを格納する。ユーザ詳細情報は、メールアドレス等を含んでもよく、アカウントに紐づけた名称等（ニックネーム、アバターアイコン等）を含んでもよく、及び／又は、例えば氏名及び所属等のより個人的、具体的な情報を含んでもよい。

[0033] 画像ＤＢは、学習用データの生成に用いる１つ以上の医用画像（例えば病理画像）を格納する。領域ＤＢは、各医用画像における領域画像を特定する情報を格納する。一例として、領域ＤＢは、領域画像ＩＤとともに、領域画像の設定先の医用画像ＩＤと、画像座標データとを格納する。画像座標データは、各領域画像が医用画像のどの範囲を占めるのかを規定する。

[0034] アノテーションＤＢは、システム１を利用して行われた各アノテーションを互いに区別するためのアノテーション情報を、アノテーションＩＤとともに格納する。アノテーション情報は、どの医用画像に対して、何時、誰が、どのタイプのアノテーションを行ったのかを特定可能な情報であることが好ましい。アノテーション情報は、好ましくは、「アノテーション主体情報」を格納可能である。アノテーション主体情報は、「誰がアノテーションを行ったか（つまりアノテーション主体）」を特定するための識別情報である。アノテーション主体情報は、例えばユーザＩＤ（例えばログインアカウント）でもよく、例えばアカウントＤＢに格納されるユーザ詳細情報中の任意の１つ以上の情報でもよい。アノテーション情報は、アノテーションの開示可否（つまり、異なるユーザＩＤの他者の選択画面２７に開示してよいかどうか）を含んでもよく、アノテーションの開示条件（つまり他者に開示する場合にどのような条件で開示するか）を含んでもよい。

[0035] ラベルＤＢは、各アノテーションにより領域の画像である領域画像に対応

付けされたラベルの情報を、アノテーションＩＤと紐づけて格納する。ラベルＤＢは、各ラベルＩＤとラベル詳細情報とを対応付けて格納している。ラベル詳細情報は、ラベルの分類種別及び分類内容（例えば種別が腫瘍分類で、ラベル分類内容が腫瘍である、等）を含んでいる。

[0036] 学習用データＤＢは、生成された学習用データを格納する。ユーザは、学習用データＤＢを参照することで、任意の学習用データをダウンロード可能である。

[0037] 少なくとも画像ＤＢと領域画像ＩＤとアノテーションＩＤとラベルＩＤが互いに紐付けられることが好ましい。これにより、アノテーションＩＤを指定すれば任意のアノテーションの内容（つまり医用画像、領域画像、ラベル等）をアノテーションＤＢから簡単に呼出可能である。なお、上記のデータ管理及びＤＢ構成は一例に過ぎない。実際のハードウェア及び／又はソフトウェア等の仕様に応じて任意のデータ管理及びＤＢ構成を採用可能である。なお、もしアノテーション単位での管理を行わない場合には、アノテーションＤＢは省略可能である。

[0038] （２－４－２．その他の各機能部の構成）

表示制御部２１は、ユーザ端末３の表示デバイス３４に各種画面を表示させる表示制御を実行する。例えば、表示制御部２１は、グリッド（領域の境界線）を医用画像に重ねた作業画面２６（図４参照）を表示可能でもよい。図４の作業画面２６から伸びる吹き出しに、医用画像Ｑがグリッドで４分割された画面を例示する。実際には、より大きな医用画像Ｑがグリッドで仕切られて、十数個×十数個の領域画像が作業画面２６に表示される。各領域画像には領域画像ＩＤが対応付けられる。図４の例では、医用画像Ｑの左上の１マスを、領域画像Q_1とも称す。その一方で、表示制御部２１は、医用画像Ｑに指定枠を重ねて表示可能でもよい。

[0039] 入力部２２は、ユーザ端末３の操作デバイス３８からの各種入力を受け付ける。例えば、入力部２２は、ラベル入力操作（つまりユーザがアノテーションを行う際の操作デバイス３８の操作）を受け付ける。ラベル入力操作は

、複数の領域の画像（領域画像）の各々にどのラベルを対応付けるかの入力操作である。入力部２２は、各領域画像に対応付けされたラベルを、ラベルＤＢに格納する。入力部２２は、例えば、各領域画像を一意に識別する領域画像ＩＤと、各領域画像に対応付けされたラベルとを対応づけてラベルＤＢに格納するようにしてもよい。ラベルの入力は、例えば、画面上に表示された各領域画像の中からラベルを入力する領域画像の選択を受け付け、当該領域画像に対応付けるラベルの指定を受け付けることで行われてもよい。

[0040] 生成部２３は、学習モデルを学習させるための学習用データを生成する。例えば、生成部２３は、ラベルＤＢから領域画像ＩＤとラベルとを取得すると共に、画像ＤＢに格納されている医用画像から、領域画像ＩＤに対応する領域画像を切り出すことで学習データ画像を生成する。生成部２３は、学習データ画像データと、領域画像ＩＤに対応付けて記憶されたラベルとを組み合わせることで、学習用データを生成する。

[0041] 図３Ｂは、生成部２３の機能ブロック構成である。生成部２３は、データ取得部２３ａと、ラベル確定部２３ｂと、学習用データ構成部２３ｃとを備える。各機能ブロックの動作は、図４及び図５の処理フロー中で説明する。

[0042] （２－５．処理フロー）

図４～図６及び図８それぞれは、第１実施形態の処理フローの具体例であり、学習用データを生成するプログラム及び生成方法を提供する。図４及び図５の処理フローでは、グリッド方式のアノテーションが採用される。図６及び図８の処理フローでは、指定枠方式のアノテーションが採用される。

[0043] 第１実施形態では、生成処理をコンピュータ（システム１）に実行させる（図４のＳ１０３～Ｓ１０６、又は図５のＳ１０３、Ｓ１１５及びＳ１０６を参照）。この生成処理は、手動操作又は自動処理によって、コンピュータ（システム１）に、領域画像に対応付けられた複数の情報により領域画像とラベルとの対応付けを行わせて、これにより学習用データを生成させる。後述するように、第１実施形態では、各領域画像に第１及び第２情報が対応付けられる。領域画像に対応付けられた複数の情報により当該領域画像のラベ

ル対応付けを行うことができるので、高い利便性で学習用データを生成可能である。

[0044] ここでいう「複数の情報」は、第１実施形態と第２実施形態とで異なる情報を採用している。第１実施形態では、上記の「複数の情報」が第１及び第２情報を含み、この第１及び第２情報として第１及び第２ラベルが採用される。具体的には、第１アノテーションによる第１ラベル及び第２アノテーションによる第２ラベルを取得する処理を、コンピュータ（システム１）に実行させる。この処理は、図４ではステップＳ１００で包括的に記述されており、以下に詳述する。

[0045] 第１実施形態では、医用画像における領域画像に対応付けられた複数のラベルが、互いに同一の分類内容である場合がある。このような場合を単に「ラベル一致」とも称し、領域画像に対応付けられて且つ分類内容が同じ複数のラベルを「一致ラベル」とも称す。その一方で、領域画像に、異なる分類内容のラベルが対応付けられる場合、つまり複数のアノテーションによる判断が分かれる場合がありうる。このような場合を単に「ラベル不一致」とも称し、領域画像に対応付けられて且つ分類内容が異なる複数のラベルを「不一致ラベル」とも称す。

[0046] （２－５－１．共通の処理ステップＳ１００）

図４～図６及び図８の処理フローでは、先ずシステム１にログイン（Ｓ３００ａ）したユーザＵ００１が、システム１（表示制御部２１）から作業画面２６を受信する（Ｓ１０１ａ）。作業画面２６はアノテーション作業つまりラベル入力操作用のＧＵＩである。ユーザＵ００１がこの作業画面２６を介して「第１アノテーション」を行うことで、各領域画像へのラベル入力操作が実施される（Ｓ３０１ａ）。ステップＳ３０１ａでユーザＵ００１の操作により例えば第１ラベルLa_1とグリッド領域画像Q_1とが対応付けされると、システム１（入力部２２）は当該対応付けをラベルＤＢに登録する（Ｓ１０２ａ）。医用画像Ｑの更に他のグリッド領域画像Q_2、Q_3・・・にも同様のラベル対応付けが行われる。

[0047] 同様に、システム１にログイン（Ｓ３００ｂ）したユーザＵ００２は、システム１から作業画面２６を受信する（Ｓ１０１ｂ）。ユーザＵ００２がこの作業画面２６を介して「第２アノテーション」を行うことで、各領域画像へのラベル入力操作が行われる（Ｓ３０１ｂ）。ステップＳ３０１ｂでは、ユーザＵ００２の操作により例えば第２ラベルLb_1と領域画像Q_1とが対応付けされて、システム１は当該対応付けをラベルＤＢに登録する（Ｓ１０２ｂ）。医用画像Ｑの更に他のグリッド領域画像Q_2、Q_3・・・にも同様のラベル対応付けが行われる。

[0048] 図４及び図５の処理フローの場合、ステップＳ１００でグリッド領域画像にラベル入力が行われる。図６及び図８の処理フローの場合、ステップＳ１００で、指定枠により指定枠領域画像が設定され、この指定枠領域画像にラベル入力が行われる。

[0049] ここで、ステップＳ３０１ａ、Ｓ３０１ｂ各々でラベル入力が完了した領域画像については、入力されたラベルの分類内容を示す情報が領域画像に重ねて表示可能とされる。第１実施形態では、ラベルLa_1又はLb_1の一例として、「Ｔ」と「Ｎ」という文字でラベル分類内容を表示する（Ｓ３０１ａ、Ｓ３０ｂ）。領域画像が腫瘍細胞の存在が疑われる画像であることを示すラベルが付与された領域については、「Ｔ」の文字が表示される。領域が腫瘍細胞の存在が疑われる画像ではないことを示すラベルが付与された領域画像については、「Ｎ」の文字が表示される。図４の領域画像Q_1では、第１及び第２アノテーションで付与したラベル内容が相違しており、ユーザＵ００１、Ｕ００２で判断が分かれた場合を例示している。

[0050] 図４では説明便宜上１つのグリッド領域画像Q_1を例示的に図示している。しかし実際には多数のグリッド領域画像Q_k（即ちQ_1、Q_2・・・Q_n）を対象にアノテーションが行われる。ｋは、説明の便宜上導入した領域画像識別子であり、具体的にはｋ＝１、２、・・・ｎ（ｎは整数）であり、領域画像ＩＤと対応している。第１アノテーションにより、一つ以上の領域画像Q_1～Q_nの各々に、第１ラベルLa_1～La_nが対応付けられる。第２アノテーショ

ンにより、一つ以上の領域画像Q_1～Q_nの各々に、第２ラベルLb_1～Lb_nが対応付けられる。以下、任意の領域画像やラベルを指すことを目的として、領域画像Q_k、第１ラベルLa_k、又は第２ラベルLb_kのように記載することがある。

[0051] （２－５－２．第１の処理フロー例）

図４は、第１実施形態の制御動作を表す処理フローの第１の例（手動確定処理）である。図４の例は、ユーザの手動選択操作を受け付けることでラベル確定を行う。前述のステップＳ１００により第１及び第２アノテーションによる第１及び第２ラベルが、すでにラベルＤＢに登録されている。この段階で、学習用データを欲するユーザ（図４の例ではユーザＵ００２）が、操作デバイス３８を介してシステム１に学習用データ生成指令を発する（Ｓ３０２）。この学習用データ生成指令では、今回の学習用データ生成に必要な入力データを指定可能である。一例として、学習用データの採用対象として、医用画像ＩＤが指定可能されてもよく、領域画像ＩＤの範囲が指定可能とされてもよく、アノテーション呼出モードが設定可能とされてもよい。

[0052] 学習用データ生成指令（Ｓ３０２）に応答して、システム１は、学習用データの生成処理を開始する（Ｓ１０３）。図４の例では、生成処理は、ステップＳ１０３～Ｓ１０６を含む。まず生成部２３のデータ取得部２３ａ（図３Ｂ参照）が、記憶部２０の各ＤＢを参照する。データ取得部２３ａは、学習用データ生成指令の内容に応じて、今回の学習用データに含むべき各領域画像Q_kを読み出し、その一方で、領域画像ＩＤでラベルＤＢを検索することで各領域画像Q_kに対応付けられた各ラベルを読み出す。ここで読み出されるラベルは第１及び第２ラベルLa_k、Lb_kを含む。

[0053] データ取得部２３ａは、アノテーション呼出モードに応じて異なるデータ取得処理をする。アノテーション呼出モードは、一例として指定呼出モード、全部呼出モード、又は層別呼出モードを採用可能であり、これらのうち２つ以上をユーザが選択的に使用可能でもよい。

[0054] 指定呼出モードでは、アノテーションＩＤ等で所望のアノテーションが指

定され、指定されたアノテーションの領域画像ＩＤ及びラベルＩＤ等が参照される。全部呼出モードは、呼出時点で医用画像に対応付けられた全アノテーションが参照され、当該全アノテーションの領域画像ＩＤ及びラベルＩＤ等が参照される。

[0055] 層別呼出モードでは、指定条件に該当するアノテーションのみを抽出して、抽出したアノテーションの領域画像ＩＤ及びラベルＩＤ等が参照される。層別の種類は様々に設定可能である。例えば、アノテーション主体情報が所定条件に該当するか否かに基づく層別でもよい。例えば実績値が所定値以上であるアノテーションのみを抽出してもよく、これにより高実績で信頼の置けるアノテーションのみを利用可能である。例えば、呼出条件の一つとして期間指定が可能とされてもよく、指定期間内に登録（具体的には、新規登録又は編集最終登録）があったアノテーションのみを抽出してもよい。他の例として指標値を利用してもよく、例えばアノテーション内での全ラベルの要約統計量（例えば、平均指標値或いは全ラベルの指標値の中央値）が所定閾値以上のアノテーションのみを抽出してもよく、これによりアノテーション主体がある程度の確信度を持って作成したアノテーションのみを利用可能となる。

[0056] 次に、ラベル確定部２３ｂが、ラベル確定処理（Ｓ１０４、Ｓ１０５）を実行する。ラベル確定部２３ｂは、ステップＳ１０３で取得したデータに基づいて選択画面２７を生成し、ユーザ端末３に送信する（Ｓ１０４）。選択画面２７は、各領域画像$Ｑ_{k}$に対応付けられた複数のラベルのうち何れを確定ラベルとすべきかを、ユーザＵ００２に選択させるためのＧＵＩである。図４では、一例として領域画像$Ｑ_{1}$において第１及び第２ラベル$Ｌａ_{1}$、$Ｌｂ_{1}$のどちらかを選択するようにシステム１が操作を受け付ける。例えばマウスポインタＰでクリックする等でユーザＵ００２が所望のラベルを確定可能である（Ｓ３０３）。実際には選択画面２７に複数の領域画像$Ｑ_{k}$が網羅的に並べられてもよい。図４の例では、第２ラベル$Ｌｂ_{1}$がクリックされたものとする。ラベル確定部２３ｂは、ユーザが選択したラベル（この例では第２ラベル

Lb_1）を確定ラベルとして確定し、この確定ラベルと当該領域画像Q_1とを対応付ける。このような操作及び処理が、学習用データに必要な領域画像Q_k各々について繰り返される。

[0057] 必要な領域画像Q_kと確定ラベルとが揃った段階で、処理はステップＳ１０６に進む。ステップＳ１０６では、生成部２３の学習用データ構成部２３ｃが、医用画像Ｑをグリッドに従って切り出して、グリッドによる領域画像と同一サイズの学習データ画像を生成する。学習用データ構成部２３ｃは、学習データ画像とラベルとを対応付けた学習用データを構成してＤＢに登録する（Ｓ１０６）。ユーザＵ００１は学習用データを任意にダウンロード可能となる。

[0058] 以上説明したように、図４の処理フローは生成処理をコンピュータ（システム１）に実行させる。生成処理は、操作デバイス３８を介して受け付けた手動選択操作（Ｓ３０３）に従って、コンピュータ（システム１）に、領域画像Q_kに対応付けられた複数のラベルLa_k、Lb_kにより領域画像Q_kごとの確定ラベルを決定させて（Ｓ１０５）、確定ラベルと領域画像とを対応付けさせることで、医用画像から学習用データを生成させる（Ｓ１０６）。第１アノテーション（Ｓ３０１ａ）及び第２アノテーション（Ｓ３０１ｂ）の各々によるラベル（例えばLa_1、Lb_1）に基づいて各領域画像Q_kの確定ラベルを手動で決定可能な仕組みを、ユーザにとって利便性の高い形で提供可能である。複数のアノテーションを考慮に入れつつ学習用データを効率よく生成可能であり、これにより利便性が向上する。

[0059] なお、ステップＳ１０４で、ラベル一致とラベル不一致との両方を区別することなく、各領域画像Q_kに対してユーザのラベル選択操作を受け付けてもよいが、変形例として一括確定処理が設けられてもよい。一括確定処理は、ユーザの指示（例えば所定ボタンのクリック等）に基づいて、ラベル一致の各領域画像については当該一致ラベルの分類内容に一括で確定させる処理である。

[0060] なお、図４では、過去にアノテーション作業済の第２ユーザＵ００２によ

り、学習用データ生成指令（Ｓ３０２）がシステム１に送信される。しかしこれは一例でありこれに限られない。他の例として、他の第３ユーザが、医用画像Ｑに対するラベル入力操作（Ｓ３０１ａ、Ｓ３０１ｂ）を行うこと無く、確定操作（Ｓ３０３）のみを実施してもよい。この場合も、確定操作（Ｓ３０３）による各領域画像のラベル選択がアノテーション作業に相当している。

[0061]（２－５－３．第２の処理フロー例）

図５は、第１実施形態の制御動作を表す処理フローの第２の例（自動確定処理）である。図５ではコンピュータ（システム１であり、具体的にはラベル確定部２３ｂ）が、予め設定した規則に従って自動的にラベル確定処理（Ｓ１１５）を実行する。この点が図４の例との違いである。他のステップＳ１００（Ｓ１０１ａ～Ｓ１０２ｂ）、Ｓ１０３及びＳ１０６は図４のフローと同様である。

[0062] ステップＳ１０３の次に、ラベル確定部２３ｂは、ラベル確定処理（Ｓ１１５）を実行する。

[0063] ラベル一致時のラベル確定処理は、一例として次のようなものでもよい。第１及び第２ラベルLa_k、Lb_kが互いに同じ結果を示す領域画像Q_kがある場合を想定する。確定ラベルを第１及び第２ラベルLa_k、Lb_kと同じ内容とする処理を、コンピュータ（システム１）に実行させる。具体的には、もし仮に領域画像Q_2に対応付けられた第１及び第２ラベルLa_2、Lb_2（何れも不図示）が、両方とも「Ｔ」（つまり腫瘍）であったとすると、この場合は領域画像Q_2の確定ラベルが「Ｔ」とされる。

[0064] その一方で、ラベル不一致時のラベル確定処理は、一例として次のようなものでもよい。第１及び第２ラベルLa_k、Lb_kが互いに異なる結果を示す領域画像Q_kがある場合を想定する。この場合、ラベル確定処理として予め設定した処理を実行する。一例として、予め定めた規則に従って第１及び第２ラベルLa_k、Lb_kのうち片方の結果を確定ラベルとする処理を、コンピュータ（システム１）に実行させてもよい。具体例を挙げると、図５の例のよ

うに、領域画像Q_1に対応付けられた第１及び第２ラベルLa_1、Lb_1が「Ｔ」と「Ｎ」とで相違している場合を想定する。この場合に、ラベル確定部２３ｂは、一例として、アノテーション主体（第１実施形態ではユーザＵ００１、Ｕ００２）の情報に基づいて確定ラベルを決定してもよい。アノテーション主体の情報の一例は、「アノテーション実績値」である。一例として、ラベル確定部２３ｂが、第１ラベルLa_1を付与したユーザＵ００１のアノテーション実績値と第２ラベルLb_1を付与したユーザＵ００２のアノテーション実績値とを記憶部２０から読み出して、それらの実績値を比較してもよい。比較の結果、アノテーション実績値が相対的に高いほうのユーザにより付与されたラベルを確定ラベルとしてもよい。他の例として、ラベル不一致の領域画像Q_kがある場合に、当該領域画像Q_kを学習用データの対象外とする処理をコンピュータ（システム１）に実行させてもよい。更に他の例として、ラベル不一致の領域画像Q_kについて手動操作選択処理（例えば図４のＳ１０４，Ｓ１０５を流用可）を採用してもよく、これによりユーザからのラベル選択を受け付ける処理をコンピュータ（システム１）に実行させてもよい。

[0065] 以上説明したように図５の処理フローは生成処理をコンピュータに実行させる。生成処理は、コンピュータ（システム１）の自動処理（Ｓ１１５のラベル確定処理）に従って、コンピュータ（システム１）に、領域画像Q_kに対応付けられた複数のラベル（第１及び第２ラベルLa_k、Lb_k）により領域画像Q_kごとの確定ラベルを決定させて、確定ラベルと領域画像Q_kとを対応付けさせることで医用画像から学習用データを生成させる。第１及び第２アノテーション各々によるラベルに基づいて、各領域画像の確定ラベルを自動で決定することができる。複数のアノテーションを考慮に入れた学習用データを効率よく且つ少ない作業負担で生成することができ、これにより利便性が向上する。

[0066] （２－５－４．第３の処理フロー例）

図６の処理フローは多くの点で図４の処理フローに類似している。ただし

大きな相違点として、図６の例では指定枠方式のアノテーションを採用している。ステップＳ１０１ａで、作業画面２６はグリッド無しの医用画像Ｑを表示する。ステップＳ３０１ａ、Ｓ３０１ｂで、ユーザＵ００１、Ｕ００２は、作業画面２６を介して、指定枠により医用画像Ｑに指定枠領域画像Q_{1a}、Q_{1b}を設定し、ラベルLa_1、Lb_1を入力する。

[0067] 選択画面提示処理（Ｓ１０４）では、指定枠領域画像Q_{1a}、Q_{1b}及びラベルLa_1、Lb_1を医用画像Ｑに重ねて表示するように、選択画面２７を生成する。確定操作（Ｓ３０３）の一例として、ユーザＵ００２は指定枠領域画像Q_{1a}、Q_{1b}の指定枠の何れかにマウスポインタＰを合わせてクリック操作をする。この操作より、指定枠領域画像Q_{1a}及びラベルLa_1の組と、指定枠領域画像Q_{1b}及びラベルLb_1の組と、のうち何れかの組が採用される。医用画像Ｑに対応付けられたすべての指定枠についてユーザＵ００２の選択操作が同様に行われる。

[0068] その後、ステップＳ１０６で学習用データ構成のための処理が実行される。学習用データ構成部２３ｃは、医用画像Ｑを指定枠に従って切り出して、指定枠による領域画像と同一サイズの学習データ画像を生成可能である。

[0069] 変形例として、コンピュータ（システム１）に、ステップＳ１０４において、以下に述べる領域画像変換処理を実行させてもよい。ここでいう「領域画像変換」は、指定枠領域の画像とグリッド領域の画像との間での変換を行うことを指している。まず、図７Ａ及び図７Ｂに例示するように、指定枠に対してグリッドを重畳する。領域画像変換のために、一例として、図７Ａのグリッドが規定する各領域画像（幅Ｗｇ×高さＨｇ）に対して、指定枠が及ぶ範囲で、指定枠画像のラベルを対応づける。一例として、指定枠領域とグリッド領域との重なり面積が「所定重なり量」以上である場合には、グリッド領域画像に指定枠領域画像と同一のラベルを対応付けてもよい。「所定重なり量」は、グリッド領域の面積の例えば２０％、３０％、４０％、５０％、６０％、７０％、８０％、９０％又は１００％であり、ここで例示した数値の何れか２つの間の範囲内であってもよい。所定重なり量が１００％であ

る場合、あるグリッド領域と指定枠領域とが完全に重なっている。領域画像変換の後である図７Ｂでは、３行２列のＴラベル付領域群と、２行３列のＮラベル付領域群とが生成されている。図７Ｂにおける２つのグリッド領域画像は、２つの指定枠の論理積（ＡＮＤ）領域画像であり、ＴラベルとＮラベルの両方が対応付けられる。ユーザＵ００２は、図７Ｂの各グリッド領域画像について、図４のフローと同様の確定操作（Ｓ３０３）を行ってもよい。その後、ステップＳ１０６では、図７Ｂのグリッドに従って学習データ画像を生成可能である。なお、図９Ａ及び図９Ｂで３つの指定枠がある場合を例示しているので、こちらも参照されたい。

[0070]（２－５－５．第４の処理フロー例）

図８の処理フローは多くの点で図５及び図６の処理フローに類似している。図５の処理フローとの相違点は、図８の例では指定枠方式のアノテーションを採用することである。図６の処理フローとの相違点は、図８の例はラベル自動確定処理を行うことである。これらの相違点を除いては、図８の処理フローは、図５及び図６各々の処理フローと同様である。

[0071] 図８のフローでは、ラベル確定処理（Ｓ１２５）において、医用画像Ｑに設定された指定枠領域画像のラベルを確定する。ステップＳ１２５の処理は、グリッド領域と指定枠領域との差異を補うことで、ステップＳ１１５の処理を流用可能である。ステップＳ１２５で、領域画像変換を行うこと無く、指定枠の自動選択を行ってもよい。この場合、一例として、アノテーション実績値の高いほうの指定枠画像及びラベルを自動採用してもよい。その後、ステップＳ１０６では、指定枠画像に従って学習データ画像を生成可能である。

[0072] ステップＳ１２５の処理で、生成部２３は、領域画像変換を行ってもよい。領域画像変換後は、図７Ｂのように、グリッド領域画像にラベル対応付けがされた状態がつくりだされる。この後は、ステップＳ１１５の処理を同様に適用可能である。その後、ステップＳ１０６では、図７Ｂのグリッドに従って学習データ画像を生成可能である。

[0073] 領域画像変換が適用される場合、一例として、コンピュータ（システム１）に、ステップＳ１２５において、以下に述べるラベル確定処理を実行させてもよい。領域画像変換では、指定枠ごとに、指定枠領域画像とグリッド領域画像との重なりが異なることがある。例えば、図７Ｂには２つのグリッド領域画像にＴ及びＮラベルが対応付けられているが、しかし図７Ａを参照すると図右側のグリッド領域画像では、Ｎラベルの指定枠領域画像がグリッド領域画像の全体を覆っているのに対し、Ｔラベルの指定枠領域画像はグリッド領域画像の全体までには及んでいない。このような、グリッド領域画像と指定枠領域画像の重なりの違いを、ラベル確定処理に算入してもよい。一例として、複数の指定枠領域画像同士でグリッド領域との重なり具合に違いがある場合、最大重なり面積を有する指定枠領域画像のラベルを、そのグリッド領域画像のラベルに採用してもよい。他の例として、一例として、最大重なり面積の指定枠領域画像とそれ以外の指定領域画像とで、重なり面積の相違量が閾値以上である場合には、最大面積の大きいほうの指定枠領域画像のラベルをそのグリッド領域画像のラベルに採用してもよく、重なり面積の相違量が閾値未満である場合には、重なり面積に基づくラベル確定は採用せず他の基準で確定処理が実行されてもよい。これにより、重なり面積がそれほど大きくない指定枠領域画像がある場合に、重要度を下げることができる。ここで述べた処理は、図９Ａ及び図９Ｂのように指定枠画像が３個であったり、或いは指定枠画像が４個以上である場合にも、同様に採用可能である。

[0074] （２－６．第１実施形態の変形例等）

第１実施形態は、少なくとも下記に述べるように様々に変形可能である。

[0075] （２－６－１．指標値）

変形例として、第１実施形態の制御に「指標値」を取り入れてもよい。指標値は、複数の情報の各々に対応付け可能なパラメータである。指標値は、複数の情報各々の内容の確からしさを表す。指標値に基づいてより精度の高い学習用データを生成可能な利点がある。複数の情報（第１実施形態では複数のラベル）は、指標値そのものであってもよく、又は指標値を備えてもよ

い。以下、具体的に説明する。

[0076]（２－６－１－１．指標値の採用例）

図１０Ａはラベルが指標値を備えるラベル画像の例であり、図１０Ｂはラベルが指標値であるラベル画像の例である。図１０Ａ及び図１０Ｂのように、アノテーション対象としての領域画像に対応付けられる第１ラベルLa_k及び／又は第２ラベルLa_kが指標値を備えるか、又は指標値であってもよい。この例の指標値はラベル分類結果がどの程度確からしいかを表す。指標値は、例えばユーザが作業画面２６でのアノテーション作業時に手動入力可能であってもよい。指標値は、グリッド方式と指定枠方式の何れのアノテーションでも採用可能である。例えば図１０Ａの第１ラベルLa_kは「Ｔ９０」と表記されているが、これはラベル分類内容が「Ｔ」（腫瘍）であり、その指標値（確からしさ）は９０（％）である。例えば図１０Ａの第２ラベルLb_kは「Ｎ７０」と表記されているが、これはラベル分類内容が「Ｎ」（非腫瘍）であり、その指標値（確からしさ）は７０（％）である。指標値は、図４及び図５のステップＳ３０１ａ、Ｓ３０２ｂで、ユーザＵ００１、Ｕ００２が手動操作で入力可能としてもよい。指標値は、任意の刻み幅（例えば１％刻みや５％刻み等の一定幅）でユーザが任意設定可能としてもよいし、或いは、指標値が例えば複数の選択肢（例えば三段階や四段階等の任意個数の段階）として提示されてもよく、そこからユーザが所望の指標値を選択可能としてもよい。例えば三段階の指標値として、「高」（例えば７０％以上）と「中」（例えば３１％～６９％）と「低」（例えば３０％以下）等としてもよい。ユーザの指標値入力の利便性が向上する利点がある。

[0077]（２－６－１－２．図４又は図６の処理フローへの適用例）

図４又は図６の例（手動操作）で指標値を採用する場合は、選択画面２７において、図１０Ａ及び図１０Ｂのように指標値付きラベルが領域画像に対応付けられてもよい。ユーザは各ラベルの指標値を視認可能及び対比可能である。ユーザは、指標値の高低も考慮に入れながら、個々の領域画像Q_kの確定ラベルを決定できる。なお、複数のラベルのうち最も指標値の高いラベル

を目立たせて表示するなどしてもよい。目立たせ方は、例えば、任意のバッジをラベルに貼付してもよく、ハイライト表示（強調表示）でもよく、或いはラベル自体を相対的に大きく表示するなどでもよい。

[0078]（１－６－１－３．図５又は図８の処理フローへの適用例）

図５又は図８の例（自動処理）の場合は、ステップＳ１１５において、指標値に基づくラベル確定処理をコンピュータに実行させてもよい。一例として、複数のラベルの中で最大の指標値を持つラベルを、確定ラベルに採用してもよい。他の例として、不一致ラベル同士で指標値の差分を算出してもよく、その差分が予め定めた基準値以上であれば基準値が相対的に大きい方のラベルを確定ラベルに採用してもよい。その一方で、不一致ラベル間で指標値の差分が当該基準値未満である場合は、当該ラベルが対応付けられた領域画像を学習用データに含めないようにしてもよく、又は当該領域画像について手動操作選択処理（例えば図４のＳ１０４，Ｓ１０５を流用可）でどのラベルを採用するかをユーザに決定させてもよい。指標値を算入しつつ複数のアノテーション結果を統合することで、より精度の高い学習用データを自動生成することができる。

[0079]（２－６－２．「複数の情報」の個数）

（２－６－２－１．アノテーション数及びラベル数について）

第１実施形態では、図４～図９に例示するように、「複数の情報」の一例として第１及び第２情報（第１及び第２ラベル）という２個の情報を採用している。しかしこれは最小限の構成を規定しているにすぎず、３つ以上の任意の個数の情報（ラベル）を追加可能であることに注意されたい。例えば、第１～３情報により、領域画像とラベルとの対応付けを行ってもよい。例えば他の第３ユーザの第３アノテーションにより領域画像Q_kに第３ラベルが対応付けられてもよく、ステップＳ１０３でこの第３ラベルが第１及び第２ラベルとともに取得されてもよい。各領域画像Q_kで第１～第３ラベルにより確定ラベルを決定する処理（図４のステップＳ１０４、Ｓ１０５又は図５のＳ１１５）をコンピュータ（システム１）に実行させて、確定ラベルと領域画

像とを対応付けた学習用データを生成する処理（Ｓ１０６）をコンピュータ（システム１）に実行させてもよい。第１実施形態では、「第１～第３ラベルにより確定ラベルを決定する」という記載は、「第１～第３ラベルにより提示される選択肢（「Ｔ」又は「Ｎ」）から確定ラベルを決定する」という意味を含んでいる。

[0080] 上記の例では便宜的に第１～第３アノテーションを例示して、第１～第３ラベルを「複数の情報」に採用可能としている。しかし第１実施形態で採用可能な「複数の情報」の個数に上限はない。任意数のユーザによる任意個数のアノテーションが図４～図６及び図８の処理フローに採用されてもよい。システム１において図４又は図５の処理フローに採用可能なアノテーション数を「Ｚ」とした場合、Ｚの値は、例えば２～１００個であり、例えば３～２０個であり、例えば４～１０個でもよい。Ｚの値は、具体的には例えば２、３、４、５、６、７、８、９、１０、１５、２０、３０、４０、５０、７５、１００個であり、ここで例示した数値の何れか２つの範囲内であってもよい。

[0081] 実施形態のシステム１は、限定されたメンバーで使用されてもよく、例えばある企業、研究室、又は研究グループ内のメンバーのみで使用されてもよい。しかしながら、これに限られず、システム１は、より多くの様々な研究機関、大学又は企業のメンバーからのユーザアカウント登録を受付可能でもよく、世界各国のあらゆる地域のユーザがアクセス可能でもよく、そのために通信ネットワーク２を通じて広く一般に公開されてもよい。こうすることでより多くの様々なユーザからアノテーションデータを収集可能となる利点がある。ただしその場合、必ずしも全ユーザが自己のユーザ情報を開示することに好意的とは限らない。この点に対処するために、実施形態のシステム１は、アノテーション開示に関する制限機能、及び／又は、後述する「アノテーション主体情報の表示／非表示機能」を備えている。

[0082] システム１は、領域画像Q_kに対応付けられた全ラベルを採用して確定処理（図４及び図６のＳ１０４、Ｓ１０５、図５のＳ１１５及び図８のＳ１２５

）を実行してもよいが、これに限定されない。例えば学習用データ生成指令（Ｓ３０２）において、学習用データの生成処理に採用すべきラベル群をユーザが指定可能としてもよく、つまり誰が何時行ったアノテーションを採用するかを指定可能としてもよい。例えば、システム１は、学習用データ生成指令に応答して、学習用データ生成対象となる領域画像ＩＤに対応付けられる各ラベルのユーザＩＤをユーザ端末３に表示してもよい。ユーザ端末３でユーザＩＤの選択指定を受け付けて、指定したユーザＩＤをシステム１に送信してもよい。

[0083] （２－６－２－２．図４又は図６の処理フローへの適用例）

一例として、図４の処理フローの選択画面２７において、第１～第３ラベルを１つのグリッド領域画像Q_kに表示させる処理を、コンピュータ（システム１、具体的には表示制御部２１）に実行させてもよい。ユーザは、第１～第３ラベルのうち１つのラベルを選択可能である。４つ以上の情報を採用する場合も、同様である。

[0084] 他の例として、３つ以上のラベルのうち、ラベル分類内容の一方（例えば「Ｔ」）と他方（例えば「Ｎ」）とにそれぞれ属するラベル数の違いを、ユーザに対して視認可能な形態で、選択画面２７を提示してもよい。この場合、選択画面２７では、各グリッド領域画像Q_kに１つの「Ｔ」ラベルと１つの「Ｎ」ラベルとが重ねて表示され、各ラベル画像にラベル数の違いが表現されてもよい。例えば選択画面２７において、「Ｔ」ラベルの数と「Ｎ」ラベルの数とをそれぞれ数値等で提示してもよいし（当該数値をグリッド領域画像に重ねて表示してもよい）、或いは数が多い方のラベルを目立たせて表示してもよい。目立たせ方は、前述の「１－６－１－２．図４の処理フローへの適用例」で述べた例と同様でもよい。

[0085] 他の例として、図６の処理フローの選択画面２７において、３つの指定枠領域画像が、それぞれ第１～第３ラベルを有してもよい（図９Ａ参照）。他の例として、図６の処理フローで領域画像変換を行う場合には、領域画像変換後のグリッド領域画像（図９Ｂ参照）に上記の図４の各適用例を採用可能

である。

[0086]（２－６－２－３．図５又は図８の処理フローへの適用例）

一例として、図５のラベル確定処理（Ｓ１１５）において第１～３情報を採用する場合に、多数決処理をコンピュータ（システム１、具体的にはラベル確定部２３ｂ）に実行させてもよい。多数決処理では、例えば、各領域画像で第１～第３ラベルの「Ｔ」ラベル個数と「Ｎ」ラベル個数とをカウントし、各領域画像で「Ｔ」と「Ｎ」のうち数が多い方を確定ラベルに採用してもよい。多数決処理は、図８の処理フローで領域画像変換を行った場合（図９Ｂ参照）にも同様に適用可能である。

[0087] ４つ以上の情報（つまり第１～４ラベル等）を使用する場合も多数決処理を採用可能である。しかし、もし仮に「複数の情報」の数を偶数個とする場合には、不一致ラベルが半数ずつで均衡する場合がある。均衡する場合には、一例としてアノテーション実績値を用いて確定ラベルを決定してもよい。具体的には、例えば、アノテーション実績値が最大であるラベルの分類内容を確定ラベルとして決定してもよい。均衡する場合の他の例として、「Ｔ」ラベルを付与したユーザのアノテーション実績値の合計値と、「Ｎ」ラベルを付与したユーザのアノテーション実績値の合計値と、を比較して、合計値が大きい方のラベルを確定ラベルとしてもよい。均衡する場合の更に他の例として、そのラベル不一致の領域画像は学習用データに含めないようにしてもよい。均衡する場合の更に他の例として、ラベル不一致の領域画像Q_kについて手動操作選択処理（例えば図４のＳ１０４，Ｓ１０５を流用可）でユーザからのラベル選択を受け付ける処理を、コンピュータ（システム１）に実行させてもよい。

[0088] 一例として、図５のラベル確定処理（Ｓ１１５）において３つ以上の情報（ラベル）を採用する場合に、指標値に基づく演算処理をコンピュータ（システム１、具体的にはラベル確定部２３ｂ）に実行させてもよい。指標値に基づく演算は、例えば、四則演算（例えば加算等）、或いは要約統計量計算（例えば平均値算出や最大値算出）などでもよい。例えば第１～第３ラベル

のうち、「Ｔ」ラベルの指標値を加算して第１加算値を求め、「Ｎ」ラベルの指標値を加算して第２加算値を求め、第１及び第２加算値のうち大きい方のラベルを確定ラベルに採用してもよい。或いは、第１及び第２加算値の代わりに第１及び第２平均値を求めて、第１及び第２平均値のうち大きい方のラベルを確定ラベルに採用してもよい。あるいは、例えば第１～第３ラベルのうち最大の指標値を持つラベルを、確定ラベルに採用してもよい。４つ以上の情報（ラベル）でも同様である。指標値に基づく演算処理は、図８の処理フローで領域画像変換を行った場合（図９Ｂ参照）にも同様に適用可能である。

[0089]（２－６－２－３．ラベル欠落発生時の対処法）

複数の領域画像Q_kに、ラベルの欠落（つまりアノテーション漏れ）がある領域画像が含まれている場合がある。具体的には、例えば、他の多く（例えば過半数）の領域画像ではＮ個のラベルが対応付け済であるにもかかわらず、一部の領域画像ではＮ個未満のラベルしか対応付けられていないか又はラベル数がゼロであることがある。この場合、一例として、ラベル欠落が発生した領域画像において少なくとも１つのラベルが対応付けられていれば、その少なくとも１つのラベルにより確定ラベルを決定する処理を行ってもよい。他の例として、ラベルの欠落がある場合及び／又はラベル数がゼロである領域画像については、その領域画像を学習用データ生成の対象外としてもよいし、或いはその領域画像を対象にしてアノテーション受付処理（例えば図４のＳ１０１ａ、Ｓ３０１ａ、Ｓ１０２ａを流用可）によりユーザからのラベル補充入力を促してもよい。

[0090]（２－７．アノテーション主体情報の開示／非開示）

アノテーション主体情報が、選択画面２７に表示されてもよい。アノテーション主体情報は、各ラベルのラベル入力を行ったアノテーション主体が誰なのかを特定可能な情報である。アノテーション主体情報はユーザＩＤなどでもよく、或いは、ユーザの氏名名称等、所属、ニックネーム又はアバターアイコン等でもよい。しかしながら、その一方で、アノテーション主体情報

は、非表示でもよい。一例として、選択画面２７に呼出したアノテーションのアノテーション主体情報が表示されなくともよい。ユーザが、アノテーションの「開示可」又は「開示不可」を設定可能でもよい。開示可の場合、条件付き（つまりアノテーション主体情報が非表示であることを条件）で、選択画面２７に表示とされてもよい。アノテーション主体情報を非表示とすることで、アノテーション主体を秘匿可能である。ユーザが自己のユーザ情報を開示することに好意的ではない場合であっても、匿名性を担保した上で、アノテーションデータを活用可能な利点がある。

[0091] 第１及び第２アノテーションを行う主体（アノテーション主体）は、人に限らずコンピュータも含んでもよい。ルールベースの分類処理により又は学習済判別モデルの機械推論分類によりコンピュータが画像検出処理を行うと、各領域画像に分類結果が対応付けされるので、これをアノテーションとみなしてもよい。複数のアノテーションは、主体が異なっていても同一でもよい。アカウントＤＢにコンピュータ用ユーザＩＤが設定されてもよく、ラベルＤＢへの登録時に各コンピュータ固有のアノテーション主体ＩＤを付与しておけばよい。

[0092] ＜３．第２実施形態＞

第２実施形態のうち、第１実施形態と同一又は同様の構成については説明を省略するか又は簡略化する。以下、相違点を中心に説明する。

[0093] （３－１．システム１および生成部２３の機能ブロック構成）

図１１Ａに示す第２実施形態のシステム１は、記憶部２０と、表示制御部２１と、入力部２２と、生成部２３と、判別部２５とを備える。判別部２５を備える点で、第１実施形態のシステム１（図３Ａ参照）と異なる。図１１Ｂに示す第２実施形態の生成部２３は、データ取得部２３ａと、ラベル確定部２２３ｂと、学習用データ構成部２３ｃとを備える。第２実施形態は、ユーザが選択したラベルを確定ラベルとして確定するラベル確定部２３ｂに代えて自動ラベル付け処理を実行するラベル確定部２２３ｂを備える点で、第１実施形態のシステム１（図３Ａ参照）と異なる。

[0094] （３－２．判別部２５の説明）

判別部２５は、任意の機械学習アルゴリズムで構築された学習済判別モデルである。判別部２５を構築する機械学習アルゴリズムに限定はなく、各種の公知技術を採用可能であり、例えばニューラルネットワーク、決定木、サポートベクターマシン、又はロジスティック回帰分析等でもよく、例えば任意の深層学習アルゴリズム（ＤＮＮ）でもよく、例えば任意のＣＮＮモデルでもよい。判別部２５は、既に他の学習用データを用いて学習が完了している。判別部２５は、入力画像データに対する画像認識処理を実行可能であり、入力画像データを分類する機械推論を実行可能に構築される。

[0095] （３－３．処理フロー）

図１２及び図１３それぞれは、第２実施形態の処理フローの具体例であり、学習用データを生成するプログラム及び生成方法を提供する。第２実施形態では、第１実施形態とは異なる生成処理をコンピュータ（システム１）に実行させる（図１２のＳ１５２～Ｓ１５７、又は図１３のＳ１５２，Ｓ１５３，Ｓ１６４，Ｓ１５６及びＳ１５７を参照）。第２実施形態の生成処理は、自動処理で、コンピュータ（システム１）に、領域画像に対応付けられた複数の情報（第１及び第２情報）により当該領域画像とラベルとの対応付けを行わせて、学習用データを生成させる。後述するように、第２実施形態でも、各領域画像において、第１及び第２情報が領域画像に対応付けられる。領域画像に対応付けられた複数の情報により当該領域画像にラベル対応付けを自動処理で行うことができるので、高い利便性で学習用データを生成可能である。

[0096] 第２実施形態では、上記の「複数の情報」が第１及び第２情報を含み、第１情報として「画像付属情報Ｇ」が採用され（図１２参照）、第２情報として「学習済判別モデルによる推論分類結果」が採用される（図１２にラベル$Ｌｃ_{k}$、$Ｌｄ_{k}$で例示）。画像付属情報Ｇは、医用画像（医用画像の全体）に付属させられる情報である。典型例としては医用画像の詳細を説明する各種の医用情報でもよく、医用画像の取得元の患者の情報等でもよく、具体的な

バリエーションは後述する。

[0097] （３－３－１．共通の処理ステップＳ１５１）

図１２及び図１３の共通点として、複数の領域画像を含む入力データを受け付ける処理を、コンピュータ（システム１）に実行させる（Ｓ１５１）。第２実施形態では、記憶部２０が、第１医用画像Ｑと第２医用画像Ｒとを記憶している。複数の領域画像$ＱＭ_k$、$ＲＭ_k$は、第１医用画像Ｑ及び第２医用画像Ｒの各々から生成可能である。一例として、複数の領域画像$ＱＭ_k$、$ＲＭ_k$は、グリッドに従って第１及び第２医用画像Ｑ，Ｒの各々を均一サイズに切り出したものでもよい。他の例として、複数の領域画像$ＱＭ_k$、$ＲＭ_k$は、第１及び第２医用画像Ｑ，Ｒの各々に設定された１つ以上の指定枠に従って、第１及び第２医用画像Ｑ，Ｒの各々を切り出したものでもよい。第１及び第２医用画像Ｑ、Ｒは第１情報（画像付属情報Ｇ）と対応付けられる。ここでは画像付属情報Ｇの一例として、遺伝子変異タイプが採用される。画像ＤＢに登録される各医用画像の遺伝子変異タイプは、遺伝子変異検査やがんパネル検査等と呼ばれる各種検査によって記憶部２０（画像ＤＢ）への登録段階で予め既知である。記憶部２０は、遺伝子変異タイプ（Ａタイプ該当／非該当、Ｂタイプ該当／非該当・・・等）と各医用画像とを対応付けて記憶する。

[0098] 第２実施形態では、一例として第１情報（画像付属情報Ｇ）が「遺伝子変異タイプ」であり、予め定めた基準（ラベル付基準）は「遺伝子変異タイプＡに該当するか否か」である。第１医用画像Ｑは、第１情報が予め定めた基準を満たす画像である。第２医用画像Ｒは、第１情報が当該基準を満たさない画像である。図１２には、画像付属情報Ｇを模式的にテーブルで図示しているが、実際には画像ＤＢ内で画像ＩＤに対応付けて画像付属情報Ｇが記憶される。

[0099] ここで「遺伝子変異タイプＡ」と便宜的に記載しているが、タイプＡが具体的にどのような遺伝子変異であるかは予め設定済であるものとする。遺伝子変異の種類として、がんの部位に応じて、例えば、ＥＧＦＲ遺伝子変異、

ＨＥＲ２遺伝子増幅（ＥＲＢＢ２コピー数異常）、ＲＡＳ遺伝子変異（ＫＲＡＳおよびＮＲＡＳ）、ＢＲＡＦ遺伝子変異（Ｖ６００ＥまたはＶ６００Ｋ）、ＡＬＫ融合遺伝子、ＲＯＳ１融合遺伝子、ＮＴＲＫ融合遺伝子、高頻度マイクロサテライト不安定性（ＭＳＩ－Ｈｉｇｈ）、ＢＲＣＡ遺伝子変異、又は遺伝子変異量（ＴＭＢ）等がある。各種の遺伝子変異において質的に及び／又は量的に具体的にどのようなものを「遺伝子変異タイプＡ」とするのかを予め規定しておけばよく、これによりタイプＡの該当／非該当を区別可能となる。記憶部２０に各タイプ（Ａ、Ｂ、Ｃ・・・）を登録しておき、記憶部２０への医用画像の登録時に各タイプの該当有無を選択的に設定可能としてもよい。

[0100] （３－３－２．第１の処理フロー例）

図１２に示す第１の処理フロー例では、例えばユーザＵ００１からの学習用データ生成指令があると（Ｓ３５１）、システム１は、学習用データの生成処理を開始する。図１２の処理フローは、先ず、複数の領域画像を含む入力データを取得する処理を、コンピュータ（システム１）に実行させる（Ｓ１５１）。具体的には、記憶部２０の各ＤＢが参照されて、今回の学習用データの生成に必要な入力データが取得される（Ｓ１５２）。図１２の例では、ユーザＵ００１が第１及び第２医用画像Ｑ，Ｒを指定し、第１情報として「遺伝子変異タイプＡ該当／非該当」が指定され、第２情報として「腫瘍／非腫瘍の判別」が指定され、これらから生成された領域画像$ＱＭ_k$、$ＲＭ_k$が処理されるものとする。

[0101] なお、任意の医用画像や任意の領域画像を指定可能としてもよく、及び／又は、第１及び第２情報を任意に指定可能としてもよい。これにより更に利便性が向上する。例えばステップＳ１５１で、ユーザＵ００１に幾つかの設定項目を要求してもよい。例えばユーザＵ００１に対して、医用画像又は領域画像の指定要求と第１情報の指定要求と第２情報の指定要求とのうち少なくとも１つの要求をユーザが任意に行えるようにしてもよく、ユーザ端末３からシステム１に当該指定要求の内容が伝達されてもよい。

[0102] ステップＳ１５３で、入力データの領域画像ＱＭ$_k$、ＲＭ$_k$を分類する推論を学習済判別モデル（判別部２５）に行わせる処理を、コンピュータ（システム１）に実行させる。各領域画像に対する推論分類結果は「Ｔ」又は「Ｎ」であり、この推論分類結果が第２実施形態の「第２情報」である。ここで、図１２では説明を分かりやすくする目的で、便宜上、「Ｔ」（腫瘍）のラベルＬｃ$_k$と「Ｎ」（非腫瘍）ラベルＬｄ$_k$とを領域画像ＱＭ$_k$、ＲＭ$_k$に対応付けているけれども、実際の処理ではこのようなラベル付けは必須ではなく（つまりラベルＤＢ登録は不要であり）、このような図示は本発明の範囲を限定しない。ステップＳ１５３で単に各領域画像が「Ｔ」又は「Ｎ」に分類されて、各領域画像の分類結果が後段の処理（Ｓ１５４）に提供されるだけでもよい。

[0103] ステップＳ１５４で、ラベル確定部２２３ｂが、領域画像ＱＭ$_k$、ＲＭ$_k$の選別処理を実行する。図１２のフロー例では、「Ｔ」ラベルが付与された領域画像のみを抽出する。ここでは一例として、「Ｔ」ラベルが付与された領域画像ＱＭ$_k$及び領域画像ＲＭ$_k$のみを次のステップＳ１５６に渡す。なお、図１２では、説明の便宜上、「Ｎ」ラベルが付与された領域画像を「ＱＭ$_j$、ＲＭ$_m$」と例示的に記載し、これらに対応付けられたラベルをＬｃ$_j$、Ｌｄ$_m$と例示的に記載している（ｊ、ｍは整数）。「Ｎ」ラベルが付与されたこれらの領域画像はステップＳ１５４で破棄される。変形例として、「Ｎ」ラベルが付与された領域画像のみを抽出することも可能である。

[0104] ステップＳ１５６で、ラベル確定部２２３ｂが、自動ラベル付け処理を実行する。すなわちステップＳ１５６の処理では、コンピュータ（システム１、具体的にはラベル確定部２２３ｂ）に、第１医用画像Ｑから生成され、分類結果として第１分類グループ（「Ｔ」）に分類された１つ以上の領域画像（Ｓ１５４の選別後の領域画像ＱＭ$_k$）を対象にしてラベルＬｅ$_k$を対応付けさせ、且つ、第２医用画像Ｒから生成され、分類結果として第１分類グループ（「Ｔ」）に分類された１つ以上の領域画像（Ｓ１５４の選別後の領域画像ＲＭ$_k$）を対象にしてラベルＬｆ$_k$を対応付けさせる。各領域画像ＱＭ$_k$、Ｒ

M_kが第１及び第２医用画像Ｑ，Ｒの何れから生成されたか等の情報は、画像ＤＢに既に記憶されている（Ｓ１５１参照）。

[0105] ラベルLe_kは、領域画像QM_kの第１情報が前述した予め定めた基準（遺伝子変異タイプＡ該当）を満たすことを表す。ラベルLf_kは、領域画像RM_kの第１情報が当該基準を満たさないことを表す。生成後の学習用データにおいて、ラベルLe_k、Lf_kが正解データとなる。このようにして、入力データの領域画像を学習済判別モデルで分類し、その分類結果（第２情報）として第１グループに分類された領域画像に対して、第１情報（画像付属情報Ｇ）により決定したラベルを自動的に対応付けすることができる。なお遺伝子変異タイプ該当性が不明の医用画像又は領域画像は、ステップＳ１５２の段階で除外してもよい。

[0106] 第２実施形態において、第１情報（画像付属情報Ｇ）は様々に設定可能である。第２実施形態に採用される医用画像が病理画像である場合に、第１情報は、遺伝子変異タイプ、ＴＭＢ－Ｈ（高腫瘍遺伝子変異量）タイプ、生化学的解析の結果、又は患者の予後予測であってもよい。ステップＳ１５６で第１情報により決定されるラベルの分類種別及び分類内容は、様々に規定可能である。ラベルの分類内容は、例えば遺伝子変異タイプが所定タイプ（例えばタイプＡ）に該当するか否か、ＴＭＢ－Ｈ（高腫瘍遺伝子変異量）タイプに該当するか否か、生化学的解析の結果が所定基準に該当するか否か、又は患者の予後予測が所定基準値以上か否か等であってもよい。患者の予後予測は、例えば生命予後が所定期間以上か否か等でもよい。第１情報は、患者の年齢、性別、病歴、免疫組織化学的解析の結果または生活習慣であってもよい。記憶部２０にこれらの各種の分類種別及び分類内容を予め登録しておき、記憶部２０への医用画像の登録時に選択的に設定可能としてもよい。

[0107] ステップＳ１５７では、ステップＳ１５６により対応付けられた領域画像及びラベルにより学習用データを生成する処理を、コンピュータ（システム１）に実行させる。生成した学習用データは学習用データＤＢに登録され、ユーザＵ００１はこれを任意にダウンロード可能となる。

[0108] 以上説明したように、図１２の処理フローは生成処理をコンピュータ（システム１）に実行させる。この生成処理では、コンピュータ（システム１）に、第１情報（遺伝子変異タイプＡの該当有無：Ｓ１５１）により決定したラベル（図１２のＳ１５６のラベルLe_k、Lf_k参照）を、第２情報（腫瘍／非腫瘍の推論分類結果：Ｓ１５３）による選別を行った領域画像（Ｓ１５４の領域画像QM_k、RM_k参照）に対応付けさせて、学習用データを生成させる。領域画像に対して第１情報によるラベル決定と第２情報による選別とを施した学習用データを、自動生成することができる。

[0109] （３－３－３．第２の処理フロー例）

図１３は、第２実施形態の制御動作を表す処理フローの第２の例（指標値を用いる処理）である。具体的には、ステップＳ１５４とステップＳ１６４とが置換されている点が図１２及び図１３の各フローの違いである。他のステップＳ１５１～Ｓ１５３、Ｓ１５６及びＳ１５７は図１２のフローと同様である。

[0110] 図１３に示す第２の処理フロー例では、ステップＳ１５３で判別部２５が確率値を出力する点に留意されたい。公知の幾つかの機械学習アルゴリズムによる画像認識モデルは、画像認識による分類結果の確からしさ（つまり確率又は一致度）を出力可能である。図１３の例では判別部２５が深層学習アルゴリズムで構築されたニューラルネットワークモデルであるものとする。ニューラルネットワークの出力層は、推論結果として、入力画像データが各分類結果にどの程度の確率で属するかを表す確率値を出力する。図１３の例ではこの確率値が指標値に使用される。ステップＳ１５３の推論により、学習済判別モデル（判別部２５）が、領域画像QM_k、RM_kの分類結果の確からしさを表す確率値を出力する。

[0111] ステップＳ１６４では、コンピュータ（システム１であり、具体的にはラベル確定部２２３ｂ）が、領域画像QM_k、RM_kの選別処理を実行する。具体的には、図１３のフロー例では、コンピュータ（システム１）に、推論分類処理（Ｓ１５３）の分類結果が「Ｔ」（腫瘍）であり且つ指標値（確率値

）が所定値以上と推論された領域画像のみを抽出させる。図１３の例では所定値が０．８０である。

[0112] この所定値は、例えば、０．７０～１．００であり、例えば０．８０～１．００でもよい。この所定値は、具体的には、例えば０．５０、０．５５、０．６０、０．６５、０．７０、０．７５、０．８０、０．８５、０．９０、又は０．９５でもよく、ここに例示した何れか２つの間の範囲内であってもよい。ここに例示した数値は、１．００を１００％に換算したうえで百分率によって表記されてもよい。

[0113] その後、図１２の処理フローと同様に、自動ラベル付け処理（Ｓ１５６）及び学習用データ構成処理（Ｓ１５７）が実行される。推論結果として得られる確率値を用いて自動選別を行うことで、より精度の高い学習用データを自動生成することができる。

[0114] （３－４．第２実施形態の変形例等）

第２実施形態ではシステム１に判別部２５が内蔵されているが、これは一例にすぎず、このような形態に限定されない。判別部２５はシステム１に内蔵されていなくともよい。学習済み判別モデルを有する任意の外部システムと連携して、この外部システムを判別部２５として利用してステップＳ１５３の処理を実行させてもよい。

[0115] ＜４．その他＞

（４－１．画像の実質的同一性の基準、及びＤＢ管理）

実施形態において、複数の医用画像の間の同一性について、次のような「実質的同一」の基準を採用可能である。ある２つの医用画像が実質的同一である場合の例として、少なくとも次の第１又は第２状況を挙げることができる。第１状況では、２つの医用画像が完全同一であり、同一の画像データを有している。第２状況では、医用画像間の差異（例えば画像データ及び／又は画像ＩＤの差異）があっても、その差異がアノテーション用途上は同一視可能である。

[0116] 例えば、同一の医用画像対象物を撮影することで、ある医用画像を取得し

たとする。その医用画像データを複製したり、複製後に画像フォーマットを変更したり（例えば圧縮形式で保存等）、複製後に画像処理（リサイズ、明るさ、コントラスト、彩度、色調等の変更）を適用したりして、新たな医用画像を生成したとする。元の医用画像と新たな医用画像は、画像ＩＤや画像データ等に違いがあっても、実質的に同一である。医用画像の撮影対象物が同一であり、アノテーション用途上はその差異を無視可能だからである。これは領域画像でも同様であり、実質的同一の医用画像から生成した領域画像についても、実質的同一の基準を採用可能である。

[0117] 或いは、例えば、いわゆる「バーチャルスライド」と称される仕組みを利用することがある。バーチャルスライドでは、同じ撮影対象（病理標本、つまり組織）を様々な拡大率で撮影し、これにより例えば数枚～数十枚程度の異なる拡大率の病理画像を取得し、その多数の病理画像を画像ＤＢに登録する。デジタル化した病理標本データを利用可能なので病理医の観察精度を向上しやすいなどの利点がある。バーチャルスライドにおける一群の病理画像は、同じ病理標本を異なる拡大率で表しているにすぎないので、実質的に同一である。この考え方は、同じ撮影対象を異なる拡大率で撮影するようなあらゆる医用画像でも同様である。これら実質的同一の医用画像から生成した領域画像についても、実質的同一の基準を採用可能である。

[0118] また、例えば組織を連続的に切り出して連続切片を生成した場合、ある病理切片と、その直上又は直下にある病理切片とで、構造がほとんど同じであることがある。その場合、それらの各病理切片の病理画像は、実質的に同一である。病理画像以外のあらゆる医用画像でも、同様のケースが想定される。例えば３次元構造をスライスしたときに、隣接する各層の平面画像が実質的に同一であることが想定される。これらの場合も、複数の医用画像は実質的に同一である。これら実質的同一の医用画像から生成した領域画像についても、実質的同一の基準を採用可能である。

[0119] 実施形態は、少なくとも上記各例で述べた「画像の実質的同一性の基準」を採用可能である。例えば、実施形態において、記憶部２０の画像ＤＢ内で

、実質的に同一として取り扱いたい医用画像群をひとまとめに扱うように、それらの画像ＩＤをユーザ手動操作又はコンピュータ自動処理により互いに紐づけ可能であってもよい。

[0120] 一例として、第１ユーザ（例えばＵ００１）が第３医用画像にアノテーションを行ったとする。第２ユーザ（例えばＵ００２）が、第４医用画像にアノテーションを行ったとする。第４医用画像は、第３医用画像を複製したものであるか、又は第３医用画像に複製し且つ画像処理を施したものであるとする。第３及び第４医用画像は画像ＤＢ内で紐付けられていて、両画像は同一と扱われてもよい。

[0121] 実施形態では、実質的に同一の医用画像の各領域画像に複数のラベル（例えば第１及び第２ラベル）が対応付けられる場合に、記憶部２０の各ＤＢにおけるＩＤ管理には幾つかの例が考えられる。幾つかの例では、実質的に同一の複数の医用画像にグリッド方式のアノテーションが行われるか、或いは指定枠方式のアノテーションでグリッド設定がなされてもよい。この場合に、各医用画像のグリッドは同一サイズ且つ同一座標で設定されてもよい。

[0122] ＩＤ管理の第１の例として、記憶部２０内のＤＢは、１つの領域画像ＩＤに、複数のアノテーション各々によるラベルＩＤを対応付けてもよい。この場合、複数の医用画像で領域画像ＩＤの群が共通採用されてもよい。つまり単純に、あるラベル（第１ラベル）と対応付けられる領域画像ＩＤが、他のラベル（第２ラベル）と対応付けられる領域画像ＩＤにも採用されてもよい。

[0123] しかしこれに限られず、ＩＤ管理の第２の例は次のようなものでもよい。第３医用画像（マスタ医用画像）をコピーして１つ以上の第４医用画像（コピー医用画像）を生成する場合がありうる。更に、第４医用画像に画像処理が施される場合もある。少なくともこれらの場合、第３及び第４医用画像は、実質的には同一の画像であるけれども、異なる画像データを有し、異なる画像ＩＤが設定され得る。そこで、ＩＤ管理の第２の例として、第３及び第４医用画像の画像ＩＤを画像ＤＢ内で紐づけされてもよい。マスタ医用画像

ＩＤ及び１つ以上のコピー医用画像ＩＤの群からなる画像群において、１つの領域画像ＩＤにラベルＩＤが対応付けられた場合には、この関連付けを辿って残りの画像ＩＤの医用画像にも同じラベルＩＤを自動的に対応付けてもよい。

[0124] 少なくとも、第１の例のように直接的に、又は第２の例のように間接的に、医用画像ＩＤと領域画像ＩＤとラベルＩＤとの対応付けを行ってもよい。これにより、互いに実質的同一な医用画像に対して複数のアノテーションを実施してもよく、一つの領域画像に複数のアノテーションにより複数のラベルを対応付け可能としてもよい。

[0125] なお、例えば第２実施形態の第１及び第２医用画像Ｑ，Ｒは、互いに異なる画像であり、実質的同一には該当しないことに注意されたい。

[0126] （４－２．サーバ構成等のバリエーション）

図１の例では、システム１は１つの情報処理装置として図示されているが、実施形態はこれに限定されない。例えば、システム１は、１又は複数の物理的なサーバ等から構成されていてもよいし、ハイパーバイザー（hypervisor）上で動作する仮想的なサーバを用いて構成されていてもよいし、クラウドサーバを用いて構成されていてもよい。また、選択画面２７の生成及び表示に関するプログラムの処理は、システム１（サーバ１）とユーザ端末３とで任意に分担可能である。図４～図６、図８、図１２及び図１３の各処理フローにおける処理の分担のさせ方は一例であることに留意されたい。

[0127] 本願の明細書及び特許請求の範囲において、「コンピュータ（a computer）」は、特に異なる意味を明示的に限定しない限りは、「少なくとも１つのコンピュータ(at least one computer)」を意味すると理解されたい。「コンピュータ」は、ユーザ端末及び／又はサーバのうちの任意の１つ以上を含んでもよい。「コンピュータ」があるプログラムを実行する場合、その実行処理の全てが単一のコンピュータで完結されてもよいし、或いは、当該実行処理の一部が第１コンピュータ（例えばサーバ）で実行され且つ当該処理の残りが他の１つ以上のコンピュータ（例えばユーザ端末又は他のサーバ）で実

行されてもよい。当該他の１つ以上のコンピュータは、第２～第ｎコンピュータのうち任意の１つ以上でもよい（ｎは任意の３以上の整数）。選択画面の生成処理は、ユーザ端末で実行されてもよい。或いは、選択画面の生成処理はサーバで実行されてもよく、この場合は完成した選択画面データをユーザ端末に送信してもよい。選択画面の生成処理は、一部の処理をユーザ端末で実行し、残りの処理をサーバで実行してもよい。ただし、上記の解釈と異なる意味で「コンピュータ（a computer）」を解釈すべきことが明記された場合には、当該記載に限定される。

[0128] 本発明に係る種々の実施形態を説明したが、これらは、例として提示したものであり、発明の範囲を限定することは意図していない。当該新規な実施形態は、その他の様々な形態で実施されることが可能であり、発明の要旨を逸脱しない範囲で、種々の省略、置き換え、変更を行うことができる。当該実施形態やその変形は、発明の範囲や要旨に含まれるとともに、特許請求の範囲に記載された発明とその均等の範囲に含まれるものである。

符号の説明

[0129] １：システム(サーバ)、２：通信ネットワーク、３：ユーザ端末、４：診断装置、１１：プロセッサ、１２：記憶装置、１３：通信ＩＦ（インタフェース）、２０：記憶部、２１：表示制御部、２２：入力部、２３：生成部、２３ａ：データ取得部、２３ｂ、２２３ｂ：ラベル確定部、２３ｃ：学習用データ構成部、２５、２２５：判別部、２６：作業画面、２７：選択画面、３１：制御部、３２：記憶部、３３：通信部、３４：表示デバイス、３５：スピーカ、３６：マイク、３７：カメラ、３８：操作デバイス、La_1、La_k：第１ラベル（第１実施形態の第１情報）、Lb_1、Lb_k：第２ラベル（第１実施形態の第２情報）、Ｑ：第１医用画像、Ｒ：第２医用画像、Ｇ：画像付属情報（第２実施形態の第１情報）、Lc_k、Ld_k：推論分類結果（第２実施形態の第２情報）、Le_k、Lf_k：ラベル、Ｐ：マウスポインタ、Q_1、Q_2、Q_3、Q_4、Q_k、QM_k、RM_k、QM_j、RM_m：領域画像、Ｕ００１、Ｕ００２：ユーザ

請求の範囲

[請求項1] 医用画像から学習用データを生成するためのプログラムであって、

前記プログラムは、生成処理を、コンピュータに実行させ、

前記医用画像における少なくとも１つの領域の画像である領域画像に複数の情報が対応付けられ、

前記生成処理は、操作入力部を介して受け付けた手動操作又はコンピュータの自動処理によって、コンピュータに、前記複数の情報により前記領域画像へのラベル対応付けを行わせて、前記医用画像から前記学習用データを生成させる、プログラム。

[請求項2] 請求項１に記載のプログラムであって、

前記複数の情報は第１及び第２情報を含み、前記第１情報が第１ラベルであり、前記第２情報が第２ラベルであり、

第１アノテーション及び第２アノテーションにより１つの前記領域画像に対応付けられた前記第１ラベル及び前記第２ラベルを取得する処理を、コンピュータに実行させる、

プログラム。

[請求項3] 請求項１又は請求項２に記載のプログラムであって、

前記複数の情報の各々がラベルであり、

前記生成処理は、操作入力部を介して受け付けた手動操作又はコンピュータの自動処理によって、コンピュータに、１つの前記領域画像に対応付けられた複数のラベルから前記１つの前記領域画像の確定ラベルを決定させて、前記医用画像から生成した学習データ画像と前記確定ラベルとを対応付けさせることで、前記学習用データを生成させる、プログラム。

[請求項4] 請求項１に記載のプログラムであって、

前記複数の情報は、第１及び第２情報を含み、

前記領域画像は、第１領域画像および第２領域画像を含み、

前記生成処理は、コンピュータに、前記第１情報により前記第１領

域画像について決定したラベルを前記第２情報による選別を行った前記第２領域画像に対応付けさせて、前記学習用データを生成させる、プログラム。

[請求項5] 請求項４に記載のプログラムであって、

複数の前記領域画像を含む入力データを取得する処理を、コンピュータに実行させ、

前記複数の前記領域画像に含まれる複数の前記第１領域画像は、第１及び第２医用画像の各々から生成され、

前記第１及び第２医用画像は前記第１情報と対応付けられ、

前記第１医用画像は、前記第１情報が予め定めた基準を満たす画像であり、前記第２医用画像は、前記第１情報が前記基準を満たさない画像であり、

前記入力データの前記複数の前記領域画像に含まれる複数の前記第２領域画像を分類する機械推論を学習済判別モデルに行わせる判別処理を、コンピュータに実行させ、

前記機械推論による前記領域画像の分類結果が、前記第２情報であり、

前記生成処理は、コンピュータに、前記第１医用画像から生成された前記第１領域画像に含まれ、前記分類結果として第１分類グループに分類された前記第２領域画像に前記第１情報が前記基準を満たすことを表すラベルを対応付けさせ、且つ、前記第２医用画像から生成された前記第１領域画像に含まれ、前記分類結果として前記第１分類グループに分類された前記第２領域画像に前記第１情報が前記基準を満たさないことを表すラベルを対応付けさせて、前記学習用データを生成させる、プログラム。

[請求項6] 請求項４又は請求項５に記載のプログラムであって、

前記医用画像は病理画像を含む、プログラム。

[請求項7] 請求項６に記載のプログラムであって、

前記第１情報は、遺伝子変異タイプ、ＴＭＢ－Ｈ（高腫瘍遺伝子変異量）タイプ、生化学的解析の結果、又は患者の年齢、性別、病歴、免疫組織化学的解析の結果、生活習慣もしくは予後予測の少なくとも１つを含む、プログラム。

[請求項8] 請求項１、２、４、５のいずれか１つに記載のプログラムであって、

前記複数の情報の少なくとも１つは、指標値を備えるか又は指標値であり、

前記指標値は情報の内容の確からしさを表す、プログラム。

[請求項9] 請求項２を引用する、請求項８に記載のプログラムであって、

前記領域画像と対応付けられるラベルが、指標値を備えるか又は指標値であり、

前記指標値はラベルの分類結果がどの程度確からしいかを表す、プログラム。

[請求項10] 請求項５を引用する請求項８に記載のプログラムであって、

前記機械推論により前記学習済判別モデルが前記領域画像の分類結果の確からしさを表す確率値を出力し、前記確率値を前記指標値に使用する、プログラム。

[請求項11] 医用画像から学習用データを生成する生成方法であって、

生成ステップを、コンピュータに実行させ、

前記医用画像における少なくとも１つの領域の画像である領域画像に複数の情報が対応付けられ、

前記生成ステップは、操作入力部を介して受け付けた手動操作又はコンピュータの自動処理によって、コンピュータに、前記複数の情報により前記領域画像へのラベル対応付けを行わせて、前記医用画像から前記学習用データを生成させる、方法。

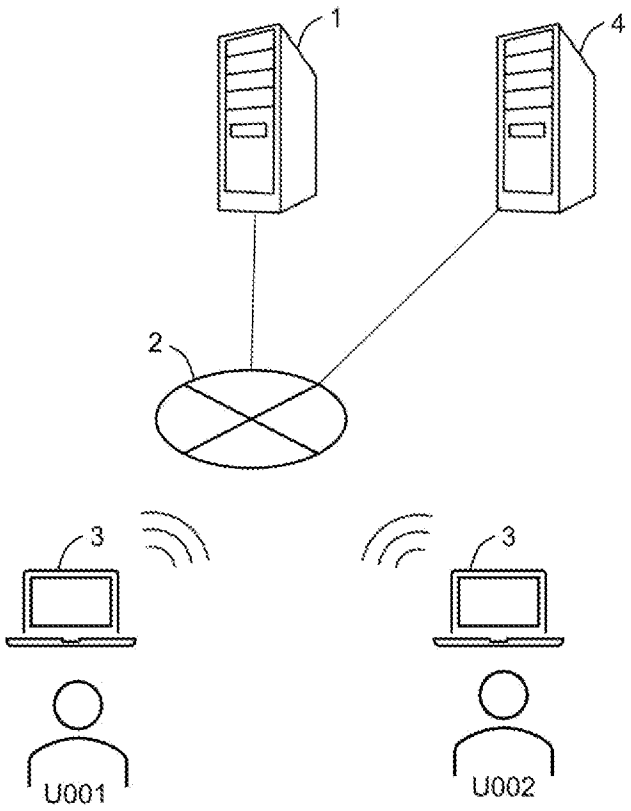
[請求項12] 医用画像から学習用データを生成するシステムであって、

生成部を備え、

前記医用画像における少なくとも１つの領域の画像である領域画像に複数の情報が対応付けられ、

前記生成部は、操作入力部を介して受け付けた手動操作又はコンピュータの自動処理によって、前記複数の情報により前記領域画像へのラベル対応付けを行うことで、前記医用画像から前記学習用データを生成する、システム。

[図1]
図1



[図2]

図2A

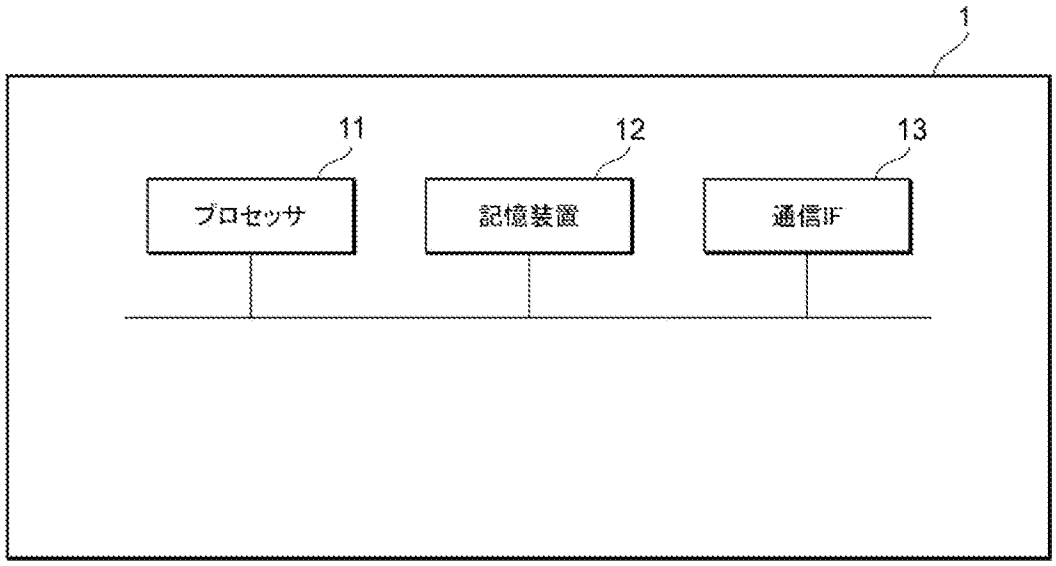
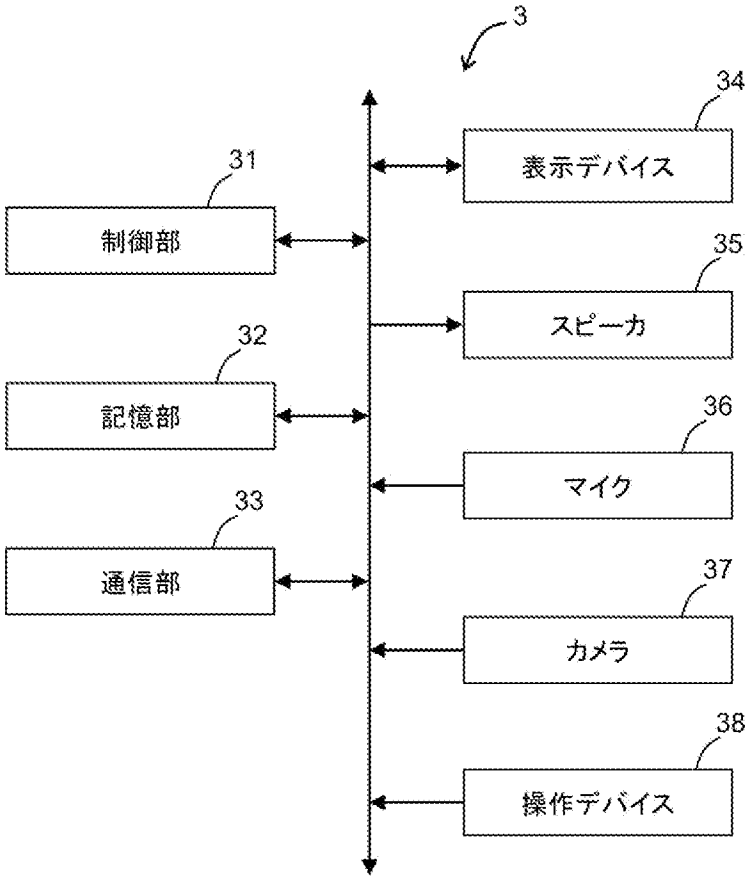

1
11
12
13
プロセッサ
記憶装置
通信IF

図2B


3
34
表示デバイス
31
制御部
35
スピーカ
32
記憶部
36
マイク
33
通信部
37
カメラ
38
操作デバイス

[図3]

図3A

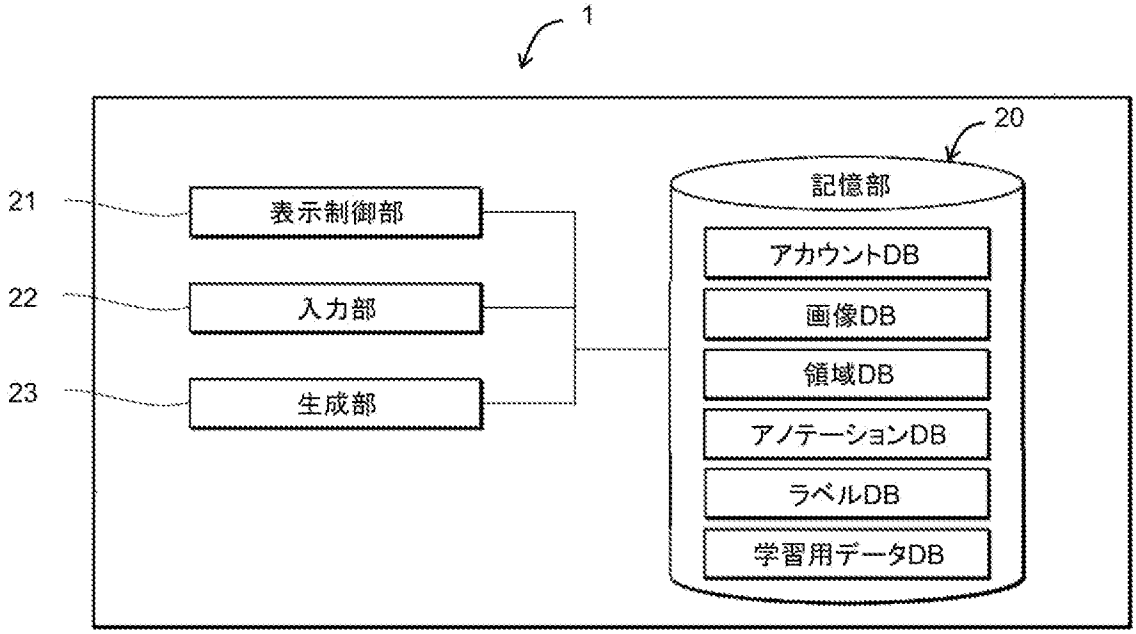
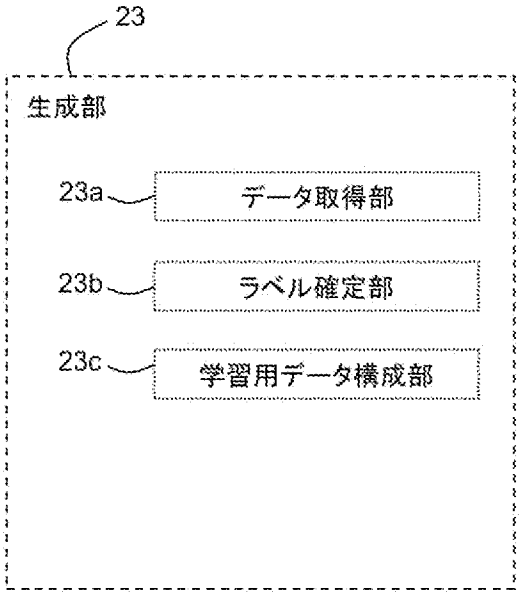
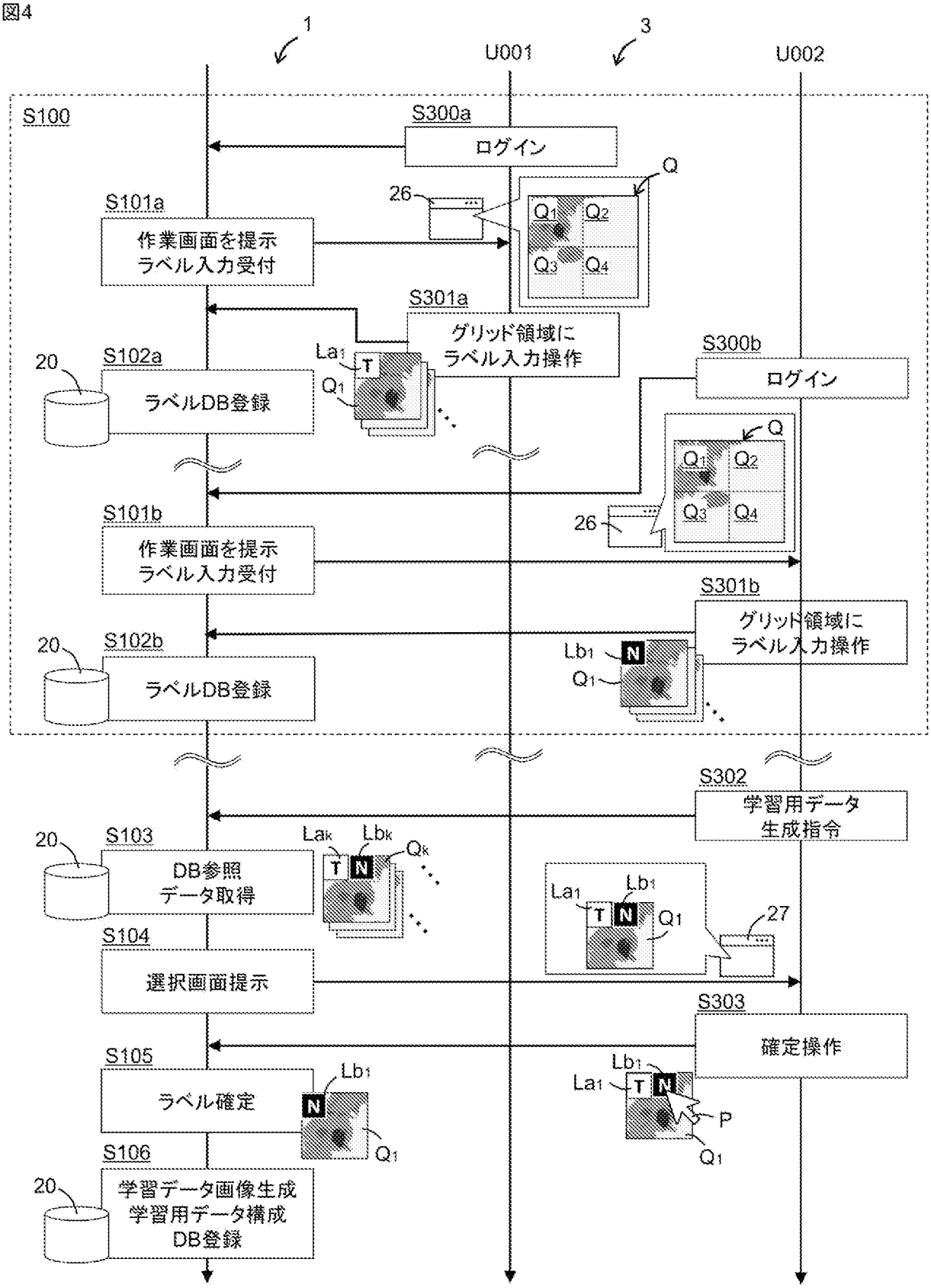

1
20
記憶部
21
表示制御部
22
入力部
23
生成部
アカウントDB
画像DB
領域DB
アノテーションDB
ラベルDB
学習用データDB

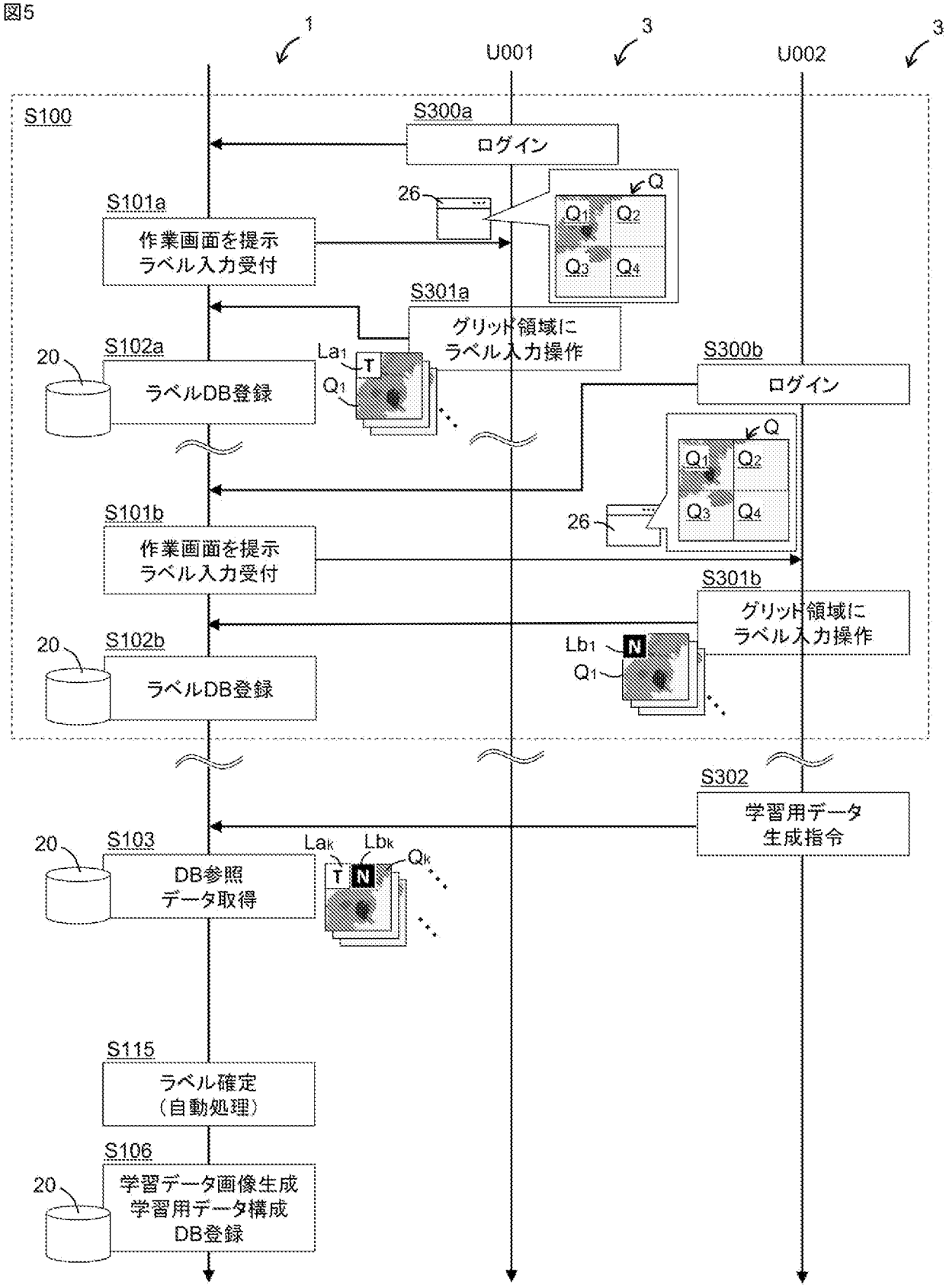
図3B


23
生成部
23a
データ取得部
23b
ラベル確定部
23c
学習用データ構成部

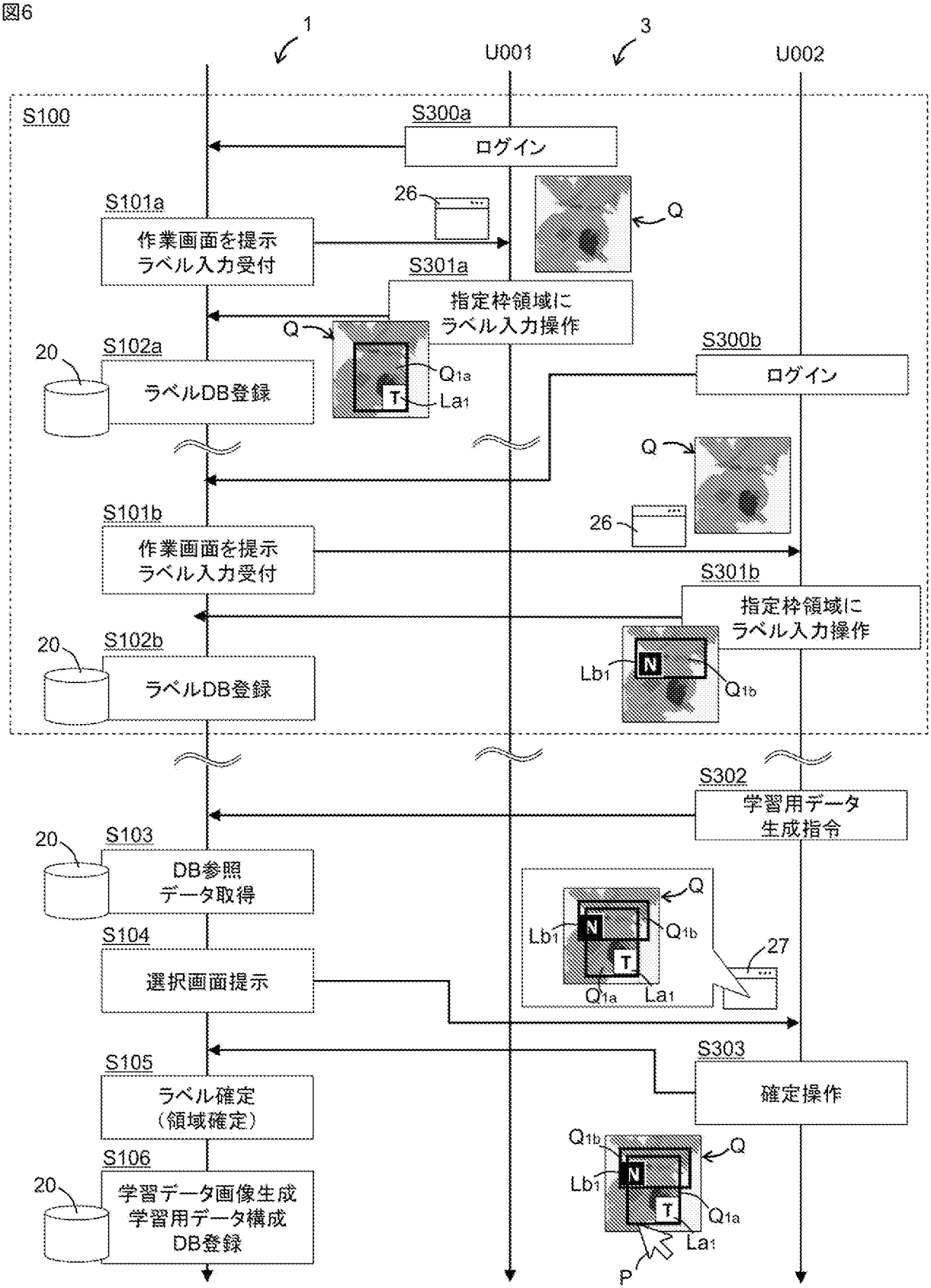
[図4]
図4


1
U001
3
U002
S100
S300a
ログイン
Q
26
S101a
作業画面を提示
ラベル入力受付
Q1 Q2 Q3 Q4
S301a
グリッド領域に
ラベル入力操作
S300b
ログイン
20
S102a
ラベルDB登録
La1
T
Q1
Q
Q1 Q2 Q3 Q4
26
S101b
作業画面を提示
ラベル入力受付
S301b
グリッド領域に
ラベル入力操作
20
S102b
ラベルDB登録
Lb1
N
Q1
S302
学習用データ
生成指令
20
S103
DB参照
データ取得
Lak
Lbk
Qk
T
N
La1
Lb1
T
N
Q1
27
S104
選択画面提示
S303
確定操作
S105
ラベル確定
Lb1
N
Q1
Lb1
La1
T
N
P
Q1
S106
20
学習データ画像生成
学習用データ構成
DB登録

[図5]
図5


1
U001
3
U002
3
S100
S300a
ログイン
26
Q
Q1
Q2
Q3
Q4
S101a
作業画面を提示
ラベル入力受付
S301a
グリッド領域に
ラベル入力操作
20
S102a
ラベルDB登録
La1
T
Q1
S300b
ログイン
S101b
作業画面を提示
ラベル入力受付
S301b
グリッド領域に
ラベル入力操作
Lb1
N
S102b
ラベルDB登録
S302
学習用データ
生成指令
S103
DB参照
データ取得
Lak
Lbk
Qk
S115
ラベル確定
（自動処理）
S106
学習データ画像生成
学習用データ構成
DB登録

[図6]
図6


1
3
U001
U002
S100
S300a
ログイン
S101a
作業画面を提示
ラベル入力受付
26
Q
S301a
指定枠領域に
ラベル入力操作
Q
Q1a
T
La1
20
S102a
ラベルDB登録
S300b
ログイン
Q
26
S101b
作業画面を提示
ラベル入力受付
S301b
指定枠領域に
ラベル入力操作
20
S102b
ラベルDB登録
Lb1
N
Q1b
S302
学習用データ
生成指令
20
S103
DB参照
データ取得
Q
Lb1
N
Q1b
T
Q1a
La1
27
S104
選択画面提示
S303
確定操作
S105
ラベル確定
（領域確定）
Q1b
Q
Lb1
N
T
Q1a
La1
P
S106
20
学習データ画像生成
学習用データ構成
DB登録

[図7]

図7A

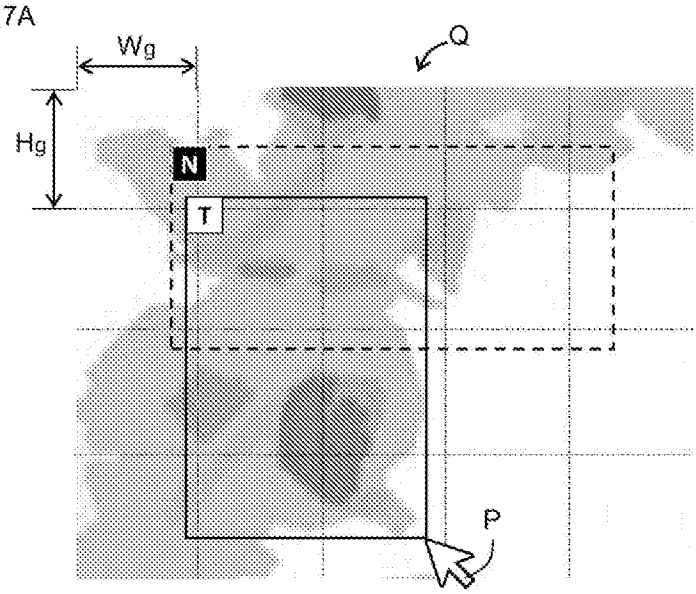
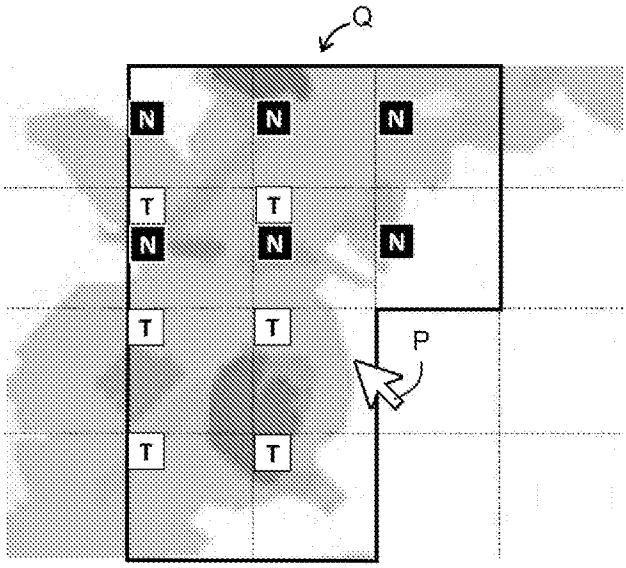
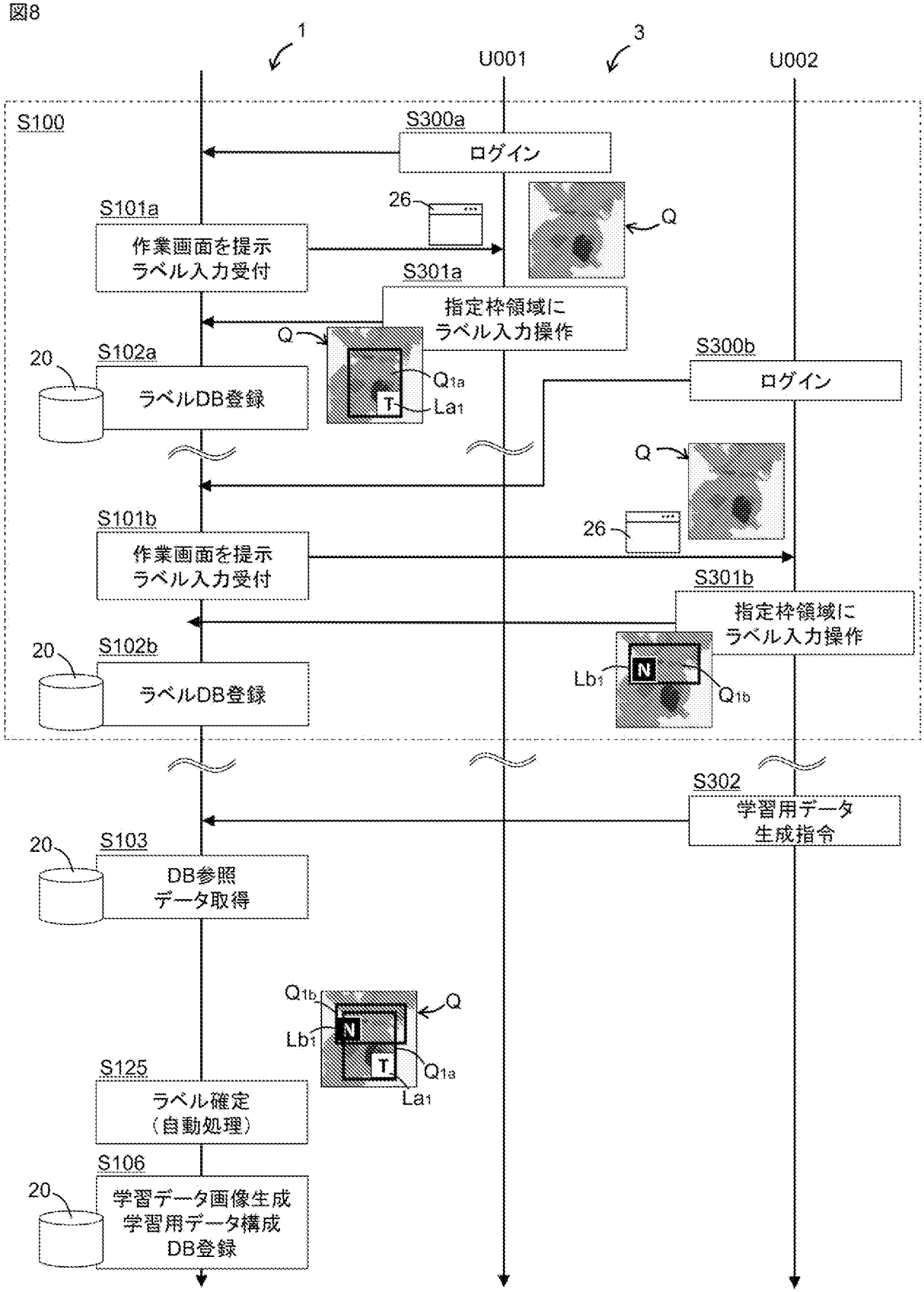


図7B



[図8]
図8


1
U001
3
U002
S100
S300a
ログイン
26
Q
S101a
作業画面を提示
ラベル入力受付
S301a
指定枠領域に
ラベル入力操作
Q
Q1a
T
La1
20
S102a
ラベルDB登録
S300b
ログイン
Q
26
S101b
作業画面を提示
ラベル入力受付
S301b
指定枠領域に
ラベル入力操作
20
S102b
ラベルDB登録
Lb1
N
Q1b
S302
学習用データ
生成指令
20
S103
DB参照
データ取得
Q1b
Q
Lb1
N
T
Q1a
La1
S125
ラベル確定
（自動処理）
S106
20
学習データ画像生成
学習用データ構成
DB登録

[図9]

図9A

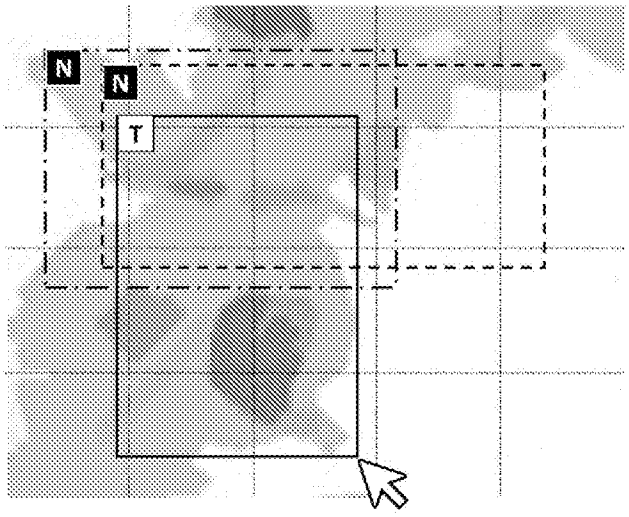
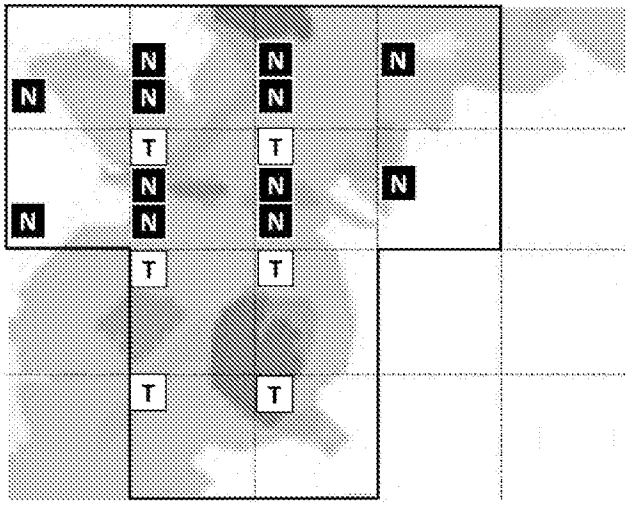


図9B



[図10]

図10A

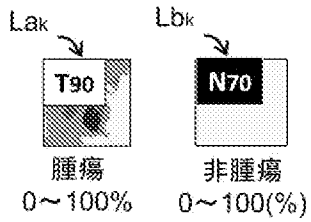
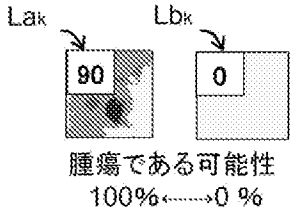


図10B



[図11]

図11A

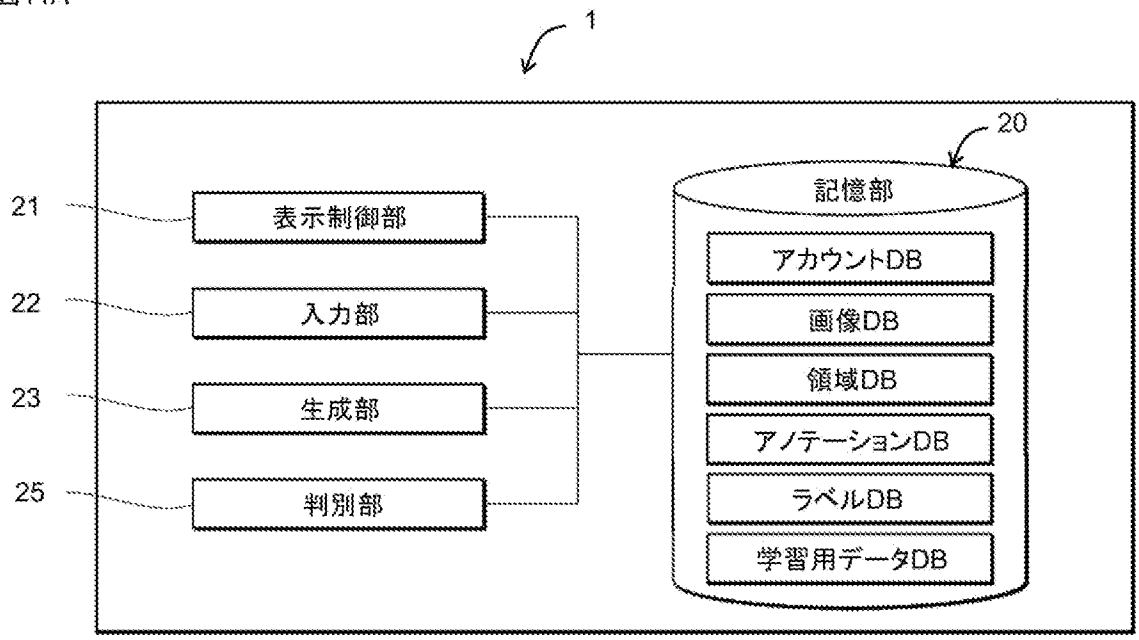
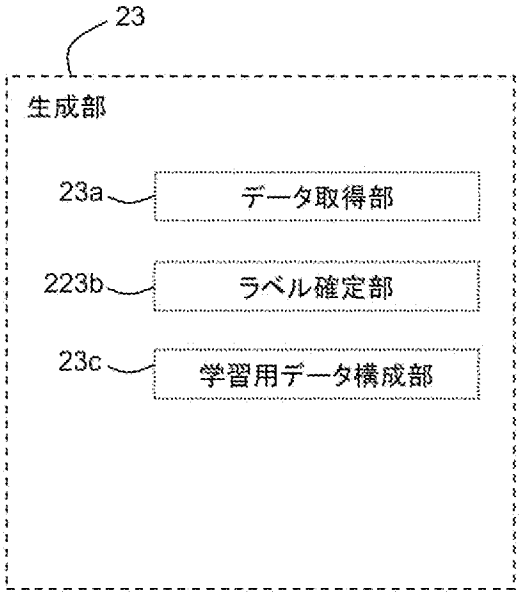
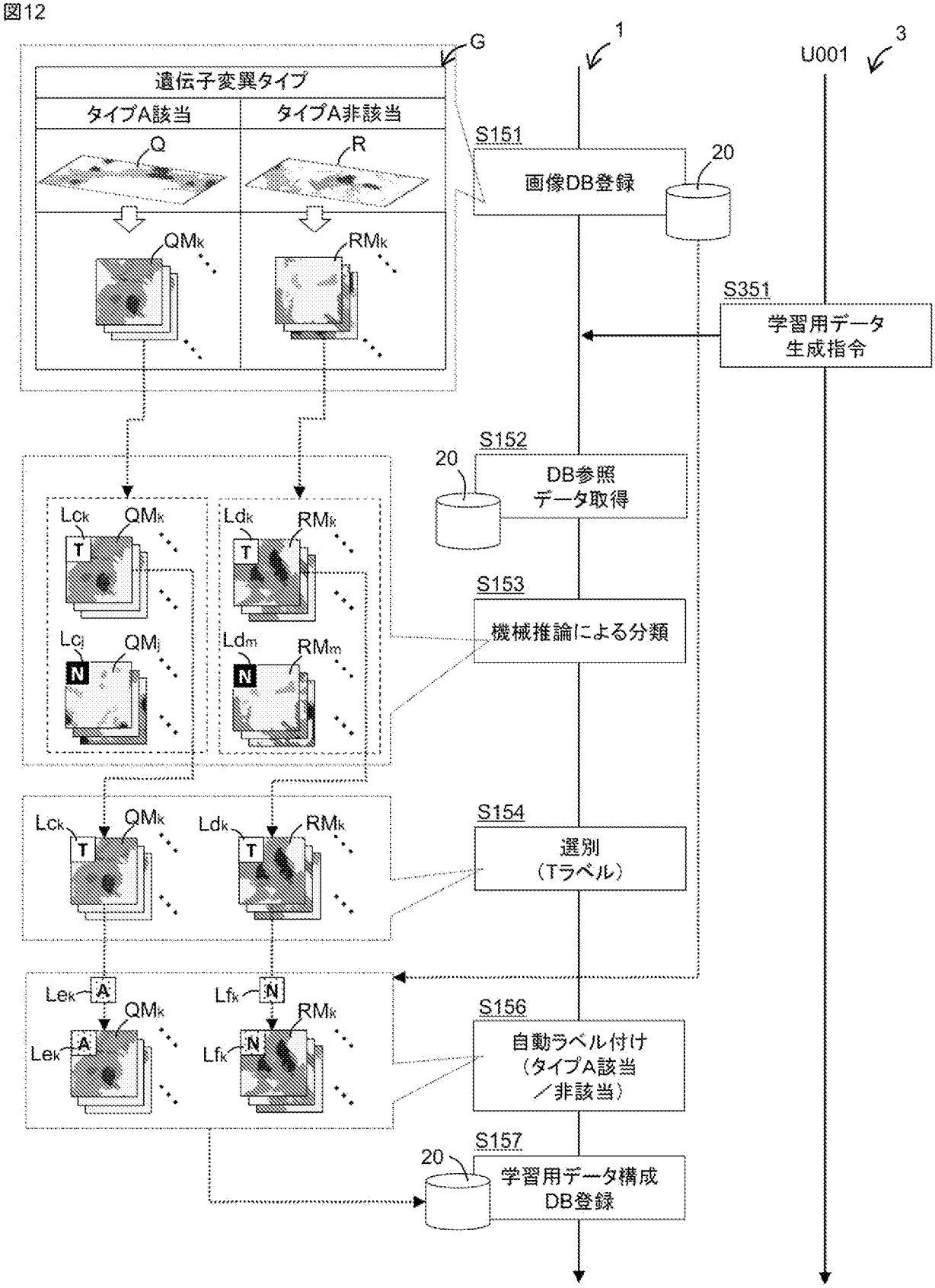


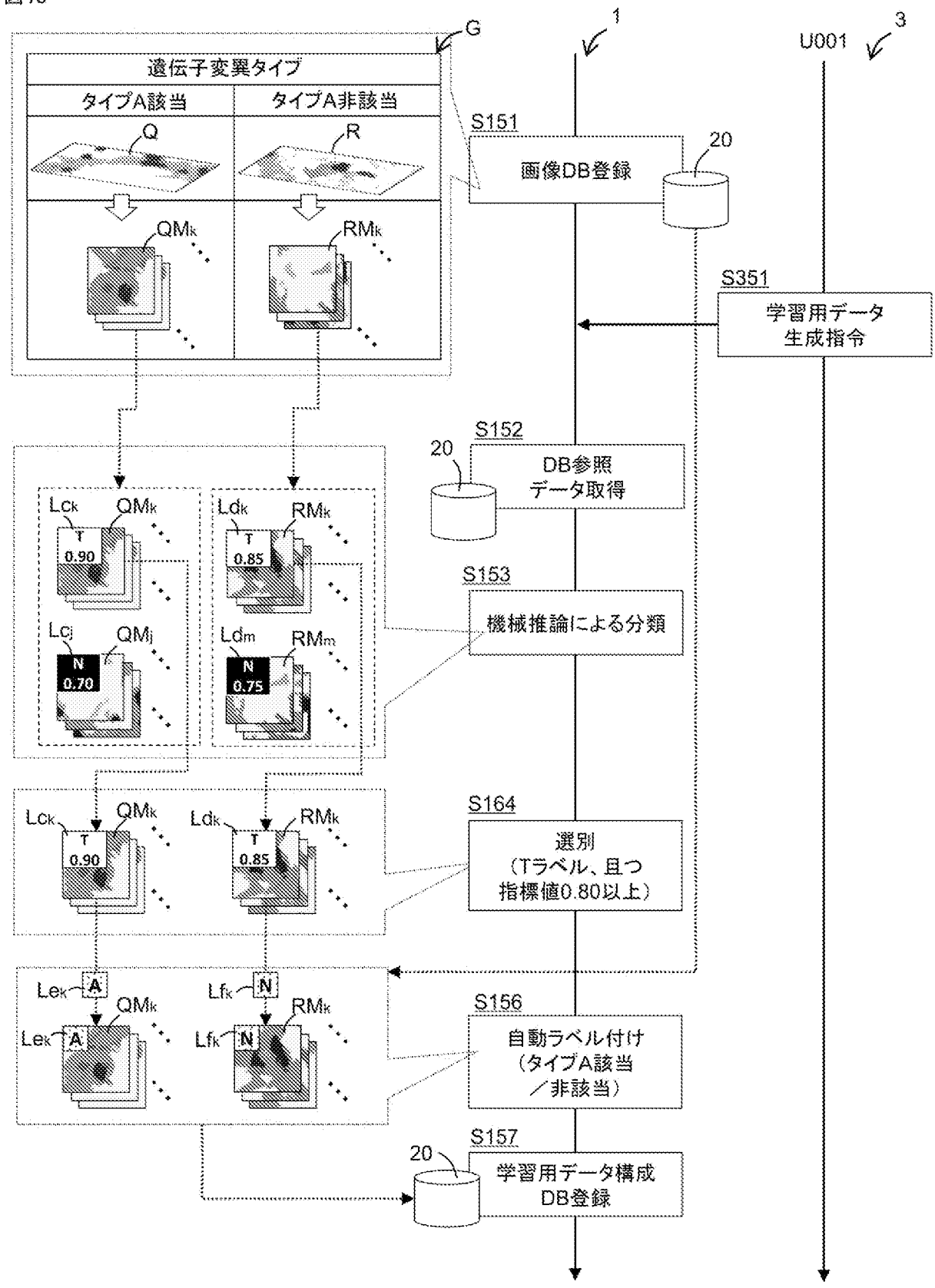
図11B



[図12]
図12


G
1
U001
3
遺伝子変異タイプ
タイプA該当
タイプA非該当
Q
R
QMk
RMk
S151
画像DB登録
20
S351
学習用データ
生成指令
S152
20
DB参照
データ取得
Lck
QMk
Ldk
RMk
T
Lcj
QMj
Ldm
RMm
N
S153
機械推論による分類
S154
選別
（Tラベル）
Lek
A
Lfk
N
S156
自動ラベル付け
（タイプA該当
／非該当）
S157
20
学習用データ構成
DB登録

[図13]
図13


G
1
U001
3
遺伝子変異タイプ
タイプA該当
タイプA非該当
Q
R
QMk
RMk
S151
画像DB登録
20
S351
学習用データ
生成指令
S152
DB参照
データ取得
Lck
Ldk
T
0.90
0.85
Lcj
Ldm
N
0.70
0.75
QMj
RMm
S153
機械推論による分類
S164
選別
（Tラベル、且つ
指標値0.80以上）
Lek
Lfk
A
N
S156
自動ラベル付け
（タイプA該当
／非該当）
S157
学習用データ構成
DB登録

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2024/019506

A. CLASSIFICATION OF SUBJECT MATTER

G06V 10/774(2022.01)i; ***A61B 5/00***(2006.01)i; ***G06T 7/00***(2017.01)i; ***G16H 30/20***(2018.01)i
FI: G06V10/774; G16H30/20; A61B5/00 D; G06T7/00 350B

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)

G06V10/774; A61B5/00; G06T7/00; G16H30/20

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Published examined utility model applications of Japan 1922-1996
Published unexamined utility model applications of Japan 1971-2024
Registered utility model specifications of Japan 1996-2024
Published registered utility model applications of Japan 1994-2024

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2021/206053 A1 (NIIGATA UNIVERSITY, et al.) 14 October 2021 (2021-10-14) paragraphs [0010]-[0012], [0019]-[0021], [0031]-[0035], fig. 5	1-3, 11-12
A		4-10
X	WO 2022/209501 A1 (FUJIFILM CORPORATION) 06 October 2022 (2022-10-06) paragraphs [0010]-[0015], [0032]-[0038], [0075]-[0083], fig. 3	1-3, 8-9, 11-12
A		4-7, 10
A	WO 2021/095698 A1 (SONY GROUP CORPORATION) 20 May 2021 (2021-05-20) paragraphs [0012]-[0035]	1-12

☐ Further documents are listed in the continuation of Box C. ☑ 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
"D" document cited by the applicant in the international application
"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	Date of mailing of the international search report
24 July 2024	**06 August 2024**

Name and mailing address of the ISA/JP	Authorized officer
Japan Patent Office (ISA/JP) **3-4-3 Kasumigaseki, Chiyoda-ku, Tokyo 100-8915** **Japan**	Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/JP2024/019506

Patent document cited in search report	Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
WO 2021/206053 A1	14 October 2021	US 2023/0148243 A1 paragraphs [0015]-[0017], [0024]-[0026], [0036]-[0040], fig. 5 EP 4134909 A1 paragraphs [0010]-[0012], [0019]-[0021], [0031]-[0035], fig. 5 CN 115380307 A paragraphs [0020]-[0023], [0032]-[0034], [0045]-[0049], fig. 5 KR 10-2022-0164755 A paragraphs [0010]-[0013], [0022]-[0024], [0036]-[0039], fig. 5	
WO 2022/209501 A1	06 October 2022	US 2024/0013524 A1 paragraphs [0039]-[0044], [0061]-[0067], [0104]-[0112], fig. 3	
WO 2021/095698 A1	20 May 2021	US 2022/0399118 A1 paragraphs [0056]-[0081] CN 114586061 A paragraphs [0061]-[0086]	

国際調査報告

国際出願番号
PCT/JP2024/019506

A. 発明の属する分野の分類（国際特許分類（ＩＰＣ））

G06V 10/774(2022.01)i; A61B 5/00(2006.01)i; G06T 7/00(2017.01)i; G16H 30/20(2018.01)i
FI: G06V10/774; G16H30/20; A61B5/00 D; G06T7/00 350B

B. 調査を行った分野

調査を行った最小限資料（国際特許分類（ＩＰＣ））

G06V10/774; A61B5/00; G06T7/00; G16H30/20

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報	１９２２－１９９６年
日本国公開実用新案公報	１９７１－２０２４年
日本国実用新案登録公報	１９９６－２０２４年
日本国登録実用新案公報	１９９４－２０２４年

国際調査で使用した電子データベース（データベースの名称、調査に使用した用語）

C. 関連すると認められる文献

引用文献のカテゴリー*	引用文献名　及び一部の箇所が関連するときは、その関連する箇所の表示	関連する請求項の番号
X	WO 2021/206053 A1 (国立大学法人新潟大学　外１名) 14.10.2021 (2021-10-14) 段落0010-0012, 0019-0021, 0031-0035, 図5	1-3, 11-12
A		4-10
X	WO 2022/209501 A1 (富士フイルム株式会社) 06.10.2022 (2022-10-06) 段落0010-0015, 0032-0038, 0075-0083, 図3	1-3, 8-9, 11-12
A		4-7, 10
A	WO 2021/095698 A1 (ソニーグループ株式会社) 20.05.2021 (2021-05-20) 段落0012-0035	1-12

☐ Ｃ欄の続きにも文献が列挙されている。

☑ パテントファミリーに関する別紙を参照。

* 引用文献のカテゴリー
“A” 特に関連のある文献ではなく、一般的技術水準を示すもの
“D” 国際出願で出願人が先行技術文献として記載した文献
“E” 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの
“L” 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す）
“O” 口頭による開示、使用、展示等に言及する文献
“P” 国際出願日前で、かつ優先権の主張の基礎となる出願の日の後に公表された文献
“T” 国際出願日又は優先日後に公表された文献であって出願と抵触するものではなく、発明の原理又は理論の理解のために引用するもの
“X” 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの
“Y” 特に関連のある文献であって、当該文献と他の１以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの
“&” 同一パテントファミリー文献

国際調査を完了した日	国際調査報告の発送日
24.07.2024	06.08.2024

名称及びあて先

日本国特許庁(ISA/JP)
〒100-8915
日本国
東京都千代田区霞が関三丁目４番３号

権限のある職員（特許庁審査官）

稲垣　良一 5P 4055

電話番号 03-3581-1101 内線 3569

国際調査報告
パテントファミリーに関する情報

国際出願番号
PCT/JP2024/019506

引用文献	公表日	パテントファミリー文献	公表日
WO 2021/206053 A1	14.10.2021	US 2023/0148243 A1 段落0015-0017, 0024-0026, 0036-0040, 図5 EP 4134909 A1 段落0010-0012, 0019-0021, 0031-0035, 図5 CN 115380307 A 段落0020-0023, 0032-0034, 0045-0049, 図5 KR 10-2022-0164755 A 段落0010-0013, 0022-0024, 0036-0039, 図5	
WO 2022/209501 A1	06.10.2022	US 2024/0013524 A1 段落0039-0044, 0061-0067, 0104-0112, 図3	
WO 2021/095698 A1	20.05.2021	US 2022/0399118 A1 段落0056-0081 CN 114586061 A 段落0061-0086	