

(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
2020年2月13日(13.02.2020)



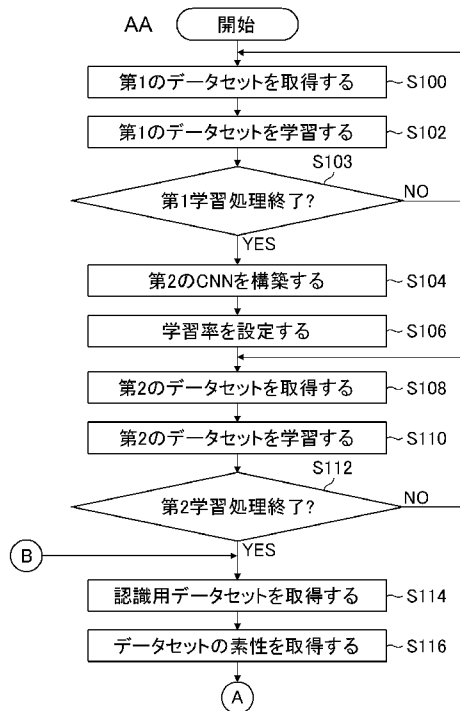
(10) 国際公開番号

WO 2020/031851 A1

- (51) 国際特許分類:
G06T 7/00 (2017.01) *G06N 3/08* (2006.01)
- (21) 国際出願番号: PCT/JP2019/030255
- (22) 国際出願日: 2019年8月1日(01.08.2019)
- (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 2018-149388 2018年8月8日(08.08.2018) JP
- (71) 出願人: 富士フイルム株式会社 (FUJIFILM CORPORATION) [JP/JP]; 〒1068620 東京都港区西麻布2丁目2番30号 Tokyo (JP).
- (72) 発明者: 加門 駿平 (KAMON, Shumpei); 〒2588538 神奈川県足柄上郡開成町宮台798番地 富士フイルム株式会社内 Kanagawa (JP).
- (74) 代理人: 松浦 憲三 (MATSUURA, Kenzo); 〒1630223 東京都新宿区西新宿二丁目6番1号 新宿住友ビル23階 私書箱第176号 新都心国際特許事務所 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, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY,

(54) Title: IMAGE PROCESSING METHOD AND IMAGE PROCESSING DEVICE

(54) 発明の名称: 画像処理方法及び画像処理装置



S100 Acquire first dataset
S102 Learn first data set
S103 Is first learning process terminated?
S104 Construct second CNN
S106 Set learning rate
S108 Acquire second dataset
S110 Learn second data set
S112 Is second learning process terminated?
S114 Acquire recognition data set
S116 Acquire background of dataset
AA Start

(57) Abstract: The purpose of the present invention is to provide an image processing method and an image processing device capable of efficiently learning images with different backgrounds. In learning and recognition using a hierarchical network, it is experientially known that a layer close to an input functions as a feature extractor for extracting a feature which is necessary for recognition, and a layer close to an output performs recognition by combining extracted features. Therefore, as in an embodiment of the present invention, learning by setting a learning rate on the layer close to the input

MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ,
NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT,
QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL,
SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA,
UG, US, UZ, VC, VN, 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, GH, GM, KE, LR, LS, MW, MZ, NA, RW, 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, 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)）

side of the hierarchical network to be larger relative to a learning rate on the layer close to the output side in a second learning process corresponds to relearning (adjusting) a feature extraction part principally with respect to data sets with different backgrounds. As a result, it is possible to absorb a difference between the data sets, and it is possible to achieve efficient learning as compared to simply performing transfer learning.

(57) 要約：本発明は、素性の異なる画像を効率的に学習できる画像処理方法及び画像処理装置を提供することを目的とする。階層型ネットワークを用いた学習及び認識では、入力に近い層は認識に必要な特徴を抽出する特徴抽出器として機能し、出力に近い層は抽出した特徴を組み合わせることで認識を行っていることが経験的に知られている。したがって、本発明の一の態様のように第2学習処理の際に階層型ネットワークの入力側に近い層での学習率を出力側に近い層の学習率に対し大きく設定して学習を行うことは、素性の異なるデータセットに対して主に特徴抽出部分を再学習（調整）することに相当する。これによりデータセット間の差を吸収することができ、単純に転移学習を行うよりも効率的な学習が可能となる。

明 細 書

発明の名称： 画像処理方法及び画像処理装置

技術分野

[0001] 本発明は画像処理方法及び画像処理装置に関し、特に階層型ネットワークを用いた機械学習を行う画像処理方法及び画像処理装置に関する。

背景技術

[0002] 機械学習の分野では、階層型ネットワークを用いて画像を始め多様なデータの学習や認識を行うことが知られている。階層型ネットワークは一般に特徴抽出や認識等を行う複数の層から構成され、層構成によって多数のパラメータが設定される。これらのパラメータは大量のデータを学習することで最適化する必要がある。

[0003] このような学習を効率的に行うための技術が提案されている。例えば特許文献1には、第1の学習工程を行い、その後学習率を大きくして第2の学習工程を行うことで学習が短時間で完了することが記載されている。

先行技術文献

特許文献

[0004] 特許文献1：特開2017-16414号公報

発明の概要

発明が解決しようとする課題

[0005] 機械学習により構築された認識器の性能は学習に用いたデータセット（画像を学習する場合は画像群）に依存し、認識時に扱うデータの素性が学習時と異なる場合は認識の精度が低下する。このような素性が異なるデータに対しても精度の良い認識器を構築するためにデータの素性ごとに別途学習を行うことが考えられるが、素性ごとに大量のデータを用意することで開発コストの増加に繋がる可能性がある。このような問題に対し上述した特許文献1では「素性の異なる画像をどのように学習するのか」という点については何ら考慮されていない。

[0006] このように、従来の技術では素性の異なる画像を効率的に学習することが困難であった。

[0007] 本発明はこのような事情に鑑みてなされたもので、素性の異なる画像を効率的に学習できる画像処理方法及び画像処理装置を提供することを目的とする。

課題を解決するための手段

[0008] 上述した目的を達成するため、本発明の第1の態様に係る画像処理方法は、画像を受け付けて画像の特徴量を出力する階層型ネットワークを備える画像処理装置の画像処理方法であって、第1画像群から選択された第1データセットを階層型ネットワークに入力して学習させる第1学習処理を実行し、第1学習処理の実行後に、第1画像群とは素性が異なる第2画像群から選択された第2データセットを階層型ネットワークに入力して学習させる第2学習処理であって、階層型ネットワークの入力側に近い層での学習率を出力側に近い層の学習率に対し大きく設定して学習を行う第2学習処理を実行する。

[0009] 階層型ネットワークを用いた学習及び認識では、入力に近い層は認識に必要な特徴を抽出する特徴抽出器として機能し、出力に近い層は抽出した特徴を組み合わせることで認識を行っていることが経験的に知られている。したがって、第1の態様のように、第2学習処理の際に階層型ネットワークの入力側に近い層での学習率を出力側に近い層の学習率に対し大きく設定して学習を行うことは、素性の異なるデータセットに対して主に特徴抽出部分を再学習（調整）することに相当する。これによりデータセット間の差違を吸収することができ、単純に転移学習を行う（第1データセットについて学習を行った状態で、学習率を調整せずに第2データセットについて学習を行う）よりも効率的な学習が可能となる。その結果、少ないデータでの学習で高い認識精度が得られる。

[0010] 第1の態様によれば、このようにして素性の異なる画像を効率的に学習することができる。なお、第1の態様及び以下の各態様において「データセッ

トの素性が異なる」とは、例えば被写体や被検体、撮像デバイス（撮像装置）、撮像条件、画像処理等の内容や特性が異なる場合を含む。また第1の態様及び以下の各態様において、第1データセットと第2データセットとでタスク（特定の被写体や注目領域の検出等、学習及び認識が目的とする処理）が共通であることが好ましい。

[0011] 第2の態様に係る画像処理方法は第1の態様において、第1学習処理と第2学習処理とでは撮像デバイスが異なる画像により構成されるデータセットを入力する。第2の態様は「素性の異なる画像」の一態様を規定するもので、「撮像デバイスが異なる」とはモダリティ（機器の分類や様式）が同一で機種、光学系特性等（撮像デバイスが内視鏡の場合、プロセッサ、スコープ、光源、センサなど）が異なることを意味する。例えば、内視鏡装置とCT装置（CT：Computed Tomography）ではモダリティが異なる。

[0012] 第3の態様に係る画像処理方法は第1または第2の態様において、第1学習処理と第2学習処理とでは異なる画像処理が施された画像により構成されるデータセットを入力する。第3の態様は「素性の異なる画像」の他の態様を規定するもので、「画像処理」の例としては、撮像素子（イメージセンサ）が出力するデータから画像を生成する処理、特定の周波数帯域の成分を強調または低減する処理、特定の対象を強調または目立たなくする処理等を挙げることができるが、これらに限定されるものではない。

[0013] 第4の態様に係る画像処理方法は第1から第3の態様のいずれか1つにおいて、第1学習処理と第2学習処理とでは被検体の異なる部位を撮像して得られた画像により構成されるデータセットを入力する。第4の態様は「素性の異なる画像」のさらに他の態様を規定するもので、「異なる部位」の例としては、人体が被検体の場合の異なる臓器（あるいは同じ臓器の異なる部分）を挙げることができるが、これらに限定されない。また、被検体は人体に限定されるものではない。

[0014] 第5の態様に係る画像処理方法は第1から第4の態様のいずれか1つにおいて、第1学習処理と第2学習処理とでは被検体の属性が異なる画像により

構成されるデータセットを入力する。第5の態様は「素性の異なる画像」のさらに他の態様を規定するもので、「被検体の属性」の例としては人体が被検体の場合の年齢、性別、人種等を挙げることができるが、これらに限定されるものではない。

[0015] 第6の態様に係る画像処理方法は第1から第5の態様のいずれか1つにおいて、第1学習処理及び第2学習処理では医療画像により構成されるデータセットを入力する。医療画像を学習や認識の対象とする場合、それぞれの素性に対し十分な量の画像を用意するのが困難な場合があるが、本発明の画像処理方法により素性の異なる画像を効率的に学習することができる。なお、医療画像とは診断、治療、計測等を目的として人体等の生体を撮影、計測等した結果を画像化したものをいい、例えば内視鏡画像、超音波画像、CT画像、MRI画像（MRI：Magnetic Resonance Imaging）を挙げることができる。また、医療画像は医用画像ともいう。

[0016] 第7の態様に係る画像処理方法は第6の態様において、第1学習処理及び第2学習処理では観察光の波長バランスが異なる内視鏡画像により構成されるデータセットを入力する。内視鏡を用いて画像を取得する場合、画像の使用目的に応じて異なる波長バランスの観察光（「異なる素性」の一態様）が用いられるが、使用目的によって画像の取得量に差があるため、観察光によっては十分な量の画像を用意するのは困難な場合がある。しかしながら第7の態様によれば、このような場合でも効率的に学習することができる。なお、「波長バランスが異なる」とは観察光の波長帯域及び／または観察光における各波長帯域の強度の相対関係が異なることを意味し、例えば白色光と狭帯域光、波長帯域の異なる複数の狭帯域光の組み合わせ等を挙げることができる。

[0017] 第8の態様に係る画像処理方法は第1から第7の態様のいずれか1つにおいて、階層型ネットワークは畳み込みニューラルネットワークである。第8の態様は階層型ネットワークの具体的態様の一例を規定するものである。

[0018] 第9の態様に係る画像処理方法は第1から第8の態様のいずれか1つにお

いて、階層型ネットワークに入力する際の画像のサイズを第1サイズとし、第1サイズより小さいサイズを第2サイズとした場合に、第2学習処理では、階層型ネットワークを構成する層のうち第2サイズよりも大きいサイズのデータが入力される層を入力側に近い層とし、階層型ネットワークを構成する層のうち第2サイズ以下のサイズのデータが入力される層を出力側に近い層として、入力側に近い層での学習率を出力側に近い層での学習率に対し大きく設定して学習を行う。第9の態様は、階層型ネットワークでは入出力されるデータのサイズは層構成によって異なることに鑑みて「入力側に近い層」と「出力側に近い層」の一態様を規定するものである。なお、第9の態様において「データ」は一番初めの層に入力される画像だけでなく、階層型ネットワークで生成される特徴マップ等を含む。

[0019] 第10の態様に係る画像処理方法は第1から第9の態様のいずれか1つにおいて、階層型ネットワークは畳み込み層と全結合層とを有し、第2学習処理では、階層型ネットワークを構成する層のうち全結合層を含まない層を入力側に近い層とし、階層型ネットワークを構成する層のうち全結合層を含む層を出力側に近い層として、入力側に近い層での学習率を出力側に近い層での学習率に対し大きく設定して学習を行う。第10の態様は、「入力側に近い層」と「出力側に近い層」の他の態様を規定するものである。なお、「全結合層を含む層」は全結合層以外の層を含んでいてもよい。

[0020] 第11の態様に係る画像処理方法は第1から第10の態様のいずれか1つにおいて、階層型ネットワークは畳み込み層と全結合層とを有し、第2学習処理では、階層型ネットワークを構成する層のうち畳み込み層を含む層を入力側に近い層とし、階層型ネットワークを構成する層のうち畳み込み層を含まない層を出力側に近い層として、入力側に近い層での学習率を出力側に近い層での学習率に対し大きく設定して学習を行う。第11の態様は、「入力側に近い層」と「出力側に近い層」のさらに他の態様を規定するものである。

[0021] 第12の態様に係る画像処理方法は第1から第11の態様のいずれか1つ

において、第1学習処理では、第2データセットより多くの数の画像により構成されるデータセットを第1データセットとして学習を行う。画像の素性によっては十分な量のデータが準備できない場合があるため、第12の態様では、第2データセットより多くの数の画像により構成されるデータセットを第1データセットとして第1学習処理を行い、この第1学習処理の実行後に少ない数の画像で構成される第2データセットを用いて第2学習処理を行う。画像数の多い第1データセットを用いた第1学習処理により十分に学習された状態で第2学習処理を行うことにより、画像の数が少ない第2データセットについても効率的に学習することができる。

[0022] 第13の態様に係る画像処理方法は第1から第12の態様のいずれか1つにおいて、画像で構成されるデータセットを取得するデータセット取得処理と、画像の素性を取得する素性取得処理と、第1学習処理により得られる第1学習済みモデルを有し、入力されるデータセットの認識結果を出力する第1認識器と、第1学習処理及び第2学習処理により得られる第2学習済みモデルを有し、入力されるデータセットの認識結果を出力する第2認識器と、のうち少なくとも一方に、素性に応じて取得したデータセットを入力する入力制御処理と、を行う。第13の態様では、第1から第12の態様のいずれか1つに係る画像処理方法により構築される第1、第2認識器を用いて高い精度で認識を行うことができる。例えば、認識のために入力される画像が第1画像群と同一または類似の素性を有する場合は、少なくとも第1認識器に入力することにより高い精度で認識を行うことができる。なお、この場合第1、第2認識器の双方に入力して一方の認識からの結果のみ出力（表示、記録等）してもよい。

[0023] なお、第1から第13の態様に係る画像処理方法を画像処理装置あるいはコンピュータに実行させるプログラム、並びにそのプログラムのコンピュータ読み取り可能なコードを記録した非一時的記録媒体も本発明の態様として挙げることができる。

[0024] 上述した目的を達成するため、本発明の第14の態様に係る画像処理装置

は第1画像群から選択された第1データセット及び第1画像群とは異なる第2画像群から選択された第2データセットを取得する学習データ取得部と、第1データセットまたは第2データセットを受け付けて特徴量を出力する階層型ネットワークと、第1データセットまたは第2データセットを階層型ネットワークに入力して学習させる学習制御部と、を備え、学習制御部は、第1データセットを階層型ネットワークに入力して学習させる第1学習処理を実行し、第1学習処理の実行後に、第2データセットを階層型ネットワークに入力して学習させる第2学習処理であって、階層型ネットワークの入力側に近い層での学習率を出力側に近い層の学習率に対し大きく設定して学習させる第2学習処理を実行する。第14の態様によれば、第1の態様と同様に素性の異なる画像を効率的に学習することができる。

[0025] 第15の態様に係る画像処理装置は第14の態様において、第1学習処理では、第2データセットより多くの数の画像により構成されるデータセットを第1データセットとして学習を行う。第15の態様によれば、第12の態様と同様に画像の数が少ない第2データセットについても効率的に学習することができる。

[0026] 第16の態様に係る画像処理装置は第14または第15の態様において、画像で構成されるデータセットを取得するデータセット取得部と、データセットの素性を取得する素性取得部と、第1学習処理により得られる第1学習済みモデルを有し、入力されるデータセットの認識結果を出力する第1認識器と、第1学習処理及び第2学習処理により得られる第2学習済みモデルを有し、入力されるデータセットの認識結果を出力する第2認識器と、第1認識器と、第2認識器と、のうち少なくとも一方に、素性に応じて取得したデータセットを入力する入力制御部と、を備える。第16の態様によれば、第13の態様と同様に高い精度で認識を行うことができる。

[0027] なお、第14から第16の態様に係る画像処理装置に対し、第2から第11の態様と同様の構成をさらに含めてもよい。

発明の効果

[0028] 以上説明したように、本発明の画像処理方法及び画像処理装置によれば、素性の異なる画像を効率的に学習することができる。

図面の簡単な説明

[0029] [図1]図1は、第1の実施形態に係る画像処理システムの構成を示す図である。

[図2]図2は、畳み込みニューラルネットワークの層構成の例を示す図である。

[図3]図3は、フィルタによる畳み込み処理の様子を示す図である。

[図4]図4は、内視鏡システムの構成を示す図である。

[図5]図5は、内視鏡システムの構成を示す他の図である。

[図6]図6は、第1の実施形態に係る画像処理方法を示すフローチャートである。

[図7]図7は、第1の実施形態に係る画像処理方法を示すフローチャート（図6の続き）である。

[図8]図8は、第1，第2学習処理の様子を示す図である。

[図9]図9は、データサイズで入力側層及び出力側層を規定する様子を示す図である。

[図10]図10は、全結合層の有無により入力側層及び出力側層を規定する様子を示す図である。

[図11]図11は、畳み込み層の有無により入力側層及び出力側層を規定する様子を示す図である。

発明を実施するための形態

[0030] 以下、添付図面を参照しつつ、本発明に係る画像処理方法及び画像処理装置の実施形態について詳細に説明する。

[0031] <第1の実施形態>

<画像処理システムの構成>

図1は、第1の実施形態に係る画像処理システム1（画像処理装置）の構成を示すブロック図である。画像処理システム1は、被検体内に挿入される

内視鏡で撮像された画像に基づく学習処理及び／または認識処理を行う画像処理装置 500（画像処理装置）と、1つ以上の内視鏡システム 10と、第1の画像データベース 601と、第2の画像データベース 602とを備える。

[0032] <第1, 第2の画像データベース>

<通常光画像及び特殊光画像>

第1の画像データベース 601及び第2の画像データベース 602は、ハードディスク等の記録媒体により構成され、内視鏡システム 10で取得した内視鏡画像が記録される。第1の画像データベース 601には、例えば通常光を観察光（第1の観察光）として撮影された複数の内視鏡画像（医療画像）を記録し、第2の画像データベース 602には、通常光と異なる観察光（第2の観察光；例えば特殊光）として撮影された複数の内視鏡画像を記録することができる。すなわち、第1の画像データベース 601に記録された複数の通常光画像は本発明における「第1画像群」の一態様であり、第2の画像データベース 602に記録された複数の特殊光画像は本発明における「第1画像群とは素性が異なる第2画像群」の一態様である。特殊光画像を撮影する特殊光（狭帯域光）は以下では短波長狭帯域光（青色狭帯域光）として説明するが、長波長狭帯域光（赤色狭帯域光）等他の波長でもよい。また、上述の例では第1, 第2の観察光が白色光と狭帯域光である場合について説明しているが、波長帯域及び／または強度が異なる第1, 第2の狭帯域光を観察光として取得された内視鏡画像等の医用画像を用いてもよい。

[0033] <素性が異なるデータの例>

上述した通常光画像と特殊光画像のように観察光の波長バランスが異なる内視鏡画像だけでなく、第1の画像データベース 601と第2の画像データベース 602とで撮像デバイス、解像度、及び画像に施す画像処理のうち1または複数の異なる内視鏡画像を記録してもよい。「撮像デバイスが異なる」には異なるスコープ、光源や光学系の特性、プロセッサの性能が異なる内視鏡を用いていることが含まれる（複数の内視鏡システム 10の間で特性が

異なる場合もある）が、これに限定されるものではない。また、「画像に施す画像処理が異なる」には、注目領域等特定の領域を強調または目立たなくする処理、特定の波長成分の影響を強調または低減する処理の有無及び／または程度が異なることが含まれるが、これらに限定されるものではない。また、これらのデータに加えて、またはこれに代えて被検体の異なる部位を撮像して得られた画像（内視鏡画像）を第１の画像データベース６０１と第２の画像データベース６０２に記録してもよいし、被検体の属性が異なる画像（内視鏡画像）を記録してもよい。「被検体の属性」の例としては年齢、性別、人種等を挙げることができるが、これらに限定されるものではない。

[0034] ＜データ取得条件によるデータ数の違い＞

内視鏡を用いた観察や検査を行う場合、ユーザは通常光（白色光）を観察光として取得された画像をモニタに表示させて確認するケースが多い。観察や検査の目的、状況（例えば、通常光では病変の構造が観察しづらい）により狭帯域光等の特殊光を観察光として画像をケースもあるが、通常光と比較すると観察光としての使用頻度が低く、そのため特殊光画像は通常光画像よりも著しく数が少ない場合が多い。機械学習により画像の学習及び／または認識を行う場合、特殊光画像についても学習及び／または認識を行う必要があるが、データ数が少ないと通常光画像と比較して学習及び／または認識の精度が低下するおそれがある。このような状況に鑑み、第１の実施形態では後述する学習処理により素性が異なるデータを効率的に学習及び／または認識できるようにしている。

[0035] ＜内視鏡画像の正解データ＞

第１の画像データベース６０１及び第２の画像データベース６０２は、上述した内視鏡画像に加え、注目領域（ROI: Region of Interest）を識別するための「正解データ」を画像と対応させて記憶することができる。例えば、第１の画像データベース６０１は複数の通常光画像にそれぞれ対応する複数の正解データを記憶し、第２の画像データベース６０２は複数の特殊光画像にそれぞれ対応する複数の正解データを記憶する。正解データは、内視

鏡画像に対して医師が指定した注目領域や鑑別結果であることが好ましい。
また、このような観察光の波長バランスが異なる内視鏡画像の正解データに代えて、またはこれに加えて、撮像デバイス、解像度、及び画像に施す画像処理、被検体の部位（撮像部位）や属性についての正解データを記録してもよい。

[0036] <学習認識装置の構成>

画像取得部510は、外部サーバ、データベース等とネットワークを介して通信する装置等により構成され、学習や認識に用いる内視鏡画像及び／または正解データを第1の画像データベース601、第2の画像データベース602から取得する。画像取得部510は、図示せぬネットワークで画像処理システム1と接続された内視鏡システム、病院内サーバ等からも内視鏡画像を取得することができる。操作部520は図示せぬキーボード、マウス等の入力デバイスを備え、ユーザはこれらデバイスを介して画像取得、学習や認識等の処理に必要な操作を行うことができる。制御部530は記録部550に記録された各種プログラムを読み込み、操作部520から入力される指令にしたがって、画像処理システム1全体の動作を制御する。また制御部530は、後述する誤差算出部568が算出した誤差（損失）を第1のCNN562及び／または第2のCNN563（CNN：Convolutional Neural Network、畳み込みニューラルネットワーク）に逆伝搬することにより、これらCNNの重みパラメータを更新する。

[0037] 表示部540はモニタ542（表示装置）を備え、内視鏡画像、学習結果、認識結果、処理条件設定画面等を表示する。記録部550は図示せぬROM（Read Only Memory）、RAM（Random Access Memory）、ハードディスク等で構成され、画像取得部510が取得したデータ、処理部560での学習結果及び／または認識結果等を記録する。また、記録部550は内視鏡画像（医用画像）の学習、認識を行うためのプログラム（本発明の画像処理方法を画像処理システム1に実行させるプログラムを含む）を記録する。処理部560は、第1のCNN562及び第2のCNN563と、これらCNN

の出力（認識結果）と上述した「正解データ」とに基づいて損失（誤差）を算出する誤差算出部568とを備える。

[0038] <第1のCNN及び第2のCNN>

第1の実施形態に係る画像処理システム1（画像処理装置500）は、畳み込みニューラルネットワーク（階層型ネットワーク）である第1のCNN562及び第2のCNN563を備える。後述する第1、第2学習処理が実行された状態において、第1のCNN562は第1学習処理により構築されたネットワークであり、第2のCNN563は第1のCNN562と同一の層構成を有するネットワークに対して第1のCNN562の状態（重みパラメータ等の値）を適用し、その状態で第2学習処理を行って構築されたネットワークである。すなわち、第2のCNN563は第1学習処理後に第2学習処理を行って構築したのと同様のCNNである。

[0039] <CNNの層構成>

第1のCNN562及び第2のCNN563の層構成について説明する。以下では主として第1のCNN562について説明するが、第2のCNN563も同様の構成を採用することができる。なお、第1のCNN562の重みパラメータ等（第1学習処理の結果）を第2のCNN563に反映すること（詳細は後述する）を考慮すると、第1のCNN562と第2のCNN563とで層構成（層の種類、数、及び順序、各層におけるノードの数等）を同一にすることが好ましい。

[0040] 図2は第1のCNN562の層構成の例を示す図である。図2の（a）部分に示す例では、第1のCNN562は入力層562Aと、中間層562Bと、出力層562Cとを含む。入力層562Aは第1の画像データベース601に記憶された通常光画像から選出された画像を入力して特徴量を出力する。中間層562Bは畳み込み層564及びプーリング層565を含み、入力層562Aが出力する特徴量を入力して他の特徴量を算出する。これらの層は複数の「ノード」が「エッジ」で結ばれた構造となっており、複数の重みパラメータを保持している。重みパラメータの値は、学習が進むにつれて

変化していく。第1のCNN 562は、図2の(b)部分に示す例のように全結合層566を含んでいてもよい。第1のCNN 562の層構成は畳み込み層564とプーリング層565とが1つずつ繰り返される場合に限らず、いずれかの層（例えば、畳み込み層564）が複数連続して含まれていてもよい。また、全結合層566が複数連続して含まれていてもよい。

[0041] <中間層における処理>

中間層562Bは、畳み込み演算及びプーリング処理によって特徴量を算出する。畳み込み層564で行われる畳み込み演算はフィルタを使用した畳み込み演算により特徴マップを取得する処理であり、画像からのエッジ抽出等の特徴抽出の役割を担う。このフィルタを用いた畳み込み演算により、1つのフィルタに対して1チャンネル(1枚)の「特徴マップ」が生成される。

「特徴マップ」のサイズは、畳み込みによりダウンスケーリングされ、各層で畳み込みが行われるにつれて小さくなって行く。プーリング層565で行われるプーリング処理は畳み込み演算により出力された特徴マップを縮小（または拡大）して新たな特徴マップとする処理であり、抽出された特徴が、平行移動などによる影響を受けないようにロバスト性を与える役割を担う。中間層562Bは、これらの処理を行う1または複数の層により構成することができる。

[0042] 図3は、図2に示した第1のCNN 562の中間層562Bの構成例を示す模式図である。中間層562Bの最初(1番目)の畳み込み層では、学習用(認識時は認識用)の画像セットとフィルタ F_1 との畳み込み演算が行われる。画像セットは、縦がH、横がWの画像サイズを有するN枚(Nチャンネル)の画像である。通常光画像を入力する場合、データセットを構成する画像はR(赤色)、G(緑色)、B(青色)の3チャンネルの画像である。この画像セットと畳み込み演算されるフィルタ F_1 は、画像セットがNチャンネル(N枚)であるため、例えばサイズ5(5×5)のフィルタの場合、フィルタサイズは5×5×Nのフィルタになる。このフィルタ F_1 を用いた畳み込み演算により、1つのフィルタ F_1 に対して1チャンネル(1枚)の「特徴マッ

プ」が生成される。2番目の畳み込み層で使用するフィルタ F_2 は、例えばサイズ3（ 3×3 ）のフィルタの場合、フィルタサイズは $3 \times 3 \times M$ になる。

[0043] 1番目の畳み込み層と同様に、2番目から n 番目の畳み込み層ではフィルタ $F_2 \sim F_n$ を用いた畳み込み演算が行われる。 n 番目の畳み込み層における「特徴マップ」のサイズが、2番目の畳み込み層における「特徴マップ」のサイズよりも小さくなっているのは、前段までの畳み込み層またはプーリング層によりダウンスケーリングされているからである。

[0044] 中間層562Bの層のうち、入力側に近い畳み込み層では低次の特徴抽出（エッジの抽出等）が行われ、出力側に近づくにつれて高次の特徴抽出（対象物の形状、構造等に関する特徴の抽出）が行われる（図8の（a）部分を参照）。なお、セグメンテーションを行う場合は後半部分の畳み込み層でアップスケーリングされ、最後の畳み込み層では、入力した画像セットと同じサイズの「特徴マップ」が得られる。一方、物体検出を行う場合は位置情報を出力すればよいのでアップスケーリングは必須ではない。第1の実施形態では、これらの点に鑑み「入力側に近い層」と「出力側に近い層」とで学習率を変えて第2の学習処理を行う（図8の（b）部分を参照；詳細は後述）。

[0045] なお、中間層562Bは畳み込み層564及びプーリング層565の他にバッチノーマライゼーションを行う層を含んでいてもよい。バッチノーマライゼーション処理は学習を行う際のミニバッチを単位としてデータの分布を正規化する処理であり、学習を速く進行させる、初期値への依存性を下げる、過学習を抑制する等の役割を担う。

[0046] <出力層における処理>

出力層562Cは、中間層562Bから出力された特徴量に基づき、入力された画像（通常光画像、特殊光画像）に映っている注目領域の位置検出を行ってその結果を出力する層である。セグメンテーションを行う場合、出力層562Cは、中間層562Bから得られる「特徴マップ」により、画像に

写っている注目領域の位置を画素レベルで把握する。すなわち、内視鏡画像の画素ごとに注目領域に属するか否かを検出し、その検出結果を出力することができる。一方、物体検出を行う場合は画素レベルでの判断は必要なく、出力層 5 6 2 C が対象物の位置情報を出力する。

[0047] 出力層 5 6 2 C は病変に関する鑑別を実行して鑑別結果を出力するものでもよい。例えば、出力層 5 6 2 C は内視鏡画像を「腫瘍性」、「非腫瘍性」、「その他」の 3 つのカテゴリに分類し、鑑別結果として「腫瘍性」、「非腫瘍性」及び「その他」に対応する 3 つのスコア（3 つのスコアの合計は 100%）として出力してもよいし、3 つのスコアから明確に分類できる場合には分類結果を出力してもよい。なお鑑別結果を出力する場合、出力層 5 6 2 C が最後の 1 層または複数の層として全結合層を有することが好ましい（図 2 の（b）部分を参照）。

[0048] <内視鏡システムの構成>

図 4 は、内視鏡システム 10（画像処理装置、診断支援装置、内視鏡システム、医療画像処理装置）の外観図であり、図 5 は内視鏡システム 10 の要部構成を示すブロック図である。図 4、5 に示すように、内視鏡システム 10 は、内視鏡本体 100（内視鏡）、プロセッサ 200（プロセッサ、画像処理装置、医療画像処理装置）、光源装置 300、及びモニタ 400（表示装置）から構成される。

[0049] <内視鏡本体の構成>

内視鏡本体 100 は、手元操作部 102 と、この手元操作部 102 に連設される挿入部 104 とを備える。術者（ユーザ）は手元操作部 102 を把持して操作し、挿入部 104 を被検体（生体）の体内に挿入して観察する。また、手元操作部 102 には送気送水ボタン 141、吸引ボタン 142、及び各種の機能を割り付けられる機能ボタン 143、及び撮影指示操作（静止画像、動画像）を受け付ける撮影ボタン 144 が設けられている。挿入部 104 は、手元操作部 102 側から順に、軟性部 112、湾曲部 114、先端硬質部 116 で構成されている。すなわち、先端硬質部 116 の基端側に湾曲

部 1 1 4 が接続され、湾曲部 1 1 4 の基端側に軟性部 1 1 2 が接続される。挿入部 1 0 4 の基端側に手元操作部 1 0 2 が接続される。ユーザは、手元操作部 1 0 2 を操作することにより湾曲部 1 1 4 を湾曲させて先端硬質部 1 1 6 の向きを上下左右に変えることができる。先端硬質部 1 1 6 には、撮影光学系 1 3 0、照明部 1 2 3、鉗子口 1 2 6 等が設けられる（図 4，図 5 参照）。

[0050] 観察、処置の際には、操作部 2 0 8（図 5 参照）の操作により、照明部 1 2 3 の照明用レンズ 1 2 3 A，1 2 3 B から白色光及び／または狭帯域光（赤色狭帯域光、緑色狭帯域光、青色狭帯域光、及び紫色狭帯域光のうち 1 つ以上）を照射することができる。また、送気送水ボタン 1 4 1 の操作により図示せぬ送水ノズルから洗浄水が放出されて、撮影光学系 1 3 0 の撮影レンズ 1 3 2（撮影レンズ、撮像部）、及び照明用レンズ 1 2 3 A，1 2 3 B を洗浄することができる。先端硬質部 1 1 6 で開口する鉗子口 1 2 6 には不図示の管路が連通しており、この管路に腫瘍摘出等のための図示せぬ処置具が挿通されて、適宜進退して被検体に必要な処置を施せるようになっている。

[0051] 図 4，図 5 に示すように、先端硬質部 1 1 6 の先端側端面 1 1 6 A には撮影レンズ 1 3 2（撮像部）が配設されている。撮影レンズ 1 3 2 の奥には C M O S（Complementary Metal-Oxide Semiconductor）型の撮像素子 1 3 4（撮像素子、撮像部）、駆動回路 1 3 6、A F E 1 3 8（A F E：Analog Front End）が配設されて、これらの要素により画像信号を出力する。撮像素子 1 3 4 はカラー撮像素子であり、特定のパターン配列（ベイヤー配列、X-T r a n s（登録商標）配列、ハニカム配列等）でマトリクス状に配置（2次元配列）された複数の受光素子により構成される複数の画素を備える。撮像素子 1 3 4 の各画素はマイクロレンズ、赤（R）、緑（G）、または青（B）のカラーフィルタ及び光電変換部（フォトダイオード等）を含んでいる。撮影光学系 1 3 0 は、赤、緑、青の 3 色の画素信号からカラー画像を生成することもできるし、赤、緑、青のうち任意の 1 色または 2 色の画素信号から画像を生成することもできる。なお、第 1 の実施形態では撮像素子 1 3 4 が

CMOS型の撮像素子である場合について説明するが、撮像素子134はCCD (Charge Coupled Device) 型でもよい。なお、撮像素子134の各画素は紫色光源310Vに対応した紫色カラーフィルタ、及び／または赤外光源に対応した赤外用フィルタをさらに備えていてもよい。

[0052] 被検体（腫瘍部、病変部）の光学像は撮影レンズ132により撮像素子134の受光面（撮像面）に結像されて電気信号に変換され、不図示の信号ケーブルを介してプロセッサ200に出力されて映像信号に変換される。これにより、プロセッサ200に接続されたモニタ400に観察画像が表示される。

[0053] また、先端硬質部116の先端側端面116Aには、撮影レンズ132に隣接して照明部123の照明用レンズ123A、123Bが設けられている。照明用レンズ123A、123Bの奥には、後述するライトガイド170の射出端が配設され、このライトガイド170が挿入部104、手元操作部102、及びユニバーサルケーブル106に挿通され、ライトガイド170の入射端がライトガイドコネクタ108内に配置される。

[0054] <光源装置の構成>

図5に示すように、光源装置300は、照明用の光源310、絞り330、集光レンズ340、及び光源制御部350等から構成されており、観察光をライトガイド170に入射させる。光源310は、それぞれ赤色、緑色、青色、紫色の狭帯域光を照射する赤色光源310R、緑色光源310G、青色光源310B、及び紫色光源310Vを備えており、赤色、緑色、青色、及び紫色の狭帯域光を照射することができる。光源310による観察光の照度は光源制御部350により制御され、必要に応じて観察光の照度を下げること、及び照明を停止することができる。

[0055] 光源310は赤色、緑色、青色、及び紫色の狭帯域光を任意の組合せで発光させることができる。例えば、赤色、緑色、青色、及び紫色の狭帯域光を同時に発光させて白色光（通常光）を観察光として照射することもできるし、いずれか1つもしくは2つを発光させることで狭帯域光（特殊光）を照射

することもできる。光源310は、赤外光（狭帯域光の一例）を照射する赤外光源をさらに備えていてもよい。また、白色光を照射する光源と、白色光及び各狭帯域光を透過させるフィルタとにより、白色光または狭帯域光を観察光として照射してもよい。

[0056] <光源の波長帯域>

光源310は白色帯域の光、または白色帯域の光として複数の波長帯域の光を発生する光源でもよいし、白色の波長帯域よりも狭い特定の波長帯域の光を発生する光源でもよい。特定の波長帯域は、可視域の青色帯域もしくは緑色帯域、あるいは可視域の赤色帯域であってもよい。特定の波長帯域が可視域の青色帯域もしくは緑色帯域である場合、390nm以上450nm以下、または530nm以上550nm以下の波長帯域を含み、かつ、390nm以上450nm以下または530nm以上550nm以下の波長帯域内にピーク波長を有していてもよい。また、特定の波長帯域が可視域の赤色帯域である場合、585nm以上615nm以下、または610nm以上730nm以下、の波長帯域を含み、かつ、特定の波長帯域の光は、585nm以上615nm以下または610nm以上730nm以下の波長帯域内にピーク波長を有していてもよい。

[0057] 上述した特定の波長帯域の光は、酸化ヘモグロビンと還元ヘモグロビンとで吸光係数が異なる波長帯域を含み、かつ、酸化ヘモグロビンと還元ヘモグロビンとで吸光係数が異なる波長帯域にピーク波長を有していてもよい。この場合、特定の波長帯域は、 400 ± 10 nm、 440 ± 10 nm、 470 ± 10 nm、または、600nm以上750nmの波長帯域を含み、かつ、 400 ± 10 nm、 440 ± 10 nm、 470 ± 10 nm、または600nm以上750nm以下の波長帯域にピーク波長を有していてもよい。

[0058] また、光源310が発生する光は790nm以上820nm以下、または905nm以上970nm以下の波長帯域を含み、かつ、790nm以上820nm以下または905nm以上970nm以下の波長帯域にピーク波長を有していてもよい。

- [0059] また、光源 310 は、ピークが 390 nm 以上 470 nm 以下である励起光を照射する光源を備えていてもよい。この場合、被検体（生体）内の蛍光物質が発する蛍光の情報を有する医用画像（生体内画像）を取得することができる。蛍光画像を取得する場合は、蛍光法用色素剤（フルオレスチン、アクリジンオレンジ等）を使用してもよい。
- [0060] 光源 310 の光源種類（レーザ光源、キセノン光源、LED 光源（LED : Light-Emitting Diode）等）、波長、フィルタの有無等は被写体の種類、観察の目的等に応じて構成することが好ましく、また観察の際は被写体の種類、観察の目的等に応じて観察光の波長を組合せ及び／または切り替えることが好ましい。波長を切り替える場合、例えば光源の前方に配置され特定波長の光を透過または遮光するフィルタが設けられた円板状のフィルタ（ロータリカラーフィルタ）を回転させることにより、照射する光の波長を切り替えてもよい。
- [0061] また、本発明を実施する際に用いる撮像素子は撮像素子 134 のように各画素に対しカラーフィルタが配設されたカラー撮像素子に限定されるものではなく、モノクロ撮像素子でもよい。モノクロ撮像素子を用いる場合、観察光の波長を順次切り替えて面順次（色順次）で撮像することができる。例えば出射する観察光の波長を（紫色、青色、緑色、赤色）の間で順次切り替えてもよいし、広帯域光（白色光）を照射してロータリカラーフィルタ（赤色、緑色、青色、紫色等）により出射する観察光の波長を切り替えてもよい。また、1 または複数の狭帯域光（緑色、青色、紫色等）を照射してロータリカラーフィルタ（緑色、青色、紫色等）により出射する観察光の波長を切り替えてもよい。狭帯域光は波長の異なる 2 波長以上の赤外光（第 1 狭帯域光、第 2 狭帯域光）でもよい。
- [0062] ライトガイドコネクタ 108（図 4 参照）を光源装置 300 に連結することにより、光源装置 300 から照射された観察光がライトガイド 170 を介して照明用レンズ 123 A、123 B に伝送され、照明用レンズ 123 A、123 B から観察範囲に照射される。

[0063] <プロセッサの構成>

図5に基づきプロセッサ200の構成を説明する。プロセッサ200は、内視鏡本体100から出力される画像信号を画像入力コントローラ202を介して入力し、画像処理部204で必要な画像処理を行ってビデオ出力部206を介して出力する。これによりモニタ400（表示装置）に観察画像（生体内画像）が表示される。これらの処理はCPU210（CPU: Central Processing Unit）の制御下で行われる。通信制御部205は、図示せぬ病院内システム（HIS: Hospital Information System）や病院内LAN（Local Area Network）、及び上述した画像取得部510、第1の画像データベース601、第2の画像データベース602等、あるいは外部のシステムやネットワークとの通信制御を行う。記録部207には、被検体の画像（内視鏡画像、医療画像）、注目領域の検出及び／または分類結果を示す情報等が記録される。音声処理部209は、CPU210及び画像処理部204の制御により、注目領域の検出及び／または分類の結果に応じたメッセージ（音声）等をスピーカ209Aから出力する。

[0064] また、ROM211（ROM: Read Only Memory）は不揮発性の記憶素子（非一時的記録媒体）であり、各種の画像処理方法をCPU210及び／または画像処理部204（画像処理装置、コンピュータ）に実行させるプログラムのコンピュータ読み取り可能なコードが記憶されている。RAM212（RAM: Random Access Memory）は各種処理の際の一時記憶用の記憶素子であり、また画像取得時のバッファとしても使用することができる。

[0065] <画像処理部の機能>

画像処理部204は、特徴量の算出、特定の周波数帯域の成分を強調または低減する処理、特定の対象（注目領域、所望の深さの血管等）を強調または目立たなくする処理等を行うことができる。また、画像処理部204は、白色帯域の光、または白色帯域の光として複数の波長帯域の光を照射して得る通常光画像に基づいて特定の波長帯域の情報を有する特殊光画像を取得する特殊光画像取得部を備えていてもよい。この場合、特定の波長帯域の信号

は、通常光画像に含まれるRGB（R：赤、G：緑、B：青）あるいはCMY（C：シアン、M：マゼンタ、Y：イエロー）の色情報に基づく演算により得ることができる。また、画像処理部204は、白色帯域の光、または白色帯域の光として複数の波長帯域の光を照射して得る通常光画像と、特定の波長帯域の光を照射して得る特殊光画像との少なくとも一方に基づく演算によって、特徴量画像を生成する特徴量画像生成部を備え、医用画像（医療画像）としての特徴量画像を取得及び表示してもよい。なお、これらの機能による処理はCPU210の制御下で行われる。

[0066] <画像処理システムにおける「素性の異なる画像」の取得>

上述した構成の画像処理システム1では、1つの内視鏡システム10で撮影条件や画像処理条件等を変えて「素性の異なる画像」を取得してもよいし、複数の内視鏡システム10で「素性の異なる画像」を取得してもよい。複数の内視鏡システム10どうしでシステムの特性或撮影条件等が異なるものが含まれていてもよい。また、画像取得部510等が外部のシステムからデータを取得してもよい。

[0067] <各種のプロセッサによる機能の実現>

上述した画像取得部510、制御部530、処理部560（第1のCNN562、第2のCNN563、誤差算出部568）、画像処理部204の機能は、各種のプロセッサ（processor）及び記録媒体を用いて実現できる。各種のプロセッサには、例えばソフトウェア（プログラム）を実行して各種の機能を実現する汎用的なプロセッサであるCPU（Central Processing Unit）が含まれる。また、上述した各種のプロセッサには、画像処理に特化したプロセッサであるGPU（Graphics Processing Unit）、FPGA（Field Programmable Gate Array）などの製造後に回路構成を変更可能なプロセッサであるプログラマブルロジックデバイス（Programmable Logic Device：PLD）も含まれる。本発明のように画像の学習や認識を行う場合は、GPUを用いた構成が効果的である。さらに、ASIC（Application Specific Integrated Circuit）などの特定の処理を実行させるために専用に設計された回

路構成を有するプロセッサである専用電気回路なども上述した各種のプロセッサに含まれる。

[0068] 各部の機能は1つのプロセッサにより実現されてもよいし、同種または異種の複数のプロセッサ（例えば、複数のFPGA、あるいはCPUとFPGAの組み合わせ、またはCPUとGPUの組み合わせ）で実現されてもよい。また、複数の機能を1つのプロセッサで実現してもよい。複数の機能を1つのプロセッサで構成する例としては、第1に、コンピュータに代表されるように、1つ以上のCPUとソフトウェアの組合せで1つのプロセッサを構成し、このプロセッサが複数の機能として実現する形態がある。第2に、システムオンチップ（System On Chip：SoC）などに代表されるように、システム全体の機能を1つのIC（Integrated Circuit）チップで実現するプロセッサを使用する形態がある。このように、各種の機能は、ハードウェア的な構造として、上述した各種のプロセッサを1つ以上用いて構成される。さらに、これらの各種のプロセッサのハードウェア的な構造は、より具体的には、半導体素子などの回路素子を組み合わせた電気回路（circuitry）である。

[0069] 上述したプロセッサあるいは電気回路がソフトウェア（プログラム）を実行する際は、実行するソフトウェアのプロセッサ（コンピュータ）読み取り可能なコードをROM（Read Only Memory）等の非一時的記録媒体に記憶しておき、プロセッサがそのソフトウェアを参照する。非一時的記録媒体に記憶しておくソフトウェアは、本発明に係る画像処理方法を実行するためのプログラムを含む。ROMではなく各種光磁気記録装置、半導体メモリ等の非一時的記録媒体にコードを記録してもよい。ソフトウェアを用いた処理の際には例えばRAM（Random Access Memory）が一時的記憶領域として用いられ、また例えば不図示のEEPROM（Electrically Erasable and Programmable Read Only Memory）に記憶されたデータを参照することもできる。

[0070] <画像処理方法>

上述した構成の画像処理システム1における画像処理方法について説明す

る。図 6, 7 は第 1 の実施形態に係る画像処理方法の手順を示すフローチャートである。なお、ユーザは操作部 520 を介して画像処理の実行指示や実行に必要な条件の指定を行うことができ、表示部 540 はこれら指示の際の画面表示、学習処理及び／または認識処理の過程、結果をモニタ 542 に表示することができる。

[0071] <第 1 学習処理>

画像取得部 510（学習データ取得部）は、第 1 の画像データベース 601 に記録されている通常光画像（白色光画像）の画像群から選択された第 1 データセットを取得する（ステップ S100：第 1 学習処理）。第 1 データセットは、複数のミニバッチにより構成されるデータセットでもよい。この場合、第 1 の CNN 562 では、学習を行う際のミニバッチを単位としてデータの分布を正規化する（バッチノーマライゼーション処理）。第 1 データセットの取得は、操作部 520 を介したユーザの指示入力に応じて行ってもよい。

[0072] 画像取得部 510 は、取得した第 1 データセットを第 1 の CNN 562 の入力層 562A に入力する（ステップ S102：第 1 学習処理）。第 1 の CNN 562 は、上述した構成の中間層 562B において、畳み込み演算等により特徴マップを算出する（ステップ S102：第 1 学習処理）。出力層 562C は、中間層 562B から得られる「特徴マップ」により、内視鏡画像の各画素が注目領域に属するか否かを検出し、その検出結果（上述したように、鑑別結果でもよい）を出力する（ステップ S102：第 1 学習処理）。誤差算出部 568 は、出力層 562C が出力する結果と第 1 データセットに対する正解とを比較して損失（誤差）を計算し、損失が小さくなるように、中間層 562B における重みパラメータを出力側の層から入力側の層に向けて更新していく（ステップ S102：第 1 学習処理、誤差逆伝搬）。

[0073] 制御部 530 及び第 1 の CNN 562 は、ステップ S103 の判断が肯定されるまで上述した第 1 学習処理をくり返す。判断が否定されたらステップ S100 に戻る。ステップ S103 の判断は、例えばデータセットの学習を

指定された繰り返し回数終了した場合や誤差が収束した場合に肯定することができる。なお、画像取得部510、第1のCNN562、誤差算出部568による第1学習処理は、制御部530（学習制御部）の制御の下で行われる。

[0074] 第1データセットと第2データセットとで画像の数に差がある場合、第1学習処理では、第2データセットより多くの数の画像により構成されるデータセットを第1データセットとして学習を行うことが好ましい。例えば白色光画像を用いた学習とB L I画像等の短波長狭帯域光画像を用いた学習とを行う場合、白色光画像と同程度の数の短波長狭帯域光画像を準備できない可能性があるが、この場合短波長狭帯域光画像により構成されるデータセット（第2データセット）より多くの数の白色光画像により構成されるデータセットを第1データセットとして第1学習処理を行い、この第1学習処理の実行後に少ない数の画像（短波長狭帯域光画像）で構成される第2データセットを用いて第2学習処理を行うことが好ましい。このような学習により、画像数の多い第1データセットを用いた第1学習処理により十分に学習された状態で第2学習処理を行うことができ、画像の数が少ない第2データセットについても効率的に学習することができる。

[0075] <第2学習処理>

<第2のCNNの構築>

制御部530は、第2のCNN563に第1のCNN562の重みパラメータ（上述した第1学習処理の結果）を反映して、第2のCNN563を構築する（ステップS104：第2学習処理、CNN構築処理）。これにより、第2のCNN563は第1学習処理がなされた状態になる。

[0076] <学習率の設定>

制御部530及び第2のCNN563は、第2のCNN563の各層での学習率を設定する（ステップS106：第2学習処理）。この際、第2のCNN563の入力側に近い層は学習率を大きく、出力側に近い層は学習率を小さくする（図8の（b）部分を参照）。階層型ネットワークを用いた学習

及び認識では、入力に近い層は認識に必要な特徴を抽出する特徴抽出器として機能し、出力に近い層は抽出した特徴を組み合わせることで認識を行っていることが経験的に知られている。したがって、第1の実施形態のように、第2学習処理の際に第2のCNN563の入力側に近い層での学習率を出力側に近い層の学習率に対し大きく設定して学習を行うことは、素性の異なるデータセットに対して主に特徴抽出部分を再学習（調整）することに相当する。これによりデータセット間の差違を吸収することができ、単純に転移学習を行う（第1データセットについて学習を行った状態で、学習率を調整せずに第2データセットについて学習を行う）よりも効率的な学習が可能となる。

[0077] 第1の実施形態では第1データセットが白色光画像、第2データセットがB L I画像であり、これら画像は全体としての色味は大きく異なるが検出対象（注目領域）の特性には共通点がある（周囲より隆起している、周囲と色相が異なる等）。よって白色光画像で学習した第1のCNN562（階層型ネットワーク）の特徴抽出部のみB L I画像用に調整すれば（第2学習処理における学習率の設定）、特徴抽出部以降の層における特徴の組み合わせ方は同じものが利用できる。第1の実施形態では、このように白色光画像で学習した学習済みモデルを利用することで、B L I画像のように少ないデータの場合でも高い認識精度が得られる。

[0078] <学習率の設定例>

以下、第2のCNN563における学習率の設定例について説明する。

[0079] （例1）データサイズによる設定

図9は、各層に入力するデータのサイズに基づいて、第2のCNN563を「入力側に近い層」と「出力側に近い層」に分ける様子を示す概念図である。図9は、第2のCNN563（階層型ネットワーク）に入力する際の画像のサイズを第1サイズとし、第1サイズより小さいサイズを第2サイズとした場合に、第2のCNN563を構成する層のうち第2サイズよりも大きいサイズのデータが入力される層（領域563Lに含まれる層）を「入力側

に近い層」とし、第2サイズ以下のサイズのデータが入力される層を出力側に近い層（領域563Sに含まれる層）とする様子を示している。図中の曲線は、中間層563Bでの畳み込み処理によりデータサイズが変化していく様子を示している。

[0080] 制御部530及び第2のCNN563は、領域563Lに含まれる層での学習率を領域563Sに含まれる層での学習率に対し大きく設定する。なお、第1のCNN562及び第2のCNN563がセグメンテーションを行うCNNである場合は、上述のように中間層562B、563Bの後半部分（出力側）の畳み込み層でアップスケーリングされデータサイズが大きくなるが、アップスケーリングによりデータサイズが再度第2サイズを超える領域563Eに含まれる層では学習率を大きくしなくてよい。なお、第1のCNN562及び第2のCNN563が物体検出を行う場合は、上述のアップスケーリングは必須ではない。

[0081] （例2）全結合層の有無による設定

図10は、入力層563A、中間層563B、及び出力層563Cにより構成される第2のCNN563を「全結合層を含むか否か」に基づいて「入力側に近い層」と「出力側に近い層」に分ける様子を示す図である。図10の（a）部分は、全結合層566以外の層を全て「入力側に近い層」（領域563L）とし、全結合層566のみを「出力側に近い層」（領域563S）とする様子を示している。一方、図10の（b）部分は、全結合層566に加えて畳み込み層564、プーリング層565を含む層を「出力側に近い層」（領域563S）とし、残りの層を「入力側に近い層」（領域563L）とする様子を示している。制御部530及び第2のCNN563は、領域563Lに含まれる層での学習率を領域563Sに含まれる層での学習率に対し大きく設定する。

[0082] （例3）畳み込み層の有無による設定

図11は、第2のCNN563を「畳み込み層を含むか否か」に基づいて「入力側に近い層」と「出力側に近い層」に分ける様子を示す図である。畳

み込み層を含む層を「入力側に近い層」（領域563L）とし、全結合層566のみを「出力側に近い層」（領域563S）とする（図11に示す層構成の場合、結果として図10の（a）部分に示す例と同一の分け方になる）。制御部530及び第2のCNN563は、領域563Lに含まれる層での学習率を領域563Sに含まれる層での学習率に対し大きく設定する。

[0083] 学習率の設定において、出力側に近い層の学習率をゼロにする（固定する）ことができる。学習率は各層に固定値を設定してもよいし、出力側の層から入力側の層にかけて連続的に大きくなるように設定してもよい。また、エポック（学習の繰り返し回数が何回目であるか）に応じて学習率を設定する際に、特定のタイミングのみ上述のような大小関係となるように学習率を設定してもよい（例えば、学習の初期段階では入力側に近い層の学習率を大きく設定して、あるエポック以降は全て同一の学習率に設定する）。

[0084] なお、本発明における学習率設定の際の層の分け方は上述した例1～例3の態様に限定されるものではない。また、制御部530及び第2のCNN563は、複数の分け方のうちからユーザの指示入力に基づいて決定した層の分け方を用いて学習率を設定してもよい。

[0085] <第2データセットの取得及び学習>

ステップS106で学習率を設定したら、上述した第1学習処理と同様の手順で第2学習処理を実行する。ステップS104で第2のCNN563が構築され第1学習処理がなされた状態になっているので、以下の処理により「第1学習処理の実行後に第2学習処理を実行する」ことになる。

[0086] 画像取得部510は、第2の画像データベース602に記録されている短波長狭帯域光画像（BLI画像、BLI（Blue Laser Imaging）は登録商標）の画像群（第2画像群）から選択された第2データセット（第1画像群とは素性が異なる第2画像群から選択されたデータセットである）を取得し（ステップS108：第2学習処理）、取得した第2データセットを第2のCNN563に入力して学習させる（ステップS110：第2学習処理）。第2学習処理においても、第1学習処理と同様に第2データセットを複数のミ

ニバッチにより構成されるデータセットとすることができる。第2のCNN 563では第1のCNN 562と同様に畳み込み演算による特徴マップの算出が行われ、注目領域の検出結果（鑑別結果でもよい）が出力される。第2のCNN 563及び誤差算出部568は誤差逆伝搬処理を実行してもよい。制御部530及び第2のCNN 563は、ステップS112の判断が肯定されるまで第2学習処理をくり返し、判断が否定されたらステップS108に戻る。ステップS112の判断は、例えばデータセットの学習を指定された繰り返し回数終了した場合や誤差が収束した場合に肯定することができる。

[0087] 第1の実施形態によれば、上述した第1、第2学習処理により素性の異なる画像を効率的に学習することができる。また、第1の実施形態のように第1画像群（第1データセット）が白色光画像で第2画像群（第2データセット）が短波長狭帯域光画像（BLI画像）である場合第2画像群を十分な量用意できない可能性があるが、上述した第1、第2学習処理によりデータ量の少ない第2画像群（第2データセット）についても効率的に学習することができる。

[0088] <認識処理>

第1の実施形態に係る画像処理システム1では、上述した第1、第2学習処理により構築されたCNNによる認識処理を行うことができる。認識処理においては、第1の学習処理が実行された状態の第1のCNN 562が第1学習済みモデル（第1認識器）となり入力されるデータセットの認識結果（例えば、注目領域の検出結果や鑑別結果）を出力する。同様に、第1、第2の学習処理が実行された状態の第2のCNN 563が第2学習済みモデル（第2認識器）となり、入力されるデータセットの認識結果を出力する。以下、認識処理の手順を説明する。

[0089] <データセット取得処理>

画像取得部510（データセット取得部）は、制御部530の制御により、画像で構成される認識用データセットを取得する（ステップS114：認識処理、データセット取得処理）。認識用データセットは第1の画像データ

ベース 601 あるいは第 2 の画像データベース 602 に記録されている画像（白色光画像、B L I 画像）で構成されていてもよいし、画像取得部 510 が内視鏡システム 10 や他のシステムから取得した画像で構成されていてもよい。画像取得部 510 は、「いずれのデータベース、装置、システムから認識用データセットを取得するか」を、操作部 520 を介したユーザの指示入力に基づいて決定してもよい。

[0090] <素性取得処理>

画像取得部 510（素性取得部）は、制御部 530 の制御により、ステップ S114 で取得したデータセットの素性を取得する（ステップ S116：認識処理、素性取得処理）。「素性」は、具体的には例えば「データセットが、内視鏡システム 10 で取得され第 1 の画像データベース 601 に記録されている通常光画像（白色光画像）により構成される」という情報である。このような情報に加えて、またはこれに代えて撮像デバイスの識別情報、光源の情報、画像処理の内容、撮像部位、被検体の属性等の情報を含んでいてもよい。素性の情報は第 1 の画像データベース 601 あるいは第 2 の画像データベース 602 に記録されているものを取得してよいし、画像取得部 510 が内視鏡システム 10 や他のシステムから取得してもよい。例えば、操作部 520 を介してユーザが入力する、機器の設定情報から取得する（光源モードや画像処理パラメータの場合）、画像から素性を認識する認識器（例えば、画像から被検体の部位を認識する認識器）の出力結果を利用する、診断時に入力される患者情報（被検体の年齢、性別等の素性の場合）を利用する、等の態様が可能である。

[0091] <入力制御処理>

制御部 530（入力制御部）は、ステップ S116 で取得した素性に応じて、第 1 認識器（第 1 の CNN 562）と第 2 認識器（第 2 の CNN 563）の少なくとも一方にデータセットを入力する（図 7 のステップ S118，S120，S124：認識処理、入力制御処理）。具体的には、制御部 530 は「取得したデータセットの素性が、第 2 学習処理に用いた第 2 データセ

ットの素性よりも第1学習処理に用いた第1データセットの素性に近いか否か」を判断する（ステップS118：認識処理、入力制御処理）。「取得したデータセットの素性が第1学習処理に用いた第1データセットの素性に近い」場合（上述した例では、認識用データセットが白色光画像により構成される場合）はステップS118の判断が肯定され、制御部530は第1認識器（第1のCNN562）にデータセットを入力する（ステップS120：認識処理、入力制御処理）。ステップS118の判断が否定された場合は、制御部530は第2認識器（第2のCNN563）にデータセットを入力する（ステップS124：認識処理、入力制御処理）。

[0092] なお、認識処理時のデータセットの素性が学習処理時のデータセットの素性と全く同一である必要はない。例えば、上述のように白色光画像、短波長狭帯域光画像（BLI画像）をそれぞれ第1、第2データセットとして第1、第2学習処理を行った場合において、認識用データセットがLCI画像（LCI：Linked Color Imaging、登録商標）により構成される場合は第1認識器に入力することができ、また認識用データセットがBLI-b r t画像（BLI-b r i g h t画像）により構成される場合は第2認識器に入力することができる。このような入力ができるのは、LCI画像やBLI-b r t画像はそれぞれ白色光画像、BLI画像に色特性が近いため、上述のような制御則で認識精度の向上が期待できるからである。

[0093] <認識結果の出力>

データセットが入力された第1認識器または第2認識器は、認識結果を出力する（ステップS122、ステップS126：認識処理、出力処理）。認識結果は、例えば注目領域の検出結果や鑑別結果をモニタ542に表示することにより、及び／または記録部550に記録することにより出力される。出力の内容は認識器の構成（注目領域の検出、鑑別、その他）により異なる。認識結果を出力すると、制御部530は認識処理が終了したか否かを判断し（ステップS128：認識処理）、判断が肯定されるまでステップS114からS126の処理をくり返す。ステップS128の判断が肯定されたら

認識処理を終了する。

[0094] なお、図7のフローチャートでは入力制御処理において第1, 第2認識器の一方にデータセットを入力する場合について説明したが、「制御部530が両方の認識器にデータセットを入力し、一方の認識器（認識用データセットの素性が学習処理に用いたデータセットの素性に近い方の認識器）についての認識結果を出力する」態様を採用してもよい。この場合、制御部530は例えば白色光画像から構成される認識用データセットを第1, 第2の認識器に入力し、第1認識器についての認識結果を出力（表示）させる。

[0095] 以上説明したように、第1の実施形態に係る画像処理システム1によれば、第1, 第2学習処理により素性の異なる画像を効率的に学習することができ、また構築された認識器によりデータセットの素性に応じて高精度に画像を認識することができる。

[0096] <その他>

上述した第1の実施形態では、画像処理装置500で学習及び認識を行う場合について説明したが、画像処理装置500の構成を内視鏡システム10（例えばプロセッサ200の画像処理部204）に含めてもよい。また、第1の画像データベース601、第2の画像データベース602をプロセッサ200の記録部207に含めてもよい。また、第1の実施形態では通常光画像（白色光画像）及び短波長狭帯域光画像（BLI画像）を用いた学習及び認識について説明したが、通常光画像と他の狭帯域光画像（例えば、LCI画像等の長波長狭帯域光画像）を用いて学習及び認識を行ってもよい。また、素性の異なる複数の通常光画像、あるいは素性が異なる複数の狭帯域光画像を用いて学習及び認識を行ってもよい。複数の狭帯域光画像を用いる場合、狭帯域光の組み合わせとしては複数の青色狭帯域光、青色狭帯域光と紫色狭帯域光、複数の赤色狭帯域光等を用いることができる。

[0097] また、上述した第1の実施形態では医療画像（医用画像）の一態様である内視鏡画像を用いて学習及び認識を行う場合について説明したが、本発明に係る画像処理装置及び画像処理方法はCT装置（Computed Tomography）、M

R I (Magnetic Resonance Imaging) 装置、超音波画像診断装置等、内視鏡画像以外の医療画像を用いる場合にも適用することができる。

[0098] さらに、本発明に係る画像処理装置及び画像処理方法は医用画像以外の画像（例えば、人物、動物、風景、建築物、工業用製品等他の画像）を用いる場合にも適用することができる。例えば、撮像デバイスや観察光（可視光と赤外光等）、露出条件、画像処理パラメータ等（これらは「素性」の一例である）が異なる画像を用いて建築物の損傷画像（ひび割れ、剥離、浮き等の損傷を撮影した画像）の学習及び認識を行うことができる。

[0099] 以上で本発明の実施形態及び他の例に関して説明してきたが、本発明は上述した態様に限定されず、本発明の精神を逸脱しない範囲で種々の変形が可能である。

符号の説明

[0100]	1	画像処理システム
	1 0	内視鏡システム
	1 0 0	内視鏡本体
	1 0 2	手元操作部
	1 0 4	挿入部
	1 0 6	ユニバーサルケーブル
	1 0 8	ライトガイドコネクタ
	1 1 2	軟性部
	1 1 4	湾曲部
	1 1 6	先端硬質部
	1 1 6 A	先端側端面
	1 2 3	照明部
	1 2 3 A	照明用レンズ
	1 2 3 B	照明用レンズ
	1 2 6	鉗子口
	1 3 0	撮影光学系

1 3 2	撮影レンズ
1 3 4	撮像素子
1 3 6	駆動回路
1 3 8	A F E
1 4 1	送気送水ボタン
1 4 2	吸引ボタン
1 4 3	機能ボタン
1 4 4	撮影ボタン
1 7 0	ライトガイド
2 0 0	プロセッサ
2 0 2	画像入力コントローラ
2 0 4	画像処理部
2 0 5	通信制御部
2 0 6	ビデオ出力部
2 0 7	記録部
2 0 8	操作部
2 0 9	音声処理部
2 0 9 A	スピーカ
2 1 0	C P U
2 1 1	R O M
2 1 2	R A M
3 0 0	光源装置
3 1 0	光源
3 1 0 B	青色光源
3 1 0 G	緑色光源
3 1 0 R	赤色光源
3 1 0 V	紫色光源
3 3 0	絞り

3 4 0	集光レンズ
3 5 0	光源制御部
4 0 0	モニタ
5 0 0	画像処理装置
5 1 0	画像取得部
5 2 0	操作部
5 3 0	制御部
5 4 0	表示部
5 4 2	モニタ
5 5 0	記録部
5 6 0	処理部
5 6 2	第 1 の C N N
5 6 2 A	入力層
5 6 2 B	中間層
5 6 2 C	出力層
5 6 3	第 2 の C N N
5 6 3 A	入力層
5 6 3 B	中間層
5 6 3 C	出力層
5 6 3 E	領域
5 6 3 L	領域
5 6 3 S	領域
5 6 4	畳み込み層
5 6 5	プーリング層
5 6 6	全結合層
5 6 8	誤差算出部
6 0 1	第 1 の画像データベース
6 0 2	第 2 の画像データベース

F 1 フィルタ

F 2 フィルタ

F n フィルタ

S 1 0 0 ～ S 1 2 8 画像処理方法の各ステップ

請求の範囲

- [請求項1] 画像を受け付けて前記画像の特徴量を出力する階層型ネットワークを備える画像処理装置の画像処理方法であって、
- 第1画像群から選択された第1データセットを前記階層型ネットワークに入力して学習させる第1学習処理を実行し、
- 前記第1学習処理の実行後に、前記第1画像群とは素性が異なる第2画像群から選択された第2データセットを前記階層型ネットワークに入力して学習させる第2学習処理であって、前記階層型ネットワークの入力側に近い層での学習率を出力側に近い層の学習率に対し大きく設定して学習を行う第2学習処理を実行する、
- 画像処理方法。
- [請求項2] 前記第1学習処理と前記第2学習処理とでは撮像デバイスが異なる画像により構成されるデータセットを入力する請求項1に記載の画像処理方法。
- [請求項3] 前記第1学習処理と前記第2学習処理とでは異なる画像処理が施された画像により構成されるデータセットを入力する請求項1または2に記載の画像処理方法。
- [請求項4] 前記第1学習処理と前記第2学習処理とでは被検体の異なる部位を撮像して得られた画像により構成されるデータセットを入力する請求項1から3のいずれか1項に記載の画像処理方法。
- [請求項5] 前記第1学習処理と前記第2学習処理とでは被検体の属性が異なる画像により構成されるデータセットを入力する請求項1から4のいずれか1項に記載の画像処理方法。
- [請求項6] 前記第1学習処理及び前記第2学習処理では医療画像により構成されるデータセットを入力する請求項1から5のいずれか1項に記載の画像処理方法。
- [請求項7] 前記第1学習処理及び前記第2学習処理では観察光の波長バランスが異なる内視鏡画像により構成されるデータセットを入力する請求項

6 に記載の画像処理方法。

[請求項8] 前記階層型ネットワークは畳み込みニューラルネットワークである請求項 1 から 7 のいずれか 1 項に記載の画像処理方法。

[請求項9] 前記階層型ネットワークに入力する際の画像のサイズを第 1 サイズとし、前記第 1 サイズより小さいサイズを第 2 サイズとした場合に、
前記第 2 学習処理では、前記階層型ネットワークを構成する層のうち前記第 2 サイズよりも大きいサイズのデータが入力される層を前記入力側に近い層とし、前記階層型ネットワークを構成する層のうち前記第 2 サイズ以下のサイズのデータが入力される層を前記出力側に近い層として、前記入力側に近い層での前記学習率を前記出力側に近い層での前記学習率に対し大きく設定して学習を行う請求項 1 から 8 のいずれか 1 項に記載の画像処理方法。

[請求項10] 前記階層型ネットワークは畳み込み層と全結合層とを有し、
前記第 2 学習処理では、前記階層型ネットワークを構成する層のうち前記全結合層を含まない層を前記入力側に近い層とし、前記階層型ネットワークを構成する層のうち前記全結合層を含む層を前記出力側に近い層として、前記入力側に近い層での前記学習率を前記出力側に近い層での前記学習率に対し大きく設定して学習を行う請求項 1 から 9 のいずれか 1 項に記載の画像処理方法。

[請求項11] 前記階層型ネットワークは畳み込み層と全結合層とを有し、
前記第 2 学習処理では、前記階層型ネットワークを構成する層のうち前記畳み込み層を含む層を前記入力側に近い層とし、前記階層型ネットワークを構成する層のうち前記畳み込み層を含まない層を前記出力側に近い層として、前記入力側に近い層での前記学習率を前記出力側に近い層での前記学習率に対し大きく設定して学習を行う請求項 1 から 10 のいずれか 1 項に記載の画像処理方法。

[請求項12] 前記第 1 学習処理では、前記第 2 データセットより多くの数の画像により構成されるデータセットを前記第 1 データセットとして前記学

習を行う請求項 1 から 1 1 のいずれか 1 項に記載の画像処理方法。

[請求項13] 画像で構成されるデータセットを取得するデータセット取得処理と、

、
前記画像の素性を取得する素性取得処理と、

前記第 1 学習処理により得られる第 1 学習済みモデルを有し、入力されるデータセットの認識結果を出力する第 1 認識器と、前記第 1 学習処理及び前記第 2 学習処理により得られる第 2 学習済みモデルを有し、入力されるデータセットの認識結果を出力する第 2 認識器と、のうち少なくとも一方に、前記素性に応じて前記取得した前記データセットを入力する入力制御処理と、

を行う請求項 1 から 1 2 のいずれか 1 項に記載の画像処理方法。

[請求項14] 第 1 画像群から選択された第 1 データセット及び前記第 1 画像群とは素性が異なる第 2 画像群から選択された第 2 データセットを取得する学習データ取得部と、

前記第 1 データセットまたは前記第 2 データセットを受け付けて特徴量を出力する階層型ネットワークと、

前記第 1 データセットまたは前記第 2 データセットを前記階層型ネットワークに入力して学習させる学習制御部と、

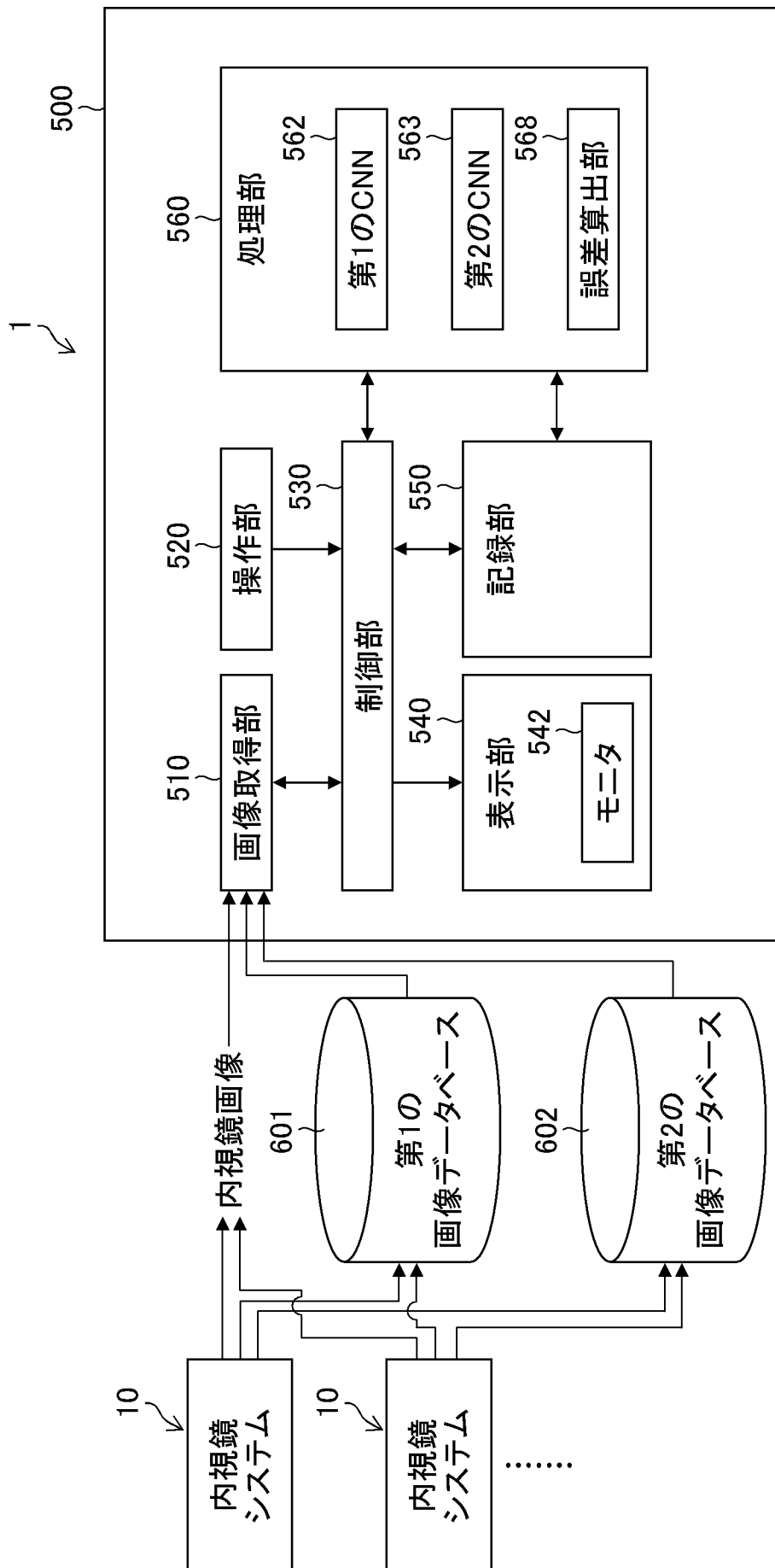
を備え、

前記学習制御部は、前記第 1 データセットを前記階層型ネットワークに入力して学習させる第 1 学習処理を実行し、前記第 1 学習処理の実行後に、前記第 2 データセットを前記階層型ネットワークに入力して学習させる第 2 学習処理であって、前記階層型ネットワークの入力側に近い層での学習率を出力側に近い層の学習率に対し大きく設定して学習させる第 2 学習処理を実行する画像処理装置。

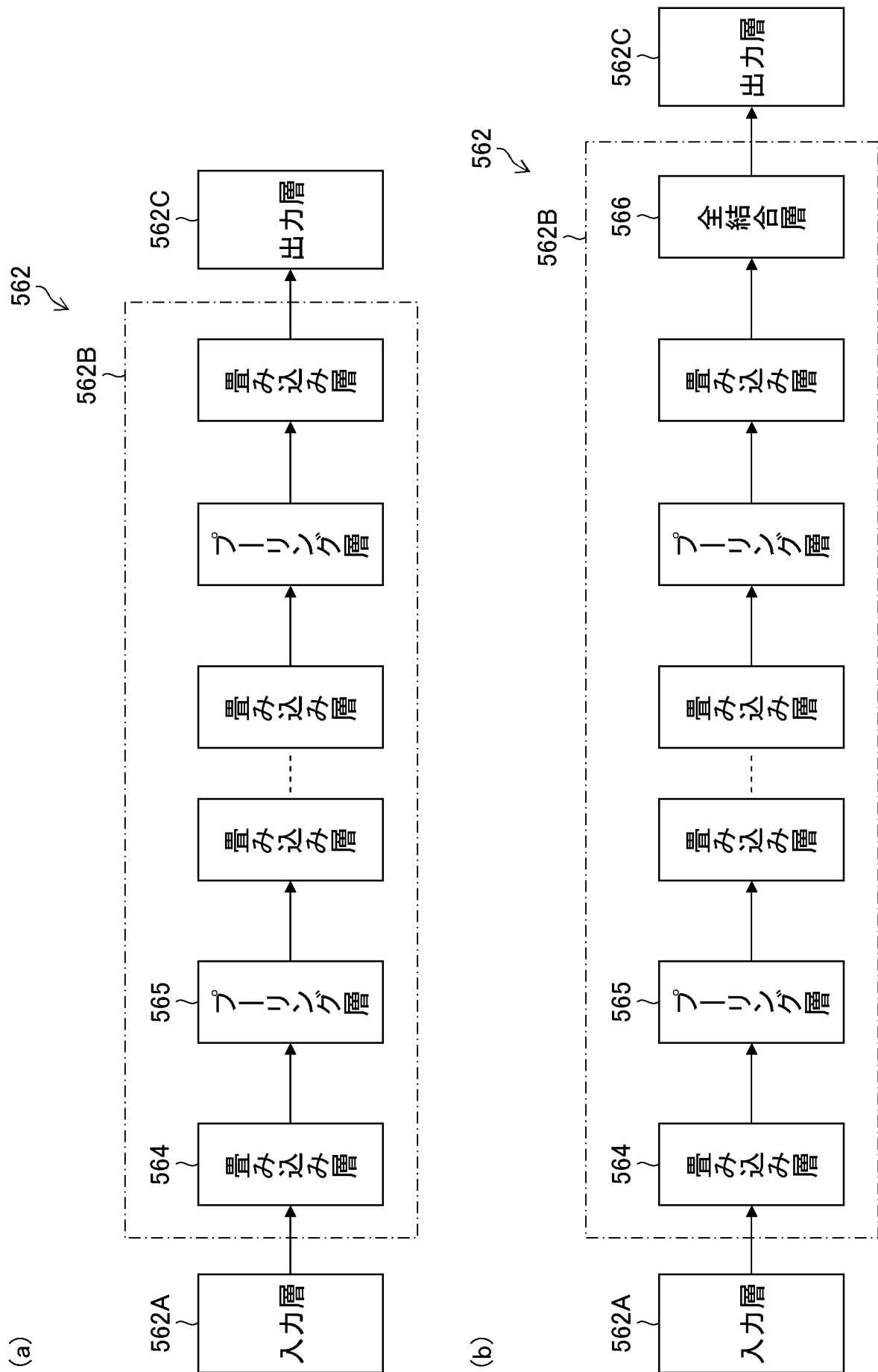
[請求項15] 前記第 1 学習処理では、前記第 2 データセットより多くの数の画像により構成されるデータセットを前記第 1 データセットとして前記学習を行う請求項 1 4 に記載の画像処理装置。

- [請求項16] 画像で構成されるデータセットを取得するデータセット取得部と、
前記データセットの素性を取得する素性取得部と、
前記第1学習処理により得られる第1学習済みモデルを有し、入力されるデータセットの認識結果を出力する第1認識器と、
前記第1学習処理及び前記第2学習処理により得られる第2学習済みモデルを有し、入力されるデータセットの認識結果を出力する第2認識器と、
前記第1認識器と、前記第2認識器と、のうち少なくとも一方に、前記素性に応じて前記取得した前記データセットを入力する入力制御部と、
を備える請求項14または15に記載の画像処理装置。

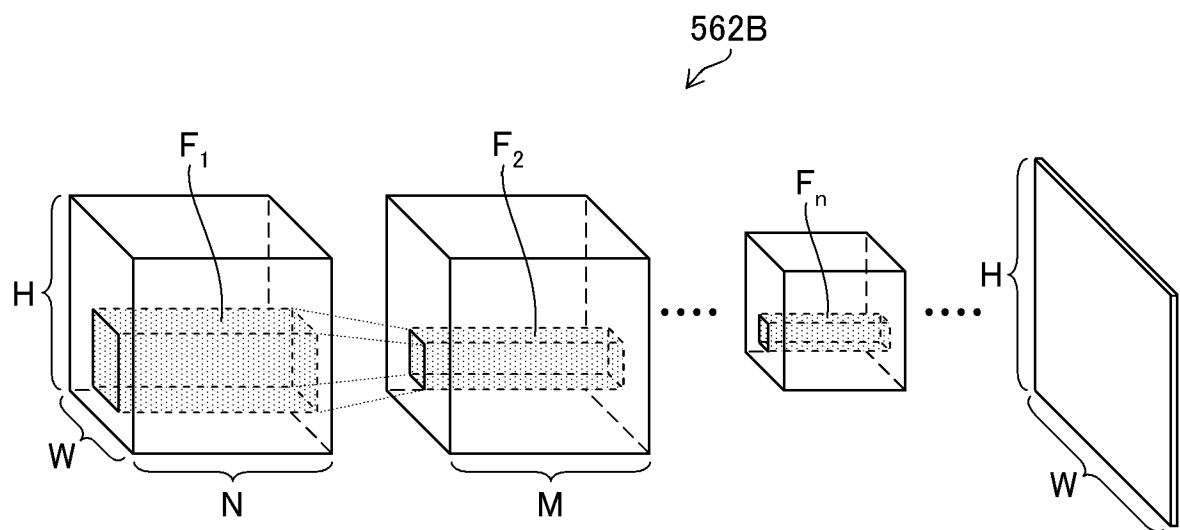
[図1]



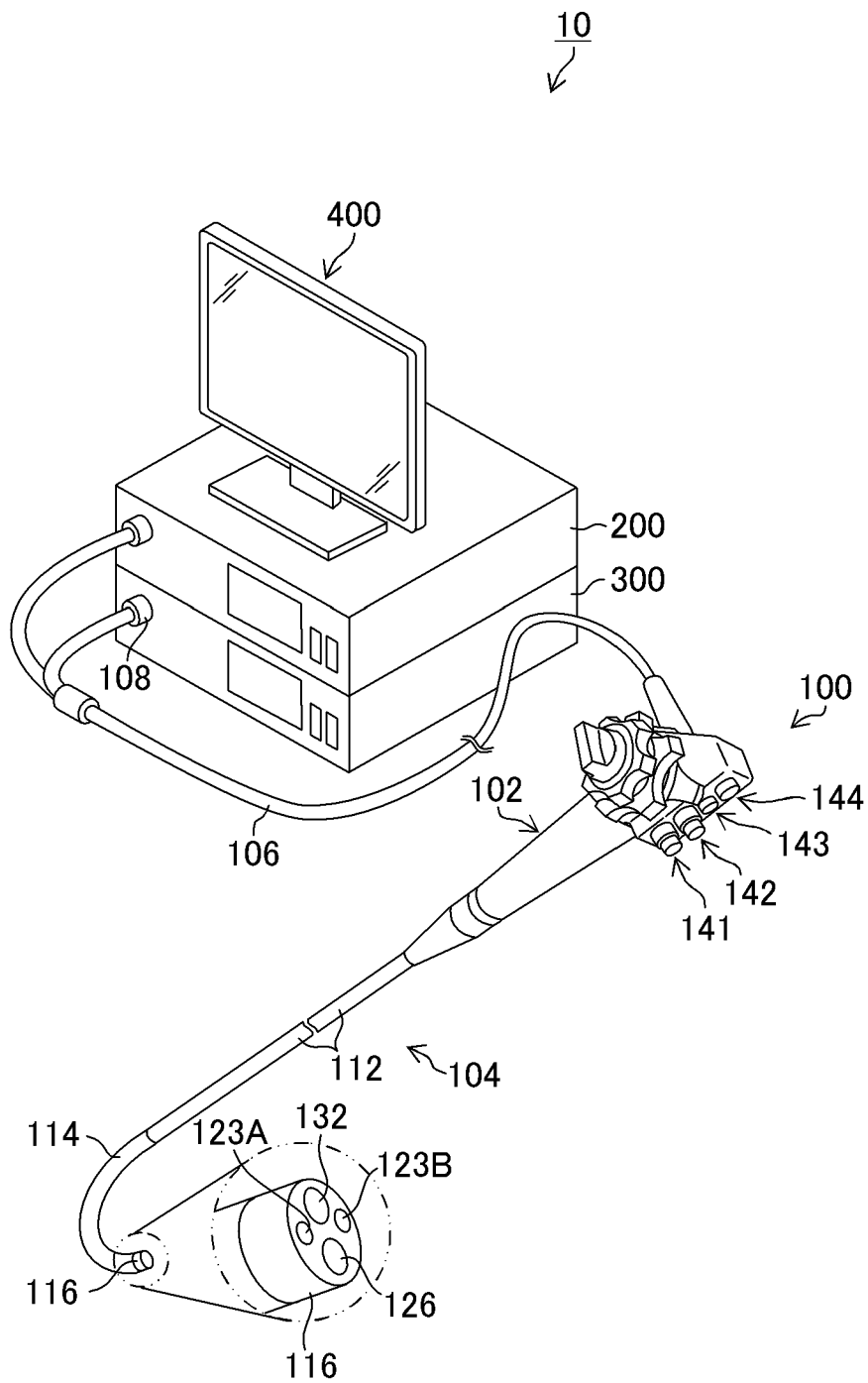
[図2]



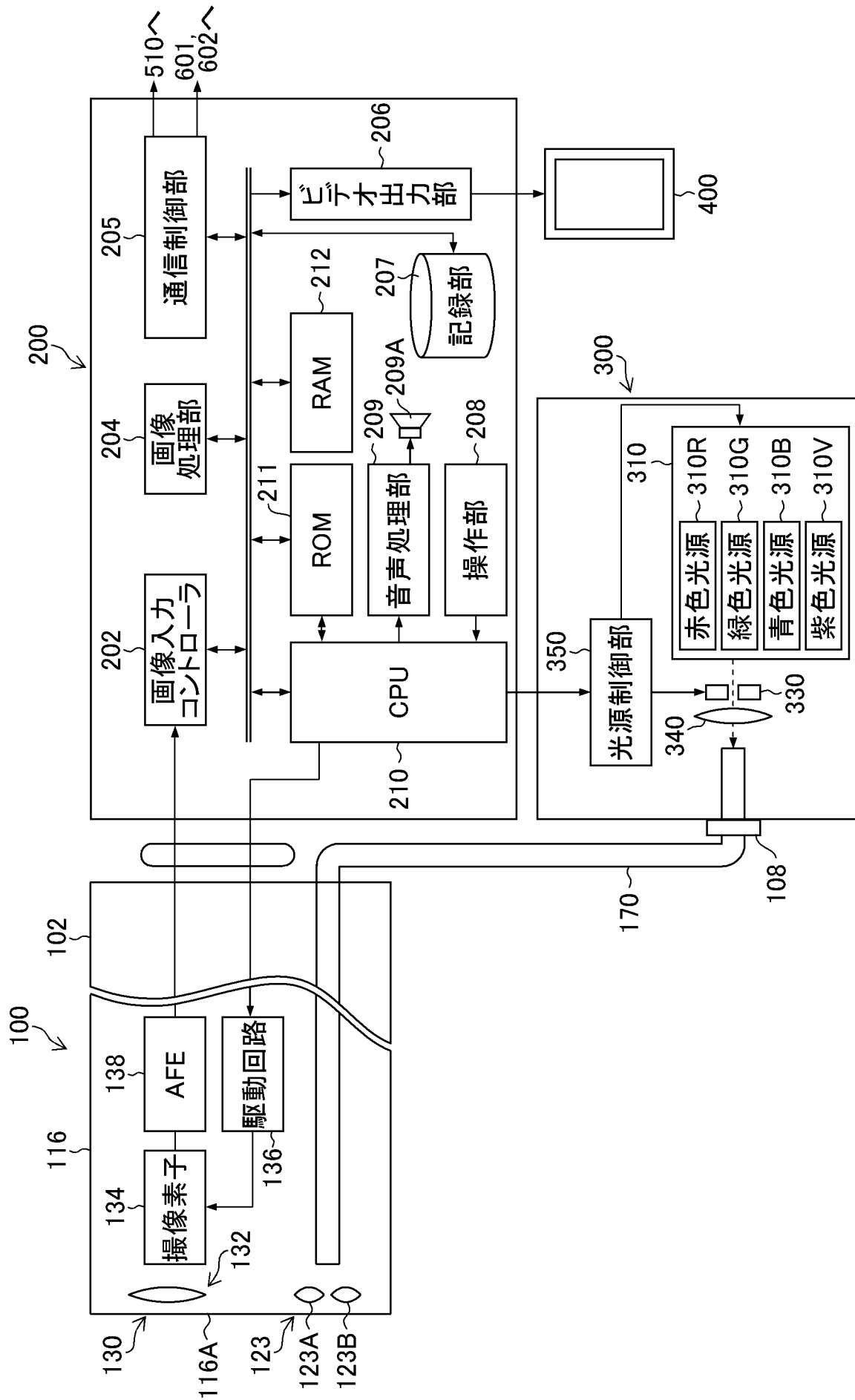
[図3]



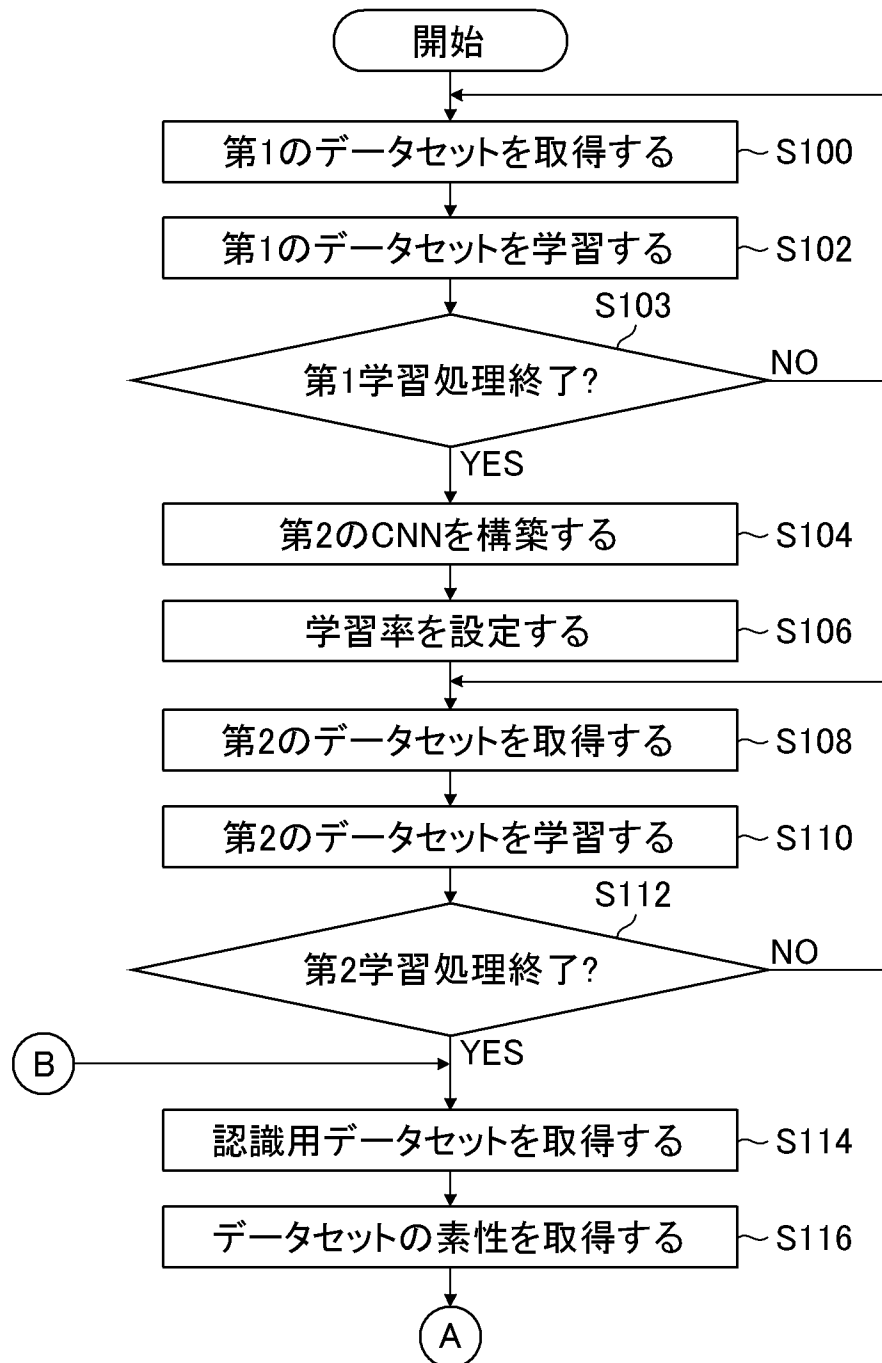
[図4]



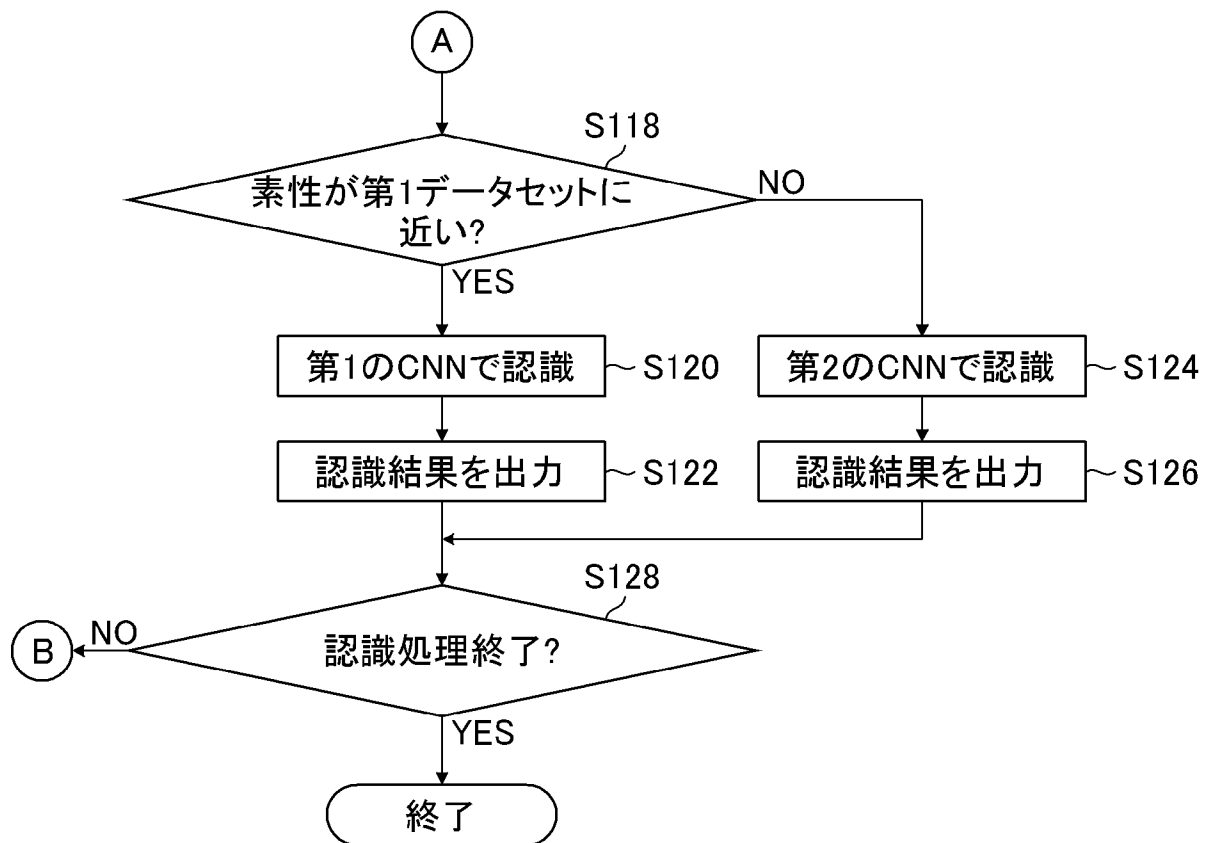
[図5]



[図6]

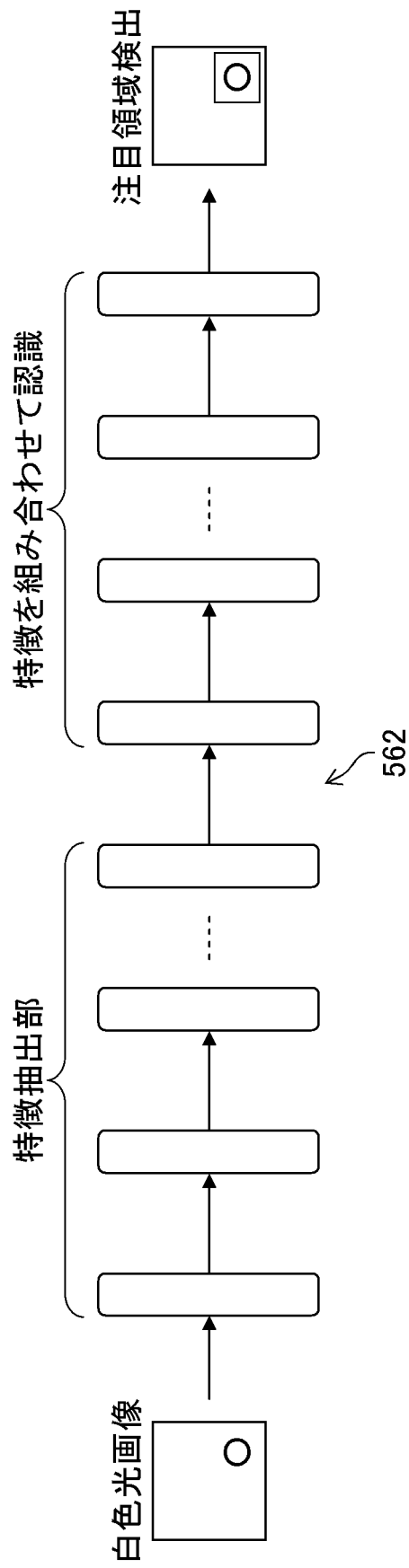


[図7]

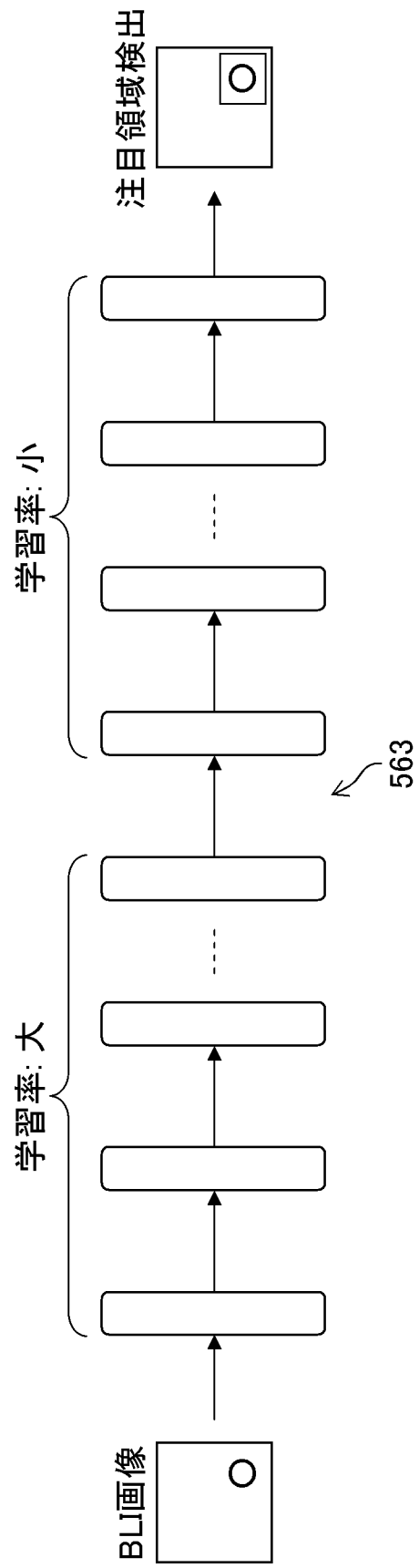


[図8]

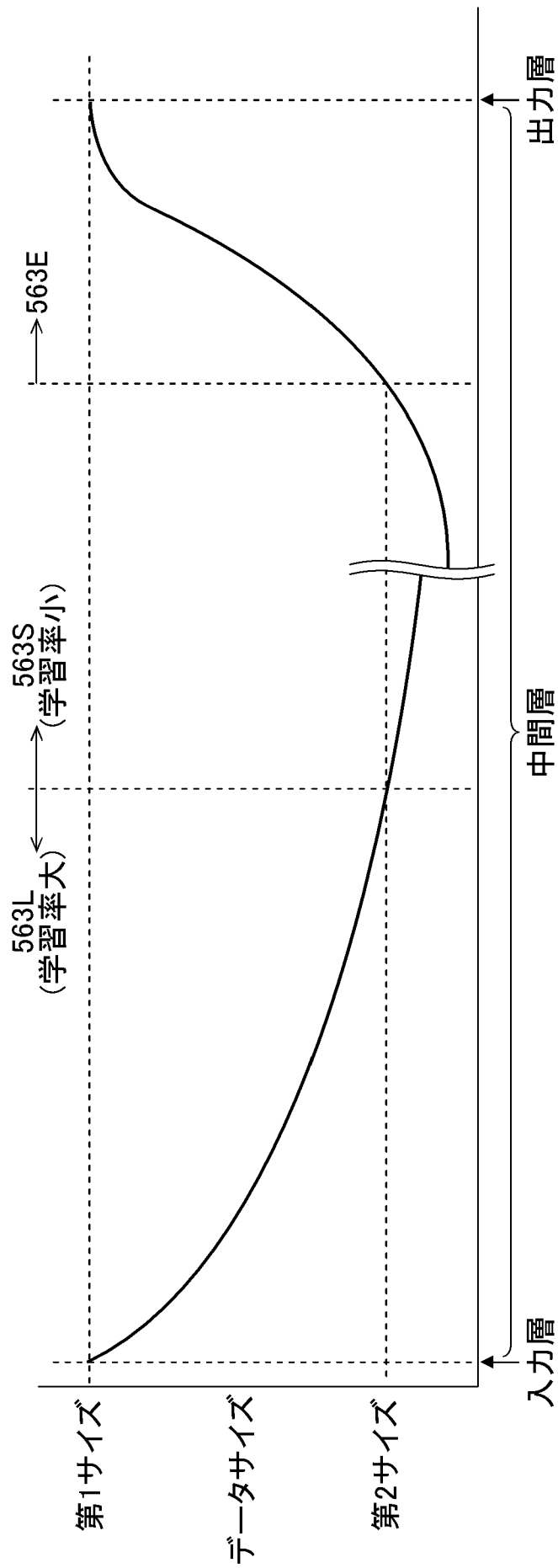
(a)



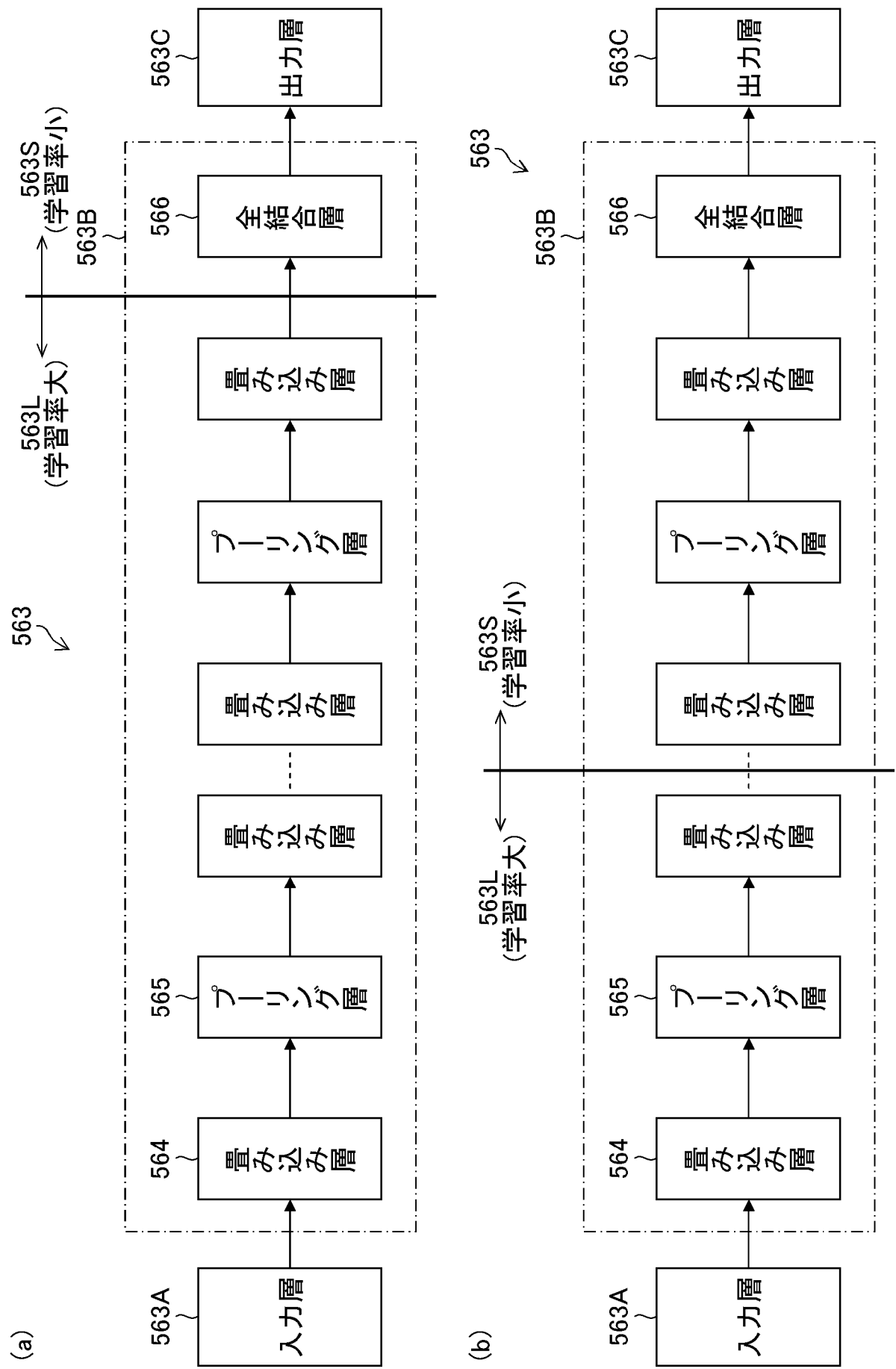
(b)



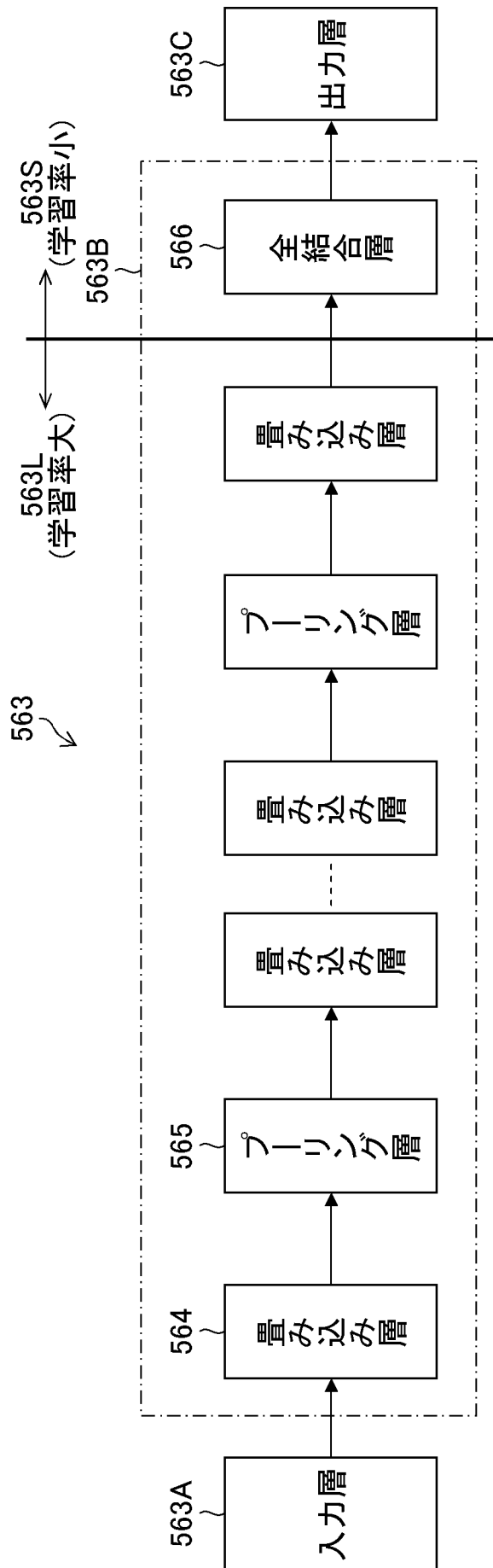
[図9]



[図10]



[図11]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2019/030255

A. CLASSIFICATION OF SUBJECT MATTER

Int.Cl. G06T7/00 (2017.01) i, G06N3/08 (2006.01) i

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)

Int.Cl. G06T7/00, G06N3/08, G06T1/00, G06K9/00

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	1971-2019
Registered utility model specifications of Japan	1996-2019
Published registered utility model applications of Japan	1994-2019

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.
A	WO 2017/175282 A1 (OLYMPUS CORP.) 12 October 2017, paragraphs [0014]-[0038], [0071]-[0073] & US 2019/0034800 A1, paragraphs [0020]-[0050], [0087]-[0090]	1-16
A	JP 2002-042107 A (FUJI ELECTRIC CO., LTD.) 08 February 2002, paragraphs [0021]-[0035], [0045]-[0046] (Family: none)	1-16
A	US 2017/0032222 A1 (XEROX CORPORATION) 02 February 2017, paragraphs [0027]-[0048] (Family: none)	1-16

☐ 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

"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

"I" 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
16 October 2019 (16.10.2019)

Date of mailing of the international search report
29 October 2019 (29.10.2019)

Name and mailing address of the ISA/
Japan Patent Office
3-4-3, Kasumigaseki, Chiyoda-ku,
Tokyo 100-8915, Japan

Authorized officer

Telephone No.

A. 発明の属する分野の分類（国際特許分類（IPC））

Int.Cl. G06T7/00(2017.01)i, G06N3/08(2006.01)i

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））

Int.Cl. G06T7/00, G06N3/08, G06T1/00, G06K9/00

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	1971-2019年
日本国実用新案登録公報	1996-2019年
日本国登録実用新案公報	1994-2019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
A	WO 2017/175282 A1（オリンパス株式会社）2017.10.12, [0014] - [0038], [0071] - [0073] & US 2019/0034800 A1, [0020] - [0050], [0087] - [0090]	1-16
A	JP 2002-042107 A（富士電機株式会社）2002.02.08, [0021] - [0035], [0045] - [0046]（ファミリーなし）	1-16
A	US 2017/0032222 A1（XEROX CORPORATION）2017.02.02, [0027] - [0048]（ファミリーなし）	1-16

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

「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矛盾するものではなく、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

16.10.2019

国際調査報告の発送日

29.10.2019

国際調査機関の名称及びあて先

日本国特許庁（ISA/J P）

郵便番号100-8915

東京都千代田区霞が関三丁目4番3号

特許庁審査官（権限のある職員）

松浦 功

5H

9181

電話番号 03-3581-1101 内線 3531