

(19) 日本国特許庁(JP)

再公表特許(A1)

(11) 国際公開番号

W02017/175282

発行日 平成31年2月28日(2019.2.28)

(43) 国際公開日 平成29年10月12日(2017.10.12)

(51) Int.Cl.	F I	テーマコード (参考)
G06T 7/00 (2017.01)	G06T 7/00 350C	4C117
A61B 5/00 (2006.01)	A61B 5/00 G	4C161
A61B 1/00 (2006.01)	A61B 1/00 C	5L096
A61B 1/045 (2006.01)	A61B 1/045 614	
G06N 3/08 (2006.01)	G06N 3/08 140	

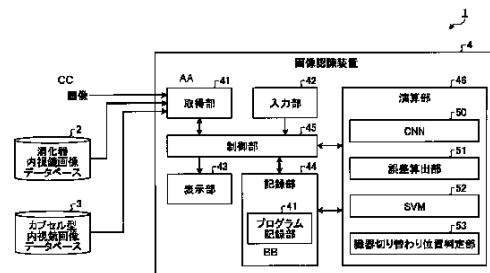
審査請求 未請求 予備審査請求 未請求 (全 22 頁)

出願番号	特願2018-510033 (P2018-510033)	(71) 出願人	000000376
(21) 国際出願番号	PCT/JP2016/061045		オリンパス株式会社
(22) 国際出願日	平成28年4月4日(2016.4.4)		東京都八王子市石川町2951番地
(81) 指定国	AP (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), EA (AM, AZ, BY, KG, KZ, RU, TJ, TM), EP (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OA (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG), AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US	(74) 代理人	110002147 特許業務法人酒井国際特許事務所
		(72) 発明者	白谷 文行 東京都八王子市石川町2951番地 オリンパス株式会社内
		Fターム(参考)	4C117 XA04 XB09 XD27 XE34 XE43 XJ01 XK33 XK42 XL12 4C161 DD07 FF14 SS21 TT15 YY07 5L096 AA06 BA06 BA13 CA02 DA02 EA39 GA51 HA09 HA11 JA11 KA04

(54) 【発明の名称】 学習方法、画像認識装置およびプログラム

(57) 【要約】

学習用画像が不足する認識対象に対して、高精度な画像認識を実現することができる学習方法、画像認識装置およびプログラムを提供する。画像認識装置4は、外部から入力された第1フレームレートによって撮像された第1画像群を用いて前記コンボリューショナルニューラルネットワークによる第1の学習を行い、第1の学習を行った後に、外部から入力された前記第1フレームレートよりも低い第2フレームレートで撮像された第2画像群であって、第1画像群より枚数が少ない第2画像群を用いて第2の学習を行う。



2 Digestive organ endoscope image database
3 Capsule type endoscope image database
4 Image recognition device
42 Input unit
43 Display unit
44 Recording unit
45 Control unit
46 Calculation unit
51 Error calculation unit
52 Organ switching position determination unit
AA Acquisition unit
BB Program recording unit
CC Image

【特許請求の範囲】**【請求項 1】**

コンボリユーショナルニューラルネットワークを含む画像認識装置が行う学習方法であって、

外部から入力された第 1 フレームレートによって撮像された第 1 画像群を用いて前記コンボリユーショナルニューラルネットワークによる第 1 の学習を行い、

前記第 1 の学習を行った後に、外部から入力された前記第 1 フレームレートよりも低い第 2 フレームレートで撮像された第 2 画像群であって、前記第 1 画像群より枚数が少ない第 2 画像群を用いて第 2 の学習を行う

ことを特徴とする学習方法。

10

【請求項 2】

前記第 1 の学習は、

前記コンボリユーショナルニューラルネットワークの出力と前記第 1 画像群に付与される正解ラベルとの誤差を小さくする誤差逆伝搬法によって前記コンボリユーショナルニューラルネットワークの重みを更新するプレトレーニングであり、

前記第 2 の学習は、

前記プレトレーニングによって前記重みが更新された前記コンボリユーショナルニューラルネットワークを用いて前記重みを更新するファインチューニングである

ことを特徴とする請求項 1 に記載の学習方法。

20

【請求項 3】

前記ファインチューニングによって前記重みが更新された前記コンボリユーショナルニューラルネットワークの中間出力を前記第 2 画像群の特徴量として抽出し、前記第 2 画像群の正解ラベルおよび前記特徴量を用いて学習器による入力画像の識別または認識対象物の検出を学習する第 3 の学習を行う

ことを特徴とする請求項 2 に記載の学習方法。

【請求項 4】

前記第 3 の学習を行った後に、前記第 2 画像群を用いて前記学習器による入力画像に対する分類または認識対象物の検出を行う

ことを特徴とする請求項 3 に記載の学習方法。

【請求項 5】

前記第 1 の学習は、

前記コンボリユーショナルニューラルネットワークの出力と前記第 1 画像群に含まれる正解ラベルとの誤差を小さくする誤差逆伝搬法によって前記コンボリユーショナルニューラルネットワークの重みを更新するプレトレーニングであり、

前記第 2 の学習は、

前記プレトレーニングによって前記コンボリユーショナルニューラルネットワークの重みが更新された前記コンボリユーショナルニューラルネットワークを用いて、前記第 2 画像群の特徴量を抽出し、前記第 2 画像群の正解ラベルおよび前記特徴量を用いて学習器による入力画像の識別または認識対象物の学習を行う

ことを特徴とする請求項 1 に記載の学習方法。

40

【請求項 6】

前記第 2 の学習を行った後に、前記第 2 画像群を用いて前記学習器による前記入力画像の分類または認識対象物の検出を行う

ことを特徴とする請求項 5 に記載の学習方法。

【請求項 7】

前記第 2 画像群は、インターバル撮影された画像群であることを特徴とする請求項 1 ～ 6 のいずれか一つに記載の学習方法。

【請求項 8】

前記第 1 画像群は、被検体に挿入可能な内視鏡によって撮像された画像群であり、

前記第 2 画像群は、被検体に飲み込まれて所定時間後に排出されるカプセル型内視鏡に

50

よって撮像された画像群であることを特徴とする請求項 1～7 のいずれか一つに記載の学習方法。

【請求項 9】

コンボリューショナルニューラルネットワークを含む画像認識装置であって、

第 1 フレームレートによって撮像された第 1 画像群および前記第 1 フレームレートよりも低い第 2 フレームレートで撮像された第 2 画像群であって、前記第 1 画像群より枚数が少ない第 2 画像群それぞれを取得する取得部と、

前記第 1 画像群が前記コンボリューショナルニューラルネットワークに入力され、前記コンボリューショナルニューラルネットワークの出力結果と前記第 1 画像群に付与された正解ラベルとの第 1 誤差を算出する誤差算出部と、

前記第 1 誤差が前記コンボリューショナルニューラルネットワークに逆伝搬することによって、前記コンボリューショナルニューラルネットワークの重みを更新する制御部と、
を備え、

前記誤差算出部は、前記重みが更新された後に、第 2 画像群が前記コンボリューショナルニューラルネットワークに入力され、前記コンボリューショナルニューラルネットワークの出力結果と前記第 2 画像群に付与された正解ラベルとの第 2 誤差を算出し、

前記制御部は、前記第 2 誤差が前記コンボリューショナルニューラルネットワークに逆伝搬することによって、前記コンボリューショナルニューラルネットワークの重みを更新することを特徴とする画像認識装置。

【請求項 10】

前記重みが更新された後に、前記コンボリューショナルニューラルネットワークによって前記第 2 画像群の特徴量として抽出された前記コンボリューショナルニューラルネットワークの中間出力と前記第 2 画像群の正解ラベルとを用いて入力画像の識別または認識対象物の検出を学習する学習器をさらに備えたことを特徴とする請求項 9 に記載の画像認識装置。

【請求項 11】

前記学習器は、前記第 2 画像群を学習した後に、前記第 2 画像群を用いて入力画像に対する分類または認識対象物の検出を行う

ことを特徴とする請求項 10 に記載の画像認識装置。

【請求項 12】

前記学習器は、サポートベクターマシンであることを特徴とする請求項 11 に記載の画像認識装置。

【請求項 13】

コンボリューショナルニューラルネットワークを含む画像認識装置であって、

第 1 フレームレートによって撮像された第 1 画像群および前記第 1 フレームレートよりも低い第 2 フレームレートで撮像された第 2 画像群であって、前記第 1 画像群より枚数が少ない第 2 画像群それぞれを取得する取得部と、

前記第 1 画像群が前記コンボリューショナルニューラルネットワークに入力され、前記コンボリューショナルニューラルネットワークの出力結果と前記第 1 画像群に付与された正解ラベルとの第 1 誤差を算出する誤差算出部と、

前記第 1 誤差を前記コンボリューショナルニューラルネットワークに逆伝搬することによって、前記コンボリューショナルニューラルネットワークの重みを更新する制御部と、

前記第 2 画像群を用いて入力画像の識別または画像中からの認識対象物の検出を学習する学習器と、

を備え、

前記学習器は、前記重みが更新された後に、前記コンボリューショナルニューラルネットワークによって前記第 2 画像群の特徴量として抽出された前記コンボリューショナルニューラルネットワークの中間出力と前記第 2 画像群の正解ラベルとを用いて前記入力画像の識別または画像中からの認識対象物の検出を学習することを特徴とする画像認識装置。

【請求項 14】

前記学習器は、サポートベクターマシンであることを特徴とする請求項 1 3 に記載の画像認識装置。

【請求項 1 5】

コンボリューショナルニューラルネットワークを含む画像認識装置に、

外部から入力された第 1 フレームレートによって撮像された第 1 画像群を用いて前記コンボリューショナルニューラルネットワークによる第 1 の学習を行う第 1 の学習手順と、

前記第 1 の学習を行った後に、外部から入力された前記第 1 フレームレートよりも低い第 2 フレームレートで撮像された第 2 画像群であって、前記第 1 画像群より枚数が少ない第 2 画像群を用いて第 2 の学習を行う第 2 の学習手順と、

を実行させることを特徴とするプログラム。

10

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、画像群を学習して認識するコンボリューショナルニューラルネットワークを含む学習方法、画像認識装置およびプログラムに関する。

【背景技術】

【0002】

近年、医療装置の分野において、被検体が撮像された画像群をコンピュータなどにより機械を用いて学習することによって、画像群から自動的に臓器を分類したり、病変を検出したりするコンピュータ診断支援技術が知られている。例えば、CT 装置が被検体を撮像した画像群から関連する画像同士の位置合わせや重ね合わせを行うために、基準点として用いられるランドマークの最も確からしい位置を自動的に特定する技術が知られている（特許文献 1 参照）。この技術では、ランドマークの位置が既知である第 1 の医用画像を用いて分類器に学習（トレーニング）させた後に、分類器にランドマークの位置が未知である第 2 の医用画像に対してランドマークの位置を特定させる。

20

【先行技術文献】

【特許文献】

【0003】

【特許文献 1】特開 2015-36123 号公報

【発明の概要】

30

【発明が解決しようとする課題】

【0004】

ところで、機械学習を用いた診断支援技術の一例として、ポリープや腫瘍等の病変を、機械学習する際には、豊富な数の学習用画像が必要となる。しかしながら、病変が写っている画像の出現頻度は、一般的に低いため、学習用画像が不足するという問題点があった。上述した特許文献 1 には、医療画像の第 1 領域より狭い第 2 領域を用いて分類器を調整することによってランドマーク位置の特定精度を向上させる方法が提示されているが、学習用画像が不足する問題に対する解決策については記載されていない。

【0005】

また、例えばカプセル型内視鏡のようにインターバル撮影された画像群の場合、胃の幽門や大腸の回盲弁等のランドマークが写っている画像と写っていない画像とが存在する。そのため、ランドマークの学習用画像が不足するという問題点があった。

40

【0006】

本発明は、上記に鑑みてなされたものであって、学習用画像が不足する認識対象に対して、高精度の画像認識を実現することができる学習方法、画像認識装置およびプログラム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】

【0007】

上述した課題を解決し、目的を達成するために、本発明に係る学習方法は、コンボリューショナルニューラルネットワークを含む画像認識装置が行う学習方法であって、外部か

50

ら入力された第1フレームレートによって撮像された第1画像群を用いて前記コンボリユーショナルニューラルネットワークによる第1の学習を行い、前記第1の学習を行った後に、外部から入力された前記第1フレームレートよりも低い第2フレームレートで撮像された第2画像群であって、前記第1画像群より枚数が少ない第2画像群を用いて第2の学習を行うことを特徴とする。

【0008】

また、本発明に係る画像認識装置は、コンボリユーショナルニューラルネットワークを含む画像認識装置であって、第1フレームレートによって撮像された第1画像群および前記第1フレームレートよりも低い第2フレームレートで撮像された第2画像群であって、前記第1画像群より枚数が少ない第2画像群それぞれを取得する取得部と、前記第1画像群が前記コンボリユーショナルニューラルネットワークに入力され、コンボリユーショナルニューラルネットワークの出力結果と前記第1画像群に付与された正解ラベルとの第1誤差を算出する誤差算出部と、前記第1誤差がコンボリユーショナルニューラルネットワークに逆伝搬することによって、前記コンボリユーショナルニューラルネットワークの重みを更新する制御部と、を備え、前記誤差算出部は、前記重みが更新された後に、第2画像群がコンボリユーショナルニューラルネットワークに入力され、コンボリユーショナルニューラルネットワークの出力結果と前記第2画像群に付与された正解ラベルとの第2誤差を算出し、前記制御部は、前記第2誤差がコンボリユーショナルニューラルネットワークに逆伝搬することによって、前記コンボリユーショナルニューラルネットワークの重みを更新することを特徴とする。

【0009】

また、本発明に係る画像認識装置は、コンボリユーショナルニューラルネットワークを含む画像認識装置であって、第1フレームレートによって撮像された第1画像群および前記第1フレームレートよりも低い第2フレームレートで撮像された第2画像群であって、前記第1画像群より枚数が少ない第2画像群それぞれを取得する取得部と、前記第1画像群が前記コンボリユーショナルニューラルネットワークに入力され、前記コンボリユーショナルニューラルネットワークの出力結果と前記第1画像群に付与された正解ラベルとの第1誤差を算出する誤差算出部と、前記第1誤差を前記コンボリユーショナルニューラルネットワークに逆伝搬することによって、前記コンボリユーショナルニューラルネットワークの重みを更新する制御部と、前記第2画像群を用いて入力画像の識別または画像中からの認識対象物の検出を学習する学習器と、を備え、前記学習器は、前記重みが更新された後に、前記コンボリユーショナルニューラルネットワークによって前記第2画像群の特徴量として抽出された前記コンボリユーショナルニューラルネットワークの中間出力と前記第2画像群の正解ラベルとを用いて前記入力画像の識別または画像中からの認識対象物の検出を学習することを特徴とする。

【0010】

また、本発明に係るプログラムは、コンボリユーショナルニューラルネットワークを含む画像認識装置に、外部から入力された第1フレームレートによって撮像された第1画像群を用いて前記コンボリユーショナルニューラルネットワークによる第1の学習を行う第1の学習手順と、前記第1の学習を行った後に、外部から入力された前記第1フレームレートよりも低い第2フレームレートで撮像された第2画像群であって、前記第1画像群より枚数が少ない第2画像群を用いて第2の学習を行う第2の学習手順と、を実行させることを特徴とする。

【発明の効果】

【0011】

本発明によれば、学習用画像が不足する認識対象に対して、高精度の画像認識を実現することができるという効果を奏する。

【図面の簡単な説明】

【0012】

【図1】図1は、本発明の実施の形態1に係る学習システムの構成を示すブロック図であ

る。

【図 2】図 2 は、本発明の実施の形態 1 に係る画像認識装置における演算部の概略構成を示す模式図である。

【図 3】図 3 は、本発明の実施の形態 1 に係る画像認識装置が実行する学習処理の概要を示すフローチャートである。

【図 4】図 4 は、本発明の実施の形態 1 に係る画像認識装置が実行する認識処理の概要を示すフローチャートである。

【図 5】図 5 は、本発明の実施の形態 1 に係る画像認識装置が実行するランドマーク処理において用いるステップ関数の一例を模式的に示す図である。

【図 6】図 6 は、本発明の実施の形態 1 に係る画像認識装置によるランドマーク処理において胃と小腸の時系列画像群を自動分類した一例を示す図である。

【図 7】図 7 は、本発明の実施の形態 1 に係る画像認識装置が実行するランドマーク処理の概要を示すフローチャートである。

【図 8】図 8 は、本発明の実施の形態 1 の変形例 1 に係る画像認識装置が実行する学習処理の概要を示すフローチャートである。

【図 9】図 9 は、本発明の実施の形態 1 の変形例 2 に係る画像認識装置が実行する学習処理の概要を示すフローチャートである。

【図 10】図 10 は、本発明の実施の形態 2 に係る画像認識装置が実行する学習処理の概要を示すフローチャートである。

【発明を実施するための形態】

【0013】

以下、本発明の実施の形態に係る学習方法、画像認識装置およびプログラムについて、図面を参照しながら説明する。なお、これらの実施の形態によって本発明が限定されるものではない。また、各図面の記載において、同一の部分には同一の符号を付して説明する。

【0014】

(実施の形態 1)

〔学習システムの構成〕

図 1 は、本発明の実施の形態 1 に係る学習システムの構成を示すブロック図である。本実施の形態 1 に係る学習システム 1 は、一例として、被検体に挿入することによって被検体の体内を撮像する消化器内視鏡（軟性内視鏡や硬性内視鏡等の内視鏡スコープ）および被写体に飲み込まれて所定時間後に排出されるカプセル型内視鏡によって、被検体の管腔を撮像することにより取得された管腔内画像に基づいて学習を行う学習処理、認識を行う認識処理、胃の幽門や大腸の回盲弁等の臓器区分を表すランドマークの位置や臓器区分を分類するランドマーク処理を実行するシステムである。

【0015】

図 1 に示す学習システム 1 は、消化器内視鏡によって被検体内を撮像した画像データ群およびランドマークを認識するための正解ラベルを記録する消化器内視鏡画像データベース 2 と、カプセル型内視鏡によって被検体内を撮像した画像データ群およびランドマークを認識するための正解ラベルを記録するカプセル型内視鏡画像データベース 3 と、コンボリユーショナルニューラルネットワーク（Convolutional Neural Network）を含み、消化器内視鏡画像データベース 2、カプセル型内視鏡画像データベース 3 および外部から管腔内画像のいずれかから画像データを取得して学習または認識する画像認識装置 4 と、を備える。なお、消化器内視鏡画像データベース 2、カプセル型内視鏡画像データベース 3 および画像認識装置 4 は、ネットワークを介して双方向に通信可能に接続されてもよい。この場合には、無線および有線のいずれかによって接続されればよい。

【0016】

〔消化器内視鏡画像データベース〕

消化器内視鏡画像データベース 2 は、ハードディスク等の記録媒体を用いて構成される。消化器内視鏡画像データベース 2 は、被検体に挿入可能な消化器内視鏡が第 1 フレーム

レートによって撮像した画像データ群（以下、「第1画像群」という）を記録するとともに、ランドマークを認識するための正解ラベル（以下、「第1正解ラベル」という）を記録する。ここで、第1フレームレートとは、例えば60fpsまたは30fpsのいずれかである。また、第1正解ラベルとは、画像認識装置4にランドマークを学習させるために消化器内視鏡によって撮像された画像群からランドマークが写る画像とそうでない画像を予め選択し、それぞれに付与した識別子（例えば：+1、-1）のことである。

【0017】

〔カプセル型内視鏡画像データベース〕

カプセル型内視鏡画像データベース3は、ハードディスク等の記録媒体を用いて構成される。カプセル型内視鏡画像データベース3は、カプセル型内視鏡が第2フレームレートによって撮像した画像データ群（以下、「第2画像群」という）であって、第1画像群より撮影枚数が少ない第2画像群を記録するとともに、ランドマークを認識するための正解ラベル（以下、「第2正解ラベル」という）を記録する。ここで、第2フレームレートとは、例えば2fpsである。また、第2画像群は、インターバル撮影された画像群である。さらに、第2正解ラベルとは、画像認識装置4にランドマークを学習させるためにカプセル型内視鏡によって撮像された画像群からランドマークが写る画像とそうでない画像を予め選択し、それぞれに付与した識別子（例えば：+1と-1）のことである。さらに、第1画像群より撮影枚数が少ないとは、例えば第1画像群が50万枚に対して、第2画像群が5万枚であり、あくまでも一例にすぎない。

【0018】

〔画像認識装置の詳細な構成〕

画像認識装置4は、消化器内視鏡画像データベース2、カプセル型内視鏡画像データベース3および外部のいずれかから画像群を取得する取得部41と、外部からの操作によって入力された入力信号を受け付ける入力部42と、画像や各種情報を表示する表示部43と、取得部41によって取得された画像群や各種のプログラムを記録する記録部44と、画像認識装置4全体の動作を制御する制御部45と、画像群に対して学習および認識を行う演算部46と、を備える。

【0019】

取得部41は、画像群の受け渡しに可搬型の記録媒体が使用される場合、この記録媒体を着脱自在に装着し、記録された画像群を読み出すリーダ装置として構成される。また、取得部41は、画像群を記録するサーバまたはデータベースを用いる場合、このサーバまたはデータベースとネットワークを介して通信可能な通信装置等で構成され、サーバまたはデータベースとデータ通信を行うことによって画像群を取得する。さらにまた、取得部41は、消化器内視鏡からケーブルを介して画像群が入力されるインターフェース装置等で構成してもよい。

【0020】

入力部42は、例えばキーボード、マウス、タッチパネルおよび各種スイッチ等の入力デバイスによって実現される。入力部42は、外部からの操作に応じて受け付けた入力信号を制御部45へ出力する。

【0021】

表示部43は、液晶や有機EL（Electro Luminescence）等の表示パネルによって実現される。表示部43は、制御部45の制御のもと、画像を含む各種画面を表示する。

【0022】

記録部44は、フラッシュメモリ、ROM（Read Only Memory）およびRAM（Random Access Memory）といった各種メモリ、および内蔵若しくはデータ通信端子で接続されたハードディスク等によって実現される。記録部44は、取得部41によって取得された画像データその他、画像認識装置4を動作させるとともに、種々の機能を画像認識装置4に実行させるためのプログラム、このプログラムの実行中に使用されるデータ等を記録する。例えば、記録部44は、取得部41から入力された画像群を学習するプログラムを記録するプログラム記録部441を記録する。

【0023】

制御部45は、CPU (Central Processing Unit) 等を用いて実現される。制御部45は、記録部44に記録された各種プログラムを読み込むことによって、取得部41から入力される画像群や入力部42から入力される入力信号等に従って、画像認識装置4を構成する各部への指示や転送等を行い、画像認識装置4全体の動作を統括的に制御する。また、制御部45は、後述する誤差算出部51が算出した第1誤差をコンボリユーションニューラルネットワーク50 (以下、「CNN50」という) に逆伝搬することによって、CNN50のパラメータの重みを更新する。さらに、制御部45は、後述する誤差算出部51が算出した第2誤差をCNN50に逆伝搬することによって、CNN50のパラメータの重みを更新する。

10

【0024】

演算部46は、CPU等によって実現される。演算部46は、プログラム記録部441が記録するプログラムを読み込むことによって画像群に対する学習処理、認識処理およびランドマーク処理を実行する。

【0025】

〔演算部の詳細な構成〕

次に、演算部46の詳細な構成について説明する。

図2は、演算部46の概略構成を示す模式図である。図1および図2に示す演算部46は、CNN50と、誤差算出部51と、サポートベクターマシン (Support Vector Machine) 52 (以下、「SVM52」という) と、臓器切り替わり位置判定部53と、を有する。

20

【0026】

CNN50は、取得部41が消化器内視鏡画像データベース2から取得した第1画像群を用いて学習した後に、取得部41がカプセル型内視鏡画像データベース3から取得した第2画像群を用いて学習する。具体的には、CNN50は、第1画像群と第1正解ラベルとを用いて、第1画像群と第1正解ラベルとの誤差を小さくする誤差逆伝搬法によってパラメータの重みを更新する第1の学習としてのプレトレーニング (Pre-training) を行った後に、第2画像群と第2正解ラベルとを用いて、第2画像群と第2正解ラベルとの誤差を小さくする誤差逆伝搬法によってパラメータを更新する第2の学習としてのファインチューニング (Fine-tuning) を行う。

30

【0027】

CNN50は、コンボリユーション層501と、プーリング層502 (サブサンプリング層) と、全結合層503と、を有する。本実施の形態1において、CNN50は、コンボリユーション層501とプーリング層502とを3回繰り返す構成によって説明するが、少なくともコンボリユーション層501とプーリング層502とを1回以上繰り返す構成であればよく、例えばコンボリユーション層501とプーリング層502とを4回繰り返す構成であってもよい。また、本実施の形態1において、CNN50は、全結合層503を2回繰り返す構成によって説明するが、少なくとも全結合層503を1回以上繰り返す構成であればよく、例えば全結合層503を3回繰り返す構成であってもよい。

【0028】

誤差算出部51は、第1画像群がCNN50に入力され、CNN50の出力結果と第1正解ラベルとの第1誤差を算出する。また、誤差算出部51は、CNN50の重みが更新された後に、第2画像群がCNN50に入力され、CNN50の出力結果と第2画像群に付与された第2正解ラベルとの第2誤差を算出する。

40

【0029】

SVM52は、CNN50のパラメータの重みが更新された後に、CNN50によって第2画像群の特徴量として抽出されたCNN50の中間出力と第2画像群の正解ラベルとを用いて入力画像の識別または認識対象物の検出を学習する。さらに、SVM52は、第2画像群を学習した後に、第2画像群を用いて入力画像に対する分類または認識対象物の検出を行う。ここで、入力画像は、例えば胃や大腸等の管腔臓器の内壁を撮影した臓器画

50

像である。認識対象物は、例えば管腔臓器の内壁表面に存在するポリープ等の病変部である。なお、本実施の形態1では、SVM52が学習器として機能するが、これに限定することなく、他の学習器を組み合わせてもよい。

【0030】

臓器切り替わり位置判定部53は、SVM52によって臓器の区分が判定（分類）された判定結果と後述するステップ関数との類似度とによって、臓器の切り替わり位置、例えば、胃と小腸との切り替わり位置の候補を抽出（決定）し、幽門や回盲弁などの臓器の切替りの正解な位置を決定する。

【0031】

〔学習処理〕

次に、画像認識装置4が実行する学習処理について説明する。図3は、画像認識装置4が実行する学習処理の概要を示すフローチャートである。

【0032】

図3に示すように、まず、制御部45は、CNN50の重みを乱数で初期化する（ステップS101）。

【0033】

続いて、制御部45は、取得部41を介して消化器内視鏡画像データベース2から消化器内視鏡画像をCNN50に入力する（ステップS102）。

【0034】

続いて、制御部45は、CNN50の出力と消化器内視鏡画像データベース2からの正解ラベル（+1，-1等）との誤差を小さくする誤差逆伝搬法によってCNN50の重みを更新する（ステップS103）。具体的には、図2に示すように、CNN50は、消化器内視鏡画像に対して、コンボリューション層501によるコンボリューション処理およびプーリング層502によるプーリング処理それぞれを順伝搬によって3回行い、全結合層503による全結合処理を順伝搬によって2回行った後、この出力と第1正解ラベルとの第1誤差を誤差算出部51が算出する。そして、制御部45は、誤差算出部51が算出した第1誤差を全結合層503、プーリング層502およびコンボリューション層501の順で逆伝搬することによって、CNN50パラメータの重みを更新するプレトレーニング（第1の学習）を行う。

【0035】

その後、CNN50によるプレトレーニングの学習枚数が所定の閾値Th1を超えた場合（ステップS104：Yes）、画像認識装置4は、ステップS105へ移行する。これに対して、CNN50によるプレトレーニングの学習枚数が所定の閾値Th1を超えていない場合（ステップS104：No）、画像認識装置4は、上述したステップS102へ戻る。

【0036】

ステップS105において、制御部45は、取得部41を介してカプセル型内視鏡画像データベース3からカプセル型内視鏡画像をCNN50に入力する。

【0037】

続いて、制御部45は、CNN50の出力とカプセル型内視鏡画像データベース3からの正解ラベル（+1，-1等）との誤差を小さくする誤差逆伝搬法によってCNN50の重みを更新する（ステップS106）。具体的には、図2に示すように、CNN50は、カプセル型内視鏡画像に対して、コンボリューション層501によるコンボリューション処理およびプーリング層502によるプーリング処理それぞれを順伝搬によって3回行い、全結合層503による全結合処理を順伝搬によって2回行った後、この出力と第2正解ラベルとの第2誤差を誤差算出部51が算出する。そして、制御部45は、誤差算出部51が算出した第2誤差を全結合層503、プーリング層502およびコンボリューション層501の順で逆伝搬することによって、CNN50の重みを更新するファインチューニング（第2の学習）を行う。

【0038】

10

20

30

40

50

その後、CNN50によるファインチューニングの学習枚数が所定の閾値Th2を超えた場合（ステップS107：Yes）、画像認識装置4は、ステップS108へ移行する。これに対して、CNN50によるファインチューニングの学習枚数が所定の閾値Th2を超えていない場合（ステップS107：No）、画像認識装置4は、上述したステップS105へ戻る。

【0039】

ステップS108において、制御部45は、取得部41を介してカプセル型内視鏡画像データベース3からカプセル型内視鏡画像をCNN50に入力する。

【0040】

続いて、CNN50は、全結合層503の手前の中間出力を第2画像群のCNN特徴量として抽出する（ステップS109）。具体的には、CNN50は、取得部41を介して入力されたカプセル型内視鏡画像に対して、コンボリューション層501によるコンボリューション処理およびプーリング層502によるプーリング処理それぞれを順伝搬によって3回行い、この結果を第2画像群のCNN特徴量として抽出する。

【0041】

その後、SVM52は、CNN50から入力されたCNN特徴量とカプセル型内視鏡画像データベース3から入力された第2正解ラベルとに基づいて、正例と負例との判別境界のマージンが最大となるようサポートベクター係数を決定する（ステップS110）。ステップS110の後、画像認識装置4は、本処理を終了する。

【0042】

このように画像認識装置4は、CNN50が消化器内視鏡画像群と第1正解ラベルとを用いて、消化器内視鏡画像群と第1正解ラベルとの誤差を小さくする誤差逆伝搬法によってパラメータの重みを更新するプレトレーニング（第1の学習）を行った後、CNN50がカプセル型内視鏡画像群と第2正解ラベルとを用いて、カプセル型内視鏡画像群と第2正解ラベルとの誤差を小さくする誤差逆伝搬法によってパラメータを更新するファインチューニング（第2の学習）を行う。その後、CNN50は、CNN50の中間出力をカプセル型内視鏡画像群のCNN特徴量として抽出する。続いて、SVM52は、CNN50によって抽出されたCNN特徴量と第2正解ラベルとを用いてカプセル型内視鏡画像群に対する入力画像の識別または認識対象物の検出を学習する（第3の学習）。そして、SVM52は、取得部41から入力される未知のカプセル型内視鏡画像群に対して分類または認識対象物の検出を行う。これにより、学習用画像が不足する認識対象に対して、高精度の画像認識を実現することができる。

【0043】

〔認識処理〕

次に、学習済みのCNN50およびSVM52を用いて、画像認識装置4が未知の画像を認識する認識処理について説明する。図4は、画像認識装置4が実行する認識処理の概要を示すフローチャートである。なお、以下においては、未知の画像として、カプセル型内視鏡装置によって撮像されたカプセル型内視鏡画像を例に説明する。

【0044】

図4に示すように、まず、制御部45は、取得部41を介して外部から未知のカプセル型内視鏡画像をCNN50に入力する（ステップS201）。

【0045】

続いて、CNN50は、全結合層503の手前の出力をCNN特徴量として抽出する（ステップS202）。具体的には、CNN50は、取得部41を介して入力されたカプセル型内視鏡画像に対して、コンボリューション層501によるコンボリューション処理およびプーリング層502によるプーリング処理それぞれを順伝搬によって3回行い、この結果をCNN特徴量として抽出する。

【0046】

その後、SVM52は、CNN50からCNN特徴量を入力し、分類結果（胃、小腸および大腸のいずれか）を出力する（ステップS203）。例えば、SVM52は、未知の

カプセル型内視鏡画像が胃であると出力する。ステップ S 2 0 3 の後、画像認識装置 4 は、本処理を終了する。

【 0 0 4 7 】

このように画像認識装置 4 は、学習済みの C N N 5 0 および学習済みの S V M 5 2 を用いて、未知のカプセル型内視鏡画像を識別して分類することができる。

【 0 0 4 8 】

〔ランドマーク処理〕

次に、画像認識装置 4 が実行する時間的に連続する時系列画像の臓器分類結果系列に対して、臓器の切り替え位置を推定するランドマーク処理について説明する。以下においては、画像の 2 分類として、胃と小腸の画像を自動的に分類（判定）する例を説明する。学習済みの S V M 5 2 は、時間的に連続するカプセル型内視鏡画像の対象物が胃であるか小腸であるかを判定（分類）する。続いて、臓器切り替わり位置判定部 5 3 は、この判定結果とステップ関数との類似度により胃と小腸との切り替わり位置の候補を抽出（決定）する。そして、後述するように、臓器切り替わり位置判定部 5 3 は、幽門や回盲弁などの臓器の切替りの正解な位置を決定する。ここで、時間的に連続するカプセル型内視鏡画像とは、撮影された時間の順に並んだ未分類（未知のラベル）の画像であり、具体的には、所定のフレームレート（例えば 6 0 f p s や 3 0 f p s）で撮像された動画画像または所定の間隔（例えば 2 f p s や 4 f p s）で連続的にインターバル撮像された画像である。

【 0 0 4 9 】

図 5 は、画像認識装置 4 が実行するランドマーク処理において用いるステップ関数の一例を模式的に示す図である。図 6 は、胃と小腸の時系列画像群を自動分類（判定結果）した一例を示す図である。図 6 において、縦軸が S V M 5 2 の判定結果を示し、横軸が時系列画像の番号を示す。また、図 6 において、S V M 5 2 の判定結果 $S(x) = -1$ が胃を示し、 $S(x) = +1$ が小腸を表す。画像には、時系列の順に $img 0 0 0 0 2 3 \sim img 0 0 9 6 2 3$ の番号が付されおり、その時系列の順に判定結果（分類結果） $S(x)$ が並んでいる。さらに、図 6 において、線 B 1 が胃から小腸への切り替わりの正解位置である幽門を表す。さらにまた、図 6 において、画像の番号は、所定の間隔でサンプリングされたものとする。

【 0 0 5 0 】

臓器切り替わり位置判定部 5 3 は、図 5 に示すステップ関数 $H(x)$ と図 6 に示す判定結果 $S(x)$ とのコンボリューション処理（畳み込み演算処理）を行うことによって、以下の式（1）における評価値 $Conv(x)$ を求める。そして、評価値 $Conv(x)$ と式（2）とに基づいて、臓器切り替わり位置を抽出する。

【数 1】

$$Conv(x) = \frac{1}{2L} \sum_{y=-L}^{y=L} S(x-y)H(y) \quad \cdots (1)$$

$$\text{臓器切り替わり位置} = \underset{x}{\text{Arg}} \left[\underset{L=L0}{\overset{L=L1}{\text{MIN}}} \underset{L}{Conv}(x) \right] \quad \cdots (2)$$

【 0 0 5 1 】

ここで、図 5 に示すステップ関数 $H(x)$ は、 $x = -L \sim +L$ の範囲で定義された関数であり、その窓幅 L は、可変である。また、臓器切り替わり位置判定部 5 3 は、上式（1）において、窓幅 L を任意に固定し、 $2L$ で除算した値を正規化した中でコンボリューション処理による評価値 $Conv(x)$ を算出する。

【 0 0 5 2 】

また、上式（2）の $[\]$ 内において、臓器切り替わり位置判定部 5 3 は、位置 x を任意に固定し、窓幅 L を $L0$ （例えば 1 0 とする） $\sim L1$ まで動かす中での $Conv(x)$ の最小値を算出する。ここで、画像の探索開始位置を $M0$ 、画像の探索終了位置を $M1$ とす

ると、 $L1 = \min [x - M0, M1 - x]$ となる。例えば、 $M0 = 1$ 、 $M1 = 1000$ とする。

【0053】

そして、臓器切り替わり位置判定部53は、Arg関数により、最小値の位置 x を臓器切り替え位置として抽出する。なお、臓器切り替わり位置判定部53は、上述した式(2)において、Conv(x)の最小値を用いていたが、これに限定されることなく、Conv(x)の最大値を用いてもよい。

【0054】

以下において、画像認識装置4が実行するランドマーク処理の詳細な処理について説明する。図7は、画像認識装置4が実行するランドマーク処理の概要を示すフローチャートである。

10

【0055】

図7に示すように、まず、取得部41は、外部から入力されたカプセル型内視鏡画像列から疎な間隔(例えば50枚間隔)でサンプリングする(ステップS300)。

【0056】

その後、制御部45は、取得部41がカプセル型内視鏡画像列から疎な間隔でサンプリングしたそれぞれの画像を学習済みのCNN50に入力することによって、胃、小腸および大腸に分類させる(ステップS301)。

【0057】

続いて、臓器切り替わり位置判定部53は、は、上述したステップS301において、CNN50による分類結果列から、胃と小腸および小腸と大腸の切り替え概略位置を上述の算出方法によって推定する(ステップS302)。

20

【0058】

その後、取得部41は、ステップS302において臓器切り替わり位置判定部53が推定した切り替え概略位置の近傍の区間から密な間隔(例えば概略推定位置の前後各200枚全て)で外部から入力されたカプセル型内視鏡画像列をサンプリングする(ステップS303)。ここで、切り替え概略位置の近傍の区間から密な間隔でサンプリングしたカプセル型内視鏡の画像群をM画像群と呼ぶ。

【0059】

続いて、制御部45は、取得部41がサンプリングしたM画像群を学習済みのCNN50に入力することによって、胃、小腸および大腸それぞれに分類させる(ステップS304)。

30

【0060】

その後、臓器切り替わり位置判定部53は、ステップS304においてCNN50が分類した分類結果に基づいて、胃と小腸および小腸と大腸の切り替わりの位置を上述の方法によってより高精度に推定する(ステップS305)。

【0061】

続いて、制御部45は、M画像群を、胃と小腸の切り替わりの正確な位置を表す幽門、小腸と大腸の切り替わりの正確な位置を表す回盲弁、およびそれ以外の3つのカテゴリを消化器内視鏡画像を用いてプレトレーニングを行った後に、カプセル型内視鏡画像を用いてファインチューニングを行ったCNN50に入力することによって、幽門、回盲弁およびそれ以外に分類させる(ステップS306)。この場合、CNN50は、幽門の信頼度に相当する幽門スコアと、回盲弁の信頼度に相当する回盲弁スコアを算出する。

40

【0062】

その後、幽門スコアが所定の閾値 $Th4$ より大きい場合(ステップS307:Yes)、CNN50は、幽門スコアの最大位置の画像を幽門ランドマークとして表示部43に表示する(ステップS308)。ステップS308の後、画像認識装置4は、後述するステップS310へ移行する。

【0063】

ステップS307において、幽門スコアが所定の閾値 $Th4$ 以下である場合(ステップ

50

S 3 0 7 : N o)、C N N 5 0 は、M 画像群に幽門が写っていない可能性があるため、ステップ S 3 0 5 において推定した胃と小腸の切り替わり推定位置を幽門ランドマークとして表示部 4 3 に表示する (ステップ S 3 0 9)。ステップ S 3 0 9 の後、画像認識装置 4 は、後述するステップ S 3 1 0 へ移行する。

【0064】

続いて、回盲弁スコアが所定の閾値 Th_5 より大きい場合 (ステップ S 3 1 0 : Y e s)、C N N 5 0 は、回盲弁スコアの最大位置の画像を回盲弁ランドマークとして表示部 4 3 に表示する (ステップ S 3 1 1)。ステップ S 3 1 1 の後、画像認識装置 4 は、本処理を終了する。

【0065】

ステップ S 3 1 0 において、回盲弁スコアが所定の閾値 Th_5 以下である場合 (ステップ S 3 1 0 : N o)、C N N 5 0 は、M 画像群に回盲弁が写っていない可能性があるため、ステップ S 3 0 5 において推定した小腸と大腸の切り替わり推定位置を回盲弁ランドマークとして表示部 4 3 に表示する (ステップ S 3 1 2)。ステップ S 3 1 2 の後、画像認識装置 4 は、本処理を終了する。

【0066】

以上説明した本実施の形態 1 によれば、C N N 5 0 が消化器内視鏡画像群と第 1 正解ラベルとを用いて、消化器内視鏡画像群と第 1 正解ラベルとの誤差を小さくする誤差逆伝搬法によってパラメータの重みを更新するプレトレーニングを行った後、C N N 5 0 がカプセル型内視鏡画像群と第 2 正解ラベルとを用いて、カプセル型内視鏡画像群と第 2 正解ラベルとの誤差を小さくする誤差逆伝搬法によってパラメータを更新するファインチューニングを行い、C N N 5 0 の中間出力をカプセル型内視鏡画像群の特徴量として抽出し、S V M 5 2 が C N N 5 0 によって抽出されたカプセル型内視鏡画像群の特徴量と第 2 正解ラベルとを用いて入力画像の識別または認識対象物の検出を学習する第 3 の学習を行う。これにより、学習用画像が不足する認識対象に対して、高精度な画像認識を実現することができる。

【0067】

また、本実施の形態 1 によれば、消化管の蠕動運動によって運ばれ、液体 (例えば水や消化液等を介して撮像されたカプセル型内視鏡画像群の代わりに、医師によって撮像方向が調整され、かつ、空気を介して撮像された消化器内視鏡によるカプセル型内視鏡画像と見え方が異なる消化器内視鏡画像群を間接的にプレトレーニング用の画像として利用することで、デジタルカメラや他の医療装置による映像画像をプレトレーニングに用いる場合よりも、高精度な学習を行うことができる。

【0068】

(実施の形態 1 の変形例 1)

次に、本発明の実施の形態 1 の変形例 1 について説明する。本実施の形態 1 の変形例 1 は、上述した実施の形態 1 と同様の構成を有し、画像認識装置が実行する学習処理が異なる。以下においては、本実施の形態 1 の変形例 1 に係る画像認識装置が実行する学習処理について説明する。なお、上述した実施の形態 1 に係る画像認識装置 4 と同一の構成には同一の符号を付して説明を省略する。

【0069】

〔学習処理〕

図 8 は、本実施の形態 1 の変形例 1 に係る画像認識装置 4 が実行する学習処理の概要を示すフローチャートである。図 8 において、ステップ S 4 0 1 ~ ステップ S 4 0 4 は、上述した図 3 のステップ S 1 0 1 ~ ステップ S 1 0 4 それぞれに対応し、C N N 5 0 による消化器内視鏡画像群を用いたプレトレーニング (第 1 の学習) に対応する。また、ステップ S 4 0 5 ~ ステップ S 4 0 7 は、上述した図 3 のステップ S 1 0 8 ~ ステップ S 1 1 0 それぞれに対応し、S V M 5 2 によるカプセル型内視鏡画像群を用いた学習処理 (第 2 の学習) に対応する。ステップ S 4 0 7 の後、画像認識装置 4 は、本処理を終了する。

【0070】

以上説明した本実施の形態 1 の変形例 1 によれば、CNN 50 によるカプセル型内視鏡画像を用いたファインチューニングを省略することによって、CNN 50 による学習時間を短縮することができる。

【0071】

(実施の形態 1 の変形例 2)

次に、本発明の実施の形態 1 の変形例 2 について説明する。本実施の形態 1 の変形例 2 は、上述した実施の形態 1 と同様の構成を有し、画像認識装置が実行する学習処理が異なる。以下においては、本実施の形態 1 の変形例 2 に係る画像認識装置が実行する学習処理について説明する。なお、上述した実施の形態 1 に係る画像認識装置 4 と同一の構成には同一の符号を付して説明を省略する。

【0072】

〔学習処理〕

図 9 は、本実施の形態 1 の変形例 2 に係る画像認識装置 4 が実行する学習処理の概要を示すフローチャートである。図 9 において、ステップ S 501～ステップ S 504 は、上述した図 3 のステップ S 101～ステップ S 104 それぞれに対応し、CNN 50 による消化器内視鏡画像を用いたプレトレーニング（第 1 の学習）に対応する。また、ステップ S 505～ステップ S 507 は、上述した図 3 のステップ S 105～ステップ S 107 それぞれに対応し、CNN 50 によるカプセル型内視鏡画像を用いたファインチューニング（第 2 の学習）に対応する。

【0073】

以上説明した本実施の形態 1 の変形例 2 によれば、SVM 52 によるカプセル型内視鏡画像を用いた学習を省略することができる。

【0074】

(実施の形態 2)

次に、本発明の実施の形態 2 について説明する。本実施の形態 2 は、上述した実施の形態 1 と同様の構成を有し、画像認識装置が実行する学習処理が異なる。具体的には、本実施の形態 1 では、カプセル型内視鏡画像群に対する画像分類を説明していたが、本実施の形態 2 では、病変部等の病変部の概略位置および大きさに関する情報も学習する。なお、上述した実施の形態 1 と同一の構成には同一の符号を付して説明を省略する。

【0075】

〔学習処理〕

図 10 は、本実施の形態 2 に係る画像認識装置 4 が実行する学習処理の概要を示すフローチャートである。なお、図 10 においては、CNN 50 のタイプを Faster R-CNN（以下、「R-CNN 50」という）の例として説明する。また、図 10 においては、R-CNN 50 のネットワークを上述した実施の形態 1 において説明した全結合層 503 に代えて、物体が存在する矩形領域（Bounding Box）の位置および大きさ（矩形の 4 頂点）を学習して回帰および分類するネットワークを設定した場合について説明する。さらに、図 10 においては、画像上でなく、R-CNN 50 のコンボリューション層とプーリング層との最終層の出力上において、矩形領域をスライドさせる代わりに、Fully-CNN として実現する。

【0076】

図 10 において、まず、制御部 45 は、R-CNN 50 の重みを乱数で初期化する（ステップ S 601）。

【0077】

続いて、制御部 45 は、取得部 41 を介して消化器内視鏡画像データベース 2 から消化器内視鏡によるポリープ画像および正常画像を R-CNN 50 に入力する。

【0078】

その後、制御部 45 は、R-CNN 50 の出力と消化器内視鏡画像データベース 2 からの正解ラベル（+1，-1 等）との誤差を小さくする誤差逆伝搬法によって R-CNN 50 の重みを更新する（ステップ S 603）。これにより、R-CNN 50 は、消化器内視

10

20

30

40

50

鏡によるポリープ画像および正常画像のプレトレーニング（第1の学習）を行うことができる。

【0079】

続いて、R-CNN50によるプレトレーニングの学習枚数が所定の閾値Th1を超えた場合（ステップS604：Yes）、画像認識装置4は、ステップS605へ移行する。これに対して、R-CNN50によるプレトレーニングの学習枚数が所定の閾値Th1を超えていない場合（ステップS604：No）、画像認識装置4は、上述したステップS602へ戻る。

【0080】

ステップS605において、制御部45は、取得部41を介してカプセル型内視鏡によるポリープ画像および正常画像をR-CNN50に入力する。

【0081】

続いて、制御部45は、R-CNN50の出力とカプセル型内視鏡画像データベース3からの正解ラベル（+1，-1等）との誤差を小さくする誤差逆伝搬法によってR-CNN50の重みを更新する（ステップS606）。これにより、R-CNN50は、カプセル型内視鏡によるポリープ画像および正常画像のファインチューニング（第2の学習）を行うことによって、病変部の位置（座標）および大きさを認識することができる回帰および分類用のネットワークを構成することができる。

【0082】

その後、R-CNN50によるファインチューニングの学習枚数が所定の閾値Th2を超えた場合（ステップS607：Yes）、画像認識装置4は、本処理を終了する。これに対して、R-CNN50によるファインチューニングの学習枚数が所定の閾値Th2を超えていない場合（ステップS607：No）、画像認識装置4は、上述したステップS605へ戻る。

【0083】

以上説明した本実施の形態2によれば、学習用画像が不足する認識対象に対して、高精度な画像認識を実現することができる。

【0084】

さらに、本実施の形態2によれば、カプセル型内視鏡画像列をR-CNN50に入力した場合、ポリープ候補の矩形領域を出力することができる。

【0085】

なお、本実施の形態2では、R-CNN50をFaster R-CNNを用いて説明したが、OverFeat、Spatial Pyramid Pooling Network（SPP-Net）やそれ以外のCNNを用いてもよい。

【0086】

（その他の実施の形態）

本発明に係るCNNとして、周知のネットワークを用いて構成してもよい。例えば、CNNをVGG-16によって構成してもよい。この場合、VGG-16は、3×3のサイズのコンボリューションフィルタを用いて、入力画像とのコンボリューション結果を非線形関数ReLUに適用する。そして、VGG-16は、コンボリューション層を2層または3層続けた後、Max Pooling（前層の2×2の出力の中の最大値を選択するサブサンプリング）が用いられる。なお、VGG-16では、層の数え方の中にプーリング層が入っていない。即ち、VGG-16では、13層のコンボリューション層、5層のMax Poolingおよび3層の全結合層から構成される。CNNとしてVGG-16を用いる場合であっても、消化器内視鏡画像を用いてプレトレーニング学習（第1学習）を行った後に、カプセル型内視鏡画像を用いてファインチューニング（第2学習）を行うことによって、上述した実施の形態と同様の構成を奏する。なお、CNNの構成例としてVGG-16を例に説明したが、例えば、Alex-net、Alex-netの改良版であるZF-net、GoogLeNetおよびNetwork in Networkのいずれかを用いて構成してもよいし、これらのネットワークを適宜組み合わせ構成してもよい

10

20

30

40

50

。

【0087】

また、本発明では、記録装置に記録された画像処理プログラムをパーソナルコンピュータやワークステーション等のコンピュータシステムで実行することによって実現することができる。また、このようなコンピュータシステムを、ローカルエリアネットワーク（LAN）、広域エリアネットワーク（WAN）、または、インターネット等の公衆回線を介して、他のコンピュータシステムやサーバ等の機器に接続して使用しても良い。この場合、実施の形態1，2およびこれらの変形例に係る画像認識装置は、これらのネットワークを介して管腔内画像の画像データを取得したり、これらのネットワークを介して接続されたビューアーやプリンタ等の種々の出力機器に画像処理結果を出力したり、これらのネットワークを介して接続された記憶装置、例えばネットワークに接続された読取装置によって読み取り可能な記録媒体等に画像処理結果を格納するようにしても良い。

10

【0088】

なお、本発明は、実施の形態1，2およびこれらの変形例に限定されるものではなく、各実施の形態や変形例に開示されている複数の構成要素を適宜組み合わせることによって、種々の発明を形成できる。例えば、各実施の形態や変形例に示される全構成要素からいくつかの構成要素を除外して形成しても良いし、異なる実施の形態や変形例に示した構成要素を適宜組み合わせて形成しても良い。

【符号の説明】

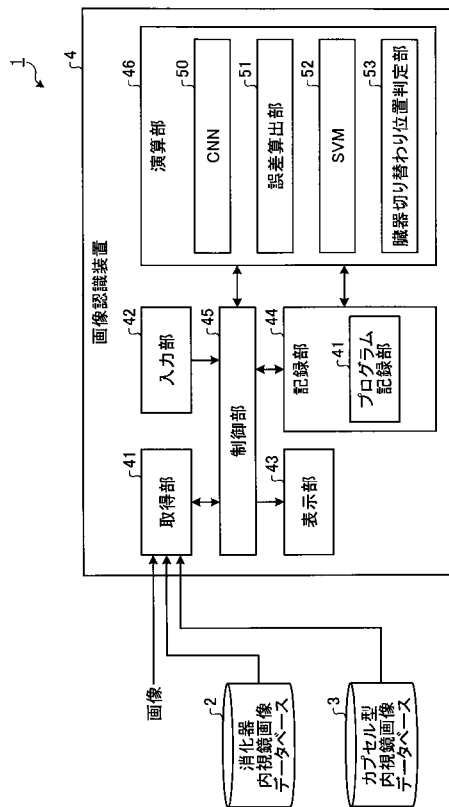
【0089】

20

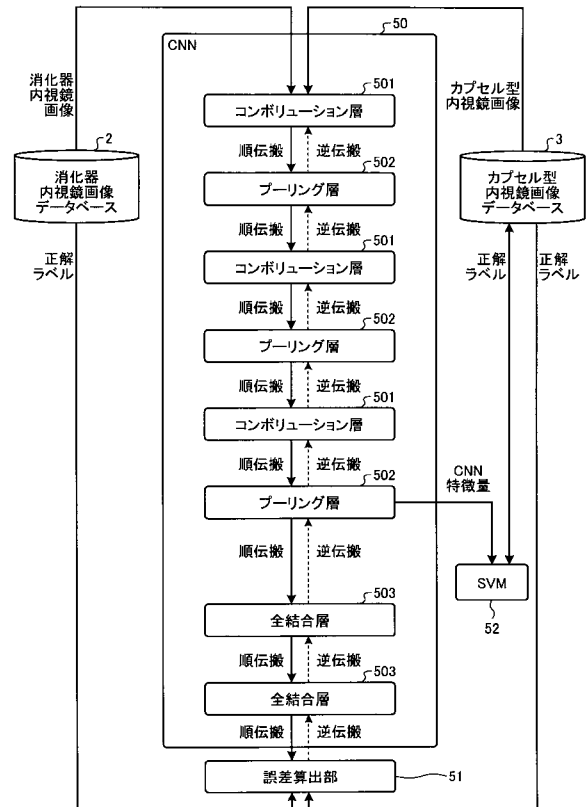
- 1 学習システム
- 2 消化器内視鏡画像データベース
- 3 カプセル型内視鏡画像データベース
- 4 画像認識装置
 - 41 取得部
 - 42 入力部
 - 43 表示部
 - 44 記録部
 - 45 制御部
 - 46 演算部
- 50 コンボリユーショナルニューラルネットワーク
 - 51 誤差算出部
 - 52 サポートベクターマシン
 - 53 臓器切り替わり位置判定部
- 441 プログラム記録部
- 501 コンボリユーション層
- 502 プーリング層
- 503 全結合層

30

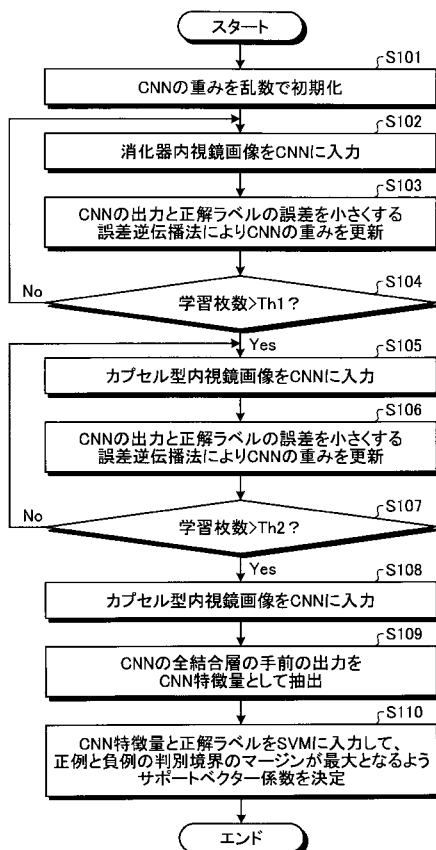
【図 1】



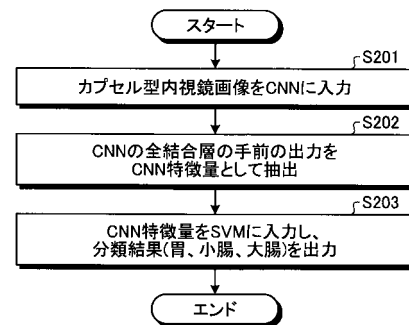
【図 2】



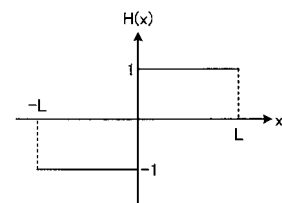
【図 3】



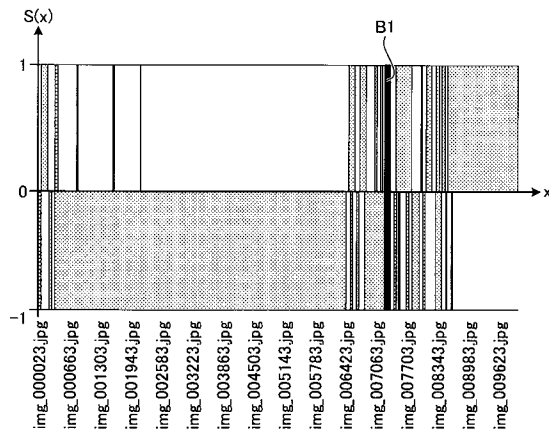
【図 4】



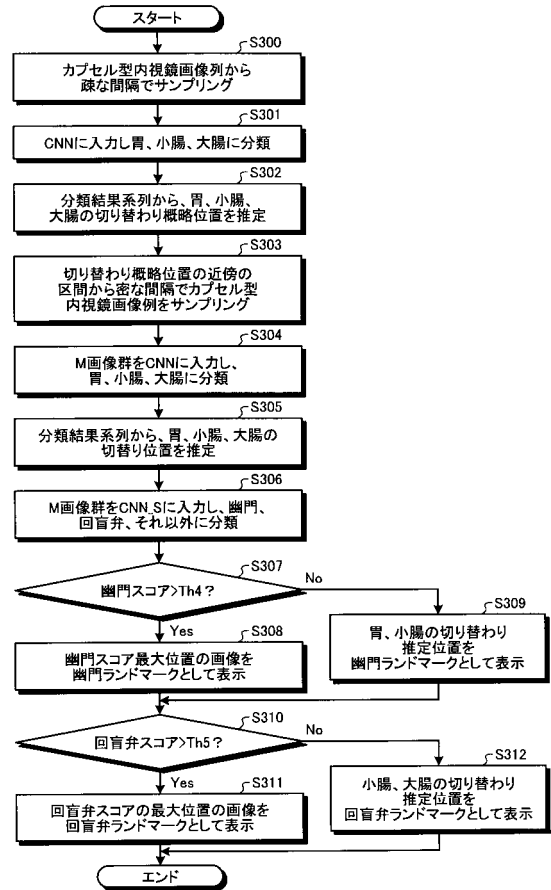
【図 5】



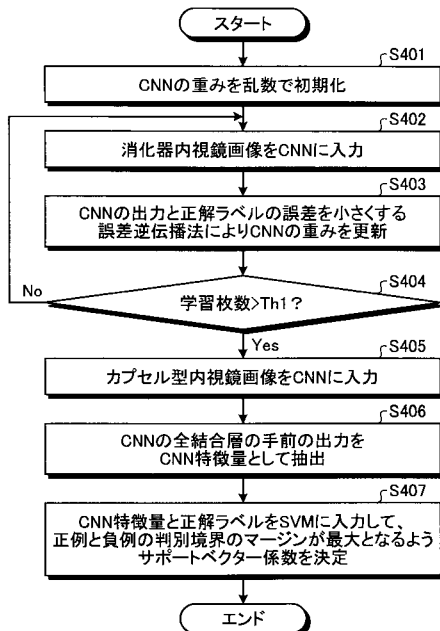
【図 6】



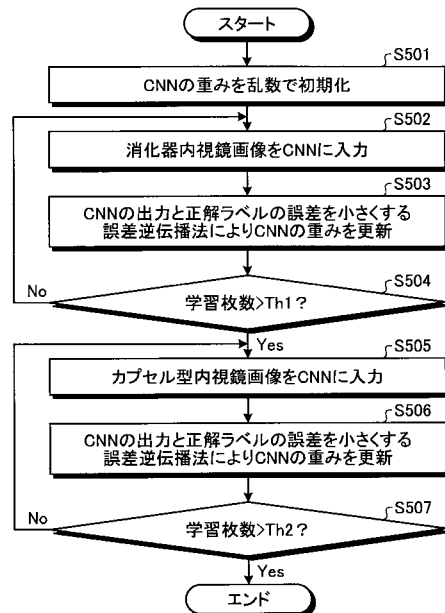
【図 7】



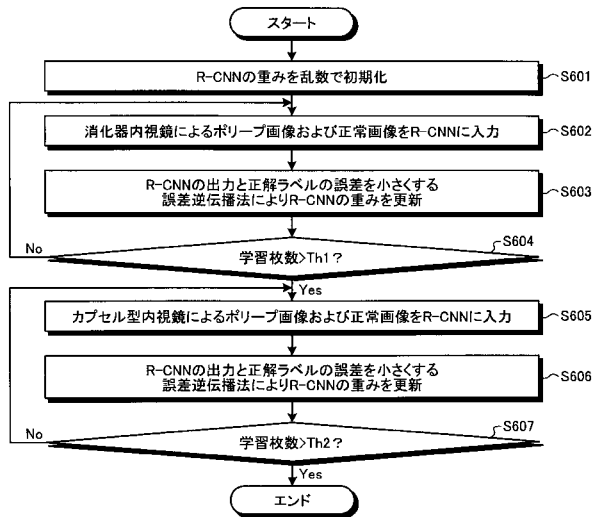
【図 8】



【図 9】



【図 10】



【国際調査報告】

INTERNATIONAL SEARCH REPORT		International application No. PCT/JP2016/061045									
A. CLASSIFICATION OF SUBJECT MATTER A61B5/00(2006.01)i, A61B1/00(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) A61B5/00-5/01, A61B1/00, G06T1/00, G06T7/00-7/60 Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Jitsuyo Shinan Koho 1922-1996 Jitsuyo Shinan Toroku Koho 1996-2016 Kokai Jitsuyo Shinan Koho 1971-2016 Toroku Jitsuyo Shinan Koho 1994-2016 Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) IEEE Xplore & keyword: CNN, Convolutional Neural Network, learning, frame rate											
C. DOCUMENTS CONSIDERED TO BE RELEVANT <table border="1"> <thead> <tr> <th>Category*</th> <th>Citation of document, with indication, where appropriate, of the relevant passages</th> <th>Relevant to claim No.</th> </tr> </thead> <tbody> <tr> <td>A</td> <td>JP 2007-87345 A (Canon Inc.), 05 April 2007 (05.04.2007), paragraphs [0022], [0029]; fig. 1 & US 2007/0122036 A1 & US 2013/0322770 A1 & EP 1768058 A2 paragraphs [0068], [0075]; fig. 1 & CN 1940965 A & KR 10-2007-0034966 A</td> <td>1-15</td> </tr> <tr> <td>A</td> <td>Nima Tajbakhsh et al., "Convolutional Neural Networks for Medical Image Analysis: Full Training or Fine Tuning?", IEEE Transactions on Medical Imaging, 2016.05, Vol.35, No.5, p.1299-1312</td> <td>1-15</td> </tr> </tbody> </table>			Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.	A	JP 2007-87345 A (Canon Inc.), 05 April 2007 (05.04.2007), paragraphs [0022], [0029]; fig. 1 & US 2007/0122036 A1 & US 2013/0322770 A1 & EP 1768058 A2 paragraphs [0068], [0075]; fig. 1 & CN 1940965 A & KR 10-2007-0034966 A	1-15	A	Nima Tajbakhsh et al., "Convolutional Neural Networks for Medical Image Analysis: Full Training or Fine Tuning?", IEEE Transactions on Medical Imaging, 2016.05, Vol.35, No.5, p.1299-1312	1-15
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.									
A	JP 2007-87345 A (Canon Inc.), 05 April 2007 (05.04.2007), paragraphs [0022], [0029]; fig. 1 & US 2007/0122036 A1 & US 2013/0322770 A1 & EP 1768058 A2 paragraphs [0068], [0075]; fig. 1 & CN 1940965 A & KR 10-2007-0034966 A	1-15									
A	Nima Tajbakhsh et al., "Convolutional Neural Networks for Medical Image Analysis: Full Training or Fine Tuning?", IEEE Transactions on Medical Imaging, 2016.05, Vol.35, No.5, p.1299-1312	1-15									
☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.											
<table border="0"> <tr> <td> * 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 </td> <td> "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 </td> </tr> </table>			* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier application or patent but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family							
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier application or patent but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family										
Date of the actual completion of the international search 13 June 2016 (13.06.16)		Date of mailing of the international search report 21 June 2016 (21.06.16)									
Name and mailing address of the ISA/ Japan Patent Office 3-4-3, Kasumigaseki, Chiyoda-ku, Tokyo 100-8915, Japan		Authorized officer Telephone No.									

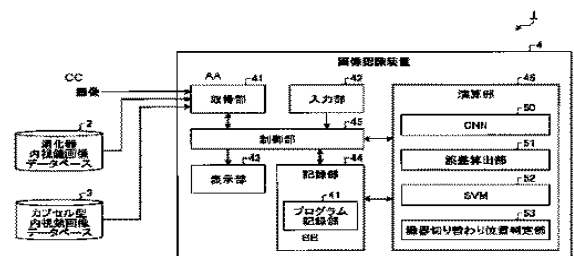
国際調査報告		国際出願番号 PCT/J P 2 0 1 6 / 0 6 1 0 4 5	
A. 発明の属する分野の分類 (国際特許分類 (I P C)) Int.Cl. A61B5/00(2006.01)i, A61B1/00(2006.01)i			
B. 調査を行った分野 調査を行った最小限資料 (国際特許分類 (I P C)) Int.Cl. A61B5/00-5/01, A61B1/00, G06T1/00, G06T7/00-7/60			
最小限資料以外の資料で調査を行った分野に含まれるもの 日本国実用新案公報 1922-1996年 日本国公開実用新案公報 1971-2016年 日本国実用新案登録公報 1996-2016年 日本国登録実用新案公報 1994-2016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) IEEE Xplore & キーワード: CNN, Convolutional Neural Network, learning, frame rate			
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	JP 2007-87345 A (キヤノン株式会社) 2007.04.05, 段落【0022】、【0029】、図1 & US 2007/0122036 A1 & US 2013/0322770 A1 & EP 1768058 A2, 段落[0068], [0075], 図1 & CN 1940965 A & KR 10-2007-0034966 A	1-15	
A	Nima Tajbakhsh et al., "Convolutional Neural Networks for Medical Image Analysis: Full Training or Fine Tuning?", IEEE Transactions on Medical Imaging, 2016.05, Vol.35, No.5, p.1299-1312	1-15	
☐ 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、一般的技术水準を示すもの 「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 13.06.2016		国際調査報告の発送日 21.06.2016	
国際調査機関の名称及びあて先 日本国特許庁 (I S A / 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) 伊知地 和之 電話番号 03-3581-1101 内線 3292	

(注) この公表は、国際事務局(WIPO)により国際公開された公報を基に作成したものである。なおこの公表に係る日本語特許出願(日本語実用新案登録出願)の国際公開の効果は、特許法第184条の10第1項(実用新案法第48条の13第2項)により生ずるものであり、本掲載とは関係ありません。

专利名称(译)	学习方法，图像识别装置和程序		
公开(公告)号	JPWO2017175282A1	公开(公告)日	2019-02-28
申请号	JP2018510033	申请日	2016-04-04
[标]申请(专利权)人(译)	奥林巴斯株式会社		
申请(专利权)人(译)	奥林巴斯公司		
[标]发明人	白谷文行		
发明人	白谷 文行		
IPC分类号	G06T7/00 A61B5/00 A61B1/00 A61B1/045 G06N3/08		
CPC分类号	A61B1/00009 A61B1/041 A61B1/00 G06K9/4628 G06K9/6274 G06K2209/051 G06N3/0427 G06N3/0454 G06N3/084 G06N20/10 G06T7/70 G06T2207/10068 G06T2207/20081 G06T2207/20084 G06T2207/30028 G06T2207/30092 G06K9/6269 G06T7/0012		
FI分类号	G06T7/00.350.C A61B5/00.G A61B1/00.C A61B1/045.614 G06N3/08.140		
F-TERM分类号	4C117/XA04 4C117/XB09 4C117/XD27 4C117/XE34 4C117/XE43 4C117/XJ01 4C117/XK33 4C117/XK42 4C117/XL12 4C161/DD07 4C161/FF14 4C161/SS21 4C161/TT15 4C161/YY07 5L096/AA06 5L096/BA06 5L096/BA13 5L096/CA02 5L096/DA02 5L096/EA39 5L096/GA51 5L096/HA09 5L096/HA11 5L096/JA11 5L096/KA04		
其他公开文献	JP6656357B2		
外部链接	Espacenet		

摘要(译)

提供一种学习方法，图像识别装置以及程序，其能够针对缺乏学习图像的识别对象实现高精度的图像识别。图像识别设备4，通过卷积神经网络的第一次学习使用从外部输入的第一帧率捕获的第一图像组，在执行了第一次学习之后，外部通过使用第二图像组执行第二学习，该第二图像组以低于从帧输入的第一帧速率的第二帧速率被捕获并且在数量上小于第一图像组。



- 2 Digestive organ endoscope image database
- 3 Capsule type endoscope image database
- 4 Image recognition device
- 42 Input unit
- 43 Display unit
- 44 Recording unit
- 45 Control unit
- 46 Calculation unit
- 51 Error calculation unit
- 53 Organ switching position determination unit
- AA Acquisition unit
- BB Program recording unit
- CC Image