

(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
2022年3月17日(17.03.2022)



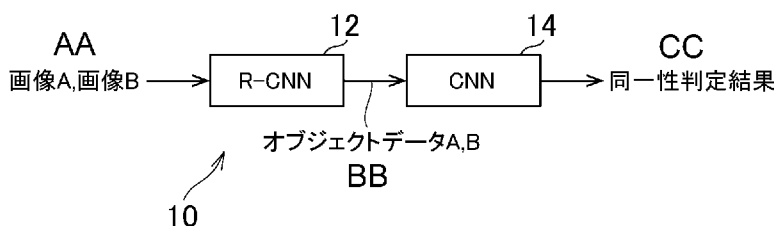
(10) 国際公開番号
WO 2022/054124 A1

- (51) 国際特許分類:
G06T 7/00 (2017.01) *G06N 20/00* (2019.01)
- (21) 国際出願番号: PCT/JP2020/033921
- (22) 国際出願日: 2020年9月8日(08.09.2020)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (71) 出願人: 楽天グループ株式会社 (**RAKUTEN GROUP, INC.**) [JP/JP]; 〒1580094 東京都世田谷区玉川一丁目14番1号 Tokyo (JP).
- (72) 発明者: 蔡 永男 (**CHAE, Yeongnam**); 〒1580094 東京都世田谷区玉川一丁目14番1号 楽天株式会社内 Tokyo (JP). プラキャッシャー プリーサム (**PRAKASHA, Preetham**); 〒1580094 東京都世田谷区玉川一丁目14番1号 楽天株式会社内 Tokyo (JP).
- (74) 代理人: 特許業務法人はるか国際特許事務所 (**HARUKA PATENT & TRADEMARK ATTORNEYS**); 〒1020085 東京都千代田区六番町3六番町SKビル5階 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, IT, JO, JP, KE, KG, KH, KN, KP, KR, KW, 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, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(54) **Title:** IMAGE DETERMINATION DEVICE, IMAGE DETERMINATION METHOD, AND PROGRAM

(54) 発明の名称: 画像判定装置、画像判定方法及びプログラム

[図1]



AA ... Image A, image B
BB ... Object data
CC ... Identity determination result

(57) **Abstract:** The purpose of the present invention is to easily and accurately determine substantial identity of a plurality of images, each having objects arranged therein. Provided is an image determination method comprising: a step for acquiring, from a first image, first object data indicating the attributes and arrangement of the objects in the first image, using an R-CNN 12, which is a first machine learning model; a step for acquiring, from a second image, second object data indicating the attributes and arrangement of the objects in the second image, using the R-CNN 12; and a step for determining substantial identity of the first image and the second image from the first object data and the second object data using a CNN that is a second machine learning model.



WO 2022/054124 A1

(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))

(57) 要約 : それぞれオブジェクトが配置されている複数の画像の実質同一性を簡易且つ正しく判定すること。第1の機械学習モデルであるR-CNN12を用いて、第1画像から該第1画像中のオブジェクトの属性及び配置を示す第1オブジェクトデータを取得するステップと、R-CNN12を用いて、第2画像から該第2画像中のオブジェクトの属性及び配置を示す第2オブジェクトデータを取得するステップと、第2の機械学習モデルであるCNNを用いて、前記第1オブジェクトデータ及び前記第2オブジェクトデータから、前記第1画像及び前記第2画像の実質同一性を判定するステップと、を含む画像判定方法が提供される。

明 細 書

発明の名称：画像判定装置、画像判定方法及びプログラム

技術分野

[0001] 本発明は画像判定装置、画像判定方法及びプログラムに関し、特に複数の画像の実質的同一性を判定する技術に関する。

背景技術

[0002] コンピュータ技術の分野において、それぞれオブジェクトが配置されている複数の画像が実質的に同一であるか否かを判定するニーズは多い。例えば、コンピュータアプリケーションのスクリーン画像は、特定のスマートフォンなどの選択された1つの実行環境で表示されることを前提に、人手によりデザインされる。こうしたスクリーン画像には、ボタン、画像、入力フォームなどのオブジェクトが配置されている。1つの実行環境のためのスクリーン画像のデザインが完成すると、実質的に同一のデザインのスクリーン画像が、他の実行環境にも移植される。これにより、実質的に同一のスクリーン画像を様々な実行環境で提供できるようになる。例えば、特定の製造者の特定のスマートフォンに向けてデザインされたスクリーン画像は、同一の製造者の他のスマートフォンや、異なる製造者のスマートフォンに向けて、人手又はコンピュータソフトウェアによる自動変換により移植される。このようにして得られた複数のスクリーン画像は、従来、目視でその実質同一性が確認されることが多かった。しかしながら、コンピュータアプリケーションの実行環境の数が増えるにつれ、そのような確認作業を自動化する必要性が高まっている。

発明の概要

発明が解決しようとする課題

[0003] コンピュータアプリケーションの実行環境により、画面のサイズ、アスペクト比、解像度が異なる。またオペレーションシステムなどの実行環境により提供され、各スクリーン画像に含められるボタン等のオブジェクトは、そ

の外観が少なからず異なっている。このため、複数のスクリーン画像を画素単位で比較しても、それらの実質同一性を確認することは困難である。また、機械学習モデルにそれらスクリーン画像を入力することで、それら画像の実質同一性を判定することも考えられるが、学習量が膨大になってしまうという問題が懸念される。

[0004] なお、コンピュータアプリケーションのスクリーン画像の他にも、各種環境で閲覧される電子書籍のページ画像や、各種環境で閲覧されるウェブコンテンツ画像など、オブジェクトが配置された複数画像の実質同一性を判定するニーズは多数存在する。

本発明は上記課題に鑑みてなされたものであって、その目的は、それぞれオブジェクトが配置されている複数の画像の実質同一性を簡易且つ正しく判定できる画像判定装置、画像判定方法及びプログラムを提供することにある。

課題を解決するための手段

[0005] 上記課題を解決するために、本発明の一側面に係る画像判定装置は、画像が入力され、該画像中のオブジェクトの属性及び配置を示すオブジェクトデータを出力する第1の機械学習モデルを用いて、第1画像から該第1画像中のオブジェクトの属性及び配置を示す第1オブジェクトデータを取得するとともに、第2画像から該第2画像中のオブジェクトの属性及び配置を示す第2オブジェクトデータを取得するオブジェクトデータ取得手段と、前記第1オブジェクトデータ及び前記第2オブジェクトデータが入力され、前記第1画像及び前記第2画像の実質同一性を出力する第2の機械学習モデルを用いて、前記第1画像及び前記第2画像の実質同一性を判定する判定手段と、を含む。

[0006] ここで、前記第1の機械学習モデルはR-CNNを含んでよい。

[0007] また、前記第1の機械学習モデルは、所与のベース画像に複数のオブジェクトの一部又は全部を重畳させることで生成されるトレーニング画像により学習されてよい。

- [0008] また、前記第2機械学習モデルは全結合層を含んでよい。
- [0009] 前記第2機械学習モデルは、前記全結合層の上流側に、前記第1オブジェクトデータ及び前記第2オブジェクトデータに基づく入力データの次元数を削減する畳込み層及びプーリング層を含んでよい。
- [0010] また、前記第2の機械学習モデルは、所定配置規則に従って、同一又は類似の第1及び第2のベース画像のそれぞれに所与のオブジェクトを重畳させることで生成される第1トレーニング画像及び第2トレーニング画像から得られる、前記第1トレーニング画像中のオブジェクトの属性及び配置を示す第1学習オブジェクトデータ、及び前記第2トレーニング画像中のオブジェクトの属性及び配置を示す第2学習オブジェクトデータにより学習されてよい。
- [0011] また、本発明の他の側面に係る画像判定方法は、第1の機械学習モデルを用いて、第1画像から該第1画像中のオブジェクトの属性及び配置を示す第1オブジェクトデータを取得するステップと、前記第1の機械学習モデルを用いて、第2画像から該第2画像中のオブジェクトの属性及び配置を示す第2オブジェクトデータを取得するステップと、第2の機械学習モデルを用いて、前記第1オブジェクトデータ及び前記第2オブジェクトデータから、前記第1画像及び前記第2画像の実質同一性を判定するステップと、を含む。
- [0012] ここで、前記第1の機械学習モデルはR-CNNを含んでよい。
- [0013] また、上記方法は、所与のベース画像に複数のオブジェクトの一部又は全部を重畳させることで生成されるトレーニング画像により、前記第1機械学習モデルの学習を実行するステップをさらに含んでよい。
- [0014] また、前記第2機械学習モデルは全結合層を含んでよい。
- [0015] また、前記第2機械学習モデルは、前記全結合層の上流側に、前記第1オブジェクトデータ及び前記第2オブジェクトデータに基づく入力データの次元数を削減する畳込み層及びプーリング層を含んでよい。
- [0016] また、上記方法は、所定配置規則に従って、同一又は類似の第1及び第2のベース画像のそれぞれに所与のオブジェクトを重畳させることで第1トレ

ーニング画像及び第2トレーニング画像を生成するステップと、前記第1トレーニング画像及び前記第2トレーニング画像のそれぞれを前記第1機械学習モデルに入力することにより、前記第1トレーニング画像中のオブジェクトの属性及び配置を示す第1学習オブジェクトデータ、及び前記第2トレーニング画像中のオブジェクトの属性及び配置を示す第2学習オブジェクトデータを取得するステップと、前記第1学習オブジェクトデータ及び前記第2学習オブジェクトデータにより前記第2機械学習モデルの学習を実行するステップと、をさらに含んでよい。

[0017] また、本発明のさらに他の側面に係るプログラムは、第1の機械学習モデルを用いて、第1画像から該第1画像中のオブジェクトの属性及び配置を示す第1オブジェクトデータを取得するステップと、前記第1の機械学習モデルを用いて、第2画像から該第2画像中のオブジェクトの属性及び配置を示す第2オブジェクトデータを取得するステップと、第2の機械学習モデルを用いて、前記第1オブジェクトデータ及び前記第2オブジェクトデータから、前記第1画像及び前記第2画像の実質同一性を判定するステップと、をコンピュータに実行させるためのプログラムである。このプログラムは、光磁気ディスクや半導体メモリなどのコンピュータ可読情報記憶媒体に格納されてよい。

図面の簡単な説明

[0018] [図1]本発明の実施形態に係る画像判定装置のブロック図である。

[図2]2つのスクリーン画像の同一性判定例を示す図である。

[図3]他の実施形態に係る画像判定装置のブロック図である。

[図4]CNNの構成例を示す図である。

[図5]CNNの他の構成例を示す図である。

[図6]R-CNNの学習に用いる学習装置のブロック図である。

[図7]ベース画像の例を示す図である。

[図8]オブジェクト画像の例を示す図である。

[図9]オブジェクト属性テーブルの例を示す図である。

[図10] トレーニング画像の例を示す図である。

[図11] 図10に示されるトレーニング画像に対応する正解データを示す図である。

[図12] R-CNNの学習処理を示すフロー図である。

[図13] CNNの学習に用いる学習装置のブロック図である。

[図14] CNNの学習に用いる正例及び負例に係るトレーニングデータを示す図である。

[図15] CNNの学習処理を示すフロー図である。

発明を実施するための形態

[0019] 本発明の一実施形態について図面に基づき詳細に説明する。なお、以下の説明においては、同一部分については各図面において同一符号を付し、適宜説明を省略する。

[0020] 図1は、本発明の一実施形態に係る画像判定装置の機能ブロック図である。同図に示される各機能ブロックは、CPUやメモリを中心に構成されたコンピュータ上で、本発明の一実施形態に係る画像判定プログラムが実行されることにより、実現される。同図に示される画像判定装置10は、2つのスクリーン画像の実質同一性を判定するものである。

[0021] 各スクリーン画像には、上述のようにボタン等のオブジェクトが配置されている。画像判定装置10は、2つのスクリーン画像に含まれるオブジェクトの数、それぞれのオブジェクトの属性、当該スクリーン画像中でのそれぞれのオブジェクトの配置に基づき、それら2つのスクリーン画像の実質同一性を判定する。ここで、オブジェクトの属性は、例えばオブジェクトの種類や色情報である。オブジェクトの種類としては、ボタン、ロゴ画像、商標画像、入力フォームなどが挙げられる。オブジェクトの色情報としては、当該オブジェクトの1又は複数の代表色の情報、平均色の情報などを採用してよい。

[0022] ここで、2つの画像に「実質同一性がある」とは、各画像に含まれるオブジェクト画像の数、それぞれの属性及び配置が、事前に定義された正例規則

に合致する場合をいう。また、2つの画像に「実質同一性がない」とは、各画像に含まれるオブジェクト画像の数、それぞれの属性及び配置が、事前に定義された負例規則に合致する場合をいう。

[0023] 図1に示される画像判定装置10によれば、図2(a)に示すようにスクリーン画像の左上にオブジェクトAが配置されており、右下にオブジェクトBが配置されている場合には、同図(b)に示すように、オブジェクトAだけが同じ位置に配置されており、オブジェクトBが失われているスクリーン画像は、同図(a)に示されるスクリーン画像と実質同一性がないと判断される。すなわち、オブジェクト数が不一致のスクリーン画像は実質同一性がないと判断される。

[0024] また、同図(c)に示すように、オブジェクトAは同じ位置に配置されているものの、オブジェクトBは水平方向にずれて配置されているスクリーン画像も、同図(a)に示されるスクリーン画像と実質同一性がないと判断される。すなわち、対応するオブジェクトが水平方向にずれて配置されているスクリーン画像も、同図(a)に示されるスクリーン画像と実質同一性がないと判断される。

[0025] 一方、同図(d)に示すように、オブジェクトBの大きさが異なるスクリーン画像や、同図(e)に示すように、オブジェクトAの大きさが異なるスクリーン画像は、同図(a)に示されるスクリーン画像と実質同一性があると判断される。すなわち、2つのスクリーン画像を比較したとき、オブジェクトの数、属性、配置が同じであれば、オブジェクトの大きさが異なっても、両スクリーン画像には実質同一性があると判断される。

[0026] このような実質同一性の判断を実現するため、図1に示すように、画像判定装置10は、第1の機械学習モデルであるR-CNN (Regions with Convolutional Neural Networks) 12と、第2の機械学習モデルであるCNN (Convolutional Neural Networks) 14と、を含んで構成されている。R-CNN 12には、比較対象である2つのスクリーン画像のうちスクリーン画像Aが入力されると、該スクリーン画像Aに含まれるオブジェクトの属性及び

配置を示すオブジェクトデータAが出力される。また、それら2つのスクリーン画像のうちスクリーン画像Bが入力されると、該スクリーン画像Bに含まれるオブジェクトの属性及び配置を示すオブジェクトデータBが出力される。これらオブジェクトデータA及びオブジェクトデータBはCNN14に入力される。CNN14では、これらオブジェクトデータA及びオブジェクトデータBに基づいて、スクリーン画像A及びスクリーン画像Bに実質同一性があるか否かを出力する。

[0027] なお、ここではR-CNN12にスクリーン画像A及びスクリーン画像Bを順次入力し、オブジェクトデータA及びオブジェクトデータBを順次得るようにしたが、図3に示すように、2つのR-CNN12を並列するようにしてCNN14の前段に設け、スクリーン画像AからオブジェクトデータAを生成するのと並行して、スクリーン画像BからオブジェクトデータBを生成するようにしてもよい。

[0028] 図1において、R-CNN12は、オブジェクト検出のための機械学習モデルであり、例えばオリジナルのR-CNN、Fast R-CNN、Faster R-CNN、MASK R-CNNといった公知のオブジェクト検出用機械学習モデルにより構成されてよい。本実施形態では、R-CNN12としてFaster R-CNNを一例として採用する。R-CNN12に画像が入力されると、該画像に含まれるオブジェクトのそれぞれについて、入力画像における配置及び当該オブジェクトの属性を示すオブジェクトデータが出力される。入力画像中の各オブジェクトの配置は、オブジェクトデータ中、例えばアンカーボックスにより特定される。

[0029] CNN14は、図4に示すように、データ統合部15、それぞれ畳込み層及びプーリング層の組からなる、複数段の次元削減部16、一次元化部17、及び複数段の全結合層18を含んでいる。データ統合部15は、R-CNN12から出力されるオブジェクトデータA及びオブジェクトデータBを結合(Concatenate)して1の入力データを生成する。例えば、オブジェクトデータA及びオブジェクトデータBは、オブジェクトの序数を示す1つの次元

(例えば300個のデータを含んでよい。)、オブジェクトの大きさ及び配置を示す2つの次元(例えば各次元に4個のデータを含んでよい。)、オブジェクトの属性を示す1つの次元(例えば1024個のデータを含んでよい。)からなる、4次元テンソルデータであってよい。データ統合部15は、例えばオブジェクトの序数の次元を2倍のサイズに伸長させることで、オブジェクトデータA及びオブジェクトBを結合し、入力データを得てよい。この場合、オブジェクトの序数1~300には、スクリーン画像Aに含まれるオブジェクトが割り当てられ、オブジェクトの序数301~600には、スクリーン画像Bに含まれるオブジェクトが割り当てられる。

[0030] データ統合部15により生成される入力データは複数段の次元削減部16により次元削減が行われ、最終段の次元削減部16から2次元の中間データが出力される。一次元化部17は、中間データを一次元化し、一次元化された中間データを最初の段の全結合層18に入力する。最終段の全結合層18は、一次元化された中間データから1次元(2個のデータを含んでよい。)の同一性判定結果を出力する。同一性判定結果は、スクリーン画像A及びスクリーン画像Bの同一性がある程度を示すデータと、そのような同一性がない程度を示すデータと、を含む。

[0031] 図4に示すCNN14によれば、スクリーン画像A及びスクリーン画像Bのそれぞれに含まれるオブジェクトの属性及び配置から、スクリーン画像A及びスクリーン画像Bの実質同一性の有無を判定できるようになる。

[0032] なお、図4ではオブジェクトデータA及びオブジェクトデータBを結合して入力データを得て、この入力データの次元削減を行うようにしたが、オブジェクトデータA及びオブジェクトデータBのそれぞれに対して次元削減を行うようにしてよい。すなわち、図5に示すCNN14のように、畳込み層及びプーリング層をそれぞれ含む複数段の次元削減部16aにより、オブジェクトデータAの次元削減を行うとともに、次元削減部16aと同一機能の次元削減部16bにより、オブジェクトデータAの次元削減と並行して、オブジェクトデータBの次元削減を行ってよい。次元削減部16a, 16bの

出力はそれぞれ2次元の中間データであり、データ統合部19は、各中間データを1次元化し、それら結合 (Concatenate) して1の1次元の中間データを出力する。データ統合部19から出力される中間データは複数段の全結合層18に入力され、最終段からは同一性判定結果が出力される。

[0033] 図5に示すCNN14を用いても、スクリーン画像A及びスクリーン画像Bのそれぞれに含まれるオブジェクトの属性及び配置から、スクリーン画像A及びスクリーン画像Bの実質同一性の有無を判定できるようになる。

[0034] ここで、R-CNN12及びCNN14の学習について説明する。

[0035] 図6は、R-CNN12の学習を行う学習装置20の構成を示す図である。同図に示される各機能ブロックは、CPUやメモリを中心に構成されたコンピュータ上で、本発明の一実施形態に係るプログラムが実行されることにより、実現される。同図に示すように、学習装置20は、R-CNN12の内部パラメータを更新することによりR-CNN12の学習を実行するトレーニング部21、R-CNN12の学習に用いるトレーニングデータを生成するトレーニングデータ生成部22、トレーニングデータを生成する際に用いるベース画像を記憶するベース画像記憶部24、同様にトレーニングデータを生成する際に用いるオブジェクト画像を記憶するオブジェクト画像記憶部26を含んでいる。

[0036] 図7は、ベース画像記憶部24に記憶されるベース画像の例を示している。ベース画像は、スクリーン画像の基礎となるラスタ形式の画像データであり、同図(a)は、全面に文章が表わされたベース画像を示している。同図(b)は白色無地のベース画像を示している。同図(c)は、有色無地のページ上部及び文書が表わされた下部からなるベース画像を示している。同図(d)は、複数ブロックに区分され、各ブロックに文書が表わされたベース画像を示している。ベース画像記憶部24は、こうしたラスタ形式のベース画像を多数記憶している。

[0037] 図8は、オブジェクト画像記憶部26に記憶されるオブジェクト画像の例を示している。オブジェクト画像は、スクリーン画像の基礎となるラスタ形

式の画像データであり、スクリーン画像中に配置される、ボタンや入力フォームなどのオブジェクトの外観を示している。同図（a）は、オブジェクト画像の一種であるロゴ画像を示している。同図に示すオブジェクト画像には、オブジェクト番号「001」が付与されている。同図（b）は、オブジェクト画像の一種である入力フォーム画像を示している。同図に示すオブジェクト画像には、オブジェクト番号「002」が付与されている。同図（c）は、オブジェクト画像の一種であるボタン画像を示している。同図に示すオブジェクト画像には、オブジェクト番号「003」が付与されている。同図（d）は、オブジェクト画像の一種である商標画像を示している。同図に示すオブジェクト画像には、オブジェクト番号「004」が付与されている。オブジェクト画像記憶部26は、こうしたラスタ形式のオブジェクト画像を多数記憶している。

[0038] オブジェクト画像記憶部26では、多数のオブジェクト画像に加えて、図9に示されるオブジェクト属性テーブルを記憶している。オブジェクト属性テーブルは、オブジェクト画像記憶部26に記憶される各オブジェクト画像の属性を記憶するものである。具体的には、オブジェクト属性テーブルは、各オブジェクト番号に関連づけて、該番号に係るオブジェクト画像の種類及び色情報を記憶するものである。例えば、オブジェクト属性テーブルでは、図8（a）に示されるオブジェクト画像のオブジェクト番号「001」に関連づけて、該オブジェクト画像の種類「ロゴ」及び該オブジェクト画像の色情報「赤・白」が記憶されている。

[0039] トレーニングデータ生成部22は、ベース画像記憶部24及びオブジェクト画像記憶部26に記憶されたデータに基づいて、多数のトレーニングデータを生成する。トレーニングデータは、トレーニング画像と正解データを含んでいる。

[0040] 図10は、トレーニングデータ生成部22により生成されるトレーニング画像の例を示している。同図（a）に示されるトレーニング画像は、図7（b）に示されるベース画像に、図8（a）に示されるオブジェクト画像、図

8 (d) に示されるオブジェクト画像及び図 8 (b) に示されるオブジェクト画像を重畳させることにより、生成されたものである。また、図 10 (b) に示されるトレーニング画像は、図 7 (a) に示されるベース画像に、図 8 (b) に示されるオブジェクト画像及び図 8 (c) に示されるオブジェクト画像を重畳させることにより、生成されたものである。

[0041] トレーニング画像を生成する際に用いるベース画像は、ベース画像記憶部 24 に記憶される多数のベース画像の中から、トレーニングデータ生成部 22 によりランダムに選択される。また、ベース画像に重畳させるオブジェクト画像も、オブジェクト画像記憶部 24 に記憶される多数のベース画像の中から、トレーニングデータ生成部 22 によりランダムに選択される。さらに、各オブジェクト画像の配置（位置及び大きさ）も、トレーニングデータ生成部 22 によりランダムに決定される。

[0042] 図 11 は、トレーニングデータ生成部 22 により生成される正解データの例を示している。同図 (a) は、図 10 (a) に示されるトレーニング画像に対する正解データを示している。この正解データには、トレーニング画像に 3 つのオブジェクト画像が含まれていること、1 番目のオブジェクト画像の配置は「 $(X A 1, Y A 1) - (X B 1, Y B 1)$ 」であり、属性は「ロゴ、赤・白」であること、2 番目のオブジェクト画像の配置は「 $(X A 2, Y A 2) - (X B 2, Y B 2)$ 」であり、属性は「マーク、黄」であること、3 番目のオブジェクト画像の配置は「 $(X A 3, Y A 3) - (X B 3, Y B 3)$ 」であり、属性は「フォーム、白・黒」であること、が示されている。すなわち、正解データは、トレーニング画像に含まれるオブジェクト画像の数、及び各オブジェクト画像の配置及び属性を示している。同様に、図 11 (b) は、図 10 (b) に示されるトレーニング画像に対する正解データを示している。

[0043] トレーニングデータ生成部 22 は、トレーニング画像を生成する際、ベース画像に重畳させるオブジェクト画像をランダムに選択し、その配置もランダムに決定している。トレーニングデータ生成部 22 は、選択されたオブジ

ェクト画像の属性を、図9に示されるオブジェクト属性テーブルから読み出し、図11に示す正解データに含めるとともに、決定されたオブジェクト画像の配置も、図11に示す正解データに含めることで、トレーニング画像に対応する正解データを生成している。

[0044] トレーニング部21は、トレーニングデータ生成部22により生成されるトレーニングデータを用いてR-CNN12の学習処理を実行する。具体的には、トレーニングデータに含まれるトレーニング画像を順次R-CNN12に入力し、オブジェクトデータの出力を得る。この出力と、トレーニングデータに含まれる正解データとの差分を演算し、この差分が小さくなるようにR-CNN12の内部パラメータを更新する。

[0045] 図12は、学習装置20の動作フロー図である。学習装置20は、まずトレーニングデータ生成部22が、オブジェクト画像記憶部26に記憶されたオブジェクト画像をランダムに選択する(S101)。ここで、選択するオブジェクト画像の数もランダムに決定される。次に、トレーニングデータ生成部22が、ベース画像記憶部24からランダムに1つのベース画像を選択する(S102)。

[0046] トレーニングデータ生成部22は、選択されたベース画像に対し、選択されたオブジェクト画像を配置することで、トレーニング画像を生成する(S103)。このとき、トレーニングデータ生成部22は、各オブジェクト画像について、配置及びサイズをランダムに決定する。

[0047] トレーニングデータ生成部22は、さらに、S101乃至S103での処理内容に応じて、図11に例示される正解データを生成し、S103で生成されたトレーニング画像とともに、トレーニングデータに含める(S104)。

[0048] S101乃至S104の処理を所定数のトレーニングデータが生成されるまで繰り返した後(S105)、生成されたトレーニングデータを用いてトレーニング部21がR-CNN12の学習処理を実行する(S106)。

[0049] 次に、図13は、CNN14の学習を行う学習装置30の構成を示す図で

ある。同図に示される各機能ブロックは、CPUやメモリを中心に構成されたコンピュータ上で、本発明の一実施形態に係るプログラムが実行されることにより、実現される。同図に示すように、学習装置30は、CNN14の内部パラメータを更新することによりCNN14の学習を実行するトレーニング部31、CNN14の学習に用いるトレーニングデータを生成するトレーニングデータ生成部32、ベース画像記憶部24及びオブジェクト画像記憶部26を含んでいる。トレーニングデータ生成部32は、ベース画像記憶部24に記憶されているベース画像、及びオブジェクト画像記憶部26に記憶されているオブジェクト画像を用いて、正例及び負例に係るトレーニングデータを生成する。各トレーニングデータには、正例、すなわち実質同一性のあるトレーニング画像のペア、又は負例、すなわち実質同一性のないトレーニング画像のペアが含まれる。

[0050] 図14は、CNN14の学習に用いる、正例及び負例に係るトレーニング画像を示す図である。同図(a)にはトレーニング画像のペアの一方の例が示されている。当該一方のトレーニング画像は、ランダムに選択されるベース画像に対し、ランダムに選択された複数のオブジェクト画像を、ランダムな位置に、ランダムな大きさで、それぞれ配置することにより生成される。同図(a)では、全体に文書が表わされたベース画像に対し、オブジェクト画像A及びBが配置されている。

[0051] 同図(b)は、第1の負例規則を適用して生成された他方のトレーニング画像を示している。第1の負例規則では、一方のトレーニング画像に含まれたオブジェクト画像の一部が除去される。例えば、同図(b)のトレーニング画像は、同図(a)のトレーニング画像とは異なり、オブジェクト画像Bを含んでいない。

[0052] 同図(c)は、第2の負例規則を適用して生成された他方のトレーニング画像を示している。第2の負例規則では、一方のトレーニング画像に含まれたオブジェクト画像の一部が水平方向に移動される。例えば、同図(c)のトレーニング画像では、同図(a)のトレーニング画像に比べて、オブジ

ェクト画像Bが左方向にずれている。

[0053] 同図(d)は、正例規則を適用して生成された他方のトレーニング画像を示している。正例規則では、一方のトレーニング画像に含まれたオブジェクト画像の一部が拡大又は縮小される。例えば、同図(d)のトレーニング画像では、同図(a)のトレーニング画像に比べて、オブジェクト画像Aが拡大されている。

[0054] トレーニング部31は、こうして生成されるトレーニングデータを用いて、CNN14の学習処理を実行する。具体的には、トレーニングデータに含まれる2つのトレーニング画像を順次R-CNN12に入力し、2つのオブジェクトデータを取得する。そして、それらオブジェクトデータをCNN14に入力する。トレーニング部31は、CNN14から出力される同一性判定結果を取得し、同一性判定結果が正しいものとなるよう、CNN14の内部パラメータを更新する。すなわち、正例に係るトレーニング画像をR-CNN12に入力した場合には、CNN14から出力される同一性判定結果が実質的同一であることを示すものになるよう、内部パラメータを更新する。逆に、負例に係るトレーニング画像をR-CNN12に入力した場合には、CNN14から出力される同一性判定結果が実質的同一でないことを示すものになるよう、内部パラメータを更新する。

[0055] 図15は、学習装置30の動作フロー図である。同図に示すように、学習装置30では、まずトレーニングデータ生成部32がオブジェクト画像記憶部26からランダムにオブジェクト画像を選択する(S201)。さらにトレーニングデータ生成部32は、ベース画像記憶部24からランダムにベース画像を選択する(S202)。そして、トレーニングデータ生成部32は、選択されたベース画像に対し、各オブジェクト画像を重畳させることで、トレーニング画像Aを生成する(S203)。このとき、各オブジェクト画像の位置及び大きさをランダムに決定する。

[0056] 次に、トレーニングデータ生成部32は、正例規則又は負例規則を適用して、選択されたベース画像に対し、選択されたオブジェクト画像の全部又は

一部を重畳させることで、トレーニング画像Bを生成する（S204）。例えば、第1の負例規則を適用する場合には、選択されたオブジェクト画像のうち一部を、選択されたベース画像に重畳させずに、トレーニング画像Bを生成する。第2の負例規則を適用する場合には、選択されたオブジェクト画像のうち一部を、右方向又は左方向に移動させてから、選択されたベース画像に重畳させることで、トレーニング画像Bを生成する。正例規則を適用する場合には、選択されたオブジェクト画像のうち一部を、拡大又は縮小させてから、選択されたベース画像に重畳させることで、トレーニング画像Bを生成する。

[0057] その後、トレーニングデータ生成部32は、生成されたトレーニング画像A及びBのペアと、それらが正例であるか負例であるかの区別と、を含むトレーニングデータを生成し、記憶させる（S205）。

[0058] S201乃至S205の処理を所定数のトレーニングデータが生成されるまで繰り返した後（S206）、生成されたトレーニングデータを用いてトレーニング部31がCNN14の学習処理を実行する（S207）。

[0059] 以上説明した画像判定装置10によれば、比較対象となる2つのスクリーン画像のそれぞれから、それらに含まれるオブジェクト画像の属性及び位置を示すオブジェクトデータが得られる。そして、2つのオブジェクトデータに基づいて、2つのスクリーン画像の実質同一性の有無が判定される。実質同一性の有無の判定のために、正例規則及び負例規則に従って生成された多数のトレーニング画像ペアにより、CNN14が事前に訓練される。本実施形態によれば、2つのスクリーン画像の実質同一性を好適に判定することができる。

[0060] 具体的には、本発明者らは、500枚のベース画像と33種類のオブジェクト画像を用いて、R-CNN12のために5000のトレーニングデータ、CNN14のために8000のトレーニングデータを生成し、R-CNN12及びCNN14の学習を実施したところ、スクリーン画像の実質同一性の判定精度はおよそ86%となり、十分実用に耐えることが分かった。

[0061] なお、本発明の範囲は上記実施形態に限定されず、種々の変形を包含する。例えば、本発明はスクリーン画像のみならず、電子書籍のページ画像やウェブコンテンツ画像など、様々な画像にも同様に適用することができる。

符号の説明

[0062] 10 画像判定装置、12 R-CNN、14 CNN、15, 19 データ統合部、16, 16a 畳み込み層及びプーリング層、17 一次元化部、18 全結合層、20, 30 学習装置、21, 31 トレーニング部、22, 32 トレーニングデータ生成部、24 ベース画像記憶部、26 オブジェクト画像記憶部。

請求の範囲

- [請求項1] 画像が入力され、該画像中のオブジェクトの属性及び配置を示すオブジェクトデータを出力する第1の機械学習モデルを用いて、第1画像から該第1画像中のオブジェクトの属性及び配置を示す第1オブジェクトデータを取得するとともに、第2画像から該第2画像中のオブジェクトの属性及び配置を示す第2オブジェクトデータを取得するオブジェクトデータ取得手段と、
- 前記第1オブジェクトデータ及び前記第2オブジェクトデータが入力され、前記第1画像及び前記第2画像の実質同一性を出力する第2の機械学習モデルを用いて、前記第1画像及び前記第2画像の実質同一性を判定する判定手段と、
- を含むことを特徴とする画像判定装置。
- [請求項2] 請求項1に記載の画像判定装置において、
- 前記第1の機械学習モデルはR-CNNを含む、
- ことを特徴とする画像判定装置。
- [請求項3] 請求項1又は2に記載の画像判定装置において、
- 前記第1の機械学習モデルは、所与のベース画像に複数のオブジェクトの一部又は全部を重畳させることで生成されるトレーニング画像により学習される、
- ことを特徴とする画像判定装置。
- [請求項4] 請求項1乃至3のいずれかに記載の画像判定装置において、
- 前記第2機械学習モデルは全結合層を含む、
- ことを特徴とする画像判定装置。
- [請求項5] 請求項4に記載の画像判定装置において、
- 前記第2機械学習モデルは、前記全結合層の上流側に、前記第1オブジェクトデータ及び前記第2オブジェクトデータに基づく入力データの次元数を削減する畳込み層及びプーリング層を含む、
- ことを特徴とする画像判定装置。

- [請求項6] 請求項4又は5に記載の画像判定装置において、
前記第2の機械学習モデルは、所定配置規則に従って、同一又は類似の第1及び第2のベース画像のそれぞれに所与のオブジェクトを重畳させることで生成される第1トレーニング画像及び第2トレーニング画像から得られる、前記第1トレーニング画像中のオブジェクトの属性及び配置を示す第1学習オブジェクトデータ、及び前記第2トレーニング画像中のオブジェクトの属性及び配置を示す第2学習オブジェクトデータにより学習される、
ことを特徴とする画像判定装置。
- [請求項7] 第1の機械学習モデルを用いて、第1画像から該第1画像中のオブジェクトの属性及び配置を示す第1オブジェクトデータを取得するステップと、
前記第1の機械学習モデルを用いて、第2画像から該第2画像中のオブジェクトの属性及び配置を示す第2オブジェクトデータを取得するステップと、
第2の機械学習モデルを用いて、前記第1オブジェクトデータ及び前記第2オブジェクトデータから、前記第1画像及び前記第2画像の実質同一性を判定するステップと、
を含むことを特徴とする画像判定方法。
- [請求項8] 請求項7に記載の画像判定方法において、
前記第1の機械学習モデルはR-CNNを含む、
ことを特徴とする画像判定方法。
- [請求項9] 請求項7又は8に記載の画像判定方法において、
所与のベース画像に複数のオブジェクトの一部又は全部を重畳させることで生成されるトレーニング画像により、前記第1機械学習モデルの学習を実行するステップをさらに含む、
ことを特徴とする画像判定方法。
- [請求項10] 請求項7乃至9のいずれかに記載の画像判定方法において、

前記第2機械学習モデルは全結合層を含む、
ことを特徴とする画像判定方法。

[請求項11]

請求項10に記載の画像判定方法において、

前記第2機械学習モデルは、前記全結合層の上流側に、前記第1オブジェクトデータ及び前記第2オブジェクトデータに基づく入力データの次元数を削減する畳込み層及びプーリング層を含む、
ことを特徴とする画像判定方法。

[請求項12]

請求項10又は11に記載の画像判定方法において、

所定配置規則に従って、同一又は類似の第1及び第2のベース画像のそれぞれに所与のオブジェクトを重畳させることで第1トレーニング画像及び第2トレーニング画像を生成するステップと、

前記第1トレーニング画像及び前記第2トレーニング画像のそれぞれを前記第1機械学習モデルに入力することにより、前記第1トレーニング画像中のオブジェクトの属性及び配置を示す第1学習オブジェクトデータ、及び前記第2トレーニング画像中のオブジェクトの属性及び配置を示す第2学習オブジェクトデータを取得するステップと、

前記第1学習オブジェクトデータ及び前記第2学習オブジェクトデータにより前記第2機械学習モデルの学習を実行するステップと、
をさらに含むことを特徴とする画像判定方法。

[請求項13]

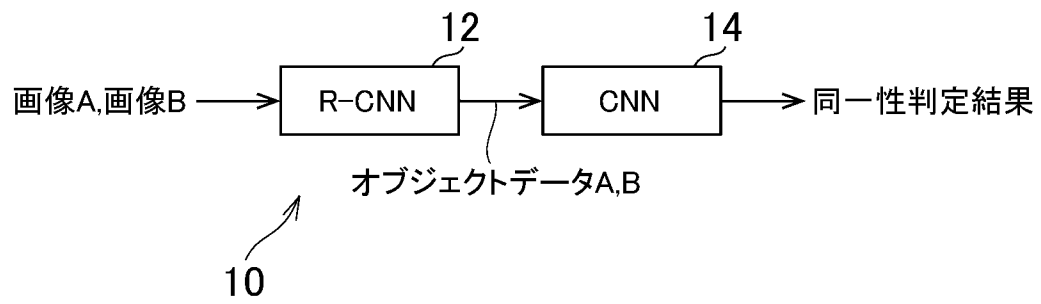
第1の機械学習モデルを用いて、第1画像から該第1画像中のオブジェクトの属性及び配置を示す第1オブジェクトデータを取得するステップと、

前記第1の機械学習モデルを用いて、第2画像から該第2画像中のオブジェクトの属性及び配置を示す第2オブジェクトデータを取得するステップと、

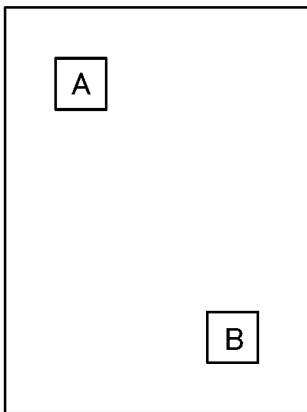
第2の機械学習モデルを用いて、前記第1オブジェクトデータ及び前記第2オブジェクトデータから、前記第1画像及び前記第2画像の実質同一性を判定するステップと、

をコンピュータに実行させるためのプログラム。

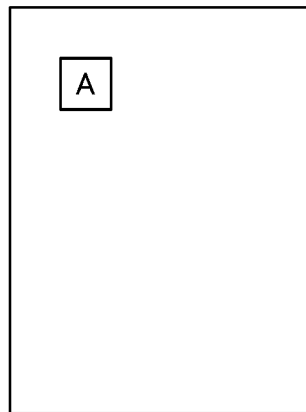
[図1]



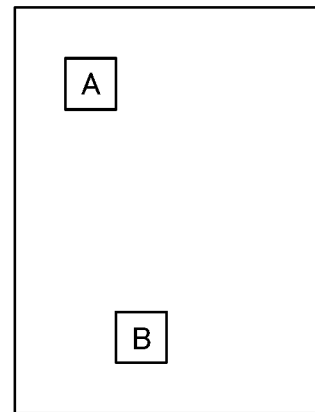
[図2]



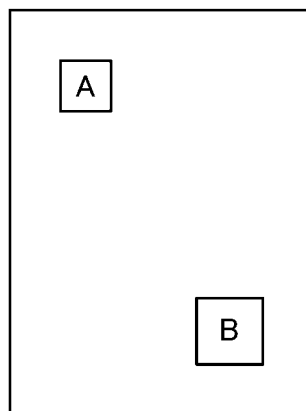
(a)



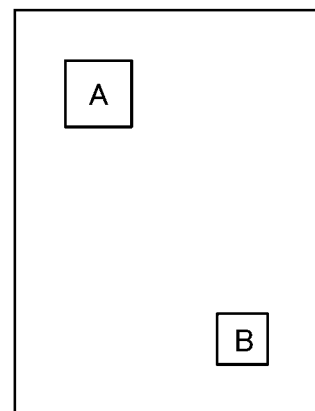
(b)



(c)

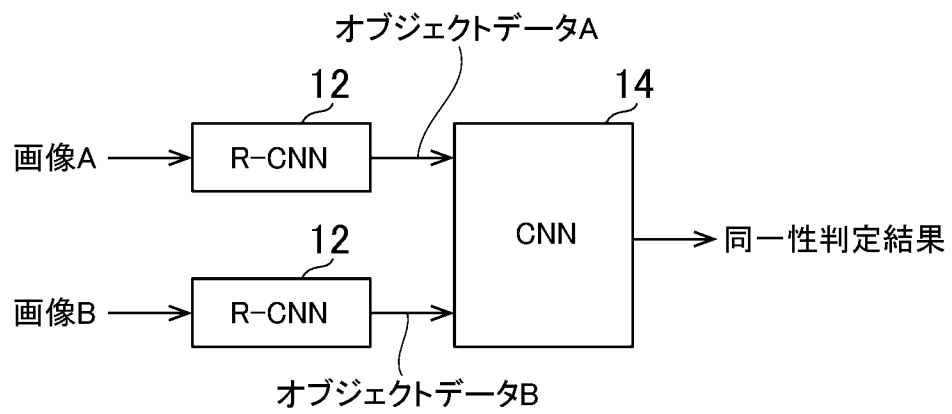


(d)

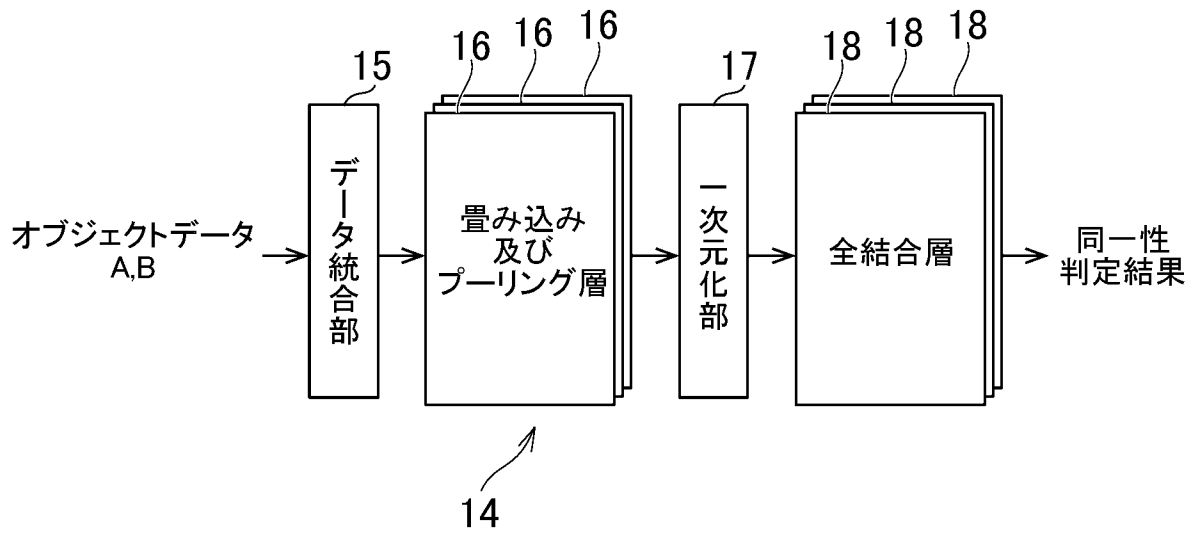


(e)

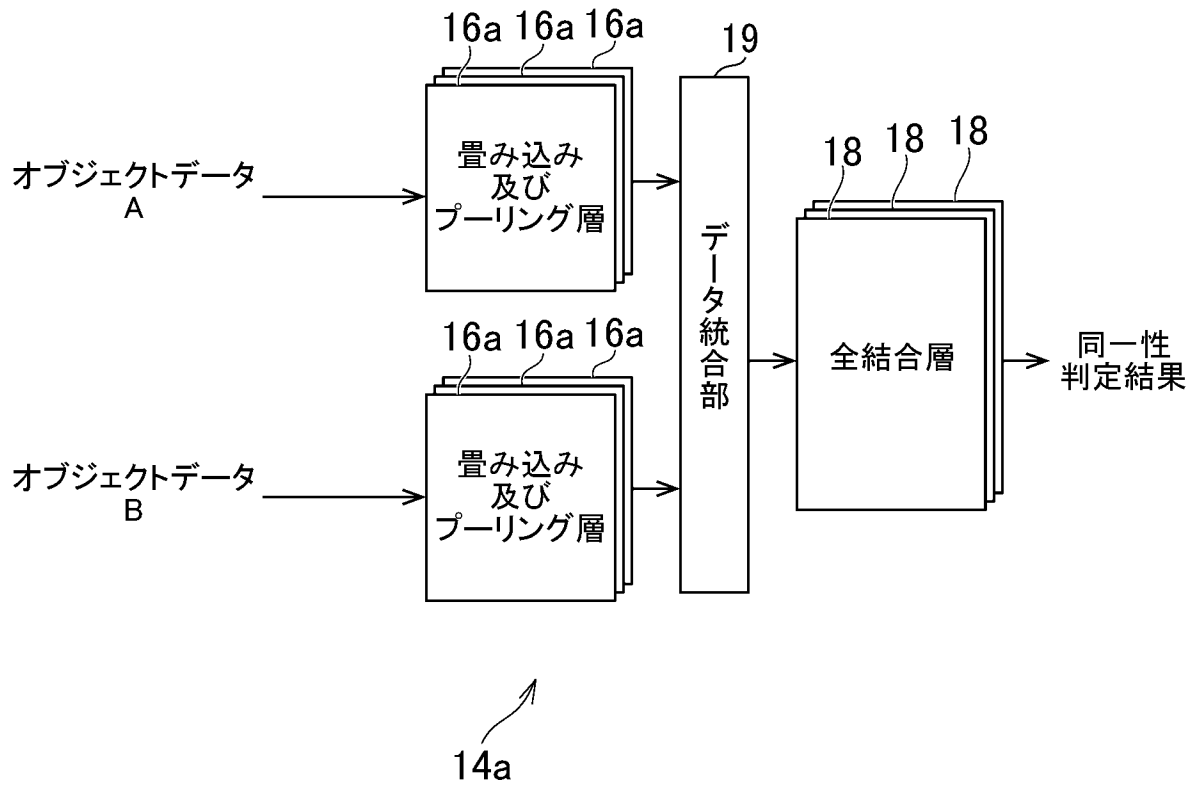
[図3]



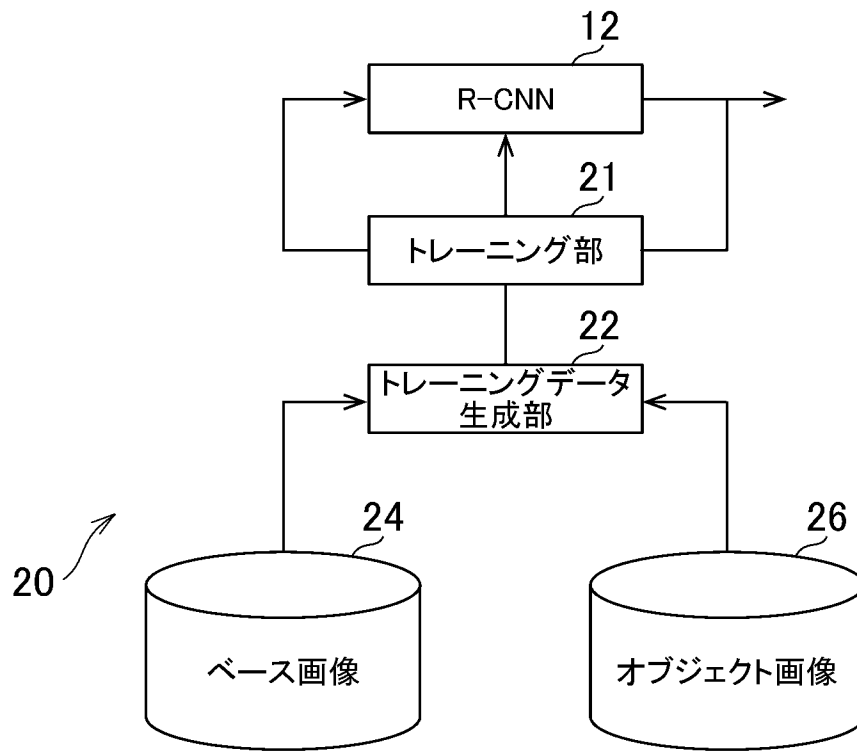
[図4]



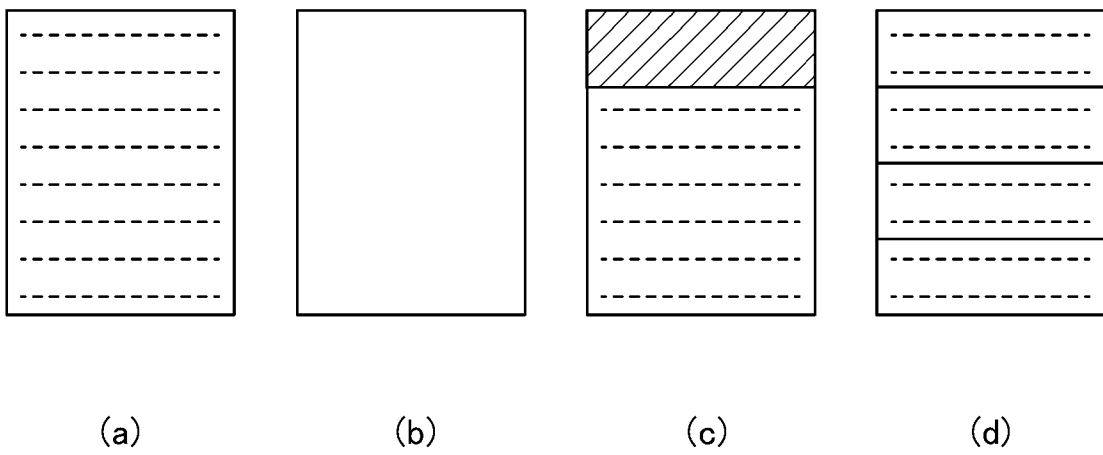
[図5]



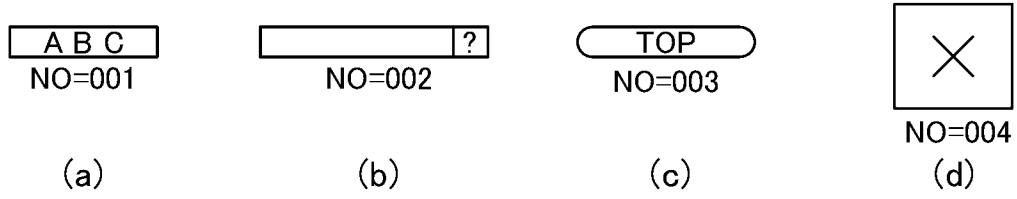
[図6]



[図7]



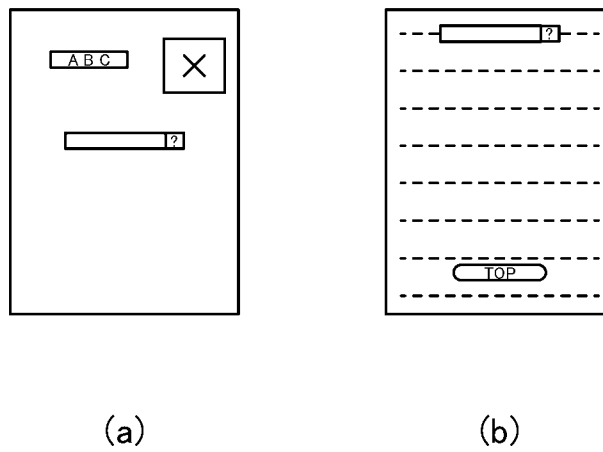
[図8]



[図9]

オブジェクト番号	種類	色
001	ロゴ	赤・白
002	フォーム	白・黒
003	ボタン	緑
004	マーク	黄

[図10]



[図11]

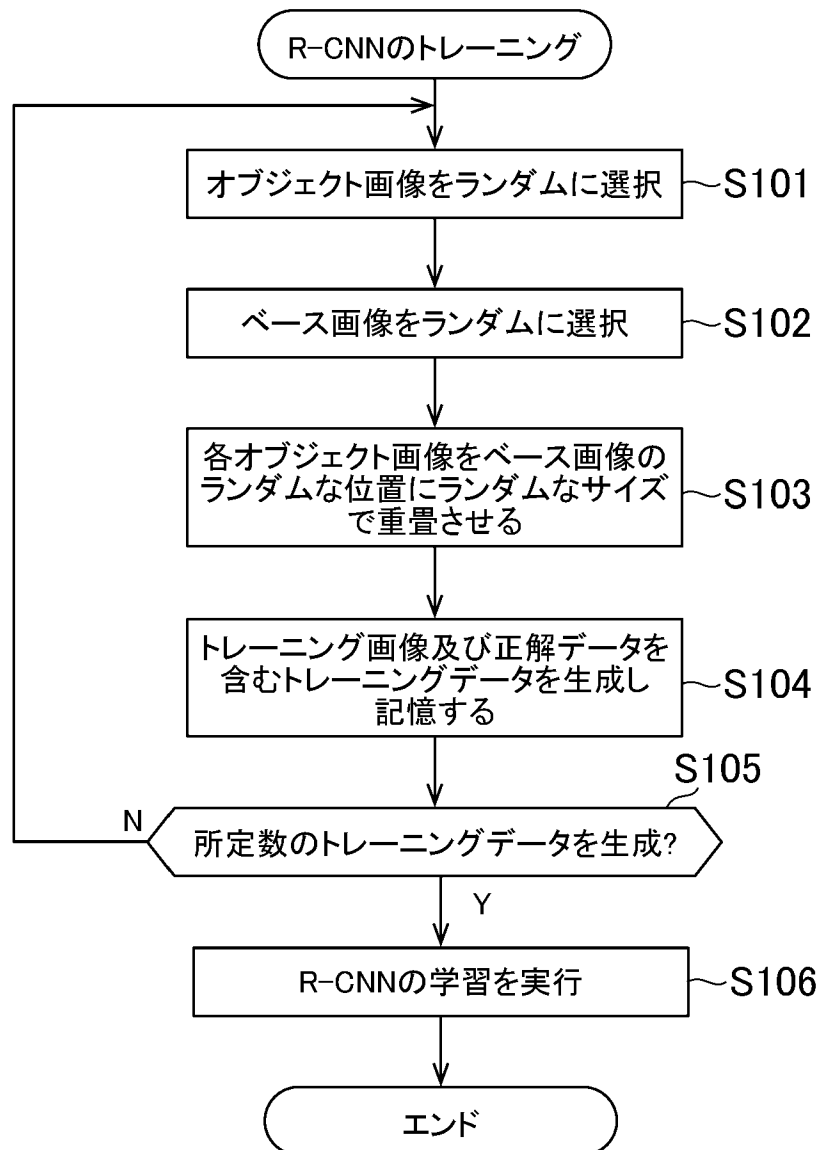
(a)

1	$(XA_1, YA_1)-(XB_1, YB_1)$ ロゴ, 赤・白
2	$(XA_2, YA_2)-(XB_2, YB_2)$ マーク, 黄
3	$(XA_3, YA_3)-(XB_3, YB_3)$ フォーム, 白・黒

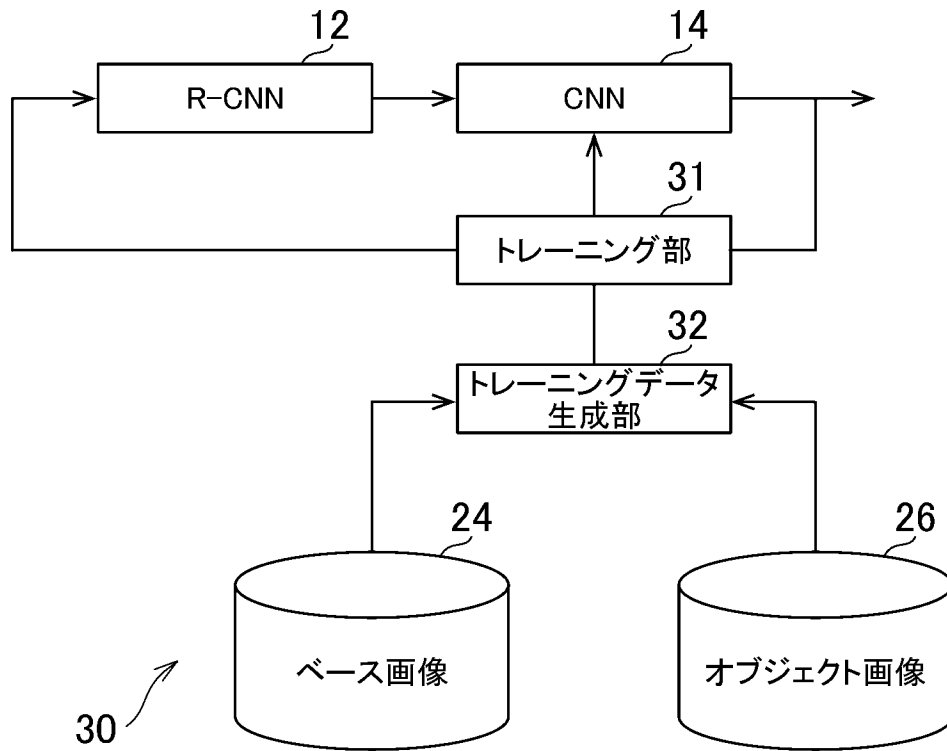
(b)

1	$(XA_1, YA_1)-(XB_1, YB_1)$ フォーム, 白・黒
2	$(XA_2, YA_2)-(XB_2, YB_2)$ ボタン, 緑

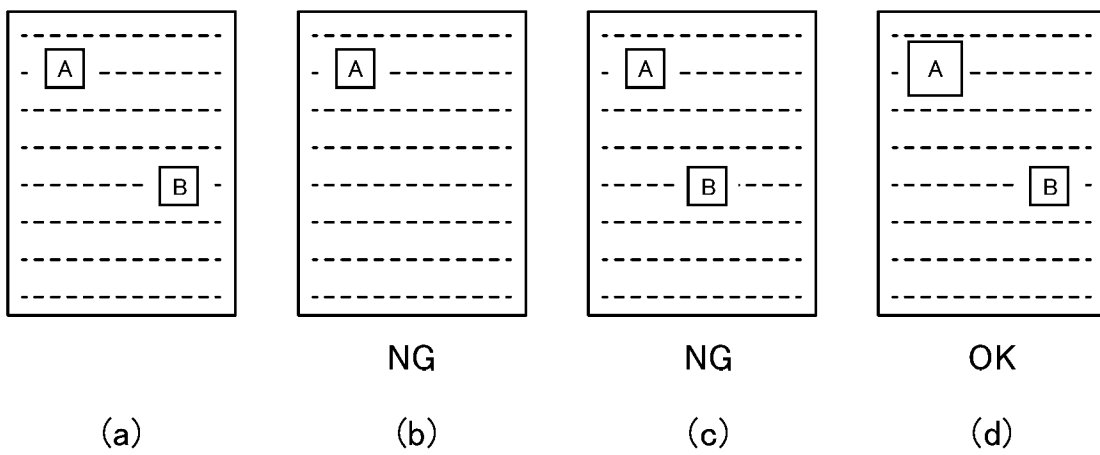
[図12]



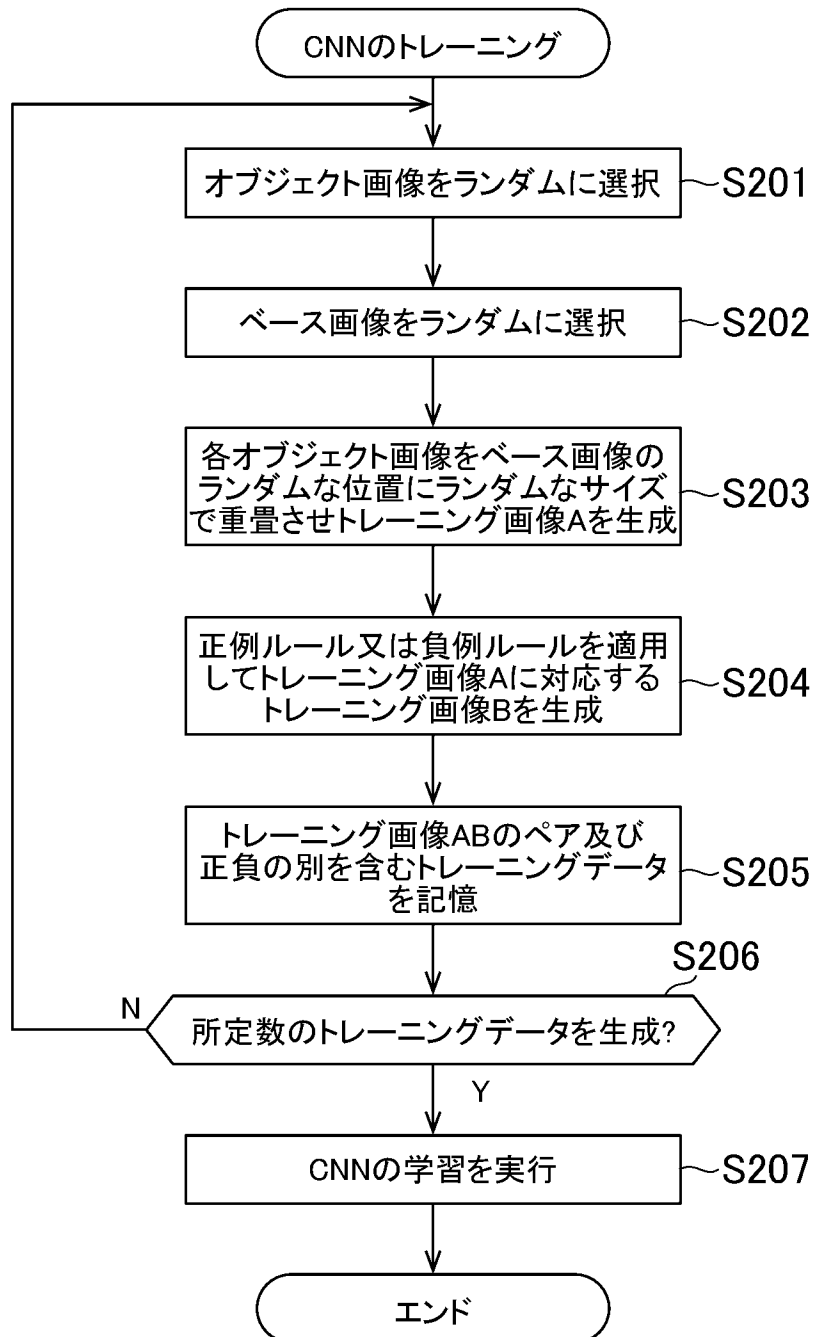
[図13]



[図14]



[図15]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2020/033921

A. CLASSIFICATION OF SUBJECT MATTER

Int. Cl. G06T7/00 (2017.01) i, G06N20/00 (2019.01) i
 FI: G06T7/00 350C, G06N20/00

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, G06N20/00

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 1971-2020
 Registered utility model specifications of Japan 1996-2020
 Published registered utility model applications of Japan 1994-2020

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 Y	JP 2019-219766 A (LIXIL CORP.) 26 December 2019, paragraphs [0067]-[0079], fig. 2	1-2, 7-8, 13 3-6, 9-12
Y	JP 2018-169690 A (NIPPON TELEGRAPH AND TELEPHONE CORP.) 01 November 2018, paragraph [0075]	3-6, 9-12
Y	JP 2020-107185 A (KDDI CORP.) 09 July 2020, paragraph [0021], fig. 1	4-6, 10-12
A	JP 2020-88416 A (SHARP CORP.) 04 June 2020, paragraphs [0013]-[0018]	1-13
A	US 2019/0073560 A1 (MATEI et al.) 07 March 2019, paragraph [0066], fig. 4	1-13

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

- “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
21.10.2020

Date of mailing of the international search report
02.11.2020

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

Authorized officer

 Telephone No.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/JP2020/033921

Patent Documents referred to in the Report	Publication Date	Patent Family	Publication Date
JP 2019-219766 A	26.12.2019	(Family: none)	
JP 2018-169690 A	01.11.2018	(Family: none)	
JP 2020-107185 A	09.07.2020	(Family: none)	
JP 2020-88416 A	04.06.2020	(Family: none)	
US 2019/0073560 A1	07.03.2019	(Family: none)	

A. 発明の属する分野の分類（国際特許分類（IPC）） G06T 7/00(2017.01)i; G06N 20/00(2019.01)i FI: G06T7/00 350C; G06N20/00		
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）） G06T7/00; G06N20/00 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 1971-2020年 日本国実用新案登録公報 1996-2020年 日本国登録実用新案公報 1994-2020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	JP 2019-219766 A (株式会社 LIXIL) 26.12.2019 (2019-12-26) 段落[0067]-[0079], 第2図	1-2, 7-8, 13
Y		3-6, 9-12
Y	JP 2018-169690 A (日本電信電話株式会社) 01.11.2018 (2018-11-01) 段落[0075]	3-6, 9-12
Y	JP 2020-107185 A (KDDI株式会社) 09.07.2020 (2020-07-09) 段落[0021], 第1図	4-6, 10-12
A	JP 2020-88416 A (シャープ株式会社) 04.06.2020 (2020-06-04) 段落[0013]-[0018]	1-13
A	US 2019/0073560 A1 (MATEI et al.) 07.03.2019 (2019-03-07) 段落[0066], 第4図	1-13
<input type="checkbox"/> C欄の続きにも文献が列挙されている。 <input checked="" type="checkbox"/> 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 21.10.2020	国際調査報告の発送日 02.11.2020	
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） 千葉 久博 5H 3991 電話番号 03-3581-1101 内線 3531	

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/JP2020/033921

引用文献	公表日	パテントファミリー文献	公表日
JP 2019-219766 A	26.12.2019	(ファミリーなし)	
JP 2018-169690 A	01.11.2018	(ファミリーなし)	
JP 2020-107185 A	09.07.2020	(ファミリーなし)	
JP 2020-88416 A	04.06.2020	(ファミリーなし)	
US 2019/0073560 A1	07.03.2019	(ファミリーなし)	