

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7635170号
(P7635170)

(45)発行日 令和7年2月25日(2025.2.25)

(24)登録日 令和7年2月14日(2025.2.14)

(51)国際特許分類

F I

G 0 6 T 7/00 (2017.01)

G 0 6 T 7/00 3 5 0 C

G 0 6 N 3/10 (2006.01)

G 0 6 T 7/00 3 0 0 F

G 0 6 N 3/10

請求項の数 21 (全34頁)

(21)出願番号	特願2022-19857(P2022-19857)	(73)特許権者	000003078
(22)出願日	令和4年2月10日(2022.2.10)		株式会社東芝
(65)公開番号	特開2023-117247(P2023-117247 A)	(74)代理人	東京都港区芝浦一丁目1番1号
(43)公開日	令和5年8月23日(2023.8.23)		110003708
審査請求日	令和6年3月1日(2024.3.1)		弁理士法人鈴榮特許総合事務所
		(72)発明者	井田 孝
			東京都港区芝浦一丁目1番1号 株式会
			社東芝内
		(72)発明者	小野 利幸
			東京都港区芝浦一丁目1番1号 株式会
			社東芝内
		審査官	吉田 千裕

最終頁に続く

(54)【発明の名称】 画像処理装置、学習装置、推論装置、および画像処理方法

(57)【特許請求の範囲】

【請求項1】

入力画像に基づくN個(N≧3)の処理画像について、ニューラルネットワークを用いた特徴量抽出処理を行うことによってN個の特徴量を生成する特徴量抽出部と、
前記特徴量抽出処理の過程で発生する処理過程データを保持するメモリと、
前記N個の特徴量のうちの2個以上N-1個以下であるM個の組み合わせで2回以上の比較を行うことによって最大特徴量を選択する最大特徴量選択部と、
前記2回以上の比較毎に、選択されなかったM-1個以下の特徴量に対応するM-1個以下の処理過程データを前記メモリから解放させる最適化部と
を具備する、画像処理装置。

10

【請求項2】

前記特徴量抽出部は、前記入力画像の一部を切り出すことによって前記N個の処理画像を生成する、
請求項1に記載の画像処理装置。

【請求項3】

前記特徴量抽出部は、前記N個の特徴量を生成する際に、同じパラメータの前記ニューラルネットワークを用いて前記特徴量抽出処理を行う、
請求項1または請求項2に記載の画像処理装置。

【請求項4】

前記特徴量抽出部は、前記入力画像を縮小させることによって縮小率の異なる前記N個

20

の処理画像を生成する、

請求項 1 に記載の画像処理装置。

【請求項 5】

前記特徴量抽出部は、前記 N 個の特徴量を生成する際に、それぞれ異なるパラメータの前記ニューラルネットワークを用いて前記特徴量抽出処理を行う、

請求項 4 に記載の画像処理装置。

【請求項 6】

前記 N 個の処理画像は、前記入力画像に対して畳み込み処理を行って生成されうる中間画像を縦横 1 画素以上の N 個のブロックに分解した前記 N 個のブロックにそれぞれ対応する前記入力画像における N 個の領域に相当し、

10

前記特徴量抽出処理は、前記畳み込み処理であり、

前記特徴量抽出部は、前記 N 個の領域について、前記畳み込み処理を行うことによって前記 N 個のブロックのそれぞれに対応する前記 N 個の特徴量を生成する、

請求項 1 に記載の画像処理装置。

【請求項 7】

前記特徴量抽出部は、前記 N 個のブロックのそれぞれについて、加算、平均、または最大値を選択することによって前記 N 個の特徴量を生成する、

請求項 6 に記載の画像処理装置。

【請求項 8】

前記 N 個の特徴量のそれぞれは、複数の要素を持つベクトルである、

20

請求項 1 から請求項 5 までのいずれか一項に記載の画像処理装置。

【請求項 9】

前記最大特徴量選択部は、前記 2 回以上の比較のそれぞれにおいて、前記複数の要素毎に比較を行うことによって大きい要素を選択し、

前記最大特徴量は、前記 N 個の特徴量のうちの最も大きい要素を組み合わせたベクトルに相当する、

請求項 8 に記載の画像処理装置。

【請求項 10】

前記最適化部は、更に、前記最大特徴量を選択するための比較毎に、選択されなかった要素に対応する処理過程データを前記メモリから解放させる、

30

請求項 9 に記載の画像処理装置。

【請求項 11】

前記特徴量抽出部は、前記 N 個の特徴量をそれぞれシーケンシャルに、または複数の特徴量毎に生成する、

請求項 1 から請求項 10 までのいずれか一項に記載の画像処理装置。

【請求項 12】

前記メモリは、M 個の特徴量に対応する M 個の処理過程データを上限として保持する、

請求項 1 から請求項 11 までのいずれか一項に記載の画像処理装置。

【請求項 13】

前記最大特徴量選択部は、前記 2 回以上の比較において、前記 M 個の組み合わせの個数が異なる比較を含む、

40

請求項 1 から請求項 12 までのいずれか一項に記載の画像処理装置。

【請求項 14】

前記 N 個の処理画像は、第 1 の処理画像と、第 2 の処理画像と、第 3 の処理画像とを含み、

前記特徴量抽出処理は、第 1 の抽出処理と、第 2 の抽出処理と、第 3 の抽出処理とを含み、

前記 N 個の特徴量は、第 1 の特徴量と、第 2 の特徴量と、第 3 の特徴量とを含み、

前記特徴量抽出部は、第 1 の抽出部と、第 2 の抽出部と、第 3 の抽出部とを含み、

前記最大特徴量選択部は、第 1 の選択部と、第 2 の選択部とを含み、

50

前記第 1 の抽出部は、前記第 1 の処理画像について前記第 1 の抽出処理を行うことによって前記第 1 の特徴量を生成し、

前記メモリは、前記第 1 の抽出処理の過程で発生する第 1 の処理過程データを保持し、

前記第 2 の抽出部は、前記第 2 の処理画像について前記第 2 の抽出処理を行うことによって前記第 2 の特徴量を生成し、

前記メモリは、前記第 2 の抽出処理の過程で発生する第 2 の処理過程データを保持し、

前記第 1 の選択部は、前記第 1 の特徴量と前記第 2 の特徴量とを比較することによって大きい方を第 1 の選択特徴量として選択し、

前記最適化部は、前記第 1 の選択部による選択において、選択されなかった特徴量に対応する処理過程データを前記メモリから解放させ、

前記第 3 の抽出部は、前記第 3 の処理画像について前記第 3 の抽出処理を行うことによって前記第 3 の特徴量を生成し、

前記メモリは、前記第 3 の抽出処理の過程で発生する第 3 の処理過程データを保持し、

前記第 2 の選択部は、前記第 1 の選択特徴量と前記第 3 の特徴量とを比較することによって大きい方を第 2 の選択特徴量として選択し、

前記最適化部は、前記第 2 の選択部による選択において、選択されなかった特徴量に対応する処理過程データを前記メモリから解放させる、

請求項 1 に記載の画像処理装置。

【請求項 15】

前記 N は 4 以上であり、

前記 N 個の処理画像は、第 1 の処理画像と、第 2 の処理画像と、第 3 の処理画像と、第 4 の処理画像とを含み、

前記特徴量抽出処理は、第 1 の抽出処理と、第 2 の抽出処理と、第 3 の抽出処理と、第 4 の抽出処理とを含み、

前記 N 個の特徴量は、第 1 の特徴量と、第 2 の特徴量と、第 3 の特徴量と、第 4 の特徴量とを含み、

前記特徴量抽出部は、第 1 の抽出部と、第 2 の抽出部と、第 3 の抽出部と、第 4 の抽出部とを含み、

前記最大特徴量選択部は、第 1 の選択部と、第 2 の選択部と、第 3 の選択部とを含み、

前記第 1 の抽出部は、前記第 1 の処理画像について前記第 1 の抽出処理を行うことによって前記第 1 の特徴量を生成し、

前記メモリは、前記第 1 の抽出処理の過程で発生する第 1 の処理過程データを保持し、

前記第 2 の抽出部は、前記第 2 の処理画像について前記第 2 の抽出処理を行うことによって前記第 2 の特徴量を生成し、

前記メモリは、前記第 2 の抽出処理の過程で発生する第 2 の処理過程データを保持し、

前記第 1 の選択部は、前記第 1 の特徴量と前記第 2 の特徴量とを比較することによって大きい方を第 1 の選択特徴量として選択し、

前記最適化部は、前記第 1 の選択部による選択において、選択されなかった特徴量に対応する処理過程データを前記メモリから解放させ、

前記第 3 の抽出部は、前記第 3 の処理画像について前記第 3 の抽出処理を行うことによって前記第 3 の特徴量を生成し、

前記メモリは、前記第 3 の抽出処理の過程で発生する第 3 の処理過程データを保持し、

前記第 4 の抽出部は、前記第 4 の処理画像について前記第 4 の抽出処理を行うことによって前記第 4 の特徴量を生成し、

前記メモリは、前記第 4 の抽出処理の過程で発生する第 4 の処理過程データを保持し、

前記第 2 の選択部は、前記第 3 の特徴量と前記第 4 の特徴量とを比較することによって大きい方を第 2 の選択特徴量として選択し、

前記最適化部は、前記第 2 の選択部による選択において、選択されなかった特徴量に対応する処理過程データを前記メモリから解放させ、

前記第 3 の選択部は、前記第 1 の選択特徴量と前記第 2 の選択特徴量とを比較すること

10

20

30

40

50

によって大きい方を第３の選択特徴量として選択し、

前記最適化部は、前記第３の選択部による選択において、選択されなかった特徴量に対応する処理過程データを前記メモリから解放させる、

請求項１に記載の画像処理装置。

【請求項１６】

請求項１から請求項１５までのいずれか一項に記載の画像処理装置と、

前記最大特徴量と前記入力画像に対応する正解特徴量とに基づいて誤差値を算出する誤差算出部と、

前記メモリが最終的に保持している前記最大特徴量に関する処理過程データと前記誤差値とに基づいて前記ニューラルネットワークを学習する学習部と

10

を具備する、学習装置。

【請求項１７】

請求項１から請求項１５までのいずれか一項に記載の画像処理装置と、

前記最大特徴量に基づいて前記入力画像において認識対象の物体が存在しているか否かを表す推論結果を出力する出力部と

を具備する、推論装置。

【請求項１８】

前記最大特徴量選択部は、前記２回以上の比較毎に、前記Ｍ個の特徴量それぞれとしきい値とを更に比較し、

前記最適化部は、前記２回以上の比較毎に、前記選択されなかったＭ－１個以下の特徴量うち、前記しきい値以上の特徴量に対応する処理過程データを前記メモリから解放させない、

20

請求項１７に記載の推論装置。

【請求項１９】

前記出力部は、前記最大特徴量と前記しきい値以上の特徴量とに基づいて前記推論結果を出力する、

請求項１８に記載の推論装置。

【請求項２０】

前記処理過程データは、前記推論結果を可視化した推論画像であり、

前記出力部は、更に、前記入力画像と前記推論画像とに基づいて前記入力画像に写っている前記認識対象の物体の画素を強調させた画像を出力する、

30

請求項１７から請求項１９までのいずれか一項に記載の推論装置。

【請求項２１】

入力画像に基づくＮ個（ $N \geq 3$ ）の処理画像について、ニューラルネットワークを用いた特徴量抽出処理を行うことによってＮ個の特徴量を生成することと、

前記特徴量抽出処理の過程で発生する処理過程データをメモリに保持することと、

前記Ｎ個の特徴量のうちの２個以上 $N - 1$ 個以下であるＭ個の組み合わせで２回以上の比較を行うことによって最大特徴量を選択することと、

前記２回以上の比較毎に、選択されなかった $M - 1$ 個以下の特徴量に対応する $M - 1$ 個以下の処理過程データを前記メモリから解放させることと

40

を具備する、画像処理方法。

【発明の詳細な説明】

【技術分野】

【０００１】

本発明の実施形態は、画像処理装置、学習装置、推論装置、および画像処理方法に関する。

【背景技術】

【０００２】

製造工場における製造品の撮影画像を用いた外観検査や、Ｘ線透視画像やＣＴ画像などの医用画像を用いた医療診断などにおいて、異常の有無の認識にニューラルネットワーク

50

を用いると、他の画像処理を用いるよりも一般的に高い認識精度が得られることが知られている。また、このような外観検査や医療診断においては、認識対象画像の全体に対して異常が写っているのはごく一部分であることが多い。そのため、認識対象画像を複数の処理画像に分割して、分割した複数の処理画像をそれぞれ個別にニューラルネットワークで処理する技術が知られている。この技術を用いることにより、個別のニューラルネットワークの処理量は、認識対象画像をそのまま処理する場合の処理量よりも少なくできる。

【0003】

ところで、上記技術におけるニューラルネットワークの学習には、理想的には、複数の処理画像それぞれに異常の有無を正解値として教示する手法が望ましい。しかし、この手法では、認識対象画像に対して異常の有無を教示するよりも、分割する処理画像の数に比例して正解値のデータを作成する手間が多くかかるという問題がある。この問題に対して、複数の処理画像を個別にニューラルネットワークで処理して得られる各出力の最大値を求め、この最大値から得られる異常の有無の推定値に対して、分割前の認識対象画像における異常の有無を正解値として教示することによって個別ニューラルネットワークを学習する技術が知られている。

10

【0004】

上記技術で用いられているニューラルネットワークの学習過程では、複数の処理画像のそれぞれをニューラルネットワークで処理し、各々の処理過程における変換画像の画素値やニューラルネットワークのその時点の重みパラメータなどの処理過程データを全てメモリに保持している。そして、ニューラルネットワークの出力として得られる推定値の正解値に対する誤差と、学習に寄与する処理過程データを用いて、逆伝播の処理によりニューラルネットワークの重みパラメータを更新した後、全ての処理過程データを解放する構成である。

20

【0005】

しかしながら、上記構成では、逆伝播の処理を行うまでの間に、全ての処理過程データを保持する必要があるため、メモリ容量を低減できないという問題があった。

【先行技術文献】

【非特許文献】

【0006】

【文献】Maximilian Ilse, 外2名, "Attention - based Deep Multiple Instance Learning", Volume 80 : International Conference on Machine Learning, (スウェーデン), 2018, PMLR 80:2127-2136.

30

【発明の概要】

【発明が解決しようとする課題】

【0007】

本発明が解決しようとする課題は、ニューラルネットワークを用いた画像処理に必要なメモリ容量を低減することができる画像処理装置、学習装置、推論装置、および方法を提供することである。

【課題を解決するための手段】

40

【0008】

一実施形態に係る画像処理装置は、特徴量抽出部と、メモリと、最大特徴量選択部と、最適化部とを備える。特徴量抽出部は、入力画像に基づくN個(N≧3)の処理画像について、ニューラルネットワークを用いた特徴量抽出処理を行うことによってN個の特徴量を生成する。メモリは、特徴量抽出処理の過程で発生する処理過程データを保持する。最大特徴量選択部は、N個の特徴量のうちの2個以上N-1個以下であるM個の組み合わせで2回以上の比較を行うことによって最大特徴量を選択する。最適化部は、2回以上の比較毎に、選択されなかったM-1個以下の特徴量に対応するM-1個以下の処理過程データをメモリから解放させる。

【図面の簡単な説明】

50

【 0 0 0 9 】

【図 1】第 1 の実施形態に係る画像処理装置を含む学習装置の構成を例示するブロック図。

【図 2】入力画像を切り出すことによって 3 つの処理画像に分割する例を示す説明図。

【図 3】図 1 の画像処理装置の詳細な構成を例示するブロック図。

【図 4】入力画像を切り出すことによって 4 つの処理画像に分割する例を示す説明図。

【図 5】入力画像を重複して切り出すことによって 4 つの処理画像に分割する例を示す説明図。

【図 6】図 3 の画像処理装置における特徴量抽出部および最大特徴量選択部の他の第 1 の構成例を示すブロック図。

【図 7】第 1 の実施形態に係る画像処理装置の動作を例示するフローチャート。

10

【図 8】図 3 の画像処理装置における特徴量抽出部および最大特徴量選択部の他の第 2 の構成例を示すブロック図。

【図 9】入力画像を縮小することによって 2 つの処理画像を生成する例を示す説明図。

【図 10】入力画像に対する 2 回の畳み込み処理における変換画像、中間画像、および受容野を例示する説明図。

【図 11】図 1 の画像処理装置の他の構成例を示すブロック図。

【図 12】第 1 の実施形態に係る画像処理装置の他の動作を例示するフローチャート。

【図 13】図 6 の画像処理装置における特徴量抽出部および最大特徴量選択部の他の第 3 の構成例を示すブロック図。

【図 14】入力画像に対する畳み込み処理における複数の変換画像、中間画像、および受容野を例示する説明図。

20

【図 15】畳み込み処理の処理単位毎に分割した中間画像。

【図 16】複数のチャンネルを有する中間画像とチャンネル毎の特徴量との関係を例示する説明図。

【図 17】第 2 の実施形態に係る画像処理装置を含む推論装置の構成を例示するブロック図。

【図 18】畳み込み処理済みの部分画像とメモリに保持される部分画像データとの関係を例示する説明図。

【図 19】メモリから解放される部分画像データを例示する説明図。

【図 20】メモリに保持される新たな部分画像データを例示する説明図。

30

【図 21】部分画像から生成された補間中間画像。

【図 22】一実施形態に係るコンピュータのハードウェア構成を例示するブロック図。

【発明を実施するための形態】

【 0 0 1 0 】

以下、図面を参照しながら、画像処理装置を含む学習装置および推論装置に関する実施形態について詳細に説明する。

【 0 0 1 1 】

(第 1 の実施形態)

第 1 の実施形態では、画像に認識対象の物体が含まれているか否かを認識するニューラルネットワークを学習することについて説明される。上記の物体は、例えば、外観検査における製造品のひびや汚れ、或いは医療診断における腫瘍や内出血した血管などを想定する。

40

【 0 0 1 2 】

図 1 は、第 1 の実施形態に係る画像処理装置 110 を含む学習装置 100 の構成を例示するブロック図である。学習装置 100 は、画像処理装置 110 (画像処理部) と、誤差算出部 120 と、学習部 130 とを備える。画像処理装置 110 は、特徴量抽出部 111 と、メモリ 112 と、最大特徴量選択部 113 と、最適化部 114 とを備える。

【 0 0 1 3 】

なお、学習装置 100 は、ニューラルネットワークの学習に必要な入力画像と、この入力画像に対応する正解ラベル (正解値) とを組とした学習データセットを取得する取得部

50

を備えてもよい。また、学習装置 100 は、各部を制御するための制御部を備えてもよい。

【0014】

特徴量抽出部 111 は、他の装置（図示せず）から入力画像を受け取る。特徴量抽出部 111 は、入力画像に基づく N 個（ $N \geq 3$ ）の処理画像について、ニューラルネットワークを用いた特徴量抽出処理を行うことによって N 個の特徴量を生成する。特徴量抽出部 111 は、特徴量抽出処理の過程で発生する処理過程データをメモリ 112 へと出力し、 N 個の特徴量を最大特徴量選択部 113 へと出力する。

【0015】

具体的には、特徴量抽出部 111 は、 N 個の処理画像それぞれに対してシーケンシャルに特徴量抽出処理を行う。即ち、特徴量抽出部 111 は、第 1 の処理画像に対する処理が終了した後、後続する第 2 の処理画像に対する処理を行い、これを第 N の処理画像まで繰り返す。また、特徴量抽出部 111 は、処理画像に対する特徴量抽出処理が行われる度に、処理過程データをメモリ 112 へと出力し、特徴量を最大特徴量選択部 113 へと出力する。

10

【0016】

上記の特徴量抽出処理は、畳み込み処理、活性化処理、全結合処理、およびプーリング処理などを含む。具体的には、特徴量抽出部 111 は、処理画像に対して畳み込み処理および活性化処理などの変換を行った後、全結合処理やプーリング処理などによりスカラー値に変換することによって特徴量を生成する。言い換えると、特徴量抽出部 111 は、処理画像を入力すると特徴量を出力するようなニューラルネットワークによって構成される。また、特徴量抽出部 111 は、 N 個の処理画像それぞれに対応する N 個のニューラルネットワークを有する。尚、 N 個のニューラルネットワークのそれぞれは、個別ニューラルネットワークと呼ばれてもよい。

20

【0017】

なお、特徴量抽出処理は、特徴量を生成する直前に最終活性化処理を行ってもよい。この最終活性化処理は、例えばシグモイド関数を適用することによって「0」から「1」までの値に変換する処理である。特徴量抽出処理において最終活性化処理が行われない場合、後述する最大特徴量選択部 113 において最大特徴量が選択された後から、誤差算出部 120 において誤差算出処理が行われるまでの間で、任意のユニットにおいて最終活性化処理が行われるものとする。また、任意のユニットとして、最終活性化処理を行う活性化部を、最大特徴量選択部 113 と誤差算出部 120 との間に設けてもよい。

30

【0018】

上記の処理過程データは、例えば、変換後の処理画像（変換画像または中間画像）の画素値などのデータ値、変換時の処理に設定されていた重みパラメータの値、および変換時の処理に設定されていたシフトパラメータの値である。この処理過程データは、後述するニューラルネットワークの学習時に利用されるため、学習に必要なデータに言い換えられてもよい。

【0019】

更に、特徴量抽出部 111 は、入力画像に基づいて N 個の処理画像を生成してもよい。例えば、特徴量抽出部 111 は、入力画像の一部を切り出すことによって N 個の処理画像を生成する。入力画像と N 個の処理画像との関係について、図 2 を用いて説明する。

40

【0020】

図 2 は、入力画像 200 を切り出すことによって 3 つの処理画像 210 から 230 までに分割する例を示す説明図である。特徴量抽出部 111 は、入力画像 200 を所定のサイズで切り出すことによって 3 つの処理画像 210 から 230 までを生成する。3 つの処理画像 210 から 230 までは、それぞれ等しいサイズでもよいし、異なるサイズでもよい。

【0021】

なお、入力画像を切り出す処理は、入力画像から特定の領域を選択する処理に言い換えられてもよい。即ち、特徴量抽出部 111 は、入力画像のうちのそれぞれ異なる領域を選択することによって N 個の処理画像を生成する。

50

【 0 0 2 2 】

メモリ 1 1 2 は、特徴量抽出部 1 1 1 から処理過程データを入力しこれを保持する。また、メモリ 1 1 2 は、最適化部 1 1 4 から解放指示情報を入力する。メモリ 1 1 2 は、解放指示情報に従って、保持している複数の処理過程データのうちの不要な処理過程データ（以降、不要データと称する）を解放する。メモリ 1 1 2 は、他の各部における処理画像に関する一連の処理が終了した後、最終的に保持している処理過程データ、即ち最大特徴量に関する処理過程データを学習部 1 3 0 へと出力する。

【 0 0 2 3 】

具体的には、メモリ 1 1 2 は、特徴量抽出部 1 1 1 からシーケンシャルに入力される処理過程データを保持しつつ、同様に最適化部 1 1 4 からシーケンシャルに入力される解放指示情報に従って、不要データを解放する。このような動作により、メモリ 1 1 2 は、常に不要データを解放するため、全ての処理過程データを保持する必要がなくなる。

10

【 0 0 2 4 】

上記の不要データは、後述する選択処理において、選択されなかった特徴量に対応する処理過程データである。選択されなかった特徴量はニューラルネットワークの学習時に考慮されないため、不要データは、ニューラルネットワークの学習に寄与しない処理過程データであると言える。

【 0 0 2 5 】

最大特徴量選択部 1 1 3 は、特徴量抽出部 1 1 1 から N 個の特徴量を入力する。最大特徴量選択部 1 1 3 は、N 個の特徴量のうちの 2 個以上 N - 1 個以下である M 個の組み合わせで 2 回以上の比較を行うことによって最大特徴量を選択する。最大特徴量選択部 1 1 3 は、選択処理により選択されなかった特徴量に関する非選択情報を生成して最適化部 1 1 4 へと出力し、最大特徴量を誤差算出部 1 2 0 へと出力する。

20

【 0 0 2 6 】

具体的には、最大特徴量選択部 1 1 3 は、特徴量抽出部 1 1 1 からシーケンシャルに特徴量を入力し、入力された特徴量が選択処理に必要な個数（上記の M 個）となった後、選択処理を行い、最も大きい特徴量を選択する。その後、最大特徴量選択部 1 1 3 は、再び特徴量抽出部 1 1 1 からシーケンシャルに特徴量を入力し、入力された特徴量が再び選択処理に必要な個数となった後、選択処理を行い、以降これを繰り返す。

【 0 0 2 7 】

選択処理によって選択された特徴量は、後続する選択処理がある場合には、後続する選択処理において再度用いられ、後続する選択処理がない場合には、最大特徴量として誤差算出部 1 2 0 へと出力される。また、最大特徴量選択部 1 1 3 は、選択処理が行われる度に、非選択情報を生成して最適化部 1 1 4 へと出力する。

30

【 0 0 2 8 】

なお、最大特徴量選択部 1 1 3 は、2 回以上の比較において、組み合わせの個数が異なる比較を含んでもよい。例えば、全部で 2 回の比較を行う場合、最大特徴量選択部 1 1 3 は、1 回目の比較において 3 つの特徴量を比較し、後続する 2 回目の比較において 2 つの特徴量を比較してもよい。

【 0 0 2 9 】

最適化部 1 1 4 は、最大特徴量選択部 1 1 3 から非選択情報を入力する。最適化部 1 1 4 は、非選択情報に基づいて解放指示情報を生成し、メモリ 1 1 2 へと出力する。解放指示情報は、メモリ 1 1 2 に保持されている不要データを解放させるための情報である。換言すると、最適化部 1 1 4 は、最大特徴量選択部 1 1 3 における 2 回以上の比較毎（即ち、最大特徴量を選択するための比較毎）に、選択されなかった M - 1 個以下の特徴量に対応する M - 1 個以下の処理過程データをメモリから解放させる。

40

【 0 0 3 0 】

誤差算出部 1 2 0 は、最大特徴量選択部 1 1 3 から最大特徴量を入力し、他の装置から入力画像に対応する正解値（正解特徴量）を入力する。誤差算出部 1 2 0 は、最大特徴量と正解特徴量とに基づいて誤差値を算出する。誤差算出部 1 2 0 は、誤差値を学習部 1 3

50

0 へと出力する。

【0031】

具体的には、誤差算出部120は、最大特徴量と正解特徴量とを比較し、バイナリクロスエントロピーなどに代表される誤差値を算出する。正解特徴量は、例えば、入力画像に認識対象の物体が含まれていれば「1」、含まれていなければ「0」とする値である。

【0032】

学習部130は、メモリ112から最大特徴量に関する処理過程データを入力し、誤差算出部120から誤差値を入力する。学習部130は、最大特徴量に関する処理過程データと誤差値とに基づいて特徴量抽出部111を構成しているニューラルネットワークを学習する。

10

【0033】

具体的には、学習部130は、最大特徴量に関する処理過程データと誤差値とを用いて、誤差逆伝播法により最大特徴量が抽出された個別ニューラルネットワークを学習する。この誤差逆伝播法による学習は、入力画像を入力してから最大特徴量を得るまでの順方向処理におけるデータの繋がりを逆方向にたどることによって、個別ニューラルネットワークにおける種々の処理に設定されていた重みパラメータの値およびシフトパラメータの値を順次更新することによって行われる。このことから、選択されなかった特徴量に対応する個別ニューラルネットワークは、データの繋がりが途中で途切れてしまっているため、学習対象にはならない。即ち、選択されなかった特徴量に関する処理過程データは、最大特徴量が抽出された個別ニューラルネットワークの学習には寄与しないものである。

20

【0034】

以上、第1の実施形態に係る画像処理装置110を含む学習装置100の構成について説明した。次に、画像処理装置110の詳細な構成について、図3を用いて説明する。尚、図3の画像処理装置110では、図2に示すような3つの処理画像を用いることを想定する。

【0035】

図3は、画像処理装置110の詳細な構成を例示するブロック図である。図3の特徴量抽出部111は、処理画像生成部310と、第1の抽出部320-1と、第2の抽出部320-2と、第3の抽出部320-3とを備える。図3の最大特徴量選択部113は、第1の選択部330-1と、第2の選択部330-2とを備える。

30

【0036】

処理画像生成部310は、入力画像に基づいて3つの処理画像を生成する。処理画像生成部310は、3つの処理画像のうちの第1の処理画像を第1の抽出部320-1へと出力し、第2の処理画像を第2の抽出部320-2へと出力し、第3の処理画像を第3の抽出部320-3へと出力する。

【0037】

第1の抽出部320-1は、処理画像生成部310から第1の処理画像を入力する。第1の抽出部320-1は、第1の処理画像について、特徴量抽出処理に相当する第1の抽出処理を行うことによって第1の特徴量を生成する。第1の抽出部320-1は、第1の特徴量を第1の選択部330-1へと出力し、第1の抽出処理の過程で発生する第1の処理過程データをメモリ112へと出力する。

40

【0038】

第1の特徴量が抽出された後、メモリ112は、第1の抽出部320-1から第1の処理過程データを入力し、これを保持する。この時点において、メモリ112は、1つの処理過程データを保持している。

【0039】

第2の抽出部320-2は、処理画像生成部310から第2の処理画像を入力する。第2の抽出部320-2は、第2の処理画像について、特徴量抽出処理に相当する第2の抽出処理を行うことによって第2の特徴量を生成する。第2の抽出部320-2は、第2の特徴量を第1の選択部330-1へと出力し、第2の抽出処理の過程で発生する第2の処

50

理過程データをメモリ 1 1 2 へと出力する。

【 0 0 4 0 】

第 2 の特徴量が抽出された後、メモリ 1 1 2 は、第 2 の抽出部 3 2 0 - 2 から第 2 の処理過程データを入力し、これを保持する。この時点において、メモリ 1 1 2 は、2 つの処理過程データを保持している。

【 0 0 4 1 】

第 1 の選択部 3 3 0 - 1 は、第 1 の抽出部 3 2 0 - 1 から第 1 の特徴量を入力し、第 2 の抽出部 3 2 0 - 2 から第 2 の特徴量を入力する。第 1 の選択部 3 3 0 - 1 は、第 1 の特徴量と第 2 の特徴量とを比較することによって大きい方を第 1 の選択特徴量として選択する。第 1 の選択部 3 3 0 - 1 は、選択されなかった特徴量に関する第 1 の非選択情報を生成して最適化部 1 1 4 へと出力し、第 1 の選択特徴量を第 2 の選択部 3 3 0 - 2 へと出力する。

10

【 0 0 4 2 】

第 1 の非選択情報が生成された後、最適化部 1 1 4 は、第 1 の選択部 3 3 0 - 1 から第 1 の非選択情報を入力する。最適化部 1 1 4 は、第 1 の非選択情報に基づいて第 1 の解放指示情報を生成し、メモリ 1 1 2 へと出力する。

【 0 0 4 3 】

第 1 の解放指示情報が生成された後、メモリ 1 1 2 は、第 1 の解放指示情報を入力する。メモリ 1 1 2 は、第 1 の解放指示情報に従って、保持している 2 つの処理過程データのうちの不要データを解放する。この時点において、メモリ 1 1 2 は、1 つの処理過程データを保持している。

20

【 0 0 4 4 】

第 3 の抽出部 3 2 0 - 3 は、処理画像生成部 3 1 0 から第 3 の処理画像を入力する。第 3 の抽出部 3 2 0 - 3 は、第 3 の処理画像について、特徴量抽出処理に相当する第 3 の抽出処理を行うことによって第 3 の特徴量を生成する。第 3 の抽出部 3 2 0 - 3 は、第 3 の特徴量を第 2 の選択部 3 3 0 - 2 へと出力し、第 3 の抽出処理の過程で発生する第 3 の処理過程データをメモリ 1 1 2 へと出力する。

【 0 0 4 5 】

なお、第 3 の抽出部 3 2 0 - 3 における特徴量抽出処理は、第 3 の処理過程データをメモリ 1 1 2 へ出力する際に、メモリ 1 1 2 に 1 つの処理過程データしか保持されていないタイミングで行われる。または、第 3 の抽出部 3 2 0 - 3 における特徴量抽出処理は、メモリ 1 1 2 において 1 つの処理過程データを保持している状態で行われる。

30

【 0 0 4 6 】

第 3 の特徴量が抽出された後、メモリ 1 1 2 は、第 3 の抽出部 3 2 0 - 3 から第 3 の処理過程データを入力し、これを保持する。この時点において、メモリ 1 1 2 は、2 つの処理過程データを保持している。

【 0 0 4 7 】

第 2 の選択部 3 3 0 - 2 は、第 1 の選択部 3 3 0 - 1 から第 1 の選択特徴量を入力し、第 3 の抽出部 3 2 0 - 3 から第 3 の特徴量を入力する。第 2 の選択部 3 3 0 - 2 は、第 1 の選択特徴量と第 3 の選択特徴量とを比較することによって大きい方を第 2 の選択特徴量として選択する。第 2 の選択部 3 3 0 - 2 は、選択されなかった特徴量に関する第 2 の非選択情報を生成して最適化部 1 1 4 へと出力し、第 2 の選択特徴量を最大特徴量として誤差算出部 1 2 0 へと出力する。

40

【 0 0 4 8 】

第 2 の非選択情報が生成された後、最適化部 1 1 4 は、第 2 の選択部 3 3 0 - 2 から第 2 の非選択情報を入力する。最適化部 1 1 4 は、第 2 の非選択情報に基づいて第 2 の解放指示情報を生成し、メモリ 1 1 2 へと出力する。

【 0 0 4 9 】

第 2 の解放指示情報が生成された後、メモリ 1 1 2 は、第 2 の解放指示情報を入力する。メモリ 1 1 2 は、第 2 の解放指示情報に従って、保持している 2 つの処理過程データの

50

うちの不要データを解放する。この時点において、メモリ 1 1 2 は、最大特徴量に関する処理過程データのみを保持している。そして、メモリ 1 1 2 は、最大特徴量に関する処理過程データを学習部 1 3 0 へと出力する。

【 0 0 5 0 】

図 3 の構成を概括すると、メモリ 1 1 2 は、第 1 の選択部 3 3 0 - 1 または第 2 の選択部 3 3 0 - 2 において選択処理の対象となっている 2 つの特徴量にそれぞれ対応する 2 つの処理過程データのみを保持する。即ち、メモリ 1 1 2 は、2 つの処理過程データを上限として保持する。特徴量抽出部 1 1 1 では合計で 3 つの処理過程データが発生するが、メモリ 1 1 2 は、選択処理毎に不要データを解放するため、3 つの処理過程データを全て保持する必要がなく、メモリ容量を削減することができる。

10

【 0 0 5 1 】

以上、入力画像から生成された 3 つの処理画像を用いた処理の例を述べた。以下では、入力画像から 4 つの処理画像を生成する例について、図 4 を用いて説明する。

【 0 0 5 2 】

図 4 は、入力画像 4 0 0 を切り出すことによって 4 つの処理画像 4 1 0 から 4 4 0 までに分割する例を示す説明図である。特徴量抽出部 1 1 1 は、入力画像 4 0 0 を所定のサイズで切り出すことによって 4 つの処理画像 4 1 0 から 4 4 0 までを生成する。

【 0 0 5 3 】

図 4 では、入力画像 4 0 0 を単純に分割することによって 4 つの処理画像 4 1 0 から 4 4 0 までを生成した。しかし、認識対象の物体が隣接する処理画像の境界付近に存在すると、認識対象の物体が境界で分断されてしまう可能性がある。そこで、複数の処理画像同士を重複するように切り出すことについて、図 5 を用いて説明する。

20

【 0 0 5 4 】

図 5 は、入力画像 5 0 0 を重複して切り出すことによって 4 つの処理画像 5 1 0 から 5 4 0 までに分割する例を示す説明図である。特徴量抽出部 1 1 1 は、入力画像 5 0 0 を、複数の処理画像同士が重複するように切り出すことによって 4 つの処理画像 5 1 0 から 5 4 0 までを生成する。

【 0 0 5 5 】

図 5 では、入力画像 5 0 0 の左上頂点を含む処理画像 5 1 0 と、右上頂点を含む処理画像 5 2 0 と、左下頂点を含む処理画像 5 3 0 と、右下頂点を含む処理画像 5 4 0 とが示されている。これら 4 つの処理画像 5 1 0 から 5 4 0 までは、それぞれ一部の領域が重複している。このように、複数の処理画像が重複していることにより、一方の処理画像において認識対象の物体が分断されてしまったとしても、他方の処理画像において認識対象の物体が分断されないようにすることができる。よって、画像処理装置 1 1 0 は、認識対象の非検出を防ぐことができる。

30

【 0 0 5 6 】

以上、入力画像から 4 つの処理画像を生成する例を述べた。しかし、分割する処理画像の数は 3、或いは 4 に限らない。以下では、分割する処理画像を N 個まで拡張した場合の画像処理装置の構成例について、図 6 を用いて説明する。

【 0 0 5 7 】

40

図 6 は、図 3 の画像処理装置 1 1 0 における特徴量抽出部 1 1 1 および最大特徴量選択部 1 1 3 の他の第 1 の構成例を示すブロック図である。第 1 の構成例は、図 3 で示した 3 つの処理画像を用いた処理を N 個の処理画像を用いた処理まで拡張させたものである。よって、図 6 では、特徴量抽出部 1 1 1 を特徴量抽出部 1 1 1 A とし、最大特徴量選択部 1 1 3 を最大特徴量選択部 1 1 3 A として説明する。尚、図 6 では、画像処理装置 1 1 0 におけるメモリ 1 1 2 および最適化部 1 1 4 の図示を省略している。

【 0 0 5 8 】

特徴量抽出部 1 1 1 A は、処理画像生成部 6 1 0 と、第 1 の抽出部 6 2 0 - 1 から第 N の抽出部 6 2 0 - N までとを備える。最大特徴量選択部 1 1 3 A は、第 1 の選択部 6 3 0 - 1 から第 L の選択部 6 3 0 - L までを備える。ここで、L は N - 1 である。

50

【 0 0 5 9 】

処理画像生成部 6 1 0 は、入力画像に基づいて N 個の処理画像を生成する。処理画像生成部 6 1 0 は、N 個の処理画像のそれぞれを第 1 の抽出部 6 1 0 - 1 から第 N の抽出部 6 1 0 - N までへと出力する。

【 0 0 6 0 】

第 1 の抽出部 6 2 0 - 1、第 2 の抽出部 6 2 0 - 2、第 1 の選択部 6 3 0 - 1、第 3 の抽出部 6 2 0 - 3、および第 2 の選択部 6 3 0 - 2 は、図 3 の第 1 の抽出部 3 2 0 - 1、第 2 の抽出部 3 2 0 - 2、第 1 の選択部 3 3 0 - 1、第 3 の抽出部 3 2 0 - 3、および第 2 の選択部 6 3 0 - 2 と同様の処理であるため説明を省略する。

【 0 0 6 1 】

第 4 の抽出部 6 2 0 - 4 および第 3 の選択部 6 3 0 - 3 は、第 3 の抽出部 6 2 0 - 3 および第 2 の選択部 6 3 0 - 2 と略同様の処理である。尚、以降の抽出部および選択部についても同様である。

【 0 0 6 2 】

図 6 の構成を概括すると、第 2 の選択部 6 3 0 - 2 以降の選択部は、直前の選択部において選択された選択特徴量と、未選択の特徴量との 2 つの特徴量を順次比較する構成である。また、第 3 の抽出部 6 2 0 - 3 以降の抽出部は、処理過程データをメモリ 1 1 2 へ出力する際に、メモリ 1 1 2 に 1 つの処理過程データしか保持されていないタイミングで行われる。または、第 3 の抽出部 6 2 0 - 3 以降の抽出部における特徴量抽出処理は、メモリ 1 1 2 において 1 つの処理過程データを保持している状態で行われる。即ち、処理画像が N 個まで拡張されたとしても、メモリ 1 1 2 には、2 つの処理過程データを上限として保持するだけでよい。

【 0 0 6 3 】

以上、分割する処理画像を N 個まで拡張した場合の画像処理装置の構成例について説明した。次に、N 個の処理画像を用いた第 1 の実施形態に係る画像処理装置 1 1 0 の動作について、図 7 を用いて説明する。

【 0 0 6 4 】

図 7 は、第 1 の実施形態に係る画像処理装置の動作を例示するフローチャートである。図 7 のフローチャートは、1 つの入力画像についての最大特徴量選択処理の一連の流れを示している。また、図 7 のフローチャートは、図 6 で示したような、選択部において 2 つの特徴量を比較する構成を前提としている。以降では、図 1 および図 6 の各部を参照して説明する。

【 0 0 6 5 】

(ステップ S T 7 0 1)

画像処理装置 1 1 0 が入力画像を取得すると、処理画像生成部 6 1 0 は、入力画像に基づく N 個 (N ≥ 3) の処理画像を生成する。

【 0 0 6 6 】

(ステップ S T 7 0 2)

第 1 の抽出部 6 2 0 - 1 は、第 1 の処理画像について第 1 の抽出処理を行うことによって第 1 の特徴量を生成する。

【 0 0 6 7 】

(ステップ S T 7 0 3)

メモリ 1 1 2 は、第 1 の抽出処理の過程で発生する第 1 の処理過程データを保持する。

【 0 0 6 8 】

(ステップ S T 7 0 4)

第 2 の抽出部 6 2 0 - 2 は、第 2 の処理画像について第 2 の抽出処理を行うことによって第 2 の特徴量を生成する。

【 0 0 6 9 】

(ステップ S T 7 0 5)

メモリ 1 1 2 は、第 2 の抽出処理の過程で発生する第 2 の処理過程データを保持する。

10

20

30

40

50

この時、メモリ 112 は、2つの処理過程データを保持している。

【0070】

(ステップST706)

第1の選択部630-1は、第1の特徴量と第2の特徴量とを比較することによって大きい方を第1の選択特徴量として選択する。

【0071】

(ステップST707)

最適化部114は、第1の特徴量と第2の特徴量との比較において選択されなかった特徴量に対応する処理過程データをメモリ112から解放させる。これにより、メモリ112は、1つの処理過程データを保持する。

10

【0072】

(ステップST708)

画像処理装置110は、変数*i*および変数*j*を定義し、それぞれ3および1を代入する。

【0073】

(ステップST709)

第*i*の抽出部620-*i*は、第*i*の処理画像について第*i*の抽出処理を行うことによって第*i*の特徴量を生成する。

【0074】

(ステップST710)

メモリ112は、第*i*の抽出処理の過程で発生する第*i*の処理過程データを保持する。
この時、メモリ112は、2つの処理過程データを保持している。

20

【0075】

(ステップST711)

第(*i*-1)の選択部630-(*i*-1)は、第*j*の選択特徴量と第*i*の特徴量とを比較することによって大きい方を第(*j*+1)の選択特徴量として選択する。

【0076】

(ステップST712)

最適化部114は、第*j*の選択特徴量と第*i*の特徴量との比較において選択されなかった特徴量に対応する処理過程データをメモリ112から解放させる。これにより、メモリ112は、1つの処理過程データだけを保持する。

30

【0077】

(ステップST713)

画像処理装置110は、変数*i*が*N*であるか否かを判定する。変数*i*が*N*ではない場合、処理はステップST714へ進む。他方、変数*i*が*N*である場合、画像処理装置110は、直前の選択処理において選択された選択特徴量を最大特徴量として誤差算出部120へと出力し、最大特徴量に関する処理過程データを学習部130へと出力し、処理は終了する。

【0078】

(ステップST714)

画像処理装置110は、変数*i*および変数*j*にそれぞれ1を加算する。ステップST714の処理の後、処理はステップST709へ戻る。

40

【0079】

以上、*N*個の処理画像を用いた第1の実施形態に係る画像処理装置110の動作について説明した。上記までは、選択部において2つの特徴量を比較する構成について述べた。しかし、比較する特徴量の数は2に限らない。以下では、選択部において3つの特徴量を比較する例について、図8を用いて説明する。尚、図8の説明の際、比較対象として、選択部において2つの特徴量を比較する構成である図6を参照する。

【0080】

図8は、図3の画像処理装置における特徴量抽出部および最大特徴量選択部の他の第2の構成例を示すブロック図である。第2の構成例は、図3または図6で示した選択部にお

50

ける2つの特徴量の比較を3つの特徴量の比較に拡張させたものである。よって、図8では、特徴量抽出部111を特徴量抽出部111Bとし、最大特徴量選択部113を最大特徴量選択部113Bとして説明する。尚、図8では、画像処理装置110におけるメモリ112および最適化部114の図示を省略している。

【0081】

特徴量抽出部111Bは、処理画像生成部810と、第1の抽出部820-1から第Nの抽出部820-Nまでとを備える。最大特徴量選択部113Bは、第1の選択部830-1から第Lの選択部830-Lまでを備える。ここで、Lは $(N-1)/2$ である。

【0082】

前述の通り、図8と図6との違いは、選択部において比較する特徴量の数である。具体的には、図8の第1の選択部830-1は、第1の抽出部820-1から第1の特徴量を入力し、第2の抽出部820-2から第2の特徴量を入力し、第3の抽出部820-3から第3の特徴量を入力する。そして、第1の選択部830-1は、第1の特徴量から第3の特徴量までの3つの特徴量を比較することによって最も大きい特徴量を第1の選択特徴量として選択する。

10

【0083】

また、図8では、メモリ112が保持する処理過程データの数も異なる。例えば、第1の選択部830-1において3つの特徴量を比較することから、第3の特徴量が抽出された時点において、メモリ112は、3つの処理過程データを保持している。

【0084】

20

第1の選択部830-1において第1の選択特徴量が選択されると、メモリ112は、保持している3つの処理過程データのうちの不要データ（ここでは、選択されなかった2つの特徴量に対応する2つの処理過程データ）を解放する。この時点において、メモリ112は、1つの処理過程データを保持している。

【0085】

次いで、第4の抽出部820-4において第4の特徴量が抽出され、第5の抽出部820-5において第5の特徴量が抽出されることによって、メモリ112は、再び3つの処理過程データを保持する状態となる。

【0086】

更に、第2の選択部830-2は、第1の選択部830-1から第1の選択特徴量を入力し、第4の抽出部820-4から第4の特徴量を入力し、第5の抽出部820-5から第5の特徴量を入力する。そして、第2の選択部830-2は、第1の選択特徴量と第4の特徴量と第5の特徴量との3つの特徴量を比較することによって最も大きい特徴量を第2の選択特徴量として選択する。

30

【0087】

第2の選択特徴量が選択されると、メモリ112は、保持している3つの処理過程データのうちの不要データを解放する。この時点において、メモリ112は、再び1つの処理過程データのみを保持する状態となる。

【0088】

図8の構成を概括すると、第2の選択部830-2以降の選択部は、直前の選択部において選択された選択特徴量と、未選択の2つの特徴量との3つの特徴量を順次比較する構成である。また、第4の抽出部820-4以降の抽出部は、処理過程データをメモリ112へ出力する際に、メモリ112に多くとも2つの処理過程データを保持している状態で行われる。即ち、メモリ112は、3つの処理過程データを上限として保持する。

40

【0089】

なお、図8では、画像処理装置110は、選択部において3つの特徴量を比較する構成の例を示したがこれに限らない。例えば、画像処理装置110は、選択部において4つ以上の特徴量を比較する構成でもよい。

【0090】

以上のように、メモリ112に保持する処理過程データの上限と選択部において比較す

50

る特徴量の数とを一致させることによって、処理画像の個数に関わらず、メモリ 1 1 2 には、選択部において比較する特徴量の数と同数の処理過程データを上限として保持するだけでよい。

【 0 0 9 1 】

さらに、図 8 の構成では、第 1 の抽出部 8 2 0 - 1、第 2 の抽出部 8 2 0 - 2、および第 3 の抽出部 8 2 0 - 3 のそれぞれの抽出処理を並列して行ってよい。その後、第 1 の選択部 8 3 0 - 1 による選択処理が完了し、メモリ 1 1 2 から不要データが削除された後、第 4 の抽出部 8 2 0 - 4 および第 5 の抽出部 8 2 0 - 5 のそれぞれの抽出処理も並列して行ってよく、以降も同様である。よって、図 8 の構成によれば、抽出処理が並列して行えることにより、画像処理装置 1 1 0 は、図 6 の構成に比べて全体の処理時間を短縮することができる。

10

【 0 0 9 2 】

図 8 における抽出処理の並列化は、選択部において比較する特徴量の数が増えたとしても同様である。例えば、画像処理装置 1 1 0 は、選択部において M 個の特徴量を比較する場合、第 1 の選択処理で用いられる M 個の特徴量に係る抽出処理を並列して行い、第 2 の選択処理以降で用いられる M - 1 個の特徴量に係る抽出処理を並列して行ってよい。換言すると、画像処理装置 1 1 0 は、選択部へ入力される抽出処理を経た直後の特徴量が複数ある場合、これら複数の特徴量を同時に生成してよい。

【 0 0 9 3 】

(処理画像の他の実施例)

20

上記では、入力画像を分割することによって複数の処理画像を生成する例について説明した。しかし、複数の処理画像は、入力画像の分割に限らない。以下では、入力画像を縮小することによって複数の処理画像を生成する例について、図 9 を用いて説明する。

【 0 0 9 4 】

図 9 は、入力画像 9 0 0 を縮小することによって 2 つの処理画像 9 1 0 および 9 2 0 を生成する例を示す説明図である。特徴量抽出部 1 1 1 は、縮小処理によって入力画像 9 0 0 を縮小させることによって縮小率の異なる 2 つの処理画像 9 1 0 および 9 2 0 を生成する。例えば、特徴量抽出部 1 1 1 は、入力画像 9 0 0 を $1/2$ に縮小することによって処理画像 9 1 0 を生成し、入力画像 9 0 0 を $1/4$ に縮小することによって処理画像 9 2 0 を生成する。尚、画像処理装置 1 1 0 の処理可能な画像サイズを満たす場合、入力画像 9 0 0 は、処理画像として用いてよい。

30

【 0 0 9 5 】

上記の縮小処理は、例えばバイリニアおよびバイキュービックなどの固定フィルタを用いてニューラルネットワークの処理とは別に行われてもよいし、上記固定フィルタを畳み込みフィルタとしてニューラルネットワークの一部として行われてもよい。後者の場合、畳み込みフィルタのパラメータを新たに学習する必要があるため、ニューラルネットワーク全体の学習速度が低下するものの、前者の場合に比べて、認識精度を向上させることが期待できる。

【 0 0 9 6 】

次に、図 9 のような縮小率の異なる処理画像を用いることのメリットについて説明する。以下では、2 つの観点について説明する。

40

【 0 0 9 7 】

1 つ目の観点は、縮小画像自体を学習することのメリットである。例えば、画像処理装置 1 1 0 を用いた推論時において、学習時とは異なる大きさの認識対象の物体が入力画像に含まれている場合、縮小画像を用いていない複数の処理画像を用いて学習したニューラルネットワークは、その物体を認識できない可能性がある。

【 0 0 9 8 】

一方、図 9 のような縮小率の異なる処理画像を用いた場合、例えば、処理画像としての入力画像 9 0 0 では物体をサイズ A で学習し、処理画像 9 1 0 では物体をサイズ $A/2$ で学習し、処理画像 9 2 0 では物体をサイズ $A/4$ で学習することとなる。このとき、個別

50

ニューラルネットワークのパラメータを共有していれば、画像処理装置 110 は、上記の何れのサイズの物体であっても認識させることができる。

【0099】

更に、処理画像としての入力画像 900 にサイズ 2A の物体が写っていた場合であっても、処理画像 910 においてはサイズ A に縮小されることから、画像処理装置 110 は、サイズ 2A の物体も認識させることができる。このことは、処理画像としての入力画像 900 にサイズ 4A の物体写っていた場合でも同様である。

【0100】

よって、画像処理装置 110 は、縮小率の異なる処理画像を用いた学習および推論において個別ニューラルネットワークのパラメータを共有させることにより、推論時において学習時とは異なる大きさの認識対象の物体を認識させることができる。

10

【0101】

2 つ目の観点は、個別ニューラルネットワークに畳み込み処理を含めることのメリットである。そのために、まずニューラルネットワークによる畳み込み処理における受容野の概念について、図 10 を用いて説明する。

【0102】

図 10 は、入力画像 1010 に対する 2 回の畳み込み処理における変換画像 1020、中間画像 1030、および受容野を例示する説明図である。図 10 では、入力画像 1010 に対して 3×3 画素の畳み込み処理を行って変換画像 1020 を生成し、変換画像 1020 に対しても同様に 3×3 画素の畳み込み処理を行って中間画像 1030 を生成する例が示されている。なお活性化処理などの図示は省略している。ここで、受容野とは、中間画像 1030 の 1 画素（例えば、画素 1031）に影響を与える、変換画像 1020 の画素範囲 1021 および入力画像 1010 の画素範囲 1011 のことである。図 10 では、画素範囲 1021 は 3×3 画素であり、画素範囲 1011 は 5×5 画素である例が示されている。画素 1031 は、受容野である画素範囲 1021 および画素範囲 1011 のみに依存するため、この受容野以外の画素値がいくら変化しても影響はない。

20

【0103】

なお、入力画像 1010 の受容野（画素範囲 1011）は、中間画像 1030 のサイズ、および畳み込み処理のカーネル（例えば、 3×3 画素）が変わらない場合、変換画像を生成する畳み込み処理を増やすほど、即ち、畳み込み層の数を増やすほどその画素範囲が

30

【0104】

以上を踏まえて、図 9 のような縮小率の異なる処理画像それぞれについての畳み込み処理について考える。まず、3 つの個別ニューラルネットワークは、いずれも複数の畳み込み層で構成される。また、これら複数の畳み込み層において、最後に畳み込み処理された変換画像を中間画像と呼ぶこととする。画像処理装置 110 は、中間画像に対して全結合処理やグローバルプーリング処理を適用して特徴量を生成する。

【0105】

上記のように構成されたニューラルネットワークは、一般的に、中間画像までの処理において入力画像の特徴量の抽出を行い、その後の処理（上記の全結合処理およびグローバルプーリング処理）において特徴量を用いた識別を行う。従って、中間画像の各画素には、認識対象の物体の特徴が十分に反映されていることが望ましい。しかし、例えば、学習時よりも大きな物体が入力画像に含まれていると、入力画像の受容野よりも物体が大きくなってしまい、認識対象の物体の特徴が中間画像において十分に反映されないことがある。このことに対して、畳み込み層を増やすことにより入力画像の受容野の領域を広くすることが考えられるが、畳み込み層を増やした分だけ処理量が増えるという問題がある。

40

【0106】

一方、縮小画像を用いてニューラルネットワークを学習することにより、複数の個別ニューラルネットワークの構造が同じであれば、それぞれの個別ニューラルネットワークにおいて入力画像に対する受容野の大きさは変わらない。よって、画像処理装置 110 は、

50

学習時よりも大きな物体が入力画像に含まれていたとしても、いずれかの個別ニューラルネットワークにおいて認識対象の物体を認識させることができる。

【 0 1 0 7 】

また、先に述べたように、画像の縮小に用いる畳み込みフィルタのパラメータも個別ニューラルネットワークに学習させる場合、入力画像とは異なった様相の縮小画像になることがある。よって、画像処理装置 1 1 0 は、個別ニューラルネットワークのパラメータを異ならせて、画像の縮小率毎に個別ニューラルネットワークのパラメータを最適化させることが望ましい。

【 0 1 0 8 】

上記を概括すると、画像処理装置 1 1 0 は、入力画像を縮小した複数の処理画像を用いることにより、学習時に想定していた認識対象の物体のサイズとは異なるサイズの物体であっても認識することができる。

10

【 0 1 0 9 】

さらに、処理画像のサイズが小さくなれば処理過程データの容量も少なくすることができる。よって、画像処理装置 1 1 0 は、縮小率が最も高い処理画像から特徴量抽出処理を行うことによって、メモリ 1 1 2 に保持するデータの容量を低減することができる。

【 0 1 1 0 】

(最大特徴量選択部の他の構成例)

上記では、最大特徴量選択部における各々の選択部において、異なる選択部からの出力同士を比較する構成は例示していなかった。以下では、異なる選択部からの出力同士を比較する選択部を含む最大特徴量選択部の構成について、図 1 1 を用いて説明する。尚、図 1 1 では、説明を簡略化するため、4つの処理画像を用いることを想定する。

20

【 0 1 1 1 】

図 1 1 は、図 1 の画像処理装置 1 1 0 の他の構成例を示すブロック図である。他の構成例は、異なる選択部からの出力同士を比較する構成を含むものである。よって、図 1 1 では、特徴量抽出部 1 1 1、メモリ 1 1 2、最大特徴量選択部 1 1 3、および最適化部 1 1 4 をそれぞれ特徴量抽出部 1 1 1 C、メモリ 1 1 2 C、最大特徴量選択部 1 1 3 C、および最適化部 1 1 4 C として説明する。

【 0 1 1 2 】

特徴量抽出部 1 1 1 C は、処理画像生成部 1 1 1 0 と、第 1 の抽出部 1 1 2 0 - 1 から第 4 の抽出部 1 1 2 0 - 4 までとを備える。最大特徴量選択部 1 1 3 C は、第 1 の選択部 1 1 3 0 - 1 から第 3 の選択部 1 1 3 0 - 3 までを備える。

30

【 0 1 1 3 】

処理画像生成部 1 1 1 0 は、入力画像に基づいて 4 つの処理画像を生成する。処理画像生成部 1 1 1 0 は、4 つの処理画像のうちの第 1 の処理画像を第 1 の抽出部 1 1 2 0 - 1 へと出力し、第 2 の処理画像を第 2 の抽出部 1 1 2 0 - 2 へと出力し、第 3 の処理画像を第 3 の抽出部 1 1 2 0 - 3 へと出力し、第 4 の処理画像を第 4 の抽出部 1 1 2 0 - 4 へと出力する。

【 0 1 1 4 】

第 1 の抽出部 1 1 2 0 - 1 は、処理画像生成部 1 1 1 0 から第 1 の処理画像を入力する。第 1 の抽出部 1 1 2 0 - 1 は、第 1 の処理画像について、特徴量抽出処理に相当する第 1 の抽出処理を行うことによって第 1 の特徴量を生成する。第 1 の抽出部 1 1 2 0 - 1 は、第 1 の特徴量を第 1 の選択部 1 1 3 0 - 1 へと出力し、第 1 の抽出処理の過程で発生する第 1 の処理過程データをメモリ 1 1 2 C へと出力する。

40

【 0 1 1 5 】

第 1 の特徴量が抽出された後、メモリ 1 1 2 C は、第 1 の抽出部 1 1 2 0 - 1 から第 1 の処理過程データを入力し、これを保持する。この時点において、メモリ 1 1 2 C は、1 つの処理過程データを保持している。

【 0 1 1 6 】

第 2 の抽出部 1 1 2 0 - 2 は、処理画像生成部 1 1 1 0 から第 2 の処理画像を入力する

50

。第2の抽出部1120-2は、第2の処理画像について、特徴量抽出処理に相当する第2の抽出処理を行うことによって第2の特徴量を生成する。第2の抽出部1120-2は、第2の特徴量を第1の選択部1130-1へと出力し、第2の抽出処理の過程で発生する第2の処理過程データをメモリ112Cへと出力する。

【0117】

なお、第2の抽出部1120-2における第2の抽出処理は、第1の抽出部1120-1における第1の抽出処理と同じタイミングで行われてよい。

【0118】

第2の特徴量が抽出された後、メモリ112Cは、第2の抽出部1120-2から第2の処理過程データを入力し、これを保持する。この時点において、メモリ112Cは、2つの処理過程データを保持している。

10

【0119】

第1の選択部1130-1は、第1の抽出部1120-1から第1の特徴量を入力し、第2の抽出部1120-2から第2の特徴量を入力する。第1の選択部1130-1は、第1の特徴量と第2の特徴量とを比較することによって大きい方を第1の選択特徴量として選択する。第1の選択部1130-1は、選択されなかった特徴量に関する第1の非選択情報を生成して最適化部114Cへと出力し、第1の選択特徴量を第3の選択部1130-3へと出力する。

【0120】

第1の非選択情報が生成された後、最適化部114Cは、第1の選択部1130-1から第1の非選択情報を入力する。最適化部114Cは、第1の非選択情報に基づいて第1の解放指示情報を生成し、メモリ112Cへと出力する。

20

【0121】

第1の解放指示情報が生成された後、メモリ112Cは、第1の解放指示情報を入力する。メモリ112Cは、第1の解放指示情報に従って、保持している2つの処理過程データのうちの不要データを解放する。この時点において、メモリ112Cは1つの処理過程データを保持している。

【0122】

第3の抽出部1120-3は、処理画像生成部1110から第3の処理画像を入力する。第3の抽出部1120-3は、第3の処理画像について、特徴量抽出処理に相当する第3の抽出処理を行うことによって第3の特徴量を生成する。第3の抽出部1120-3は、第3の特徴量を第2の選択部1130-2へと出力し、第3の抽出処理の過程で発生する第3の処理過程データをメモリ112Cへと出力する。

30

【0123】

なお、第3の抽出部1120-3における特徴量抽出処理は、第3の処理過程データをメモリ112Cへ出力する際に、メモリ112Cに1つの処理過程データしか保持されていないタイミングで行われる。または、第3の抽出部1120-3における特徴量抽出処理は、メモリ112Cにおいて1つの処理過程データを保持している状態で行われる。

【0124】

第3の特徴量が抽出された後、メモリ112Cは、第3の抽出部1120-3から第3の処理過程データを入力し、これを保持する。この時点において、メモリ112Cは、2つの処理過程データを保持している。

40

【0125】

第4の抽出部1120-4は、処理画像生成部1110から第4の処理画像を入力する。第4の抽出部1120-4は、第4の処理画像について、特徴量抽出処理に相当する第4の抽出処理を行うことによって第4の特徴量を生成する。第4の抽出部1120-4は、第4の特徴量を第2の選択部1130-2へと出力し、第4の抽出処理の過程で発生する第4の処理過程データをメモリ112Cへと出力する。

【0126】

なお、第4の抽出部1120-4における第4の抽出処理は、第3の抽出部1120-

50

3における第3の抽出処理と同じタイミングで行われてよい。

【0127】

第4の特徴量が抽出された後、メモリ112Cは、第4の抽出部1120-4から第4の処理過程データを入力し、これを保持する。この時点において、メモリ112Cは、3つの処理過程データを保持している。

【0128】

図11の構成を概括すると、図6のように選択部において2つの特徴量を比較する構成の組み合わせではあるものの、異なる選択部からの出力（即ち、2つの選択特徴量）同士を比較する構成が含まれる。これにより、メモリ112Cには3つの処理過程データを保持することとなるが、特徴量抽出部111Cは、2つの抽出部によって同時に抽出処理を行うことができる。即ち、特徴量抽出部111Cは、4つの特徴量を2つの特徴量毎に生成することができる。

10

【0129】

更に、N個の処理画像について複数の特徴量の比較をする選択部まで拡張させると、特徴量抽出部111Cは、N個の特徴量を複数の特徴量毎に生成することができる。これにより、画像処理装置110は、従来よりもメモリの使用量を低減しつつ、特徴量抽出処理のスループットを向上させることができる。

【0130】

以上、異なる選択部からの出力同士を比較する選択部を含む最大特徴量選択部の構成について説明した。次に、このような構成を有する画像処理装置110の動作について、図12を用いて説明する。

20

【0131】

図12は、第1の実施形態に係る画像処理装置110の他の動作を例示するフローチャートである。図12のフローチャートは、1つの入力画像についての最大特徴量選択処理の一連の流れを示している。また、図12のフローチャートは、図11で示したような、選択部において2つの選択特徴量の比較も含む構成を前提とし、処理画像の数をN個まで拡張させている。以降では、図1および図11の各部を参照して説明する。

【0132】

（ステップST1201）

画像処理装置110が入力画像を取得すると、処理画像生成部1110は、入力画像に基づくN個（N=4）の処理画像を生成する。

30

【0133】

（ステップST1202）

ステップST1202の処理は、図7のステップST702からステップST707までの処理と同様である。具体的には、第1の抽出部1120-1は、第1の処理画像について第1の抽出処理を行うことによって第1の特徴量を生成する。メモリ112Cは、第1の抽出処理の過程で発生する第1の処理過程データを保持する。第2の抽出部1120-2は、第2の処理画像について第2の抽出処理を行うことによって第2の特徴量を生成する。メモリ112Cは、第2の抽出処理の過程で発生する第2の処理過程データを保持する。第1の選択部1130-1は、第1の特徴量と第2の特徴量とを比較することによって大きい方を第1の選択特徴量として選択する。最適化部114Cは、第1の特徴量と第2の特徴量との比較において選択されなかった特徴量に対応する処理過程データをメモリ112Cから解放させる。

40

【0134】

（ステップST1203）

画像処理装置110は、変数iおよび変数jを定義し、それぞれ3および2を代入する。

【0135】

（ステップST1204）

第iの抽出部1120-iは、第iの処理画像について第iの抽出処理を行うことによって第iの特徴量を生成する。

50

【0136】

(ステップST1205)

メモリ112Cは、第*i*の抽出処理の過程で発生する第*i*の処理過程データを保持する。この時、メモリ112Cは、2つの処理過程データを保持している。

【0137】

(ステップST1206)

第(*i* + 1)の抽出部620 - (*i* + 1)は、第(*i* + 1)の処理画像について第(*i* + 1)の抽出処理を行うことによって第(*i* + 1)の特徴量を生成する。

【0138】

(ステップST1207)

メモリ112Cは、第(*i* + 1)の抽出処理の過程で発生する第(*i* + 1)の処理過程データを保持する。この時、メモリ112Cは、3つの処理過程データを保持している。

【0139】

(ステップST1208)

第(*i* - 1)の選択部1130 - (*i* - 1)は、第*i*の特徴量と第(*i* + 1)の特徴量とを比較することによって大きい方を第*j*の選択特徴量として選択する。

【0140】

(ステップST1209)

最適化部114Cは、第*i*の特徴量と第(*i* + 1)の特徴量との比較において選択されなかった特徴量に対応する処理過程データをメモリ112Cから解放させる。これにより、メモリ112Cは、2つの処理過程データを保持する。

【0141】

(ステップST1210)

第*i*の選択部1130 - *i*は、第(*j* - 1)の選択特徴量と第*j*の選択特徴量とを比較することによって大きい方を第(*j* + 1)の選択特徴量として選択する。

【0142】

(ステップST1211)

最適化部114Cは、第(*j* - 1)の選択特徴量と第*j*の選択特徴量との比較において選択されなかった特徴量に対応する処理過程データをメモリ112Cから解放させる。これにより、メモリ112Cは、1つの処理過程データだけを保持する。

【0143】

(ステップST1212)

画像処理装置110は、変数*i*が*N* - 1であるか否かを判定する。変数*i*が*N* - 1ではない場合、処理はステップST1213へ進む。他方、変数*i*が*N* - 1である場合、画像処理装置110は、直前の選択処理において選択された選択特徴量を最大特徴量として誤差算出部120へと出力し、最大特徴量に関する処理過程データを学習部130へと出力し、処理は終了する。

【0144】

(ステップST1213)

画像処理装置110は、変数*i*および変数*j*にそれぞれ2を加算する。ステップST1213の後、処理はステップST1204へ戻る。

【0145】

なお、ステップST1204およびステップST1206の処理は、それぞれ同じタイミングで行われてもよい。

【0146】

(特徴量抽出部の他の実施例)

上記では、入力画像を基準とした複数の処理画像を用いて、複数の処理画像のそれぞれについて特徴量抽出処理を行う構成であった。換言すると、上記の構成は、複数の処理画像それぞれについて個別ニューラルネットワークを用いていた。以下では、入力画像に対して1つのニューラルネットワークを用いた処理を行いつつも、従来よりもメモリ容量を

10

20

30

40

50

低減可能な構成について、図 13 から図 15 までを用いて説明する。

【0147】

図 13 は、図 6 の画像処理装置 110 における特徴量抽出部 111 A および最大特徴量選択部 113 A の他の第 3 の構成例を示すブロック図である。第 3 の構成例は、図 6 で示した N 個の特徴量を用いた処理をベースとして、N 個の処理画像についての取り扱いを変更させたものである。よって、図 13 では、特徴量抽出部 111 A を特徴量抽出部 111 D とし、最大特徴量選択部 113 A を最大特徴量選択部 113 D として説明する。尚、図 13 では、画像処理装置 110 におけるメモリ 112 および最適化部 114 の図示を省略している。

【0148】

特徴量抽出部 111 D は、畳み込み処理部 1310 を備える。最大特徴量選択部 113 D は、第 1 の選択部 1320 - 1 から第 L の選択部 1320 - L までを備える。ここで、L は N - 1 である。

【0149】

畳み込み処理部 1310 は、特徴量抽出処理としての畳み込み処理を行うことによって入力画像から中間画像を生成し、中間画像を縦横 1 画素以上の N 個のブロックに分解し、N 個のブロックのそれぞれについて N 個の特徴量を生成する。この時、畳み込み処理部 1310 は、畳み込み処理を入力画像全体に対して一度には行わずに、特定の領域単位で行う。特定の領域とは、中間画像におけるブロックに影響を与える入力画像における領域を示す。以下では、入力画像および中間画像の関係について図 14 を用いて説明する。

【0150】

図 14 は、入力画像 1410 に対する畳み込み処理における複数の変換画像 1420、中間画像 1430、および受容野 1440 を例示する説明図である。通常、畳み込み処理によって、入力画像 1410 から変換画像 1420 を生成し、畳み込み処理を繰り返すことによって最後に生成される変換画像である中間画像 1430 が生成される。この時、中間画像 1430 におけるブロック 1431 は、変換画像 1420 における領域および入力画像 1410 における領域と対応関係（受容野 1440）がある。中間画像の例として図 15 を用いて説明する。

【0151】

図 15 は、畳み込み処理の処理単位毎に分割した中間画像 1500 である。図 15 では、縦横を 4 × 6 のブロックに分割した中間画像 1500 の例が示されている。分割された各ブロックは、上述した特定の領域に相当する。即ち、畳み込み処理部 1310 は、通常の畳み込み処理によって生成される中間画像 1500 を、特定のブロック毎に生成する。これにより、畳み込み処理部 1310 は、複数の処理画像をそれぞれ個別ニューラルネットワークで行っていた処理と同等の処理を、1 つのニューラルネットワークで行うことができる。

【0152】

具体的には、畳み込み処理部 1310 は、中間画像 1500 におけるブロック 1510 の受容野に基づいて入力画像の領域を特定し、これを第 1 の処理画像とみなして特徴量抽出処理を行う。ブロック 1510 に後続するブロック 1520 は第 2 の処理画像に対応し、更に後続するブロック 1530 は第 3 の処理画像に対応する。そして、畳み込み処理部 1310 は、最後のブロック 1540 に対応する第 24 の処理画像の特徴量抽出処理を行った後、入力画像に対する処理を終了する。

【0153】

なお、上記で説明した図 13 の特徴量抽出部 111 D は、特徴量の抽出方法が他の特徴量抽出部と異なるだけであり、以降の最大特徴量選択部 113 D による処理は、例えば図 6 の最大特徴量選択部 113 A と同様の処理を行えばよい。

【0154】

また、入力画像に対する中間画像および受容野の関係は、任意に設定可能である。例えば、中間画像の隣接するブロックのそれぞれについて、入力画像の領域を重複するような

10

20

30

40

50

受容野を設定することにより、図 5 で説明したような処理画像とみなして特徴量抽出処理を行うことができる。

【 0 1 5 5 】

(入力画像の他の実施例)

上記では、1 チャンネルの入力画像 (例えば、白黒画像) を想定して説明した。しかし、入力画像は R G B のカラー画像でもよい。入力画像をカラー画像とした場合、画像処理装置 1 1 0 は、一つの入力画像を、R e d 成分、G r e e n 成分、および B l u e 成分の縦横の画素数が同じ 3 枚の画像、いわゆる 3 チャンネルの画像として扱う。この場合、画像処理装置 1 1 0 は、3 × 3 画素 × 3 チャンネルなどの 3 次元のカーネルを用いる。また、画像処理装置 1 1 0 は、特徴量抽出処理において、2 チャンネル以上の変換処理を行ってもよい。ニューラルネットワークを用いた画像認識処理においては、一般的に変換画像のチャンネル数を多くするほど認識精度が高くなることが知られている。よって、本実施形態においても、必要に応じてチャンネル数を設定すればよい。

10

【 0 1 5 6 】

(特徴量の他の実施例)

上記では、特徴量は、スカラー値として生成されることを想定して説明した。しかし、特徴量は複数の要素を有するベクトルでもよい。例えば、ひびと汚れなど種類が異なる物体を区別して、それらを同時に認識する場合、画像処理装置 1 1 0 は、認識対象の物体の種類の数と同じ数の次元のベクトルを特徴量として生成する。

【 0 1 5 7 】

具体的には、画像処理装置 1 1 0 は、個別ニューラルネットワークにおける最後の処理で全結合を行う場合には、全結合の出力のチャンネル数を認識する種類の数に合わせ、それらを並べて特徴量とする。または、画像処理装置 1 1 0 は、個別ニューラルネットワークにおける最後の処理で平均値プーリングや最大値プーリングを行う場合には、中間画像のチャンネル数を認識する種類の数に合わせておき、チャンネルごとにプーリングした値を並べて特徴量のベクトルとする。複数のチャンネルを有する中間画像とそれぞれの特徴量とについて、図 1 6 を用いて説明する。

20

【 0 1 5 8 】

図 1 6 は、複数のチャンネルを有する中間画像とチャンネル毎の特徴量との関係を例示する説明図である。図 1 6 では、4 つのチャンネル 1 6 1 0 から 1 6 4 0 までを有する中間画像が示されている。画像処理装置 1 1 0 は、この中間画像について、チャンネル 1 6 1 0 に対応する個別特徴量 1 6 1 1、チャンネル 1 6 2 0 に対応する個別特徴量 1 6 2 1、チャンネル 1 6 3 0 に対応する個別特徴量 1 6 3 1、およびチャンネル 1 6 4 0 に対応する個別特徴量 1 6 4 1 を並べたベクトルとして特徴量を生成する。

30

【 0 1 5 9 】

次に、特徴量がベクトルの場合の選択部、最適化部、および誤差算出部における処理について説明する。以降の説明では、特徴量が 2 つの要素を有するベクトルの場合について説明する。例えば、選択部は、2 つの特徴量を比較する際、それぞれの特徴量のベクトルの要素毎に比較を行い、要素毎の大きい方を選択したベクトルを選択特徴量として出力する。この時、最適化部は、ベクトルの各要素がいずれも選択されなかった特徴量に関する処理過程データをメモリから解放させる。また、誤差算出部は、最大特徴量の各要素と、ベクトルの要素毎にそれぞれ対応する正解特徴量とに基づいてベクトルで表される誤差値を算出する。

40

【 0 1 6 0 】

以上のように、特徴量をベクトルにすることで、種類が異なる物体を区別して、同時に認識することが可能になる。また、この場合においても、画像処理装置 1 1 0 は、従来のように全ての処理過程データをメモリに保持する必要はなく、メモリの容量を少なくすることができる。

【 0 1 6 1 】

なお、個別ニューラルネットワークにおいて、各チャンネルを独立に、互いのデータ値

50

が他に影響しない構成にした場合は、選択部における要素ごとの比較において、2つの特徴量のうちの選択されなかった要素に関する処理過程データをメモリから解放させる。これにより、メモリ容量をさらに削減することができる。

【0162】

以上説明したように、第1の実施形態に係る画像処理装置は、入力画像に基づくN個（ $N \geq 3$ ）の処理画像について、ニューラルネットワークを用いた特徴量抽出処理を行うことによってN個の特徴量を生成し、特徴量抽出処理の過程で発生する処理過程データをメモリに保持し、N個の特徴量のうちの2個以上 $N - 1$ 個以下であるM個の組み合わせで2回以上の比較を行うことによって最大特徴量を選択し、2回以上の比較毎に、選択されなかった $M - 1$ 個以下の特徴量に対応する $M - 1$ 個以下の処理過程データをメモリから解放させる。

10

【0163】

従って、第1の実施形態に係る画像処理装置は、入力画像における最大特徴量を抽出するまでの過程において、不要な処理過程データをメモリから随時解放することができるため、ニューラルネットワークを用いた画像処理に必要なメモリ容量を低減することができる。

【0164】

また、第1の実施形態に係る画像処理装置を含む学習装置は、最大特徴量と入力画像に対応する正解特徴量とに基づいて誤差値を算出し、メモリが最終的に保持している最大特徴量に関する処理過程データと誤差値とに基づいてニューラルネットワークを学習する。

20

【0165】

従って、上記学習装置は、ニューラルネットワークの学習時において必要なメモリ容量を低減することができる。

【0166】

（第2の実施形態）

第1の実施形態では、画像処理装置を含む学習装置について説明した。他方、第2の実施形態では、画像処理装置を含む推論装置について説明する。第2の実施形態に係る画像処理装置の構成は、第1の実施形態に係る画像処理装置の構成と略同様である。一方で、第2の実施形態に係る画像処理装置は、メモリに保持される処理過程データの種類が第1の実施形態に係る画像処理装置と異なる。

30

【0167】

図17は、第2の実施形態に係る画像処理装置1710を含む推論装置1700の構成を例示するブロック図である。推論装置1700は、画像処理装置1710（画像処理部）と、出力部1720とを備える。画像処理装置1710は、特徴量抽出部1711と、メモリ1712と、最大特徴量選択部1713と、最適化部1714とを備える。

【0168】

なお、推論装置1700は、ニューラルネットワークによる推論に用いる入力画像を取得する取得部を備えてもよい。また、推論装置1700は、各部を制御するための制御部を備えてもよい。

【0169】

特徴量抽出部1711、メモリ1712、最大特徴量選択部1713、および最適化部1714は、例えば図1の特徴量抽出部111、メモリ112、最大特徴量選択部113、および最適化部114と略同様の構成であるため重複する説明を省略する。

40

【0170】

メモリ1712は、最大特徴量に関する処理過程データを出力部1720へと出力する点において、図1のメモリ112と異なる。最大特徴量選択部1713は、最大特徴量を出力部1720へと出力する点において、図1の最大特徴量選択部113と異なる。

【0171】

出力部1720は、最大特徴量選択部1713から最大特徴量を入力し、メモリ1712から最大特徴量に関する処理過程データを入力する。出力部1720は、最大特徴量に

50

基づいて推論結果を生成し、他の装置へと出力する。推論結果は、例えば、入力画像において認識対象の物体が存在しているか否かを表す情報である。

【 0 1 7 2 】

具体的には、出力部 1 7 2 0 は、最大特徴量としきい値とを比較することによって推論結果を生成する。例えば、出力部 1 7 2 0 は、最大特徴量がしきい値以下の場合、入力画像において認識対象の物体が存在していないことを表す推論結果を出力し、最大特徴量がしきい値よりも大きい場合、入力画像において認識対象の物体が存在していることを表す推論結果を出力する。尚、最大特徴量が「 0 」から「 1 」までの値で表されている場合、しきい値は例えば「 0 . 5 」である。

【 0 1 7 3 】

次に、画像処理装置 1 7 1 0 が扱う処理過程データの種類について説明する。第 2 の実施形態における処理過程データは、例えば、中間画像の一部である。この処理過程データは、後述する推論結果を提示する際に利用されるため、推論結果の提示に必要なデータに言い換えられてもよい。尚、第 2 の実施形態における処理過程データは、更に処理画像を含んでもよい。

【 0 1 7 4 】

中間画像を保持しておく意義として、最大特徴量に対応する中間画像は、入力画像における物体位置に対応し、その画素値が大きくなることが知られている。このことは、例えば、非特許文献「微小オブジェクト検出のためのニューラルネットワーク」(ビジョン技術の実利用ワークショップ, IS1-03, pp.32-37, Dec. 2020.) に示されている。よって、メモリに中間画像を保持しておくことにより、推論結果を提示する際に、中間画像において画素値が大きくなった部分を示した入力画像を併せて表示させることができる。この表示により、ユーザは、認識結果を目視で確認できるため、ニューラルネットワークの説明性を向上させることができる。

【 0 1 7 5 】

次に、処理過程データとしての中間画像と、メモリに記憶される中間画像のうちの部分画像との関係について図 1 8 から図 2 0 までを用いて説明する。

【 0 1 7 6 】

図 1 8 は、畳み込み処理済みの部分画像とメモリに保持される部分画像データとの関係を例示する説明図である。図 1 8 では、中間画像 1 8 0 0 のうち、最初の 2 つの部分画像 1 8 1 0 および 1 8 2 0 に対して特徴量抽出処理を行った後の状態が示されている。この例では、部分画像 1 8 2 0 は、認識対象の物体を含み、物体の位置する画素値が大きくなっている。この時、メモリ 1 7 1 2 には、部分画像 1 8 1 0 に対応する部分画像データ 1 8 1 1 と、部分画像 1 8 2 0 に対応する部分画像データ 1 8 2 1 とが保持されている。その後、最大特徴量選択部 1 7 1 3 は、選択処理によって部分画像 1 8 2 0 に対応する特徴量を選択したものとする。

【 0 1 7 7 】

図 1 9 は、メモリから解放される部分画像データを例示する説明図である。図 1 9 では、選択処理によって選択された部分画像 1 8 2 0 のみが示されている。この時、メモリ 1 7 1 2 には、部分画像データ 1 8 2 1 のみが保持され、選択処理によって選択されなかった部分画像 1 8 1 0 に対応する部分画像データ 1 8 1 1 を解放したことによる空き領域 1 9 0 0 ができている。その後、特徴量抽出部 1 7 1 1 は、新たな部分画像に対して特徴量抽出処理を行うものとする。

【 0 1 7 8 】

図 2 0 は、メモリに保持される新たな部分画像データを例示する説明図である。図 2 0 では、中間画像 1 8 0 0 のうち、部分画像 1 8 2 0 に後続する部分画像 1 8 3 0 に対して特徴量抽出処理を行った後の状態が示されている。この時、メモリ 1 7 1 2 には、部分画像データ 1 8 1 1 と、部分画像 1 8 3 0 に対応する部分画像データ 1 8 3 1 とが保持されている。

【 0 1 7 9 】

10

20

30

40

50

以降、画像処理装置 1710 は、中間画像 1800 のうちの他の部分画像についても、メモリ 1712 への部分画像データの保持と解放とを繰り返しながら処理を進める。そして、画像処理装置 1710 は、最終的にメモリ 1712 に保持されている部分画像データを処理過程データとして出力部 1720 へと出力する。

【0180】

なお、出力部 1720 は、部分画像データを含む処理過程データと入力画像とに基づいて合成画像を生成してもよい。合成画像は、例えば、ブレンド合成により入力画像に写っている認識対象の物体の画素を強調させた画像である。換言すると、合成画像は、入力画像に推論結果を可視化して反映させたものである。この時、出力部 1720 は、メモリ 1712 に保持されていない部分画像データを補間することによって補間中間画像を生成してもよい。補間中間画像について図 21 を用いて説明する。

10

【0181】

図 21 は、部分画像 1820 から生成された補間中間画像 2100 である。出力部 1720 は、部分画像 1820 以外の領域 2110 に対して、例えばゼロパディングを行うことによって補間中間画像 2100 を生成する。これにより、入力画像に対応する中間画像を復元することができるため、出力部 1720 は、合成することができる。尚、部分画像および補間中間画像は、推論結果の内容を表すものとみなせることから、推論結果を可視化した推論画像と呼ばれてもよい。

【0182】

(最大特徴量選択部の他の構成例)

20

第 1 の実施形態で述べた学習装置と異なり、推論装置におけるメモリは、所定の条件を満たすことにより、選択されなかった特徴量に関する処理過程データを保持してもよい。所定の条件とは、特徴量がしきい値以上の場合である。具体的には、最大特徴量選択部 1713 の各々の選択部は、選択されなかった特徴量に対してしきい値との比較処理を行う。そして、最大特徴量選択部 1713 は、選択されなかった特徴量がしきい値以上の場合に、その特徴量に関する非選択情報を生成しない。これにより、メモリ 1712 に複数の処理過程データが保持されることから、推論装置 1700 は、入力画像に認識対象の物体が複数含まれる場合に対応することができる。この場合、推論装置 1700 は、最大特徴量としきい値以上の特徴量とに基づいて推論結果を出力してもよい。

【0183】

30

以上説明したように、第 2 の実施形態に係る画像処理装置は、第 1 の実施形態に係る画像処理装置と同様に、入力画像に基づく N 個 ($N \geq 3$) の処理画像について、ニューラルネットワークを用いた特徴量抽出処理を行うことによって N 個の特徴量を生成し、特徴量抽出処理の過程で発生する処理過程データをメモリに保持し、 N 個の特徴量のうちの 2 個以上 $N - 1$ 個以下である M 個の組み合わせで 2 回以上の比較を行うことによって最大特徴量を選択し、2 回以上の比較毎に、選択されなかった $M - 1$ 個以下の特徴量に対応する $M - 1$ 個以下の処理過程データをメモリから解放させる。

【0184】

従って、第 2 の実施形態に係る画像処理装置は、第 1 の実施形態に係る画像処理装置と同様の効果が見込める。

40

【0185】

また、第 2 の実施形態に係る画像処理装置を含む推論装置は、最大特徴量に基づいて入力画像において認識対象の物体が存在しているか否かを表す推論結果を出力する。更に、上記推論装置は、画像処理装置における 2 回以上の比較毎に、 M 個の特徴量それぞれとしきい値とを更に比較し、2 回以上の比較毎に、選択されなかった $M - 1$ 個以下の特徴量のうち、しきい値以上の特徴量に対応する処理過程データをメモリから解放させない。更に、上記推論装置は、最大特徴量としきい値以上の特徴量とに基づいて推論結果を出力する。更に、上記推論装置は、処理過程データが推論結果を可視化した推論画像とした場合、更に、入力画像と推論画像とに基づいて入力画像に写っている認識対象の物体の画素を強調させた画像を出力する。

50

【0186】

従って、上記推論装置は、ニューラルネットワークを用いた推論時において必要なメモリ容量を低減することができる。

【0187】

(ハードウェア構成)

図22は、一実施形態に係るコンピュータ2200のハードウェア構成を例示するブロック図である。コンピュータ2200は、ハードウェアとして、CPU(Central Processing Unit)2210、RAM(Random Access Memory)2220、プログラムメモリ2230、補助記憶装置2240、入出力インタフェース2250を備える。CPU2210は、バス2260を介して、RAM2220、プログラムメモリ2230、補助記憶装置2240、および入出力インタフェース2250と通信する。

10

【0188】

CPU2210は、汎用プロセッサの一例である。RAM2220は、ワーキングメモリとしてCPU2210に使用される。RAM2220は、SDRAM(Synchronous Dynamic Random Access Memory)などの揮発性メモリを含む。プログラムメモリ2230は、最大特徴量選択処理に関するプログラム(最大特徴量選択プログラム)などを含む種々のプログラムを記憶する。プログラムメモリ2230として、例えば、ROM(Read-Only Memory)、補助記憶装置2240の一部、またはその組み合わせが使用される。補助記憶装置2240は、データを非一時的に記憶する。補助記憶装置2240は、HDDまたはSSDなどの不揮発性メモリを含む。

20

【0189】

入出力インタフェース2250は、他のデバイスと接続するためのインタフェースである。入出力インタフェース2250は、例えば、他の装置との接続に使用される。

【0190】

プログラムメモリ2230に記憶されている各プログラムはコンピュータ実行可能命令を含む。プログラム(コンピュータ実行可能命令)は、CPU2210により実行されると、CPU2210に所定の処理を実行させる。例えば、最大特徴量選択プログラムなどは、CPU2210により実行されると、CPU2210に図1、3、6、8、11、13、および17の各部に関して説明された一連の処理を実行させる。

30

【0191】

プログラムは、コンピュータで読み取り可能な記憶媒体に記憶された状態でコンピュータ2200に提供されてよい。この場合、例えば、コンピュータ2200は、記憶媒体からデータを読み出すドライブ(図示せず)をさらに備え、記憶媒体からプログラムを取得する。記憶媒体の例は、磁気ディスク、光ディスク(CD-ROM、CD-R、DVD-ROM、DVD-Rなど)、光磁気ディスク(MOなど)、半導体メモリを含む。また、プログラムを通信ネットワーク上のサーバに格納し、コンピュータ2200が入出力インタフェース2250を使用してサーバからプログラムをダウンロードするようにしてもよい。

40

【0192】

実施形態において説明される処理は、CPU2210などの汎用ハードウェアプロセッサがプログラムを実行することにより行われることに限らず、ASIC(Application Specific Integrated Circuit)などの専用ハードウェアプロセッサにより行われてもよい。処理回路(処理部)という語は、少なくとも一つの汎用ハードウェアプロセッサ、少なくとも一つの専用ハードウェアプロセッサ、または少なくとも一つの汎用ハードウェアプロセッサと少なくとも一つの専用ハードウェアプロセッサとの組み合わせを含む。図22に示す例では、CPU2210、RAM2220、およびプログラムメモリ2230が処理回路に相当する。

【0193】

50

よって、以上の各実施形態によれば、ニューラルネットワークを用いた画像処理に必要なメモリ容量を低減することができる。

【0194】

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

【符号の説明】

【0195】

100...学習装置、110...画像処理装置、111...特徴量抽出部、112...メモリ、113...最大特徴量選択部、114...最適化部、120...誤差算出部、130...学習部、200, 400, 500, 900, 1010, 1410...入力画像、210, 220, 230, 410, 420, 430, 440, 510, 520, 530, 540, 910, 920...処理画像、1011, 1021...画素範囲、1020, 1420...変換画像、1030, 1430, 1500, 1800...中間画像、1031...画素、1431, 1510, 1520, 1530, 1540...ブロック、1440...受容野、1610, 1620, 1630, 1640...チャンネル、1611, 1621, 1631, 1641...個別特徴量、1700...推論装置、1710...画像処理装置、1711...特徴量抽出部、1712...メモリ、1713...最大特徴量選択部、1714...最適化部、1720...出力部、1810, 1820, 1830...部分画像、1811, 1821, 1831...部分画像データ、1900...領域、2100...補間中間画像、2110...領域、2200...コンピュータ、2230...プログラムメモリ、2240...補助記憶装置、2250...入出力インタフェース、2260...バス。

10

20

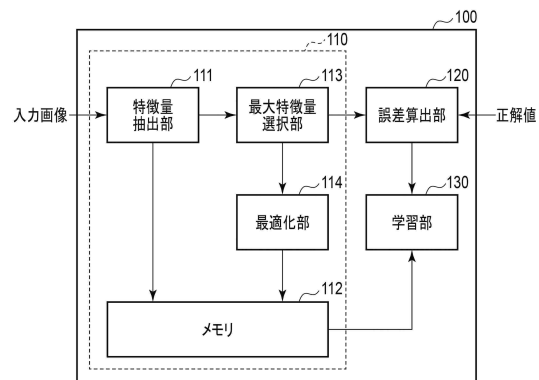
30

40

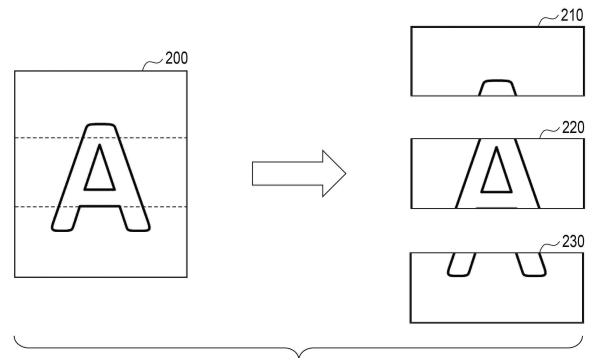
50

【図面】

【 図 1 】

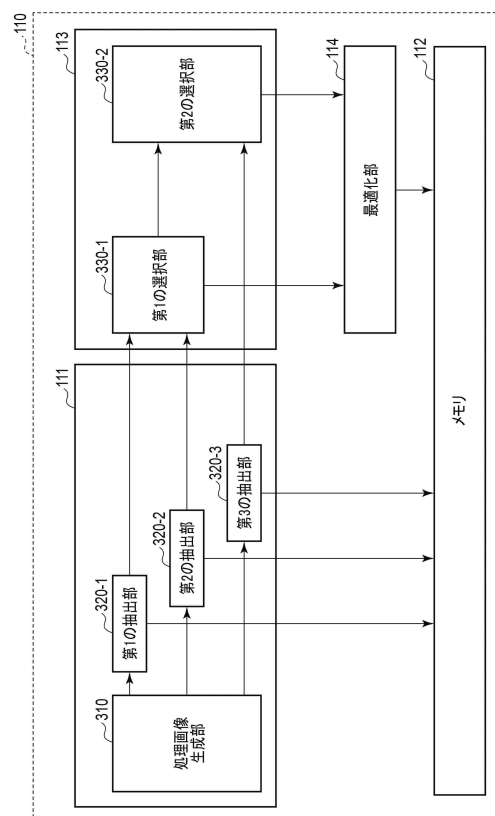


【 図 2 】

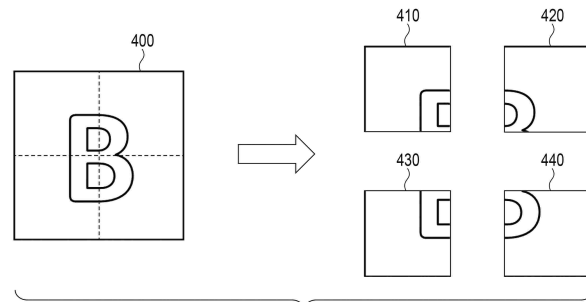


10

【 図 3 】



【 図 4 】



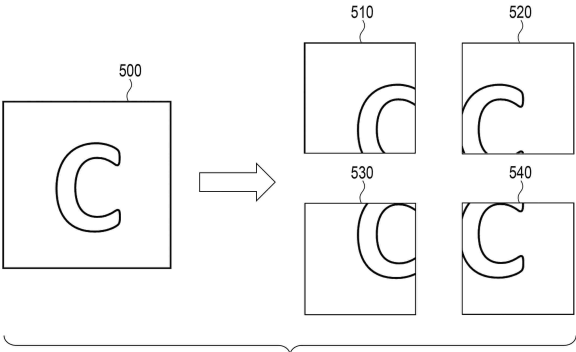
20

30

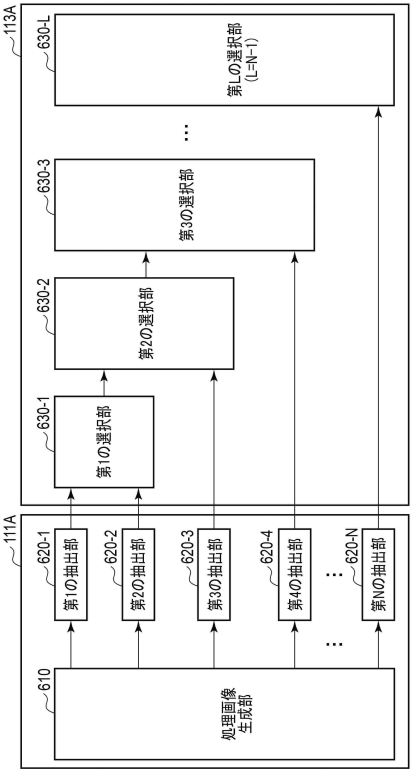
40

50

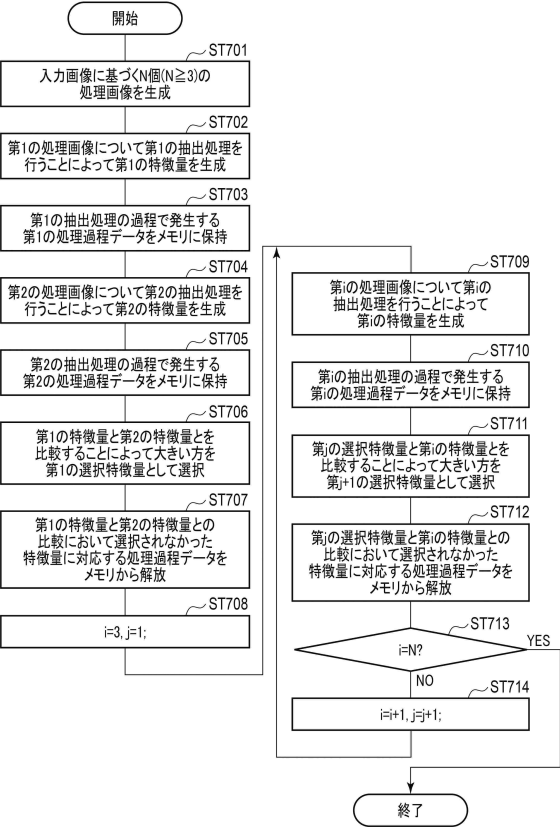
【図 5】



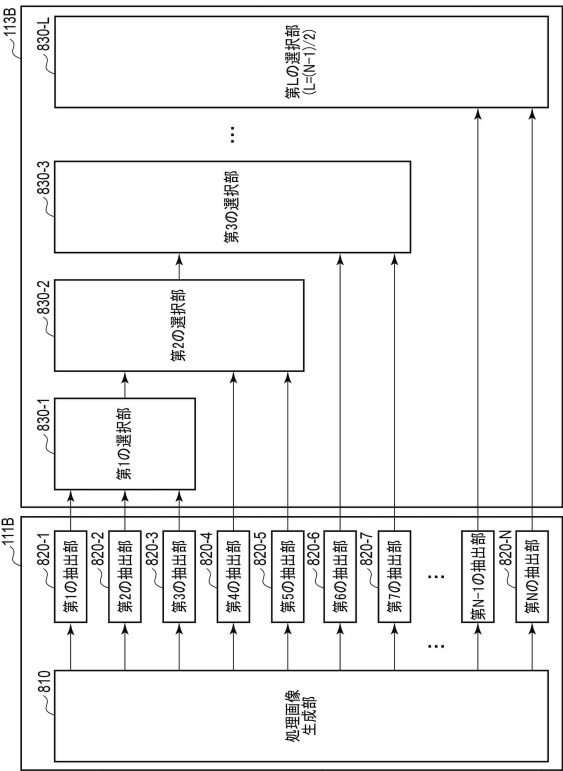
【図 6】



【図 7】



【図 8】



10

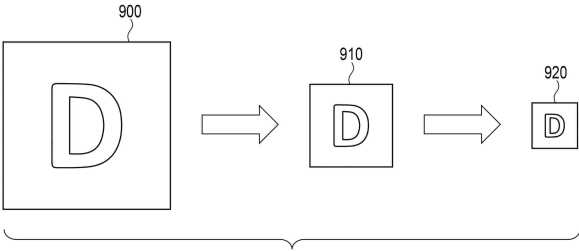
20

30

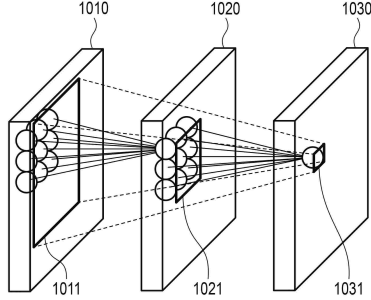
40

50

【図 9】

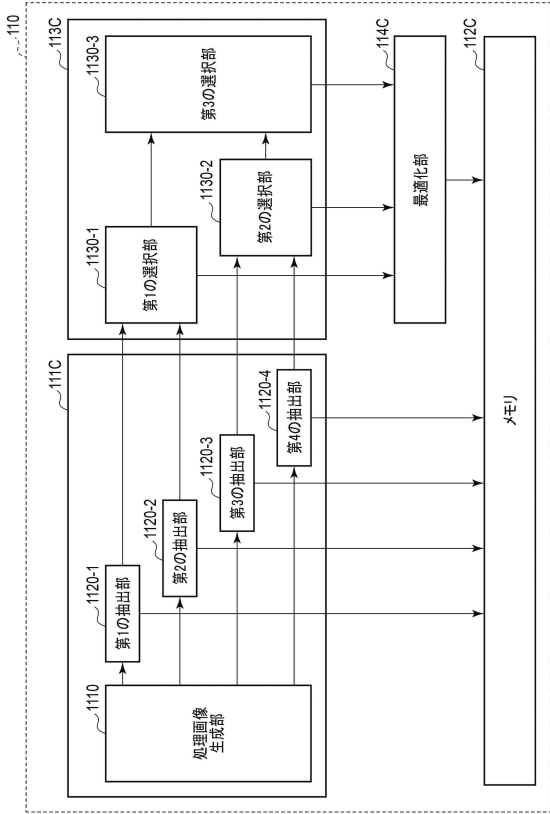


【図 10】

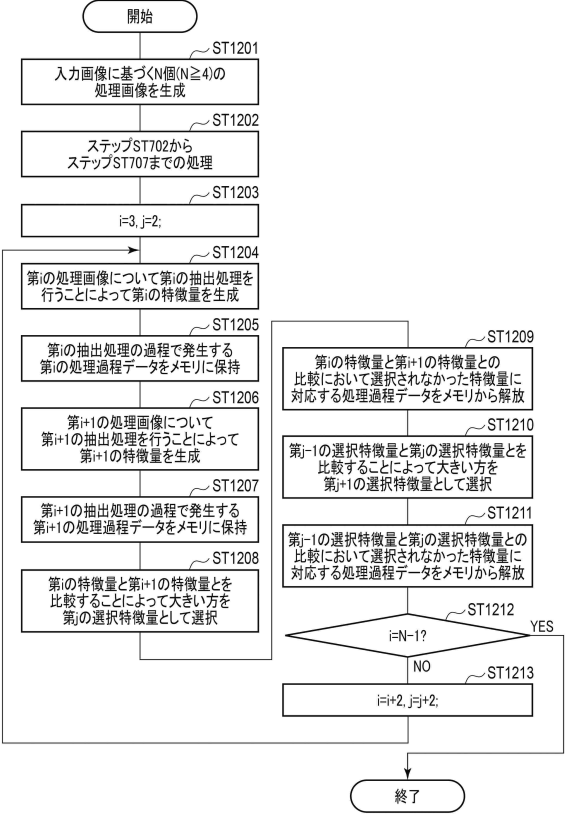


10

【図 11】



【図 12】



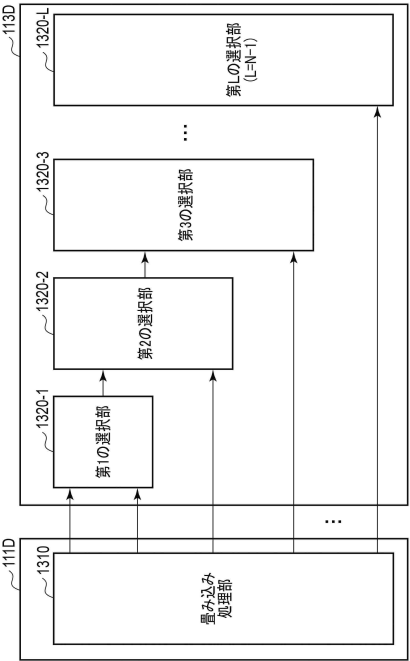
20

30

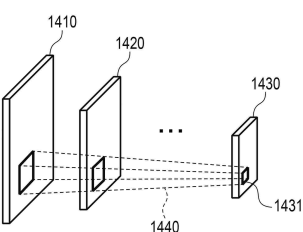
40

50

【図 1 3】



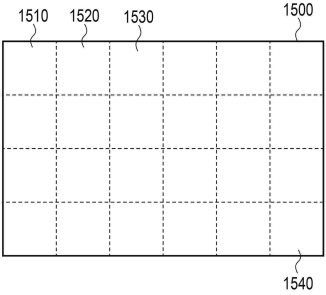
【図 1 4】



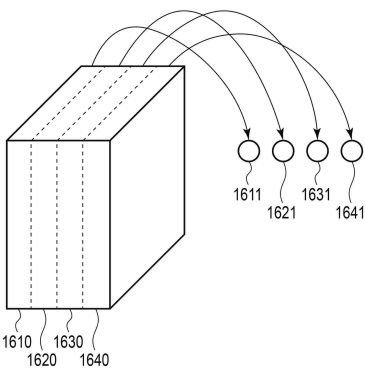
10

20

【図 1 5】



【図 1 6】

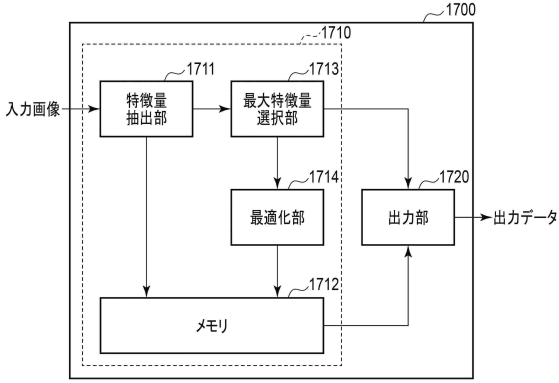


30

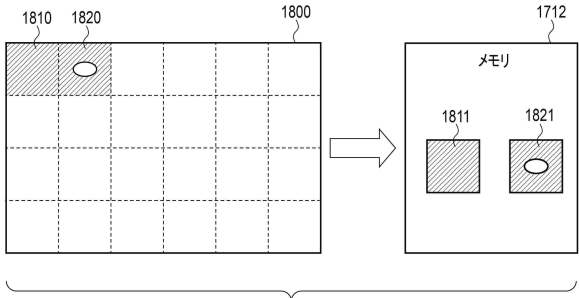
40

50

【図 17】

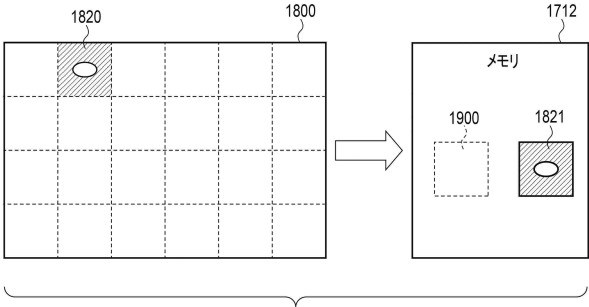


【図 18】

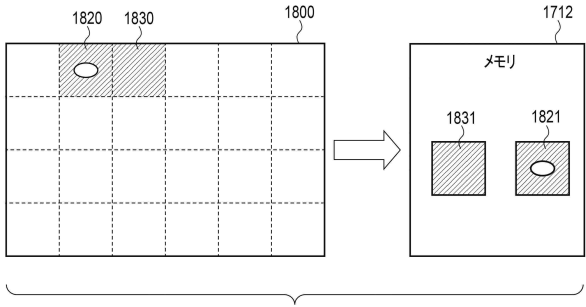


10

【図 19】



【図 20】



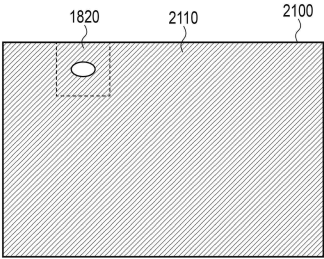
20

30

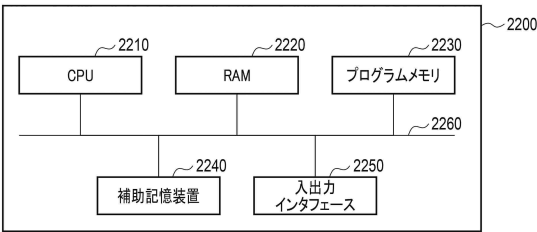
40

50

【図 2 1】



【図 2 2】



10

20

30

40

50

フロントページの続き

- (56)参考文献 特開 2 0 1 9 - 2 1 2 2 0 6 (J P , A)
 特開 2 0 2 0 - 1 4 4 6 5 9 (J P , A)
 特開 2 0 1 0 - 1 2 8 9 3 0 (J P , A)
 米国特許出願公開第 2 0 2 1 / 0 0 7 3 5 8 5 (U S , A 1)
 中国特許出願公開第 1 1 1 9 5 0 7 1 8 (C N , A)
 Detection of Prostate Cancer in Whole-Slide Images Through End-to-End Training With Image-Level Labels , IEEE Transactions on Medical Imaging , 2021年 , Volume: 40, Issue: 7 ,
 <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9380553>
- (58)調査した分野 (Int.Cl. , D B 名)
 G 0 6 T 7 / 0 0
 G 0 6 N 3 / 1 0