

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第6857368号  
(P6857368)

(45) 発行日 令和3年4月14日(2021.4.14)

(24) 登録日 令和3年3月24日(2021.3.24)

(51) Int. Cl.		F I			
<b>G06T</b>	<b>7/00</b>	<b>(2017.01)</b>	G06T	7/00	350C
<b>G06T</b>	<b>7/11</b>	<b>(2017.01)</b>	G06T	7/11	
<b>G06N</b>	<b>3/04</b>	<b>(2006.01)</b>	G06N	3/04	

請求項の数 24 (全 37 頁)

(21) 出願番号	特願2020-1957 (P2020-1957)	(73) 特許権者	517038176
(22) 出願日	令和2年1月9日(2020.1.9)		株式会社ストラドビジョン
(65) 公開番号	特開2020-119534 (P2020-119534A)		大韓民国 37668 キョンサンブク
(43) 公開日	令和2年8月6日(2020.8.6)		ド ポハン-シ ナム-グ ジゴク-ロ
審査請求日	令和2年1月9日(2020.1.9)		394 フィフス ベンチャー-ドン ス
(31) 優先権主張番号	16/255,044		イト 304-308
(32) 優先日	平成31年1月23日(2019.1.23)	(74) 代理人	110000855
(33) 優先権主張国・地域又は機関	米国 (US)		特許業務法人浅村特許事務所
		(72) 発明者	キム、ケイ - ヒョン
			大韓民国、ソウル、ソデムン - グ、ソ
			ソムン - ロ 27、ルーム ナンバー
			1004

最終頁に続く

(54) 【発明の名称】 CNN学習方法及び学習装置、これを利用したテスト方法及びテスト装置

(57) 【特許請求の範囲】

【請求項1】

イメージ認識のためのCNNのパラメータを学習する方法において、

(a) 少なくとも一つのトレーニングイメージが入力されると、学習装置が、(i) 少なくとも一つのコンボリューションレイヤをもって、前記トレーニングイメージ上でROIの領域をクロップ(Crop)し、リサイズ(Resize)して生成されたROIイメージそれぞれに対してコンボリューション演算を少なくとも一回適用させて少なくとも一つの物体に対応する少なくとも一つの前記ROIごとにROI特徴マップを生成させ、(ii) (ii-1) プーリングレイヤをもって、前記ROI特徴マップに対してプーリング演算を少なくとも一回適用させて前記ROIごとにプーリング済み特徴マップを生成させ、第1トランスポーズレイヤ(Transposing Layer)をもって、前記ROIごとの前記プーリング済み特徴マップ上の対応する同一の各位置のピクセルそれぞれを前記ROIごとにコンカチネート(Concatenating)させて統合特徴マップ(Integrated Feature Map)を生成させるか、(ii-2) 前記プーリングレイヤをもって、前記ROI特徴マップに対してプーリング演算を適用させて前記ROIごとの前記プーリング済み特徴マップを生成させ、前記プーリングレイヤをもって、前記ROIごとの前記プーリング済み特徴マップ上の対応する同一の各位置の前記ピクセルそれぞれを前記ROIごとにコンカチネートさせて前記統合特徴マップを生成させる段階；

(b) 前記学習装置が、(b1) (i) 第1リシェイプレイヤ(Reshaping

Layer)をもって、前記統合特徴マップのすべてのチャンネルのうちでそれぞれに対応するH1個のチャンネルから構成されたそれぞれのグループ内の特徴それぞれをコンカチネートさせて第1リシェイプ済み特徴マップ(Reshaped Feature Map)を生成させ、(ii)1×H1コンボリューションレイヤをもって、前記第1リシェイプ済み特徴マップに対して1×H1コンボリューション演算を適用させてボリューム(Volume)が調整された第1調整特徴マップ(Adjusted Feature Map)を生成させ、(b2)(i)第2リシェイプレイヤをもって、前記第1調整特徴マップのすべてのチャンネルのうちでそれぞれに対応するH2個のチャンネルから構成されたそれぞれのグループ内の特徴それぞれをコンカチネートさせて第2リシェイプ済み特徴マップを生成させ、(ii)1×H2コンボリューションレイヤをもって、前記第2リシェイプ済み特徴マップに対して1×H2コンボリューション演算を適用させてボリュームが調整された第2調整特徴マップを生成させる段階；及び

10

(c)前記学習装置が、(c1)(i)第2トランスポーズレイヤをもって、前記第2調整特徴マップを前記ピクセルごとに分離して前記ROIごとにピクセル別特徴マップ(Pixel-Wise Feature Map)を生成させ、分類レイヤをもって、前記ROIごとの前記ピクセル別特徴マップを利用して前記ROIそれぞれに関する物体情報を生成させるか、(ii)前記分類レイヤをもって、前記第2調整特徴マップを前記ピクセル別に分離して前記ROIごとに前記ピクセル別特徴マップを生成させ、前記分類レイヤをもって、前記ROIごとの前記ピクセル別特徴マップを利用して前記ROIそれぞれに対する前記物体情報を生成させた後、(c2)ロスレイヤをもって、前記物体情報とこれに対応する原本正解とを参照して少なくとも一つの物体ロスを算出させることにより前記物体ロスをバックプロパゲーションして前記1×H2コンボリューションレイヤ、前記1×H1コンボリューションレイヤ、及び前記コンボリューションレイヤのうち少なくとも一部のパラメータを学習する段階；

20

を含むことを特徴とする学習方法。

#### 【請求項2】

前記(b)段階で、

前記統合特徴マップのチャンネル個数が前記H1の倍数でない場合、前記学習装置は、前記第1リシェイプレイヤをもって、前記統合特徴マップに少なくとも一つの第1ダミーチャンネル(Dummy Channel)を追加させて少なくとも一つの前記第1ダミーチャンネルを含む前記統合特徴マップのチャンネル個数が前記H1の倍数になるようにした後、前記統合特徴マップの前記すべてのチャンネルのうちでそれぞれの対応するH1個のチャンネルから構成されたそれぞれの前記グループ内の前記特徴それぞれをコンカチネートさせ、

30

前記(b)段階で、

前記第1調整特徴マップのチャンネル個数が前記H2の倍数でない場合、前記学習装置は、前記第2リシェイプレイヤをもって、前記第1調整特徴マップに少なくとも一つの第2ダミーチャンネルを追加させて少なくとも一つの前記第2ダミーチャンネルを含む前記第1調整特徴マップのチャンネル個数が前記H2の倍数になるようにした後、前記第1調整特徴マップの前記すべてのチャンネルのうちでそれぞれの対応する前記H2個のチャンネルから構成されたそれぞれの前記グループ内の前記特徴それぞれをコンカチネートさせることを特徴とする請求項1に記載の学習方法。

40

#### 【請求項3】

前記ROIの個数をNとし、前記ROIごとの前記プーリング済み特徴マップの幅をM1、高さをM2とし、前記ROIごとの前記プーリング済み特徴マップのチャンネル個数をJとする場合、

前記(a)段階で、

前記学習装置は、(i)前記第1トランスポーズレイヤをもって、前記ROIごとの前記プーリング済み特徴マップを、幅N、高さ1、チャンネルM1×M2×Jである前記統合特徴マップに変換させるようにするか、(ii)前記プーリングレイヤをもって、前記

50

ROIごとの前記プーリング済み特徴マップを、幅N、高さ1、チャンネルM1×M2×Jである前記統合特徴マップに変換させるようにすることを特徴とする請求項1に記載の学習方法。

【請求項4】

前記1×H1コンボリューションレイヤのフィルタの個数をKとし、前記1×H2コンボリューションレイヤのフィルタの個数をLとする場合、

前記(c)段階で、

前記学習装置は、前記第1リシェイプレイヤをもって、N×H1のサイズと

$$\text{CEIL} \left( \frac{M1 \times M2 \times J}{H1} \right)$$

10

のチャンネル個数を有する前記第1リシェイプ済み特徴マップを生成させ、前記1×H1コンボリューションレイヤをもって、幅N、高さ1、チャンネルKであるN×1×Kのボリュームを有する前記第1調整特徴マップを生成させ、

前記第2リシェイプレイヤをもって、幅N、高さH2、チャンネル

$$\text{CEIL} \left( \frac{K}{H2} \right)$$

20

である前記第2リシェイプ済み特徴マップを生成させ、前記1×H2コンボリューションレイヤをもって、幅N、高さ1、チャンネルLであるN×1×Lのボリュームを有する前記第2調整特徴マップを生成させる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(c)段階で、

前記学習装置は、(i)前記第2トランスポーズレイヤをもって、前記第2調整特徴マップを、N個のROIそれぞれに対応する幅1、高さ1、チャンネルLである1×1×Lのボリュームを有する前記ROIごとの前記ピクセル別特徴マップに変換させるか、(ii)前記分類レイヤをもって、前記第2調整特徴マップを、前記N個のROIそれぞれに対応する幅1、高さ1、チャンネルLである1×1×Lのボリュームを有する前記ROIごとの前記ピクセル別特徴マップに変換させることを特徴とする請求項4に記載の学習方法。

30

【請求項6】

前記分類レイヤは、少なくとも一つのソフトマックス(Softmax)アルゴリズムを使用することを特徴とする請求項1に記載の学習方法。

【請求項7】

イメージ認識のためのCNNをテストする方法において、

(a)学習装置は、(1)(i)少なくとも一つのコンボリューションレイヤをもって、少なくとも一つのトレーニングイメージ上で学習用ROIの領域をクロップ(Crop)し、リサイズ(Resize)して生成された学習用ROIイメージそれぞれに対してコンボリューション演算を少なくとも一回適用させて少なくとも一つの学習用物体に対応する少なくとも一つの前記学習用ROIごとに学習用ROI特徴マップを生成させ、(ii)(ii-1)プーリングレイヤをもって、前記学習用ROI特徴マップに対してプーリング演算を少なくとも一回適用させて前記学習用ROIごとに学習用プーリング済み特徴マップを生成させ、第1トランスポーズレイヤ(Transposing Layer)をもって、前記学習用ROIごとの前記学習用プーリング済み特徴マップ上の対応する同一の各位置のピクセルそれぞれを前記学習用ROIごとにコンカチネート(Concatenating)させて学習用統合特徴マップ(Integrated Feature Map)を生成させるか、(ii-2)前記プーリングレイヤをもって、前記学習用ROI特徴マップに対してプーリング演算を適用させて前記学習用ROIごとの前記学習

40

50

用プーリング済み特徴マップを生成させ、前記プーリングレイヤをもって、前記学習用ROIごとの前記学習用プーリング済み特徴マップ上の対応する同一の各位置の前記ピクセルそれぞれを前記学習用ROIごとにコンカチネートさせて前記学習用統合特徴マップを生成させ、(2)(2-1)(i)第1リシェイプレイヤ(Reshaping Layer)をもって、前記学習用統合特徴マップのすべてのチャンネルのうちでそれぞれに対応するH1個のチャンネルから構成されたそれぞれのグループ内の特徴それぞれをコンカチネートさせて学習用第1リシェイプ済み特徴マップ(Reshaped Feature Map)を生成させ、(ii)1×H1コンボリューションレイヤをもって、前記学習用第1リシェイプ済み特徴マップに対して1×H1コンボリューション演算を適用させてボリューム(Volume)が調整された学習用第1調整特徴マップ(Adjusted Feature Map)を生成させ、(2-2)(i)第2リシェイプレイヤをもって、前記学習用第1調整特徴マップのすべてのチャンネルのうちでそれぞれに対応するH2個のチャンネルから構成されたそれぞれのグループ内の特徴それぞれをコンカチネートさせて学習用第2リシェイプ済み特徴マップを生成させ、(ii)1×H2コンボリューションレイヤをもって、前記学習用第2リシェイプ済み特徴マップに対して1×H2コンボリューション演算を適用させてボリュームが調整された学習用第2調整特徴マップを生成させ、(3)(3-1)(i)第2トランスポーズレイヤをもって、前記学習用第2調整特徴マップを前記ピクセル別に分離して前記学習用ROIごとに学習用ピクセル別特徴マップ(Pixel-Wise Feature Map)を生成させ、分類レイヤをもって、前記学習用ROIごとの前記学習用ピクセル別特徴マップを利用して前記学習用ROIそれぞれに関する学習用物体情報を生成させるか、(ii)前記分類レイヤをもって、前記学習用第2調整特徴マップを前記ピクセル別に分離して前記学習用ROIごとに前記学習用ピクセル別特徴マップを生成させ、前記分類レイヤをもって、前記学習用ROIごとの前記学習用ピクセル別特徴マップを利用して前記学習用ROIそれぞれに対する前記学習用物体情報を生成させた後、(3-2)ロスレイヤをもって、前記学習用物体情報とこれに対応する原本正解とを参照して少なくとも一つの物体ロスを算出させることにより、前記物体ロスをバックプロパゲーションして前記1×H2コンボリューションレイヤ、前記1×H1コンボリューションレイヤ、及び前記コンボリューションレイヤのうち少なくとも一部のパラメータを学習した状態で、少なくとも一つのテストイメージが取得されると、テスト装置が、(i)前記コンボリューションレイヤをもって、前記テストイメージ上でテスト用ROIの領域をクロップ(Crop)し、リサイズ(Resize)して生成されたテスト用ROIイメージそれぞれに対してコンボリューション演算を少なくとも一回適用させて少なくとも一つのテスト用物体に対応する少なくとも一つの前記テスト用ROIごとにテスト用ROI特徴マップを生成させ、(ii)(ii-1)前記プーリングレイヤをもって、前記テスト用ROI特徴マップに対してプーリング演算を少なくとも一回適用させて前記テスト用ROIごとにテスト用プーリング済み特徴マップを生成させ、前記第1トランスポーズレイヤ(Transposing Layer)をもって、前記テスト用ROIごとの前記テスト用プーリング済み特徴マップ上の対応する同一の各位置のピクセルそれぞれを前記テスト用ROIごとにコンカチネート(Concatenating)させてテスト用統合特徴マップ(Integrated Feature Map)を生成させるか、(ii-2)前記プーリングレイヤをもって、前記テスト用ROI特徴マップに対してプーリング演算を適用させて前記テスト用ROIごとの前記テスト用プーリング済み特徴マップを生成させ、前記プーリングレイヤをもって、前記テスト用ROIごとの前記テスト用プーリング済み特徴マップ上の対応する同一の各位置の前記ピクセルそれぞれを前記テスト用ROIごとにコンカチネートさせて前記テスト用統合特徴マップを生成させる段階；

(b)前記テスト装置が、(b1)(i)前記第1リシェイプレイヤ(Reshaping Layer)をもって、前記テスト用統合特徴マップのすべてのチャンネルのうちでそれぞれに対応するH1個のチャンネルから構成されたそれぞれのグループ内の特徴それぞれをコンカチネートさせてテスト用第1リシェイプ済み特徴マップ(Reshape

10

20

30

40

50

d Feature Map)を生成させ、(ii)前記1×H1コンボリューションレイヤをもって、前記テスト用第1リシェイプ済み特徴マップに対して1×H1コンボリューション演算を適用させてボリューム(Volume)が調整されたテスト用第1調整特徴マップ(Adjusted Feature Map)を生成させ、(b2)(i)前記第2リシェイプレイヤをもって、前記テスト用第1調整特徴マップのすべてのチャンネルのうちでそれぞれに対応するH2個のチャンネルから構成されたそれぞれのグループ内の特徴それぞれをコンカチネートさせてテスト用第2リシェイプ済み特徴マップを生成させ、(ii)前記1×H2コンボリューションレイヤをもって、前記テスト用第2リシェイプ済み特徴マップに対して1×H2コンボリューション演算を適用させてボリュームが調整されたテスト用第2調整特徴マップを生成させる段階；及び

10

(c)前記テスト装置が、(i)前記第2トランスポーズレイヤをもって、前記テスト用第2調整特徴マップを前記ピクセル別に分離して前記テスト用ROIごとにテスト用ピクセル別特徴マップ(Pixel-Wise Feature Map)を生成させ、前記分類レイヤをもって、前記テスト用ROIごとの前記テスト用ピクセル別特徴マップを利用して前記テスト用ROIそれぞれに関するテスト用物体情報を生成させるか、(ii)前記分類レイヤをもって、前記テスト用第2調整特徴マップを前記ピクセル別に分離して前記テスト用ROIごとに前記テスト用ピクセル別特徴マップを生成させ、前記分類レイヤをもって、前記テスト用ROIごとの前記テスト用ピクセル別特徴マップを利用して前記テスト用ROIそれぞれに対する前記テスト用物体情報を生成させる段階；

を含むことを特徴とするテスト方法。

20

【請求項8】

前記(b)段階で、

前記テスト用統合特徴マップのチャンネル個数が前記H1の倍数でない場合、前記テスト装置は、前記第1リシェイプレイヤをもって、前記テスト用統合特徴マップに少なくとも一つの第1ダミーチャンネル(Dummy Channel)を追加させて少なくとも一つの前記第1ダミーチャンネルを含む前記テスト用統合特徴マップのチャンネル個数が前記H1の倍数になるようにした後、前記テスト用統合特徴マップの前記すべてのチャンネルのうちでそれぞれの対応するH1個のチャンネルから構成されたそれぞれの前記グループ内の前記特徴それぞれをコンカチネートさせ、

前記(b)段階で、

30

前記テスト用第1調整特徴マップのチャンネル個数が前記H2の倍数でない場合、前記テスト装置は、前記第2リシェイプレイヤをもって、前記テスト用第1調整特徴マップに少なくとも一つの第2ダミーチャンネルを追加させて少なくとも一つの前記第2ダミーチャンネルを含む前記テスト用第1調整特徴マップのチャンネル個数が前記H2の倍数になるようにした後、前記テスト用第1調整特徴マップの前記すべてのチャンネルのうちでそれぞれの対応する前記H2個のチャンネルから構成されたそれぞれの前記グループ内の前記特徴それぞれをコンカチネートさせることを特徴とする請求項7に記載のテスト方法。

【請求項9】

前記テスト用ROIの個数をNとし、前記テスト用ROIごとの前記テスト用プーリング済み特徴マップの幅をM1、高さをM2とし、前記テスト用ROIごとの前記テスト用プーリング済み特徴マップのチャンネル個数をJとする場合、

40

前記(a)段階で、

前記テスト装置は、(i)前記第1トランスポーズレイヤをもって、前記テスト用ROIごとの前記テスト用プーリング済み特徴マップを幅N、高さ1、チャンネルM1×M2×Jである前記テスト用統合特徴マップに変換させるようにするか、(ii)前記プーリングレイヤをもって、前記テスト用ROIごとの前記テスト用プーリング済み特徴マップを幅N、高さ1、チャンネルM1×M2×Jである前記テスト用統合特徴マップに変換させるようにすることを特徴とする請求項7に記載のテスト方法。

【請求項10】

前記1×H1コンボリューションレイヤのフィルタの個数をKとし、前記1×H2コン

50

ポリューションレイヤのフィルタの個数を  $L$  とする場合、  
 前記 ( b ) 段階で、  
 前記テスト装置は、  
 前記第 1 リシェイプレイヤをもって、  $N \times H_1$  のサイズと

$$\text{CEIL} \left( \frac{M_1 \times M_2 \times J}{H_1} \right)$$

のチャンネル個数を有する前記テスト用第 1 リシェイプ済み特徴マップを生成させ、前記  $1 \times H_1$  コンポリューションレイヤをもって、幅  $N$ 、高さ 1、チャンネル  $K$  である  $N \times 1 \times K$  のポリュームを有する前記テスト用第 1 調整特徴マップを生成させ、  
 前記第 2 リシェイプレイヤをもって、幅  $N$ 、高さ  $H_2$  のサイズと

$$\text{CEIL} \left( \frac{K}{H_2} \right)$$

のチャンネル個数を有する前記テスト用第 2 リシェイプ済み特徴マップを生成させ、前記  $1 \times H_2$  コンポリューションレイヤをもって、幅  $N$ 、高さ 1、チャンネル個数  $L$  である  $N \times 1 \times L$  のポリュームを有する前記テスト用第 2 調整特徴マップを生成させることを特徴とする請求項 9 に記載のテスト方法。

【請求項 11】

前記 ( c ) 段階で、  
 前記テスト装置は、 ( i ) 前記第 2 トランスポーズレイヤをもって、前記テスト用第 2 調整特徴マップを、  $N$  個の前記テスト用 ROI それぞれに対応する幅 1、高さ 1、チャンネル  $L$  である  $1 \times 1 \times L$  のポリュームを有する前記テスト用 ROI ごとの前記テスト用ピクセル別特徴マップに変換させるか、 ( ii ) 前記分類レイヤをもって、前記テスト用第 2 調整特徴マップを、前記  $N$  個のテスト用 ROI それぞれに対応する幅 1、高さ 1、チャンネル  $L$  である  $1 \times 1 \times L$  のポリュームを有する前記テスト用 ROI ごとの前記テスト用ピクセル別特徴マップに変換させることを特徴とする請求項 10 に記載のテスト方法。

【請求項 12】

前記分類レイヤは、少なくとも一つのソフトマックス ( Softmax ) アルゴリズムを使用することを特徴とする請求項 7 に記載のテスト方法。

【請求項 13】

イメージ認識のための CNN を学習する学習装置において、  
 インストラクションを格納する少なくとも一つのメモリと、  
 ( I ) ( i ) 少なくとも一つのコンポリューションレイヤをもって、少なくとも一つのトレーニングイメージ上で ROI の領域をクロップ ( Crop ) し、リサイズ ( Resize ) して生成された ROI イメージそれぞれに対してコンポリューション演算を少なくとも一回適用させて少なくとも一つの物体に対応する少なくとも一つの前記 ROI ごとの ROI 特徴マップを生成させ、 ( ii ) ( ii - 1 ) プーリングレイヤをもって、前記 ROI 特徴マップに対してプーリング演算を少なくとも一回適用させて前記 ROI ごとのプーリング済み特徴マップを生成させ、第 1 トランスポーズレイヤ ( Transposing Layer ) をもって、前記 ROI ごとの前記プーリング済み特徴マップ上の対応する同一の各位置のピクセルそれぞれを前記 ROI ごとのコンカチネート ( Concatenating ) させて統合特徴マップ ( Integrated Feature Map ) を生成させるか、 ( ii - 2 ) 前記プーリングレイヤをもって、前記 ROI 特徴マップに対してプーリング演算を適用させて前記 ROI ごとの前記プーリング済み特徴マップを生成させ、前記プーリングレイヤをもって、前記 ROI ごとの前記プーリング済み特徴マップ上の対応する同一の各位置の前記ピクセルそれぞれを前記 ROI ごとのコンカチネートさせて前記統合特徴マップを生成させるプロセス、 ( II ) ( II - 1 ) ( i ) 第 1 リ

10

20

30

40

50

シェイプレイヤ (Reshaping Layer) をもって、前記統合特徴マップのすべてのチャンネルのうちでそれぞれに対応する  $H1$  個のチャンネルから構成されたそれぞれのグループ内の特徴それぞれをコンカチネートさせて第1リシェイプ済み特徴マップ (Reshaped Feature Map) を生成させ、(ii)  $1 \times H1$  コンボリューションレイヤをもって、前記第1リシェイプ済み特徴マップに対して  $1 \times H1$  コンボリューション演算を適用させてボリューム (Volume) が調整された第1調整特徴マップ (Adjusted Feature Map) を生成させ、(II-2)(i) 第2リシェイプレイヤをもって、前記第1調整特徴マップのすべてのチャンネルのうちでそれぞれに対応する  $H2$  個のチャンネルから構成されたそれぞれのグループ内の特徴それぞれをコンカチネートさせて第2リシェイプ済み特徴マップを生成させ、(ii)  $1 \times H2$  コンボリューションレイヤをもって、前記第2リシェイプ済み特徴マップに対して  $1 \times H2$  コンボリューション演算を適用させてボリュームが調整された第2調整特徴マップを生成させるプロセス、及び(III)(III-1)(i) 第2トランスポーズレイヤをもって、前記第2調整特徴マップを前記ピクセル別に分離して前記ROIごとにピクセル別特徴マップ (Pixel-Wise Feature Map) を生成させ、分類レイヤをもって、前記ROIごとの前記ピクセル別特徴マップを利用して前記ROIそれぞれに対する物体情報を生成させるか、(ii) 前記分類レイヤをもって、前記第2調整特徴マップを前記ピクセル別に分離して前記ROIごとに前記ピクセル別特徴マップを生成させ、前記分類レイヤをもって、前記ROIごとの前記ピクセル別特徴マップを利用して前記ROIそれぞれに対する前記物体情報を生成させた後、(III-2) ロスレイヤをもって、前記物体情報とこれに対応する原本正解とを参照して少なくとも一つの物体ロスを算出させることにより、前記物体ロスをバックプロパゲーションして前記  $1 \times H2$  コンボリューションレイヤ、前記  $1 \times H1$  コンボリューションレイヤ、及び前記コンボリューションレイヤのうち少なくとも一部のパラメータを学習するプロセスを遂行するための前記インスタクションを実行するように構成された少なくとも一つのプロセッサと、を含むことを特徴とする学習装置。

【請求項14】

前記(II)プロセスで、

前記統合特徴マップのチャンネル個数が前記  $H1$  の倍数でない場合、前記プロセッサは、前記第1リシェイプレイヤをもって、前記統合特徴マップに少なくとも一つの第1ダミーチャンネル (Dummy Channel) を追加させて少なくとも一つの前記第1ダミーチャンネルを含む前記統合特徴マップのチャンネル個数が前記  $H1$  の倍数になるようにした後、前記統合特徴マップの前記すべてのチャンネルのうちでそれぞれの対応する  $H1$  個のチャンネルから構成されたそれぞれの前記グループ内の前記特徴それぞれをコンカチネートさせ、

前記(II)プロセスで、

前記第1調整特徴マップのチャンネル個数が前記  $H2$  の倍数でない場合、前記プロセッサは、前記第2リシェイプレイヤをもって、前記第1調整特徴マップに少なくとも一つの第2ダミーチャンネルを追加させて少なくとも一つの前記第2ダミーチャンネルを含む前記第1調整特徴マップのチャンネル個数が前記  $H2$  の倍数になるようにした後、前記第1調整特徴マップの前記すべてのチャンネルのうちでそれぞれの対応する前記  $H2$  個の前記チャンネルから構成されたそれぞれの前記グループ内の前記特徴それぞれをコンカチネートさせることを特徴とする請求項13に記載の学習装置。

【請求項15】

前記ROIの個数を  $N$  とし、前記ROIごとの前記プーリング済み特徴マップの幅を  $M1$ 、高さを  $M2$  とし、前記ROIごとの前記プーリング済み特徴マップのチャンネル個数を  $J$  とする場合、

前記(I)プロセスで、

前記プロセッサは(i) 前記第1トランスポーズレイヤをもって、前記ROIごとの前記プーリング済み特徴マップを幅  $N$ 、高さ1、チャンネル  $M1 \times M2 \times J$  である前記統合

10

20

30

40

50

特徴マップに変換させるようにするか、( i i ) 前記プーリングレイヤをもって、前記 R O I ごとの前記プーリング済み特徴マップを幅 N、高さ 1、チャンネル M 1 × M 2 × J である前記統合特徴マップに変換させるようにすることを特徴とする請求項 1 3 に記載の学習装置。

【請求項 1 6】

前記 1 × H 1 コンボリューションレイヤのフィルタの個数を K とし、前記 1 × H 2 コンボリューションレイヤのフィルタの個数を L とする場合、

前記 ( I I ) プロセスで、

前記プロセッサは、前記第 1 リシェイプレイヤをもって、N × H 1 のサイズと

$$\text{CEIL} \left( \frac{M 1 \times M 2 \times J}{H 1} \right)$$

10

のチャンネル個数を有する前記第 1 リシェイプ済み特徴マップを生成させ、前記 1 × H 1 コンボリューションレイヤをもって、幅 N、高さ 1、チャンネル K である N × 1 × K のボリュームを有する前記第 1 調整特徴マップを生成させ、

前記プロセッサは、前記第 2 リシェイプレイヤをもって、幅 N、高さ H 2、チャンネル

$$\text{CEIL} \left( \frac{K}{H 2} \right)$$

20

である前記第 2 リシェイプ済み特徴マップを出力させ、前記 1 × H 2 コンボリューションレイヤをもって、幅 N、高さ 1、チャンネル L である N × 1 × L のボリュームを有する前記第 2 調整特徴マップを生成させることを特徴とする請求項 1 5 に記載の学習装置。

【請求項 1 7】

前記 ( I I I ) プロセスで、

前記プロセッサは、( i ) 前記第 2 トランスポーズレイヤをもって、前記第 2 調整特徴マップを、N 個の R O I それぞれに対応する幅 1、高さ 1、チャンネル L である 1 × 1 × L のボリュームを有する前記 R O I ごとの前記ピクセル別特徴マップに変換させるか、( i i ) 前記分類レイヤをもって、前記第 2 調整特徴マップを、N 個の R O I それぞれに対応する幅 1、高さ 1、チャンネル L である 1 × 1 × L のボリュームを有する前記 R O I ごとの前記ピクセル別特徴マップに変換させることを特徴とする請求項 1 6 に記載の学習装置。

30

【請求項 1 8】

前記分類レイヤは、少なくとも一つのソフトマックス ( S o f t m a x ) アルゴリズムを使用することを特徴とする請求項 1 3 に記載の学習装置。

【請求項 1 9】

イメージ認識のための CNN をテストするテスト装置において、

インストラクションを格納する少なくとも一つのメモリと、

学習装置は、( 1 ) ( i ) 少なくとも一つのコンボリューションレイヤをもって、少なくとも一つのトレーニングイメージ上で学習用 R O I の領域をクロップ ( C r o p ) し、リサイズ ( R e s i z e ) して生成された学習用 R O I イメージそれぞれに対してコンボリューション演算を少なくとも一回適用させて少なくとも一つの学習用物体に対応する少なくとも一つの前記学習用 R O I ごとに学習用 R O I 特徴マップを生成させ、( i i ) ( i i - 1 ) プーリングレイヤをもって、前記学習用 R O I 特徴マップに対してプーリング演算を少なくとも一回適用させて前記学習用 R O I ごとに学習用プーリング済み特徴マップを生成させ、第 1 トランスポーズレイヤ ( T r a n s p o s i n g L a y e r ) をもって、前記学習用 R O I ごとの前記学習用プーリング済み特徴マップ上の対応する同一の各位置のピクセルそれぞれを前記学習用 R O I ごとにコンカチネート ( C o n c a t e n a t i n g ) させて学習用統合特徴マップ ( I n t e g r a t e d F e a t u r e M

40

50

ap)を生成させるか、(ii-2)前記プーリングレイヤをもって、前記学習用ROI特徴マップに対してプーリング演算を適用させて前記学習用ROIごとの前記学習用プーリング済み特徴マップを生成させ、前記プーリングレイヤをもって、前記学習用ROIごとの前記学習用プーリング済み特徴マップ上の対応する同一の各位置の前記ピクセルそれぞれを前記学習用ROIごとにコンカチネートさせて前記学習用統合特徴マップを生成させ、(2)(2-1)(i)第1リシェイプレイヤ(Reshaping Layer)をもって、前記学習用統合特徴マップのすべてのチャンネルのうちでそれぞれに対応するH1個のチャンネルから構成されたそれぞれのグループ内の特徴それぞれをコンカチネートさせて学習用第1リシェイプ済み特徴マップ(Reshaped Feature Map)を生成させ、(ii)1xH1コンボリューションレイヤをもって、前記学習用第1リシェイプ済み特徴マップに対して1xH1コンボリューション演算を適用させてボリューム(Volume)が調整された学習用第1調整特徴マップ(Adjusted Feature Map)を生成させ、(2-2)(i)第2リシェイプレイヤをもって、前記学習用第1調整特徴マップのすべてのチャンネルのうちでそれぞれに対応するH2個のチャンネルから構成されたそれぞれのグループ内の特徴それぞれをコンカチネートさせて学習用第2リシェイプ済み特徴マップを生成させ、(ii)1xH2コンボリューションレイヤをもって、前記学習用第2リシェイプ済み特徴マップに対して1xH2コンボリューション演算を適用させてボリュームが調整された学習用第2調整特徴マップを生成させ、(3)(3-1)(i)第2トランスポーズレイヤをもって、前記学習用第2調整特徴マップを前記ピクセルごとに分離して前記学習用ROIごとの学習用ピクセル別特徴マップ(Pixel-Wise Feature Map)を生成させ、分類レイヤをもって、前記学習用ROIごとの前記学習用ピクセル別特徴マップを利用して前記学習用ROIそれぞれに関する学習用物体クラス情報を生成させるか、(ii)前記分類レイヤをもって、前記学習用第2調整特徴マップを前記ピクセルごとに分離して前記学習用ROIごとに前記学習用ピクセル別特徴マップを生成させ、前記分類レイヤをもって、前記学習用ROIごとの前記学習用ピクセル別特徴マップを利用して前記学習用ROIそれぞれに関する前記学習用物体クラス情報を生成させた後、(3-2)ロスレイヤをもって、前記学習用物体情報とこれに対応する原本正解とを参照して少なくとも一つの物体ロスを算出させることにより、前記物体ロスをバックプロパゲーションして前記1xH2コンボリューションレイヤ、前記1xH1コンボリューションレイヤ、及び前記コンボリューションレイヤのうち少なくとも一部のパラメータを学習した状態で、(I)(i)前記コンボリューションレイヤをもって、少なくとも一つのテストイメージ上でテスト用ROIの領域をクロップ(Crop)し、リサイズ(Resize)して生成されたテスト用ROIイメージそれぞれに対してコンボリューション演算を少なくとも一回適用させて少なくとも一つのテスト用物体に対応する少なくとも一つの前記テスト用ROIごとにテスト用ROI特徴マップを生成させ、(ii)(ii-1)前記プーリングレイヤをもって、前記テスト用ROI特徴マップに対してプーリング演算を少なくとも一回適用させて前記テスト用ROIごとにテスト用プーリング済み特徴マップを生成させ、前記第1トランスポーズレイヤ(Transposing Layer)をもって、前記テスト用ROIごとの前記テスト用プーリング済み特徴マップ上の対応する同一の各位置のピクセルそれぞれを前記テスト用ROIごとにコンカチネート(Concatenating)させてテスト用統合特徴マップ(Integrated Feature Map)を生成させるか、(ii-2)前記プーリングレイヤをもって、前記テスト用ROI特徴マップに対してプーリング演算を適用させて前記テスト用ROIごとの前記テスト用プーリング済み特徴マップを生成させ、前記プーリングレイヤをもって、前記テスト用ROIごとの前記テスト用プーリング済み特徴マップ上の対応する同一の各位置の前記ピクセルそれぞれを前記テスト用ROIごとにコンカチネートさせて前記テスト用統合特徴マップを生成させるプロセス、(II)(II-1)(i)前記第1リシェイプレイヤ(Reshaping Layer)をもって、前記テスト用統合特徴マップのすべてのチャンネルのうちでそれぞれに対応するH1個のチャンネルから構成されたそれぞれのグループ内の特徴それぞれをコン

10

20

30

40

50

カチネートさせてテスト用第1リシェイプ済み特徴マップ (Reshaped Feature Map) を生成させ、(ii) 前記  $1 \times H1$  コンボリューションレイヤをもって、前記テスト用第1リシェイプ済み特徴マップに対して  $1 \times H1$  コンボリューション演算を適用させてボリューム (Volume) が調整されたテスト用第1調整特徴マップ (Adjusted Feature Map) を生成させ、(II-2) (i) 前記第2リシェイプレイヤをもって、前記テスト用第1調整特徴マップのすべてのチャンネルのうちでそれぞれに対応する  $H2$  個のチャンネルから構成されたそれぞれのグループ内の特徴それぞれをコンカチネートさせてテスト用第2リシェイプ済み特徴マップを生成させ、(ii) 前記  $1 \times H2$  コンボリューションレイヤをもって、前記テスト用第2リシェイプ済み特徴マップに対して  $1 \times H2$  コンボリューション演算を適用させてボリュームが調整されたテスト用第2調整特徴マップを生成させるプロセス、及び(III) (i) 前記第2トランスポーズレイヤをもって、前記テスト用第2調整特徴マップを前記ピクセル別に分離して前記テスト用ROIごとにテスト用ピクセル別特徴マップ (Pixel-Wise Feature Map) を生成させ、前記分類レイヤをもって、前記テスト用ROIごとの前記テスト用ピクセル別特徴マップを利用して前記テスト用ROIそれぞれに関するテスト用物体情報を生成させるか、(ii) 前記分類レイヤをもって、前記テスト用第2調整特徴マップを前記ピクセル別に分離して前記テスト用ROIごとの前記テスト用ピクセル別特徴マップを生成させ、前記分類レイヤをもって、前記テスト用ROIごとの前記テスト用ピクセル別特徴マップを利用して前記テスト用ROIそれぞれに対する前記テスト用物体情報を生成させるプロセスを遂行するための前記インストラクションを実行するように構成された少なくともひとつのプロセッサと、

10

20

を含むことを特徴とするテスト装置。

【請求項20】

前記(II)プロセスで、

前記テスト用統合特徴マップのチャンネル個数が前記  $H1$  の倍数でない場合、前記プロセッサは、前記第1リシェイプレイヤをもって、前記テスト用統合特徴マップに少なくともひとつの第1ダミーチャンネル (Dummy Channel) を追加させて少なくともひとつの前記第1ダミーチャンネルを含む前記テスト用統合特徴マップのチャンネル個数が前記  $H1$  の倍数になるようにした後、前記テスト用統合特徴マップの前記すべてのチャンネルのうちでそれぞれの対応する  $H1$  個のチャンネルから構成されたそれぞれの前記グループ内の前記特徴それぞれをコンカチネートさせ、

30

前記(II)プロセスで、

前記テスト用第1調整特徴マップのチャンネル個数が前記  $H2$  の倍数でない場合、前記プロセッサは、前記第2リシェイプレイヤをもって、前記テスト用第1調整特徴マップに少なくともひとつの第2ダミーチャンネルを追加させて少なくともひとつの前記第2ダミーチャンネルを含む前記テスト用第1調整特徴マップのチャンネル個数が前記  $H2$  の倍数になるようにした後、前記テスト用第1調整特徴マップの前記すべてのチャンネルのうちでそれぞれの対応する前記  $H2$  個の前記チャンネルから構成されたそれぞれの前記グループ内の前記特徴それぞれをコンカチネートさせることを特徴とする請求項19に記載のテスト装置。

40

【請求項21】

前記テスト用ROIの個数を  $N$  とし、前記テスト用ROIごとの前記テスト用プーリング済み特徴マップの幅を  $M1$ 、高さを  $M2$  とし、前記テスト用ROIごとの前記テスト用プーリング済み特徴マップのチャンネル個数を  $J$  とする場合、

前記(I)プロセスで、

前記プロセッサは、(i) 前記第1トランスポーズレイヤをもって、前記テスト用ROIごとの前記テスト用プーリング済み特徴マップを幅  $N$ 、高さ1、チャンネル  $M1 \times M2 \times J$  である前記テスト用統合特徴マップに変換させるようにするか、(ii) 前記プーリングレイヤをもって、前記テスト用ROIごとの前記テスト用プーリング済み特徴マップを幅  $N$ 、高さ1、チャンネル  $M1 \times M2 \times J$  である前記テスト用統合特徴マップに変換さ

50

せるようにすることを特徴とする請求項 19 に記載のテスト装置。

【請求項 22】

前記 1 × H1 コンボリューションレイヤのフィルタの個数を K とし、前記 1 × H2 コンボリューションレイヤのフィルタの個数を L とする場合、

前記 (II) プロセスで、

前記プロセッサは、前記第 1 リシェイプレイヤをもって、N × H1 のサイズと

$$\text{CEIL} \left( \frac{M1 \times M2 \times J}{H1} \right)$$

10

のチャンネル個数を有する前記テスト用第 1 リシェイプ済み特徴マップを生成させ、前記 1 × H1 コンボリューションレイヤをもって、幅 N、高さ 1、チャンネル K である N × 1 × K のボリュームを有する前記テスト用第 1 調整特徴マップを生成させ、

前記プロセッサは、前記第 2 リシェイプレイヤをもって、幅 N、高さ H2、チャンネル

$$\text{CEIL} \left( \frac{K}{H2} \right)$$

である前記テスト用第 2 リシェイプ済み特徴マップを生成させ、前記 1 × H2 コンボリューションレイヤをもって、幅 N、高さ 1、チャンネル L である N × 1 × L のボリュームを有する前記テスト用第 2 調整特徴マップを生成させることを特徴とする請求項 21 に記載のテスト装置。

20

【請求項 23】

前記 (III) プロセスで、

前記プロセッサは、(i) 前記第 2 トランスポーズレイヤをもって、前記テスト用第 2 調整特徴マップを、N 個のテスト用 ROI それぞれに対応する幅 1、高さ 1、チャンネル L である 1 × 1 × L のボリュームを有する前記テスト用 ROI ごとの前記テスト用ピクセル別特徴マップに変換させるか、(ii) 前記分類レイヤをもって、前記テスト用第 2 調整特徴マップを、前記 N 個のテスト用 ROI それぞれに対応する幅 1、高さ 1、チャンネル L である 1 × 1 × L のボリュームを有する前記テスト用 ROI ごとの前記テスト用ピクセル別特徴マップに変換させることを特徴とする請求項 22 に記載のテスト装置。

30

【請求項 24】

前記分類レイヤは、少なくとも一つのソフトマックス (Softmax) アルゴリズムを使用することを特徴とする請求項 19 に記載のテスト装置。

【発明の詳細な説明】

【技術分野】

【0001】

ハードウェア最適化に使用されるイメージ認識のための 1 × H コンボリューションを利用した CNN 学習方法及び学習装置、これを利用したテスト方法及びテスト装置 { LEARNING METHOD AND LEARNING DEVICE FOR CONVOLUTIONAL NEURAL NETWORK USING 1 × H CONVOLUTION FOR IMAGE RECOGNITION TO BE USED FOR HARDWARE OPTIMIZATION, AND TESTING METHOD AND TESTING DEVICE USING THE SAME } を提供する。

40

【0002】

本発明はハードウェア最適化に使用されるイメージ認識のための CNN を学習する方法に関し、より詳細には、イメージ認識のための前記 CNN のパラメータを学習する方法において、(a) 少なくとも一つのトレーニングイメージが入力されると、学習装置が、(i) 少なくとも一つのコンボリューションレイヤをもって、前記トレーニングイメージ上

50

でROIの領域をクロップ(Crop)し、リサイズ(Resize)して生成されたROIイメージそれぞれに対してコンボリューション演算を少なくとも一回適用させて少なくとも一つの物体に対応する少なくとも一つの前記ROIごとにROI特徴マップを生成させ、(ii)(i-1)プーリングレイヤをもって、前記ROI特徴マップに対してプーリング演算を少なくとも一回適用させて前記ROIごとにプーリング済み特徴マップを生成させ、第1トランスポーズレイヤ(Transposing Layer)をもって、前記ROIごとの前記プーリング済み特徴マップ上の対応する同一の各位置のピクセルそれぞれを前記ROIごとにコンカチネート(Concatenating)させて統合特徴マップ(Integrated Feature Map)を生成させるか、(ii-2)前記プーリングレイヤをもって、前記ROI特徴マップに対してプーリング演算を適用させて前記ROIごとの前記プーリング済み特徴マップを生成させ、前記プーリングレイヤをもって、前記ROIごとの前記プーリング済み特徴マップ上の対応する同一の各位置の前記ピクセルそれぞれを前記ROIごとにコンカチネートさせて前記統合特徴マップを生成させる段階；(b)前記学習装置が、(b1)(i)第1リシェイプレイヤ(Reshaping Layer)をもって、前記統合特徴マップのすべてのチャンネルのうちでそれぞれに対応するH1個のチャンネルから構成されたそれぞれのグループ内の特徴それぞれをコンカチネートさせて第1リシェイプ済み特徴マップ(Reshaped Feature Map)を生成させ、(ii)1×H1コンボリューションレイヤをもって、前記第1リシェイプ済み特徴マップに対して1×H1コンボリューション演算を適用させてボリューム(Volume)が調整された第1調整特徴マップ(Adjusted Feature Map)を生成させ、(b2)(i)第2リシェイプレイヤをもって、前記第1調整特徴マップのすべてのチャンネルのうちでそれぞれに対応するH2個のチャンネルから構成されたそれぞれのグループ内の特徴それぞれをコンカチネートさせて第2リシェイプ済み特徴マップを生成させ、(ii)1×H2コンボリューションレイヤをもって、前記第2リシェイプ済み特徴マップに対して1×H2コンボリューション演算を適用させてボリュームが調整された第2調整特徴マップを生成させる段階；及び(c)前記学習装置が、(c1)(i)第2トランスポーズレイヤをもって、前記第2調整特徴マップを前記ピクセル別に分離して前記ROIごとにピクセル別特徴マップ(Pixel-Wise Feature Map)を生成させ、分類レイヤをもって、前記ROIごとの前記ピクセル別特徴マップを利用して前記ROIそれぞれに対する物体情報を生成させるか、(ii)前記分類レイヤをもって、前記第2調整特徴マップを前記ピクセル別に分離して前記ROIごとに前記ピクセル別特徴マップを生成させ、前記分類レイヤをもって、前記ROIごとの前記ピクセル別特徴マップを利用して前記ROIそれぞれに対する前記物体情報を生成させた後、(c2)ロスレイヤをもって、前記物体情報とこれに対応する原本正解とを参照して少なくとも一つの物体ロスを算出させることにより、前記物体ロスをバックプロパゲーションして前記1×H2コンボリューションレイヤ、前記1×H1コンボリューションレイヤ、及び前記コンボリューションレイヤのうち少なくとも一部のパラメータを学習する段階；を含むことを特徴とする学習方法、学習装置及びこれを利用したテスト方法、テスト装置に関する。

【背景技術】

【0003】

ディープコンボリューションニューラルネットワーク(Deep Convolutional Neural Networks; Deep CNN)は、ディープラーニング分野で起きた驚くべき発展の核心である。CNNは、文字の認識問題を解決するために90年代にすでに使用されていたが、現在のように広く使用されるようになったのは最近の研究結果の賜物である。このようなディープCNNは、2012年ImageNetイメージ分類コンテストで他の競争相手に勝って優勝を収めた。そして、コンボリューションニューラルネットワークは機械学習(Machine Learning)分野で非常に有用なツールとなった。

【0004】

そして、CNNはイメージから特徴を抽出する特徴抽出器 (Feature Extractor) 及びイメージ内の物体を検出するか、特徴抽出器が抽出した特徴を参照してイメージを認識する特徴分類機 (Feature Classifier) を含む。

【0005】

そして、CNNの特徴抽出器はコンボリューションレイヤから構成され、特徴分類器は特徴抽出器が抽出した特徴に対してFC演算 (Fully Connected Operation) を適用することができるFCレイヤから構成される。

【0006】

しかしながら、このようなFCレイヤは入力された全ての特徴に対して重み付け値を生成しなければならない不便さがあり、入力されたすべての特徴に対してFC演算を遂行しなければならないため、コンピューティング演算量が多くなる短所がある。

10

【0007】

また、FCレイヤではFCレイヤに入力されるイメージのサイズがFCレイヤモデルの所定サイズと一致しなければならない。そのため、このようなFCレイヤを含むCNNに予め設定されたサイズと異なるトレーニングイメージまたはテストイメージが入力される場合、コンボリューションレイヤがトレーニングイメージやテストイメージに対して当該演算を適用することができるが、FCレイヤでは前記所定サイズと異なるサイズの入力イメージが入力されるため、トレーニングイメージやテストイメージに該当演算を適用することができない。

【0008】

20

したがって、本発明者は、前記のようなFCレイヤの問題点を克服することができるようにするイメージ認識のためのCNNを提案することにする。

【発明の概要】

【発明が解決しようとする課題】

【0009】

本発明は、前述した問題点を全て解決することをその目的とする。

【0010】

本発明は、入力イメージのサイズに関係なしに、少なくとも一つの入力イメージ内に位置する少なくとも一つの物体を認識することができるようにするイメージ認識のためのCNNを提供することを他の目的とする。

30

【0011】

本発明は、特徴分類に使用される重み付け値の個数を最小化し得るようにするイメージ認識のためのCNNを提供することをまた他の目的とする。

【0012】

本発明は、特徴分類のためのコンピューティング演算量を減少させ得るようにするイメージ認識のためのCNNを提供することをまた他の目的とする。

【課題を解決するための手段】

【0013】

前記のような本発明の目的を達成し、後述する本発明の特徴的な効果を実現するための本発明の特徴的な構成は以下の通りである。

40

【0014】

本発明の一態様によると、イメージ認識のためのCNNのパラメータを学習する方法において、(a) 少なくとも一つのトレーニングイメージが入力されると、学習装置が、(i) 少なくとも一つのコンボリューションレイヤをもって、前記トレーニングイメージ上でROIの領域をクロップ (Crop) し、リサイズ (Resize) して生成されたROIイメージそれぞれに対してコンボリューション演算を少なくとも一回適用させて少なくとも一つの物体に対応する少なくとも一つの前記ROIごとにROI特徴マップを生成させ、(ii) (ii-1) プーリングレイヤをもって、前記ROI特徴マップに対してプーリング演算を少なくとも一回適用させて前記ROIごとにプーリング済み特徴マップを生成させ、第1トランスポーズレイヤ (Transposing Layer) をもつ

50

て、前記ROIごとの前記プーリング済み特徴マップ上の対応する同一の各位置のピクセルそれぞれを前記ROIごとにコンカチネート(Concatenating)させて統合特徴マップ(Integrated Feature Map)を生成させるか、(ii-2)前記プーリングレイヤをもって、前記ROI特徴マップに対してプーリング演算を適用させて前記ROIごとの前記プーリング済み特徴マップを生成させ、前記プーリングレイヤをもって、前記ROIごとの前記プーリング済み特徴マップ上の対応する同一の各位置の前記ピクセルそれぞれを前記ROIごとにコンカチネートさせて前記統合特徴マップを生成させる段階；(b)前記学習装置が、(b1)(i)第1リシェイプレイヤ(Reshaping Layer)をもって、前記統合特徴マップのすべてのチャンネルのうちでそれぞれに対応するH1個のチャンネルから構成されたそれぞれのグループ内の特徴それぞれをコンカチネートさせて第1リシェイプ済み特徴マップ(Reshaped Feature Map)を生成させ、(ii)1×H1コンボリューションレイヤをもって、前記第1リシェイプ済み特徴マップに対して1×H1コンボリューション演算を適用させてボリューム(Volume)が調整された第1調整特徴マップ(Adjusted Feature Map)を生成させ、(b2)(i)第2リシェイプレイヤをもって、前記第1調整特徴マップのすべてのチャンネルのうちでそれぞれに対応するH2個のチャンネルから構成されたそれぞれのグループ内の特徴それぞれをコンカチネートさせて第2リシェイプ済み特徴マップを生成させ、(ii)1×H2コンボリューションレイヤをもって、前記第2リシェイプ済み特徴マップに対して1×H2コンボリューション演算を適用させてボリュームが調整された第2調整特徴マップを生成させる段階；及び(c)前記学習装置が、(c1)(i)第2トランスポーズレイヤをもって、前記第2調整特徴マップを前記ピクセルごとに分離して前記ROIごとにピクセル別特徴マップ(Pixel-Wise Feature Map)を生成させ、分類レイヤをもって、前記ROIごとの前記ピクセル別特徴マップを利用して前記ROIそれぞれに関する物体情報を生成させるか、(ii)前記分類レイヤをもって、前記第2調整特徴マップを前記ピクセル別に分離して前記ROIごとに前記ピクセル別特徴マップを生成させ、前記分類レイヤをもって、前記ROIごとの前記ピクセル別特徴マップを利用して前記ROIそれぞれに対する前記物体情報を生成させた後、(c2)ロスレイヤをもって、前記物体情報とこれに対応する原本正解とを参照して少なくとも一つの物体ロスを算出させることにより前記物体ロスをバックプロパゲーションして前記1×H2コンボリューションレイヤ、前記1×H1コンボリューションレイヤ、及び前記コンボリューションレイヤのうち少なくとも一部のパラメータを学習する段階；を含むことを特徴とする。

#### 【0015】

一実施例において、前記(b)段階で、前記統合特徴マップのチャンネル個数が前記H1の倍数でない場合、前記学習装置は、前記第1リシェイプレイヤをもって、前記統合特徴マップに少なくとも一つの第1ダミーチャンネル(Dummy Channel)を追加させて少なくとも一つの前記第1ダミーチャンネルを含む前記統合特徴マップのチャンネル個数が前記H1の倍数になるようにした後、前記統合特徴マップの前記すべてのチャンネルのうちでそれぞれの対応するH1個のチャンネルから構成されたそれぞれの前記グループ内の前記特徴それぞれをコンカチネートさせ、前記(b)段階で、前記第1調整特徴マップのチャンネル個数が前記H2の倍数でない場合、前記学習装置は、前記第2リシェイプレイヤをもって、前記第1調整特徴マップに少なくとも一つの第2ダミーチャンネルを追加させて少なくとも一つの前記第2ダミーチャンネルを含む前記第1調整特徴マップのチャンネル個数が前記H2の倍数になるようにした後、前記第1調整特徴マップの前記すべてのチャンネルのうちでそれぞれの対応する前記H2個のチャンネルから構成されたそれぞれの前記グループ内の前記特徴それぞれをコンカチネートさせることを特徴とする。

#### 【0016】

一実施例において、前記ROIの個数をNとし、前記ROIごとの前記プーリング済み特徴マップの幅をM1、高さをM2とし、前記ROIごとの前記プーリング済み特徴マッ

プのチャンネル個数を  $J$  とする場合、前記 ( a ) 段階で、前記学習装置は、( i ) 前記第 1 トランスポーズレイヤをもって、前記 ROI ごとの前記プーリング済み特徴マップを、幅  $N$ 、高さ 1、チャンネル  $M_1 \times M_2 \times J$  である前記統合特徴マップに変換させるようにするか、( i i ) 前記プーリングレイヤをもって、前記 ROI ごとの前記プーリング済み特徴マップを、幅  $N$ 、高さ 1、チャンネル  $M_1 \times M_2 \times J$  である前記統合特徴マップに変換させるようにすることを特徴とする。

【 0 0 1 7 】

一実施例において、前記  $1 \times H_1$  コンボリューションレイヤのフィルタの個数を  $K$  とし、前記  $1 \times H_2$  コンボリューションレイヤのフィルタの個数を  $L$  とする場合、前記 ( b ) 段階で、前記学習装置は、前記第 1 リシェイプレイヤをもって、 $N \times H_1$  のサイズと

10

$$\text{CEIL} \left( \frac{M_1 \times M_2 \times J}{H_1} \right)$$

のチャンネル個数を有する前記第 1 リシェイプ済み特徴マップを生成させ、前記  $1 \times H_1$  コンボリューションレイヤをもって、幅  $N$ 、高さ 1、チャンネル  $K$  である  $N \times 1 \times K$  のボリュームを有する前記第 1 調整特徴マップを生成させ、前記第 2 リシェイプレイヤをもって、幅  $N$ 、高さ  $H_2$ 、チャンネル

20

$$\text{CEIL} \left( \frac{K}{H_2} \right)$$

である前記第 2 リシェイプ済み特徴マップを生成させ、前記  $1 \times H_2$  コンボリューションレイヤをもって、幅  $N$ 、高さ 1、チャンネル  $L$  である  $N \times 1 \times L$  のボリュームを有する前記第 2 調整特徴マップを生成させることを特徴とする。

【 0 0 1 8 】

一実施例において、前記 ( c ) 段階で、前記学習装置は ( i ) 前記第 2 トランスポーズレイヤをもって、前記第 2 調整特徴マップを、 $N$  個の ROI それぞれに対応する幅 1、高さ 1、チャンネル  $L$  である  $1 \times 1 \times L$  のボリュームを有する前記 ROI ごとの前記ピクセル別特徴マップに変換させるか、( i i ) 前記分類レイヤをもって、前記第 2 調整特徴マップを、前記  $N$  個の ROI それぞれに対応する幅 1、高さ 1、チャンネル  $L$  である  $1 \times 1 \times L$  のボリュームを有する前記 ROI ごとの前記ピクセル別特徴マップに変換させることを特徴とする。

30

【 0 0 1 9 】

一実施例において、前記分類レイヤは、少なくとも一つのソフトマックス ( Softmax ) アルゴリズムを使用することを特徴とする。

【 0 0 2 0 】

本発明の他の態様によると、イメージ認識のための CNN をテストする方法において、( a ) 学習装置は、( 1 ) ( i ) 少なくとも一つのコンボリューションレイヤをもって、少なくとも一つのトレーニングイメージ上で学習用 ROI の領域をクロップ ( Crop ) し、リサイズ ( Resize ) して生成された学習用 ROI イメージそれぞれに対してコンボリューション演算を少なくとも一回適用させて少なくとも一つの学習用物体に対応する少なくとも一つの前記学習用 ROI ごとに学習用 ROI 特徴マップを生成させ、( i i ) ( i i - 1 ) プーリングレイヤをもって、前記学習用 ROI 特徴マップに対してプーリング演算を少なくとも一回適用させて前記学習用 ROI ごとに学習用プーリング済み特徴マップを生成させ、第 1 トランスポーズレイヤ ( Transposing Layer ) をもって、前記学習用 ROI ごとの前記学習用プーリング済み特徴マップ上の対応する同一の各位置のピクセルそれぞれを前記学習用 ROI ごとにコンカチネート ( Concatenating ) させて学習用統合特徴マップ ( Integrated Feature

40

50

Map)を生成させるか、(ii-2)前記プーリングレイヤをもって、前記学習用ROI特徴マップに対してプーリング演算を適用させて前記学習用ROIごとの前記学習用プーリング済み特徴マップを生成させ、前記プーリングレイヤをもって、前記学習用ROIごとの前記学習用プーリング済み特徴マップ上の対応する同一の各位置の前記ピクセルそれぞれを前記学習用ROIごとにコンカチネートさせて前記学習用統合特徴マップを生成させ、(2)(2-1)(i)第1リシェイプレイヤ(Reshaping Layer)をもって、前記学習用統合特徴マップのすべてのチャンネルのうちでそれぞれに対応するH1個のチャンネルから構成されたそれぞれのグループ内の特徴それぞれをコンカチネートさせて学習用第1リシェイプ済み特徴マップ(Reshaped Feature Map)を生成させ、(ii)1×H1コンボリューションレイヤをもって、前記学習用第1リシェイプ済み特徴マップに対して1×H1コンボリューション演算を適用させてボリューム(Volume)が調整された学習用第1調整特徴マップ(Adjusted Feature Map)を生成させ、(2-2)(i)第2リシェイプレイヤをもって、前記学習用第1調整特徴マップのすべてのチャンネルのうちでそれぞれに対応するH2個のチャンネルから構成されたそれぞれのグループ内の特徴それぞれをコンカチネートさせて学習用第2リシェイプ済み特徴マップを生成させ、(ii)1×H2コンボリューションレイヤをもって、前記学習用第2リシェイプ済み特徴マップに対して1×H2コンボリューション演算を適用させてボリュームが調整された学習用第2調整特徴マップを生成させ、(3)(3-1)(i)第2トランスポーズレイヤをもって、前記学習用第2調整特徴マップを前記ピクセル別に分離して前記学習用ROIごとに学習用ピクセル別特徴マップ(Pixel-Wise Feature Map)を生成させ、分類レイヤをもって、前記学習用ROIごとの前記学習用ピクセル別特徴マップを利用して前記学習用ROIそれぞれに関する学習用物体情報を生成させるか、(ii)前記分類レイヤをもって、前記学習用第2調整特徴マップを前記ピクセル別に分離して前記学習用ROIごとに前記学習用ピクセル別特徴マップを生成させ、前記分類レイヤをもって、前記学習用ROIごとの前記学習用ピクセル別特徴マップを利用して前記学習用ROIそれぞれに対する前記学習用物体情報を生成させた後、(3-2)ロスレイヤをもって、前記学習用物体情報とこれに対応する原本正解とを参照して少なくとも一つの物体ロスを算出させることにより、前記物体ロスをバックプロパゲーションして前記1×H2コンボリューションレイヤ、前記1×H1コンボリューションレイヤ、及び前記コンボリューションレイヤのうち少なくとも一部のパラメータを学習した状態で、少なくとも一つのテストイメージが取得されると、テスト装置が、(i)前記コンボリューションレイヤをもって、前記テストイメージ上でテスト用ROIの領域をクロップ(Crop)し、リサイズ(Resize)して生成されたテスト用ROIイメージそれぞれに対してコンボリューション演算を少なくとも一回適用させて少なくとも一つのテスト用物体に対応する少なくとも一つの前記テスト用ROIごとにテスト用ROI特徴マップを生成させ、(ii)(ii-1)前記プーリングレイヤをもって、前記テスト用ROI特徴マップに対してプーリング演算を少なくとも一回適用させて前記テスト用ROIごとにテスト用プーリング済み特徴マップを生成させ、前記第1トランスポーズレイヤ(Transposing Layer)をもって、前記テスト用ROIごとの前記テスト用プーリング済み特徴マップ上の対応する同一の各位置のピクセルそれぞれを前記テスト用ROIごとにコンカチネート(Concatenating)させてテスト用統合特徴マップ(Integrated Feature Map)を生成させるか、(ii-2)前記プーリングレイヤをもって、前記テスト用ROI特徴マップに対してプーリング演算を適用させて前記テスト用ROIごとの前記テスト用プーリング済み特徴マップを生成させ、前記プーリングレイヤをもって、前記テスト用ROIごとの前記テスト用プーリング済み特徴マップ上の対応する同一の各位置の前記ピクセルそれぞれを前記テスト用ROIごとにコンカチネートさせて前記テスト用統合特徴マップを生成させる段階；(b)前記テスト装置が、(b1)(i)前記第1リシェイプレイヤ(Reshaping Layer)をもって、前記テスト用統合特徴マップのすべてのチャンネルのうちでそれぞれに対応するH1個のチャンネルから構成されたそ

10

20

30

40

50

それぞれのグループ内の特徴それぞれをコンカチネートさせてテスト用第1リシェイプ済み特徴マップ (Reshaped Feature Map) を生成させ、(ii) 前記  $1 \times H1$  コンボリューションレイヤをもって、前記テスト用第1リシェイプ済み特徴マップに対して  $1 \times H1$  コンボリューション演算を適用させてボリューム (Volume) が調整されたテスト用第1調整特徴マップ (Adjusted Feature Map) を生成させ、(b2)(i) 前記第2リシェイプレイヤをもって、前記テスト用第1調整特徴マップのすべてのチャンネルのうちでそれぞれに対応する  $H2$  個のチャンネルから構成されたそれぞれのグループ内の特徴それぞれをコンカチネートさせてテスト用第2リシェイプ済み特徴マップを生成させ、(ii) 前記  $1 \times H2$  コンボリューションレイヤをもって、前記テスト用第2リシェイプ済み特徴マップに対して  $1 \times H2$  コンボリューション演算を適用させてボリュームが調整されたテスト用第2調整特徴マップを生成させる段階；及び(c) 前記テスト装置が、(i) 前記第2トランスポーズレイヤをもって、前記テスト用第2調整特徴マップを前記ピクセル別に分離して前記テスト用ROIごとにテスト用ピクセル別特徴マップ (Pixel-Wise Feature Map) を生成させ、前記分類レイヤをもって、前記テスト用ROIごとの前記テスト用ピクセル別特徴マップを利用して前記テスト用ROIそれぞれに関するテスト用物体情報を生成させるか、(ii) 前記分類レイヤをもって、前記テスト用第2調整特徴マップを前記ピクセル別に分離して前記テスト用ROIごとに前記テスト用ピクセル別特徴マップを生成させ、前記分類レイヤをもって、前記テスト用ROIごとの前記テスト用ピクセル別特徴マップを利用して前記テスト用ROIそれぞれに対する前記テスト用物体情報を生成させる段階；を含むことを特徴とする。

10

20

## 【0021】

一実施例において、前記(b)段階で、前記テスト用統合特徴マップのチャンネル個数が前記  $H1$  の倍数でない場合、前記テスト装置は、前記第1リシェイプレイヤをもって、前記テスト用統合特徴マップに少なくとも一つの第1ダミーチャンネル (Dummy Channel) を追加させて少なくとも一つの前記第1ダミーチャンネルを含む前記テスト用統合特徴マップのチャンネル個数が前記  $H1$  の倍数になるようにした後、前記テスト用統合特徴マップの前記すべてのチャンネルのうちでそれぞれの対応する  $H1$  個のチャンネルから構成されたそれぞれの前記グループ内の前記特徴それぞれをコンカチネートさせ、前記(b)段階で、前記テスト用第1調整特徴マップのチャンネル個数が前記  $H2$  の倍数でない場合、前記テスト装置は、前記第2リシェイプレイヤをもって、前記テスト用第1調整特徴マップに少なくとも一つの第2ダミーチャンネルを追加させて少なくとも一つの前記第2ダミーチャンネルを含む前記テスト用第1調整特徴マップのチャンネル個数が前記  $H2$  の倍数になるようにした後、前記テスト用第1調整特徴マップの前記すべてのチャンネルのうちでそれぞれの対応する前記  $H2$  個のチャンネルから構成されたそれぞれの前記グループ内の前記特徴それぞれをコンカチネートさせることを特徴とする。

30

## 【0022】

一実施例において、前記テスト用ROIの個数を  $N$  とし、前記テスト用ROIごとの前記テスト用プーリング済み特徴マップの幅を  $M1$ 、高さを  $M2$  とし、前記テスト用ROIごとの前記テスト用プーリング済み特徴マップのチャンネル個数を  $J$  とする場合、前記(a)段階で、前記テスト装置は、(i) 前記第1トランスポーズレイヤをもって、前記テスト用ROIごとの前記テスト用プーリング済み特徴マップを幅  $N$ 、高さ1、チャンネル  $M1 \times M2 \times J$  である前記テスト用統合特徴マップに変換させるようにするか、(ii) 前記プーリングレイヤをもって、前記テスト用ROIごとの前記テスト用プーリング済み特徴マップを幅  $N$ 、高さ1、チャンネル  $M1 \times M2 \times J$  である前記テスト用統合特徴マップに変換させるようにすることを特徴とする。

40

## 【0023】

一実施例において、前記  $1 \times H1$  コンボリューションレイヤのフィルタの個数を  $K$  とし、前記  $1 \times H2$  コンボリューションレイヤのフィルタの個数を  $L$  とする場合、前記(b)段階で、前記テスト装置は、前記第1リシェイプレイヤをもって、 $N \times H1$  のサイズと

50

$$\text{CEIL} \left( \frac{M1 \times M2 \times J}{H1} \right)$$

のチャンネル個数を有する前記テスト用第1リシェイプ済み特徴マップを生成させ、前記1×H1コンボリューションレイヤをもって、幅N、高さ1、チャンネルKであるN×1×Kのボリュームを有する前記テスト用第1調整特徴マップを生成させ、前記第2リシェイプレイヤをもって、幅N、高さH2のサイズと

$$\text{CEIL} \left( \frac{K}{H2} \right)$$

10

のチャンネル個数を有する前記テスト用第2リシェイプ済み特徴マップを生成させ、前記1×H2コンボリューションレイヤをもって、幅N、高さ1、チャンネル個数LであるN×1×Lのボリュームを有する前記テスト用第2調整特徴マップを生成させることを特徴とする。

#### 【0024】

一実施例において、前記(c)段階で、前記テスト装置は、(i)前記第2トランスポーズレイヤをもって、前記テスト用第2調整特徴マップを、N個のテスト用ROIそれぞれに対応する幅1、高さ1、チャンネルLである1×1×Lのボリュームを有する前記テスト用ROIごとの前記テスト用ピクセル別特徴マップに変換させるか、(ii)前記分類レイヤをもって、前記テスト用第2調整特徴マップを、前記N個のテスト用ROIそれぞれに対応する幅1、高さ1、チャンネルLである1×1×Lのボリュームを有する前記テスト用ROIごとの前記テスト用ピクセル別特徴マップに変換させることを特徴とする。

20

#### 【0025】

一実施例において、前記分類レイヤは、少なくとも一つのソフトマックス(Softmax)アルゴリズムを使用することを特徴とする。

#### 【0026】

本発明のまた他の態様によると、イメージ認識のためのCNNを学習する学習装置において、インストラクションを格納する少なくとも一つのメモリと、(I)(i)少なくとも一つのコンボリューションレイヤをもって、少なくとも一つのトレーニングイメージ上でROIの領域をクロップ(Crop)し、リサイズ(Resize)して生成されたROIイメージそれぞれに対してコンボリューション演算を少なくとも一回適用させて少なくとも一つの物体に対応する少なくとも一つの前記ROIごとにROI特徴マップを生成させ、(ii)(ii-1)プーリングレイヤをもって、前記ROI特徴マップに対してプーリング演算を少なくとも一回適用させて前記ROIごとにプーリング済み特徴マップを生成させ、第1トランスポーズレイヤ(Transposing Layer)をもって、前記ROIごとの前記プーリング済み特徴マップ上の対応する同一の各位置のピクセルそれぞれを前記ROIごとにコンカチネート(Concatenating)させて統合特徴マップ(Integrated Feature Map)を生成させるか、(ii-2)前記プーリングレイヤをもって、前記ROI特徴マップに対してプーリング演算を適用させて前記ROIごとの前記プーリング済み特徴マップを生成させ、前記プーリングレイヤをもって、前記ROIごとの前記プーリング済み特徴マップ上の対応する同一の各位置の前記ピクセルそれぞれを前記ROIごとにコンカチネートさせて前記統合特徴マップを生成させるプロセス、(II)(II-1)(i)第1リシェイプレイヤ(Reshaping Layer)をもって、前記統合特徴マップのすべてのチャンネルのうちでそれぞれに対応するH1個のチャンネルから構成されたそれぞれのグループ内の特徴それぞれをコンカチネートさせて第1リシェイプ済み特徴マップ(Reshaped Feature Map)を生成させ、(ii)1×H1コンボリューションレイヤをもって

30

40

50

、前記第1リシェイプ済み特徴マップに対して $1 \times H1$ コンボリューション演算を適用させてボリューム(Volume)が調整された第1調整特徴マップ(Adjusted Feature Map)を生成させ、(II-2)(i)第2リシェイプレイヤをもって、前記第1調整特徴マップのすべてのチャンネルのうちでそれぞれに対応する $H2$ 個のチャンネルから構成されたそれぞれのグループ内の特徴それぞれをコンカチネートさせて第2リシェイプ済み特徴マップを生成させ、(ii) $1 \times H2$ コンボリューションレイヤをもって、前記第2リシェイプ済み特徴マップに対して $1 \times H2$ コンボリューション演算を適用させてボリュームが調整された第2調整特徴マップを生成させるプロセス、及び(III)(III-1)(i)第2トランスポーズレイヤをもって、前記第2調整特徴マップを前記ピクセル別に分離して前記ROIごとにピクセル別特徴マップ(Pixel-Wise Feature Map)を生成させ、分類レイヤをもって、前記ROIごとの前記ピクセル別特徴マップを利用して前記ROIそれぞれに対する物体情報を生成させるか、(ii)前記分類レイヤをもって、前記第2調整特徴マップを前記ピクセル別に分離して前記ROIごとに前記ピクセル別特徴マップを生成させ、前記分類レイヤをもって、前記ROIごとの前記ピクセル別特徴マップを利用して前記ROIそれぞれに対する前記物体情報を生成させた後、(III-2)ロスレイヤをもって、前記物体情報とこれに対応する原本正解とを参照して少なくとも一つの物体ロスを算出させることにより、前記物体ロスをバックプロパゲーションして前記 $1 \times H2$ コンボリューションレイヤ、前記 $1 \times H1$ コンボリューションレイヤ、及び前記コンボリューションレイヤのうち少なくとも一部のパラメータを学習するプロセスを遂行するための前記インストラクションを実行するように構成された少なくとも一つのプロセッサと、を含むことを特徴とする。

【0027】

一実施例において、前記(II)プロセスで、前記統合特徴マップのチャンネル個数が前記 $H1$ の倍数でない場合、前記プロセッサは、前記第1リシェイプレイヤをもって、前記統合特徴マップに少なくとも一つの第1ダミーチャンネル(Dummy Channel)を追加させて少なくとも一つの前記第1ダミーチャンネルを含む前記統合特徴マップのチャンネル個数が前記 $H1$ の倍数になるようにした後、前記統合特徴マップの前記すべてのチャンネルのうちでそれぞれの対応する $H1$ 個のチャンネルから構成されたそれぞれの前記グループ内の前記特徴それぞれをコンカチネートさせ、前記(II)プロセスで、前記第1調整特徴マップのチャンネル個数が前記 $H2$ の倍数でない場合、前記プロセッサは、前記第2リシェイプレイヤをもって、前記第1調整特徴マップに少なくとも一つの第2ダミーチャンネルを追加させて少なくとも一つの前記第2ダミーチャンネルを含む前記第1調整特徴マップのチャンネル個数が前記 $H2$ の倍数になるようにした後、前記第1調整特徴マップの前記すべてのチャンネルのうちでそれぞれの対応する前記 $H2$ 個の前記チャンネルから構成されたそれぞれの前記グループ内の前記特徴それぞれをコンカチネートさせることを特徴とする。

【0028】

一実施例において、前記ROIの個数を $N$ とし、前記ROIごとの前記プーリング済み特徴マップの幅を $M1$ 、高さを $M2$ とし、前記ROIごとの前記プーリング済み特徴マップのチャンネル個数を $J$ とする場合、前記(I)プロセスで、前記プロセッサは(i)前記第1トランスポーズレイヤをもって、前記ROIごとの前記プーリング済み特徴マップを幅 $N$ 、高さ1、チャンネル $M1 \times M2 \times J$ である前記統合特徴マップに変換させるようにするか、(ii)前記プーリングレイヤをもって、前記ROIごとの前記プーリング済み特徴マップを幅 $N$ 、高さ1、チャンネル $M1 \times M2 \times J$ である前記統合特徴マップに変換させるようにすることを特徴とする。

【0029】

一実施例において、前記 $1 \times H1$ コンボリューションレイヤのフィルタの個数を $K$ とし、前記 $1 \times H2$ コンボリューションレイヤのフィルタの個数を $L$ とする場合、前記(II)プロセスで、前記プロセッサは、前記第1リシェイプレイヤをもって、 $N \times H1$ のサイズと

10

20

30

40

50

$$\text{CEIL} \left( \frac{M1 \times M2 \times J}{H1} \right)$$

のチャンネル個数を有する前記第1リシェイプ済み特徴マップを生成させ、前記1×H1コンポリューションレイヤをもって、幅N、高さ1、チャンネルKであるN×1×Kのボリュームを有する前記第1調整特徴マップを生成させ、前記プロセッサは、前記第2リシェイプレイヤをもって、幅N、高さH2、チャンネル

$$\text{CEIL} \left( \frac{K}{H2} \right)$$

10

である前記第2リシェイプ済み特徴マップを出力させ、前記1×H2コンポリューションレイヤをもって、幅N、高さ1、チャンネルLであるN×1×Lのボリュームを有する前記第2調整特徴マップを生成させることを特徴とする。

【0030】

一実施例において、前記(III)プロセスで、前記プロセッサは、(i)前記第2トランスポーズレイヤをもって、前記第2調整特徴マップを、N個のROIそれぞれに対応する幅1、高さ1、チャンネルLである1×1×Lのボリュームを有する前記ROIごとの前記ピクセル別特徴マップに変換させるか、(ii)前記分類レイヤをもって、前記第2調整特徴マップを、前記N個のROIそれぞれに対応する幅1、高さ1、チャンネルLである1×1×Lのボリュームを有する前記ROIごとの前記ピクセル別特徴マップに変換させることを特徴とする。

20

【0031】

一実施例において、前記分類レイヤは、少なくとも一つのソフトマックス(Softmax)アルゴリズムを使用することを特徴とする。

【0032】

本発明のまた他の態様によると、イメージ認識のためのCNNをテストするテスト装置において、インストラクションを格納する少なくとも一つのメモリと、学習装置は、(1) (i)少なくとも一つのコンポリューションレイヤをもって、少なくとも一つのトレーニングイメージ上で学習用ROIの領域をクロップ(Crop)し、リサイズ(Resize)して生成された学習用ROIイメージそれぞれに対してコンポリューション演算を少なくとも一回適用させて少なくとも一つの学習用物体に対応する少なくとも一つの前記学習用ROIごとに学習用ROI特徴マップを生成させ、(ii) (ii-1)プーリングレイヤをもって、前記学習用ROI特徴マップに対してプーリング演算を少なくとも一回適用させて前記学習用ROIごとに学習用プーリング済み特徴マップを生成させ、第1トランスポーズレイヤ(Transposing Layer)をもって、前記学習用ROIごとの前記学習用プーリング済み特徴マップ上の対応する同一の各位置のピクセルそれぞれを前記学習用ROIごとにコンカチネート(Concatenating)させて学習用統合特徴マップ(Integrated Feature Map)を生成させるか、(ii-2)前記プーリングレイヤをもって、前記学習用ROI特徴マップに対してプーリング演算を適用させて前記学習用ROIごとの前記学習用プーリング済み特徴マップを生成させ、前記プーリングレイヤをもって、前記学習用ROIごとの前記学習用プーリング済み特徴マップ上の対応する同一の各位置の前記ピクセルそれぞれを前記学習用ROIごとにコンカチネートさせて前記学習用統合特徴マップを生成させ、(2) (2-1) (i)第1リシェイプレイヤ(Reshaping Layer)をもって、前記学習用統合特徴マップのすべてのチャンネルのうちでそれぞれに対応するH1個のチャンネルから構成されたそれぞれのグループ内の特徴それぞれをコンカチネートさせて学習用第1リシェイプ済み特徴マップ(Reshaped Feature Map)を生成させ、

30

40

50

( i i ) 1 x H 1 コンボリユーションレイヤをもって、前記学習用第 1 リシェイプ済み特徴マップに対して 1 x H 1 コンボリユーション演算を適用させてボリューム ( Volume ) が調整された学習用第 1 調整特徴マップ ( Adjusted Feature Map ) を生成させ、( 2 - 2 ) ( i ) 第 2 リシェイプレイヤをもって、前記学習用第 1 調整特徴マップのすべてのチャンネルのうちでそれぞれに対応する H 2 個のチャンネルから構成されたそれぞれのグループ内の特徴それぞれをコンカチネートさせて学習用第 2 リシェイプ済み特徴マップを生成させ、( i i ) 1 x H 2 コンボリユーションレイヤをもって、前記学習用第 2 リシェイプ済み特徴マップに対して 1 x H 2 コンボリユーション演算を適用させてボリュームが調整された学習用第 2 調整特徴マップを生成させ、( 3 ) ( 3 - 1 ) ( i ) 第 2 トランスポーズレイヤをもって、前記学習用第 2 調整特徴マップを前記ピクセルごとに分離して前記学習用 ROI ごとの学習用ピクセル別特徴マップ ( Pixel - Wise Feature Map ) を生成させ、分類レイヤをもって、前記学習用 ROI ごとの前記学習用ピクセル別特徴マップを利用して前記学習用 ROI それぞれに関する学習用物体クラス情報を生成させるか、( i i ) 前記分類レイヤをもって、前記学習用第 2 調整特徴マップを前記ピクセルごとに分離して前記学習用 ROI ごとに前記学習用ピクセル別特徴マップを生成させ、前記分類レイヤをもって、前記学習用 ROI ごとの前記学習用ピクセル別特徴マップを利用して前記学習用 ROI それぞれに関する前記学習用物体クラス情報を生成させた後、( 3 - 2 ) ロスレイヤをもって、前記学習用物体情報とこれに対応する原本正解とを参照して少なくとも一つの物体ロスを算出させることにより、前記物体ロスをバックプロパゲーションして前記 1 x H 2 コンボリユーションレイヤ、前記 1 x H 1 コンボリユーションレイヤ、及び前記コンボリユーションレイヤのうち少なくとも一部のパラメータを学習した状態で、( I ) ( i ) 前記コンボリユーションレイヤをもって、少なくとも一つのテストイメージ上でテスト用 ROI の領域をクロップ ( Crop ) し、リサイズ ( Resize ) して生成されたテスト用 ROI イメージそれぞれに対してコンボリユーション演算を少なくとも一回適用させて少なくとも一つのテスト用物体に対応する少なくとも一つの前記テスト用 ROI ごとにテスト用 ROI 特徴マップを生成させ、( i i ) ( i i - 1 ) 前記プーリングレイヤをもって、前記テスト用 ROI 特徴マップに対してプーリング演算を少なくとも一回適用させて前記テスト用 ROI ごとにテスト用プーリング済み特徴マップを生成させ、前記第 1 トランスポーズレイヤ ( Transposing Layer ) をもって、前記テスト用 ROI ごとの前記テスト用プーリング済み特徴マップ上の対応する同一の各位置のピクセルそれぞれを前記テスト用 ROI ごとにコンカチネート ( Concatenating ) させてテスト用統合特徴マップ ( Integrated Feature Map ) を生成させるか、( i i - 2 ) 前記プーリングレイヤをもって、前記テスト用 ROI 特徴マップに対してプーリング演算を適用させて前記テスト用 ROI ごとの前記テスト用プーリング済み特徴マップを生成させ、前記プーリングレイヤをもって、前記テスト用 ROI ごとの前記テスト用プーリング済み特徴マップ上の対応する同一の各位置の前記ピクセルそれぞれを前記テスト用 ROI ごとにコンカチネートさせて前記テスト用統合特徴マップを生成させるプロセス、( I I ) ( I I - 1 ) ( i ) 前記第 1 リシェイプレイヤ ( Reshaping Layer ) をもって、前記テスト用統合特徴マップのすべてのチャンネルのうちでそれぞれに対応する H 1 個のチャンネルから構成されたそれぞれのグループ内の特徴それぞれをコンカチネートさせてテスト用第 1 リシェイプ済み特徴マップ ( Reshaped Feature Map ) を生成させ、( i i ) 前記 1 x H 1 コンボリユーションレイヤをもって、前記テスト用第 1 リシェイプ済み特徴マップに対して 1 x H 1 コンボリユーション演算を適用させてボリューム ( Volume ) が調整されたテスト用第 1 調整特徴マップ ( Adjusted Feature Map ) を生成させ、( I I - 2 ) ( i ) 前記第 2 リシェイプレイヤをもって、前記テスト用第 1 調整特徴マップのすべてのチャンネルのうちでそれぞれに対応する H 2 個のチャンネルから構成されたそれぞれのグループ内の特徴それぞれをコンカチネートさせてテスト用第 2 リシェイプ済み特徴マップを生成させ、( i i ) 前記 1 x H 2 コンボリユーションレイヤをもって、前記テスト用第 2 リシェイプ済み特徴マップに対して

10

20

30

40

50

1 × H 2 コンボリユーション演算を適用させてボリュームが調整されたテスト用第 2 調整特徴マップを生成させるプロセス、及び ( I I I ) ( i ) 前記第 2 トランスポーズレイヤをもって、前記テスト用第 2 調整特徴マップを前記ピクセル別に分離して前記テスト用 R O I ごとにテスト用ピクセル別特徴マップ ( P i x e l - W i s e F e a t u r e M a p ) を生成させ、前記分類レイヤをもって、前記テスト用 R O I ごとの前記テスト用ピクセル別特徴マップを利用して前記テスト用 R O I それぞれに関するテスト用物体情報を生成させるか、( i i ) 前記分類レイヤをもって、前記テスト用第 2 調整特徴マップを前記ピクセル別に分離して前記テスト用 R O I ごとの前記テスト用ピクセル別特徴マップを生成させ、前記分類レイヤをもって、前記テスト用 R O I ごとの前記テスト用ピクセル別特徴マップを利用して前記テスト用 R O I それぞれに対する前記テスト用物体情報を生成させるプロセスを遂行するための前記インストラクションを実行するように構成された少なくともひとつのプロセッサと、を含むことを特徴とする。

10

## 【 0 0 3 3 】

一実施例において、前記 ( I I ) プロセスで、前記テスト用統合特徴マップのチャンネル個数が前記 H 1 の倍数でない場合、前記プロセッサは、前記第 1 リシェイプレイヤをもって、前記テスト用統合特徴マップに少なくともひとつの第 1 ダミーチャンネル ( D u m m y C h a n n e l ) を追加させて少なくともひとつの前記第 1 ダミーチャンネルを含む前記テスト用統合特徴マップのチャンネル個数が前記 H 1 の倍数になるようにした後、前記テスト用統合特徴マップの前記すべてのチャンネルのうちでそれぞれの対応する H 1 個のチャンネルから構成されたそれぞれの前記グループ内の前記特徴それぞれをコンカチネートさせ、前記 ( I I ) プロセスで、前記テスト用第 1 調整特徴マップのチャンネル個数が前記 H 2 の倍数でない場合、前記プロセッサは、前記第 2 リシェイプレイヤをもって、前記テスト用第 1 調整特徴マップに少なくともひとつの第 2 ダミーチャンネルを追加させて少なくともひとつの前記第 2 ダミーチャンネルを含む前記テスト用第 1 調整特徴マップのチャンネル個数が前記 H 2 の倍数になるようにした後、前記テスト用第 1 調整特徴マップの前記すべてのチャンネルのうちでそれぞれの対応する前記 H 2 個の前記チャンネルから構成されたそれぞれの前記グループ内の前記特徴それぞれをコンカチネートさせることを特徴とする。

20

## 【 0 0 3 4 】

一実施例において、前記テスト用 R O I の個数を N とし、前記テスト用 R O I ごとの前記テスト用プーリング済み特徴マップの幅を M 1、高さを M 2 とし、前記テスト用 R O I ごとの前記テスト用プーリング済み特徴マップのチャンネル個数を J とする場合、前記 ( I ) プロセスで、前記プロセッサは、( i ) 前記第 1 トランスポーズレイヤをもって、前記テスト用 R O I ごとの前記テスト用プーリング済み特徴マップを幅 N、高さ 1、チャンネル M 1 × M 2 × J である前記テスト用統合特徴マップに変換させるようにするか、( i i ) 前記プーリングレイヤをもって、前記テスト用 R O I ごとの前記テスト用プーリング済み特徴マップを幅 N、高さ 1、チャンネル M 1 × M 2 × J である前記テスト用統合特徴マップに変換させるようにすることを特徴とする。

30

## 【 0 0 3 5 】

一実施例において、前記 1 × H 1 コンボリユーションレイヤのフィルタの個数を K とし、前記 1 × H 2 コンボリユーションレイヤのフィルタの個数を L とする場合、前記 ( I I ) プロセスで、前記プロセッサは、前記第 1 リシェイプレイヤをもって、N × H 1 のサイズと

40

$$C E I L \left( \frac{M 1 \times M 2 \times J}{H 1} \right)$$

のチャンネル個数を有する前記テスト用第 1 リシェイプ済み特徴マップを生成させ、前記 1 × H 1 コンボリユーションレイヤをもって、幅 N、高さ 1、チャンネル K である N × 1 × K のボリュームを有する前記テスト用第 1 調整特徴マップを生成させ、前記プロセッサ

50

は、前記第2リシェイプレイヤをもって、幅N、高さH2、チャンネル

$$C E I L \left( \frac{K}{H^2} \right)$$

である前記テスト用第2リシェイプ済み特徴マップを生成させ、前記1×H2コンボリューションレイヤをもって、幅N、高さ1、チャンネルLであるN×1×Lのボリュームを有する前記テスト用第2調整特徴マップを生成させることを特徴とする。

【0036】

一実施例において、前記(III)プロセスで、前記プロセッサは、(i)前記第2トランスポートレイヤをもって、前記テスト用第2調整特徴マップを、N個のテスト用ROIそれぞれに対応する幅1、高さ1、チャンネルLである1×1×Lのボリュームを有する前記テスト用ROIごとの前記テスト用ピクセル別特徴マップに変換させるか、(ii)前記分類レイヤをもって、前記テスト用第2調整特徴マップを、前記N個のテスト用ROIそれぞれに対応する幅1、高さ1、チャンネルLである1×1×Lのボリュームを有する前記テスト用ROIごとの前記テスト用ピクセル別特徴マップに変換させることを特徴とする。

10

【0037】

一実施例において、前記分類レイヤは、少なくとも一つのソフトマックス(Softmax)アルゴリズムを使用することを特徴とする。

20

【0038】

その他にも、本発明の方法を実行するためのコンピュータプログラムを格納するためのコンピュータ読取り可能な記録媒体がさらに提供される。

【発明の効果】

【0039】

本発明は、コンボリューションレイヤのみでイメージ認識のためのCNNを具現することにより、入力イメージのサイズに関係なしに入力イメージ内に位置する物体を認識することができる効果がある。

【0040】

また、本発明は、コンボリューションレイヤのみでイメージ認識のためのCNNを具現することにより、従来のFCレイヤを利用する場合に比べて特徴分類のために使用される重み付け値の個数を最小化することができる他の効果がある。

30

【0041】

また、本発明は、コンボリューションレイヤのみでイメージ認識のためのCNNを具現することにより、従来のFCレイヤを利用する場合に比べて特徴分類のためのコンピューティング演算量を減少させることができるまた他の効果がある。

【0042】

本発明の実施例の説明に利用されるために添付された以下の図面は、本発明の実施例のうち単に一部であるに過ぎず、本発明の属する技術分野において通常の知識を有する者(以下「通常の技術者」)にとっては、発明的作業が行われずにこれらの図面に基づいて他の各図面が得られ得る。

40

【図面の簡単な説明】

【0043】

【図1】本発明の一実施例によるイメージ認識のためのCNNを学習する学習装置を簡略に示したものである。

【図2】本発明の一実施例によるイメージ認識のためのCNNを学習する方法を簡略に示したものである。

【図3】本発明の一実施例によるイメージ認識のためのCNNを学習するのに使用され、プーリングレイヤによって生成されたROIごとのプーリング済み特徴マップを簡略に示したものである。

50

【図4】本発明の一実施例によるイメージ認識のためのCNNを学習するのに使用され、ROIごとのプリーング済み特徴マップを利用して生成された統合特徴マップを簡略に示したものである。

【図5a】本発明の一実施例による(i)統合特徴マップをリシェイプして生成された第1リシェイプ済み特徴マップと、(ii)第1リシェイプ済み特徴マップに対して $1 \times H_1$ コンボリューション演算を適用して生成された第1調整特徴マップを簡略に示したものである。

【図5b】本発明の一実施例による(i)統合特徴マップをリシェイプして生成された第1リシェイプ済み特徴マップと、(ii)第1リシェイプ済み特徴マップに対して $1 \times H_1$ コンボリューション演算を適用して生成された第1調整特徴マップを簡略に示したものである。

10

【図6a】本発明の一実施例による(i)第1調整特徴マップをリシェイプして生成された第2リシェイプ済み特徴マップと、(ii)第2リシェイプ済み特徴マップに対して $1 \times H_2$ コンボリューション演算を適用して生成された第2調整特徴マップを簡略に示したものである。

【図6b】本発明の一実施例による(i)第1調整特徴マップをリシェイプして生成された第2リシェイプ済み特徴マップと、(ii)第2リシェイプ済み特徴マップに対して $1 \times H_2$ コンボリューション演算を適用して生成された第2調整特徴マップを簡略に示したものである。

【図7】本発明の一実施例によるイメージ認識のためのCNNを学習するのに使用され、第2調整特徴マップを利用して生成されたROIごとのピクセル別特徴マップを簡略に示したものである。

20

【図8】本発明の一実施例によるイメージ認識のためのCNNをテストするテスト装置を簡略に示したものである。

【図9】本発明の一実施例によるイメージ認識のためのCNNをテストする方法を簡略に示したものである。

【発明を実施するための形態】

【0044】

後述する本発明に関する詳細な説明は、本発明の各目的、各技術的解法、及び各長所を明らかにするために本発明が実施され得る特定の実施例を例示として示す添付図面を参照する。これらの実施例は、当業者が本発明を実施することができるように十分詳細に説明される。

30

【0045】

また、本発明の詳細な説明及び各請求項にわたって、「含む」という単語及びそれらの変形は、他の技術的各特徴各付加物、構成要素又は段階を除外することを意図したものではない。通常の技術者にとって本発明の他の各目的、長所及び各特性が、一部は本説明書から、また一部は本発明の実施から明らかになるであろう。以下の例示及び図面は実例として提供され、本発明を限定することを意図したものではない。

【0046】

さらに、本発明は、本明細書に示された実施例のすべての可能な組合せを網羅する。本発明の多様な実施例は互いに異なるが、相互に排他的である必要はないことが理解されるべきである。例えば、ここに記載されている特定の形状、構造及び特性は、一実施例に関連して本発明の精神及び範囲を逸脱せず、かつ他の実施例で具現され得る。また、それぞれの開示された実施例内の個別の構成要素の位置又は配置は、本発明の精神及び範囲を逸脱せず、かつ変更され得ることが理解されるべきである。したがって、後述の詳細な説明は、限定的な意味として受け取ろうとするものではなく、本発明の範囲は適切に説明されるのであれば、その請求項が主張することと均等な全ての範囲とともに添付された請求項によってのみ限定される。図面において類似の参照符号は、様々な側面にわたって同一であるか、又は類似の機能を指す。

40

【0047】

50

本発明で言及している各種イメージは、舗装または非舗装道路関連のイメージを含み得、この場合、道路環境で登場し得る物体（例えば、自動車、人、動物、植物、物、建物、飛行機やドローンのような飛行体、その他の障害物）を想定し得るが、必ずしもこれに限定されるものではなく、本発明で言及している各種イメージは、道路と関係のないイメージ（例えば、非舗装道路、路地、空き地、海、湖、川、山、森、砂漠、空、室内と関連したイメージ）でもあり得、この場合、非舗装道路、路地、空き地、海、湖、川、山、森、砂漠、空、室内環境で登場し得る物体（例えば、自動車、人、動物、植物、物、建物、飛行機やドローンのような飛行体、その他の障害物）を想定し得るが、必ずしもこれに限定されるものではない。

【0048】

以下、本発明の属する技術分野において通常の知識を有する者が本発明を容易に実施し得るようにするために、本発明の好ましい実施例について、添付された図面を参照して詳細に説明することにする。

【0049】

図1は、本発明の一実施例によるイメージ認識のためのCNNの学習装置100を簡略に示したものであり、図1を参照すると、学習装置100は通信部110とプロセッサ120とを含むことができる。

【0050】

先ず、通信部110は、少なくとも一つのトレーニングイメージを受信することができる。

【0051】

この際、トレーニングイメージはデータベース300に格納されていてもよく、データベース300にはトレーニングイメージに対応して少なくとも一つの物体に対するクラス情報の少なくとも一つの原本正解（Ground Truth）及びそれぞれの物体に対する位置情報の少なくとも一つの原本正解が格納されていてもよい。

【0052】

また、学習装置は、次のプロセスを遂行するためのコンピュータ読取り可能なインストラクション（Instruction）を格納することができるメモリ115をさらに含むことができる。一実施例によると、プロセッサ、メモリ、媒体等は統合プロセッサ（Integrated Processor）として統合され得る。

【0053】

次に、プロセッサ120は、(i)少なくとも一つのコンボリューションレイヤをもって、トレーニングイメージ上でROIの領域をクロップ（Crop）し、リサイズ（Resize）して生成されたROIイメージそれぞれに対してコンボリューション演算を少なくとも一回適用させて物体に対応する少なくとも一つのROIごとにROI特徴マップを生成させ、(ii)(ii-1)プーリングレイヤをもって、ROI特徴マップに対してプーリング演算を少なくとも一回適用させてROIごとにプーリング済み特徴マップを生成させ、第1トランスポーズレイヤ（Transposing Layer）をもって、ROIごとのプーリング済み特徴マップ上の対応する同一の各位置のピクセルそれぞれをROIごとにコンカチネート（Concatenating）させて統合特徴マップ（Integrated Feature Map）を生成させるか、(ii-2)プーリングレイヤをもって、ROI特徴マップに対してプーリング演算を適用させてROIごとのプーリング済み特徴マップを生成させ、プーリングレイヤをもって、ROIごとのプーリング済み特徴マップ上の対応する同一の各位置のピクセルそれぞれをROIごとにコンカチネートさせて統合特徴マップを生成させるプロセスを遂行することができる。そして、プロセッサ120は、(i)第1リシェイプレイヤ（Reshaping Layer）をもって、統合特徴マップのすべてのチャンネルのうちでそれぞれに対応するH1個のチャンネルから構成されたそれぞれのグループ内の特徴それぞれをコンカチネートさせて第1リシェイプ済み特徴マップ（Reshaped Feature Map）を生成させ、(ii)1×H1コンボリューションレイヤをもって、第1リシェイプ済み特徴マッ

10

20

30

40

50

プに対して $1 \times H1$ コンボリューション演算を適用させてボリューム (Volume) が調整された第1調整特徴マップ (Adjusted Feature Map) を生成させ、(iii) 第2リシェイプレイヤをもって、第1調整特徴マップのすべてのチャンネルのうちでそれぞれに対応する $H2$ 個のチャンネルから構成されたそれぞれのグループ内の特徴それぞれをコンカチネートさせて第2リシェイプ済み特徴マップを生成させ、(iv)  $1 \times H2$ コンボリューションレイヤをもって、第2リシェイプ済み特徴マップに対して $1 \times H2$ コンボリューション演算を適用させてボリュームが調整された第2調整特徴マップを生成させるプロセスを遂行することができる。以後、プロセッサ120は、(i) 第2トランスポーズレイヤをもって、第2調整特徴マップを前記ピクセル別に分離してROIごとにピクセル別特徴マップ (Pixel-Wise Feature Map) を生成させ、分類レイヤをもって、ROIごとのピクセル別特徴マップを利用してROIそれぞれに対する物体情報を生成させるか、(ii) 分類レイヤをもって、第2調整特徴マップをピクセル別に分離してROIごとにピクセル別特徴マップを生成させ、分類レイヤをもって、ROIごとのピクセル別特徴マップを利用してROIそれぞれに対する物体情報を生成させるプロセスを遂行することができる。そして、プロセッサ120は、ロスレイヤをもって、物体情報とこれに対応する原本正解とを参照して少なくとも一つの物体ロスを算出させることにより、物体ロスをバックプロパゲーションして $1 \times H2$ コンボリューションレイヤ、 $1 \times H1$ コンボリューションレイヤ、及びコンボリューションレイヤのうち少なくとも一部のパラメータを学習するプロセスを遂行することができる。

【0054】

この際、本発明の一実施例による学習装置100は、コンピューティング装置であって、プロセッサを搭載して演算能力を備えた装置であればいくらかでも本発明による学習装置100として採択され得る。また、図1では一つの学習装置100のみを示したが、これに限定されず、学習装置は複数個に分けて役割を遂行することもできる。

【0055】

このように構成された本発明の一実施例による学習装置100を利用してイメージ認識のためのCNNを学習する方法を、図2を参照して説明すると以下のとおりである。

【0056】

まず、トレーニングイメージが入力されると、学習装置100が、コンボリューションレイヤ121をもって、トレーニングイメージ上で物体に対応するROIの領域をクロップし、リサイズして生成されたROIイメージそれぞれに対してコンボリューション演算を適用して、少なくとも一つの物体に対応する少なくとも一つのROIごとにROI特徴マップを生成させることができる。この際、コンボリューションレイヤ121は、ROIイメージ上に順次にコンボリューション演算を適用するために単一のコンボリューションレイヤ又は複数個のコンボリューションレイヤであり得る。

【0057】

次に、学習装置100はプーリングレイヤ122をもって、ROIごとにROI特徴マップに対してプーリング演算を適用させてROIごとのプーリング済み特徴マップ $P1$ 、 $P2$ 、...、 $PN$ を生成させることができる。

【0058】

一例として、図3を参照すると、ROIの個数を $N$ とし、ROIごとのプーリング済み特徴マップ $P1$ 、 $P2$ 、...、 $PN$ の幅を $M1$ 、ROIごとのプーリング済み特徴マップ $P1$ 、 $P2$ 、...、 $PN$ の高さを $M2$ とし、ROIごとのプーリング済み特徴マップ $P1$ 、 $P2$ 、...、 $PN$ のチャンネル個数を $J$ とする場合、ROIごとのプーリング済み特徴マップ $P1$ 、 $P2$ 、...、 $PN$ はそれぞれ $M1 \times M2 \times J$ のボリュームを有することができる。

【0059】

次に、学習装置100は、第1トランスポーズレイヤ123をもって、プーリングレイヤ122から出力されたROIごとのプーリング済み特徴マップ $P1$ 、 $P2$ 、...、 $PN$ 上の対応する同一の各位置のそれぞれのピクセルをROIごとにコンカチネートさせて統合特徴マップ400を生成させることができる。

## 【 0 0 6 0 】

一例として、図 3 と図 4 を参照すると、図 3 における R O I ごとのそれぞれのプーリング済み特徴マップ P 1、P 2、...、P N 上における最も左側上端のピクセルをそれぞれ R 1 F 1、R 2 F 1、...、R N F 1 とする場合、R O I ごとのプーリング済み特徴マップ P 1、P 2、...、P N それぞれにおける最も左側上端のピクセルである R 1 F 1、R 2 F 1、...、R N F 1 をコンカチネートして、図 4 においてすべてのチャンネルのうち最初のチャンネルに対応する最も前面にあるラインのように幅 N、高さ 1 である統合特徴マップ 4 0 0 の第 1 部分が生成されるようにし、この過程を R O I ごとのプーリング済み特徴マップ P 1、P 2、...、P N それぞれのピクセルの残りの部分にも繰り返し実行する。そして、R O I ごとのプーリング済み特徴マップ P 1、P 2、...、P N それぞれにおけるピクセル個数と同一にチャンネル個数が M 1 × M 2 × J である統合特徴マップ 4 0 0 を生成することができる。

10

## 【 0 0 6 1 】

一方、前記では、学習装置 1 0 0 が第 1 トランスポーズレイヤ 1 2 3 をもって、プーリングレイヤ 1 2 2 から出力された R O I ごとのプーリング済み特徴マップ P 1、P 2、...、P N を利用して統合特徴マップ 4 0 0 を生成させた。その他の例として、第 1 トランスポーズレイヤ 1 2 3 を使用せずにプーリングレイヤ 1 2 2 をもって統合特徴マップ 4 0 0 を生成させることができる。すなわち、その他の例として、学習装置 1 0 0 はプーリングレイヤ 1 2 2 をもって、R O I ごとの R O I 特徴マップ上で R O I それぞれに対応する領域それぞれに対してプーリング演算を適用させて R O I ごとのプーリング済み特徴マップ P 1、P 2、...、P N を生成させ、R O I ごとのプーリング済み特徴マップ P 1、P 2、...、P N 上の対応する同一の各位置のそれぞれのピクセルを R O I ごとにコンカチネートさせて統合特徴マップ 4 0 0 を生成させることができる。

20

## 【 0 0 6 2 】

次に、学習装置 1 0 0 は第 1 リシェイプレイヤ 1 2 4 をもって、統合特徴マップ 4 0 0 のすべてのチャンネルのうち、それぞれの対応する H 1 個のチャンネルから構成されたそれぞれのグループ内のそれぞれの特徴をコンカチネートさせて第 1 リシェイプ済み特徴マップ 5 0 0 A を生成させることができる。

## 【 0 0 6 3 】

一例として、図 4 と図 5 a を参照すると、学習装置 1 0 0 は、第 1 リシェイプレイヤ 1 2 4 をもって、統合特徴マップ 4 0 0 の第 1 チャンネルそれぞれの特徴 R 1 C H 1、R 2 C H 1、...、R N C H 1 ないし第 H 1 チャンネルのそれぞれの特徴 R 1 C H ( H 1 )、R 2 C H ( H 1 )、...、R N C H ( H 1 ) をコンカチネートさせ、統合特徴マップ 4 0 0 の第 H 1 + 1 チャンネルのそれぞれの特徴 R 1 C H ( H 1 + 1 )、R 2 C H ( H 1 + 1 )、...、R N C H ( H 1 + 1 ) ないし第 2 × H 1 チャンネルのそれぞれの特徴 R 1 C H ( 2 × H 1 )、R 2 C H ( 2 × H 1 )、...、R N C H ( 2 × H 1 ) をコンカチネートさせることができる。つまり、学習装置 1 0 0 は、この過程を繰り返すことにより、幅 N、高さ H 1、チャンネル

30

$$\text{CEIL} \left( \frac{M1 \times M2 \times J}{H1} \right)$$

40

である第 1 リシェイプ済み特徴マップ 5 0 0 A を生成させることができる。

## 【 0 0 6 4 】

この際、統合特徴マップ 4 0 0 のチャンネル個数が H 1 の倍数でない場合、学習装置 1 0 0 は第 1 リシェイプレイヤ 1 2 4 をもって、統合特徴マップ 4 0 0 に少なくとも一つの第 1 ダミーチャンネルを追加させ、少なくとも一つの第 1 ダミーチャンネルを含む統合特徴マップ 4 0 0 のチャンネル個数が H 1 の倍数になるようにした後、統合特徴マップ 4 0 0 のすべてのチャンネルのうちそれぞれの対応する H 1 個のチャンネルから構成されたそれぞれのグループ内のそれぞれの特徴をコンカチネートさせることができる。つまり、統

50

合特徴マップ400のチャンネル個数がH1の倍数にならないことにより、

$$\text{CEIL} \left( \frac{M1 \times M2 \times J}{H1} \right)$$

番目のチャンネルをコンカチネートして生成された統合特徴マップ400のサイズが幅N、高さH1にならない場合、統合特徴マップ400の幅はN、高さはH1になるように少なくとも一つのゼロパディング領域を追加することができる。

【0065】

次に、学習装置100は1×H1コンボリューションレイヤ125をもって、第1リシェイプ済み特徴マップ500Aに対して1×H1コンボリューション演算を適用させてボリュームが調整された第1調整特徴マップ500Bを生成させることができる。

10

【0066】

一例として、図5aと図5bを参照すると、1×H1コンボリューションレイヤ125のフィルタの個数をKとする場合、学習装置100は1×H1コンボリューションレイヤ125をもって、図5aの第1リシェイプ済み特徴マップ500Aに対して1×H1コンボリューション演算を適用させて幅N、高さ1、チャンネルKのN×1×Kのボリュームを有する第1調整特徴マップ500Bを生成させることができる。これを通じてコンピューティング演算量を1/H1に減少させることができるようになり、それに伴ってコンピューティング演算速度及び演算能力を増大させることができるようになる。

20

【0067】

この際、第1調整特徴マップ500B上の特徴R1'CH1、R2'CH1、...、RN'CH1は、それぞれ第1リシェイプ済み特徴マップ500A上の特徴R1CH1、R1CH2、...、R1CH(H1)を局地的に統合(Locally Connecting)し、第1リシェイプ済み特徴マップ500A上の特徴R2CH1、R2CH2、...、R2CH(H1)を局地的に統合し、第1リシェイプ済み特徴マップ500A上の特徴RNCH1、RNCH2、...、RNCH(H1)を局地的に統合した結果であり得る。参考までに、前記局地的統合(Locally Connecting)は、1×H1コンボリューション演算を適用することを意味する。

30

【0068】

次に、学習装置100は、第2リシェイプレイヤ126をもって、第1調整特徴マップ500Bのすべてのチャンネルのうちそれぞれの対応するH2個のチャンネルから構成されたそれぞれのグループ内のそれぞれの特徴をコンカチネートして第2リシェイプ済み特徴マップ600Aを生成させることができる。

【0069】

一例として、図5bと図6aを参照すると、学習装置100は、第2リシェイプレイヤ126をもって、第1調整特徴マップ500Bで第1チャンネルのそれぞれの特徴R1'CH1、R2'CH1、...、RN'CH1ないし第H1チャンネル内のそれぞれの特徴R1'CH(H1)、R2'CH(H1)、...、RN'CH(H1)をコンカチネートさせ、第1調整特徴マップ500Bの第H1+1チャンネルのそれぞれの特徴R1'CH(H1+1)、R2'CH(H1+1)、...、RN'CH(H1+1)ないし第2×H1チャンネルのそれぞれの特徴R1'CH(2×H1)、R2'CH(2×H1)、...、RN'CH(2×H1)をコンカチネートさせることができる。つまり、学習装置100は、このコンカチネートの過程を繰り返すことにより、幅N、高さH2、チャンネル

40

$$\text{CEIL} \left( \frac{K}{H2} \right)$$

である第2リシェイプ済み特徴マップ600Aを生成させることができる。

【0070】

50

この際、第1調整特徴マップ500Bのチャンネル個数がH2の倍数でない場合、学習装置100は第2リシェイプレイヤ126をもって、第1調整特徴マップ500Bに少なくとも一つの第2ダミーチャンネルを追加させ、少なくとも一つの第2ダミーチャンネルを含む第1調整特徴マップ500Bのチャンネル個数がH2の倍数になるようにした後、第1調整特徴マップ500Bのすべてのチャンネルのうちそれぞれの対応するH2個のチャンネルから構成されたそれぞれのグループ内のそれぞれの特徴をコンカチネートさせることができる。すなわち、第1調整特徴マップ500Bのチャンネル個数がH2の倍数にならないことにより、

$$\text{CEIL} \left( \frac{K}{H^2} \right)$$

10

番目のチャンネルをコンカチネートして生成された第1調整特徴マップ500Bのサイズが幅はN、高さはH2にならない場合、第1調整特徴マップ500Bの幅はN、高さはH2になるように少なくとも一つのゼロパディング領域を追加することができる。

【0071】

次に、学習装置100は、 $1 \times H^2$ コンボリューションレイヤ127をもって、第2リシェイプ済み特徴マップ600Aに対して $1 \times H^2$ コンボリューション演算を適用させてボリュームが調整された第2調整特徴マップ600Bを生成させることができる。

【0072】

20

一例として、図6aと図6bを参照すると、 $1 \times H^2$ コンボリューションレイヤ127のフィルタの個数をLとする場合、学習装置100は $1 \times H^2$ コンボリューションレイヤ127をもって、図6aの第2リシェイプ済み特徴マップ600Aに対して $1 \times H^2$ コンボリューション演算を適用させて幅N、高さ1、チャンネルLである $N \times 1 \times L$ のボリュームを有する第2調整特徴マップ600Bを生成させることができる。これを通じて、コンピューティング演算量を $1/H^2$ に減少させることができようになり、それに伴ってコンピューティング演算速度及び演算能力を増大させることができるようになる。この場合、第2調整特徴マップ600B上の特徴 $R_1 \text{ " CH } 1$ 、 $R_2 \text{ " CH } 1$ 、...、 $R_N \text{ " CH } 1$ はそれぞれ第2リシェイプ済み特徴マップ600A上の特徴 $R_1 \text{ ' CH } 1$ 、 $R_1 \text{ ' CH } 2$ 、...、 $R_1 \text{ ' CH } (H^2)$ を局地的に統合(Locally Connecting)し、第2リシェイプ済み特徴マップ600A上の特徴 $R_2 \text{ ' CH } 1$ 、 $R_2 \text{ ' CH } 2$ 、...、 $R_2 \text{ ' CH } (H^2)$ を局地的に統合し、第2リシェイプ済み特徴マップ600A上の特徴 $R_N \text{ ' CH } 1$ 、 $R_N \text{ ' CH } 2$ 、...、 $R_N \text{ ' CH } (H^2)$ を局地的に統合した結果であり得る。参考までに、前記局地的統合(Locally Connecting)は、 $1 \times H^2$ コンボリューション演算を適用することを意味する。

30

【0073】

次に、学習装置100は、第2トランスポーズレイヤ128をもって、第2調整特徴マップ600Bをピクセル別に分離してROIごとのピクセル別特徴マップPW1、PW2、...、PWNを生成させることができる。

【0074】

40

一例として、図7を参照すると、学習装置100は第2トランスポーズレイヤ128をもって、図6bの幅N、高さ1、チャンネルLである $N \times 1 \times L$ のボリュームを有する第2調整特徴マップ600Bをピクセル別に分離して幅1、高さ1、チャンネルLであるROIごとのピクセル別特徴マップPW1、PW2、...、PWNを生成させる。

【0075】

次に、学習装置100は、分類レイヤ129をもって、ROIごとのピクセル別特徴マップPW1、PW2、...、PWNを利用してROIそれぞれに対する物体情報を生成させることができる。この際、分類レイヤ129はソフトマックス(Softmax)アルゴリズムを使用することができ、それぞれのROI内の物体それぞれがどのような種類の物体であるのか認知された結果を生成し、認知された結果が正答である可能性に関する確率情

50

報を生成させることができる。これに加えて、物体情報はトレーニングイメージ内のそれぞれの物体位置情報を含むこともできる。

【0076】

一方、前記では、第2トランスポーズレイヤ128が第2調整特徴マップ600Bを利用してROIごとのピクセル別特徴マップPW1、PW2、...、PWNを生成することができる。その他の例として、分類レイヤ129は、第2トランスポーズレイヤ128を使用せずに第2調整特徴マップ600Bを利用してROIごとのピクセル別特徴マップPW1、PW2、...、PWNを生成させることができる。つまり、学習装置100は分類レイヤ129をもって、第2調整特徴マップ600Bをピクセル別に分離してROIごとのピクセル別特徴マップPW1、PW2、...、PWNを生成させ、分類レイヤ129をもって、ROIごとのピクセル別特徴マップを利用してROIそれぞれに対する物体情報を生成させることができる。

10

【0077】

次に、学習装置100は、ロスレイヤ130をもって、物体情報とこれに対応する原本正解とを参照して少なくとも一つの物体ロスを算出させることにより、物体ロスを利用したバックプロパゲーションを通じて1xH2コンボリューションレイヤ127、1xH1コンボリューションレイヤ125、及びコンボリューションレイヤ121のうち少なくとも一部のパラメータを学習することができる。

【0078】

前記のように、本発明の一実施例によるイメージ認識のためのCNNは、コンボリューションレイヤのみで構成されることにより、入力イメージのサイズに関係なしに正常に入力イメージ内の物体を認識することができる。また、本発明の一実施例によるイメージ認識のためのCNNは、1xHコンボリューションの局地的統合の演算を適用することにより、従来のFCレイヤのFC(Fully Connected)演算に比べてコンピューティング演算量を減少させることができるようになる。

20

【0079】

図8は、本発明の一実施例によるイメージ認識のためのCNNのテスト装置200を簡略に示したものであり、図8を参照すると、テスト装置200は通信部210とプロセッサ220とを含むことができる。

【0080】

また、テスト装置200は、次のプロセスを遂行するためのコンピュータ読取り可能なインストラクション(Instruction)を格納することができるメモリ215をさらに含むことができる。一実施例によると、プロセッサ、メモリ、媒体等は統合プロセッサ(Integrated Processor)として統合され得る。

30

【0081】

先ず、通信部210は、少なくとも一つのテストイメージを取得するか、他の装置をもって取得するように支援することができる。

【0082】

この際、本発明の一実施例によるイメージ認識のためのCNNは、図1ないし図7を参照して説明した学習方法によって学習された状態であり得る。

40

【0083】

参考までに、以下の説明で混同を避けるために「学習用」という文句は、先に説明した学習プロセスに関する用語について追加され、「テスト用」という文句はテストプロセスに関する用語について追加される。

【0084】

つまり、少なくとも一つのトレーニングイメージが取得されると、学習装置は、(a)(i)コンボリューションレイヤをもって、トレーニングイメージ上で学習用ROIの領域をクロップ(Crop)し、リサイズ(Resize)して生成された学習用ROIイメージそれぞれに対してコンボリューション演算を適用させて少なくとも一つの学習用物体に対応する少なくとも一つの学習用ROIごとの学習用ROI特徴マップを生成させ、

50

( i i ) ( i i - 1 ) プーリングレイヤをもって、学習用 R O I 特徴マップに対してプーリング演算を適用させて学習用 R O I ごとの学習用プーリング済み特徴マップを生成させ、第 1 トランスポーズレイヤ ( T r a n s p o s i n g L a y e r ) をもって、学習用 R O I ごとの学習用プーリング済み特徴マップ上の対応する同一の各位置のピクセルそれぞれを学習用 R O I ごとにコンカチネート ( C o n c a t e n a t i n g ) させて学習用統合特徴マップ ( I n t e g r a t e d F e a t u r e M a p ) を生成させるか、( i i - 2 ) プーリングレイヤをもって、学習用 R O I 特徴マップに対してプーリング演算を適用させて学習用 R O I ごとの学習用プーリング済み特徴マップを生成させ、プーリングレイヤをもって、学習用 R O I ごとの学習用プーリング済み特徴マップ上の対応する同一の各位置のピクセルそれぞれを学習用 R O I ごとにコンカチネートさせて学習用統合特徴マップを生成させ、( b ) ( b 1 ) ( i ) 第 1 リシェイプレイヤ ( R e s h a p i n g L a y e r ) をもって、学習用統合特徴マップのすべてのチャンネルのうちでそれぞれに対応する H 1 個のチャンネルから構成されたそれぞれのグループ内の特徴それぞれをコンカチネートさせて学習用第 1 リシェイプ済み特徴マップ ( R e s h a p e d F e a t u r e M a p ) を生成させ、( i i ) 1 x H 1 コンボリューションレイヤをもって、学習用第 1 リシェイプ済み特徴マップに対して 1 x H 1 コンボリューション演算を適用させてボリューム ( V o l u m e ) が調整された学習用第 1 調整特徴マップ ( A d j u s t e d F e a t u r e M a p ) を生成させ、( b 2 ) ( i ) 第 2 リシェイプレイヤをもって、学習用第 1 調整特徴マップのすべてのチャンネルのうちでそれぞれに対応する H 2 個のチャンネルから構成されたそれぞれのグループ内の特徴それぞれをコンカチネートさせて学習用第 2 リシェイプ済み特徴マップを生成させ、( i i ) 1 x H 2 コンボリューションレイヤをもって、学習用第 2 リシェイプ済み特徴マップに対して 1 x H 2 コンボリューション演算を適用させてボリュームが調整された学習用第 2 調整特徴マップを生成させ、( c ) ( c 1 ) ( i ) 第 2 トランスポーズレイヤをもって、学習用第 2 調整特徴マップをピクセル別に分離して学習用 R O I ごとの学習用ピクセル別特徴マップ ( P i x e l - W i s e F e a t u r e M a p ) を生成させ、分類レイヤをもって、学習用 R O I ごとの学習用ピクセル別特徴マップを利用して学習用 R O I それぞれに対する学習用物体情報を生成させるか、( i i ) 分類レイヤをもって、学習用第 2 調整特徴マップをピクセル別に分離して学習用 R O I ごとに学習用ピクセル別特徴マップを生成させ、分類レイヤをもって、学習用 R O I ごとの学習用ピクセル別特徴マップを利用して学習用 R O I それぞれに対する学習用物体情報を生成させた後、( c 2 ) ロスレイヤをもって、学習用物体情報とこれに対応する原本正解とを参照して少なくとも一つの物体ロスを算出させることにより、物体ロスをバックプロパゲーションして 1 x H 2 コンボリューションレイヤ、1 x H 1 コンボリューションレイヤ、及びコンボリューションレイヤのうち少なくとも一部のパラメータを学習するプロセスを遂行した状態であり得る。次に、プロセッサ 2 2 0 は、( i ) コンボリューションレイヤをもって、テストイメージ上でテスト用 R O I の領域をクロップ ( C r o p ) し、リサイズ ( R e s i z e ) して生成されたテスト用 R O I イメージそれぞれに対してコンボリューション演算を少なくとも一回適用させて少なくとも一つのテスト用物体に対応する少なくとも一つのテスト用 R O I ごとにテスト用 R O I 特徴マップを生成させ、( i i ) ( i i - 1 ) プーリングレイヤをもって、テスト用 R O I 特徴マップに対してプーリング演算を少なくとも一回適用させてテスト用 R O I ごとにテスト用プーリング済み特徴マップを生成させ、第 1 トランスポーズレイヤ ( T r a n s p o s i n g L a y e r ) をもって、テスト用 R O I ごとのテスト用プーリング済み特徴マップ上の対応する同一の各位置のピクセルそれぞれをテスト用 R O I ごとにコンカチネート ( C o n c a t e n a t i n g ) させてテスト用統合特徴マップ ( I n t e g r a t e d F e a t u r e M a p ) を生成させるか、( i i - 2 ) プーリングレイヤをもって、テスト用 R O I 特徴マップに対してプーリング演算を適用させてテスト用 R O I ごとのテスト用プーリング済み特徴マップを生成させ、プーリングレイヤをもって、テスト用 R O I ごとのテスト用プーリング済み特徴マップ上の対応する同一の各位置のピクセルそれぞれをテスト用 R O I ごとにコンカチネートさせてテスト用統合特徴マップを生成させるプ

10

20

30

40

50

ロセスを遂行することができる。そして、プロセッサ(220)は、(i)第1リシェイプレイヤ(Reshaping Layer)をもって、テスト用統合特徴マップのすべてのチャンネルのうちでそれぞれに対応するH1個のチャンネルから構成されたそれぞれのグループ内の特徴それぞれをコンカチネートさせてテスト用第1リシェイプ済み特徴マップ(Reshaped Feature Map)を生成させ、(ii)1×H1コンボリューションレイヤをもって、テスト用第1リシェイプ済み特徴マップに対して1×H1コンボリューション演算を適用させてボリューム(Volume)が調整されたテスト用第1調整特徴マップ(Adjusted Feature Map)を生成させ、(iii)第2リシェイプレイヤをもって、テスト用第1調整特徴マップのすべてのチャンネルのうちでそれぞれに対応するH2個のチャンネルから構成されたそれぞれのグループ内の特徴それぞれをコンカチネートさせてテスト用第2リシェイプ済み特徴マップを生成させ、(vi)1×H2コンボリューションレイヤをもって、テスト用第2リシェイプ済み特徴マップに対して1×H2コンボリューション演算を適用させてボリュームが調整されたテスト用第2調整特徴マップを生成させるプロセスを遂行することができる。以後、プロセッサ220は、(i)第2トランスポーズレイヤをもって、テスト用第2調整特徴マップを前記ピクセル別に分離してテスト用ROIごとにテスト用ピクセル別特徴マップ(Pixel-Wise Feature Map)を生成させ、分類レイヤをもって、テスト用ROIごとのテスト用ピクセル別特徴マップを利用してテスト用ROIそれぞれに関するテスト用物体情報を生成させるか、(ii)分類レイヤをもって、テスト用第2調整特徴マップをピクセル別に分離してテスト用ROIごとにテスト用ピクセル別特徴マップを生成させ、分類レイヤをもって、テスト用ROIごとのテスト用ピクセル別特徴マップを利用してテスト用ROIそれぞれに関するテスト用物体情報を生成するようにプロセスを遂行することができる。

10

20

**【0085】**

この際、本発明の一実施例によるテスト装置200は、コンピューティング装置であって、プロセッサを搭載して演算能力を備えた装置であればいくらかでも本発明によるテスト装置200として採択され得る。また、図8では一つのテスト装置200のみを示したが、これに限定されず、テスト装置は複数個に分けて役割を遂行することもできる。

**【0086】**

このように構成された本発明の一実施例によるテスト装置200を利用してイメージ認識のためのCNNをテストする方法を、図9を参照して説明すると次のとおりである。以下の説明で、図1ないし図7を参照して説明した学習方法から容易に理解可能な部分については詳細な説明を省略することとする。

30

**【0087】**

まず、図1ないし図7を参照して説明した学習方法によってコンボリューションレイヤ221、1×H1コンボリューションレイヤ225、及び1×H2コンボリューションレイヤ227が学習された状態でテストイメージが入力されると、テスト装置200がコンボリューションレイヤ221をもって、テストイメージ上でテスト用物体に対応するテスト用ROIの領域をクロップしてリサイズしたテスト用ROIイメージそれぞれに対してコンボリューション演算を適用してテスト用ROIごとのテスト用ピクセル別特徴マップを生成させる。この際、コンボリューションレイヤ221は、テスト用ROIイメージに対して順次にコンボリューション演算のための単一コンボリューションレイヤ又は複数個のコンボリューションレイヤであり得る。

40

**【0088】**

次に、テスト装置200は、プーリングレイヤ222をもって、テスト用ROIごとのテスト用ピクセル別特徴マップに対してプーリング演算を適用してテスト用ROIごとのテスト用プーリング済み特徴マップを生成させることができる。

**【0089】**

次に、テスト装置200は、第1トランスポーズレイヤ223をもって、プーリングレイヤ222から出力されたテスト用ROIごとのテスト用プーリング済み特徴マップ上の

50

対応する同一の各位置のそれぞれのピクセルをテスト用ROIごとにコンカチネートさせてテスト用統合特徴マップを生成させることができる。

【0090】

一方、前記では、テスト装置200が第1トランスポートレイヤ223をもって、プーリングレイヤ222から出力されたテスト用ROIごとのテスト用プーリング済み特徴マップを利用してテスト用統合特徴マップを生成させることができる。その他の例として、テスト装置200は、第1トランスポートレイヤ223を使用せずに、プーリングレイヤ222をもってテスト用統合特徴マップを生成させることができる。つまり、テスト装置200は、プーリングレイヤ222をもって、テスト用ROIごとのテスト用ROI特徴マップに対してプーリング演算を適用させてテスト用ROIごとのテスト用プーリング済み特徴マップを生成させ、テスト用ROIごとのテスト用プーリング済み特徴マップ上の対応する同一の各位置のそれぞれのピクセルをテスト用ROIごとにコンカチネートしてテスト用統合特徴マップを生成させることができる。

10

【0091】

次に、テスト装置200は、第1リシェイプレイヤ224をもって、テスト用統合特徴マップのすべてのチャンネルのうちでそれぞれの対応するH1個のチャンネルから構成されたそれぞれのグループ内のそれぞれの特徴をコンカチネートさせてテスト用第1リシェイプ済み特徴マップを生成させることができる。

【0092】

この際、テスト用統合特徴マップのチャンネル個数がH1の倍数でない場合、テスト装置200は第1リシェイプレイヤ224をもって、テスト用統合特徴マップに少なくとも一つの第1ダミーチャンネルを追加して少なくとも一つの第1ダミーチャンネルを含むテスト用統合特徴マップのチャンネル個数がH1の倍数になるようにした後、テスト用統合特徴マップのすべてのチャンネルのうちでそれぞれの対応するH1個のチャンネルから構成されたそれぞれのグループ内のそれぞれの特徴をコンカチネートさせることができる。つまり、テスト用統合特徴マップのチャンネル個数がH1の倍数にならないことにより、

20

$$\text{CEIL} \left( \frac{M1 \times M2 \times J}{H1} \right)$$

30

番目のチャンネルをコンカチネートして生成されたテスト用統合特徴マップのサイズが幅N、高さH1にならない場合、テスト用統合特徴マップの幅がN、高さがH1になるように少なくとも一つのゼロパディング領域を追加することができる。

【0093】

次に、テスト装置200は、1×H1コンボリューションレイヤ225をもって、テスト用第1リシェイプ済み特徴マップに対して1×H1コンボリューション演算を適用させてボリュームが調整されたテスト用第1調整特徴マップを生成させることができる。

【0094】

次に、テスト装置200は第2リシェイプレイヤ226をもって、テスト用第1調整特徴マップのすべてのチャンネルのうちでそれぞれの対応するH2個のチャンネルから構成されたそれぞれのグループ内のそれぞれの特徴をコンカチネートさせてテスト用第2リシェイプ済み特徴マップを生成させることができる。

40

【0095】

この際、テスト用第1調整特徴マップのチャンネル個数がH2の倍数でない場合、テスト装置200は第2リシェイプレイヤ226をもって、テスト用第1調整特徴マップに少なくとも一つの第2ダミーチャンネルを追加させて少なくとも一つの第2ダミーチャンネルを含むテスト用第1調整特徴マップのチャンネル個数がH2の倍数になるようにした後、テスト用第1調整特徴マップのすべてのチャンネルのうちでそれぞれの対応するH2個のチャンネルから構成されたそれぞれのグループ内のそれぞれの特徴をコンカチネートさせることができる。つまり、テスト用第1調整特徴マップのチャンネル個数がH2の倍数

50

にならないことにより、

$$\text{CEIL} \left( \frac{K}{H^2} \right)$$

番目のチャンネルをコンカチネートして生成されたテスト用第1調整特徴マップのサイズが幅N、高さH2にならない場合、テスト用第1調整特徴マップの幅がN、高さがH2になるように少なくとも一つのゼロパディング領域を追加することができる。

【0096】

次に、テスト装置200は、1×H2コンボリユーションレイヤ227をもって、テスト用第2リシェイプ特徴マップに対して1×H2コンボリユーション演算を適用させてボリユームが調整されたテスト用第2調整特徴マップを生成させることができる。

10

【0097】

次に、テスト装置200は第2トランスポーズレイヤ228をもって、テスト用第2調整特徴マップをピクセル別に分離させてテスト用ROIごとのテスト用ピクセル別特徴マップを生成させることができる。

【0098】

次に、テスト装置200は分類レイヤ229をもって、テスト用ROIごとのテスト用ピクセル別特徴マップを利用してテスト用ROIに関するテスト用物体情報を生成させることができる。この際、分類レイヤ229は、ソフトマックスアルゴリズムを使用することができ、それぞれのテスト用ROI内のテスト用物体がどのような種類の物体であるのか認知した結果を生成することができ、認知した結果が正答である可能性に関する確率情報を生成させることができる。これに加えて、テスト用物体情報は、テストイメージ内にテスト用物体それぞれの位置情報を含むこともできる。

20

【0099】

一方、前記では、第2トランスポーズレイヤ228は、テスト用第2調整特徴マップを利用してテスト用ROIごとのテスト用ピクセル別特徴マップを生成させることができる。その他の例として、分類レイヤ229は、第2トランスポーズレイヤ228を使用せずに、テスト用第2調整特徴マップを利用してテスト用ROIごとのテスト用ピクセル別特徴マップを生成させることができる。つまり、テスト装置200は分類レイヤ229をもって、テスト用第2調整特徴マップをピクセル別に分離させてテスト用ROIごとのテスト用ピクセル別特徴マップを生成し、分類レイヤ229をもって、テスト用ROIごとのテスト用ピクセル別特徴マップを利用してテスト用ROIそれぞれに関するテスト用物体情報を生成させることができる。

30

【0100】

前記学習方法とテスト方法は、KPI(Key Performance Index、重要業績評価指標)を満たすハードウェアの最適化に使用されるために提供され得る。

【0101】

また、以上にて説明された本発明による実施例は、多様なコンピュータの構成要素を通じて遂行することができるプログラム命令語の形態で具現されて、コンピュータ読取り可能な記録媒体に格納され得る。前記コンピュータ読取り可能な記録媒体はプログラム命令語、データファイル、データ構造などを単独で又は組み合わせて含むことができる。前記コンピュータ読取り可能な記録媒体に格納されるプログラム命令語は、本発明のために特別に設計され、構成されたものであるか、コンピュータソフトウェア分野の当業者に公知にされて使用可能なものであり得る。コンピュータ読取り可能な記録媒体の例には、ハードディスク、フロッピーディスク及び磁気テープのような磁気媒体、CD-ROM、DVDのような光記録媒体、フロプティカル・ディスク(Floptical Disk)のような磁気-光メディア(Magneto-Optical Media)、及びROM、RAM、フラッシュメモリなどのようなプログラム命令語を格納して遂行するように特別に構成されたハードウェア装置が含まれる。プログラム命令語の例には、コンパイラによって作られるもののような機械語コードだけでなく、インタープリターなどを使用して

40

50

コンピュータによって実行される高級言語コードも含まれる。前記ハードウェア装置は、本発明による処理を実行するために一つ以上のソフトウェアモジュールとして作動するように構成され得、その反対も同様である。

【0102】

以上にて本発明が具体的な構成要素などのような特定事項と限定された実施例及び図面によって説明されたが、これは本発明のより一般的な理解の一助とするために提供されたものであるに過ぎず、本発明が前記実施例に限られるものではなく、本発明が属する技術分野において通常の知識を有する者であれば、かかる記載から多様な修正及び変形が行われ得る。

【0103】

従って、本発明の思想は、前記説明された実施例に局限されて定められてはならず、後述する特許請求の範囲だけでなく、本特許請求の範囲と均等または等価的に変形されたものすべては、本発明の思想の範囲に属するといえる。

【符号の説明】

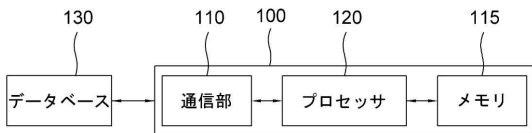
【0104】

- 100：学習装置
- 200：テスト装置
- 110、210：通信部
- 120、220：プロセッサ
- 300：データベース

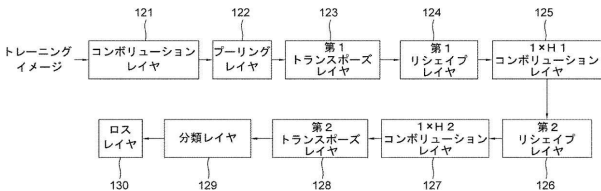
10

20

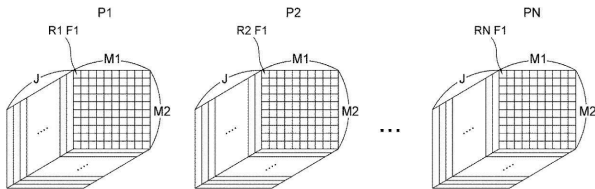
【図1】



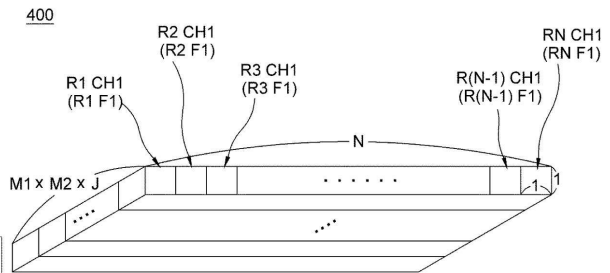
【図2】



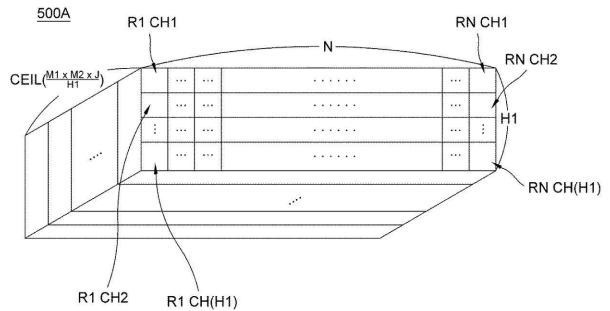
【図3】



【図4】

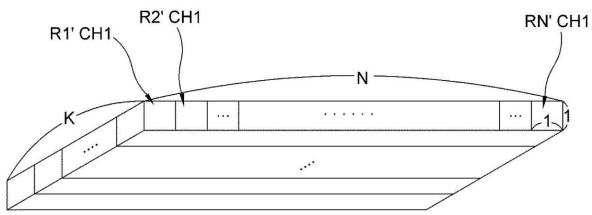


【図5a】



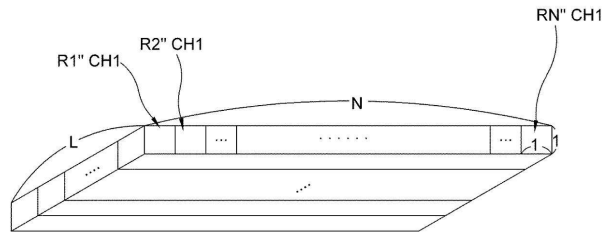
【図 5 b】

500B



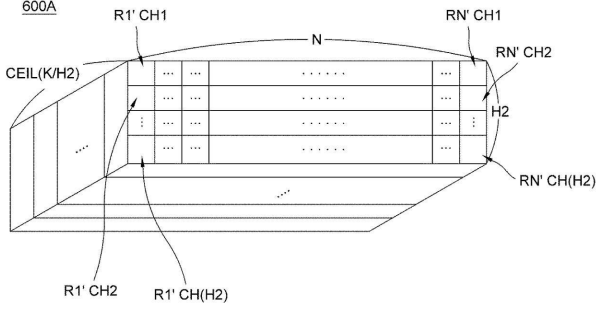
【図 6 b】

600B

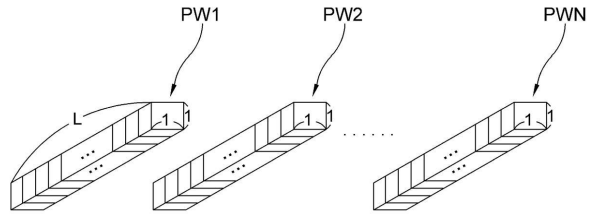


【図 6 a】

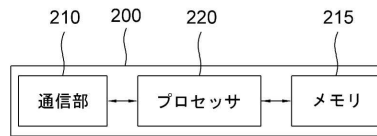
600A



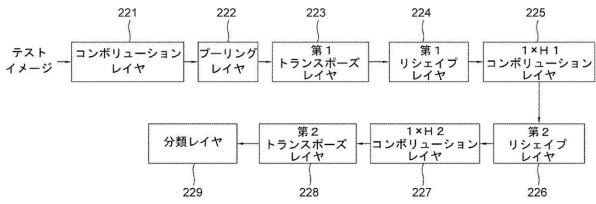
【図 7】



【図 8】



【図 9】



## フロントページの続き

- (72)発明者 キム、ヨンジュン  
大韓民国、キョンサンブク - ド、ポハン - シ、ナム - グ、チゴク - 口、394、  
フィフス ベンチャー - ドン、スウィート 304-308
- (72)発明者 キム、インスー  
大韓民国、キョンサンブク - ド、ポハン - シ、ナム - グ、チゴク - 口、394、  
フィフス ベンチャー - ドン、スウィート 304-308
- (72)発明者 キム、ハク - キョン  
大韓民国、キョンサンブク - ド、ポハン - シ、ナム - グ、チゴク - 口、394、  
フィフス ベンチャー - ドン、スウィート 304-308
- (72)発明者 ナム、ウヒョン  
大韓民国、キョンサンブク - ド、ポハン - シ、ナム - グ、ヨーニル - ユー、ユガ  
ン - ギル 9ベオン - ギル、57、ナンバー 203-803
- (72)発明者 ブー、ソッフン  
大韓民国、キョンギ - ド、アンヤン - シ、マナン - グ、マナン - 口 55ベオン  
- ギル、20-9、ナンバー ビー02
- (72)発明者 ソン、ミュンチュル  
大韓民国、キョンサンブク - ド、ポハン - シ、ブク - グ、チャンニャン - 口 1  
74ベオン - ギル、13
- (72)発明者 ヨー、ドンフン  
大韓民国、キョンサンブク - ド、ポハン - シ、ナム - グ、チゴク - 口、394、  
フィフス ベンチャー - ドン、スウィート 304-308
- (72)発明者 リュー、ウジュ  
大韓民国、キョンサンブク - ド、ポハン - シ、ナム - グ、チゴク - 口、394、  
フィフス ベンチャー - ドン、スウィート 304-308
- (72)発明者 チャン、テウン  
大韓民国、ソウル、カンナム - グ、オンジュ - 口 113-ギル、18-5
- (72)発明者 ジョン、キュンチョン  
大韓民国、キョンサンブク - ド、ポハン - シ、ナム - グ、チゴク - 口、294、  
ヒュンダイグリーン アpartment、ナンバー 232-501
- (72)発明者 チェ、ホンモ  
大韓民国、キョンサンブク - ド、ポハン - シ、ナム - グ、チゴク - 口、20、ス  
ンニ アpartment、ナンバー 5-1805
- (72)発明者 チョウ、ホジン  
大韓民国、キョンサンブク - ド、ポハン - シ、ナム - グ、チゴク - 口、394、  
フィフス ベンチャー - ドン、スウィート 304-308

審査官 岡本 俊威

(56)参考文献 特表2018-503152(JP,A)

(58)調査した分野(Int.Cl., DB名)

G06T 7/00 - 7/90  
G06N 3/04