

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第6869559号
(P6869559)

(45) 発行日 令和3年5月12日(2021.5.12)

(24) 登録日 令和3年4月16日(2021.4.16)

(51) Int.Cl. F 1
G 0 6 T 7 / 0 0 (2017.01) G 0 6 T 7 / 0 0 3 5 0 C

請求項の数 28 (全 29 頁)

<p>(21) 出願番号 特願2019-160143 (P2019-160143) (22) 出願日 令和1年9月3日(2019.9.3) (65) 公開番号 特開2020-38663 (P2020-38663A) (43) 公開日 令和2年3月12日(2020.3.12) 審査請求日 令和1年9月3日(2019.9.3) (31) 優先権主張番号 16/121, 681 (32) 優先日 平成30年9月5日(2018.9.5) (33) 優先権主張国・地域又は機関 米国 (US)</p>	<p>(73) 特許権者 517038176 株式会社ストラドビジョン 大韓民国 37668 キョンサンブク ド ポハン-シ ナム-グ ジゴク-ロ 394 フィフス ベンチャー-ドン ス イート 304-308 (74) 代理人 110000039 特許業務法人アイ・ピー・ウィン (74) 復代理人 100120628 弁理士 岩田 慎一 (72) 発明者 金桂賢 大韓民国ソウル特別市西大門区西小門路2 7、1004</p>
---	--

最終頁に続く

(54) 【発明の名称】 障害物を検出する学習方法及び学習装置そしてこれを利用したテスト方法及びテスト装置 {LEARNING METHOD, LEARNING DEVICE FOR DETECTIN

(57) 【特許請求の範囲】

【請求項1】

少なくとも一つのトレーニングイメージから少なくとも一つの障害物を検出するためのCNNのパラメータを学習する方法において、

(a) 学習装置が、前記トレーニングイメージを獲得し、第1コンボリューションレイヤないし第n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から逐次的に第1エンコード済み特徴マップないし第nエンコード済み特徴マップを各々生成するようにする段階；

(b) 前記学習装置が、第nデコンボリューションレイヤないし第1デコンボリューションレイヤをもって、前記第nエンコード済み特徴マップから逐次的に第nデコード済み特徴マップないし第1デコード済み特徴マップを生成するようにする段階；

(c) 前記少なくとも一つのデコード済み特徴マップの行方向を第1方向とし、その列方向を第2方向として、前記少なくとも一つのデコード済み特徴マップを前記第1方向に対し第1間隔ずつ区画し、前記第2方向に対し第2間隔ずつ区画することで、複数の行と複数の列を有するグリッドの各セルが生成される際、前記学習装置が、前記各々の列ごとに前記各々の行の特徴各々をチャンネル方向へコンカチネートして、少なくとも一つの修正済み特徴マップを生成する段階；

(d) 前記学習装置が前記修正済み特徴マップとそれに対応する原本正解(Ground Truth) イメージとを参照して少なくとも一つのロスを算出するものの、前記原本正解イメージは、前記各々の列をそれに対応する一番下のセルから前記第2方向に沿っ

10

20

て確認した際、近接障害物各々が実際に位置する原本正解ポジション各々に対応する前記各々の行を示すイメージであることを特徴とする段階；及び

(e) 前記学習装置が、前記ロスをバックプロパゲーションして前記CNNの前記パラメータを学習する段階；
を含むことを特徴とする方法。

【請求項2】

前記(b)段階で、

前記学習装置は、前記第nコンボリューションレイヤに対応する前記第nデコンボリューションレイヤをもって前記第nエンコード済み特徴マップを前記第1方向へサイズを所定倍数に拡大し、前記第nデコード済み特徴マップを生成するようにすることを特徴とする請求項1に記載の方法。

10

【請求項3】

前記(c)段階は、

前記学習装置が、前記各々の列をそれに対応する一番下のセルから始まり前記第2方向に沿って確認することで、前記近接障害物各々が位置すると判断される、前記少なくとも一つのデコード済み特徴マップ内の推定位置各々を示すセグメンテーション結果を生成する段階；

をさらに含み、

前記(d)段階で、

前記セグメンテーション結果と前記原本正解イメージとを参照して前記ロスが算出されることを特徴とする請求項1に記載の方法。

20

【請求項4】

前記(c)段階で、

前記学習装置が、前記修正済み特徴マップの前記各々の列に対してコンカチネートされたチャンネル上で前記障害物各々の一番下の部分各々の前記推定位置各々を確認することにより、前記セグメンテーション結果を生成することを特徴とする請求項3に記載の方法。

【請求項5】

前記各々の列は、前記第1方向に少なくとも一つのピクセルを含み、前記各々の行は前記第2方向に少なくとも一つのピクセルを含むことを特徴とする請求項1に記載の方法。

【請求項6】

前記原本正解イメージ及び前記セグメンテーション結果は、前記トレーニングイメージを前記第2方向へNc個の行に分割して、前記列ごとに前記障害物各々が前記行のうちどの行にあるかを示す情報を含み、前記修正済み特徴マップはNc個のチャンネルを含むことを特徴とする請求項3に記載の方法。

30

【請求項7】

前記修正済み特徴マップは、前記デコード済み特徴マップを区画して生じた前記各々の行をチャンネル方向へ前記列ごとにコンカチネートして生成されることを特徴とする請求項1に記載の方法。

【請求項8】

前記(c)段階は、

(c1) 前記学習装置が、 $C_i \times W_i \times H_i$ のサイズを有する前記デコード済み特徴マップを利用して $C_i \cdot H_i \times W_i \times 1$ のサイズを有する第1修正済み特徴マップを生成するものの、前記 C_i は、前記デコード済み特徴マップのチャンネル数であり、前記 W_i は、前記デコード済み特徴マップの列のサイズであり、前記 H_i は、前記デコード済み特徴マップの行のサイズである段階；

40

(c2) 前記学習装置が、前記 $C_i \cdot H_i \times W_i \times 1$ のサイズを有する前記第1修正済み特徴マップを $(W_I / W_i) \times N_c \times W_i \times 1$ のサイズを有する前記第1修正済み特徴マップへ変更するコンボリューション演算を少なくとも一度遂行するものの、前記 W_I は、前記トレーニングイメージの列のサイズであり、前記 W_i は、前記デコード済み特徴マップの前記列のサイズである段階；

50

(c3) 前記学習装置が、 $((W I / W i) \times N c) \times W i \times 1$ のサイズを有する前記第1修正済み特徴マップを利用して $N c \times W I \times 1$ のサイズを有する第2修正済み特徴マップを生成するものの、前記 $N c$ は前記障害物各々の前記列ごとの位置を特定するために前記トレーニングイメージを分割した行の数である段階；及び

(c4) 前記学習装置が、前記第2修正済み特徴マップの前記列ごとに前記 $N c$ 個のチャンネルをノーマライジング(normalizing)するソフトマックス出力プロセスを遂行する段階；

を含み、

前記(d)段階で、

前記(c4)段階で出力された、前記原本正解イメージに対応する各々のソフトマックス出力と前記原本正解イメージを参照して、前記ロスが算出されることを特徴とする請求項7に記載の方法。

10

【請求項9】

前記第1ないし前記第 n コンボリューションレイヤは、コンボリューション、マックスプーリング(max pooling)、ReLUの中から少なくとも一つの演算を遂行して前記トレーニングイメージまたは任意の特徴マップの大きさを縮小させることで、少なくとも一つの縮小された(reduced)特徴マップを生成し、前記第 n ないし第1デコンボリューションレイヤは、デコンボリューション及びReLUの中から少なくとも一つの演算を遂行して入力された特徴マップの大きさを拡大した、少なくとも一つの拡大された(enlarged)特徴マップを生成することを特徴とする請求項1に記載の方法。

20

【請求項10】

少なくとも一つのテストイメージから少なくとも一つの障害物を検出するためのCNNテスト方法において、

(a) 学習装置が、(i)トレーニングイメージを受信し、第1コンボリューションレイヤないし第 n 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から逐次的に第1学習用エンコード済み特徴マップないし第 n 学習用エンコード済み特徴マップを各々生成するようにするプロセス、(ii)第 n デコンボリューションレイヤないし第1デコンボリューションレイヤをもって、前記第 n 学習用エンコード済み特徴マップから逐次的に第 n 学習用デコード済み特徴マップないし第1学習用デコード済み特徴マップを生成するようにするプロセス、(iii)前記少なくとも一つの学習用デコード済み特徴マップの行方向を第1方向とし、その列方向を第2方向として、前記少なくとも一つの学習用デコード済み特徴マップを前記第1方向に対し第1間隔ずつ区画し、前記第2方向に対し第2間隔ずつ区画することで、複数の行と複数の列を有するグリッドの各セルが生成される際、前記各々の列ごとに前記各々の行の特徴各々をチャンネル方向へコンカチネートして、少なくとも一つの学習用修正済み特徴マップを生成するプロセス、(iv)前記学習用修正済み特徴マップとそれに対応する原本正解イメージとを参照して少なくとも一つのロスを算出するものの、前記原本正解イメージは、前記各々の列をそれに対応する一番下のセルから前記第2方向に沿って確認した際、近接障害物各々が実際に位置する原本正解ポジション各々に対応する前記各々の行を示すイメージであることを特徴とするプロセス、及び(v)前記ロスをバックプロパゲーションして前記CNNのパラメータを学習するプロセスを遂行した状態で、テスト装置が少なくとも一つのテストイメージを獲得し、前記第1コンボリューションレイヤないし前記第 n コンボリューションレイヤをもって逐次的に前記テストイメージから第1テスト用エンコード済み特徴マップないし第 n テスト用エンコード済み特徴マップを各々生成するようにする段階；

30

40

(b) 前記テスト装置が、前記デコンボリューションレイヤのうち少なくとも一つをもって前記第 n テスト用エンコード済み特徴マップから少なくとも一つのテスト用デコード済み特徴マップを生成するようにする段階；

(c) 前記テスト装置が、前記各々の列ごとに前記各々の行のテスト用特徴各々を前記チャンネル方向にコンカチネートし、少なくとも一つのテスト用修正済み特徴マップを生成

50

する段階；及び

(d) 前記テスト装置が、前記テスト用修正済み特徴マップの前記各々の列に対してコンカチネートされたチャンネル上で前記障害物各々の一番下の部分各々の推定位置各々を確認することにより、セグメンテーション結果を生成し、前記障害物を検出する段階；
を含むことを特徴とする方法。

【請求項 11】

前記 (ii) プロセスで、

前記学習装置は、前記第 n コンボリューションレイヤに対応する前記第 n デコンボリューションレイヤをもって前記第 n 学習用エンコード済み特徴マップを前記第 1 方向へサイズを所定倍数に拡大し、前記第 n 学習用デコード済み特徴マップを生成するようにし、

前記 (b) 段階で、

前記テスト装置は、前記第 n デコンボリューションレイヤをもって前記第 n テスト用エンコード済み特徴マップを前記第 1 方向へサイズを所定倍数に拡大し、少なくとも一つの前記テスト用デコード済み特徴マップを生成することを特徴とする請求項 10 に記載の方法。

【請求項 12】

前記原本正解イメージ及び前記セグメンテーション結果は、各々前記トレーニングイメージと前記テストイメージとを前記第 2 方向に沿って N_c 個の行に分割して、前記列ごとに前記障害物各々が前記行のうちどの行にあるかを示す情報を含み、

前記学習用修正済み特徴マップは、 N_c 個のチャンネルを含み、前記テスト用修正済み特徴マップは、 N_c 個のチャンネルを含むことを特徴とする請求項 10 に記載の方法。

【請求項 13】

前記 (iii) プロセスは、

(iii-1) 前記学習装置が、 $C_i \times W_i \times H_i$ のサイズを有する前記学習用デコード済み特徴マップを利用して $C_i \cdot H_i \times W_i \times 1$ のサイズを有する第 1 学習用修正済み特徴マップを生成するものの、前記 C_i は、前記学習用デコード済み特徴マップのチャンネル数であり、前記 W_i は、前記学習用デコード済み特徴マップの列のサイズであり、前記 H_i は、前記学習用デコード済み特徴マップの行のサイズであるプロセス；

(iii-2) 前記学習装置が、前記 $C_i \cdot H_i \times W_i \times 1$ のサイズを有する前記第 1 学習用修正済み特徴マップを $(W_I / W_i) \times N_c$ のサイズを有する前記第 1 学習用修正済み特徴マップへ変更するコンボリューション演算を少なくとも一度遂行するものの、前記 W_I は、前記トレーニングイメージの列のサイズであり、前記 W_i は、前記学習用デコード済み特徴マップの前記列のサイズであるプロセス；

(iii-3) 前記学習装置が、 $(W_I / W_i) \times N_c$ のサイズを有する前記第 1 学習用修正済み特徴マップを利用して $N_c \times W_I \times 1$ のサイズを有する第 2 学習用修正済み特徴マップを生成するものの、前記 N_c は前記障害物各々の前記列ごとの位置を特定するために前記トレーニングイメージが分割された行の数であるプロセス；及び

(iii-4) 前記学習装置が、前記第 2 学習用修正済み特徴マップの前記列ごとに前記 N_c 個のチャンネルをノーマライジング (normalizing) するソフトマックス出力プロセスを遂行するプロセス；

を含み、

前記 (iv) プロセスで、

前記 (iii-4) プロセスで出力された、前記原本正解イメージに対応する各々のソフトマックス出力と前記原本正解イメージを参照して、前記ロスが算出され、

前記 (c) 段階は、

(c1) 前記テスト装置が、 $C_j \times W_j \times H_j$ のサイズを有する前記テスト用デコード済み特徴マップを利用して $C_j \cdot H_j \times W_j \times 1$ のサイズを有する第 1 テスト用修正済み特徴マップを生成するものの、前記 C_j は、前記テスト用デコード済み特徴マップのチャンネル数であり、前記 W_j は、前記テスト用デコード済み特徴マップの列のサイズであり、前記 H_j は、前記テスト用デコード済み特徴マップの行のサイズである段階；

(c2) 前記テスト装置が、前記 $C_j \cdot H_j \times W_j \times 1$ のサイズを有する前記第1テスト用修正済み特徴マップを $(W_j / W_j) \times N_c \times W_j \times 1$ のサイズを有する前記第1テスト用修正済み特徴マップへ変更するコンボリユーション演算を少なくとも一度遂行するものの、前記 W_j は、前記テストイメージの列のサイズであり、前記 W_j は、前記テスト用デコード済み特徴マップの前記列のサイズである段階；

(c3) 前記テスト装置が、 $(W_j / W_j) \times N_c \times W_j \times 1$ のサイズを有する前記第1テスト用修正済み特徴マップを利用して $N_c \times W_j \times 1$ のサイズを有する第2テスト用修正済み特徴マップを生成するものの、前記 N_c は前記障害物各々の前記列ごとの位置を特定するために前記テストイメージを分割した行の数である段階；及び

(c4) 前記テスト装置が、前記第2テスト用修正済み特徴マップの前記列ごとに前記 N_c 個のチャンネルをノーマライジング (normalizing) するソフトマックス出力プロセスを遂行する段階；

を含むことを特徴とする請求項 10 に記載の方法。

【請求項 14】

前記 (c4) 段階で、

前記チャンネルのうち、前記列ごとに最も大きい値を有する特定チャンネル各々に前記障害物各々の一番下の部分各々が存在することが決定されれば、前記 N_c 個の行のうち前記行ごとの前記一番下の部分各々の前記推定位置が各々最も大きい値を有し、前記行のうち残りの行の位置はより小さい値を有するように前記セグメンテーション結果が生成されることを特徴とする請求項 13 に記載の方法。

【請求項 15】

少なくとも一つのテストイメージから少なくとも一つの障害物を検出するための CNN のパラメータを学習するための学習装置において、

前記トレーニングイメージを受信するための通信部；及び

(i) 第1コンボリユーションレイヤないし第 n コンボリユーションレイヤをもって、前記トレーニングイメージから逐次的に第1エンコード済み特徴マップないし第 n エンコード済み特徴マップを各々生成するようにするプロセス、(ii) 第 n デコンボリユーションレイヤないし第1デコンボリユーションレイヤをもって、前記第 n エンコード済み特徴マップから逐次的に第 n デコード済み特徴マップないし第1デコード済み特徴マップを生成するようにするプロセス、(iii) 前記少なくとも一つのデコード済み特徴マップの行方向を第1方向とし、列方向を第2方向として、前記少なくとも一つのデコード済み特徴マップを前記第1方向に対し第1間隔ずつ区画し、前記第2方向に対し第2間隔ずつ区画することで、複数の行と複数の列を有するグリッドの各セルが生成される際、前記各々の列ごとに前記各々の行の特徴各々をチャンネル方向へコンカチネートして、少なくとも一つの修正済み特徴マップを生成するプロセス、(iv) 前記修正済み特徴マップとそれに対応する原本正解イメージを参照して少なくとも一つのロスを算出するものの、前記原本正解イメージは、前記各々の列をそれに対応する一番下のセルから前記第2方向に沿って確認した際、近接障害物各々が実際に位置する原本正解ポジション各々に対応する前記各々の行を示すイメージであることを特徴とするプロセス、及び(v) 前記ロスをバックプロパゲーションして前記 CNN の前記パラメータを学習するプロセスを遂行するための

を含むことを特徴とする装置。

【請求項 16】

前記 (ii) プロセスで、

前記プロセッサは、前記第 n コンボリユーションレイヤに対応する前記第 n デコンボリユーションレイヤをもって前記第 n エンコード済み特徴マップを前記第1方向へサイズを所定倍数に拡大し、前記第 n デコード済み特徴マップを生成するようにすることを特徴とする請求項 15 に記載の装置。

【請求項 17】

前記 (iii) プロセスは、

前記各々の列をそれに対応する一番下のセルから始まり前記第2方向に沿って確認することで、前記近接障害物各々が位置すると判断される、前記少なくとも一つのデコード済み特徴マップ内の推定位置各々を示すセグメンテーション結果を生成するプロセス；

をさらに含み、

前記(i v)プロセスで、

前記セグメンテーション結果と前記原本正解イメージとを参照して前記ロスが算出されることを特徴とする請求項15に記載の装置。

【請求項18】

前記(i i i)プロセスで、

前記プロセッサは、前記修正済み特徴マップの前記各々の列に対してコンカチネートされたチャンネル上で前記障害物各々の一番下の部分各々の前記推定位置各々を確認することにより、前記セグメンテーション結果を生成することを特徴とする請求項17に記載の装置。

10

【請求項19】

前記各々の列は、前記第1方向に少なくとも一つのピクセルを含み、前記各々の行は前記第2方向に少なくとも一つのピクセルを含むことを特徴とする請求項15に記載の装置。

【請求項20】

前記原本正解イメージ及び前記セグメンテーション結果は、前記トレーニングイメージを前記第2方向へN c個の行に分割して、前記列ごとに前記障害物各々が前記行のうちどの行にあるかを示す情報を含み、前記修正済み特徴マップはN c個のチャンネルを含むことを特徴とする請求項17に記載の装置。

20

【請求項21】

前記修正済み特徴マップは、前記デコード済み特徴マップを区画して生じた前記各々の行をチャンネル方向へ前記列ごとにコンカチネートして生成されることを特徴とする請求項15に記載の装置。

【請求項22】

前記(i i i)プロセスは

(i i i - 1) $C_i \times W_i \times H_i$ のサイズを有する前記デコード済み特徴マップを利用して $C_i \cdot H_i \times W_i \times 1$ のサイズを有する第1修正済み特徴マップを生成するものの、前記 C_i は、前記デコード済み特徴マップのチャンネル数であり、前記 W_i は、前記デコード済み特徴マップの列のサイズであり、前記 H_i は、前記デコード済み特徴マップの行のサイズであるプロセス；

30

(i i i - 2) 前記 $C_i \cdot H_i \times W_i \times 1$ のサイズを有する前記第1修正済み特徴マップを $(W_I / W_i) \times N_c \times W_i \times 1$ のサイズを有する前記第1修正済み特徴マップへ変更するコンボリューション演算を少なくとも一度遂行するものの、前記 W_I は、前記トレーニングイメージの列のサイズであり、前記 W_i は、前記デコード済み特徴マップの前記列のサイズであるプロセス；

(i i i - 3) $(W_I / W_i) \times N_c \times W_i \times 1$ のサイズを有する前記第1修正済み特徴マップを利用して $N_c \times W_I \times 1$ のサイズを有する第2修正済み特徴マップを生成するものの、前記 N_c は前記障害物各々の前記列ごとの位置を特定するために前記トレーニングイメージを分割した行の数であるプロセス；及び

40

(i i i - 4) 前記第2修正済み特徴マップの前記列ごとに前記N c個のチャンネルをノーマライジングするソフトマックス出力プロセスを遂行するプロセス；

を含み、

前記(i v)プロセスで、

前記(i i i - 4)プロセスで出力された、前記原本正解イメージに対応する各々のソフトマックス出力と前記原本正解イメージとを参照して、前記ロスが算出されることを特徴とする請求項21に記載の装置。

【請求項23】

50

前記第1ないし前記第nコンボリューションレイヤは、コンボリューション、マックスプーリング(max pooling)、ReLUの中から少なくとも一つの演算を遂行して前記トレーニングイメージまたは任意の特徴マップの大きさを縮小させることで、少なくとも一つの縮小された(reduced)特徴マップを生成し、前記第nないし第1デコンボリューションレイヤは、デコンボリューション及びReLUの中から少なくとも一つの演算を遂行して入力された特徴マップの大きさを拡大した、少なくとも一つの拡大された(enlarged)特徴マップを生成することを特徴とする請求項15に記載の装置。

【請求項24】

少なくとも一つのテストイメージから少なくとも一つの障害物を検出するためのCNNテスト装置において、

学習装置が、(i)トレーニングイメージを受信し、第1コンボリューションレイヤないし第n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から逐次的に第1学習用エンコード済み特徴マップないし第n学習用エンコード済み特徴マップを各々生成するようにするプロセス、(ii)第nデコンボリューションレイヤないし第1デコンボリューションレイヤをもって、前記第n学習用エンコード済み特徴マップから逐次的に第n学習用デコード済み特徴マップないし第1学習用デコード済み特徴マップを生成するようにするプロセス、(iii)前記少なくとも一つの学習用デコード済み特徴マップの行方向を第1方向とし、その列方向を第2方向として、前記少なくとも一つの学習用デコード済み特徴マップを前記第1方向に対し第1間隔ずつ区画し、前記第2方向に対し第2間隔ずつ区画することで、複数の行と複数の列を有するグリッドの各セルが生成される際、前記学習装置が、前記各々の列ごとに前記各々の行の特徴各々をチャネル方向へコンカチネートして、少なくとも一つの学習用修正済み特徴マップを生成するプロセス、(iv)前記学習用修正済み特徴マップとそれに対応する原本正解イメージを参照して少なくとも一つのロスを算出するものの、前記原本正解イメージは、前記各々の列をそれに対応する一番下のセルから前記第2方向に沿って確認した際、近接障害物各々が実際に位置する原本正解ポジション各々に対応する前記各々の行を示すイメージであることを特徴とするプロセス、及び(v)前記ロスをバックプロパゲーションして前記CNNのパラメータを学習するプロセスを遂行した状態で、前記テストイメージを獲得する通信部；及び

(I)前記第1コンボリューションレイヤないし前記第nコンボリューションレイヤをもって、逐次的に前記テストイメージから第1テスト用エンコード済み特徴マップないし第nテスト用エンコード済み特徴マップを各々生成するようにするプロセス、(II)前記デコンボリューションレイヤのうち少なくとも一つをもって前記第nテスト用エンコード済み特徴マップから少なくとも一つのテスト用デコード済み特徴マップを生成するようにするプロセス、(III)前記各々の列ごとに前記各々の行のテスト用特徴各々を前記チャネル方向にコンカチネートし、少なくとも一つのテスト用修正済み特徴マップを生成するプロセス、及び(IV)前記テスト用修正済み特徴マップの前記各々の列に対してコンカチネートされたチャネル上で前記障害物各々の一番下の部分各々の推定位置各々を確認することにより、セグメンテーション結果を生成し、前記障害物を検出するプロセスを遂行するプロセッサ；

を含むことを特徴とする装置。

【請求項25】

前記(ii)プロセスで、

前記学習装置は、前記第nコンボリューションレイヤに対応する前記第nデコンボリューションレイヤをもって前記第n学習用エンコード済み特徴マップを前記第1方向へサイズを所定倍数に拡大し、前記第n学習用デコード済み特徴マップを生成するようにし、

前記(II)プロセスで、

前記プロセッサは、前記第nデコンボリューションレイヤをもって前記第nテスト用エンコード済み特徴マップを前記第1方向へサイズを所定倍数に拡大し、少なくとも一つの前記テスト用デコード済み特徴マップを生成することを特徴とする請求項24に記載の装

10

20

30

40

50

置。

【請求項 26】

前記原本正解イメージ及び前記セグメンテーション結果は、各々前記トレーニングイメージと前記テストイメージとを前記第2方向に沿って N_c 個の行に分割して、前記列ごとに前記障害物各々が前記行のうちどの行にあるかを示す情報を含み、

前記学習用修正済み特徴マップは、 N_c 個のチャンネルを含み、前記テスト用修正済み特徴マップは、 N_c 個のチャンネルを含むことを特徴とする請求項24に記載の装置。

【請求項 27】

前記(iii)プロセスは、

(iii-1) $C_i \times W_i \times H_i$ のサイズを有する前記学習用デコード済み特徴マップを利用して $C_i \cdot H_i \times W_i \times 1$ のサイズを有する第1学習用修正済み特徴マップを生成するものの、前記 C_i は、前記学習用デコード済み特徴マップのチャンネル数であり、前記 W_i は、前記学習用デコード済み特徴マップの列のサイズであり、前記 H_i は、前記学習用デコード済み特徴マップの行のサイズであるプロセス；

(iii-2) 前記 $C_i \cdot H_i \times W_i \times 1$ のサイズを有する前記第1学習用修正済み特徴マップを $(W_I / W_i) \times N_c \times W_i \times 1$ のサイズを有する前記第1学習用修正済み特徴マップへ変更するコンボリューション演算を少なくとも一度遂行するものの、前記 W_I は、前記トレーニングイメージの列のサイズであり、前記 W_i は、前記学習用デコード済み特徴マップの前記列のサイズであるプロセス；

(iii-3) $(W_I / W_i) \times N_c \times W_i \times 1$ のサイズを有する前記第1学習用修正済み特徴マップを利用して $N_c \times W_I \times 1$ のサイズを有する第2学習用修正済み特徴マップを生成するものの、前記 N_c は前記障害物各々の前記列ごとの位置を特定するために前記トレーニングイメージが分割された行の数であるプロセス；及び

(iii-4) 前記第2学習用修正済み特徴マップの前記列ごとに前記 N_c 個のチャンネルをノーマライジングするソフトマックス出力プロセスを遂行するプロセス；

を含み、

前記(iv)プロセスで、

前記(iii-4)プロセスで出力された、前記原本正解イメージに対応する各々のソフトマックス出力と前記原本正解イメージとを参照して前記ロスが算出され、

前記(III)プロセスは、

(III-1) $C_j \times W_j \times H_j$ のサイズを有する前記テスト用デコード済み特徴マップを利用して $C_j \cdot H_j \times W_j \times 1$ のサイズを有する第1テスト用修正済み特徴マップを生成するものの、前記 C_j は、前記テスト用デコード済み特徴マップのチャンネル数であり、前記 W_j は、前記テスト用デコード済み特徴マップの列のサイズであり、前記 H_j は、前記テスト用デコード済み特徴マップの行のサイズであるプロセス；

(III-2) 前記 $C_j \cdot H_j \times W_j \times 1$ のサイズを有する前記第1テスト用修正済み特徴マップを $(W_J / W_j) \times N_c \times W_j \times 1$ のサイズを有する前記第1テスト用修正済み特徴マップへ変更するコンボリューション演算を少なくとも一度遂行するものの、前記 W_J は、前記テストイメージの列のサイズであり、前記 W_j は、前記テスト用デコード済み特徴マップの前記列のサイズであるプロセス；

(III-3) $(W_J / W_j) \times N_c \times W_j \times 1$ のサイズを有する前記第1テスト用修正済み特徴マップを利用して $N_c \times W_J \times 1$ のサイズを有する第2テスト用修正済み特徴マップを生成するものの、前記 N_c は前記障害物各々の前記列ごとの位置を特定するために前記テストイメージを分割した行の数であるプロセス；及び

(III-4) 前記第2テスト用修正済み特徴マップの前記列ごとに前記 N_c 個のチャンネルをノーマライジング(normalizing)するソフトマックス出力プロセスを遂行するプロセス；

を含むことを特徴とする請求項24に記載の装置。

【請求項 28】

前記(III-4)プロセスで、

前記チャンネルのうち、前記列ごとに最も大きい値を有する特定チャンネル各々に前記障害物各々の一番下の部分各々が存在することが決定されれば、前記N c 個の行のうち前記行ごとの前記一番下の部分各々の前記推定位置が各々最も大きい値を有し、前記行のうち残りの行の位置はより小さい値を有するように前記セグメンテーション結果が生成されることを特徴とする請求項 27 に記載の装置。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、少なくとも一つのトレーニングイメージから少なくとも一つの障害物を検出するためのCNNのパラメータ学習方法に関し；より詳細には、前記トレーニングイメージから前記障害物を検出するための前記CNNのパラメータ学習方法において、(a)学習装置が、前記トレーニングイメージを獲得し、第1コンボリューションレイヤないし第n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から逐次的に第1エンコード済み特徴マップないし第nエンコード済み特徴マップを各々生成するようにする段階；(b)第nデコンボリューションレイヤないし第1デコンボリューションレイヤをもって、前記第nエンコード済み特徴マップから逐次的に第nデコード済み特徴マップないし第1デコード済み特徴マップを生成するようにする段階；(c)前記少なくとも一つのデコード済み特徴マップの前記行方向を第1方向とし、その列方向を第2方向として、前記少なくとも一つのデコード済み特徴マップを前記第1方向に対し第1間隔ずつ区画し、前記第2方向に対し第2間隔ずつ区画することで、複数の行と複数の列を有するグリッドの各セルが生成される際、前記各々の列ごとに前記各々の行の特徴各々をチャンネル方向へコンカチネートして、少なくとも一つの修正済み特徴マップを生成する段階；(d)前記学習装置が前記修正済み特徴マップとそれに対応する原本正解(Ground Truth)イメージとを参照して少なくとも一つのロスを算出するものの、前記原本正解イメージは、前記各々の列をそれに対応する一番下のセルから前記第2方向に沿って確認した際、近接障害物各々が実際に位置する原本正解ポジション各々に対応する前記各々の行を示すイメージであることを特徴とする段階；及び(e)前記ロスをバックプロパゲーションして前記CNNの前記パラメータを学習する段階；を含むことを特徴とする方法及びこれを利用した学習装置、テスト方法及び装置に関する。

【背景技術】

【0002】

ディープラーニング(Deep Learning)は、モノやデータを群集化・分類するのに用いられる技術である。例えば、コンピュータは写真だけで犬と猫を区別することができない。しかし、人はとても簡単に区別できる。このため「機械学習(Machine Learning)」という方法が考案された。多くのデータをコンピュータに入力し、類似したものを分類するようにする技術である。保存されている犬の写真と似たような写真が入力されると、これを犬の写真だとコンピュータが分類するようにしたのである。

【0003】

データをどのように分類するかをめぐり、すでに多くの機械学習アルゴリズムが登場した。「決定木」や「ベイジアンネットワーク」「サポートベクターマシン(SVM)」「人工神経網」などが代表的だ。このうち、ディープラーニングは人工神経網の後裔だ。

【0004】

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 Convolution 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くために90年代にすでに使われたが、現在のように広く使われるようになったのは最近の研究結果のおかげだ。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分野で非常に有用なツールとなった。

【 0 0 0 5 】

図 1 は従来の CNN を用いた一般的なセグメンテーションの従来のプロセスを簡略的に示した図面である。

【 0 0 0 6 】

図 1 を参照すれば、従来の車線検出方法では、学習装置が、入力イメージの入力を受けて、複数のコンボリューションレイヤで数回のコンボリューション演算と ReLU などの非線形演算を遂行して特徴マップを生成し、前記特徴マップの最後のマップに対して複数のデコンボリューションレイヤで複数回のデコンボリューション演算及びソフトマックス (SoftMax) 演算を行うことでセグメンテーション結果を生成する。

【 0 0 0 7 】

一方、従来の道路セグメンテーション方法では、前記入力イメージからすべてのピクセルをセグメンテーションして、すべてのピクセルを見てこのピクセルが道路に該当するピクセルか、道路に該当しないピクセルかを区別しなければならなかった。この方法では、すべてのピクセルに対して判断を行うため演算量が多いという問題点が存在する。

【 0 0 0 8 】

一方、自動車の自律走行のために前記道路セグメンテーションをする際には、前記入力イメージから全ての物体や車線上の全ての物体をセグメンテーションする必要はなく、自律走行の妨げとなる障害物 (obstacle) だけを検出すれば充分である。

【 0 0 0 9 】

従って、入力イメージから前記障害物のみを検出する新たな手法の提示が求められている。

【 発明の概要 】

【 発明が解決しようとする課題 】

【 0 0 1 0 】

本発明は、前述した問題点をすべて解決することをその目的とする。

【 0 0 1 1 】

また、本発明は、自動車の自律走行のために、障害物を検出する新たな手法を提供することを目的とする。

【 0 0 1 2 】

また、本発明は、入力イメージ内の全てのピクセルを検討することなく、少ない演算量で前記障害物の位置だけを素早く把握し得る方法を提示することをまたの目的とする。

【 課題を解決するための手段 】

【 0 0 1 3 】

本発明の一態様によれば、少なくとも一つのトレーニングイメージから少なくとも一つの障害物を検出するための CNN のパラメータを学習する方法において、(a) 学習装置が、前記トレーニングイメージを獲得し、第 1 コンボリューションレイヤないし第 n 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から逐次的に第 1 エンコード済み特徴マップないし第 n エンコード済み特徴マップを各々生成するようにする段階；(b) 前記学習装置が、第 n デコンボリューションレイヤないし第 1 デコンボリューションレイヤをもって、前記第 n エンコード済み特徴マップから逐次的に第 n デコード済み特徴マップないし第 1 デコード済み特徴マップを生成するようにする段階；(c) 前記少なくとも一つのデコード済み特徴マップの前記行方向を第 1 方向とし、その列方向を第 2 方向として、前記少なくとも一つのデコード済み特徴マップを前記第 1 方向に対し第 1 間隔ずつ区画し、前記第 2 方向に対し第 2 間隔ずつ区画することで、複数の行と複数の列を有するグリッドの各セルが生成される際、前記学習装置が、前記各々の列ごとに前記各々の行の特徴各々をチャネル方向へコンカチネートして、少なくとも一つの修正済み特徴マップを生成する段階；(d) 前記学習装置が、前記修正済み特徴マップとそれに対応する原本正解 (Ground Truth) イメージとを参照して少なくとも一つのロスを算出するものの、前記原本正解イメージは、前記各々の列をそれに対応する一番下のセルから前記第 2 方向に沿って確認した際、近接障害物各々が実際に位置する原本正解ポジション各

10

20

30

40

50

々に対応する前記各々の行を示すイメージであることを特徴とする段階；及び(e)前記学習装置が、前記ロスをバックプロパゲーションして前記CNNの前記パラメータを学習する段階；を含むことを特徴とする。

【0014】

一例として、前記(b)段階で、前記学習装置は、前記第nコンボリューションレイヤに対応する前記第nデコンボリューションレイヤをもって前記第nエンコード済み特徴マップを前記第1方向へサイズを所定倍数に拡大し、前記第nデコード済み特徴マップを生成するようにすることを特徴とする。

【0015】

一例として、前記(c)段階は、前記学習装置が、前記各々の列をそれに対応する一番下のセルから始まり前記第2方向に沿って確認することで、前記近接障害物各々が位置すると判断される、前記少なくとも一つのデコード済み特徴マップ内の推定位置各々を示すセグメンテーション結果を生成する段階；をさらに含み、前記(d)段階で、前記セグメンテーション結果と前記原本正解イメージとを参照して前記ロスが算出されることを特徴とする。

10

【0016】

一例として、前記(c)段階で、前記学習装置が、前記修正済み特徴マップの前記各々の列に対してコンカチネートされたチャンネル上で前記障害物各々の一番下の部分各々の前記推定位置各々を確認することにより、前記セグメンテーション結果を生成することを特徴とする。

20

【0017】

一例として、前記各々の列は、前記第1方向に少なくとも一つのピクセルを含み、前記各々の行は前記第2方向に少なくとも一つのピクセルを含むことを特徴とする。

【0018】

一例として、前記原本正解イメージ及び前記セグメンテーション結果は、前記トレーニングイメージを前記第2方向へ N_c 個の行に分割して、前記列ごとに前記障害物各々が前記行のうちどの行にあるかを示す情報を含み、前記修正済み特徴マップは N_c 個のチャンネルを含むことを特徴とする。

【0019】

一例として、前記修正済み特徴マップは、前記デコード済み特徴マップを区画して生じた前記各々の行をチャンネル方向へ前記列ごとにコンカチネートして生成されることを特徴とする。

30

【0020】

一例として、前記(c)段階は、(c1)前記学習装置が $C_i \times W_i \times H_i$ のサイズを有する前記デコード済み特徴マップを利用して $C_i \cdot H_i \times W_i \times 1$ のサイズを有する第1修正済み特徴マップを生成するものの、前記 C_i は、前記デコード済み特徴マップのチャンネル数であり、前記 W_i は、前記デコード済み特徴マップの列のサイズであり、前記 H_i は、前記デコード済み特徴マップの行のサイズである段階；(c2)前記学習装置が、前記 $C_i \cdot H_i \times W_i \times 1$ のサイズを有する前記第1修正済み特徴マップを $(W_I / W_i) \times N_c \times W_i \times 1$ のサイズを有する前記第1修正済み特徴マップへ変更するコンボリューション演算を少なくとも一度遂行するものの、前記 W_I は、前記トレーニングイメージの列のサイズであり、前記 W_i は、前記デコード済み特徴マップの前記列のサイズである段階；(c3)前記学習装置が、 $(W_I / W_i) \times N_c \times W_i \times 1$ のサイズを有する前記第1修正済み特徴マップを利用して $N_c \times W_I \times 1$ のサイズを有する第2修正済み特徴マップを生成するものの、前記 N_c は前記障害物各々の前記列ごとの位置を特定するために前記トレーニングイメージを分割した行の数である段階；及び(c4)前記学習装置が、前記第2修正済み特徴マップの前記列ごとに前記 N_c 個のチャンネルをノーマライジング(normalizing)するソフトマックス出力プロセスを遂行する段階；を含み、前記(d)段階で、前記(c4)段階で出力された、前記原本正解イメージに対応する各々のソフトマックス出力と前記原本正解イメージを参照して、前記ロスが算出され

40

50

ることを特徴とする。

【0021】

一例として、前記(c4)段階は、前記チャンネルのうち、前記列ごとに最も大きい値を有する特定チャンネル各々に前記障害物各々の一番下の部分各々が存在することが決定されれば、前記N_c個の行のうち前記行ごとの前記一番下の部分の各々の前記推定位置が各々最も大きい値を有し、前記行のうち残りの行の位置はより小さい値を有するように前記セグメンテーション結果が生成されることを特徴とする。

【0022】

一例として、前記第1ないし前記第nコンボリューションレイヤは、コンボリューション、マックスプーリング(max pooling)、ReLUの中から少なくとも一つの演算を遂行して前記トレーニングイメージまたは任意の特徴マップの大きさを縮小させることで、少なくとも一つの縮小された(reduced)特徴マップを生成し、前記第nないし第1デコンボリューションレイヤは、デコンボリューション及びReLUの中から少なくとも一つの演算を遂行して入力された特徴マップの大きさを拡大した、少なくとも一つの拡大された(enlarged)特徴マップを生成することを特徴とする。

【0023】

本発明の他の態様によれば、少なくとも一つのテストイメージから少なくとも一つの障害物を検出するためのCNNテスト方法において、(a)学習装置が、(i)トレーニングイメージを受信し、第1コンボリューションレイヤないし第n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から逐次的に第1学習用エンコード済み特徴マップないし第n学習用エンコード済み特徴マップを各々生成するようにするプロセス、(ii)第nデコンボリューションレイヤないし第1デコンボリューションレイヤをもって、前記第n学習用エンコード済み特徴マップから逐次的に第n学習用デコード済み特徴マップないし第1学習用デコード済み特徴マップを生成するようにするプロセス、(iii)前記少なくとも一つの学習用デコード済み特徴マップの前記行方向を第1方向とし、その列方向を第2方向として、前記少なくとも一つの学習用デコード済み特徴マップを前記第1方向に対し第1間隔ずつ区画し、前記第2方向に対し第2間隔ずつ区画することで、複数の行と複数の列を有するグリッドの各セルが生成される際、前記各々の列ごとに前記各々の行の特徴各々をチャンネル方向へコンカチネートして、少なくとも一つの学習用修正済み特徴マップを生成するプロセス、(iv)前記学習用修正済み特徴マップとそれに対応する原本正解イメージとを参照して少なくとも一つのロスを算出するものの、前記原本正解イメージは、前記各々の列をそれに対応する一番下のセルから前記第2方向に沿って確認した際、近接障害物各々が実際に位置する原本正解ポジション各々に対応する前記各々の行を示すイメージであることを特徴とするプロセス、及び(v)前記ロスをバックプロパゲーションして前記CNNのパラメータを学習するプロセスを遂行した状態で、テスト装置が少なくとも一つのテストイメージを獲得し、前記第1コンボリューションレイヤないし前記第nコンボリューションレイヤをもって逐次的に前記テストイメージから第1テスト用エンコード済み特徴マップないし第nテスト用エンコード済み特徴マップを各々生成するようにする段階；(b)前記テスト装置が、前記デコンボリューションレイヤのうち少なくとも一つのテスト用デコード済み特徴マップを生成するようにする段階；(c)前記テスト装置が、前記各々の列ごとに前記各々の行のテスト用特徴各々を前記チャンネル方向にコンカチネートし、少なくとも一つのテスト用修正済み特徴マップを生成する段階；及び(d)前記テスト装置が、前記テスト用修正済み特徴マップの前記各々の列に対してコンカチネートされたチャンネル上で前記障害物各々の一番下の部分各々の推定位置各々を確認することにより、セグメンテーション結果を生成し、前記障害物を検出する段階；を含むことを特徴とする。

【0024】

一例として、前記(ii)プロセスで、前記学習装置は、前記第nコンボリューションレイヤに対応する前記第nデコンボリューションレイヤをもって前記第n学習用エンコ

10

20

30

40

50

ード済み特徴マップを前記第1方向へサイズを所定倍数に拡大し、前記第n学習用デコード済み特徴マップを生成するようにし、前記(b)段階で、前記テスト装置は、前記第nデコンボリューションレイヤをもって前記第nテスト用エンコード済み特徴マップを前記第1方向へサイズを所定倍数に拡大し、前記第nテスト用デコード済み特徴マップを生成することを特徴とする。

【0025】

一例として、前記原本正解イメージ及び前記セグメンテーション結果は、各々前記トレーニングイメージと前記テストイメージとを前記第2方向に沿って N_c 個の行に分割して、前記列ごとに前記障害物各々が前記行のうちどの行にあるかを示す情報を含み、前記学習用修正済み特徴マップは、 N_c 個のチャンネルを含み、前記テスト用修正済み特徴マップは、 N_c 個のチャンネルを含むことを特徴とする。

10

【0026】

一例として、前記(iii)プロセスは、(iii-1)前記学習装置が $C_i \times W_i \times H_i$ のサイズを有する前記学習用デコード済み特徴マップを利用して $C_i \cdot H_i \times W_i \times 1$ のサイズを有する前記第1学習用修正済み特徴マップを生成するものの、前記 C_i は、前記学習用デコード済み特徴マップのチャンネル数であり、前記 W_i は、前記学習用デコード済み特徴マップの列のサイズであり、前記 H_i は、前記学習用デコード済み特徴マップの行のサイズであるプロセス；(iii-2)前記学習装置が、前記 $C_i \cdot H_i \times W_i \times 1$ のサイズを有する前記第1学習用修正済み特徴マップを $((W_I / W_i) \times N_c) \times W_i \times 1$ のサイズを有する前記第1学習用修正済み特徴マップへ変更するコンボリューション演算を少なくとも一度遂行するものの、前記 W_I は、前記トレーニングイメージの列のサイズであり、前記 W_i は、前記学習用デコード済み特徴マップの前記列のサイズであるプロセス；(iii-3)前記学習装置が、 $((W_I / W_i) \times N_c) \times W_i \times 1$ のサイズを有する前記第1学習用修正済み特徴マップを利用して $N_c \times W_I \times 1$ のサイズを有する第2学習用修正済み特徴マップを生成するものの、前記 N_c は前記障害物各々の前記列ごとの位置を特定するために前記トレーニングイメージが分割された行の数であるプロセス；及び(iii-4)前記学習装置が、前記第2学習用修正済み特徴マップの前記列ごとに前記 N_c 個のチャンネルをノーマライジング(normalizing)するソフトウェア出力プロセスを遂行するプロセス；を含み、前記(iv)プロセスで、前記(iii-4)プロセスで出力された、前記原本正解イメージに対応する各々のソフトウェア出力と前記原本正解イメージを参照して、前記ロスが算出され、前記(c)段階は、(c1)前記テスト装置が、 $C_j \times W_j \times H_j$ のサイズを有する前記テスト用デコード済み特徴マップを利用して $C_j \cdot H_j \times W_j \times 1$ のサイズを有する第1テスト用修正済み特徴マップを生成するものの、前記 C_j は、前記テスト用デコード済み特徴マップのチャンネル数であり、前記 W_j は、前記テスト用デコード済み特徴マップの列のサイズであり、前記 H_j は、前記テスト用デコード済み特徴マップの行のサイズである段階；(c2)前記テスト装置が、前記 $C_j \cdot H_j \times W_i \times 1$ のサイズを有する前記第1テスト用修正済み特徴マップを $((W_J / W_j) \times N_c) \times W_j \times 1$ のサイズを有する前記第1テスト用修正済み特徴マップへ変更するコンボリューション演算を少なくとも一度遂行するものの、前記 W_J は、前記テストイメージの列のサイズであり、前記 W_j は、前記テスト用デコード済み特徴マップの前記列のサイズである段階；(c3)前記テスト装置が、 $((W_J / W_j) \times N_c) \times W_j \times 1$ のサイズを有する前記第1テスト用修正済み特徴マップを利用して $N_c \times W_J \times 1$ のサイズを有する第2テスト用修正済み特徴マップを生成するものの、前記 N_c は前記障害物各々の前記列ごとの位置を特定するために前記テストイメージを分割した行の数である段階；及び(c4)前記テスト装置が、前記第2テスト用修正済み特徴マップの前記列ごとに前記 N_c 個のチャンネルをノーマライジング(normalizing)するソフトウェア出力プロセスを遂行する段階；を含むことを特徴とする。

20

30

40

【0027】

一例として、前記(c4)段階で、前記チャンネルのうち、前記列ごとに最も大きい値を有する特定チャンネル各々に前記障害物各々の一番下の部分各々が存在することが決定され

50

れば、前記 N_c 個の行のうち前記行ごとの前記一番下の部分各々の前記推定位置が各々最も大きい値を有し、前記行のうち残りの行の位置はより小さい値を有するように前記セグメンテーション結果が生成されることを特徴とする。

【0028】

本発明のまた他の態様によれば、少なくとも一つのテストイメージから少なくとも一つの障害物を検出するためのCNNのパラメータを学習するための学習装置において、前記トレーニングイメージを受信するための通信部；及び(i)第1コンボリューションレイヤないし第n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から逐次的に第1エンコード済み特徴マップないし第nエンコード済み特徴マップを各々生成するようにするプロセス、(ii)第nデコンボリューションレイヤないし第1デコンボリューションレイヤをもって、前記第nエンコード済み特徴マップから逐次的に第nデコード済み特徴マップないし第1デコード済み特徴マップを生成するようにするプロセス、(iii)前記少なくとも一つのデコード済み特徴マップの前記行方向を第1方向とし、前記列方向を第2方向として、前記少なくとも一つのデコード済み特徴マップを前記第1方向に対し第1間隔ずつ区画し、前記第2方向に対し第2間隔ずつ区画することで、複数の行と複数の列を有するグリッドの各セルが生成される際、前記各々の列ごとに前記各々の行の特徴各々をチャンネル方向へコンカチネートして、少なくとも一つの修正済み特徴マップを生成するプロセス、(iv)前記修正済み特徴マップとそれに対応する原本正解イメージを参照して少なくとも一つのロスを算出するものの、前記原本正解イメージは、前記各々の列をそれに対応する一番下のセルから前記第2方向に沿って確認した際、近接障害物各々が実際に位置する原本正解ポジション各々に対応する前記各々の行を示すイメージであることを特徴とするプロセス、及び(v)前記ロスをバックプロパゲーションして前記CNNの前記パラメータを学習するプロセスを遂行するためのプロセッサ；を含むことを特徴とする。

【0029】

一例として、前記(iii)プロセスで、前記プロセッサは、前記第nコンボリューションレイヤに対応する前記第nデコンボリューションレイヤをもって前記第nエンコード済み特徴マップを前記第1方向へサイズを所定倍数に拡大し、前記第nデコード済み特徴マップを生成するようにすることを特徴とする。

【0030】

一例として、前記(iii)プロセスは、前記各々の列をそれに対応する一番下のセルから始まり前記第2方向に沿って確認することで、前記近接障害物各々が位置すると判断される、前記少なくとも一つのデコード済み特徴マップ内の推定位置各々を示すセグメンテーション結果を生成するプロセス；をさらに含み、前記(iv)プロセスで、前記セグメンテーション結果と前記原本正解イメージとを参照して前記ロスが算出されることを特徴とする。

【0031】

一例として、前記(iii)プロセスで、前記プロセッサは、前記修正済み特徴マップの前記各々の列に対してコンカチネートされたチャンネル上で前記障害物各々の一番下の部分各々の前記推定位置各々を確認することにより、前記セグメンテーション結果を生成することを特徴とする。

【0032】

一例として、前記各々の列は、前記第1方向に少なくとも一つのピクセルを含み、前記各々の行は前記第2方向に少なくとも一つのピクセルを含むことを特徴とする。

【0033】

一例として、前記原本正解イメージ及び前記セグメンテーション結果は、前記トレーニングイメージを前記第2方向へ N_c 個の行に分割して、前記列ごとに前記障害物各々が前記行のうちどの行にあるかを示す情報を含み、前記修正済み特徴マップは N_c 個のチャンネルを含むことを特徴とする。

【0034】

10

20

30

40

50

一例として、前記修正済み特徴マップは、前記デコード済み特徴マップを区画して生じた前記各々の行をチャンネル方向へ前記列ごとにコンカチネートして生成されることを特徴とする。

【0035】

一例として、前記 ($i i i$) プロセスは ($i i i - 1$) $C_i \times W_i \times H_i$ のサイズを有する前記デコード済み特徴マップを利用して $C_i \cdot H_i \times W_i \times 1$ のサイズを有する第1修正済み特徴マップを生成するものの、前記 C_i は、前記デコード済み特徴マップのチャンネル数であり、前記 W_i は、前記デコード済み特徴マップの列のサイズであり、前記 H_i は、前記デコード済み特徴マップの行のサイズであるプロセス； ($i i i - 2$) 前記 $C_i \cdot H_i \times W_i \times 1$ のサイズを有する前記第1修正済み特徴マップを ($(W_I / W_i) \times N_c$) $\times W_i \times 1$ のサイズを有する前記第1修正済み特徴マップへ変更するコンボリューション演算を少なくとも一度遂行するものの、前記 W_I は、前記トレーニングイメージの列のサイズであり、前記 W_i は、前記デコード済み特徴マップの前記列のサイズであるプロセス； ($i i i - 3$) ($(W_I / W_i) \times N_c$) $\times W_i \times 1$ のサイズを有する前記第1修正済み特徴マップを利用して $N_c \times W_I \times 1$ のサイズを有する第2修正済み特徴マップを生成するものの、前記 N_c は前記障害物各々の前記列ごとの位置を特定するために前記トレーニングイメージを分割した行の数であるプロセス；及び ($i i i - 4$) 前記第2修正済み特徴マップの前記列ごとに前記 N_c 個のチャンネルをノーマライジングするソフトマックス出力プロセスを遂行するプロセス；を含み、前記 ($i v$) プロセスで、前記 ($i i i - 4$) プロセスで出力された、前記原本正解イメージに対応する各々のソフトマックス出力と前記原本正解イメージとを参照して、前記ロスが算出されることを特徴とする

【0036】

一例として、前記 ($i i i - 4$) プロセスは、前記チャンネルのうち、前記列ごとに最も大きい値を有する特定チャンネル各々に前記障害物各々の一番下の部分各々が存在することが決定されれば、前記 N_c 個の行のうち前記行ごとの前記一番下の部分各々の前記推定位置が各々最も大きい値を有し、前記行のうち残りの行の位置はより小さい値を有するように前記セグメンテーション結果が生成されることを特徴とする。

【0037】

一例として、前記第1ないし前記第 n コンボリューションレイヤは、コンボリューション、マックスプーリング ($max\ pooling$)、 $ReLU$ の中から少なくとも一つの演算を遂行して前記トレーニングイメージまたは任意の特徴マップの大きさを縮小させることで、少なくとも一つの縮小された ($reduced$) 特徴マップを生成し、前記第 n ないし第1デコンボリューションレイヤは、デコンボリューション及び $ReLU$ の中から少なくとも一つの演算を遂行して入力された特徴マップの大きさを拡大した、少なくとも一つの拡大された ($enlarged$) 特徴マップを生成することを特徴とする。

【0038】

本発明のまた他の態様によれば、少なくとも一つのテストイメージから少なくとも一つの障害物を検出するための CNN テスト装置において、学習装置が、 (i) トレーニングイメージを受信し、第1コンボリューションレイヤないし第 n 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から逐次的に第1学習用エンコード済み特徴マップないし第 n 学習用エンコード済み特徴マップを各々生成するようにするプロセス、 ($i i$) 第 n デコンボリューションレイヤないし第1デコンボリューションレイヤをもって、前記第 n 学習用エンコード済み特徴マップから逐次的に第 n 学習用デコード済み特徴マップないし第1学習用デコード済み特徴マップを生成するようにするプロセス、 ($i i i$) 前記少なくとも一つの学習用デコード済み特徴マップの前記行方向を第1方向とし、その列方向を第2方向として、前記少なくとも一つの学習用デコード済み特徴マップを前記第1方向に対し第1間隔ずつ区画し、前記第2方向に対し第2間隔ずつ区画することで、複数の行と複数の列を有するグリッドの各セルが生成される際、前記学習装置が、前記各々の列ごとに前記各々の行の特徴各々をチャンネル方向へコンカチネートして、少なくとも一つの学習用修正済み特徴マップを生成するプロセス、 ($i v$) 前記学習用修正済み特徴マッ

プとそれに対応する原本正解イメージを参照して少なくとも一つのロスを算出するものの、前記原本正解イメージは、前記各々の列をそれに対応する一番下のセルから前記第2方向に沿って確認した際、近接障害物各々が実際に位置する原本正解ポジション各々に対応する前記各々の行を示すイメージであることを特徴とするプロセス、及び(v)前記ロスをバックプロパゲーションして前記CNNのパラメータを学習するプロセスを遂行した状態で、前記テストイメージを獲得する通信部；及び(I)前記テストイメージを獲得し、前記第1コンボリューションレイヤないし前記第nコンボリューションレイヤをもって、逐次的に前記テストイメージから第1テスト用エンコード済み特徴マップないし第nテスト用エンコード済み特徴マップを各々生成するようにするプロセス、(II)前記デコンボリューションレイヤのうち少なくとも一つをもって前記第nテスト用エンコード済み特徴マップから少なくとも一つのテスト用デコード済み特徴マップを生成するようにするプロセス、(III)前記各々の列ごとに前記各々の行のテスト用特徴各々を前記チャネル方向にコンカチネートし、少なくとも一つのテスト用修正済み特徴マップを生成するプロセス、及び(IV)前記テスト用修正済み特徴マップの前記各々の列に対してコンカチネートされたチャネル上で前記障害物各々の一番下の部分各々の推定位置各々を確認することにより、セグメンテーション結果を生成し、前記障害物を検出するプロセスを遂行するプロセッサ；を含むことを特徴とする。

10

【0039】

一例として、前記(ii)プロセスで、前記学習装置は、前記第nコンボリューションレイヤに対応する前記第nデコンボリューションレイヤをもって前記第n学習用エンコード済み特徴マップを前記第1方向へサイズを所定倍数に拡大し、前記第n学習用デコード済み特徴マップを生成するようにし、前記(II)プロセスで、前記プロセッサは、前記第nデコンボリューションレイヤをもって前記第nテスト用エンコード済み特徴マップを前記第1方向へサイズを所定倍数に拡大し、前記第nテスト用デコード済み特徴マップを生成することを特徴とする。

20

【0040】

一例として、前記原本正解イメージ及び前記セグメンテーション結果は、各々前記トレーニングイメージと前記テストイメージとを前記第2方向に沿って N_c 個の行に分割して、前記列ごとに前記障害物各々が前記行のうちどの行にあるかを示す情報を含み、前記学習用修正済み特徴マップは、 N_c 個のチャネルを含み、前記テスト用修正済み特徴マップは、 N_c 個のチャネルを含むことを特徴とする

30

【0041】

一例として、前記(iii)プロセスは、 $(iii-1)$ $C_i \times W_i \times H_i$ のサイズを有する前記学習用デコード済み特徴マップを利用して $C_i \cdot H_i \times W_i \times 1$ のサイズを有する前記第1学習用修正済み特徴マップを生成するものの、前記 C_i は、前記学習用デコード済み特徴マップのチャネル数であり、前記 W_i は、前記学習用デコード済み特徴マップの列のサイズであり、前記 H_i は、前記学習用デコード済み特徴マップの行のサイズであるプロセス； $(iii-2)$ 前記 $C_i \cdot H_i \times W_i \times 1$ のサイズを有する前記第1学習用修正済み特徴マップを $(W_I / W_i) \times N_c \times W_i \times 1$ のサイズを有する前記第1学習用修正済み特徴マップへ変更するコンボリューション演算を少なくとも一度遂行するものの、前記 W_I は、前記トレーニングイメージの列のサイズであり、前記 W_i は、前記学習用デコード済み特徴マップの前記列のサイズであるプロセス； $(iii-3)$ $(W_I / W_i) \times N_c \times W_i \times 1$ のサイズを有する前記第1学習用修正済み特徴マップを利用して $N_c \times W_I \times 1$ のサイズを有する第2学習用修正済み特徴マップを生成するものの、前記 N_c は前記障害物各々の前記列ごとの位置を特定するために前記トレーニングイメージが分割された行の数であるプロセス；及び $(iii-4)$ 前記第2学習用修正済み特徴マップの前記列ごとに前記 N_c 個のチャネルをノーマライジングするソフトマックス出力プロセスを遂行するプロセス；を含み、前記(iv)プロセスで、前記 $(iii-4)$ プロセスで出力された、前記原本正解イメージに対応する各々のソフトマックス出力と前記原本正解イメージとを参照して前記ロスが算出され、前記(III)プロセスは、(I

40

50

$I I - 1$) $C_j \times W_j \times H_j$ のサイズを有する前記テスト用デコード済み特徴マップを利用して $C_j \cdot H_j \times W_j \times 1$ のサイズを有する第 1 テスト用修正済み特徴マップを生成するものの、前記 C_j は、前記テスト用デコード済み特徴マップのチャンネル数であり、前記 W_j は、前記テスト用デコード済み特徴マップの列のサイズであり、前記 H_j は、前記テスト用デコード済み特徴マップの行のサイズであるプロセス；
 $(I I I - 2)$ 前記 $C_j \cdot H_j \times W_j \times 1$ のサイズを有する前記第 1 テスト用修正済み特徴マップを $(W_j / W_j) \times N_c \times W_j \times 1$ のサイズを有する前記第 1 テスト用修正済み特徴マップへ変更するコンポリューション演算を少なくとも一度遂行するものの、前記 W_j は、前記テストイメージの列のサイズであり、前記 W_j は、前記テスト用デコード済み特徴マップの前記列のサイズであるプロセス；
 $(I I I - 3)$ $(W_j / W_j) \times N_c \times W_j \times 1$ のサイズを有する前記第 1 テスト用修正済み特徴マップを利用して $N_c \times W_j \times 1$ のサイズを有する第 2 テスト用修正済み特徴マップを生成するものの、前記 N_c は前記障害物各々の前記列ごとの位置を特定するために前記テストイメージを分割した行の数であるプロセス；及び
 $(I I I - 4)$ 前記第 2 テスト用修正済み特徴マップの前記列ごとに前記 N_c 個のチャンネルをノーマライジング (normalizing) するソフトマックス出力プロセスを遂行するプロセス；を含むことを特徴とする。

10

【0042】

一例として、前記 $(I I I - 4)$ プロセスで、前記チャンネルのうち、前記列ごとに最も大きい値を有する特定チャンネル各々に前記障害物各々の一番下の部分各々が存在することが決定されれば、前記 N_c 個の行のうち前記行ごとの前記一番下の部分各々の前記推定位置が各々最も大きい値を有し、前記行のうち残りの行の位置はより小さい値を有するように前記セグメンテーション結果が生成されることを特徴とする。

20

【発明の効果】

【0043】

本発明によれば、前記入力イメージを所定の間隔に分けることで生成されたグリッド X 軸上の前記列に前記障害物がどこにあるかを検出し、前記自律走行自動車が走行可能な経路を決定し得る効果がある。

【0044】

また、本発明によれば、前記入力イメージ内の全てのピクセルを検討しなくても、少ない演算量で前記障害物を検出し得るまたの効果がある。

30

【0045】

また、本発明は、低解像度の特徴から高解像度の情報を出力して、演算量が少なくなるまた他の効果がある。

【図面の簡単な説明】

【0046】

【図 1】従来の CNN を用いて一般的なセグメンテーションプロセスを簡略的に示した図面である。

【図 2】本発明に係る障害物検出のための CNN の学習方法を示したフローチャートである。

【図 3】本発明に係る前記障害物検出のための CNN の学習方法を説明するために入力イメージ上の演算プロセスを例示的に示した図面である。

40

【図 4】本発明に係る前記障害物検出のための修正プロセスを簡略的に表した図面である。

【図 5】本発明に係る前記障害物検出のための前記入力イメージ及びこれに対応する原本正解イメージを例示的に示した図面である。

【図 6】本発明に係る障害物検出のための前記 CNN のテスト方法を説明するために前記入力イメージの演算プロセスを例示的に示した図面である。

【図 7 a】従来の障害物検出結果を簡略的に示した図面である。

【図 7 b】本発明に係る障害物検出結果を簡略的に示した図面である。

【図 7 c】本発明に係る障害物検出結果を簡略的に示した図面である。

50

【発明を実施するための形態】**【0047】**

後述する本発明に対する詳細な説明は、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充分詳細に説明される。本発明の多様な実施例は相互異なるが、相互排他的である必要はないことを理解されたい。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係る本発明の精神及び範囲を逸脱せずに他の実施例で具現され得る。また、各々の開示された実施例内の個別構成要素の位置または配置は本発明の精神及び範囲を逸脱せずに変更され得ることを理解されたい。従って、後述する詳細な説明は限定的な意味で捉えようとするものではなく、本発明の範囲は、適切に説明されると、その請求項が主張することと均等なすべての範囲と、併せて添付された請求項によってのみ限定される。図面で類似する参照符号はいくつかの側面にかけて同一か類似する機能を指称する。

10

【0048】

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り、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）でもあり得り、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。

20

【0049】

以下、本発明が属する技術分野で通常の知識を有する者が本発明を容易に実施することができ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の図面を参照して詳細に説明することとする。

【0050】

本発明は、速くて少ない演算により高解像度のイメージから近接障害物を検出できるアルゴリズムを開発して提示された技術である。本発明の技術に係る前記近接障害物の検出方法は、入力イメージから道路と少なくとも一つの障害物の間の境界を見つけることを目標とする。このために、前記入力イメージの行方向を第1方向、列方向を第2方向とした場合、前記第1方向に第1間隔ずつ区画して複数の列を形成し、前記第2方向に第2間隔ずつ区画することで複数の行を形成した結果、グリッドが生成され得る。前記各々の列をこれに該当する前記グリッドの一番低いセルから始まって前記第2方向へ確認することで前記近接障害物が各々存在すると推定される前記各々の列に対して特定の行の情報を用いて前記近接障害物の道路上の位置を検出するものと言えるだろう。また、本発明は、(i)高解像度情報を利用するマルチロス(Multi-loss)学習プロセス及び(ii)低解像度特徴だけを利用するテストプロセスによって演算量を減らし得るようにする。

30

【0051】

図2は、本発明に係る前記近接障害物検出のためのCNNの学習方法を示したフローチャートである。図3は本発明に係る前記近接障害物検出のための前記CNNの前記学習方法を説明するために前記入力イメージの演算プロセスを例示的に示した図面である。

40

【0052】

図2及び図3を参照して、本発明に係る前記近接障害物検出のための前記CNNの前記学習方法を具体的に説明すると次の通りである。

【0053】

本発明に係る前記近接障害物検出プロセスは、少なくとも一つの入力イメージからエンコード済み特徴マップ及びデコード済み特徴マップを生成する段階S01から始まる。前記S01段階では、学習装置がトレーニングイメージとして前記イメージを受信すると、前記学習装置は、第1ないし第nコンボリューションレイヤをもって、前記トレーニング

50

イメージから逐次的に第1エンコード済み特徴マップないし第nエンコード済み特徴マップを各々生成するようにする。ここで、前記第1ないし第nコンボリューションレイヤは前記近接障害物検出用に用いられる前記CNNに含まれる。また、前記近接障害物検出用に用いられる前記CNNは、前記第1ないし第nコンボリューションレイヤに対応する第nないし第1デコンボリューションレイヤを含むが、前記学習装置は前記第nないし第1デコンボリューションレイヤをもって、前記第nエンコード済み特徴マップから逐次的に第nデコード済み特徴マップないし第1デコード済み特徴マップを生成するようにする。

【0054】

例えば、図3を参照すれば、前記近接障害物検出用に用いられる前記CNNは、前記第1コンボリューションレイヤないし第5コンボリューションレイヤ(11ないし15)及び第5デコンボリューションレイヤないし前記第1デコンボリューションレイヤ(16ないし20)を含み得り、前記学習装置は、3ch、640×256サイズの前記トレーニングイメージ100を受信し得る。この入力イメージは、前記第1コンボリューションレイヤ11に入力され、8ch、320×128サイズの前記第1エンコード済み特徴マップ110が生成され、第2コンボリューションレイヤ12に入力され、16ch、160×64サイズの第2エンコード済み特徴マップ120が生成され、第3コンボリューションレイヤ13に入力され、32ch、80×32サイズの第3エンコード済み特徴マップ130が生成され、第4コンボリューションレイヤ14に入力され、64ch、40×16サイズの第4エンコード済み特徴マップ140が生成され、第5コンボリューションレイヤ15に入力され、128ch、20×8サイズの第5エンコード済み特徴マップ150が生成される。

【0055】

このように、前記コンボリューションレイヤは、前記入力イメージまたは特徴マップのチャンネルは増やし、横及び縦のサイズは小さくして、前記エンコード済み特徴マップを生成する機能をする。例えば、前記第2コンボリューションレイヤ12ないし前記第5コンボリューションレイヤ15は、前記入力される特徴マップの前記チャンネルは2倍に増やし、横や縦サイズは各々1/2に減らして前記エンコード済み特徴マップを生成する。

【0056】

一方、前記学習装置は、前記第nコンボリューションレイヤに対応する前記第nデコンボリューションレイヤをもって、前記第nエンコード済み特徴マップの横サイズを所定倍数に拡大し、第nデコード済み特徴マップを生成する。例えば、図3に示した例で、前記学習装置は、前記第5デコンボリューションレイヤ16をもって、128ch、20×8サイズの前記第5エンコード済み特徴マップ150から、64ch、40×8サイズの第5デコード済み特徴マップ160を生成するようにする。

【0057】

一般的にデコンボリューションレイヤは、チャンネル数は減らし、横及び縦サイズは大きくするが、本発明に係る前記第nデコンボリューションレイヤは、前記第nエンコード済み特徴マップのチャンネルを減らして、前記横サイズを所定の倍数(例えば2倍)に大きくするが、前記特徴マップの前記縦サイズは変更させないことができる。その理由は、前述のように、本発明は前記グリッドの前記列の中でどの位置が最も高いスコアを有するかを区別することで充分だからである。すなわち、本発明では、従来のセグメンテーションと異なり、すべてのピクセルを確認する必要がなく、前記縦サイズを大きくする必要もない。本発明で提案している方法は、入力と出力の横の解像度が同一の効果があり、従来の横の解像度が低くなる問題がない。縦の解像度も高ければ更によいだろうが、そうすると多くの演算量が必要であるという問題がある。従って、本発明では、少ない演算量で前記近接障害物の検出をするために、前記横の解像度だけを増加させる方法を提示するものである。このため、前述したように、前記第nデコンボリューションレイヤは、前記第nエンコード済み特徴マップの前記チャンネル数を減らして、前記横サイズのみ所定倍数(例えば2倍)に増加させるが、前記縦サイズは変更させないのである。

【0058】

再度、図3に示したデコーディングプロセスを見ると、前記学習装置は、前記第4デコンボリューションレイヤ17をもって、 $64ch$ 、 40×8 サイズの前記第5デコード済み特徴マップ160から $32ch$ 、 80×16 サイズの前記第4デコード済み特徴マップ170を生成するようにし、前記第3デコンボリューションレイヤ18をもって $32ch$ 、 80×16 サイズの前記第4デコード済み特徴マップ170から $16ch$ 、 160×32 サイズの前記第3デコード済み特徴マップ180を生成するようにし、前記第2デコンボリューションレイヤ19をもって $16ch$ 、 160×32 サイズの前記第3デコード済み特徴マップ180から $8ch$ 、 320×64 サイズの前記第2デコード済み特徴マップ190を生成するようにし、前記第1デコンボリューションレイヤ20をもって $8ch$ 、 320×64 サイズの前記第2デコード済み特徴マップ190から $4ch$ 、 640×128 サイズの前記第1デコード済み特徴マップ200を生成するようにする。

10

【0059】

このように、前記デコンボリューションレイヤは、入力される特徴マップのチャンネルを減らし、前記横及び縦のサイズは大きくしてデコード済み特徴マップを生成する機能をする。例えば、前記第4デコンボリューションレイヤ17ないし前記第1デコンボリューションレイヤ20は、前記チャンネル数を $1/2$ に減らし、前記入力される特徴マップの前記横や縦サイズは各々2倍にして前記デコード済み特徴マップを生成する。

【0060】

一方、前記コンボリューションレイヤは、コンボリューション、マックスプーリング(max pooling)、ReLUのうち少なくとも一つの演算を遂行し、前記デコンボリューションレイヤは、デコンボリューション及びReLUのうち少なくとも一つの演算を遂行し得る。

20

【0061】

その後、図2を参照すると、S02段階において前記学習装置は、 $C_i \times W_i \times H_i$ サイズを有する前記デコード済み特徴マップを利用して $C_i H_i \times W_i \times 1$ サイズを有する第1修正済み特徴マップを生成し得り、この際 C_i は、前記チャンネルの数を意味し、前記 W_i は、前記列のサイズ、前記 H_i は、前記デコード済み特徴マップの前記行のサイズを意味する。

【0062】

すなわち、本発明に係る修正(reshaping)プロセスにおいて、前記少なくとも一つのデコード済み特徴マップを前記第1方向に第1間隔で区画し、前記第2方向に第2間隔で区画することで、複数の列と複数の行を有する前記グリッドの各セルが生成されるとした場合、前記学習装置は、前記列ごとに前記各々の行の特徴各々をチャンネル方向へコンカチネート(concatenate)して、少なくとも一つの修正済み特徴マップを生成する。

30

【0063】

図4は、本発明に係る前記近接障害物検出のための前記修正プロセスを簡略的に表した図面である。

【0064】

図4を参照すると、前記修正プロセスで、図面符号410で表示された特徴マップに示したように、デコード済み特徴マップが行に分かれた後、図面符号420で表示された特徴マップに示したように、前記列ごとに前記各々の行の前記特徴各々が前記チャンネルの方向にコンカチネートされる。これによって、 $(C \times W \times H)$ サイズの特徴マップは $(C * H) \times W \times 1$ サイズの特徴マップに変換される。

40

【0065】

図4の例で、前記図面符号410で表示された特徴マップ上で、太線で描かれた四角形各々は、前記デコード済み特徴マップの第1列に対応する各行の特徴各々を示す。仮に、前記図面符号410で表示された特徴マップが8つの行を有するなら、前記図面符号420で表示された特徴マップは、前記チャンネル数の8倍に増えたチャンネル数と、前記図面符号410で表示された特徴マップの高さの $1/8$ の高さを有し得る。

50

【0066】

図3の例で、 $64ch$ 、 40×8 サイズの前記第5デコード済み特徴マップ160は、第1修正プロセス($reshape5-1$)によって、 $64 * 8ch$ 、 40×1 サイズの前記第1修正済み特徴マップ161に変換され、 $32ch$ 、 80×16 サイズの前記第4デコード済み特徴マップ170は、第1修正プロセス($reshape4-1$)によって、 $32 * 16ch$ 、 80×1 サイズの前記第1修正済み特徴マップ171に変換され、 $16ch$ 、 160×32 サイズの前記第3デコード済み特徴マップ180は、第1修正プロセス($reshape3-1$)によって、 $16 * 32ch$ 、 160×1 サイズの前記第1修正済み特徴マップ181に変換され、 $8ch$ 、 320×64 サイズの前記第2デコード済み特徴マップ190は、第1修正プロセス($reshape2-1$)によって、 $8 * 64ch$ 、 320×1 サイズの前記第1修正済み特徴マップ191に変換され、 $4ch$ 、 640×128 サイズの前記第1デコード済み特徴マップ200は第1修正プロセス($reshape1-1$)によって、 $4 * 128ch$ 、 640×1 サイズの前記第1修正済み特徴マップ201に変換される。

10

【0067】

参考までに、図3では、すべてのデコード済み特徴マップに対して前記第1修正プロセスを遂行するものと説明したが、すべてのデコード済み特徴マップについて修正プロセスを遂行する必要はなく、前記デコード済み特徴マップの一部に対してのみ修正プロセスを遂行しても充分である。

20

【0068】

その後、S03段階では、 $C_i H_i \times W_i \times 1$ サイズ有する前記第1修正済み特徴マップを $(W_I / W_i) \times N_c \times W_i \times 1$ サイズの前記第1修正済み特徴マップに変更するコンポリューション演算を遂行し得る。ここで、前記 W_I は、前記トレーニングのイメージの列サイズであり、前記 W_i は、前記デコード済み特徴マップの列サイズである。このコンポリューション演算は、 1×1 コンポリューションであり、これは、横、縦は1マスのみ含まれるが、すべてのチャンネルにわたっている前記グリッド内のセルを被演算子($operand$)とする演算であり、各第1修正済み特徴マップの N_c 個の列の各々において前記近接障害物の下段ライン各々がどこに位置するかを知るための過程であり、前記 N_c は、前記入力イメージの前記第2方向を所定の大きさに分割した数である。すでに前記第1修正プロセスで前記デコード済み特徴マップのすべての前記列方向の情報を、同時に演算できるように前記チャンネルに統合した状態であるため、前記コンポリューション演算を通じて前記チャンネルの情報をすべて確認し、各列ごとにどの位置に前記近接障害物の前記下段ライン各々が位置するかどうかを確認し得る。

30

【0069】

もちろん、前記第5デコード済み特徴マップ160の例のように、修正プロセスなしで 8×1 コンポリューションが遂行される場合、前記第1修正演算と前記 1×1 コンポリューション演算を一度で行い得る。つまり、特定の特徴マップの高さが N である場合、 $N \times 1$ コンポリューションを利用し得る。しかしながら、一般的にハードウェア上、 1×1 コンポリューション演算は素早くに計算できるが、あまり利用されない形である 8×1 カーネルまたは $N \times 1$ カーネルは演算速度が著しく遅いため、前記修正プロセス演算と前記 1×1 コンポリューション演算を分けたほうが効果的である。

40

【0070】

前記 1×1 コンポリューション演算結果を参照すれば、前記入力された特徴マップの前記列サイズが W_i 、前記元のイメージの前記列サイズが W_I とした場合、 $(W_I / W_i) \times N_c$ くらいのチャンネルを有するように前記入力特徴マップが変換される。

【0071】

図3の例で、 $64 * 8ch$ 、 40×1 サイズの前記第5デコード済み特徴マップの第1修正済み特徴マップ161は、前記 1×1 コンポリューション演算によって $N_c * 16ch$ 、 40×1 サイズの第1修正済み特徴マップ162に変更され、 $32 * 16ch$ 、 80×1 サイズの前記第4デコード済み特徴マップの第1修正済み特徴マップ171は、前記

50

1 × 1 コンボリユーション演算によって $N_c \times 8 \text{ ch}$ 、 80×1 サイズの第 1 修正済み特徴マップ 172 に変更され、 $16 \times 32 \text{ ch}$ 、 160×1 サイズの前記第 3 デコード済み特徴マップの第 1 修正済み特徴マップ (181) は、前記 1 × 1 コンボリユーション演算によって $N_c \times 4 \text{ ch}$ 、 160×1 サイズの第 1 修正済み特徴マップ 182 に変更され、 $8 \times 64 \text{ ch}$ 、 320×1 サイズの前記第 2 デコード済み特徴マップの第 1 修正済み特徴マップ 191 は、前記 1 × 1 コンボリユーション演算によって $N_c \times 2 \text{ ch}$ 、 320×1 サイズの第 1 修正済み特徴マップ 192 に変更され、 $4 \times 128 \text{ ch}$ 、 640×1 サイズの前記第 1 デコード済み特徴マップの第 1 修正済み特徴マップ 201 は、前記 1 × 1 コンボリユーション演算によって $N_c \times 1 \text{ ch}$ 、 640×1 サイズの第 1 修正済み特徴マップ 202 に変更される。

10

【0072】

再度図 2 を参照すれば、S04 段階では、 $(W_I / W_i) \times N_c \times W_i \times 1$ サイズの前記第 1 修正済み特徴マップは $N_c \times W_I \times 1$ サイズを有する第 2 修正済み特徴マップに修正され得る。ここで前記 N_c 個は、前記近接障害物の各々の前記下段ライン各々が前記列ごとにどこに位置するかを特定するために前記入力イメージの前記第 2 方向に分割した前記行の数である。

【0073】

そして S05 段階では、前記第 2 修正済み特徴マップの前記列ごとに前記 N_c 個のチャンネルに対応する各々の値をノーマライジング (normalizing) するソフトマックス演算が遂行され、S06 段階では、前記第 2 方向に沿って前記各々の列をこれに対応する一番下のセルから確認することで前記近接障害物が存在すると推定される前記特定の行の各々から前記近接障害物各々の前記下段ライン各々の推定位置各々を示す前記入力イメージ内の前記列ごとのセグメンテーション結果が生成される。

20

【0074】

前記第 2 修正プロセス S04 で、前記出力された $(W_I / W_i) \times N_c \times W_i \times 1$ サイズの特徴マップは、データは固定されたままの形だけが変化して $N_c \times W_I \times 1$ サイズの形態に変換され得る。そして前記ソフトマックスプロセス S05 で各列ごとに前記 N_c 個のチャンネルの前記値を 0 ~ 1 の間の値へノーマライジングさせ、前記ノーマライジングされた値を参照して、前記列ごとにそれに対応する最も大きい値を有する特定のチャンネル各々を探して前記近接障害物各々の前記下段ライン各々の前記列ごとの位置を推定し得る。

30

【0075】

従って、前記 1 × 1 コンボリユーション演算 S03 と前記修正演算 S04 によって、前記列ごとに前記行の中から前記近接障害物各々の前記下段ライン各々の前記推定位置各々は、各々に対応する最も大きい値を有し得り、前記列ごとの前記行の中のうちの残りの行はそれより小さい値を有するように特徴マップが生成され得る。前記ソフトマックス演算 S05 は、前記入力イメージの列ごとに前記 N_c 個の値の中で最も大きい値を見つけ、その位置を出力して前記近接障害物の前記位置各々を探し出すために利用される。そして、前記ノーマライジングされた値を参照して前記列ごとに前記チャンネルの値のうち大きい値を有する特定チャンネル各々に前記近接障害物各々の前記下段ラインが位置すると推定されると、前記 N_c 行のうち、前記列ごとの前記行の中から前記近接障害物各々の前記下段ライン各々の推定位置各々は、対応する最も大きい値を有し、前記列ごとの前記行のうち、残りの行はそれより小さい値を有するようにするセグメンテーション結果が生成され得るようになる。

40

【0076】

このプロセスを理解するためには、最終結果 (前記ソフトマックス演算の結果) の形態についての理解が必要である。前記 CNN の前記学習方法から期待される出力は、前記入力イメージで前記列ごとに前記 N_c 個の行のうち最大の値を有する各々の行を前記近接障害物の位置として探し出すことである。このためには、列ごとに N_c 個のスコアが必要である。例えば、前記入力イメージ内の前記列の個数 (つまり前記入力イメージの前記幅)

50

が640個(つまり640個のピクセルまたは640個の列)ならば、 N_c (チャンネル) × 640(幅) × 1(高さ)サイズのスコアマップとして出力されるべきである。

【0077】

前記出力として N_c (チャンネル) × 640(幅) × 1(高さ)サイズの前記スコアマップを生成する前記プロセスを見ると次のとおりである。例えば、前記第1修正(reshape-1)プロセスによって512(64*8)(チャンネル) × 40(幅) × 1(高さ)サイズの前記第5デコード済み特徴マップの前記第1修正済み特徴マップ161が生成された場合、この第1修正済み特徴マップの列は、前記入力イメージの列(640個)の1/16の40個だけである。従って、この場合、16回の前記1×1コンボリューション演算で N_c 個のスコアマップを16回出力すれば解決し得る。従って、図3でCONV__OUT5から出た前記出力162サイズは($N_c * 16$)(チャンネル) × 40(幅) × 1(高さ)になるようにデザインされるべきだ。そして、($N_c * 16$)(チャンネル) × 40(幅) × 1(高さ)サイズの前記スコアマップを N_c (チャンネル) × 640(幅) × 1(高さ)サイズの前記スコアマップに変換するために、前記第2修正プロセス(reshape-2)が必要なのである。

10

【0078】

図3に示した例を参照に、 $N_c * 16 ch$ 、40 × 1サイズの前記特徴マップ162は、前記1×1コンボリューションであるCONV__OUT5によって $N_c ch$ 、640 × 1サイズの前記特徴マップ163に変換され、前記ソフトマックス演算によって640個の列ごとの前記 N_c 個の行のうち、前記列ごとに前記行の中から前記近接障害物各々の前記下段ライン各々の前記推定位置各々に対応する最も大きい値を有し、前記列ごとに前記行の中の前記残りの位置はより小さい値を有するように出力164が生成される。そして $N_c * 8 ch$ 、80 × 1サイズの前記特徴マップ172、 $N_c * 4 ch$ 、160 × 1サイズの前記特徴マップ182、 $N_c * 2 ch$ 、320 × 1サイズの前記特徴マップ192、 $N_c ch$ 、640 × 1サイズの前記特徴マップ202は、CONV__OUT4ないしCONV__OUT1の前記1×1コンボリューション演算各々によって $N_c ch$ 、640 × 1サイズの前記特徴マップ173、183、193、203へ各々変換され、前記ソフトマックス演算によって640個の列ごとに前記 N_c 個の行のうち、前記近接障害物各々の前記下段ライン各々の前記推定位置各々が存在する前記列ごとにこれに該当する最も大きい値を有し、前記列ごとの前記行の中の残りの位置はより小さい値を有するように各々の出力174、184、194、204が生成される。

20

30

【0079】

つまり、前記学習装置が、前記少なくとも一つのデコード済み特徴マップにおいて、前記推定位置各々を示す前記セグメンテーション結果を生成し得り、前記各々の列をそれに対応する一番下のセルから前記第2方向に(つまり、上がりながら)確認することで、前記近接障害物各々の前記下段ライン各々の前記推定位置各々が推定される。前記学習装置は、前記修正済み特徴マップの前記列ごとにコンカチネートされたチャンネル上の前記近接障害物各々の前記下段ライン各々の前記推定位置各々を確認することにより、前記セグメンテーション結果を生成し得る。

【0080】

40

再び図2を参照すると、S07段階では、前記セグメンテーション結果とそれに対応する少なくとも一つの原本正解イメージを参照して少なくとも一つのロスが算出され得り、S08段階では、前記ロスをバックプロパゲーション(backpropagation)して、前記CNNのパラメータが学習されたり、最適化され得る。

【0081】

この際、前記原本正解イメージは、前記各々の列に対してそれに対応する一番下のセルから前記第2方向に(つまり、上がりながら)確認したとき、前記近接障害物各々が実際に位置する原本正解位置各々に対応する各々の列が表示されたイメージである。

【0082】

図5は、本発明によって前記近接障害物検出のための前記入力イメージ及びこれに対応

50

する原本正解イメージを例示的に示す図面である。図5を参照すれば、前記入力イメージで列ごと(640個のピクセルを前記第1間隔で区切ってできた列ごとまたは前記640個のピクセルごと)に前記下段から上段まで確認するとき、近接障害物を前記近接障害物に指定することで、前記原本正解イメージが生成され得る。前記原本正解イメージが前記入力イメージで前記列ごとに前記近接障害物各々の前記下段ライン各々が、実際に前記 N_c 個の行の中に位置する行を示す情報を含み、前記セグメンテーション結果が前記入力イメージで前記列ごとに前記近接障害物各々の前記下段ライン各々が前記 N_c 個の行の中のどこに位置するかを推定する情報を含むため、前記デコード済み特徴マップを修正したすべての前記修正済み特徴マップ164、174、184、194、204は、前記 N_c のチャンネルを有するように生成されるのである。

10

【0083】

前記S07段階での前記ロスはクロスエントロピーロス(Cross-entropy loss)であり得る。前記ロスは、前記CNNの前記パラメータを学習したり、最適化したりするためにバックプロパゲーションされる。図3の例では、5つのデコード済み特徴マップを通じて出力を算出するため5つのロスが算出されるが、前記5つのデコード済み特徴マップのうち、少なくとも一部からの前記ロスと、前記出力を参照して前記バックプロパゲーションを遂行し得る。特に、前記第1コンボリューションレイヤ20から出力された前記第1デコード済み特徴マップを参照して算出した前記ロスを利用することが好ましいが、必須ではない。

【0084】

20

前記のようなプロセスを経て、前記CNNの前記パラメータが学習された状態で、前記学習されたパラメータを有している前記CNNを利用したテスト装置は、前記入力イメージとしての少なくとも一つのテストイメージから近接障害物を検出し得る。

【0085】

図6は、本発明に係る前記近接障害物検出のための前記CNNのテスト方法を説明するため、前記入力イメージに対する演算の過程を例示的に示す。図6を参照すれば、図3の前記学習装置と異なって、一つの出力だけを生成すれば充分で、前記第5デコード済み特徴マップを利用し、直ちに前記出力を生成し得るため、前記第4デコンボリューションレイヤないし前記第1デコンボリューションレイヤは省略しても構わない。もう一つの例として、前記省略されたデコンボリューションレイヤの一部を含んでも構わないと言える。

30

【0086】

具体的な前記プロセスは、図3で説明した内容と類似するため、図6の前記近接障害物検出過程を簡略に説明すると、次の通りだ。まず、前記テスト装置が、前記テストイメージ100を受信し、前記第1ないし前記第 n コンボリューションレイヤ(11ないし15)をもって前記テストイメージ100から逐次的にテスト用第1エンコード済み特徴マップないしテスト用第 n エンコード済み特徴マップ110、120、130、140、150を各々生成するようにし得る。そして前記テスト装置は少なくとも一つのデコンボリューションレイヤ16をもって前記テスト用第 n エンコード済み特徴マップ150からテスト用デコード済み特徴マップ160を生成するようにし得る。そして、前記テスト用デコード済み特徴マップ160から、前記グリッドを参照にし、前記第2方向に羅列された前記列ごとの前記各々の行の特徴各々を前記チャンネル方向へコンカチネートしてテスト用修正済み特徴マップ161を生成し得る。そして、前記 1×1 コンボリューション演算および追加修正プロセスを経て、チャンネルが変更されたテスト用特徴マップ162と当該チャンネルを N_c 個に合わせて、該当列の前記横軸の個数を前記テストイメージの前記横軸の個数に合わせたテスト用特徴マップ163が生成される。そして、前記テスト装置は、前記テスト用修正済み特徴マップの前記列ごとにコンカチネートされたチャンネル上で前記近接障害物各々の前記下段ライン各々の前記列ごとの前記行の中の前記推定位置各々を確認することにより、テスト用セグメンテーション結果164を生成して、前記近接障害物を検出し得る。

40

50

【 0 0 8 7 】

図 7 a は、前記従来の障害物検出結果を簡略に示す図面であり、図 7 b 及び図 7 c は本発明に係る前記障害物の検出結果を簡略に示す図面である。

【 0 0 8 8 】

図 7 a は、前記従来の検出方法によって前記近接障害物を検出した例を示すが、すべてのピクセルを見て前記ピクセルが前記道路に該当するピクセルであるか否かを区別しなければならないので、その結果、演算量が多くなってしまふ。しかし、図 7 b 及び図 7 c に示した本発明に係る方法によれば、前記近接障害物各々の前記下段ライン各々の位置（黒線部）を推測するためにイメージの下段から上段へ所定の個数（例えば、 N_c 個）の前記グリッドセルを確認することで、前記近接障害物を検出し、その結果として少ない演算によって速くて高解像度の前記近接障害物を検出し得る。

10

【 0 0 8 9 】

さらに、前記従来の技術は処理時間の関係で、前記障害物検出結果の前記横の解像度が前記入力イメージの前記解像度より低いという問題があるが、本発明で新たに提案する方法は、前記入力イメージの前記出力結果の横の解像度が同一であるためこの問題を解決し得る。

【 0 0 9 0 】

また、本発明では、マルチロスを用いた学習時に高解像度情報を利用し、実際のテストでは低解像度特徴マップだけでも高解像度の結果が出力され得る。これにより、前記低解像度の特徴マップから高解像度情報を出力し得り、演算量も少なくして処理速度も早くなり得る。

20

【 0 0 9 1 】

技術分野の通常の技術者に理解されるものとして、前記で説明されたイメージ、例えばトレーニングイメージ、テストイメージといったイメージデータの送受信が学習装置及びテスト装置の各通信部によって行われ得り、特徴マップと演算を遂行するためのデータが前記学習装置及び前記テスト装置のプロセッサ（及び/またはメモリ）によって保有/維持でき得り、コンボリユーション演算、デコンボリユーション演算、ロス値の演算過程が主に学習装置及びテスト装置のプロセッサにより遂行され得るが、本発明はこれに限定されるものではない。

【 0 0 9 2 】

以上で説明された本発明に係る実施例は、多様なコンピュータ構成要素を通じて遂行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で判読可能な記録媒体に記録され得る。前記コンピュータで判読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または組み合わせて含まれ得る。前記コンピュータ判読可能な記録媒体に記録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て構成されたものか、コンピュータソフトウェア分野の当業者に公知となって使用可能なものでもよい。コンピュータで判読可能な記録媒体の例には、ハードディスク、フロッピィ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（*f l o p t i c a l d i s k*）のような磁気-光媒体（*m a g n e t o - o p t i c a l m e d i a*）、およびROM、RAM、フラッシュメモリなどといったプログラム命令語を保存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、インタプリタなどを用いてコンピュータによって実行され得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係る処理を遂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ることがあり、その逆も同様である。

30

40

【 0 0 9 3 】

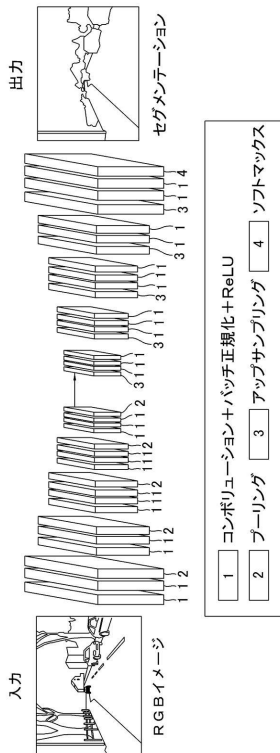
以上、本発明が具体的な構成要素などのような特定事項と限定された実施例及び図面によって説明されたが、これは本発明のより全般的な理解を助けるために提供されたものであるに過ぎず、本発明が前記実施例に限られるものではなく、本発明が属する技術分野において通常の知識を有する者であれば係る記載から多様な修正及び変形が行われ得る。

50

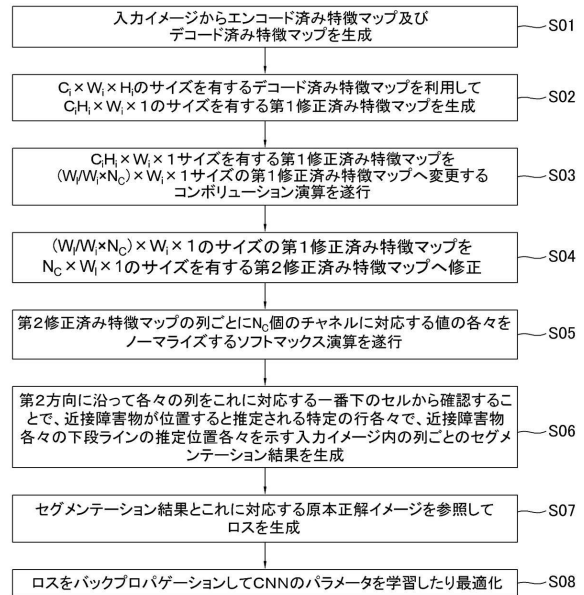
【 0 0 9 4 】

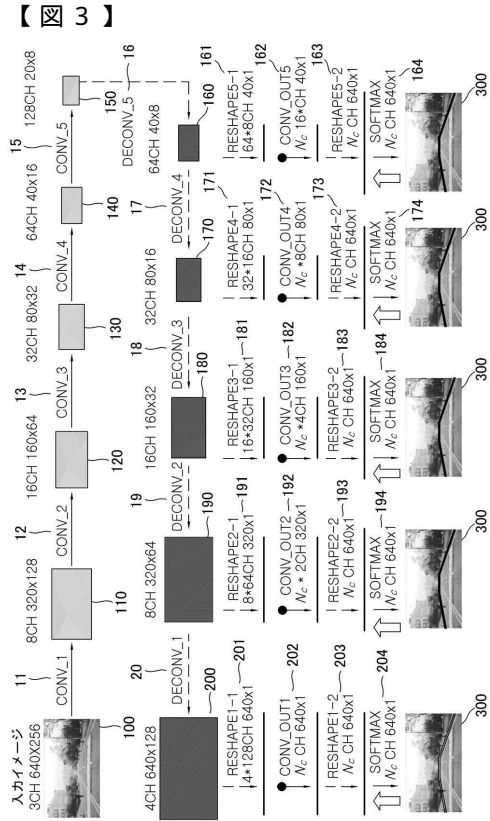
従って、本発明の思想は前記説明された実施例に極限されて定められてはならず、後述する特許請求の範囲だけでなく、本特許請求の範囲と均等または等価的に変形されたものすべては、本発明の思想の範囲に属するといえる。

【 図 1 】

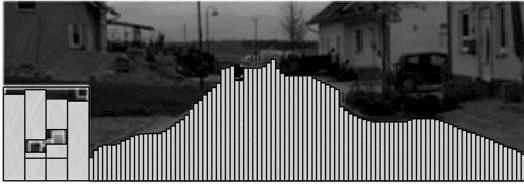


【 図 2 】





【図 7 a】



【図 7 b】



【図 7 c】



フロントページの続き

- (72)発明者 金鎔重
大韓民国慶尚北道浦項市南区チョンナム - 口67、4 - 4427
- (72)発明者 金寅洙
大韓民国慶尚北道浦項市南区チョンナム - 口67、4 - 4427
- (72)発明者 金鶴京
大韓民国慶尚北道浦項市南区チョンナム - 口67、4 - 4427
- (72)発明者 南雲鉉
大韓民国慶尚北道浦項市南区延日邑ユガン - ギル9ボン - ギル57、203 - 803
- (72)発明者 夫碩 くん
大韓民国京畿道安養市萬安区萬安路55ボン - ギル20 - 9、ビー02
- (72)発明者 成明哲
大韓民国慶尚北道浦項市北区張良路174ボン - ギル13
- (72)発明者 呂東勳
大韓民国慶尚北道浦項市南区チョンナム - 口67、4 - 4427
- (72)発明者 柳宇宙
大韓民国慶尚北道浦項市南区チョンナム - 口67、4 - 4427
- (72)発明者 張泰雄
大韓民国ソウル特別市江南区彦州路113 - ギル18 - 5
- (72)発明者 鄭景中
大韓民国慶尚北道浦項市南区芝谷路294、ヒュンダイグリーンアパート232 - 501
- (72)発明者 諸泓模
大韓民国慶尚北道浦項市南区芝谷路20、スンニアパート5 - 1805
- (72)発明者 趙浩辰
大韓民国慶尚北道浦項市南区チョンナム - 口67、4 - 4427

審査官 粕谷 満成

- (56)参考文献 特表2020 - 510463 (JP, A)
米国特許出願公開第2019 / 0286921 (US, A1)
Hyeonwoo Noh ; Seunghoon Hong ; Bohyung Han, Learning Deconvolution Network for Semantic Segmentation, 2015 IEEE International Conference on Computer Vision (ICCV), IEEE, 2015年12月7日, pp,1520-1528, URL, <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7410535>
藤吉 弘巨 他, 深層学習による画像認識, 日本ロボット学会誌, 一般社団法人日本ロボット学会, 2017年4月15日, 第35巻 第3号, pp.188-185

(58)調査した分野(Int.Cl., DB名)
G06T 7/00

(54)【発明の名称】障害物を検出する学習方法及び学習装置そしてこれを利用したテスト方法及びテスト装置 {LEARNING METHOD, LEARNING DEVICE FOR DETECTING OBSTACLES AND TESTING METHOD, TESTING DEVICE USING THE SAME}