

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第6599986号
(P6599986)

(45) 発行日 令和1年10月30日(2019.10.30)

(24) 登録日 令和1年10月11日(2019.10.11)

(51) Int.Cl.		F I			
G06T 7/00	(2017.01)	G06T	7/00	350C	
G06F 16/55	(2019.01)	G06F	16/55		
G06N 3/04	(2006.01)	G06N	3/04	154	
G06N 3/08	(2006.01)	G06N	3/08	140	
G06N 20/00	(2019.01)	G06N	20/00	130	

請求項の数 7 (全 17 頁)

(21) 出願番号	特願2017-526087 (P2017-526087)	(73) 特許権者	504080663
(86) (22) 出願日	平成27年10月16日(2015.10.16)		エヌイーシー ラボラトリーズ アメリカ
(65) 公表番号	特表2018-503161 (P2018-503161A)		インク
(43) 公表日	平成30年2月1日(2018.2.1)		NEC Laboratories Am
(86) 国際出願番号	PCT/US2015/055943		erica, Inc.
(87) 国際公開番号	W02016/077027		アメリカ合衆国 08540 ニュージャ
(87) 国際公開日	平成28年5月19日(2016.5.19)		ージー州 プリンストン スイート 20
審査請求日	平成30年7月13日(2018.7.13)		0 インディペンデンス ウェイ 4
(31) 優先権主張番号	62/079,316	(74) 代理人	100123788
(32) 優先日	平成26年11月13日(2014.11.13)		弁理士 宮崎 昭夫
(33) 優先権主張国・地域又は機関	米国 (US)	(74) 代理人	100127454
(31) 優先権主張番号	14/884,600		弁理士 緒方 雅昭
(32) 優先日	平成27年10月15日(2015.10.15)		
(33) 優先権主張国・地域又は機関	米国 (US)		

最終頁に続く

(54) 【発明の名称】 きめの細かい画像分類のためのハイパークラス拡張化および正則化深層学習

(57) 【特許請求の範囲】

【請求項 1】

学習機械を訓練する方法であって、

1 つ以上のハイパークラスによって注釈が付けられたラベル付きデータを用いて訓練に用いるデータを拡張することと、

前記拡張によって得られた拡張されたデータの前記ラベル付きデータに対して、きめの細かい分類およびハイパークラス分類の 2 つのタスクを含み、該 2 つのタスクが同じ特徴層を共有する多重タスク深層学習を実行することと、

きめの細かいクラスと前記ハイパークラスとの間の 1 つ以上の関係に基づいて前記多重タスク深層学習における正則化を適用することと、を有し、

前記 1 つ以上のハイパークラスが、1 つは、一組のきめの細かいクラスを包含するスーパークラスであり、もう 1 つは、大きいクラス内分散を説明する、自動車の異なる視点上の要因クラスと呼ばれている、2 つの共通のハイパークラスを有し、

きめの細かいデータ中の注釈が付けられたハイパークラスを識別することと、

前記識別したハイパークラスを使用して外部ソースから多数のハイパークラスラベル付き画像を獲得することと、をさらに有し、

前記拡張のために前記多数のハイパークラスラベル付き画像が用いられる、ことを特徴する方法。

【請求項 2】

前記外部ソースが画像探索エンジンを含む、請求項 1 に記載の方法。

【請求項 3】

前記きめの細かい認識とハイパークラス認識との間の正則化からの学習モデルエンジンを適用することを有する、請求項 1 に記載の方法。

【請求項 4】

学習される特徴の汎化能力を改善すべく補助画像を利用するためにデータ拡張を行うことを有する、請求項 1 に記載の方法。

【請求項 5】

関係を「有する」を捕捉するためにハイパークラスを適用することを有する、請求項 1 に記載の方法。

【請求項 6】

クラス内分散またはポーズ分散を説明するために前記ハイパークラスを適用することを有する、請求項 5 に記載の方法。

【請求項 7】

データの 2 つのソースとタスクに対応する 2 つの損失関数とを有するミニバッチ確率的勾配降下を使用する誤差逆伝播法によって深層 CNN を訓練することを有し、確率的勾配を決定するためにミニバッチにおける画像をサンプリングすることをさらに有する、請求項 1 に記載の方法。

【発明の詳細な説明】

【技術分野】

【0001】

本出願は、2014年11月13日に提出した仮出願 62/079316、および 2015年10月15日に提出した実用出願 14/884600 の優先権を主張するものであり、これらの内容は、参照によって組み込まれている。

【0002】

本出願は、きめの細かい画像分類のためのハイパークラス拡張化および正則化深層学習に関する。

【背景技術】

【0003】

深層畳み込みニューラルネットワーク (CNN) は、大規模な包括的对象認識において大成功しているが、深層畳み込みニューラルネットワークは、きめの細かい画像分類 (FGIC) においてとても成功している。包括的对象認識と比較すると、(i) 多数のきめの細かいラベル付きデータを取得するのに費用が高くつき (通常、特定分野の専門的技術が必要である)、(ii) 大きいクラス内分散および小さいクラス間分散のため、FGIC には課題がある。少しの訓練データで画像認識用の深層 CNN を使用する従来のシステムは、大規模外部データセット (例えば ImageNet) 上での深層 CNN を事前学習し、特定の分類タスクに適合するための小規模ターゲットデータ上でそれを微調整することを含む単純な戦略を採用している。しかしながら、包括的なデータセットから学習される特徴は、特定の FGIC タスクに適していないことがあり、したがってその性能を制限することがある。

【発明の概要】

【0004】

1 つ以上のハイパークラスによって注釈が付けられたラベル付きデータを使用してきめの細かい画像認識からのデータを拡張すること、多重タスク深層学習を実施すること、きめの細かい分類およびハイパークラス分類を同じ特徴層を共有し、かつ、学習できるようにすること、および、多重タスク深層学習における正則化を適用して、きめの細かいクラスとハイパークラスとの間の 1 つ以上の関係を利用することによって、学習機械を訓練するためのシステムおよび方法が開示される。

【0005】

好ましい実施態様の利点には、以下のうちの 1 つ以上を含み得る。システムは、2 つのタスク (きめの細かい分類およびハイパークラス分類) が同じ特徴層を共有し、かつ、学

10

20

30

40

50

習できる、多重タスク深層学習を提供する。多重タスク深層学習における正則化技術は、分類器レベルで学習プロセスの明確なガイダンスを提供する、きめの細かいクラスとハイパークラスとの間の関係を利用している。クラス内分散を説明する要因クラスを利用する場合、本出願人らの学習モデルエンジンは、大きいクラス内分散の問題を軽減し、汎化能力(generalization performance)を改善することができる。

【図面の簡単な説明】

【0006】

【図1】図1Aおよび1Bは、深層CNNを学習するための体系的なフレームワークを有する画像分類器を示す図である。

【図2】図2A～2Bは、ハイパークラスときめの細かいクラスとの間の2つのタイプの関係を示す図である。

【図3】図3は、図1A～1Bの画像分類器を有する自律運転システムを示す図である。

【発明を実施するための形態】

【0007】

図1Aおよび1Bは、深層CNNを学習するための体系的なフレームワークを有する画像分類器を示す。システムは、(i)きめの細かいデータに固有の注釈付きハイパークラスを容易に識別し、容易に利用することができる外部ソース(例えば画像探索エンジン)から多数のハイパークラスラベル付き画像を取得し、問題を多重タスク学習に定式化すること、(ii)きめの細かい認識モデルエンジンとハイパークラス認識モデルエンジンとの間の正則化を利用することによる学習モデルエンジン、といった2つの新しい観点から、分類課題に対処する。

【0008】

図1A～1Bは、2種類のハイパークラスを示す。図1Aは、例示的ハイパークラス拡張化深層CNNを示し、一方、図1Bは、例示的ハイパークラス拡張化および正則化深層CNNを示す。システムは、FGIC用の深層CNNを学習する課題に明確に取り組むための、原則に基づいた手法を提供する。本出願人らのシステムは、データ不足問題に対処するためのタスク特化データ拡張手法を提供する。本出願人らは、きめの細かいデータの固有属性であるいくつかのハイパークラスによって注釈が付けられた、容易に利用することができるデータを使用して、きめの細かい画像認識のデータを拡張する。本出願人らは、一方が、一組のきめの細かいクラスを包含するスーパークラスであり、もう一方が、大きいクラス内分散を説明する要因クラスと呼ばれている(例えば自動車の異なる視点)、2つの共通のタイプのハイパークラスを使用する。次に、本出願人らは、問題を多重タスク深層学習の中に定式化し、2つのタスク(きめの細かい分類およびハイパークラス分類)が同じ特徴層を共有し、かつ、学習できることを可能にする。多重タスク深層学習における正則化技術は、学習プロセスに対する明確なガイダンスを分類器レベルで提供する、きめの細かいクラスとハイパークラスとの間の関係を利用する。クラス内分散を説明する要因クラスを利用する場合、開示した学習モデルエンジンは、大きいクラス内分散の問題を軽減し、汎化能力を改善することができる。本出願人らは、本出願人らの新しいフレームワークをハイパークラス拡張化および正則化深層学習と呼んでいる。

【0009】

図1A～1Bのハイパークラス拡張化および正則化深層学習システムでは、FGICの第1の課題は、きめの細かいラベルを取得するのに費用が高くつき、集中的な労働および特定分野の専門的技術を必要とすることである。したがって、ラベル付き訓練は、通常、過剰適合を伴うことなく深層CNNを訓練するには十分に大きくない。第2の課題は、大きいクラス内分散対小さいクラス間分散である。第1の課題に対処するために、本出願人らは、データ拡張方法を使用する。この着想は、きめの細かいデータの固有属性であり、はるかに容易に注釈が付けられ得る、いくつかのハイパークラスによってラベルが付けられた多数の補助画像を使用して、きめの細かいデータを拡張することである。第2の課題に対処するために、本出願人らは、拡張データを利用する深層CNNモデルエンジンを使用する。

10

20

30

40

50

として視覚的に類似した画像を見つけることができる。本出願人らは、B a i d uによって検索される画像は、ビュー予測により適しており、一方、G o o g l e画像探索は、自動車を認識し、同じタイプの自動車を有する画像を返すことを試行することを見出した。本出願人らの実験では、本出願人らは、B a i d uから検索される画像を本出願人らの拡張データとして使用する。

【 0 0 1 4 】

次に、ハイパークラス正則化学習モデルエンジンについて説明する。本出願人らのモデルエンジンの詳細を説明する前に、本出願人らは、最初に、本書類全体を通して使用されているいくつかの表記法および用語を紹介する。

【 0 0 1 5 】

画像

【 0 0 1 6 】

【数 1】

$$\mathbf{x}_i^t$$

のきめの細かいクラスラベル（例えば自動車の製造、モデルおよび年）を示す

【 0 0 1 7 】

【数 2】

$$y_i^t \in \{1, \dots, C\}$$

で

【 0 0 1 8 】

【数 3】

$$D_t = \{(\mathbf{x}_1^t, y_1^t), \dots, (\mathbf{x}_n^t, y_n^t)\}$$

を一組のきめの細かい訓練画像とし、また、

【 0 0 1 9 】

【数 4】

$$D_a = \{(\mathbf{x}_1^a, v_1^a), \dots, (\mathbf{x}_m^a, v_m^a)\}$$

を一組の補助画像とし、

【 0 0 2 0 】

【数 5】

$$v_i^a \in \{1, \dots, K\}$$

は、画像

【 0 0 2 1 】

【数 6】

$$\mathbf{x}_i^a$$

のハイパークラスラベル（例えば自動車の視点）を示す。 がスーパークラスを表す場合、本出願人らは、 c をきめの細かいクラス c のスーパークラスとする。結果として、2つの用語「分類器」および「認識モデル」/「モデルエンジン」は、交換可能に使用される。

【 0 0 2 2 】

10

20

30

40

50

目標は、画像のきめの細かいクラスラベルを予測することができる認識モデルエンジンを学習することである。特に、本出願人らは、 $\Pr(y|x)$ によって与えられる予測関数を学習することを目的としており、異なるきめの細かいクラスに属する可能性がある入力画像が与えられる。同様に、本出願人らは、 $\Pr(y|x)$ がハイパークラス分類モデルエンジンを示すものとする。きめの細かい訓練画像および補助ハイパークラスラベル付き画像が与えられ、簡単な戦略は、共通の特徴を共有し、かつ、個別に分類器を学習することによって、多重タスク深層CNNを訓練することである。多重タスク深層学習を観察して、個々のタスクの性能を改善する。この単純な戦略をさらに改善するために、本出願人らは、きめの細かい分類器とハイパークラス分類器との間の正則化を利用することによる新規な多重タスク正則化学習フレームワークを開示する。本出願人らは、最初に、要因

10

【0023】

次に、要因クラス正則化学習について説明する。要因クラスは、きめの細かいクラスを生成するための隠れた変数と見做すことが可能であり、したがって本出願人らは、

【0024】

【数7】

$$\Pr(y|x) = \sum_{v=1}^K \Pr(y|v,x) \Pr(v|x)$$

20

(1)

によって $\Pr(y|x)$ をモデル化する。ここで、 $\Pr(v|x)$ は、任意の要因クラスの確率であり、 $\Pr(y|v,x)$ は、要因クラスおよび入力画像 x を与えられた任意のきめの細かいクラスの確率を規定する。本出願人らは、 $h(x)$ が x の高レベル特徴を示すものと仮定し、ソフトマックス関数

【0025】

【数8】

$$\Pr(v|x) = \frac{\exp(u_v^T h(x))}{\sum_{v'=1}^K \exp(u_{v'}^T h(x))}$$

30

(2)

によって確率 $\Pr(y|v,x)$ をモデル化する。ここで、 $\{u_v\}$ は、ハイパークラス分類モデルエンジンに対する重みを表す。すべての公式化において、バイアス項は本出願人の説明には無関係であるため、バイアス項を無視していることに留意されたい。しかしながら、実際にはバイアス項は含まれるべきである。 x の要因クラスおよび高レベル特徴 h が与えられると、確率 $\Pr(y|v,x)$ は、

【0026】

【数9】

$$\Pr(y=c|v,x) = \frac{\exp(w_{v,c}^T h(x))}{\sum_{c=1}^C \exp(w_{v,c}^T h(x))}$$

40

(3)

によって計算される。ここで、 $\{w_{v,c}\}$ は、要因に特化されたきめの細かい認識モデルエンジンの重みを表す。(2)および(3)を合わせ、本出願人らは、特定のきめの細かいクラスに対する以下の予測確率を有し、また、本出願人らは、この式を使用して最終予測を実施する。

50

【 0 0 2 7 】

【 数 1 0 】

$$\Pr(y = c | \mathbf{x}) = \sum_{v=1}^K \frac{\exp(\mathbf{w}_{v,c}^T \mathbf{h}(\mathbf{x}))}{\sum_{c=1}^C \exp(\mathbf{w}_{v,c}^T \mathbf{h}(\mathbf{x}))} \frac{\exp(\mathbf{u}_v^T \mathbf{h}(\mathbf{x}))}{\sum_{v'=1}^K \exp(\mathbf{u}_{v'}^T \mathbf{h}(\mathbf{x}))}$$

(4)

【 0 0 2 8 】

10

本出願人らのモデルエンジンは、その根源を混合モデルに有しているが、自由パラメータとして $\Pr(\cdot | \mathbf{x})$ を取り扱う従来のほとんどの混合モデルとは異なり、本出願人らは、それを識別モデルとして公式化していることは注目に値する。本出願人らの正確な $\{\mathbf{u}_v\}$ の学習を可能にしているのは、ハイパークラス拡張画像である。次に、本出願人らは、きめの細かい認識に対する D_t 中のデータの負の対数尤度と、ハイパークラス認識に対する D_a 中のデータの負の対数尤度を書き留めることができ、すなわち

【 0 0 2 9 】

【 数 1 1 】

$$\begin{aligned} L(\{\mathbf{w}_{v,c}\}, \{\mathbf{u}_v\}) &= -\log \Pr(D) \\ &= -\sum_{i=1}^n \sum_{c=1}^C \delta(y_i^t, c) \log \Pr(y = c | \mathbf{x}_i^t) \\ &\quad - \sum_{i=1}^m \sum_{v=1}^K \delta(v_i^a, v) \log \Pr(v | \mathbf{x}_i^a) \end{aligned}$$

(5)

20

である。

【 0 0 3 0 】

30

簡単ではない正則化を動機付けるために、要因特化重み $\mathbf{w}_{v,c}$ が、対応する要因クラス分類器 \mathbf{u}_v として、同じような高レベル要因関連特徴を捕捉すべきである。そのために、本出願人らは、 $\{\mathbf{w}_{v,c}\}$ と $\{\mathbf{u}_v\}$ の間の以下の正則化を導入する。

【 0 0 3 1 】

【 数 1 2 】

$$R(\{\mathbf{w}_{v,c}\}, \{\mathbf{u}_v\}) = \frac{\beta}{2} \sum_{v=1}^K \sum_{c=1}^C \|\mathbf{w}_{v,c} - \mathbf{u}_v\|_2^2 \quad (6)$$

40

上記正則化は、

【 0 0 3 2 】

【 数 1 3 】

$$\Pr(\mathbf{w}_{v,c} | \mathbf{u}_v) \propto \exp\left(-\frac{\beta}{2} \|\mathbf{w}_{v,c} - \mathbf{u}_v\|_2^2\right)$$

50

により $w_{v,c}$ を正規のプライアとすることによって解釈できる。

【0033】

(6)における正則化は、要因クラス認識モデルときめの細かい認識モデルの間で重みを共有する別の興味ある直観を享受する。これを調べるために、本出願人は $w'_{v,c} = w_{v,c} - u_v$ を導入し、また、(6)における正則化項 (regularizer) を

【0034】

【数14】

$$R(\{w'_{v,c}\}) = \frac{\beta}{2} \sum_{v=1}^K \sum_{c=1}^C \|w'_{v,c}\|_2^2$$

10

として記述し、また、 $\Pr(y = c | x)$ は

【0035】

【数15】

$$\Pr(y = c | x) = \frac{\sum_{v=1}^K \frac{\exp((w'_{v,c} + u_v)^T h(x))}{\sum_{c=1}^C \exp((w'_{v,c} + u_v)^T h(x))} \frac{\exp(u_v^T h(x))}{\sum_{v'=1}^K \exp(u_{v'}^T h(x))}$$

20

によって計算される。

【0036】

きめの細かい分類器は、要因クラス分類器の同じ成分 u_v を共有することが分かる。したがって、開示したモデルを従来の浅い多重タスク学習に使用されている重み共有に接続する。

【0037】

一方、スーパークラス正則化学習については、スーパークラス c がきめの細かいラベル c によって暗に示されるため、スーパークラス正則化深層学習に対する相違は、

30

【0038】

【数16】

$$\Pr(y = c | v_c, x) = \frac{\exp(w_{v_c,c}^T h(x))}{\sum_{c=1}^C \exp(w_{v_c,c}^T h(x))}$$

によって単純にモデル化できる $\Pr(y | v_c, x)$ にある。したがって、正則化は、

【0039】

【数17】

$$R(\{w_{v_c,c}\}, \{u_{v_c}\}) = \frac{\beta}{2} \sum_{c=1}^C \|w_{v_c,c} - u_{v_c}\|_2^2 \quad (7)$$

40

になる。

【0040】

同様の正則化が既に利用されていることが注目される。しかしながら、本出願人の仕

50

事との間には、スーパークラス分類用の重み u が、本出願人らのモデルエンジンでは、ハイパークラス拡張画像から差別的に学習される点で、大きな相違が存在している。

【0041】

一体化深層CNNが実施され得る。ハイパークラス拡張データおよび多重タスク正則化学習技法を使用することにより、本出願人らは、図1Bに示されている一体化深層CNNフレームワークに到達する。また、本出願人らは、最適化問題

【0042】

【数18】

$$\min_{\{\mathbf{w}_{v,c}\},\{\mathbf{u}_v\},\{\mathbf{w}_l\}} L(\{\mathbf{w}_{v,c}\},\{\mathbf{u}_v\}) + R(\{\mathbf{w}_{v,c}\},\{\mathbf{u}_v\}) + \sum_{v=1}^K r(\mathbf{u}_v) + \sum_{l=1}^H r(\mathbf{w}_l)$$

10

を提示する。ここで、 w_l , $l = 1, \dots, H$ は、高レベル特徴 $h(x)$ を決定する際のCNNのすべての重みを表し、 H は、分類器層の前の層数を表し、また、 $r(w)$ は、暗黙の正則化パラメータ（または重み崩壊パラメータ）を備えた標準ユークリッドノルム平方正則化器を表す。

【0043】

開示した深層学習モデルエンジンは、同様の設定のミニバッチ確率的勾配降下を使用した誤差逆伝播法（back propagation）によって訓練される。主な相違は、データの2つのソース、および2つのタスクに対応する2つの損失関数を有していることである。ミニバッチにおける D_t 中の画像および D_o 中の画像の両方をサンプリングして確率的勾配を計算することが、極めて重要である。2つのタスクを交互に訓練する代替手法を使用することは、極めて悪い解をもたらすことになり得る。2つのタスクは異なる方向に異なる局所最適条件を有しており、その解が悪い局所最適条件に容易にトラップされるからである。

20

【0044】

要約すると、FGIC用のハイパークラス拡張化および正則化深層学習フレームワークは、きめの細かいデータ中の、固有、かつ、容易に注釈が付けられたハイパークラスを識別し、ハイパークラスによってラベルが付けられた大量の類似した画像を収集することにより、新しいデータ拡張手法を使用する。本出願人らのシステムは、一体化された深層学習フレームワークにおける、最初の利用属性学習および情報共有である。現行法は1つの属性のみを使用するが、より多くのタスクを追加し、対での重み正則化を使用することによって、複数の属性を処理するように改善される。ハイパークラス拡張データは、多重タスク学習を深層CNNの中に組み込むことによって特徴学習を一般化することができる。汎化能力をさらに改善し、かつ、大きいクラス内分散を取り扱うために、本出願人らは、きめの細かいクラスとハイパークラスとの間の関係を利用する新規な正則化技術を開示した。開示したフレームワークの成功は、公的に入手可能な小規模のきめの細かいデータセットおよび自己収集大規模自動車データの両方に対して試験された。異なるタスク間の正則化を考慮することによって、多重タスク深層学習を考慮し得ることが予想される。

30

【0045】

図3に示すように、一態様による自律運転システム100は、様々な構成要素を有する車両101を含む。特定の態様は、特定のタイプの車両との接続にとりわけ有用であり、車両は、自動車、トラック、オートバイ、バス、ボート、航空機、ヘリコプタ、芝刈り機、レクリエーション用車両、遊園地車両、建設車両、農業器具、市街電車、ゴルフカート、列車およびトロリーを始めとする任意のタイプの車両であってもよいが、これらに限定されない。車両は、プロセッサ120、メモリ130、および汎用コンピュータに典型的に存在する他の構成要素を含んだコンピュータ110などの1つ以上のコンピュータを有する。

40

【0046】

メモリ130は、プロセッサ120によって実行され、さもなければプロセッサ120

50

によって使用される命令 1 3 2 およびデータ 1 3 4 を含む、プロセッサ 1 2 0 によるアクセスが可能な情報を記憶する。メモリ 1 3 0 は、コンピュータ可読媒体、またはハードドライブ、メモリカード、ROM、RAM、DVD または他の光ディスク、ならびに他の書き込み可能メモリおよびリードオンリメモリなどの電子デバイスの助けを借りて読み出され得るデータを記憶する他の媒体を含む、プロセッサによるアクセスが可能な情報を記憶することができる任意のタイプのメモリであってもよい。システムおよび方法は、上記の異なる組合せを含んでも良く、それにより命令およびデータの異なる部分が異なるタイプの媒体上に記憶される。

【 0 0 4 7 】

命令 1 3 2 は、プロセッサによって直接実行され（機械コードなど）、あるいは間接的に実行される（スクリプトなど）、任意のセットの命令であってもよい。例えば命令は、コンピュータコードとしてコンピュータ可読媒体上に記憶されてもよい。その点に関して、「命令」および「プログラム」という用語は、本明細書においては交換可能に使用され得る。命令は、プロセッサによって直接処理するオブジェクトコードフォーマットで、あるいはスクリプト、または要求時に解釈され、あるいは前もってコンパイルされる独立ソースコードモジュールの集合を含む任意の他のコンピュータ言語で記憶されてもよい。命令の機能、方法およびルーチンは、以下でより詳細に説明される。

【 0 0 4 8 】

データ 1 3 4 は、命令 1 3 2 に従ってプロセッサ 1 2 0 によって検索、記憶または修正され得る。例えば、システムおよび方法は、データ構造によって制限されるものではないが、データは、複数の異なる欄および記録、XML 文書または単層ファイルを有する表として、関係データベース内のコンピュータレジスタに記憶されてもよい。また、データは、任意のコンピュータ可読フォーマットでフォーマット化されてもよい。単なる他の例にすぎないが、画像データは、圧縮または解凍される、無損失の（例えば BMP）または損失のある（例えば JPEG）フォーマットに従って記憶される複数のピクセルの格子を含むビットマップ、および図形を描くためのコンピュータ命令はもちろん、ビットマップまたはベクトルベース（例えば SVG）として記憶されてもよい。データは、数、記述テキスト、専用（proprietary）コード、同じメモリまたは異なるメモリ（他のネットワーク位置を含む）の他の領域に記憶されたデータの参照、または関連データを計算するための機能によって使用される情報などの関連情報を識別するのに十分な任意の情報を含んでもよい。

【 0 0 4 9 】

プロセッサ 1 2 0 は、市販の CPU などの任意の従来のプロセッサであってもよい。別法としては、プロセッサは、ASIC などの専用デバイスであってもよい。図 1 は、コンピュータ 1 1 0 のプロセッサ、メモリおよび他の要素を同じブロック内に存在しているとして機能的に示したものであるが、プロセッサおよびメモリは、実際には、同じ物理的なハウジング内に收容され、又は、收容されなくてもよい複数のプロセッサおよびメモリから構成されてもよいことは、当業者が理解し得る。例えば、メモリは、コンピュータ 1 1 0 のハウジングとは異なるハウジング内に配置されたハードドライブであっても、あるいは他の記憶媒体であってもよい。したがって、プロセッサまたはコンピュータの参照には、並列に動作し、又は、動作しないプロセッサ、コンピュータまたはメモリの集合の参照を含むことが理解される。単一のプロセッサを使用してここに記載したステップを実行するよりはむしろ、ステアリング要素および減速要素などの複数の構成要素のうちのいくつかはそれぞれ、それらの構成要素の特定の機能に関連する計算のみを実施するそれ等自身のプロセッサを有してもよい。

【 0 0 5 0 】

本明細書において説明されている様々な態様では、プロセッサは、車両から遠隔に配置され、車両と無線で通信してもよい。他の態様では、ここに記載したプロセスのいくつかは、遠隔プロセッサによって車両内またはその他に配置されたプロセッサ上で実行されてもよく、単一の操舵を実行するのに必要なステップの踏むことを含んでもよい。

10

20

30

40

50

【 0 0 5 1 】

コンピュータ 1 1 0 は、中央処理装置（CPU）、データ 1 3 4 およびウェブブラウザなどの命令を記憶するメモリ（例えば、RAM および内部ハードドライブ）、電子ディスプレイ 1 4 2（例えば、スクリーン、小型LCDタッチスクリーン、または情報を表示するように動作することができる任意の他の電気デバイスを有するモニタ）、ユーザ入力（例えばマウス、キーボード、タッチスクリーンおよび/またはマイクロホン）、ならびに人の状態および願望についての明確な情報（例えばジェスチャ）または暗示的な情報（例えば「人が眠っている」）を収集するための様々なセンサ（例えばビデオカメラ）などの、コンピュータと関連して一般的に使用される構成要素のすべてを含んでもよい。

【 0 0 5 2 】

また、車両は、デバイスの地理的位置を決定する、コンピュータ 1 1 0 と通信する地理的位置コンポーネント 1 4 4 を含んでもよい。例えば、位置コンポーネントは、デバイスの緯度、経度および/または高度位置を決定するためのGPS受信機を含んでもよい。レーザベースローカライズシステム、慣性支援GPSまたはカメラベースローカライズなどの他の位置システムが、車両の位置を識別するために使用されてもよい。また、車両は、様々なソースから位置情報を受け取り、この情報を様々なフィルタを使用して結合して車両位置の「最良」の予測を識別してもよい。例えば、車両は、地図位置、GPS位置、および車両のその前の位置からの時間による位置の変化に基づくその車両の現在位置の予測を含む、複数の位置予測を識別してもよい。車両の位置の高度予測を識別するために、この情報を1つにまとめてもよい。本明細書において説明されている車両の「位置」は、緯度、経度および高度などの絶対な地理的位置、ならびに絶対な地理的位置より少ない雑音で多く決定される近傍の他の自動車に対する位置などの相対位置情報を含んでもよい。

【 0 0 5 3 】

また、デバイスは、車両の方向および速度、またはそれらに対する変化を決定するために、加速度計、ジャイロスコープ、または、別の方向/速度検出デバイス 1 4 6 などの、コンピュータ 1 1 0 と通信する他の特徴を含んでもよい。単なる例にすぎないが、デバイス 1 4 6 は、重力の方向または重力の方向に垂直な平面に対するそのピッチ、ヨーまたはロール（あるいはそれらに対する変化）を決定してもよい。また、デバイスは、速度の増減およびその変化の方向を追跡してもよい。本明細書において示されているデバイスの位置及び向きデータの規定は、ユーザ、コンピュータ 1 1 0、他のコンピュータおよび上記の組合せに自動的に提供されてもよい。

【 0 0 5 4 】

コンピュータは、様々な構成要素を制御することによって車両の方向および速度を制御してもよい。一例として、車両が完全自律モードで動作している場合、コンピュータ 1 1 0 は、（例えばエンジンに提供される燃料または他のエネルギーを増大することによって）車両を加速させ、（例えばエンジンに供給される燃料を減少することによって、あるいはブレーキをかけることによって）減速させ、また、（例えば前輪を向けることによって）方向を変化させてもよい。

【 0 0 5 5 】

車両は、他の車両、道路上の障害物、交通信号、標識、樹木、等々などの車両の外部の対象を検出するためのコンポーネント 1 4 8 を含んでもよい。検出システムは、レーザ、ソナー、レーダ、カメラまたは任意の他の検出デバイスを含んでもよい。例えば、車両が小型乗用車である場合、自動車は、屋根または他の便利な位置に取り付けられたレーザを含んでもよい。一態様では、レーザは、その軸上で回転させ、かつ、そのピッチを変えることによって、車両と、該車両と対向している対象表面との間の距離を測定してもよい。また、レーザを使用して、例えば暗い道路によって反射または吸収された光の量と、明るい車線ラインとを区別することによって、車線ラインを識別してもよい。また、車両は、適応巡航制御システムのために使用されるレーダ検出ユニットなどの様々なレーダ検出ユニットを含んでもよい。レーダ検出ユニットは、自動車の前後ならびにフロントバンパの両側に配置されてもよい。別の例では、図1のカメラによって例示したように、異なる画

10

20

30

40

50

像からの視差を使用して、1つ以上のカメラによって捕捉される様々な対象物までの距離を計算できるように、様々なカメラが自動車上に互いに既知の間隔で設けられてもよい。これらのセンサは、車両によるその環境の把握を可能にし、その環境中に存在する通行者ならびに対象または人の安全を最大にするために、潜在的にその環境への応答を可能にする。

【0056】

上述したセンサに加えて、コンピュータは、非自律車両に固有のセンサからの入力を使用してもよい。例えば、これらのセンサは、タイヤ圧センサ、エンジン温度センサ、ブレーキ熱センサ、ブレーキパッド状態センサ、タイヤトレッドセンサ、燃料センサ、オイルレベルおよび品質センサ、空気品質センサ（空気の温度、湿度または微粒子を検出するための）、等々を含んでもよい。

10

【0057】

これらのセンサの多くは、リアルタイムにコンピュータによって処理されるデータを提供し、すなわち、センサは、時間の範囲以上検知された環境を反映させるように、それらの出力を更新してもよく、また、コンピュータが、検知した環境に応じて、車両のその時点の方向または速度を修正すべきかどうかを決定することができるように、連続して、または、要求により、その更新した出力をコンピュータに提供してもよい。

【0058】

これらのセンサを使用して、道路上の歩行者、自転車、他の車両または対象の動きを識別し、追跡し、予測してもよい。例えば、センサは、車両の周囲の対象の位置および形状の情報をコンピュータ110に提供してもよく、対象を別の車両として識別してもよい。また、対象の現在の動きは、センサ（例えば、コンポーネントは独立型速度レーダ検出器である）によって決定されてもよく、あるいはセンサによって提供される情報に基づいて（例えば時間家かに伴う対象の位置データの変化を比較することによって）、コンピュータ110によって決定されてもよい。

20

【0059】

コンピュータは、検出された対象の存在に基づいて、車両の現在の経路および速度を変更してもよい。例えば、車両は、その現在の速度が50mphであり、また、そのカメラを使用し、かつ、光学文字認識を使用して、車両が間もなく、速度制限が35mphであることを示す標識を通過することを検出すると、自動的に減速してもよい。同様に、コンピュータは、対象が車両の意図した経路を妨害していることを決定すると、車両を操作してその障害物を迂回させてもよい。

30

【0060】

車両のコンピュータシステムは、検出された対象の期待される動きを予測してもよい。コンピュータシステム110は、単に対象の瞬時の方向、加速/減速および速度に基づいて、対象の将来の動きを単純に予測することが可能であり、例えば対象の現在の方向および動きが継続するであろうことを予測し得る。

【0061】

対象が検出されると、システムは、対象のタイプ、例えば、交通コーン、人、自動車、トラックまたは自転車を決定し、かつ、この情報を使用して対象の将来の挙動を予測してもよい。例えば、車両は、レーザによって決定される対象の形状、レーダに基づく、あるいはカメラ画像に基づくパターン一致による対象のサイズおよび速度のうちの一つ以上に基づいて対象のタイプを決定してもよい。また、対象は、対象のサイズ（自転車は、パン貯蔵箱より大きく、かつ、自動車より小さい）、対象の速度（自転車は、時速40マイルより速い速度で走り、または時速0.1マイル未満で走るといった傾向を示さない）、自転車によって発生される熱（自転車には、運転者に体熱を放出させる傾向がある）、等々のうちの一つ以上を考慮可能な対象分類器を使用することによって同じく識別され得る。

40

【0062】

いくつかの例では、車両によって識別される対象は、進路を変更するように車両に実際に要求しなくてもよい。例えば、砂嵐の間、車両は、1つ以上の対象として砂を検出して

50

もよいが、その軌道を変更する必要はなく、安全上の理由で車両自体の速度を遅くするか、または停止してもよい。

【 0 0 6 3 】

別の例では、車両の外部の場面は、様々なセンサの入力から分割する必要はなく、また、車両に応答動作を取らせるために対象を分類する必要もない。むしろ、車両は、対象の色および/または形状に基づいて1つ以上の動作を取ってもよい。

【 0 0 6 4 】

また、システムは、検出された対象の動きとは無関係の情報に依存して、対象の次のアクションを予測してもよい。一例として、車両が、別の対象が車両の前方の急な丘を登ろうとしている自転車であると決定すると、コンピュータが、自転車の速度が間もなく遅くなり、そのために、自転車が現在は比較的速い速度で走っているかどうかに関係なく、車両の速度が遅くなることを予測してもよい。

【 0 0 6 5 】

車両の外部の対象への識別、分類、及び反応の上記方法は、衝突を回避する可能性を高くするために、単独で、あるいは任意の組合せで使用されてもよいことは理解されよう。

さらなる例として、システムは、車両の近傍の対象が曲り専用車線の別の自動車であると決定してもよい(例えば、他の自動車を捕捉している画像データを解析することによって、他の自動車はその車線に存在し、該車線には左折の矢印が描かれている)。その点に関し、システムは、他の自動車が次の交差点で曲がるかもしれないことを予測してもよい。

【 0 0 6 6 】

コンピュータは、周囲の対象の予測動作に応じて、車両に特定の動作を取らせてもよい。例えば、コンピュータ110が、例えば自動車のターンシグナルまたは自動車が存在している車線に基づいて、車両に近づいている別の自動車が上で言及した次の交差点で曲ろうとしていることを決定すると、コンピュータは、その交差点に近づくと車両の速度を落としてもよい。この点に関して、他の対象の予測挙動は、対象のタイプおよびその現在の軌道のみならず、該対象が交通規則または所定の挙動を遵守する、あるいは遵守しない何らかの可能性にも基づいている。これは、車両による、法律上の、また、予測可能な挙動のみならず、不法なuターンまたは車線変更、赤信号無視、等々などの他の運転者による予期しない挙動に対する適切な応答を可能にする。

【 0 0 6 7 】

別の例では、システムは、様々な状況における対象パフォーマンスについての規則のライブラリを含んでもよい。例えば、ライト上に取り付けられた左折矢印を有する最も左側の車線にいる自動車は、矢印が緑色に変わると左に曲がる可能性が高い。ライブラリは、手動で構築されてもよく、または車両による道路上の他の車両(自律的なものであれ、非自律的なものであれ)の観察によって構築されてもよい。ライブラリは、車両観察によって改善可能な、人間が構築した規則のセットとして開始可能である。同様に、ライブラリは、車両観察から学習される規則として開始することが可能であり、人間が規則を審査し、それらを手動で改善することも可能である。この観察および学習は、例えば機械学習のツールおよび技術によって達成可能である。

【 0 0 6 8 】

様々なセンサによって提供されるデータの処理に加えて、コンピュータは、前の時点で得られた環境データであって、その環境における車両の存在には無関係に存続することが期待される環境データに依存してもよい。例えば、データ134は、詳細な地図情報136、例えば道路の形状および高さ、走行車線ライン、交差点、横断歩道、速度制限、交通信号、建物、標識、リアルタイム交通情報または他のそのような対象および情報を識別する高度詳細図を含んでもよい。走行車線ラインまたは交差点などの対象の各々は、例えば15cm、さらには1cmまで正確である地理的位置と関連付けられてもよい。また、地図情報は、例えば様々な道路セグメントと関連付けられた明確な速度制限情報を含んでもよい。速度制限データは、手動で入力されてもよく、または、例えば光学文字認識を使用

10

20

30

40

50

して、事前に取得した速度制限標識の画像から取り込まれてもよい。地図情報は、上で列挙した1つ以上の対象を組み込んだ三次元地形地図を含んでもよい。例えば、車両は、リアルタイムデータ（例えば、車両のセンサを使用して別の自動車の現在のGPS位置を決定する）、および他のデータ（例えばGPS位置を予め記憶した車線特定地図データと比較して、他の自動車が曲り車線にいるかどうかを決定する）に基づいて、別の自動車が曲がることを期待されることを決定してもよい。

【0069】

別の例では、車両は、道路の位置、属性および状態をより良好に識別するために、地図情報を使用してセンサデータを補足してもよい。例えば、道路の走行車線ラインが摩耗によって消失している場合、車両は、センサデータのみを利用するのではなく、地図情報に基づいて走行車線ラインの位置を予想可能である。

10

【0070】

また、車両センサを使用して、地図情報を収集し、かつ、補足してもよい。例えば、運転者は、道路の位置、走行車線ライン、交差点、交通信号、等々の様々なタイプの地図情報を検出し、記憶するために、非自律モードで車両を運転してもよい。後に、車両が、記憶した情報を使用して車両を操作してもよい。別の例では、車両が、時間とともに数センチメートル移動している橋、交差点における新しい交通パターンなどの環境変化を検出または観察するか、あるいは道路が舗装されて走行車線ラインが移動した場合に、この情報は、車両によって検出され、衝突を回避するための車両の操作方法に関する様々な決定を実施するために使用されるだけでなく、その車両の地図情報に組み込まれてもよい。いくつかの例では、運転者は、他の自律車両によって使用されるように、遠隔サーバに無線で送信することによって、その変化情報を中央地図データベースに報告することを任意に選択してもよい。サーバは、それに応答してデータベースを更新し、例えば情報を自動的に送信することによって、あるいは利用可能でダウンロード可能な更新を行うことによって、他の自律車両がすべての変化を利用することができるようにすることが可能である。したがって、環境変化は、遠隔サーバから多数の車両に対して更新可能である。

20

【0071】

別の例では、自律車両は、道路に沿った対象または道路の街路レベルの画像を取り込むカメラを備えてもよい。

【0072】

また、コンピュータ110は、車両の状態およびその構成要素を車両101の乗客に伝えるために、状態指示器138を制御可能である。例えば、車両101は、車両、特定のセンサまたはとりわけコンピュータ110の総合的な状態に関する情報を表示するための、図2に示されているようなディスプレイ225を備えてもよい。ディスプレイ225は、例えばコンピュータ、車両自体、道路、交差点ならびに他の対象の状態および情報を含む、車両の周囲のコンピュータ生成画像を含んでもよい。

30

【0073】

コンピュータ110は、視覚キューまたは可聴キューを使用して、コンピュータ110が様々なセンサから有効なデータを取得しているか否か、コンピュータが自動車の方向または速度、あるいは両方を部分的に、または完全に制御しているか否か、何らかの誤差が存在しているか否か、等々を識別可能である。また、車両101は、車両101の現在の状態を示すための状態バー230などの状態指示装置を含んでもよい。図2の例では、状態バー230は、車両が、現在、運転モードにあり、時速2マイルで移動していることを示す、「D」および「2mph」を表示する。その点に関して、車両は、電子ディスプレイ上にテキストを表示し、車両101の一部を照明し、あるいは様々な他のタイプの指示を提供してもよい。さらに、コンピュータは、その瞬間に、人間または自動システムが車両を制御中であるかどうかを示す、人間、他のコンピュータまたは両方による読取りが可能な外部指示器を有してもよい。

40

【0074】

一例では、コンピュータ110は、車両の様々な構成要素と通信することができる自律

50

運転計算システムであってもよい。例えば、コンピュータ110は、車両の従来の中央プロセッサ160と通信することが可能であり、また、車両101の動き、速度、等々を制御するために、車両101の様々なシステム、例えば制動180、加速182、信号発信184およびナビゲーション186のシステムに情報を送り、かつ、受け取ることができる。さらに、係合されると、コンピュータ110は、車両101のこれらの機能のうちのいくつか、またはすべてを制御することが可能であり、したがって完全に、または単に部分的に自律可能である。車両101内には様々なシステムおよびコンピュータ110が示されているが、これらの要素は、車両101の外部であっても、あるいは長い距離を隔てて物理的に分離されてもよいことは理解されよう。

【0075】

本開示の態様によるシステムおよび方法は、特定のタイプの対象の検出、または特定のタイプの車両動作あるいは環境条件の観察に限定されず、また、特定の機械学習プロセスに限定されないが、他の運転パターンとは差別化される、固有なシグネチャで運転パターンを導出し、かつ、学習するために使用されてもよい。

【0076】

説明され、かつ、図に示されているデータのサンプル値、タイプおよび構成は、単に明示を目的としたものにすぎない。その点に関し、本開示の態様によるシステムおよび方法は、様々なタイプのセンサ、通信デバイス、ユーザインタフェース、車両制御システム、データ値、データタイプおよび構成を含んでもよい。システムおよび方法は、異なる時間に（例えば異なるサーバまたはデータベースを介して）、また、異なる実体（例えばいくつかの値は事前に提案され、あるいは異なるソースから提供されてもよい）によって、提供され、受信されてもよい。

【0077】

上述の特徴のこれらおよび他の変形形態および組合せは、特許請求の範囲によって定義されるシステムおよび方法から逸脱することなく利用することができ、実施形態についての上記説明は、特許請求の範囲によって定義されている本開示を制限するものではなく、例証として解釈されるべきである。また、実施例の規定（「などの」、「例えば」、「含む」、等々のような文節句なども）は、本開示を特定の例に限定するものとして解釈されるべきではなく、どちらかと言えば、実施例は、多くの可能な態様のうちのほんのいくつかを示すことを意図していると解釈される。

【0078】

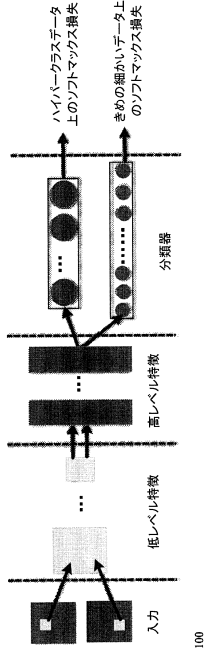
そうではないことが明確に言及されていない限り、所定の実施形態、代替または実施例におけるすべての特徴は、本明細書における他の実施形態、代替または実施例に使用可能である。例えば、車両の動きを検出するための適切なセンサは、本明細書におけるすべての構成で使用可能である。特定の運転者パターンまたはシグネチャ車両の動きを表すいかなるデータ構造も使用可能である。適切な機械学習プロセスは、本明細書における構成のいずれでも使用可能である。

10

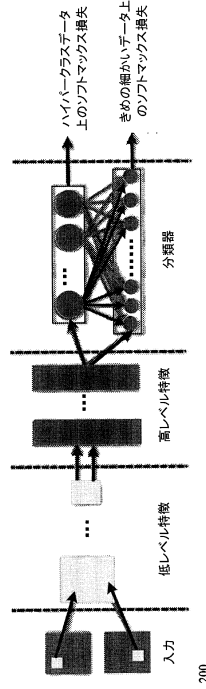
20

30

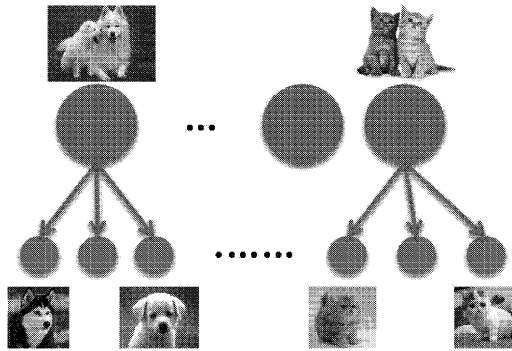
【図 1 A】



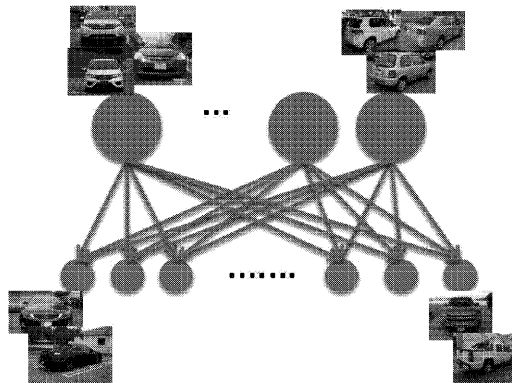
【図 1 B】



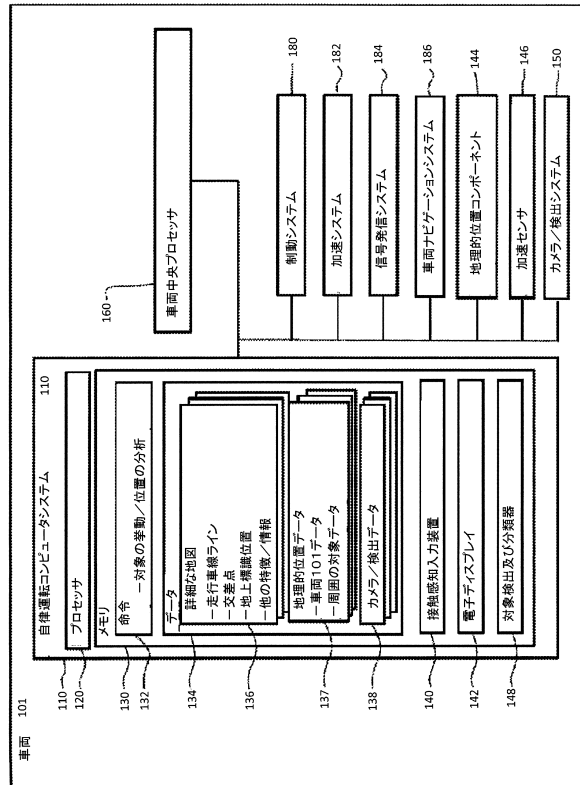
【図 2 A】



【図 2 B】



【図 3】



フロントページの続き

- (72)発明者 ヤン、 ティアンバオ
アメリカ合衆国 9 5 1 2 9 カリフォルニア州 サンノゼ アルバニー ドライブ 4 3 7 0
アパートメント 1 2
- (72)発明者 ワン、 シャオユ
アメリカ合衆国 9 4 0 8 7 カリフォルニア州 サニーヴェール アービュートス アヴェニュー
6 5 2 アpartment 4
- (72)発明者 リン、 ユエンチン
アメリカ合衆国 9 4 0 8 7 カリフォルニア州 サニーヴェール ローン ウェイ 1 0 7 5
- (72)発明者 シエ、 サニン
アメリカ合衆国 0 8 5 4 0 ニュージャージー州 プリンストン インディペンデンス ウェイ
4 スウィート 2 0 0

審査官 早川 学

- (56)参考文献 特開2005-352900(JP, A)
TETERWAK, Piotr et al., Shared Roots: Regularizing Neural Networks through Multitask Learning, Dartmouth College Computer Science Technical Report TR2014-762 [online], 2014年6月, pp.1-11, [検索日:2019.06.05], Internet<URL:https://www.cs.dartmouth.edu/~trdata/reports/abstracts/TR2014-762/>
外山託海、外1名、一般物体認識における学習データセットのサイズ拡張による影響の解析, 電子情報通信学会技術研究報告, 社団法人電子情報通信学会, 2010年5月6日, Vol. 110, No. 28, pp. 145~150

- (58)調査した分野(Int.Cl., DB名)
- | | |
|---------|-----------|
| G 0 6 T | 7 / 0 0 |
| G 0 6 F | 1 6 / 5 5 |
| G 0 6 N | 3 / 0 4 |
| G 0 6 N | 3 / 0 8 |
| G 0 6 N | 2 0 / 0 0 |