

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7539872号
(P7539872)

(45)発行日 令和6年8月26日(2024.8.26)

(24)登録日 令和6年8月16日(2024.8.16)

(51)国際特許分類 F I
G 0 6 T 7/00 (2017.01) G 0 6 T 7/00 3 5 0 B

請求項の数 15 (全14頁)

(21)出願番号	特願2021-519847(P2021-519847)	(73)特許権者	510192916 テスラ, インコーポレイテッド
(86)(22)出願日	令和1年10月10日(2019.10.10)		アメリカ合衆国 テキサス州 7 8 7 2 5 , オースティン, テスラ ロード 1
(65)公表番号	特表2022-504713(P2022-504713 A)	(74)代理人	110000659 弁理士法人広江アソシエイツ特許事務所
(43)公表日	令和4年1月13日(2022.1.13)	(72)発明者	クーバー, マシュー ジョン アメリカ合衆国 カリフォルニア州 9 4 3 0 4, パロ アルト, ディア クリーク ロード 3 5 0 0, シー/オー テスラ, インコーポレイテッド
(86)国際出願番号	PCT/US2019/055683	(72)発明者	ジャイン, バラス アメリカ合衆国 カリフォルニア州 9 4 3 0 4, パロ アルト, ディア クリーク ロード 3 5 0 0, シー/オー テスラ, インコーポレイテッド
(87)国際公開番号	WO2020/077117		
(87)国際公開日	令和2年4月16日(2020.4.16)		
審査請求日	令和3年12月6日(2021.12.6)		
審判番号	不服2023-14593(P2023-14593/J 1)		
審判請求日	令和5年8月30日(2023.8.30)		
(31)優先権主張番号	62/744,534		
(32)優先日	平成30年10月11日(2018.10.11)		
(33)優先権主張国・地域又は機関			
	最終頁に続く		最終頁に続く

(54)【発明の名称】 拡張データによって機械モデルを訓練するためのシステムおよび方法

(57)【特許請求の範囲】

【請求項1】

予測コンピュータモデルのパラメータのセットを訓練するための方法であって、
 1つまたは複数の画像収集システムに固定されている間に第1のカメラによってキャプチャされる画像のセットを特定するステップであって、前記第1のカメラが、特定のカメラ構成に従って前記1つまたは複数の画像収集システムに固定されている、ステップと、
 前記画像のセット内の各画像について、該画像の訓練出力を特定するステップと、
 前記画像のセット内の1つまたは複数の画像について、拡張画像のセットの拡張画像を生成するステップであって、
 前記画像に関連付けられた角度、スケールおよび/または姿勢が維持されるように前記画像のカメラ特性を維持する画像操作機能を用いて該画像を変更することによって拡張画像のセットの拡張画像を生成すること、および、
 前記拡張画像を前記画像の前記訓練出力と関連付けることによって、拡張画像のセットの拡張画像を生成するステップと、
 前記画像および前記拡張画像のセットを含む画像訓練セットに基づいて前記訓練出力を予測するように前記予測コンピュータモデルのパラメータのセットを訓練するステップと、
 を含み、
 訓練された前記予測コンピュータモデルは、モデル適用システムの自律または半自律制御での使用のために入力画像内の物体の存在を予測するように構成され、第2のカメラが、前記特定のカメラ構成に従って前記モデル適用システムに固定されており、

10

20

前記画像操作機能は、前記画像内の対象物体の部分に適用される切り取りを含み、前記切り取りのサイズは、前記対象物体のバウンディングボックスのサイズに基づいて決定される、方法。

【請求項 2】

前記特定のカメラ構成は、前記 1 つまたは複数の画像収集システムおよび前記モデル適用システムに関して同じ位置および/または向きを示し、一貫した画像の特性が、前記第 1 のカメラおよび前記第 2 のカメラによってキャプチャされた画像に含まれる、請求項 1 に記載の方法。

【請求項 3】

前記画像訓練セットが、画像のカメラ特性を変更する画像操作機能によって生成される画像を含まず、かつ、このような画像操作機能が、トリミング、パディング、水平もしくは垂直反転、またはアフィン変換の少なくとも 1 つを含む、請求項 1 に記載の方法。

10

【請求項 4】

前記画像操作機能が、切り取り、色相、彩度、値ジッタ、ごま塩、ドメイン移行、またはそれらの任意の組み合わせである、請求項 1 に記載の方法。

【請求項 5】

前記画像操作機能は、前記画像に適用される切り取りであり、前記切り取りの位置は、前記 1 つまたは複数の画像収集システムに関する前記第 1 のカメラの予測ビューに基づいて選択され、前記位置は、進行方向を描写する前記画像の中心を含む、請求項 1 に記載の方法。

20

【請求項 6】

前記画像操作機能が、前記画像の部分に適用される切り取りであり、前記切り取りの位置は、前記 1 つまたは複数の画像収集システムに関する前記第 1 のカメラの予測ビューに基づいて選択され、前記位置は、前記特定のカメラ構成に従って前記第 1 のカメラによってキャプチャされた画像に常に存在する、固定サイズの一貫したアーチファクトに対応する、請求項 1 に記載の方法。

【請求項 7】

前記画像操作機能が、前記画像内の前記訓練出力の位置と部分的に重なる前記画像の部分に適用される切り取りであり、前記切り取りは、ベタ色であり、前記切り取りは、前記切り取りの近くの領域と合成されている、請求項 1 に記載の方法。

30

【請求項 8】

1 つまたは複数のプロセッサと、命令を記憶している非一時的コンピュータ記憶媒体とを有するシステムであって、前記命令は、前記 1 つまたは複数のプロセッサによって実行されると、前記プロセッサに対して、

1 つまたは複数の画像収集システムに固定されている間に第 1 のカメラによってキャプチャされる画像のセットを特定するステップであって、前記第 1 のカメラが、特定のカメラ構成に従って前記 1 つまたは複数の画像収集システムに固定されている、ステップと、

前記画像のセット内の各画像について、該画像の訓練出力を特定するステップと、

前記画像のセット内の 1 つまたは複数の画像について、拡張画像のセットの拡張画像を生成するステップであって、

40

前記画像に関連付けられた角度、スケールおよび/または姿勢が維持されるように前記画像のカメラ特性を維持する画像操作機能を用いて該画像を変更することによって拡張画像のセットの拡張画像を生成すること、および、

前記拡張画像を前記画像の前記訓練出力と関連付けることによって、拡張画像のセットの拡張画像を生成するステップと、

前記画像および前記拡張画像のセットを含む画像訓練セットに基づいて前記訓練出力を予測するように予測コンピュータモデルのパラメータのセットを訓練するステップと、を含む動作を実施させ、

訓練された前記予測コンピュータモデルは、モデル適用システムの自律または半自律制御での使用のために入力画像内の物体の存在を予測するように構成され、第 2 のカメラが

50

、前記特定のカメラ構成に従って前記モデル適用システムに固定されており、

前記画像操作機能は、前記画像内の対象物体の部分に適用される切り取りを含み、前記切り取りのサイズは、前記対象物体のバウンディングボックスのサイズに基づいて決定される、システム。

【請求項 9】

前記特定のカメラ構成は、前記 1 つまたは複数の画像収集システムおよび前記モデル適用システムに関して同じ位置および/または向きを示し、一貫した画像の特性が、前記第 1 のカメラおよび前記第 2 のカメラによってキャプチャされた画像に含まれる、請求項 8 に記載のシステム。

【請求項 10】

前記画像操作機能が、切り取り、色相、彩度、値ジッタ、ごま塩、ドメイン移行、またはそれらの任意の組み合わせである、請求項 8 に記載のシステム。

【請求項 11】

前記画像操作機能が、前記画像の部分に適用される切り取りであり、前記切り取りの位置は、前記 1 つまたは複数の画像収集システムに関する前記第 1 のカメラの予測ビューに基づいて選択され、前記位置は、前記特定のカメラ構成に従って前記第 1 のカメラによってキャプチャされた画像に常に存在する、固定サイズの一貫したアーチファクトに対応する、請求項 8 に記載のシステム。

【請求項 12】

プロセッサによって実行するための命令を有する非一時的コンピュータ可読媒体であって、前記命令は、前記プロセッサによって実行されると、該プロセッサに対して、

1 つまたは複数の画像収集システムに固定されている間に第 1 のカメラによってキャプチャされる画像のセットを特定するステップであって、前記第 1 のカメラが、特定のカメラ構成に従って前記 1 つまたは複数の画像収集システムに固定されている、ステップと、

前記画像のセット内の各画像について、該画像の訓練出力を特定するステップと、

前記画像のセット内の 1 つまたは複数の画像について、拡張画像のセットの拡張画像を生成するステップであって、

前記画像に関連付けられた角度、スケールおよび/または体勢が維持されるように前記画像のカメラ特性を維持する画像操作機能を用いて該画像を変更することによって拡張画像のセットの拡張画像を生成すること、および、

前記拡張画像を前記画像の前記訓練出力と関連付けることによって、拡張画像のセットの拡張画像を生成するステップと、

前記画像および前記拡張画像のセットを含む画像訓練セットに基づいて前記訓練出力を予測するように学習するように予測コンピュータモデルを訓練するステップと、を実施させ、

訓練された前記予測コンピュータモデルは、モデル適用システムの自律または半自律制御での使用のために入力画像内の物体の存在を予測するように構成され、第 2 のカメラが、前記特定のカメラ構成に従って前記モデル適用システムに固定され、

前記特定のカメラ構成は、前記 1 つまたは複数の画像収集システムおよび前記モデル適用システムに関して同じ位置および/または向きを示し、一貫した画像の特性が、前記第 1 のカメラおよび前記第 2 のカメラによってキャプチャされた画像に含まれており、

前記画像操作機能は、前記画像内の対象物体の部分に適用される切り取りを含み、前記切り取りのサイズは、前記対象物体のバウンディングボックスのサイズに基づいて決定される、非一時的コンピュータ可読媒体。

【請求項 13】

前記画像操作機能が、前記画像の部分に適用される切り取りであり、前記切り取りの位置は、前記 1 つまたは複数の画像収集システムに関する前記第 1 のカメラの予測ビューに基づいて選択され、前記位置は、前記特定のカメラ構成に従って前記第 1 のカメラによってキャプチャされた画像に存在するアーチファクトを含む、請求項 12 に記載の非一時的コンピュータ可読媒体。

10

20

30

40

50

【請求項 1 4】

前記画像操作機能が、前記画像の部分に適用される切り取りであり、前記切り取りの位置は、前記 1 つまたは複数の画像収集システムに関する前記第 1 のカメラの予測ビューに基づいて選択され、前記位置は、前記特定のカメラ構成に従って前記第 1 のカメラによってキャプチャされた画像に常に存在する、固定サイズの一貫したアーチファクトに対応する、請求項 1 2 に記載の非一時的コンピュータ可読媒体。

【請求項 1 5】

前記画像操作機能が、前記画像内の前記訓練出力の位置と部分的に重なる領域に適用される切り取りであり、前記切り取りは、ベタ色であり、前記切り取りは、前記切り取りの近くの領域と合成されている、請求項 1 2 に記載の非一時的コンピュータ可読媒体。

10

【発明の詳細な説明】**【技術分野】****【0001】****[関連出願の相互参照]**

本出願は、米国仮特許出願第 62 / 744 , 534 号、出願日 2018 年 10 月 11 日、発明の名称「TRAINING MACHINE MODELS WITH DATA AUGMENTATION THAT RETAINS SENSOR CHARACTERISTICS」に対する優先権を主張する。米国仮特許出願第 62 / 744 , 534 号は、その全体が参照により本明細書に組み込まれる。

【0002】

20

本発明の実施形態は、一般に、機械学習環境においてデータを訓練するためのシステムおよび方法に関し、より詳細には、訓練データセットにセンサ特性などの追加のデータを含めることによって訓練データを拡張することに関する。

【背景技術】**【0003】**

典型的な機械学習アプリケーションでは、訓練データを取得するために使用されるキャプチャ機器の特性にモデルが過剰適合されることを回避するために、様々な方法でデータを拡張することができる。例えば、コンピュータモデルを訓練するために使用される典型的な画像セットでは、画像は、キャプチャされている物体に対して様々なセンサ特性を有する多くの異なるキャプチャ環境でキャプチャされる物体を表すことができる。例えば、そのような画像は、様々なスケール（例えば、画像内の有意に異なる距離）、様々な焦点距離、様々なレンズタイプ、様々な前処理または後処理、種々のソフトウェア環境、センサアレイハードウェアなどの様々なセンサ特性によってキャプチャされる場合がある。これらのセンサはまた、画像がキャプチャされるとき環境に対する撮像センサの位置および向きなどの種々の外因性パラメータに関しても異なる場合がある。これらの異なるタイプのセンサ特性のすべてが、キャプチャされる画像を画像セット内の複数の異なる画像全体にわたって別様にかつ様々に提示させ、コンピュータモデルを適切に訓練することをより困難にする可能性がある。

30

【0004】

ニューラルネットワークの多くのアプリケーションは、様々な条件でキャプチャされるデータから学習し、様々な異なるセンサ構成で（例えば、複数のタイプの携帯電話上で作動するアプリにおいて）展開される。画像をキャプチャするために使用されるセンサの差異を考慮に入れるために、開発者は、焦点距離、軸スキュー、位置、および回転などのカメラ特性に関して、開発されたモデルを一般化する画像の反転、回転、またはトリミングなどの修正によって画像訓練データを拡張することができる。

40

【0005】

これらの変動を考慮に入れ、訓練されたネットワークを様々なソース上で展開するために、訓練データを拡張または操作して、訓練されたモデルのロバスト性を高めることができる。しかしながら、これらの手法は、典型的には、拡張画像においてカメラ特性を変更する変換を適用することによって、モデルが任意の特定のカメラ構成について効果的に学

50

習することを妨げる。

【発明の概要】

【課題を解決するための手段】

【0006】

一実施形態は、予測コンピュータモデルのパラメータのセットを訓練するための方法である。この実施形態は、1つまたは複数の画像収集システムに固定されている間にカメラのセットによってキャプチャされる画像のセットを特定することと、画像のセット内の各画像について、画像の訓練出力を特定することと、画像のセット内の1つまたは複数の画像について、画像のカメラ特性を維持する画像操作機能を用いて画像を変更することによって拡張画像のセットの拡張画像を生成し、拡張訓練画像を画像の訓練出力と関連付けることとによって、拡張画像のセットの拡張画像を生成することと、画像および拡張画像のセットを含む画像訓練セットに基づいて訓練出力を予測するように予測コンピュータモデルのパラメータのセットを訓練することとを含むことができる。

10

【0007】

追加の実施形態は、1つまたは複数のプロセッサと、命令を記憶している非一時的コンピュータ記憶媒体とを有するシステムであって、命令は、1つまたは複数のプロセッサによって実行されると、プロセッサに、1つまたは複数の画像収集システムに固定されている間にカメラのセットによってキャプチャされる画像のセットを特定することと、画像のセット内の各画像について、画像の訓練出力を特定することと、画像のセット内の1つまたは複数の画像について、画像のカメラ特性を維持する画像操作機能を用いて画像を変更することによって拡張画像のセットの拡張画像を生成し、拡張訓練画像を画像の訓練出力と関連付けることとによって、拡張画像のセットの拡張画像を生成することと、画像および拡張画像のセットを含む画像訓練セットに基づいて訓練出力を予測するように予測コンピュータモデルのパラメータのセットを訓練することとを含む動作を実施させる、システムを含むことができる。

20

【0008】

別の実施形態は、プロセッサによって実行するための命令を有する非一時的コンピュータ可読媒体であって、命令は、プロセッサによって実行されると、プロセッサに、1つまたは複数の画像収集システムに固定されている間にカメラのセットによってキャプチャされる画像のセットを特定することと、画像のセット内の各画像について、画像の訓練出力を特定することと、画像のセット内の1つまたは複数の画像について、画像のカメラ特性を維持する画像操作機能を用いて画像を変更することによって拡張画像のセットの拡張画像を生成し、拡張訓練画像を画像の訓練出力と関連付けることとによって、拡張画像のセットの拡張画像を生成することと、画像および拡張画像のセットを含む画像訓練セットに基づいて訓練出力を予測するようにコンピュータモデルを訓練することとを実施させる、非一時的コンピュータ可読媒体を含むことができる。

30

【図面の簡単な説明】

【0009】

【図1】一実施形態によるコンピュータモデルの訓練および展開のための環境のブロック図である。

40

【0010】

【図2】同じカメラ特性によってキャプチャされる例示的な画像を示す図である。

【0011】

【図3】一実施形態による、モデル訓練システムの構成要素のブロック図である。

【0012】

【図4】一実施形態による、ラベル付き訓練画像に基づく拡張画像の生成の一例を示すデータフロー図である。

【発明を実施するための形態】

【0013】

図面は、例示のみを目的として本発明の様々な実施形態を示す。当業者は、以下の論述

50

から、本明細書に記載される本発明の原理から逸脱することなく、本明細書に示される構造および方法の代替的な実施形態が利用され得ることを容易に認識するであろう。

【0014】

一実施形態は、元来キャプチャされている画像のカメラ特性を維持するために拡張された画像を用いてコンピュータモデルを訓練するシステムである。これらのカメラ特性は、カメラの内因性または外因性の特性を含むことができる。そのような内因性の特性は、ダイナミックレンジ、視野、焦点距離、およびレンズ歪みなどのセンサ自体の特性を含むことができる。外因性の特性は、カメラの角度、スケール、または体勢など、キャプチャされる環境に対するカメラの構成を表すことができる。

【0015】

これらの内因性および外因性の特性は、画像内でキャプチャされる物体および他の態様、ならびにデバイスまたはシステム上でのその位置付けのためにカメラのビュー内に現れる静止物体などのアーチファクトおよび他の効果に対するカメラのビューに影響を及ぼし得る。例えば、車両に搭載されたカメラは、そのビューの一部として、多くの画像にわたって、および、同じモデルの車に同じ方法で搭載されているその構成のすべてのカメラについて現れる車のフードを含む場合がある。別の例として、これらのカメラ特性はまた、カメラのビュー内の物体から生じる反射を含み得る。反射は、カメラによってキャプチャされる画像の多くに含まれるようになる1つのタイプの一貫した特性であり得る。

【0016】

拡張画像を用いて訓練データにさらに追加しながら、データモデルを訓練するために画像のカメラ特性を維持、保存、記憶、または使用することによって、結果として得られるモデルは、同じカメラ特性を有する多くの異なるデバイスにわたって有用であり得る。さらに、拡張は、特に画像が曇っている、遮られている、または他の様態で検出可能な物体の明瞭なビューを提供しない場合に、モデル予測に一般化およびより高いロバスト性を提供することができる。これらの手法は、物体検出および自律型車両に特に有用であり得る。この手法はまた、同じカメラ構成が多くのデバイスに展開され得る他の状況にも有益であり得る。これらのデバイスは一貫した向きにある一貫したセンサセットを有することができるため、訓練データは所与の構成によって収集することができ、モデルは収集された訓練データからの拡張データによって訓練することができ、訓練されたモデルは同じ構成を有するデバイスに展開することができる。したがって、これらの技法は、この文脈における不必要な一般化をもたらし、何らかのデータ拡張を伴う他の変数の一般化を可能にする拡張を回避する。

【0017】

カメラ特性を維持するために、拡張画像を生成するために使用される画像操作機能は、カメラ特性を維持する機能である。例えば、これらの操作は、キャプチャされる環境に対するカメラの角度、スケール、または体勢への影響を回避することができる。実施形態では、カメラ特性に影響を及ぼす画像操作機能によって拡張された画像は訓練に使用されない。例えば、カメラ特性を維持するために使用され得る画像操作機能は、切り取り、色相/彩度/値ジッタ、ごま塩、ならびにドメイン移行（例えば、昼から夜に変更する）を含む。カメラ特性を変更する場合があります。したがっていくつかの実施形態では使用されない機能は、トリミング、パディング、反転（水平または垂直）、またはアフィン変換（せん断、回転、並進、およびスキューなど）を含む。

【0018】

さらなる例として、画像は、元の画像の一部を除去する「切り取り」機能によって拡張することができる。次いで、画像の除去された部分は、指定された色、ぼけ、ノイズ、または別の画像からのものなどの他の画像内容と置き換えることができる。切り取りの数、サイズ、領域、および置換内容は変更されてもよく、画像のラベル（例えば、画像内の関心領域、または物体のバウンディングボックス）に基づいてもよい。

【0019】

したがって、コンピュータモデルを、画像および拡張画像によって訓練することができ

10

20

30

40

50

、キャプチャされる画像のカメラ特性を有するデバイスに配布して、モデルをセンサ分析に使用することができる。特に、このデータ拡張およびモデル訓練は、画像内の物体または物体バウンディングボックスを検出するように訓練されたモデルに使用することができる。

【 0 0 2 0 】

図 1 は、一実施形態によるコンピュータモデルの訓練および展開のための環境である。1 つまたは複数の画像収集システム 1 4 0 は、モデル適用システムによって展開および使用され得るコンピュータモデルを訓練する際にモデル訓練システムによって使用され得る画像をキャプチャする。これらのシステムは、これらの装置がそれらを通じて通信する様々な無線または有線通信リンクを表すインターネットなどのネットワーク 1 2 0 を介して接続される。

10

【 0 0 2 1 】

モデル訓練システム 1 3 0 は、入力のセットを所与として出力を予測するために訓練可能なパラメータのセットを有するコンピュータモデルを訓練する。この例におけるモデル訓練システム 1 3 0 は、典型的には、画像に関する出力予測情報を生成するために画像入力に基づいてモデルを訓練する。例えば、様々な実施形態において、これらの出力は、画像内の物体を識別することができる（バウンディングボックスまたはセグメント化のいずれかによって物体を識別し、画像の状態（例えば、時刻、天候を識別することができる）または画像の他のタグもしくは記述子を識別することができる）。

【 0 0 2 2 】

便宜上、本明細書においては例示的なタイプのセンサデータとして画像が使用されているが、本明細書に記載の拡張およびモデル開発は、センサ構成特性を維持しながらこれらのセンサからキャプチャされる訓練データを拡張するために、様々なタイプのセンサに適用することができる。

20

【 0 0 2 3 】

画像収集システム 1 4 0 は、画像収集システム 1 4 0 の環境から情報をキャプチャするセンサのセットを有する。単一の画像収集システム 1 4 0 が示されているが、多くの画像収集システム 1 4 0 が、モデル訓練システム 1 3 0 のための画像をキャプチャすることができる。画像収集システム 1 4 0 のためのセンサは、画像収集システム 1 4 0 にわたって同じまたは実質的に同じであり得るセンサ特性を有する。一実施形態における画像収集システムは、環境内を移動し、カメラによって環境の画像をキャプチャする車両または他のシステムである。画像収集システム 1 4 0 は、手動で操作されてもよく、または、部分的もしくは完全に自動化された車両によって操作されてもよい。したがって、画像収集システム 1 4 0 が環境を移動するとき、画像収集システム 1 4 0 は、環境の画像をキャプチャし、モデル訓練システム 1 3 0 に送信することができる。

30

【 0 0 2 4 】

モデル適用システム 1 1 0 は、画像収集システムと同じまたは実質的に同じセンサ特性を有するセンサのセットを有するシステムである。いくつかの例では、モデル適用システム 1 1 0 は、画像収集システム 1 3 0 としても機能し、キャプチャされるセンサデータ（例えば、画像）を、さらなる訓練データとして使用するためにモデル訓練システム 1 3 0 に提供する。モデル適用システム 1 1 0 は、モデル訓練システム 1 3 0 から訓練されたモデルを受信し、そのセンサによって感知されるデータと共にモデルを使用する。画像収集システム 1 4 0 およびモデル適用システム 1 1 0 からキャプチャされる画像は同じカメラ構成を有するため、モデル適用システム 1 1 0 は、画像収集システムと同じ方法で、同じ視点（または実質的に同様）からその環境をキャプチャすることができる。モデルを適用した後、モデル適用システム 1 1 0 は、様々な目的のためにモデルの出力を使用することができる。例えば、モデル適用システム 1 1 0 が車両である場合、モデルは画像内の物体の存在を予測することができ、これは、安全システムの一部として、または自律（または半自律）制御システムの一部としてモデル適用システム 1 1 0 によって使用され得る。

40

【 0 0 2 5 】

50

図2は、同じカメラ特性によってキャプチャされる例示的な画像を示す図である。この例では、画像200Aは、画像収集システム130上のカメラによってキャプチャされる。別の画像200Bもまた、画像収集システム130によってキャプチャされてもよく、この画像収集システムは同じ画像収集システムであってもよく、または異なる画像収集システム130であってもよい。異なる環境および環境内の異なる物体をキャプチャしている間、これらの画像は、環境をキャプチャする画像に対するカメラ特性を維持する。カメラ特性は、環境がカメラ内でどのように見えるかに影響を与えるカメラの構成および向きを指す。例えば、これらのカメラ特性は、環境に対するカメラの角度、スケール、および体勢（例えば、見る位置）を含むことができる。画像がキャプチャされる同じ環境に対するカメラの角度、スケール、または位置を変更すると、環境の画像が変化する。例えば、より高い位置に配置されたカメラは、異なる高さから物体を見、その物体の下側の位置とは異なる部分を示す。同様に、これらの画像は、解析されるべき環境の一部ではないカメラ構成に起因する、画像内の一貫したアーチファクトおよび効果を含む。例えば、画像200Aと200Bの両方は、フロントガラスからのグレアおよび他の効果を含み、画像の右下側の物体は環境を遮蔽し、フロントガラスは画像の下部を遮蔽する。したがって、同じカメラ特性からキャプチャされる画像は、典型的には、同じアーチファクト、歪みを提示し、同じ方法で環境をキャプチャする。

10

【0026】

図3は、一実施形態による、モデル訓練システム130の構成要素を示す。モデル訓練システムは、コンピュータモデルを訓練するための様々なモジュールおよびデータストアを含む。モデル訓練システム130は、モデルの一般化を改善するために画像収集システム140からの画像を拡張することによって、モデル適用システム110によって使用されるモデルを訓練する。拡張画像は、画像のカメラ構成に影響を及ぼさない（例えば、維持する）画像操作機能を用いて生成される。これは、モデルパラメータが一貫したカメラ特性に関連する重みをより近密に学習することを可能にしながら、画像間で異なり得る画像の態様についての過剰適合をより選択的に回避するモデルパラメータの一般化を可能にしながら、より効果的なモデリングを可能にする。

20

【0027】

モデル訓練システムは、画像収集システム140から画像を受信するデータ入力モジュール310を含む。データ入力モジュール310は、これらの画像を画像データストア350に格納することができる。データ入力モジュール310は、データ収集システム140によって生成または提供されるものとしての画像を受信してもよく、または、画像収集システム140から画像を要求してもよい。

30

【0028】

ラベリングモジュール320は、画像データ350内の画像においてラベルを識別または適用することができる。いくつかの例では、画像はすでに識別された特性を有していてもよい。ラベルはまた、訓練されたモデルによって予測または出力されるデータを表すことができる。例えば、ラベルは、画像に示される環境内の特定の物体を指定することができる、または画像に関連付けられた記述子または「タグ」を含むことができる。モデルの用途に応じて、ラベルは様々な方法でこの情報を表すことができる。例えば、物体は、画像内のバウンディングボックスに関連付けられてもよく、または物体は、画像の他の部分からセグメント化されてもよい。したがって、ラベル付き画像は、モデルがそれに対して訓練されるグラウンドトゥールズを表すことができる。画像は、任意の適切な手段によってラベリングされてもよく、典型的には、教師ありラベリングプロセス（例えば、ユーザが画像を検討し、画像のラベルを指定することによってラベリングされる）によってラベリングされてもよい。次いで、これらのラベルを画像データストア350内の画像に関連付けることができる。

40

【0029】

画像拡張モジュール330は、画像収集システム140によってキャプチャされる画像に基づいて追加の画像を生成することができる。これらの画像は、モデル訓練モジュール

50

340の訓練パイプラインの一部として生成されてもよく、またはこれらの拡張画像は、モデル訓練モジュール340において訓練を開始する前に生成されてもよい。拡張画像は、画像収集システム140によってキャプチャされる画像に基づいて生成することができる。

【0030】

図4は、一実施形態による、ラベル付き訓練画像400に基づく拡張画像の生成例を示す。ラベル付き訓練画像は、画像収集システム140によってキャプチャされる画像であってもよい。訓練画像410は、ラベル付き訓練画像400内のラベル付きデータに対応する関連付けられた訓練出力420Aを有する、拡張されていない訓練画像410Aを含むことができる。

10

【0031】

画像拡張モジュール330は、ラベル付き訓練画像400に画像操作機能を適用することによって拡張画像を生成する。画像操作機能は、モデルを訓練するために画像の特性を変化させるために、ラベル付き訓練画像400の修正バージョンを生成する。訓練画像を生成するために使用される画像操作機能は、ラベル付き訓練画像400のカメラ特性を維持する。したがって、操作機能は、様々なデバイスにわたって一貫し得る環境をキャプチャするとき、カメラの物理的キャプチャ特性またはカメラの位置によって影響を受け得る環境のビューのスケール、視点、向き、および他の特性を維持することができる。したがって、画像操作機能は、物体または環境の他の特徴がどのように視認可能であるか、またはこれらがシーン内でどの程度明瞭に見えるかに影響を及ぼすことができるが、画像内の物体の位置またはサイズには影響を及ぼさないことができる。カメラ特性を維持する、適用され得る例示的な画像操作機能は、切り取り、ジッタ（例えば、色相、彩度、または色値の）、ごま塩（黒および白のドットを導入する）、ぼけ、およびドメイン移行を含む。これらの画像操作機能のうちの複数を組み合わせて適用して、拡張画像を生成することができる。切り取りとは、画像の一部を除去し、除去した部分を他の画像内容に置き換える画像操作機能を指す。ドメイン移行は、画像内の別の環境条件に対応するように画像を変更する画像操作機能を指す。例えば、日中の画像を、夜間に画像がどのように見えるかを近似するように変更することができ、または日の当たる場所で撮影された画像を、雨もしくは雪の効果を加えるように変更することができる。

20

【0032】

これらの拡張画像は、ラベル付き訓練画像400と同じ訓練出力に関連付けることができる。図4に示す例では、拡張画像410Bは、ラベル付き訓練画像400に切り取りを適用することによって生成され、拡張画像410Bは訓練出力420Bに関連付けられ得る。同様に、訓練画像410Cを生成するために、画像の複数の部分を変更するために複数の切り取りが適用される。この例では、訓練画像410Cを生成するために適用される切り取りは、異なるパターンで画像の切り取り領域を埋める。

30

【0033】

様々な実施形態において、切り取りは、訓練画像および画像内の訓練出力の位置に基づいて変化し得る様々なパラメータおよび構成を用いて適用されてもよい。したがって、切り取りの数、サイズ、位置、および置換画像内容は、異なる実施形態において、訓練出力の位置に基づいて変化してもよい。例として、切り取り機能は、同様のサイズの複数の切り取りを適用することができ、または、ある範囲内の異なる半ランダム化されたサイズのいくつかの切り取りを適用することができる。複数の切り取りを使用し、サイズを変えることによって、切り取りは、（様々なサイズの）現実世界の障害物が物体を見ることに及ぼす影響をより近密にシミュレートすることができ、訓練されたモデルが任意の1つの特定のサイズの切り取りを補償するように学習するのを妨げることができる。

40

【0034】

切り取りのサイズの範囲は、画像内の物体または他のラベルのサイズの一部分に基づいてもよい。例えば、切り取りは、画像内の物体のバウンディングボックスのサイズの40%以下であってもよく、または、最小の物体のバウンディングボックスより小さくてもよ

50

い。これにより、切り取りが対象物体を完全に不明瞭にしないこと、したがって、画像が、モデルがそこから学習することができる物体の画像データを含み続けることを保証することができる。切り取りの数はまた、ランダム化され、均一分布、ガウス分布、または指数分布などの分布から選択されてもよい。

【0035】

さらに、切り取りの位置は、画像内の物体の位置に基づいて選択されてもよい。これは、バウンディングボックスとのいくつかの、ただし過度ではない重なりをもたらし得る。物体と切り取り領域との間の交差は、切り取りによって置き換えられている物体の部分によって測定されてもよく、または、物体と切り取り領域との交差を物体の面積と切り取り領域との和集合によって除算することによって測定され得る交差オーバーユニオン（IoU）によって測定されてもよい。例えば、切り取り領域は、20%～50%の範囲内の交差オーバーユニオン値を有するように配置されてもよい。したがって、切り取りにいくつかの、ただし多大ではない量の物体を含めることによって、切り取りは、関連する画像データをあまりにも多く除去することなく物体を部分的に不明瞭にするより「困難な」例を作成することができる。同様に、画像内のカメラの予測ビューに基づいて、画像の特定の部分に切り取りを選択することもできる。例えば、底部は典型的には常に存在するアーチファクトを含み得るが、画像の中心は最も関心のある領域（例えば、車両の場合、多くの場合は、車両の進行方向である）であり得るため、切り取りは主に画像の下半分または画像の中心に位置し得る。

【0036】

切り取り領域に対する置換画像データは、ベタ色（例えば、一定）であってもよく、または、ガウスノイズなどの別のパターンであってもよい。別の例として、閉塞または他の障害物を表すために、切り取りは、同じ画像タイプまたはラベルを有する別の画像からの画像データのパッチと置き換えられてもよい。最後に、切り取りは、例えばポアソン合成によって、切り取りの近くの領域と合成することができる。背景パッチまたは合成などの様々な合成手法を使用することにより、これらは、切り取り内の置換データが環境と区別することがより困難であることを保証し、したがって現実世界の障害物により類似した例を提供することができる。

【0037】

図4においては矩形領域として示されているが、拡張画像を生成する際に適用される切り取りは、他の実施形態では種々の形状に変化してもよい。拡張画像410B、410Cを生成し、拡張画像を関連する訓練出力420B、420Cに関連付けた後、画像拡張モジュール330は、画像データストア350にこれらの画像を追加することができる。

【0038】

モデル訓練モジュール340は、画像収集システム140によってキャプチャされる画像および画像拡張モジュール330によって生成される拡張画像に基づいてコンピュータモデルを訓練する。これらの画像は、モデル訓練のための画像訓練セットとして使用することができる。一実施形態では、機械学習モデルは、訓練データに基づいてモデル訓練モジュール340によって訓練される、フィードフォワードネットワーク、畳み込みニューラルネットワーク（CNN）、ディープニューラルネットワーク（DNN）、リカレントニューラルネットワーク（RNN）、自己組織化マップ（SOM）などのニューラルネットワークモデルである。訓練後、コンピュータモデルは、訓練済みコンピュータモデルストア370に格納することができる。モデルは、センサデータ（例えば、画像）を入力として受け取り、モデルの訓練に従って出力予測を出力する。モデルを訓練する際に、モデルは、訓練データの損失関数によって評価されるものとして、入力画像に基づいて出力を予測するパラメータのセットを学習する（または「訓練する」）。すなわち、訓練中、訓練データは、予測を生成するために現在のパラメータセットに従って評価される。訓練入力についてのその予測は、（例えば、損失関数を用いて）損失を評価するために指定の出力（例えば、ラベル）と比較することができ、パラメータは、損失関数を低減するようにパラメータセットを最適化するために最適化アルゴリズムを介して修正することができる

10

20

30

40

50

。「最適化」と呼ばれているが、これらのアルゴリズムは、パラメータセットに関して損失を低減することができるが、入力の設定を所与としてパラメータの「最適な」値を求めることが保証されない場合がある。例えば、勾配降下最適化アルゴリズムは、大域的最小値ではなく、極小値を求めることができる。

【0039】

拡張訓練データに対してコンピュータモデルを訓練することにより、コンピュータモデルは、キャプチャされているデータのセンサ特性を有する環境において動作する物理センサからのセンサデータに適用されるときに、改善された正確度で実行することができる。拡張はこれらの特性を維持するため、これらのセンサ特性（例えば、カメラ特性）は、データの訓練に使用される画像内で表現される。一実施形態では、訓練データは、画像をトリミング、パディング、反転（垂直または水平）、またはアフィン変換（例えば、せん断、回転、並進、スキュー）を適用する操作など、画像のカメラ特性を変更する画像操作機能によって生成される拡張画像を含まない。

10

【0040】

訓練後、モデル配布モジュール380は、訓練されたモデルを適用するために、訓練されたモデルをシステムに配布することができる。特に、モデル配布モジュール380は、モデル適用システム110のセンサに基づいて画像の特性を検出するために使用するために、訓練されたモデル（またはそのパラメータ）をモデル適用システム110に送信することができる。したがって、モデルからの予測を、モデル適用システム110の動作において、例えば、モデル適用システム110の物体検出および制御において使用することができる。

20

【0041】

本発明の実施形態の前述の説明は、例示を目的として提示されており、網羅的であること、または本発明を開示された正確な形態に限定することは意図されていない。当業者は、上記の開示に照らして多くの変更および変形が可能であることを理解することができる。

【0042】

本明細書のいくつかの部分は、情報に対する操作のアルゴリズムおよび記号表現に関して本発明の実施形態を説明している。これらのアルゴリズム記述および表現は、データ処理分野の当業者によって、それらの作業の内容を他の当業者に効果的に伝えるために一般的に使用される。これらの動作は、機能的、計算的、または論理的に説明されているが、コンピュータプログラムまたは同等の電気回路、マイクロコードなどによって実装されると理解されたい。さらに、一般性を失うことなく、これらの動作の構成をモジュールとして参照することが時には便利であることも証明されている。記載された動作およびそれらに関連するモジュールは、ソフトウェア、ファームウェア、ハードウェア、またはそれらの任意の組み合わせにおいて具現化され得る。

30

【0043】

本明細書に記載されたステップ、動作、またはプロセスのいずれも、単独で、または他のデバイスと組み合わせて、1つまたは複数のハードウェアまたはソフトウェアモジュールによって実施または実装することができる。一実施形態では、ソフトウェアモジュールは、記載されたステップ、動作、またはプロセスのいずれかまたはすべてを実施するためにコンピュータプロセッサによって実行することができる、コンピュータプログラムコードを含むコンピュータ可読媒体を備えるコンピュータプログラム製品によって実装される。

40

【0044】

本発明の実施形態はまた、本明細書の動作を実施するための装置（例えば、システム）に関し得る。この装置は、必要な目的のために特別に構成されてもよく、および/または、コンピュータに格納されているコンピュータプログラムによって選択的に起動もしくは再構成される汎用コンピューティングデバイスを備えてもよい。コンピューティングデバイスは、1つまたは複数のプロセッサおよび/またはコンピュータシステムのシステムまたはデバイスであってもよい。そのようなコンピュータプログラムは、非一時的有形コンピュータ可読記憶媒体、またはコンピュータシステムバスに結合され得る、電子命令を記

50

憶するのに適した任意の種類媒体に記憶されてもよい。さらに、本明細書において言及される任意のコンピューティングシステムは、単一のプロセッサを含んでもよく、またはコンピューティング能力を高めるために複数のプロセッサ設計を利用するアーキテクチャであってもよい。

【 0 0 4 5 】

本発明の実施形態はまた、本明細書に記載の計算プロセスによって生成される製品に關し得る。そのような製品は、計算プロセスから生じる情報を含むことができ、情報は、非一時的有形コンピュータ可読記憶媒体に記憶され、本明細書に記載のコンピュータプログラム製品または他のデータの組み合わせの任意の実施形態を含むことができる。

【 0 0 4 6 】

最後に、本明細書において使用される文言は、主に読みやすさおよび説明目的のために選択されており、本発明の主題を描写または限定するために選択されたものではない。したがって、本発明の範囲は、この詳細な説明によってではなく、本明細書に基づく出願に関して発行される任意の特許請求項によって限定されることが意図される。したがって、本発明の実施形態の開示は、添付の特許請求の範囲に記載されている本発明の範囲の例示であるが、これを限定するものではないことが意図されている。

10

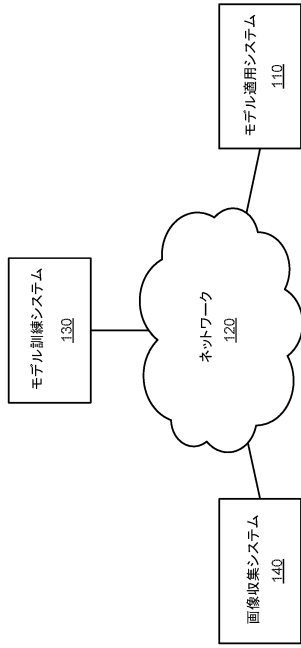
20

30

40

50

【図面】
【図 1】



100
FIG. 1

【図 2】

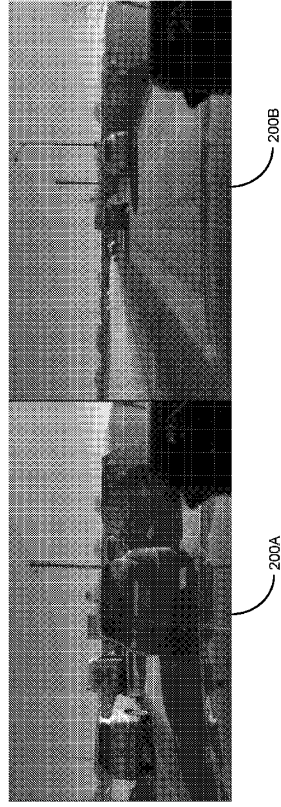


FIG. 2

【図 3】

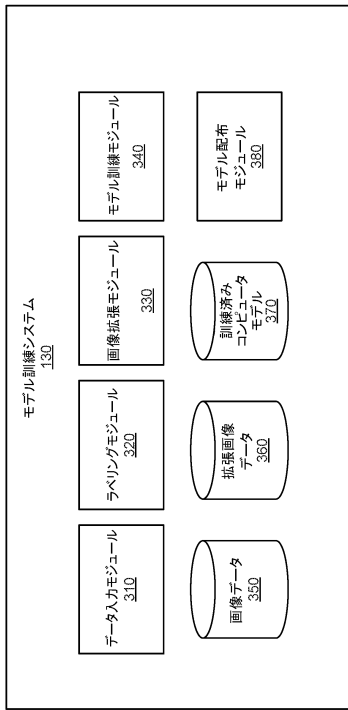


FIG. 3

【図 4】

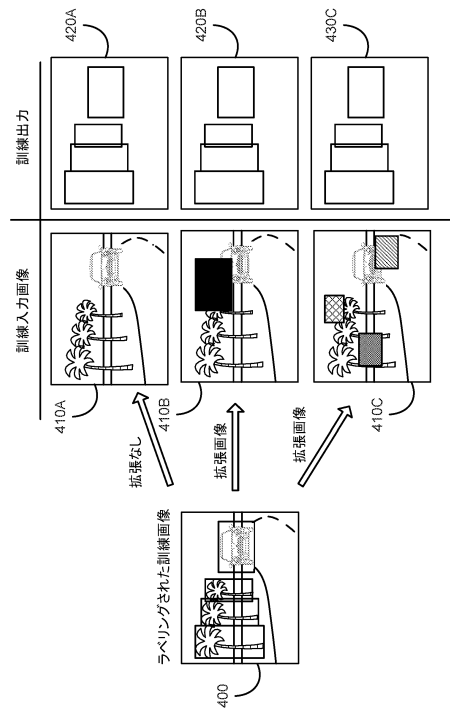


FIG. 4

フロントページの続き

- 米国(US)
インコーポレイテッド
- (72)発明者 シジュ, ハーシムラン シン
アメリカ合衆国 カリフォルニア州 94304, パロ アルト, ディア クリーク ロード 350
0, シー/オー テスラ, インコーポレイテッド
- 合議体
審判長 畑中 高行
審判官 片岡 利延
審判官 高橋 宣博
- (56)参考文献 特開2018-81404号公報
国際公開第2017/068692号
Zhun Zhong, et al., Random Erasing Data Augmentation, arxiv.org, 2017年, <https://doi.org/10.48550/arXiv.1708.04896>
内田 祐介, 外1名, 畳み込みニューラルネットワークの研究動向, 電子情報通信学会技術
研究報告 Vol.117 No.362, 日本, 一般社団法人電子情報通信学会, 20
17年, 第117巻
Terrance De Vries, et al., Improved Regularization of Convolutional Neural Networks with Cutout, arxiv.org, 2017年, <https://arxiv.org/abs/1708.04552>
- (58)調査した分野 (Int.Cl., DB名)
G06T 7/00