

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7271809号
(P7271809)

(45)発行日 令和5年5月11日(2023.5.11)

(24)登録日 令和5年4月28日(2023.4.28)

(51)国際特許分類 F I
G 0 6 T 7/00 (2017.01) G 0 6 T 7/00 3 5 0 C
G 0 6 N 20/00 (2019.01) G 0 6 N 20/00 1 3 0

請求項の数 16 (全21頁)

(21)出願番号	特願2023-513902(P2023-513902)	(73)特許権者	000006633 京セラ株式会社 京都府京都市伏見区竹田鳥羽殿町 6 番地
(86)(22)出願日	令和4年5月27日(2022.5.27)	(74)代理人	100147485 弁理士 杉村 憲司
(86)国際出願番号	PCT/JP2022/021815	(74)代理人	230118913 弁理士 杉村 光嗣
(87)国際公開番号	WO2022/250154	(74)代理人	100132045 弁理士 坪内 伸
(87)国際公開日	令和4年12月1日(2022.12.1)	(74)代理人	100195534 弁理士 内海 一成
審査請求日	令和5年2月27日(2023.2.27)	(72)発明者	浅谷 南己 京都府京都市伏見区竹田鳥羽殿町 6 番地 京セラ株式会社内
(31)優先権主張番号	特願2021-90676(P2021-90676)	審査官	新井 則和
(32)優先日	令和3年5月28日(2021.5.28)		
(33)優先権主張国・地域又は機関	日本国(JP)		
早期審査対象出願			

最終頁に続く

(54)【発明の名称】 学習済みモデル生成装置、学習済みモデル生成方法、及び認識装置

(57)【特許請求の範囲】

【請求項 1】

入力情報に含まれる認識対象の認識結果を出力する学習済みモデルを生成する制御部を備え、

前記制御部は、

前記入力情報と同一又は関連する学習対象の情報を含む教師データを用いた第1の学習を実行することによって生成された少なくとも1つのベースモデルに結合された状態で、前記学習対象の情報のうち前記第1の学習で用いられた情報と異なる情報を含む教師データを用いた第2の学習を実行することによって生成された、前記入力情報を前記少なくとも1つのベースモデルに入力する前に変換可能なアダプタを取得し、

10

前記学習対象の情報のうち前記第1の学習で用いられた情報及び前記第2の学習で用いられた情報のいずれとも異なる情報を含む教師データを用いた第3の学習を実行することによってターゲットモデルを生成し、

前記アダプタと前記ターゲットモデルとを結合することによって前記学習済みモデルを生成する、学習済みモデル生成装置。

【請求項 2】

前記ベースモデルは、前記入力情報と同一又は関連するタスクの情報として仮想的に生成された学習対象の第1疑似情報を教師データとして事前学習されたモデルであり、

前記制御部は、

前記学習対象の実際の態様を表す第1実情報及び前記第1疑似情報の少なくとも一方を

20

更に教師データとして、前記ベースモデルに結合させた前記アダプタを学習させ、

前記認識対象を表すデータとして仮想的に生成された第2疑似情報、又は、前記認識対象の実際の態様を表す第2実情報を教師データとして学習して前記ターゲットモデルを生成する、請求項1に記載の学習済みモデル生成装置。

【請求項3】

前記第1疑似情報、前記第2疑似情報、前記第1実情報、及び前記第2実情報は画像を含み、

前記アダプタは、入力される画像の態様を変換して出力する、請求項2に記載の学習済みモデル生成装置。

【請求項4】

前記ベースモデルは、前記入力情報と同一又は関連するタスクの情報として仮想的に生成された学習対象の第1疑似情報のみを教師データとして事前学習されたモデルであり、

前記制御部は、前記認識対象を表すデータとして仮想的に生成された第2疑似情報のみを教師データとして学習して前記ターゲットモデルを生成する、請求項2に記載の学習済みモデル生成装置。

【請求項5】

複数の前記ベースモデルが事前学習によって生成され、

前記アダプタは、前記複数のベースモデルそれぞれに前記入力情報を入力可能に構成され、

前記制御部は、前記アダプタの出力を前記複数のベースモデルの各々に入力させて学習させることによって、少なくとも前記アダプタを生成する、請求項1から4までのいずれか一項に記載の学習済みモデル生成装置。

【請求項6】

前記制御部は、前記アダプタの出力を前記複数のベースモデルの各々に入力させて学習させることによって、前記アダプタのみを生成又は更新する、請求項5に記載の学習済みモデル生成装置。

【請求項7】

前記制御部は、前記アダプタを生成するために、前記複数のベースモデルを複数のグループに分類して前記各グループを順番に、前記アダプタを生成するための事前学習に適用する、請求項5に記載の学習済みモデル生成装置。

【請求項8】

前記制御部は、前記各グループに1つの前記ベースモデルを分類する、請求項7に記載の学習済みモデル生成装置。

【請求項9】

前記制御部は、前記アダプタを生成するための事前学習に前記各グループを適用する順番をランダムに決定する、請求項7に記載の学習済みモデル生成装置。

【請求項10】

前記制御部は、前記認識対象を表すデータとして仮想的に生成された第2疑似情報とに基づいて学習することによって、前記アダプタに結合されるターゲットモデルを生成する、請求項1から4までのいずれか一項に記載の学習済みモデル生成装置。

【請求項11】

前記制御部は、前記アダプタを前記ターゲットモデルに結合した状態で学習することによって前記アダプタを生成する、請求項1から4までのいずれか一項に記載の学習済みモデル生成装置。

【請求項12】

前記制御部は、前記入力情報と同一又は関連するタスクの損失関数を最適化するように、前記ベースモデルに結合したアダプタを学習する、請求項1から4までのいずれか一項に記載の学習済みモデル生成装置。

【請求項13】

前記制御部は、前記入力情報と同一又は関連するタスク以外の損失関数を最適化するよ

10

20

30

40

50

うに学習することによって、前記ベースモデルに結合したアダプタを生成する、請求項 1 から 4 までのいずれか一項に記載の学習済みモデル生成装置。

【請求項 1 4】

前記ベースモデルは、前記入力情報の特徴量を抽出した結果を出力する第 1 ベースモデルと、前記第 1 ベースモデルの出力に基づいて前記入力情報についての所定の判断を行う第 2 ベースモデルとを含む、請求項 1 から 4 までのいずれか一項に記載の学習済みモデル生成装置。

【請求項 1 5】

入力情報に含まれる認識対象の認識結果を出力する学習済みモデルを生成する学習済みモデル生成装置が実行する学習済みモデル生成方法であって、

前記入力情報と同一又は関連する学習対象の情報を含む教師データを用いた第 1 の学習を実行することによって生成された少なくとも 1 つのベースモデルに結合された状態で、前記学習対象の情報のうち前記第 1 の学習で用いられた情報と異なる情報を含む教師データを用いた第 2 の学習を実行することによって生成された、前記入力情報を前記少なくとも 1 つのベースモデルに入力する前に変換可能なアダプタを取得することと、

前記学習対象の情報のうち前記第 1 の学習で用いられた情報及び前記第 2 の学習で用いられた情報のいずれとも異なる情報を含む教師データを用いた第 3 の学習を実行することによってターゲットモデルを生成することと、

前記アダプタと前記ターゲットモデルとを結合することによって前記学習済みモデルを生成することと

を含む学習済みモデル生成方法。

【請求項 1 6】

入力情報に含まれる認識対象の認識結果を出力する学習済みモデルを備えた認識装置であって、

前記学習済みモデルは、

前記入力情報と同一又は関連する学習対象の情報を含む教師データを用いた第 1 の学習を実行することによって生成された少なくとも 1 つのベースモデルに結合された状態で、前記学習対象の情報のうち前記第 1 の学習で用いられた情報と異なる情報を含む教師データを用いた第 2 の学習を実行することによって生成された、前記入力情報を前記少なくとも 1 つのベースモデルに入力する前に変換可能なアダプタと、

前記学習対象の情報のうち前記第 1 の学習で用いられた情報及び前記第 2 の学習で用いられた情報のいずれとも異なる情報を含む教師データを用いた第 3 の学習を実行することによって生成したターゲットモデルと

を含み、

前記アダプタと前記ターゲットモデルとを結合することによって構成されている、認識装置。

【発明の詳細な説明】

【関連出願へのクロスリファレンス】

【0001】

本出願は、日本国特許出願 2021-90676 号（2021 年 5 月 28 日出願）の優先権を主張するものであり、当該出願の開示全体を、ここに参照のために取り込む。

【技術分野】

【0002】

本開示は、学習済みモデル生成装置、学習済みモデル生成方法、及び認識装置に関する。

【背景技術】

【0003】

従来、対象標本を合成した学習画像を用いた学習によって生成された識別関数に被識別画像を入力することによって被識別画像に対象が含まれるか識別する装置が知られている（例えば特許文献 1 参照）。

【先行技術文献】

10

20

30

40

50

【特許文献】

【0004】

【文献】特開2016-71502号公報

【発明の概要】

【0005】

本開示の一実施形態に係る学習済みモデル生成装置は、入力情報に含まれる認識対象の認識結果を出力する学習済みモデルを生成する制御部を備える。前記制御部は、前記入力情報と同一又は関連する学習対象の情報を含む教師データを用いた第1の学習を実行することによって生成された少なくとも1つのベースモデルに結合された状態で、前記学習対象の情報のうち前記第1の学習で用いられた情報と異なる情報を含む教師データを用いた第2の学習を実行することによって生成された、前記入力情報を前記少なくとも1つのベースモデルに入力する前に変換可能なアダプタを取得する。前記制御部は、前記学習対象の情報のうち前記第1の学習で用いられた情報及び前記第2の学習で用いられた情報のいずれとも異なる情報を含む教師データを用いた第3の学習を実行することによってターゲットモデルを生成する。前記制御部は、前記アダプタと前記ターゲットモデルとを結合することによって前記学習済みモデルを生成する。

10

【0006】

本開示の一実施形態に係る学習済みモデル生成方法は、入力情報に含まれる認識対象の認識結果を出力する学習済みモデルを生成する学習済みモデル生成装置によって実行される。前記学習済みモデル生成方法は、前記入力情報と同一又は関連する学習対象の情報を含む教師データを用いた第1の学習を実行することによって生成された少なくとも1つのベースモデルに結合された状態で、前記学習対象の情報のうち前記第1の学習で用いられた情報と異なる情報を含む教師データを用いた第2の学習を実行することによって生成された、前記入力情報を前記少なくとも1つのベースモデルに入力する前に変換可能なアダプタを取得することを含む。前記学習済みモデル生成方法は、前記学習対象の情報のうち前記第1の学習で用いられた情報及び前記第2の学習で用いられた情報のいずれとも異なる情報を含む教師データを用いた第3の学習を実行することによってターゲットモデルを生成することを含む。前記学習済みモデル生成方法は、前記アダプタと前記ターゲットモデルとを結合することによって前記学習済みモデルを生成することを含む。

20

【0007】

本開示の一実施形態に係る認識装置は、入力情報に含まれる認識対象の認識結果を出力する学習済みモデルを備える。前記学習済みモデルは、前記入力情報と同一又は関連する学習対象の情報を含む教師データを用いた第1の学習を実行することによって生成された少なくとも1つのベースモデルに結合された状態で、前記学習対象の情報のうち前記第1の学習で用いられた情報と異なる情報を含む教師データを用いた第2の学習を実行することによって生成された、前記入力情報を前記少なくとも1つのベースモデルに入力する前に変換可能なアダプタを含む。前記学習済みモデルは、前記学習対象の情報のうち前記第1の学習で用いられた情報及び前記第2の学習で用いられた情報のいずれとも異なる情報を含む教師データを用いた第3の学習を実行することによって生成したターゲットモデルを含む。前記学習済みモデルは、前記アダプタと前記ターゲットモデルとを結合することによって構成されている。

30

40

【図面の簡単な説明】

【0008】

【図1】一実施形態に係る学習済みモデル生成システムの構成例を示すブロック図である。

【図2】イメージアダプタが結合される、汎用ライブラリと学習済みモデルとを示す模式図である。

【図3】イメージアダプタの一例を示す図である。

【図4】複数のベースモデルに結合されるイメージアダプタの生成と、イメージアダプタの学習済みモデルへの転移による学習済みモデルの生成とを示す模式図である。

【図5】学習済みモデル生成方法の手順例を示すフローチャートである。

50

【図 6】ロボット制御システムの構成例を示す模式図である。

【発明を実施するための形態】

【0009】

学習済みモデルを用いた認識における認識精度の向上が求められる。本開示の一実施形態に係る学習済みモデル生成装置、学習済みモデル生成方法、及び認識装置によれば、認識精度が向上され得る。

【0010】

(学習済みモデル生成装置 20 の構成例)

図 1 に示されるように、本開示の一実施形態に係る学習済みモデル生成装置 20 は、制御部 22 と、情報生成部 26 とを備える。学習済みモデル生成装置 20 は、学習済みモデル 70 (図 2 参照) を生成する。

10

【0011】

制御部 22 は、情報生成部 26 から学習に適用される対象に関する情報を取得する。学習に適用される対象は、学習対象とも称される。制御部 22 は、情報生成部 26 から取得した学習対象に関する情報を教師データとする学習を実行し、学習結果に基づく情報又はデータを出力する。例えば学習済みモデル 70 が工業部品等の特定の物体を認識するモデルとして生成される場合、その学習済みモデル 70 を生成するための学習対象は、認識する物体そのものを含んでもよいし、他の物体を含んでもよい。学習済みモデル 70 が認識できる物体は、認識対象とも称される。

【0012】

制御部 22 は、種々の機能を実行するための制御及び処理能力を提供するために、少なくとも 1 つのプロセッサを含んで構成されてよい。プロセッサは、制御部 22 の種々の機能を実現するプログラムを実行してよい。プロセッサは、単一の集積回路として実現されてよい。集積回路は、IC (Integrated Circuit) とも称される。プロセッサは、複数の通信可能に接続された集積回路及びディスクリート回路として実現されてよい。プロセッサは、他の種々の既知の技術に基づいて実現されてよい。

20

【0013】

制御部 22 は、記憶部を備えてよい。記憶部は、磁気ディスク等の電磁記憶媒体を含んでもよいし、半導体メモリ又は磁気メモリ等のメモリを含んでもよい。記憶部は、各種情報を格納する。記憶部は、制御部 22 で実行されるプログラム等を格納する。記憶部は、非一時的な読み取り可能媒体として構成されてもよい。記憶部は、制御部 22 のワークメモリとして機能してよい。記憶部の少なくとも一部は、制御部 22 とは別体として構成されてもよい。

30

【0014】

情報生成部 26 は、制御部 22 における学習で用いられる教師データを制御部 22 に出力する。情報生成部 26 は、教師データを生成してもよいし、外部装置から教師データを取得してもよい。

【0015】

情報生成部 26 は、教師データを生成又は取得するための制御及び処理能力を提供するために、少なくとも 1 つのプロセッサを含んで構成されてよい。プロセッサは、教師データを生成又は取得するプログラムを実行してよい。情報生成部 26 は、制御部 22 と同一又は類似に構成されてもよい。情報生成部 26 は、制御部 22 と一体に構成されてもよい。

40

【0016】

情報生成部 26 は、教師データとして、学習対象の実際の態様を表す情報を生成してよい。学習対象の実際の態様を表す情報は、実情報とも称される。情報生成部 26 は、学習対象の実際の画像を撮影するカメラを備えてもよい。情報生成部 26 は、学習対象の実際の画像に対してラベル等の情報を付与するアノテーションを実行してもよい。情報生成部 26 は、アノテーションに関する操作入力をユーザから受け付けてよい。情報生成部 26 は、あらかじめ準備されたアノテーションに関する学習モデルに基づいてアノテーションを実行してもよい。情報生成部 26 は、学習対象の実際の画像に対してアノテーションを

50

実行することによって実情報を生成できる。

【0017】

情報生成部26は、教師データとして、学習済みモデル70に入力する入力情報と同一又は関連するタスクの情報として学習対象に関する情報を仮想的に生成する。学習済みモデル70を、画像に含まれる物体を認識して分類するタスクに利用することが想定されている場合、入力情報は物体が撮影された画像となる。入力情報と同一又は関連するタスクは、学習済みモデル70による処理対象の入力情報を用いて実行されるタスク、又は入力情報と類似、関連する情報を用いて実行されるタスクに対応する。例えば、画像に含まれる所定の種類のネジと釘とを分類するタスクにおいて、入力情報と同一のタスクは、実際に学習済みモデル70に分類させることとなるネジと釘とを分類するタスクに対応する。10
入力情報と関連するタスクは、所定の種類のネジ及び釘と類似する他の種類のネジ若しくは釘、又はこれらと類似する物体も含めた画像から、ネジ及び釘を分類するタスクに対応する。仮想的に生成された学習対象に関する情報は、疑似情報とも称される。疑似情報は、例えばネジ又は釘の実物を実際に撮影した画像情報ではなく、認識対象となるネジ又は釘を、CG(Computer Graphics)等で描画した画像であってよい。タスクは、例えば、入力情報に含まれる認識対象を少なくとも2種類に分類する分類タスクを含んでよい。タスクは、例えば、認識対象がネジであるか釘であるかを区別するタスク、又は、入力情報に基づいて少なくとも1種類の評価値を算出する評価タスクを含んでもよい。分類タスクは、例えば認識対象が犬であるか猫であるかを区別するタスクなどに細分化され得る。20
タスクは、分類タスクに限られず、他の種々の動作を実現するタスクを含んでよい。タスクは、特定の対象物に属する画素から判定するセグメンテーションを含んでよい。タスクは、内包する矩形領域を検出する物体検出を含んでよい。タスクは、対象物の姿勢推定を含んでよい。タスクは、ある特徴点を見つけるキーポイント検出を含んでよい。

【0018】

ここで、入力情報と学習対象に関する情報とが両方とも分類タスクの情報である場合、入力情報と学習対象に関する情報との間の関係は、関連するタスクの情報になっているとする。さらに、入力情報と学習対象に関する情報とが両方とも、認識対象が犬であるか猫であるかを区別するタスクの情報である場合、入力情報と学習対象に関する情報との間の関係は、同一のタスクの情報になっているとする。入力情報と学習対象に関する情報との間の関係は、これらの例に限られず、種々の条件で定められ得る。30

【0019】

情報生成部26は、疑似情報を生成するために、学習対象の外観を仮想的に表す情報を生成してよい。情報生成部26は、学習対象の外観を仮想的に表す情報として、例えば学習対象の外観の三次元CAD(Computer Aided Design)データ等のモデリングデータを生成してもよい。情報生成部26は、学習対象の外観を仮想的に表す情報として、学習対象の画像を生成してもよい。情報生成部26は、学習対象の外観を仮想的に表すモデリングデータ又は画像等に対してラベル等の情報を付与するアノテーションを実行してもよい。情報生成部26は、生成した学習対象の外観を仮想的に表す情報にアノテーションを実行することによって疑似情報を生成できる。40

【0020】

情報生成部26は、学習対象の外観を仮想的に表す情報を外部装置から取得してもよい。情報生成部26は、モデリングデータに関する入力をユーザから受け付けてもよい。情報生成部26は、学習対象の外観を仮想的に表す情報に対してアノテーションが実行されたデータを取得してもよい。情報生成部26は、アノテーションに関する操作入力をユーザから受け付けてもよい。情報生成部26は、あらかじめ準備されたアノテーションに関する学習モデルに基づいて、学習対象の外観を仮想的に表す情報に対してアノテーションを実行してもよい。40

【0021】

(学習済みモデル70の例)

図2に示されるように、学習済みモデル生成装置20は、入力情報に含まれる認識対象

10

20

30

40

50

の認識結果を出力する学習済みモデル70を生成する。学習済みモデル70は、ターゲットモデル40の入力側にイメージアダプタ50を結合したモデルとして構成される。イメージアダプタ50は、入力情報を入力可能に構成される。イメージアダプタ50は、単にアダプタとも称される。

【0022】

学習済みモデル生成装置20は、学習済みモデル70を生成するための準備として以下の動作を実行する。学習済みモデル生成装置20は、疑似情報に基づいて学習することによってベースモデル30を生成する。ベースモデル30を生成するために実行される学習は、第1の学習とも称される。第1の学習で用いられる教師データは、入力情報と同一又は関連する学習対象の情報を含んでよい。なお、学習済みモデル生成装置20は、ベースモデル30の生成において、疑似情報の代わりに実情報を用いてもよいし、疑似情報と実情報とを併用してもよい。ベースモデル30を生成するための学習に用いられる疑似情報は、第1疑似情報とも称される。学習済みモデル生成装置20は、ベースモデル30の入力側にイメージアダプタ50を結合した状態で実情報に更に基づいて学習することによって、イメージアダプタ50を生成する。イメージアダプタ50を生成するために実行される学習は、第2の学習とも称される。第2の学習で用いられる教師データは、入力情報と同一又は関連する学習対象の情報を含んでおり、第1の学習で用いられた情報と異なる情報を含んでよい。イメージアダプタ50を生成するための学習に用いられる実情報は、第1実情報とも称される。なお、第1疑似情報及び第1実情報として、後述する第2疑似情報及び第2実情報が用いられてもよい。

【0023】

学習済みモデル生成装置20は、イメージアダプタ50を結合しない状態で疑似情報又は実情報に基づいて学習することによってターゲットモデル40を生成する。ターゲットモデル40を生成するために実行される学習は、第3の学習とも称される。第3の学習で用いられる教師データは、入力情報と同一又は関連する学習対象の情報を含んでおり、第1の学習で用いられた情報及び第2の学習で用いられた情報のいずれとも異なる情報を含んでよい。ターゲットモデル40を生成するための学習に用いられる疑似情報は、第2疑似情報とも称される。ターゲットモデル40を生成するための学習に用いられる実情報は、第2実情報とも称される。学習済みモデル生成装置20は、ベースモデル30に結合した状態で事前学習することによってあらかじめ生成したイメージアダプタ50を転移し、新たに生成したターゲットモデル40の入力側に結合することによって学習済みモデル70を生成する。なお、学習済みモデル生成装置20は、ターゲットモデル40として、事前学習に用いたベースモデル30を転移させてもよい。また、学習済みモデル生成装置20は、イメージアダプタ50とターゲットモデル40とを結合させて、第2疑似情報、第2実情報を教師データとして、更に学習を行って学習済みモデル70を生成してもよい。

【0024】

学習済みモデル生成装置20は、事前学習によってあらかじめイメージアダプタ50を生成しておくことによって、疑似情報だけに基づく学習によってターゲットモデル40を生成し、イメージアダプタ50を結合するだけで学習済みモデル70を生成できる。その結果、ターゲットモデル40を生成する作業負荷が低減され得る。なお、事前学習において、実情報若しくは疑似情報、又はこれらを併せた情報が教師データとして用いられ得る。

【0025】

ベースモデル30及びターゲットモデル40は、複数の層を有するCNN(Convolution Neural Network)として構成される。ベースモデル30及びターゲットモデル40に入力された情報に対して、CNNの各層において所定の重みづけ係数に基づく畳み込みが実行される。ベースモデル30及びターゲットモデル40の学習において、重みづけ係数が更新される。ベースモデル30及びターゲットモデル40は、VGG16又はResNet50によって構成されてもよい。ベースモデル30及びターゲットモデル40は、これらの例に限られず、他の種々のモデルとして構成されてもよい。

【0026】

10

20

30

40

50

ベースモデル 30 は、第 1 ベースモデル 31 と第 2 ベースモデル 32 とを含むとする。ターゲットモデル 40 は、第 1 ターゲットモデル 41 と第 2 ターゲットモデル 42 とを含むとする。第 1 ベースモデル 31 及び第 1 ターゲットモデル 41 は、バックボーンとも称される。第 2 ベースモデル 32 及び第 2 ターゲットモデル 42 は、ヘッドとも称される。ベースモデル 30 及びターゲットモデル 40 は、バックボーンとヘッドとを含む。なお、ターゲットモデル 40 に含まれる学習済みモデルのそれぞれは、ベースモデル 30 に含まれる学習済みモデルと異なってもよい。例えば、ターゲットモデル 40 に含まれる学習済みモデルのそれぞれは、ベースモデル 30 に含まれる学習済みモデルのそれぞれと、異なる学習処理が実行されていてもよい。より具体的には、互いに異なる情報を含む教師データによって学習処理が実行されていてもよい。なお、ターゲットモデル 40 に含まれる学習前モデルは、ベースモデル 30 に含まれる学習前モデルと、同一のモデルでもよい。

10

【0027】

バックボーンは、入力情報の特徴量を抽出した結果を出力するように構成される。特徴量は、例えば学習対象の外観の特徴を数値として表す。ヘッドは、バックボーンの出力に基づいて入力情報についての所定の判断を行うように構成される。具体的に、ヘッドは、バックボーンが出力した入力情報の特徴量に基づいて、入力情報に含まれる認識対象の認識結果を出力してよい。つまり、ヘッドは、所定の判断として、認識対象の認識を実行するように構成される。例えば、馬とシマウマとを見分けるタスクにおいて、特徴量は、体表面における縞模様の面積の割合を表すパラメータであり得る。所定の判断は、体表面における縞模様の面積の割合を閾値と比較して認識対象が馬であるかシマウマであるか判断

20

【0028】

イメージアダプタ 50 は、図 3 に例示されるように、複数の層を有する CNN として構成されてよい。イメージアダプタ 50 は、ベースモデル 30 又はターゲットモデル 40 に入力される情報を、ベースモデル 30 又はターゲットモデル 40 に入力される前に変換可能に構成される。イメージアダプタ 50 は、図 3 においてターゲットモデル 40 の入力側に結合されているが、ベースモデル 30 の入力側にも結合され得る。

30

【0029】

図 3 の例において、“Conv” と記載されているブロックは、畳み込みを実行することを表す。畳み込みは、ダウンサンプリングとも称される。また、“Conv Trans” と記載されているブロックは、転置畳み込みを実行することを表す。転置畳み込みは、アップサンプリングとも称される。転置畳み込みは、逆畳み込みと称されることもある。“Conv 4x4” と記載されているブロックは、2次元データに対して畳み込みを実行するために用いられるフィルタのサイズが 4 × 4 であることを表す。フィルタは、カーネルとも称され、ブロックに入力された情報の畳み込み又は逆畳み込みを実行する際の重みづけ係数の組に対応する。“Conv Trans 4x4” と記載されているブロックは、2次元データに対して転置畳み込みを実行するために用いられるフィルタのサイズが 4 × 4 であることを表す。“stride 2” と記載されているブロックは、畳み込み又は転置畳み込みを実行する際にフィルタを 2 要素ずつシフトさせることを表す。逆に、“stride 2” が記載されていないブロックは、畳み込み又は転置畳み込みを実行する際にフィルタを 1 要素ずつシフトさせることを表す。

40

【0030】

イメージアダプタ 50 は、ベースモデル 30 の入力側に結合された場合、学習のために入力される疑似情報又は実情報を変換してベースモデル 30 に対して出力する。疑似情報又は実情報が画像である場合、イメージアダプタ 50 は、入力される画像を変換してベースモデル 30 に対して出力する。イメージアダプタ 50 は、ターゲットモデル 40 の入力側に結合された場合、学習済みモデル 70 に入力される入力情報に含まれる認識対象の画像を変換して出力する。また、イメージアダプタ 50 は、入力される画像の態様を変換し

50

て出力してもよい。イメージアダプタ50は、入力される画像の態様を、例えば画像のエッジを強調したり、影となっている部分を明るくした態様に変換したりして出力してもよいがこれに限られない。イメージアダプタ50は、接続されるターゲットモデル40が、タスクを正しく処理できる態様に変換させる。例えば、タスクが画像に含まれる物体の認識である場合には、ベースモデル30又はターゲットモデル40が、認識対象を正しく認識した結果を出力できるように、態様を変換する。

【0031】

(学習済みモデル生成装置20の動作例)

学習済みモデル生成装置20の制御部22は、例えば図4に模式的に示される動作を実行することによって学習済みモデル70を生成できる。以下、図4を参照して学習済みモデル生成装置20の動作が説明される。

10

【0032】

制御部22は、第1ステップとして、少なくとも1つのベースモデル30を生成する。具体的に、制御部22は、情報生成部26から教師データとして第1疑似情報を取得する。制御部22は、第1疑似情報に基づいて学習することによって、ベースモデル30を生成する。制御部22は、学習中のベースモデル30から出力される情報が第1疑似情報に含まれる学習対象を表す情報となる確率を高めるように、ベースモデル30を更新する。制御部22は、ベースモデル30の重みづけ係数を更新することによってベースモデル30を更新してよい。学習開始前の状態において、ベースモデル30は、あらかじめ定められた初期状態とされてよい。つまり、ベースモデル30の重みづけ係数は、あらかじめ定められた初期値とされてよい。制御部22は、第1ステップとして説明した動作を実行することによって、第1疑似情報に基づく学習によってベースモデル30を生成できる。ベースモデル30を生成するための学習は、後述する第2ステップのイメージアダプタ50を生成するための学習に先立って実行されることから、事前学習であるともいえる。なお、第1ステップにおいて、制御部22が、情報生成部26から教師データとして第1疑似情報を取得するものとして説明したが、これに限られない。教師データとしては、第1疑似情報だけでなく第1実情報が利用され得る。また、教師データとして、第2疑似情報又は第2実情報が利用されてもよい。

20

【0033】

本実施形態において、制御部22は、 x 個のベースモデル30を生成する。 x 個のベースモデル30は、1番目のベースモデル301から x 番目のベースモデル30 x までとして区別される。制御部22は、各ベースモデル30を生成するための学習に用いる第1疑似情報として互いに異なる情報を取得する。また、1番目のベースモデル301は、第1ベースモデル311と第2ベースモデル321とを含む。 x 番目のベースモデル30 x は、第1ベースモデル31 x と第2ベースモデル32 x とを含む。

30

【0034】

制御部22は、第2ステップとして、イメージアダプタ50を生成する。具体的に、制御部22は、情報生成部26から教師データとして実情報を更に取得してよい。制御部22は、第1ステップで生成した学習済みのベースモデル30にイメージアダプタ50を結合した状態で、第1疑似情報及び実情報に基づいて学習し、イメージアダプタ50を更新する。制御部22は、イメージアダプタ50の重みづけ係数を更新することによってイメージアダプタ50を更新してよい。制御部22は、各ベースモデル30を生成するための学習に用いる実情報として互いに異なる情報を取得する。学習開始前の状態において、ベースモデル30に結合されるイメージアダプタ50は、あらかじめ定められた初期状態とされてよい。つまり、イメージアダプタ50の重みづけ係数は、あらかじめ定められた初期値とされてよい。学習によって更新する対象とされている学習中のイメージアダプタ50aは、黒塗りの矩形で表されている。なお、第2ステップにおいて、制御部22は、第1ステップで生成した学習済みのベースモデル30にイメージアダプタ50を結合した状態で、第1疑似情報及び実情報に基づいて学習し、イメージアダプタ50を更新するものとして説明したが、これに限られない。制御部22は、第1疑似情報及び実情報の一方の

40

50

みに基づいて学習を行い、イメージアダプタ 50 を更新してもよい。

【0035】

制御部 22 は、学習中のイメージアダプタ 50 a を x 個のベースモデル 30 の各々に結合した状態で、各ベースモデル 30 に対応する第 1 疑似情報又は実情報に基づいて学習する。言い換えれば、制御部 22 は、学習中のイメージアダプタ 50 a に第 1 疑似情報及び実情報をそれぞれ入力し、学習中のイメージアダプタ 50 a の出力を x 個のベースモデル 30 の各々に入力させて学習する。制御部 22 は、学習によってイメージアダプタ 50 を更新することによってイメージアダプタ 50 を生成する。制御部 22 は、イメージアダプタ 50 を経由して第 1 疑似情報を入力した各ベースモデル 30 から出力される情報と、イメージアダプタ 50 を経由して実情報を入力した各ベースモデル 30 から出力される情報とが近づくように、イメージアダプタ 50 を更新する。制御部 22 は、イメージアダプタ 50 を経由して第 1 疑似情報を入力した各ベースモデル 30 から出力される情報と、イメージアダプタ 50 を経由して実情報を入力した各ベースモデル 30 から出力される情報とが一致する確率を高めるようにイメージアダプタ 50 を更新してもよい。制御部 22 は、学習によって各ベースモデル 30 をイメージアダプタ 50 とともに更新してもよいし、イメージアダプタ 50 のみを更新してもよい。

10

【0036】

制御部 22 は、1 個のベースモデル 30 に学習中のイメージアダプタ 50 a を結合した組み合わせ 1 つずつについて学習を実行してもよい。制御部 22 は、1 個のベースモデル 30 に学習中のイメージアダプタ 50 a を結合した組み合わせを複数まとめて並列に学習を実行してもよい。

20

【0037】

制御部 22 は、第 2 ステップとして説明した動作を実行することによって、第 1 疑似情報及び実情報に基づく学習によってイメージアダプタ 50 を生成できる。イメージアダプタ 50 を生成するための学習は、後述する第 3 ステップのターゲットモデル 40 を生成するための学習と独立に実行され得る。

【0038】

制御部 22 は、第 3 ステップとして、ターゲットモデル 40 を生成する。具体的に、制御部 22 は、情報生成部 26 から教師データとして第 2 疑似情報を取得する。制御部 22 は、第 2 疑似情報として、ベースモデル 30 を生成するための学習に用いた第 1 疑似情報と同一又は関連するタスクの情報を取得する。制御部 22 は、第 2 疑似情報に基づいて学習することによって、ターゲットモデル 40 を生成する。制御部 22 は、第 2 疑似情報をイメージアダプタ 50 に入力して変換せずにターゲットモデル 40 に入力する。制御部 22 は、学習中のターゲットモデル 40 から出力される情報が第 2 疑似情報に含まれる学習対象を表す情報となる確率を高めるように、ターゲットモデル 40 を更新する。制御部 22 は、ターゲットモデル 40 の重みづけ係数を更新することによってターゲットモデル 40 を更新してよい。学習開始前の状態において、ターゲットモデル 40 は、あらかじめ定められた初期状態とされてよい。つまり、ターゲットモデル 40 の重みづけ係数は、あらかじめ定められた初期値とされてよい。学習によって更新する対象とされているターゲットモデル 40 は、学習中の第 1 ターゲットモデル 41 a と第 2 ターゲットモデル 42 a とを含み、黒塗りの矩形で表されている。制御部 22 は、第 3 ステップとして説明した動作を実行することによって、第 2 疑似情報に基づく学習によってターゲットモデル 40 を生成できる。なお、第 3 ステップにおいて、制御部 22 は、情報生成部 26 から教師データとして第 2 疑似情報を取得するものとして説明したが、これに限られない。教師データとして、第 2 疑似情報だけでなく第 2 実情報が利用されてもよい。また、第 3 ステップにおいて、制御部 22 は、第 2 疑似情報を変換せずにターゲットモデル 40 に入力して、ターゲットモデル 40 を更新するものとしたが、これに限られない。制御部 22 は、ターゲットモデル 40 とイメージアダプタ 50 とを結合させて、第 2 疑似情報若しくは第 2 実情報、又はこれらの両方を用いて学習することにより、ターゲットモデル 40 とイメージアダプタ 50 を更新してもよい。

30

40

50

【 0 0 3 9 】

制御部 2 2 は、第 4 ステップとして、イメージアダプタ 5 0 をターゲットモデル 4 0 に結合することによって、学習済みモデル 7 0 を生成する。具体的に、制御部 2 2 は、第 2 ステップで生成した学習済みのイメージアダプタ 5 0 b を、第 3 ステップで生成した学習済みの第 1 ターゲットモデル 4 1 b と第 2 ターゲットモデル 4 2 b とを含むターゲットモデル 4 0 に結合する。つまり、制御部 2 2 は、第 2 ステップで生成したイメージアダプタ 5 0 を転移してターゲットモデル 4 0 に結合する。なお、第 4 ステップでは、第 3 ステップで生成したターゲットモデル 4 0 とイメージアダプタ 5 0 とを結合するものとして説明したが、これに限られない。ターゲットモデル 4 0 としては、第 1 ステップで生成したベースモデル 3 0 を利用してもよい。この場合、第 3 ステップが実行されなくてもよい。

10

【 0 0 4 0 】

< 学習済みモデル生成方法 >

学習済みモデル生成装置 2 0 の制御部 2 2 は、以上述べてきた動作を、図 5 に例示されるフローチャートの手順を含む学習済みモデル生成方法として実行してよい。学習済みモデル生成方法は、制御部 2 2 を構成するプロセッサに実行させる学習済みモデル生成プログラムとして実現されてもよい。学習済みモデル生成プログラムは、非一時的なコンピュータ読み取り可能な媒体に格納されてよい。

【 0 0 4 1 】

制御部 2 2 は、複数のベースモデル 3 0 を取得する（ステップ S 1）。制御部 2 2 は、複数のベースモデル 3 0 を第 1 疑似情報に基づく学習によって生成してもよいし、外部装置から取得してもよい。制御部 2 2 は、イメージアダプタ 5 0 を生成するための学習に用いる複数のベースモデル 3 0 のみを取得する。

20

【 0 0 4 2 】

制御部 2 2 は、複数のベースモデル 3 0 から少なくとも 1 つのベースモデル 3 0 を選択する（ステップ S 2）。制御部 2 2 は、学習対象の情報を取得する（ステップ S 3）。具体的に、制御部 2 2 は、選択したベースモデル 3 0 を生成するための学習で用いられた疑似情報と同一又は関連するタスクの実情報を学習対象の情報として取得してよい。

【 0 0 4 3 】

制御部 2 2 は、イメージアダプタ 5 0 を選択したベースモデル 3 0 に結合した状態で学習対象の情報に基づいて学習することによってイメージアダプタ 5 0 を生成する（ステップ S 4）。具体的に、制御部 2 2 は、学習対象の情報として実情報をイメージアダプタ 5 0 に入力する。イメージアダプタ 5 0 で実情報から変換された情報は、選択したベースモデル 3 0 に入力される。制御部 2 2 は、選択したベースモデル 3 0 から出力される情報に基づいてイメージアダプタ 5 0 を更新することによってイメージアダプタ 5 0 を生成する。

30

【 0 0 4 4 】

制御部 2 2 は、全ベースモデル 3 0 が選択済みとなったか判定する（ステップ S 5）。制御部 2 2 は、全ベースモデル 3 0 が選択済みとなっていない場合（ステップ S 5：NO）、つまり少なくとも 1 つのベースモデル 3 0 が未選択である場合、ステップ S 2 の手順に戻って未選択のベースモデル 3 0 を選択する。

【 0 0 4 5 】

制御部 2 2 は、全ベースモデル 3 0 が選択済みとなった場合（ステップ S 5：YES）、認識対象の情報を取得する（ステップ S 6）。具体的に、制御部 2 2 は、選択したベースモデル 3 0 を生成するための学習で用いられた第 1 疑似情報と同一又は関連するタスクの第 2 疑似情報を認識対象の情報として取得してよい。

40

【 0 0 4 6 】

制御部 2 2 は、認識対象の情報に基づいて学習することによってターゲットモデル 4 0 を生成する（ステップ S 7）。制御部 2 2 は、イメージアダプタ 5 0 とターゲットモデル 4 0 とを結合する（ステップ S 8）。制御部 2 2 は、以上の手順を実行することによって、イメージアダプタ 5 0 とターゲットモデル 4 0 とを結合した学習済みモデル 7 0 を生成できる。制御部 2 2 は、ステップ S 8 の手順の実行後、図 5 のフローチャートの手順の実

50

行を終了する。制御部 22 は、ステップ S8 の手順の実行後、生成した学習済みモデル 70 に入力情報を入力し、入力情報に含まれる認識対象の認識精度を学習済みモデル 70 の出力に基づいて評価してもよい。制御部 22 は、ステップ S8 の手順の実行後、生成した学習済みモデル 70 を後述するロボット制御装置 110 (図 6 参照) に出力してもよい。

【0047】

<小括>

以上述べてきたように、本実施形態に係る学習済みモデル生成装置 20 は、ベースモデル 30 に結合した状態における学習によって生成したイメージアダプタ 50 を、別の学習によって新たに生成したターゲットモデル 40 に結合することによって、学習済みモデル 70 を生成できる。学習済みモデル生成装置 20 は、実情報又は疑似情報に基づいて学習することによって、イメージアダプタ 50 を生成する。学習済みモデル生成装置 20 は、疑似情報だけに基づいて学習することによってターゲットモデル 40 を生成する。実情報又は疑似情報に基づく学習によって生成されたイメージアダプタ 50 を結合した学習済みモデル 70 による認識精度は、ターゲットモデル 40 だけの場合に比べて認識精度が向上する。したがって、あらかじめ実情報又は疑似情報に基づく学習によってイメージアダプタ 50 を生成しておけば、ターゲットモデル 40 にイメージアダプタ 50 を結合させることで、高い認識精度が期待される。

【0048】

また、学習済みモデル生成装置 20 は、ターゲットモデル 40 を新規に生成した場合でもイメージアダプタ 50 を結合して学習済みモデル 70 を生成することによって、認識精度を高めることができる。言い換えれば、ターゲットモデル 40 にベースモデル 30 を転移しなくても、学習済みモデル 70 の認識精度を高めることができる。

【0049】

ここで、ベースモデル 30 の少なくとも一部を転移してターゲットモデル 40 を生成する場合、ベースモデル 30 を転移する動作そのものが学習済みモデル 70 の生成の制約となり得る。例えば、ベースモデル 30 の少なくとも一部を転移することによって、サービスのエンドユーザが所望の認識対象に合わせてターゲットモデル 40 を生成する場合に、ターゲットモデル 40 が所望の認識対象に合いにくくなることがある。これに対して、本実施形態に係る学習済みモデル生成装置 20 は、ターゲットモデル 40 にベースモデル 30 を転移しなくてもよいことによって、ターゲットモデル 40 をエンドユーザが所望するモデルに合わせやすくできる。

【0050】

複数のベースモデル 30 の各々に結合された状態で学習することによって生成されるイメージアダプタ 50 は、サービス提供者が事前に学習することによって生成することから、上流タスクとも称される。一方で、上流タスクからイメージアダプタ 50 を転移して新たに生成したターゲットモデル 40 に結合することによって生成される学習済みモデル 70 は、サービスのエンドユーザが所望の認識対象に合わせて生成することから、下流タスクとも称される。

【0051】

下流タスクにおいて、データ取得の手間が少なく、又は、短い学習時間で学習済みモデル 70 を生成して早くシステムを稼働させることが求められる。一方で、上流タスクにおいて、転移学習が高速かつ汎化性能が高い高品質なメタモデルを提供するために、事前に多くのデータと計算リソースをかけることができる。本実施形態に係る学習済みモデル生成装置 20 は、多くのデータと計算リソースをかけて上流タスクを生成することによって、少ない負荷で下流タスクを生成できる結果、システムを早期に稼働できる。

【0052】

下流タスクで実情報に基づいて学習しない場合、Sim-to-Realと呼ばれるドメインギャップが問題になり得る。本実施形態に係る学習済みモデル生成装置 20 は、ドメイン適応のためのイメージアダプタ 50 を上流タスクから下流タスクに転移することによって、実情報に基づいて学習していない下流タスクにおいても実情報に対する認識精度を高めるこ

10

20

30

40

50

とができる。具体的に、上流タスクにおいてイメージアダプタ 50 は、疑似情報に対する認識精度を高めるように生成された複数のベースモデル 30 の各々の、実情報に対する認識精度を高めるように生成される。イメージアダプタ 50 は、複数のベースモデル 30 の各々の認識精度を高めるように生成されることによって、下流タスクで新たに生成されるターゲットモデル 40 の認識精度も高めることができる。イメージアダプタ 50 が複数のベースモデル 30 の各々の認識精度を高めるように生成されることは、イメージアダプタ 50 の一般化、又は、Generalized Image Adaptor (G I A) とも称される。G I A によって、複数のベースモデル 30 でパフォーマンスが高くなる共通の特徴を強調しつつノイズ源となる特徴を抑制するような、タスクに根本的に有効な画質改善が獲得され得る。この画質改善によって、Sim-to-Real 問題に対する改善のみならず、様々なベースモデルによる認識精度の向上が期待できる。

10

【0053】

本実施形態において、学習済みモデル生成装置 20 は、上流タスクにおいてイメージアダプタ 50 を生成し、上流タスクで生成したイメージアダプタ 50 を下流タスクに転移してよい。学習済みモデル生成装置 20 は、下流タスクだけで第 2 実情報又は第 2 疑似情報に基づく学習によってイメージアダプタ 50 を生成してもよい。

【0054】

< 認識精度の比較 >

疑似情報である生成画像のみに基づく学習によって生成したモデルで実画像を含む入力情報から認識対象を認識する場合、生成画像と実画像との差異に起因して、認識精度は低下する。具体的に、生成画像に対して 100% に近い確率で認識対象を認識できるモデルにおいて、実画像に対して認識対象を認識できる確率は 70% 程度に低下し得る。

20

【0055】

本実施形態に係る学習済みモデル 70 は、複数のベースモデル 30 の各々に結合した状態で学習することによって生成したイメージアダプタ 50 を、ターゲットモデル 40 に結合したモデルとして生成される。イメージアダプタ 50 は、生成画像と実画像との差異に起因する認識結果の誤差を修正できる。その結果、実画像に対して認識対象を認識できる確率が 80% 程度にまで高められ得る。つまり、イメージアダプタ 50 を結合した場合、イメージアダプタ 50 を結合しない場合と比較して、認識対象を認識できる確率が高められ得る。また、本実施形態に係る学習済みモデル 70 は、ベースモデル 30 を転移せずに生成される。つまり、ベースモデル 30 を転移しなくても実画像に対して認識対象を認識できる確率が高められ得る。ベースモデル 30 を転移しなくてもよいことによって、ターゲットモデル 40 は、エンドユーザが所望するモデルに合いやすくなる。

30

【0056】

(ロボット制御システム 100 の構成例)

図 6 に示されるように、一実施形態に係るロボット制御システム 100 は、ロボット 2 と、ロボット制御装置 110 とを備える。本実施形態において、ロボット 2 は、作業対象物 8 を作業開始地点 6 から作業目標地点 7 へ移動させるとする。つまり、ロボット制御装置 110 は、作業対象物 8 が作業開始地点 6 から作業目標地点 7 へ移動するようにロボット 2 を制御する。作業対象物 8 は、作業対象とも称される。ロボット制御装置 110 は、ロボット 2 が作業を実施する空間に関する情報に基づいて、ロボット 2 を制御する。空間に関する情報は、空間情報とも称される。

40

【0057】

< ロボット 2 >

ロボット 2 は、アーム 2 A と、エンドエフェクタ 2 B とを備える。アーム 2 A は、例えば、6 軸又は 7 軸の垂直多関節ロボットとして構成されてよい。アーム 2 A は、3 軸又は 4 軸の水平多関節ロボット又はスカラロボットとして構成されてもよい。アーム 2 A は、2 軸又は 3 軸の直交ロボットとして構成されてもよい。アーム 2 A は、パラレルリンクロボット等として構成されてもよい。アーム 2 A を構成する軸の数は、例示したものに限られない。言い換えれば、ロボット 2 は、複数の関節で接続されるアーム 2 A を有し、関節

50

の駆動によって動作する。

【0058】

エンドエフェクタ2Bは、例えば、作業対象物8を把持できるように構成される把持ハンドを含んでよい。把持ハンドは、複数の指を有してよい。把持ハンドの指の数は、2つ以上であってよい。把持ハンドの指は、1つ以上の関節を有してよい。エンドエフェクタ2Bは、作業対象物8を吸着できるように構成される吸着ハンドを含んでもよい。エンドエフェクタ2Bは、作業対象物8を掬うことができるように構成される掬いハンドを含んでもよい。エンドエフェクタ2Bは、ドリル等の工具を含み、作業対象物8に穴を開ける作業等の種々の加工を実施できるように構成されてもよい。エンドエフェクタ2Bは、これらの例に限られず、他の種々の動作ができるように構成されてよい。図1に例示される構成において、エンドエフェクタ2Bは、把持ハンドを含むとする。

10

【0059】

ロボット2は、アーム2Aを動作させることによって、エンドエフェクタ2Bの位置を制御できる。エンドエフェクタ2Bは、作業対象物8に対して作用する方向の基準となる軸を有してもよい。エンドエフェクタ2Bが軸を有する場合、ロボット2は、アーム2Aを動作させることによって、エンドエフェクタ2Bの軸の方向を制御できる。ロボット2は、エンドエフェクタ2Bが作業対象物8に作用する動作の開始及び終了を制御する。ロボット2は、エンドエフェクタ2Bの位置、又は、エンドエフェクタ2Bの軸の方向を制御しつつ、エンドエフェクタ2Bの動作を制御することによって、作業対象物8を動かしたり加工したりすることができる。図1に例示される構成において、ロボット2は、作業開始地点6でエンドエフェクタ2Bに作業対象物8を把持させ、エンドエフェクタ2Bを作業目標地点7へ移動させる。ロボット2は、作業目標地点7でエンドエフェクタ2Bに作業対象物8を解放させる。このようにすることで、ロボット2は、作業対象物8を作業開始地点6から作業目標地点7へ移動させることができる。

20

【0060】

<センサ3>

図2に示されるように、ロボット制御システム100は、更にセンサ3を備える。センサ3は、ロボット2の物理情報を検出する。ロボット2の物理情報は、ロボット2の各構成部の現実の位置若しくは姿勢、又は、ロボット2の各構成部の速度若しくは加速度に関する情報を含んでよい。ロボット2の物理情報は、ロボット2の各構成部に作用する力に関する情報を含んでよい。ロボット2の物理情報は、ロボット2の各構成部を駆動するモータに流れる電流又はモータのトルクに関する情報を含んでよい。ロボット2の物理情報は、ロボット2の実際の動作の結果を表す。つまり、ロボット制御システム100は、ロボット2の物理情報を取得することによって、ロボット2の実際の動作の結果を把握することができる。

30

【0061】

センサ3は、ロボット2の物理情報として、ロボット2に作用する力、分布圧、若しくはすべり等を検出する力覚センサ又は触覚センサを含んでよい。センサ3は、ロボット2の物理情報として、ロボット2の位置若しくは姿勢、又は、速度若しくは加速度を検出するモーションセンサを含んでよい。センサ3は、ロボット2の物理情報として、ロボット2を駆動するモータに流れる電流を検出する電流センサを含んでよい。センサ3は、ロボット2の物理情報として、ロボット2を駆動するモータのトルクを検出するトルクセンサを含んでよい。

40

【0062】

センサ3は、ロボット2の関節、又は、関節を駆動する関節駆動部に設置されてよい。センサ3は、ロボット2のアーム2A又はエンドエフェクタ2Bに設置されてもよい。

【0063】

センサ3は、検出したロボット2の物理情報をロボット制御装置110に出力する。センサ3は、所定のタイミングでロボット2の物理情報を検出して出力する。センサ3は、ロボット2の物理情報を時系列データとして出力する。

50

【 0 0 6 4 】

< カメラ 4 >

図 1 に示される構成例において、ロボット制御システム 1 0 0 は、2 台のカメラ 4 を備えとする。カメラ 4 は、ロボット 2 の動作に影響を及ぼす可能性がある影響範囲 5 に位置する物品又は人間等を撮影する。カメラ 4 が撮影する画像は、モノクロの輝度情報を含んでもよいし、RGB (Red, Green and Blue) 等で表される各色の輝度情報を含んでもよい。影響範囲 5 は、ロボット 2 の動作範囲を含む。影響範囲 5 は、ロボット 2 の動作範囲を更に外側に広げた範囲であるとする。影響範囲 5 は、ロボット 2 の動作範囲の外側から動作範囲の内側へ向かって移動する人間等がロボット 2 の動作範囲の内側に入るまでにロボット 2 を停止できるように設定されてよい。影響範囲 5 は、例えば、ロボット 2 の動作範囲の境界から所定距離だけ外側まで拡張された範囲に設定されてもよい。カメラ 4 は、ロボット 2 の影響範囲 5 若しくは動作範囲又はこれらの周辺の領域を俯瞰的に撮影できるように設置されてもよい。カメラ 4 の数は、2 つに限られず、1 つであってもよいし、3 つ以上であってもよい。

10

【 0 0 6 5 】

< ロボット制御装置 1 1 0 >

ロボット制御装置 1 1 0 は、学習済みモデル生成装置 2 0 で生成された学習済みモデル 7 0 を取得する。ロボット制御装置 1 1 0 は、カメラ 4 で撮影した画像と学習済みモデル 7 0 とに基づいて、ロボット 2 が作業を実施する空間に存在する、作業対象物 8、又は作業開始地点 6 若しくは作業目標地点 7 等を認識する。言い換えれば、ロボット制御装置 1 1 0 は、カメラ 4 で撮影した画像に基づいて作業対象物 8 等を認識するために生成された学習済みモデル 7 0 を取得する。ロボット制御装置 1 1 0 は、認識装置とも称される。

20

【 0 0 6 6 】

ロボット制御装置 1 1 0 は、種々の機能を実行するための制御及び処理能力を提供するために、少なくとも 1 つのプロセッサを含んで構成されてよい。ロボット制御装置 1 1 0 の各構成部は、少なくとも 1 つのプロセッサを含んで構成されてもよい。ロボット制御装置 1 1 0 の各構成部のうち複数の構成部が 1 つのプロセッサで実現されてもよい。ロボット制御装置 1 1 0 の全体が 1 つのプロセッサで実現されてもよい。プロセッサは、ロボット制御装置 1 1 0 の種々の機能を実現するプログラムを実行しうる。プロセッサは、単一の集積回路として実現されてよい。集積回路は、IC (Integrated Circuit) とも称される。プロセッサは、複数の通信可能に接続された集積回路及びディスクリート回路として実現されてよい。プロセッサは、他の種々の既知の技術に基づいて実現されてよい。

30

【 0 0 6 7 】

ロボット制御装置 1 1 0 は、記憶部を備えてよい。記憶部は、磁気ディスク等の電磁記憶媒体を含んでもよいし、半導体メモリ又は磁気メモリ等のメモリを含んでもよい。記憶部は、各種情報及びロボット制御装置 1 1 0 で実行されるプログラム等を格納する。記憶部は、非一時的な読み取り可能媒体として構成されてもよい。記憶部は、ロボット制御装置 1 1 0 のワークメモリとして機能してよい。記憶部の少なくとも一部は、ロボット制御装置 1 1 0 とは別体として構成されてもよい。

40

【 0 0 6 8 】

(ロボット制御システム 1 0 0 の動作例)

ロボット制御装置 1 1 0 (認識装置) は、学習済みモデル 7 0 をあらかじめ取得する。ロボット制御装置 1 1 0 は、学習済みモデル 7 0 を記憶部に格納してよい。ロボット制御装置 1 1 0 は、カメラ 4 から作業対象物 8 を撮影した画像を取得する。ロボット制御装置 1 1 0 は、作業対象物 8 を撮影した画像を入力情報として学習済みモデル 7 0 に入力する。ロボット制御装置 1 1 0 は、学習済みモデル 7 0 から入力情報の入力に応じて出力される出力情報を取得する。ロボット制御装置 1 1 0 は、出力情報に基づいて作業対象物 8 を認識し、作業対象物 8 を把持したり移動したりする作業を実行する。

【 0 0 6 9 】

< 小括 >

50

以上述べてきたように、ロボット制御システム 100 は、学習済みモデル生成装置 20 から学習済みモデル 70 を取得し、学習済みモデル 70 によって作業対象物 8 を認識できる。

【0070】

(他の実施形態)

以下、他の実施形態が説明される。

【0071】

<損失関数>

学習済みモデル生成装置 20 は、生成した学習済みモデル 70 に対して入力情報を入力した場合における出力が教師データを入力した場合における出力に近づくように損失関数を設定してよい。本実施形態において、損失関数として交差エントロピーが用いられ得る。交差エントロピーは、2つの確率分布の間の関係を表す値として算出される。具体的に、本実施形態において、交差エントロピーは、入力する疑似情報又は実情報と、バックボーン、ヘッド又はアダプタとの間の関係を表す値として算出される。

10

【0072】

学習済みモデル生成装置 20 は、損失関数の値が小さくなるように学習する。損失関数の値が小さくなるように学習することによって生成した学習済みモデル 70 において、入力情報の入力に応じた出力が教師データの入力に応じた出力に近づき得る。

【0073】

学習済みモデル生成装置 20 の制御部 22 は、イメージアダプタ 50 がベースモデル 30 に結合した状態で入力情報と同一又は関連するタスクの損失関数を最適化するように学習することによって、イメージアダプタ 50 を生成してよい。損失関数の最適化は、例えば損失関数の値の最小化であってよい。入力情報と同一又は関連するタスクの損失関数は、ベースモデル 30 の損失関数を含む。一方で、制御部 22 は、イメージアダプタ 50 がベースモデル 30 に結合した状態で入力情報と同一又は関連するタスク以外の損失関数を最適化するように学習することによって、イメージアダプタ 50 を生成してもよい。入力情報と同一又は関連するタスク以外の損失関数は、ベースモデル 30 の損失関数以外の他の種々の有意な損失関数を含む。ベースモデル 30 の損失関数以外の損失関数として、例えば、Discrimination Loss、又は、Contrastive Loss が利用され得る。Discrimination Loss は、生成画像についての真贋を、完全に真であることを表す 1 から完全に贋であることを表す 0 までの間の数値でラベル付けして学習するために用いられる損失関数である。制御部 22 は、Discrimination Loss を損失関数とする学習において、イメージアダプタ 50 に入力情報として画像が入力された場合にイメージアダプタ 50 が出力する画像を、正解をラベルとして学習する。このようにすることで、制御部 22 は、疑似情報に基づく学習によって生成されたベースモデル 30 にとって、実情報としての画像とイメージアダプタ 50 が出力する画像との区別がつかないようにイメージアダプタ 50 を生成できる。

20

30

【0074】

<複数のベースモデル 30 の各々を学習に適用する態様>

学習済みモデル生成装置 20 の制御部 22 は、複数のベースモデル 30 の各々にイメージアダプタ 50 を結合した状態で学習することによって、イメージアダプタ 50 を生成する。つまり、制御部 22 は、複数のベースモデル 30 の各々を、イメージアダプタ 50 を生成するための事前学習に適用する。

40

【0075】

複数のベースモデル 30 が 1 番目のベースモデル 30 1 から x 番目のベースモデル 30 x までを含む場合、制御部 22 は、各ベースモデル 30 を 1 つずつ順番にイメージアダプタ 50 に結合した組み合わせを生成し、各組み合わせの各々について学習してイメージアダプタ 50 を更新することによってイメージアダプタ 50 を生成してよい。つまり、制御部 22 は、複数のベースモデル 30 の各々を 1 つずつ順番に、イメージアダプタ 50 を生成するための事前学習に適用してよい。

50

【 0 0 7 6 】

制御部 2 2 は、ベースモデル 3 0 を事前学習に適用する順番を、ランダムに決定してもよいし所定ルールに基づいて決定してもよい。制御部 2 2 は、複数の組み合わせの各々を適用した複数の事前学習を並列に実行してもよい。つまり、制御部 2 2 は、複数のベースモデル 3 0 を並列に事前学習に適用してもよい。

【 0 0 7 7 】

制御部 2 2 は、複数のベースモデル 3 0 を複数のグループに分類し、各グループを順番にイメージアダプタ 5 0 を生成するための事前学習に適用してもよい。制御部 2 2 は、1 つのグループに複数のベースモデル 3 0 を分類してもよい。この場合、制御部 2 2 は、グループに分類した複数のベースモデル 3 0 を並列に事前学習に適用してもよいし、複数のベースモデル 3 0 の各々を 1 つずつ順番に事前学習に適用してもよい。制御部 2 2 は、各グループに 1 つのベースモデル 3 0 を分類してもよい。制御部 2 2 は、各グループを事前学習に適用する順番を、ランダムに決定してもよいし所定ルールに基づいて決定してもよい。

10

【 0 0 7 8 】

以上、学習済みモデル生成システム 1 及びロボット制御システム 1 0 0 の実施形態を説明してきたが、本開示の実施形態としては、システム又は装置を実施するための方法又はプログラムの他、プログラムが記録された記憶媒体（一例として、光ディスク、光磁気ディスク、CD-ROM、CD-R、CD-RW、磁気テープ、ハードディスク、又はメモリカード等）としての実施態様をとることも可能である。

20

【 0 0 7 9 】

また、プログラムの実装形態としては、コンパイラによってコンパイルされるオブジェクトコード、インタプリタにより実行されるプログラムコード等のアプリケーションプログラムに限定されることはなく、オペレーティングシステムに組み込まれるプログラムモジュール等の形態であっても良い。さらに、プログラムは、制御基板上の CPU においてのみ全ての処理が実施されるように構成されてもされなくてもよい。プログラムは、必要に応じて基板に付加された拡張ボード又は拡張ユニットに実装された別の処理ユニットによってその一部又は全部が実施されるように構成されてもよい。

【 0 0 8 0 】

本開示に係る実施形態について、諸図面及び実施例に基づき説明してきたが、当業者であれば本開示に基づき種々の変形又は改変を行うことが可能であることに注意されたい。従って、これらの変形又は改変は本開示の範囲に含まれることに留意されたい。例えば、各構成部等に含まれる機能等は論理的に矛盾しないように再配置可能であり、複数の構成部等を 1 つに組み合わせたり、或いは分割したりすることが可能である。

30

【 0 0 8 1 】

本開示に記載された構成要件の全て、及び/又は、開示された全ての方法、又は、処理の全てのステップについては、これらの特徴が相互に排他的である組合せを除き、任意の組合せで組み合わせることができる。また、本開示に記載された特徴の各々は、明示的に否定されない限り、同一の目的、同等の目的、又は類似する目的のために働く代替の特徴に置換することができる。したがって、明示的に否定されない限り、開示された特徴の各々は、包括的な一連の同一、又は、均等となる特徴の一例にすぎない。

40

【 0 0 8 2 】

さらに、本開示に係る実施形態は、上述した実施形態のいずれの具体的構成にも制限されるものではない。本開示に係る実施形態は、本開示に記載された全ての新規な特徴、又は、それらの組合せ、あるいは記載された全ての新規な方法、又は、処理のステップ、又は、それらの組合せに拡張することができる。

【 0 0 8 3 】

本開示において「第 1 」及び「第 2 」等の記載は、当該構成を区別するための識別子である。本開示における「第 1 」及び「第 2 」等の記載で区別された構成は、当該構成における番号を交換することができる。例えば、第 1 疑似情報は、第 2 疑似情報と識別子であ

50

る「第1」と「第2」とを交換することができる。識別子の交換は同時に行われる。識別子の交換後も当該構成は区別される。識別子は削除してよい。識別子を削除した構成は、符号で区別される。本開示における「第1」及び「第2」等の識別子の記載のみに基づいて、当該構成の順序の解釈、小さい番号の識別子が存在することの根拠に利用してはならない。

【符号の説明】

【0084】

20 学習済みモデル生成装置(22:制御部、26:情報生成部)
 30 ベースモデル(31:第1ベースモデル(31a:学習中、31b:学習済み)、32:第2ベースモデル(32a:学習中、32b:学習済み)、301~30x:1番目~x番目のベースモデル、311~31x:1番目~x番目の第1ベースモデル、321~32x:1番目~x番目の第2ベースモデル)
 40 ターゲットモデル(41:第1ターゲットモデル(41a:学習中、41b:学習済み)、42:第2ターゲットモデル(42a:学習中、42b:学習済み))
 50 アダプタ(50a:学習中、50b:学習済み)
 70 学習済みモデル
 100 ロボット制御システム(2:ロボット、2A:アーム、2B:エンドエフェクタ、3:センサ、4:カメラ、5:ロボットの影響範囲、6:作業開始台、7:作業目標台、8:作業対象物、110:ロボット制御装置(認識装置))

10

20

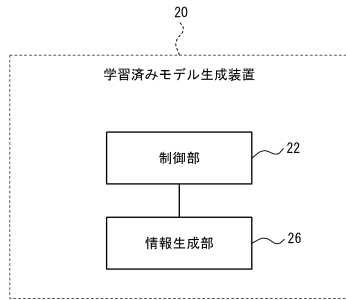
30

40

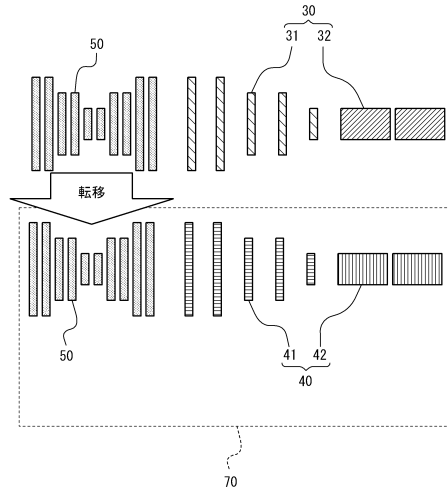
50

【図面】

【図 1】



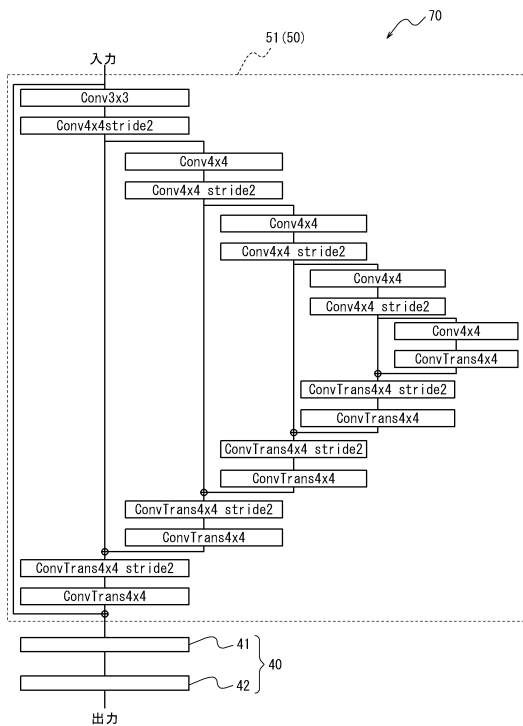
【図 2】



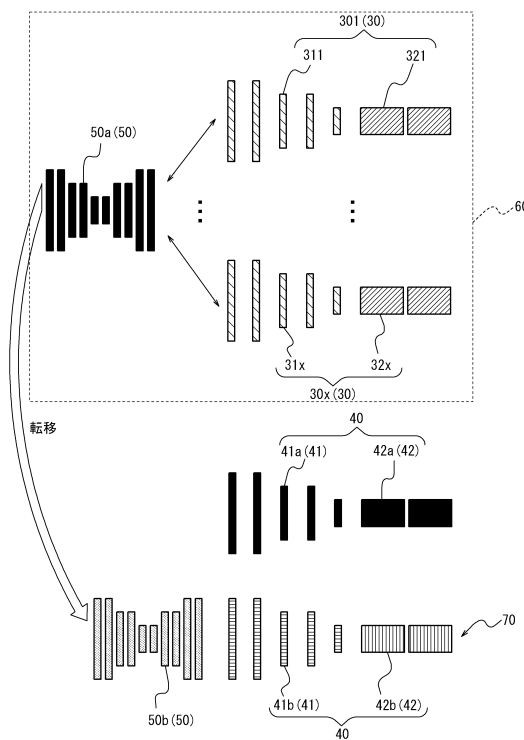
10

20

【図 3】



【図 4】

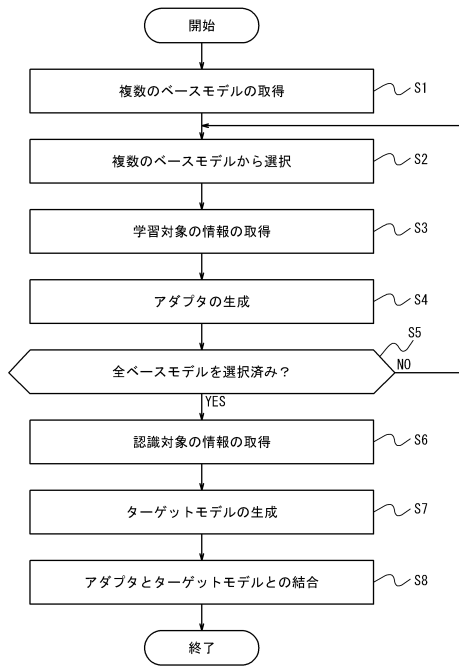


30

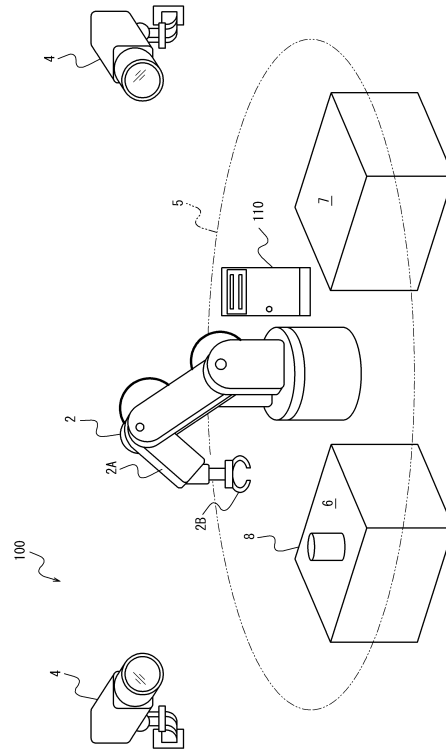
40

50

【図5】



【図6】



10

20

30

40

50

フロントページの続き

- (56)参考文献 特開 2021 - 056785 (JP, A)
特開 2020 - 144700 (JP, A)
国際公開第 2019 / 194256 (WO, A1)
米国特許第 10565471 (US, B1)
米国特許出願公開第 2020 / 0134469 (US, A1)
- (58)調査した分野 (Int.Cl., DB名)
G06T 7/00 - 7/90
G06N 20/00