

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7674211号
(P7674211)

(45)発行日 令和7年5月9日(2025.5.9)

(24)登録日 令和7年4月28日(2025.4.28)

(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 B
G 0 6 T	3/4046(2024.01)	G 0 6 T	3/4046
G 0 6 T	5/50 (2006.01)	G 0 6 T	5/50
G 0 6 N	20/00 (2019.01)	G 0 6 N	20/00
請求項の数 14 (全24頁)			
(21)出願番号	特願2021-156151(P2021-156151)	(73)特許権者	000005108 株式会社日立製作所 東京都千代田区丸の内一丁目6番6号
(22)出願日	令和3年9月24日(2021.9.24)	(74)代理人	110001689 青稜弁理士法人
(65)公開番号	特開2023-47195(P2023-47195A)	(72)発明者	平井 理宇 東京都千代田区丸の内一丁目6番6号 株式会社日立製作所内
(43)公開日	令和5年4月5日(2023.4.5)	(72)発明者	寺田 卓馬 東京都千代田区丸の内一丁目6番6号 株式会社日立製作所内
審査請求日	令和6年2月9日(2024.2.9)	(72)発明者	沖津 潤 東京都千代田区丸の内一丁目6番6号 株式会社日立製作所内
		審査官	小池 正彦
最終頁に続く			

(54)【発明の名称】 学習用データ生成システム、機械学習モデルの学習方法、および学習用データ生成方法

(57)【特許請求の範囲】

【請求項1】
画像合成部を備え、
前記画像合成部は、
背景画像と、第一の画像および該第一の画像に紐づいた正解情報と、第二の画像と、環境情報に紐づいた合成条件を入力とし、
前記合成条件を用いて、前記第一の画像、前記第二の画像、および前記背景画像から合成画像を生成し、
前記合成画像と前記正解情報から、学習用データを生成し、
前記環境情報は、階層構造の分類で定義される、
学習用データ生成システム。

【請求項2】
前記背景画像と、前記第一の画像および前記正解情報と、前記第二の画像には、それぞれ環境情報が紐づけられている、
請求項1記載の学習用データ生成システム。

【請求項3】
前記第一の画像、前記第二の画像、および前記背景画像から合成画像を生成する際に用いる合成条件に紐づいた環境情報は、
前記第一の画像、前記第二の画像、および前記背景画像に紐づいた環境情報と同一または類似のものである。

請求項 2 記載の学習用データ生成システム。

【請求項 4】

合成条件管理部を備え、
前記合成条件管理部は、
前記合成条件を、所定の分類に基づいてデータベース化する、
請求項 1 記載の学習用データ生成システム。

【請求項 5】

画像合成部を備え、
前記画像合成部は、
背景画像と、第一の画像および該第一の画像に紐づいた正解情報と、第二の画像と、環境
情報に紐づいた合成条件を入力とし、
前記合成条件を用いて、前記第一の画像、前記第二の画像、および前記背景画像から合成
画像を生成し、
前記合成画像と前記正解情報から、学習用データを生成し、
前記環境情報は、画像から得られた特徴量で構成される、
学習用データ生成システム。

10

【請求項 6】

画像合成部を備え、
前記画像合成部は、
背景画像と、第一の画像および該第一の画像に紐づいた正解情報と、第二の画像と、環境
情報に紐づいた合成条件を入力とし、
前記合成条件を用いて、前記第一の画像、前記第二の画像、および前記背景画像から合成
画像を生成し、
前記合成画像と前記正解情報から、学習用データを生成し、
環境情報管理部を備え、
前記環境情報管理部は、
画像から特徴量を抽出し、当該特徴量を前記第一の画像、前記第二の画像、および前記
背景画像を合成した合成条件に紐づけてデータベースとする、
学習用データ生成システム。

20

【請求項 7】

前記環境情報管理部は、
未知の環境から得られた前記第一の画像、前記第二の画像、および前記背景画像に関連
する画像から特徴量を抽出して、抽出した特徴量を前記データベースの特徴量と比較し、
前記画像合成部は、
比較結果に基づいて抽出された特徴量に紐づけられた合成条件を用いて、前記未知の環
境から得られた前記第一の画像、前記第二の画像、および前記背景画像の合成を行う、
請求項 6 記載の学習用データ生成システム。

30

【請求項 8】

前記合成条件は、
・合成する前記第一の画像および前記第二の画像の少なくとも一つの、種類および数の少
なくとも一つの指定
・前記背景画像に対する前記第一の画像および前記第二の画像の少なくとも一つの、位置
の指定
・前記背景画像に対する前記第一の画像および前記第二の画像の少なくとも一つの、指定
された位置からのばらつき
・前記第一の画像および前記第二の画像の重なる順序
・前記第一の画像および前記第二の画像の少なくとも一つの、大きさ
・前記第一の画像および前記第二の画像の少なくとも一つの、の向き
・合成画像の明るさ
・合成画像の解像度

40

50

・合成画像上の特殊効果
の少なくとも一つの情報を含むテンプレートとされる、
請求項 1 記載の学習用データ生成システム。

【請求項 9】

前記合成画像上の特殊効果は、前記背景画像、前記第一の画像、および前記第二の画像
の少なくとも一つの定された箇所の画像を変形させる効果である、

請求項 8 記載の学習用データ生成システム。

【請求項 10】

画像合成部を備え、

前記画像合成部は、

背景画像と、第一の画像および該第一の画像に紐づいた正解情報と、第二の画像と、環境
情報に紐づいた合成条件を入力とし、

前記合成条件を用いて、前記第一の画像、前記第二の画像、および前記背景画像から合成
画像を生成し、

前記合成画像と前記正解情報から、学習用データを生成し、

前処理部を備え、

前記前処理部は、

静止画像から、前記背景画像、前記第一の画像、および前記第二の画像を抽出し、

前記第一の画像に前記正解情報を付与して、前記第一の画像に紐づいた正解情報を生成
する、

学習用データ生成システム。

【請求項 11】

前記前処理部は、

前記背景画像と、前記第一の画像および前記正解情報と、前記第二の画像には、それぞ
れ前記静止画像に紐づけられた環境情報を紐づける、

請求項 10 記載の学習用データ生成システム。

【請求項 12】

請求項 1 記載の学習用データ生成システムで生成した学習用データを用いて、機械学習
モデルの機械学習を行う、機械学習モデルの学習方法。

【請求項 13】

環境情報に紐づけられた合成条件を準備する第 1 のステップ、

画像データから抽出した、背景画像と、対象物画像と、非対象物画像を呼び出す第 2 の
ステップ、

前記画像データに適合した環境条件に紐づけられた合成条件を用いて、前記背景画像と
、前記対象物画像と、前記非対象物画像を合成して合成画像を生成する第 3 のステップ、

前記合成画像と前記対象物画像に付された正解情報により学習用データを生成する第 4
のステップ、

を実行し、

前記画像データに適合した環境条件を選択するために、前記画像データの特徴量を用いる、
学習用データ生成方法。

【請求項 14】

環境情報に紐づけられた合成条件を準備する第 1 のステップ、

画像データから抽出した、背景画像と、対象物画像と、非対象物画像を呼び出す第 2 の
ステップ、

前記画像データに適合した環境条件に紐づけられた合成条件を用いて、前記背景画像と
、前記対象物画像と、前記非対象物画像を合成して合成画像を生成する第 3 のステップ、

前記合成画像と前記対象物画像に付された正解情報により学習用データを生成する第 4
のステップ、

を実行し、

前記画像データに適合した環境条件が階層構造の分類で定義される、

10

20

30

40

50

学習用データ生成方法。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、機械学習および機械学習に用いる学習用データ（教師データ）を生成する技術に関する。特に、深層学習（DL：Deep Learning）に用いて好適な、学習用データを容易に生成する技術に関する。

【背景技術】

【0002】

近年、深層学習を用いて機械学習モデルを生成し、画像解析、画像生成、音声認識、言語処理などの処理を行わせる技術が注目されている。機械学習モデルとしては、例えば、ディープニューラルネットワーク（DNN：Deep Neural Network）のような公知の構成が利用されている。

【0003】

画像解析を例にとると、分類、物体検出、骨格検出、セグメンテーション等の処理が知られている。分類とは、画像中の対象物にラベル付けすることである。物体検出とは、画像中の対象物にラベル付けするとともに位置（XY座標）を検出することである。骨格検出とは、画像中の対象物の特定の箇所にラベル付けするとともに位置を検出することである。セグメンテーションとは、画像のピクセルごとにラベル情報を付して領域を色分けすることである。

【0004】

分類、物体検出、骨格検出、セグメンテーション等の処理を行う機械学習モデルを生成するため、学習用データを用いてモデルの学習を行う必要がある。従来、物体検出のための検出モデルの学習には、画像および正解ラベルで構成される学習用データが必要とされていた。このとき、正解ラベルをデータ毎に人が付与した場合、作成コストの増加が問題となるため、学習用データを自動的に生成する手法が望まれていた。

【0005】

特許文献1には、特に深層学習に用いる学習用データを生成する技術に関し、背景画像上において対象領域を定める領域決定部と、前記対象領域に対象物の画像を貼り付けて合成画像を生成する画像合成部と、前記対象物の画像にかかるデータに基づいて前記合成画像の正解ラベルを作成する正解ラベル作成部と、を備える学習用データ生成装置が開示されている。

【先行技術文献】

【特許文献】

【0006】

【文献】特開2020-149086号公報

【発明の概要】

【発明が解決しようとする課題】

【0007】

特許文献1の発明では、学習用データの生成指示に従って学習用データが自動的に生成される。また、学習用データの画像上の全ての対象物に正解ラベルが付加される。このため、適切な正解ラベルを付加された学習用データを簡易に生成できる。

【0008】

しかし、実際の画像解析においては、対象物を撮影した画像にオクルージョンが発生する場合がある。オクルージョンとは、一般に手前にある物体が背後にある対象物を隠す状態をいう。オクルージョンは画像解析の結果に影響を与えるが、発生のパターンは一般に対象物が置かれた環境に依存する。ここで、環境とは対象物が置かれた空間の種々の条件、例えば、広さ、対象物、対象物以外の物体の種類や数や大きさ、対象物と対象物以外の物体の関係、背景や明るさ等の条件を含むものとする。

【0009】

この点、特許文献1では、オペレータが指示した条件に従って合成条件を決定するが、環境ごとに切り抜き画像や合成条件を決定する手間が生じる。すなわち、実際の環境に即して、実際の画像に近づくように背景や対象物や対象物以外のもの（以下、便宜上これらを「パーツ」と呼ぶことがある）の組み合わせを選ぶ必要がある。また、実際の画像に近づくようにパーツを合成する合成条件を定める必要がある。このようなパーツや合成条件を手で決定するのは負担であるし、ランダムに定めると、実際の環境とは整合しない学習用データとなり、機械学習モデルの性能が低下するおそれがある。

【0010】

そこで本願発明の課題は、環境を考慮した学習効果の改善された学習用データを、容易に生成することにある。

【課題を解決するための手段】

【0011】

本願発明の好ましい一側面は、画像合成部を備え、前記画像合成部は、背景画像と、第一の画像および該第一の画像に紐づいた正解情報と、第二の画像と、環境情報に紐づいた合成条件を入力とし、前記合成条件を用いて、前記第一の画像、前記第二の画像、および前記背景画像から合成画像を生成し、前記合成画像と前記正解情報から、学習用データを生成する、学習用データ生成システムである。

【0012】

本願発明の好ましい他の一側面は、上記の学習用データ生成システムで生成した学習用データを用いて、機械学習モデルの機械学習を行う、機械学習モデルの学習方法である。

【0013】

本願発明の好ましい他の一側面は、環境情報に紐づけられた合成条件を準備する第1のステップ、画像データから抽出した、背景画像と、対象物画像と、非対象物画像を呼び出す第2のステップ、前記画像データに適合した環境条件に紐づけられた合成条件を用いて、前記背景画像と、前記対象物画像と、前記非対象物画像を合成して合成画像を生成する第3のステップ、前記合成画像と前記対象物画像に付された正解情報により学習用データを生成する第4のステップ、を実行する学習用データ生成方法である。

【発明の効果】

【0014】

本発明によれば、環境を考慮した学習効果の改善された学習用データを、容易に生成することができる。前述した以外の課題、構成及び効果は、以下の実施例の説明によって明らかにされる。

【図面の簡単な説明】

【0015】

【図1】実施例の機械学習システムの全体ブロック図。

【図2】前処理部の処理を示す概念図。

【図3】静止画像から生成した各画像を説明する概念図。

【図4】前処理を行う際に表示するGUIの例を示す模式図。

【図5】画像合成部の処理を示す概念図。

【図6】合成条件を編集する際に表示するGUIの例を示す模式図。

【図7】動画データの環境情報を、これに関連する背景画像、対象物画像、正解情報、オクルージョン画像、合成条件、学習用データが引き継ぐ例を示す概念図。

【図8】他の実施例の機械学習システムの全体ブロック図。

【図9】階層構造で分類された合成条件を説明する表図。

【図10】他の実施例の機械学習システムの全体ブロック図。

【図11】環境情報を自動的に推定する処理を説明する概念図。

【図12】環境情報を抽出する際に表示するGUIの例を示す模式図。

【図13】環境に対して適切な合成条件を自動的に選択する処理の概念図。

【図14】骨格検出の場合に正解情報付き対象物を作成する際のGUIの説明図。

【図15】画像合成の際のGUIの説明図。

10

20

30

40

50

【図16】他の実施例のセグメンテーションの処理の概念を示す説明図。

【図17】環境変動に対応できる学習用データの画像合成の概念を示す説明図。

【図18】オクルージョン画像にも正解情報を付けた場合の作用効果を説明する概念図。

【発明を実施するための形態】

【0016】

以下、本発明の実施例について、図面を用いて詳細に説明する。なお、以下の説明は、本発明の一実施形態を説明するためのものであり、本発明の範囲を制限するものではない。従って、当業者であればこれらの各要素又は全要素をこれと同等なものに置換した実施形態を採用することが可能であり、これらの実施形態も本発明の範囲に含まれる。

【0017】

以下に説明する実施例の構成において、同一部分又は同様な機能を有する部分には同一の符号を異なる図面間で共通して用い、重複する説明は省略することがある。

【0018】

同一あるいは同様な機能を有する要素が複数ある場合には、同一の符号に異なる添字を付して説明する場合がある。ただし、複数の要素を区別する必要がない場合には、添字を省略して説明する場合がある。

【0019】

本明細書等における「第1」、「第2」、「第3」などの表記は、構成要素を識別するために付するものであり、必ずしも、数、順序、もしくはその内容を限定するものではない。また、構成要素の識別のための番号は文脈毎に用いられ、一つの文脈で用いた番号が、他の文脈で必ずしも同一の構成を示すとは限らない。また、ある番号で識別された構成要素が、他の番号で識別された構成要素の機能を兼ねることを妨げるものではない。

【0020】

図面等において示す各構成の位置、大きさ、形状、範囲などは、発明の理解を容易にするため、実際の位置、大きさ、形状、範囲などを表していない場合がある。このため、本発明は、必ずしも、図面等に開示された位置、大きさ、形状、範囲などに限定されない。

【0021】

本明細書で引用した刊行物、特許および特許出願は、そのまま本明細書の説明の一部を構成する。

【0022】

本明細書において単数形で表される構成要素は、特段文脈で明らかに示されない限り、複数形を含むものとする。

【0023】

実施例で説明される学習用データ生成システムの一例は、画像解析を行うための機械学習モデルの学習用データを自動的に生成するものである。このシステムは、解析の対象となる対象物が置かれた環境に適合した学習用データを生成する。ここで、学習用データとは、機械学習モデルを学習するために用いる、問題と解答の組からなるデータをいうものとする。画像解析の場合、問題は一般に画像データである。

【0024】

機械学習モデルの解析精度を向上させるためには、精度の劣化要因となる障害物（オクルージョン）を含んだ学習用データが必要である。オクルージョンの具体例として、対象となる人が着用する眼鏡やマスクや帽子がある。また、人がルーペなどをのぞき込んでいる場合も障害物となり得る。あるいは、複数の人が交錯して重なり合う状況が考えられる。しかし、オクルージョンの発生のパターンは、環境ごとに条件が多様であり、環境ごとの相違を考慮する必要がある。

【0025】

実施例のシステムの構成の一例は、所定環境に置かれた対象物を撮影した画像から、背景画像と、対象物画像と、対象物以外の物の画像とを抽出し、対象物画像には正解情報を紐づける。そして、背景画像と、対象物画像と、対象物以外の物の画像から合成画像を生成し、合成画像と、対象物画像に紐づいた正解情報から、学習用データを生成する。ここ

10

20

30

40

50

で、対象物とは、画像解析の場合には解析の対象となる事物である。

【0026】

合成画像を生成する際には、所定環境に紐づいた対象物画像と、対象物以外の物の画像の合成条件を用いて、画像を合成する。

【0027】

すでに述べたように、環境とは対象物が置かれた空間の種々の条件である。環境情報は環境を表現する情報であり、システム上は例えば対象物が置かれた空間をユーザが分類し、定義する環境情報タグで特定することができる。あるいは環境情報は、対象物が置かれた空間を画像データとし、当該画像データを取得した装置（例えばビデオカメラ）を分類し、定義する環境情報タグで特定することができる。あるいは環境情報は、対象物が置かれた空間を画像データとし、当該画像データから得られる特徴量を分類し、あるいはグループ化することで特定することができる。

【実施例1】

【0028】

<1. 機械学習システムの全体構成>

図1は、実施例の機械学習システムの全体ブロック図である。以下では、単一の情報処理装置で学習用データの生成と学習を行うシステムを構成する例を示す。ただし、任意の部分がネットワークで接続された、別個の情報処理装置であってもよい。あるいは、可搬型の記録媒体でデータの送付ができる構成であれば、ネットワークで接続されていなくてもよい。要するに各要素間で必要なデータの移動が可能であればよい。

【0029】

機械学習システム1は一般的なサーバーのような情報処理装置で構成することができる。一般的な情報処理装置と同様、機械学習システム1は、中央処理装置（CPU：Central Processing Unit）101、入出力インターフェース（I/O）102、メモリ200、データベース（DB）300を備える。各部を接続するバスの類は省略している。

【0030】

入出力インターフェース102には、キーボードやディスプレイなど、公知の種々の入出力装置を接続することができる。他の情報処理装置と接続するためのネットワークに対するインターフェースを備えてもよい。

【0031】

メモリ200は一般に半導体メモリのような比較的高速の記憶装置が用いられ、CPU101が実行するプログラム等が格納される。データベース300は磁気ディスク装置のような比較的大容量の記憶装置が用いられ、主にデータを記憶する。ただし、メモリ200とデータベース300の役割分担は、使用する記憶デバイスの能力により流動的であり、特に限定されるものではない。

【0032】

本実施例では、メモリ200にプログラムとして前処理部201、画像合成部202、機械学習部203を格納するものとする。本実施例では、メモリ200に格納されたプログラムをCPU101が実行することにより、所定の機能を実現する。なお、同様の機能を実現できるのであれば、ソフトウェアでなく専用のハードウェアを用いてもよい。

【0033】

データベース300には、動画データ301、背景画像302、対象物画像303、正解情報304、オクルージョン画像305、学習用データ306が格納される。各部分の役割は後に説明される。

【0034】

以下で機械学習システム1の具体的処理の例を説明する。ここでは、例えば動画像から対象物として人物を抽出し、人物の例えば、関節、目、耳等任意の部分の座標を抽出するとともにラベル付けを行う骨格検出を例に説明する。もちろんその他の画像解析処理であっても、実施例の適用は可能である。機械学習モデルにより骨格検出のような画像解析を行う場合、機械学習モデルを学習するための学習用データの質と量が重要である。以下の

実施例では、良質な学習用データを容易に生成する例を説明する。

【0035】

＜2．前処理部＞

図2は前処理部201の処理を示す概念図である。前処理部201は、例えば入出力インターフェース102を介して前処理を指示するコマンドを受信すると、以下の処理を開始する。

【0036】

前処理部201は、まずデータベース300から動画データ301を読み出す。動画データ301は、例えば対象物を撮影したビデオカメラから取得する。

【0037】

前処理部201は動画データ301を読み込み、当該動画データ301から例えば1フレームを切り出して静止画像を取得する(S2001)。動画データ301からは静止画像が複数取得でき、取得する数は任意であるが、以下の説明では1つの静止画像についての処理を説明する。

【0038】

前処理部201は、取得した静止画像から対象物(人物)やオクルージョンがない背景画像302、対象物(人物)がなくオクルージョンを含む背景画像2100、オクルージョンがなく対象物(人物)を含む背景画像2200を生成する。なお、オクルージョンとは、一般には対象物に重なる物体をさすが、本明細書では対象物以外で指定された任意の物体を総称する概念とする。

【0039】

図3は、静止画像2000から生成した各画像の概念を説明する図である。オクルージョンを含む背景画像2100には、背景(ここでは山)のほかオクルージョン画像(ここではルーベ)305が含まれ、対象物を含む背景画像2200には対象物画像303(ここでは人物)が含まれる

このような処理は、最も単純な例では、人(ユーザ)が適当なGUI(Graphical User Interface)を用いて、静止画像2000を見ながら対象物やオクルージョンを指定し、これらを背景から分離あるいは消去する。あるいは、公知の画像認識の技術を用いて自動的に分離等してもよい。あらかじめ定義された対象物やオクルージョンであれば、ある程度の精度で抽出が可能である。あるいは、自動的にこれらを抽出した後で、ユーザがチェックして修正してもよい。

【0040】

前処理部201は、次にオクルージョンを含む背景画像2100と対象物(人物)を含む背景画像2200から、オクルージョン画像305と対象物画像303を切り抜く。このためには、単純に背景画像302との差分を取ればよい。なお、上記の例では対象物やオクルージョンは一種類だが、複数あってもよい。

【0041】

次に前処理部201は、対象物画像303に個別の正解値を付与する(S2003)。最も一般的な手法では、適当なGUIを用いてユーザが正解を付与する。骨格検出の場合には、対象物画像303の画像上に目、肘関節、手首などの部位を指定して座標を記録する。以上で正解情報304が得られる。骨格検出の場合、正解情報304は、対象物例えば「人物」を特定する情報、対象物画像303、および対象物画像上の各部位の座標を含む。画像解析が分類の場合は対象物を特定する情報と対象物画像303の組を正解情報304とする。セグメンテーションの場合には、セグメンテーションした領域とカテゴリ分類の組を正解情報304とする。その他の画像解析の場合はそれらに応じた公知の正解情報とする。

【0042】

図4は、これらの前処理を行う際に、例えば入出力インターフェース102を介してディスプレイに表示するGUIの例である。

【0043】

10

20

30

40

50

前処理部 201 は、以上のように作成した背景画像 302、対象物画像 303、正解情報 304、オクルージョン画像 305 を、データベース 300 に記録する。対象物画像 303 と正解情報 304 は組になる。このとき、これらの生成したデータには環境情報を示す環境情報タグを付与する。環境情報はユーザが任意に設定して付与してもよい。

【0044】

環境情報を自動的に付与することも可能である。例えば、動画データ 301 を撮影したビデオカメラが置かれた環境を示すタグ（あるいはビデオカメラを特定する ID）を、当該動画データ 301 を元に生成した背景画像 302、対象物画像 303、正解情報 304、オクルージョン画像 305 に自動的に付与してもよい。

【0045】

例えば、動画データ 301 を撮影したビデオカメラの環境情報が「第一工場、前工程室、第一カメラ」であれば、「第一工場」あるいは「第一工場、前工程室」あるいは「第一工場、前工程室、第一カメラ」のように環境情報をタグづけする。すなわち、環境情報の定義は任意である。環境情報は、上記のように階層構造にしてもよい。

【0046】

一般には動画データ 301 からは複数種類の静止画像 2000 が取得でき、背景画像 302、対象物画像 303、正解情報 304、オクルージョン画像 305 も複数種類が取得できるから、これら複数の画像や情報に共通の環境情報が紐づけられてグループ化されることになる。以上の処理にて、学習用データを合成するためのパーツが、環境情報で分類され、データベース 300 に準備されたことになる。

【0047】

< 3. 画像合成部 >

図 5 は画像合成部 202 の処理を示す概念図である。画像合成部 202 は、例えば入出力インターフェース 102 を介して画像合成を指示するコマンドを受信すると、以下の処理を開始する。

【0048】

画像合成部 202 は、データベース 300 から、背景画像 302、対象物画像 303、正解情報 304、オクルージョン画像 305 を読み出す。このとき、例えば「第一工場、第一作業室」に設置したビデオカメラの画像を用いて機械学習モデルに画像解析を行わせたいとする。この場合、当該機械学習モデルの学習に用いる学習用データを生成するため、例えば入出力インターフェース 102 を介して「第一工場、前工程室」のように環境情報を指定する。これにより、例えば「第一工場、前工程室、第一カメラ」、「第一工場、前工程室、第二カメラ」の環境情報タグが付されたデータを呼び出すことができる。あるいは、「第一工場、前工程室、第一カメラ」のように環境情報を指定すると、「第一工場、前工程室、第一カメラ」の環境情報タグが付されたデータを呼び出すことができる。

【0049】

次に画像合成部 202 は、背景画像 302、対象物画像 303、正解情報 304、オクルージョン画像 305 を合成して、背景画像に対象物とオクルージョンが張り付けられた合成画像を生成する。このとき、ユーザは例えば入出力インターフェース 102 を介して合成条件 501 を指定することができる。

【0050】

ここで、合成条件 501 は、以下のような情報を含むテンプレートにすることができる。

- ・ 合成する対象物やオクルージョンの種類や数
- ・ 背景画像に対する対象物やオクルージョンの位置（座標）の指定
- ・ 背景画像に対する対象物やオクルージョンの指定された位置（座標）からのばらつき
- ・ 対象物やオクルージョンの重なる順序（レイヤー）
- ・ 対象物やオクルージョンの大きさ
- ・ 対象物やオクルージョンの向き（角度）
- ・ 合成画像の明るさ
- ・ 合成画像の解像度

10

20

30

40

50

・合成画像上の特殊効果

【0051】

本実施例では、一度指定した合成条件をテンプレートとして保存しておいて再利用できるようにする。具体的には、例えば「第一工場、前工程室」の環境情報タグが付されたデータを読み出して画像を合成する際には、「第一工場、前工程室」の環境を知悉しているユーザが合成条件を指定することができる。指定した合成条件は、「第一工場、前工程室」の環境情報タグを付してテンプレートとして保存する。

【0052】

例えば図3に示した背景画像302、対象物画像（人物）303、オクルージョン画像（ルーペ）305を使って合成画像を生成する際の合成条件501としては、例えば背景画像302がレイヤー1（最下層）、対象物画像303がレイヤー2、オクルージョン画像305がレイヤー3（最上層）のように重ねる順序を指定する。また、対象物画像303の配置される座標は（Xmin,Ymin）、オクルージョン画像305の配置される座標は（X'min,Y'min）で、それぞれ（ $\delta X, \delta Y$ ）、（ $\delta X', \delta Y'$ ）のばらつきを持たせる等の条件を設定する。ばらつきの条件は張り付ける領域を指定し、その中でランダムに配置させるようにしてもよい。

【0053】

図6は、ユーザが合成条件を編集する際に、入出力インターフェース102を介してディスプレイに表示するGUIの例である。ユーザは正解情報304が付された対象物画像303やオクルージョン画像305を画面上で移動させたり、重ねる順序を指定したりして合成条件を設定することができる。なお、これらのパーツ（背景画像、対象物画像、オクルージョン画像）は、同じ環境情報タグを持つ異なる複数の静止画像から生成したものを、混在して用いることができる。

【0054】

このような合成条件501を一度テンプレート化して設定したら、例えば「第一工場、前工程室」のような環境情報を示す環境情報タグを付して、画像合成部202内のメモリに記録しておく。そうすれば、同じあるいは類似の環境情報（例えば「第一工場、前工程室」あるいは「第二工場、前工程室」）を持つ背景画像302、対象物画像303、正解情報304、オクルージョン画像305を用いて画像を合成する際には、同じ合成条件を再利用できる可能性が高い。なお、合成条件501をテンプレート化して記憶する際に、当該合成条件を用いて生成した学習用データの学習精度や効率を（合成画像ではない）実画像で評価して、保存する合成条件を選択してもよい。

【0055】

合成条件の再利用を可能にするために、生成した合成条件は上記のように環境情報に紐づけておくことが望ましい。紐づける環境情報は、ユーザが任意に設定してもよいし、当該合成条件を使用して画像の合成を行ったパーツ等と同じ環境情報を自動的に設定してもよい。上記の例では、合成条件には、その合成条件を適用したパーツと同じ「第一工場、前工程室」の環境情報タグを自動的に付与する。

【0056】

かくして、画像合成部202は合成画像502と正解ラベル503の組を生成する。正解ラベル503は、この骨格検出の例では合成画像中の対象物画像303の位置とその中の目、肘関節、手首などの部位の位置座標である。合成画像502を生成した際に、対象物画像303は特定され、かつ対象物画像を張り付けた位置は判明しているので、これと正解情報304から正解ラベル503を生成することができる。合成画像502と正解ラベル503の組が1つの学習用データ306を構成する。

【0057】

画像合成部202は、生成した学習用データ306をデータベース300に記録する。記録する際には、学習用データを生成する際に使用したパーツや合成条件と同じ環境情報を付与しておいてもよい。このようにすることで、どのような環境で使用する学習モデルのためのデータが識別できる。すなわち、一つの例では、パーツ等、合成条件、および作

10

20

30

40

50

成された学習用データが、環境情報で紐づけられることになる。各データにタグ付けする代わりに、環境条件ごとに異なるデータベースあるいはファイルに分けて格納してもよい。

【0058】

図7は、動画データ301の環境情報を、これに関連する静止画像2000、背景画像302、対象物画像303、正解情報304、オクルージョン画像305、合成条件501、学習用データ306が引き継ぐ例を示している。

【0059】

環境情報タグAで識別される動画データ301からは、環境情報タグAで識別される複数種の静止画像が得られ、これから前処理部201を用いて、環境情報タグAで識別される複数種類の背景画像302、対象物画像303、正解情報304、オクルージョン画像305が得られる。これらを合成して合成画像を生成する際には、環境情報タグAで識別される共通の合成条件を適用することができる。

【0060】

すなわち、画像合成部202により環境情報タグAを持つパーツの合成のための合成条件を生成した後、当該生成した環境情報タグAで識別される合成条件501は、同じ環境情報タグAを持つパーツの合成に再利用することができる。あるいは、後に図9に示すように環境情報をあらかじめ分類しておけば、類似する環境情報タグを持つパーツの合成に再利用することができる。

【0061】

合成画像502と正解ラベル503の組からなる学習用データ306は、前述のXδ、Yδのばらつき条件により、物体相互の位置関係の異なる多数の種類の学習用データを生成することができる。また、複数種類の対象物画像303やオクルージョン画像305からランダムにパーツを選択することで、同じ対象物でも状態の異なる多数の種類の学習用データを生成することができる。これは、合成画像502を問題、正解ラベル503を正解とする教師データであり、機械学習モデルの学習用データとして用いることができる。

【0062】

<4. 機械学習部>

機械学習部203は、生成した学習用データ306を用いて、図示していない機械学習モデルの学習を行う。機械学習モデルの学習の処理自体は、公知の技術を適用可能であるため、説明を省略する。

【0063】

学習を行う際には、例えば入出力インターフェース102を介して環境情報を示すタグを検索することで、環境に応じた学習用データ306を選択することができる。例えば、「第一工場、前工程室」における人間の骨格検出処理には、「第一工場、前工程室」の環境情報タグが付された学習用データを用いて機械学習を行うのがよい。あるいは、同一の環境情報を持つデータがない場合でも、「第一工場、前工程室」のタグが付された学習用データは「第一社員寮、調理室」の環境情報タグが付された学習用データより、「第二工場、前工程室」用途の学習に適している可能性がある。このように、環境情報を考慮した学習用データを選択することが可能となる。

【0064】

本実施例によれば、機械学習モデルが行う画像解析に適した学習用データを自動的に得る際に、画像合成の合成条件を再利用することが可能となる。

【実施例2】

【0065】

図8は、他の実施例の機械学習システムの全体ブロック図である。図1の実施例と異なる部分を説明する。

【0066】

合成条件管理部204は、画像合成部202が合成条件501のテンプレートを作成した後、あるいは、そのテンプレートを用いて画像を合成した後、当該合成条件501に環境情報を示すタグを付して、データベース300に合成条件データ307として蓄積する。

【0067】

一例では、合成条件管理部204は、ユーザが入出力インターフェース102を介して指定した環境情報を示す環境情報タグを合成条件501に添付する。

【0068】

他の例では、合成条件管理部204は、その合成条件を用いて画像合成部202が合成したパーツに付されていた環境情報を示す環境情報タグを、そのまま、あるいは編集して、当該合成条件に付与する。このようにして、合成条件管理部204が、環境情報と合成条件とを紐づけて分類したうえで、データベース300に合成条件データ307として格納する。

【0069】

図9は、階層構造の環境情報で分類された合成条件データ307の一例を示す構成図である。階層構造は大分類（例えば「工場」や「道路」）、中分類（例えば「前工程室」や「交差点」）、小分類（例えば「エッチング」や「日本橋交差点」）のように環境情報を階層構造にしておくことができる。

【0070】

かかる階層構造によれば、同一の環境情報を持つ合成条件がない場合でも、類似する環境情報を持つ合成条件を抽出することができる。同様の階層構造による分類は、パーツや学習用データに対しても適用できる。

【0071】

このような階層構造は、ユーザが任意に定義することができる。あるいは、合成したパーツの元になっている動画データを採取するビデオカメラの置かれた環境や、画像解析の目的に応じて、ユーザが定義することができる。

【0072】

一つの例では、動画データを採取するビデオカメラには、環境情報と紐づけられたIDが付されており、当該ビデオカメラで撮影された動画データおよび当該動画データを元に生成された各種データおよびそのデータを合成した合成条件は、ビデオカメラと同じIDが付加される。このようにすると、一連のデータが環境情報でグループ分けされることになる。

【0073】

図9に示す合成条件データ307を構築した後、後に別の学習用データを作成する際には、ユーザが入出力インターフェース102を介して環境情報を指定し、合成条件管理部204は、指定された環境条件と同一もしくは類似の環境情報（例えば大分類のみ一致）を持つパーツを特定する。そしてパーツを合成するための合成条件として、同じく指定された環境条件と同一もしくは類似の環境情報を持つ合成条件を特定する。そして、画像合成部202は、これらを使用して学習用データの生成を行う。

【0074】

具体的な例としては、「第一工場、前工程室」の環境情報タグが付されたパーツの合成条件として生成された合成条件は、「第一工場、前工程室」の環境情報タグが付された合成条件データ307としてデータベース300に格納されており、この合成条件は、後に「第一工場、前工程室」の環境情報タグが付された動画データをもとに得られたパーツの画像合成に再利用できる。

【0075】

本実施例によれば、機械学習モデルが行う画像解析に適した学習用データを自動的に得る際に、適切な合成条件を選択して画像合成を行うことが可能となる。

【実施例3】

【0076】

図10は、他の実施例の機械学習システムの全体ブロック図である。図8の実施例と異なる部分を説明する。

【0077】

前記の実施例では、環境情報を、ユーザの定義や画像データの由来による環境情報タグ

10

20

30

40

50

で指定することを可能とした。その他の例として、オートエンコーダやCNN（Convolution Neural Network）を用いて画像の特徴量を抽出することで環境情報を特定することができる。

【0078】

環境情報管理部205は、オートエンコーダやCNNを用いて画像の特徴量を抽出する特徴量抽出器を備える。環境情報管理部205は、抽出した特徴量を環境情報データ308としてデータベース300に記録する。

【0079】

図11は、環境情報を自動的に推定する処理を説明する概念図である。合成条件管理部204は、画像合成部202が合成条件501を作成した後、動画データ301、静止画像2000、背景画像302、対象物画像303、正解情報304、オクルージョン画像305の少なくとも一つ（便宜的に「オートエンコーダ入力」ということがある）を、環境情報管理部205が備えるオートエンコーダやCNNに入力し、特徴量を抽出する（S11001）。

【0080】

得られた特徴量は環境情報データ308として、実施例1、実施例2と同様にパーツや合成条件や学習用データと紐づけることができる。合成条件データ307では、合成条件は特徴量と紐づけて保存しておく。

【0081】

具体例としては、図7の環境情報タグの代わりに、あるいは環境情報タグに加えて、特徴量を付加すればよい。このとき、特徴量は画像データが異なると異なるため、再利用を容易にするために、公知のクラスタリング等の手法によりグループ化しておくこともできる。

【0082】

オートエンコーダ入力の情報量が多いほど詳細な特徴量を得られるが、直感的な理解が難しくなるので、可視化のため主成分分析など公知の手法により低次元表現することもできる。この場合、所定の範囲の特徴量（特徴ベクトル）に対して、一つのグループのパーツや合成条件や学習用データと紐づけることができる。

【0083】

典型的な例としては、オートエンコーダ入力として静止画像2000を用いる。環境情報管理部205は、オートエンコーダ出力として環境情報（特徴量）を得たら、特徴量を静止画像のもととなる動画データ301と紐づける。また、静止画像2000から生成した、背景画像302、対象物画像303、正解情報304、オクルージョン画像305と特徴量を紐づける。また、これらのパーツを合成した合成条件と特徴量も紐づける。

【0084】

図7の例では、一つの動画データ301を起源とするデータは、同じ環境情報タグを持っているので、特徴量を抽出した静止画像と同じ環境情報タグを持つデータに、一括して抽出した環境情報（特徴量）を紐づければよい。

【0085】

環境情報管理部205は、このように他のデータと紐づけられた特徴量を、環境情報データ308としてデータベース300に記録する。

【0086】

図12は、ユーザが特徴量を抽出する際に、入出力インターフェース102を介してディスプレイに表示するGUIの例である。抽出された特徴量成分を低次元化し、グラフで可視化して示すことにより、環境情報との対応を理解しやすくなる。

【0087】

本実施例によれば、環境情報と合成条件を紐づける際に、自動的に環境情報を特定することが可能となる。また、環境情報を特徴ベクトルで表現することにより、環境情報と紐づけられた合成条件の選択を自動的に行うことが可能となる。

【0088】

10

20

30

40

50

図13は、環境条件（特徴量）の自動抽出により、環境に対して適切な合成条件等を自動的に選択する処理の概念を示す図である。オートエンコーダやCNNによる特徴量抽出器1301は、オートエンコーダ入力から特徴量を抽出し、環境情報データ308として記録する。このとき、公知手法でクラスタリングしてグループ分けしておいてもよい。

【0089】

先に説明したように、オートエンコーダ入力とした例えば静止画像2000には、抽出した特徴量を紐づけておく。また、オートエンコーダ入力とした静止画像2000に関連する動画データ301、背景画像302、対象物画像303、正解情報304、オクルージョン画像305には、抽出した特徴量を紐づけておく。これらのオートエンコーダ入力から学習用データ306を生成済みの場合には、学習用データにも抽出した特徴量を紐づけておく。また、当該学習用データを生成した際に用いた合成条件データ307にも特徴量を紐づけておく。これらの情報は予めデータベースとしておく。

【0090】

図13の例では、データベースにおいて、環境Aを示す環境情報データと環境Bを示す環境情報データがあり、それぞれに各パーツと学習用データ306と合成条件データ307が紐づけられている。

【0091】

未知の環境から得られたパーツから学習用データを生成しようとする場合には、合成条件管理部204は、例えば当該パーツの元となる例えば静止画像を環境情報管理部205に送る。

【0092】

静止画像のような画像データは、パーツが存在する環境を反映しているので、当該静止画像に適合した環境条件に紐づけられた合成条件を用いて、合成画像を生成するのがよい。静止画像から環境条件を定量化して抽出するため、特徴量を用いる。

【0093】

環境情報管理部205は、静止画像をオートエンコーダ入力として特徴量抽出器1301で特徴量を抽出する。そして、既存の環境情報データ308を検索し、近い特徴量を持つ環境情報データを抽出して合成条件管理部204に通知する。

【0094】

合成条件管理部204は、抽出された環境情報データと紐づけられた合成条件データ307を用いて未知の環境から得られたパーツの画像合成を行うことで、環境情報に適合した画像合成が可能となる。

【0095】

図13の例では、環境情報管理部205は、未知の環境Cから得られた静止画像2000から特徴量抽出器1301で特徴量を抽出し、既存の環境情報データ308である環境A、環境Bの特徴量と比較して、最も近い特徴量を持つ環境情報データ308を抽出する。

【0096】

環境情報管理部205は、抽出した環境情報データ308を合成条件管理部204に通知する。合成条件管理部204は、通知された環境情報データに紐づいた合成条件データをデータベース300から呼び出す。合成条件管理部204は、呼び出した合成条件データ307を画像合成部202に送り、そのテンプレートをを用いて未知の環境Cから得られたパーツを画像合成する。

【0097】

本実施例によれば、一度作成した合成条件を環境情報と紐づけてテンプレートに登録しておき、環境情報の差異に基づき複数候補から適した合成条件を選ぶことができる。このため、未知の環境情報を持つ画像についても、適切な合成条件を選択することが可能となる。

【0098】

特徴量を用いて環境情報を推定する上記の手法は、環境情報を階層構造としておく図9の手法と併用することもできる。例えば、大分類で合成条件を抽出した後、特徴量を用い

10

20

30

40

50

た環境情報により、さらに合成条件を絞り込んでもよい。該当手法により、さらに合成条件の選択の精度が向上することが期待できる。

【実施例 4】

【0099】

以下では、前処理部 201 において、対象画像に正解情報を付与する GUI の例と、画像合成部 202 において、画像を合成する際のテンプレートの一例である特殊効果について説明する。

【0100】

図 14 は、前処理部 201 における処理で、骨格検出の場合に対象物画像 303 に正解情報 304 を添付する際の GUI の説明図である。切り出した一つの対象物画像 303 に、ユーザが対象画像固有の相対値座標で、正解値である正解情報 304 を付与し、正解情報付き対象物画像を生成する。図 14 の例では、対象物である人物の左肩が相対値座標で (x_6, y_6) で指定されている。

【0101】

図 15 は、画像合成部 202 における処理で、図 14 で生成した正解情報付き対象物画像（対象物画像 303 および正解情報 304）を、背景画像 302 とオクルージョン画像 305 と合成する際の説明図である。

【0102】

基本的に、背景画像に対象物画像 303 を張り付ける際は、正解情報 304 の相対座標を、背景画像 302 に固有の座標系に変える。図 15 の例では、合成条件データ 307 のテンプレートに基づいて、対象物画像 303 の張り付け位置が背景画像 302 に固有の座標系で (x_{min}, y_{min}) で指定されている。テンプレートによっては、所定の位置ばらつきがランダムで与えられる。図 14 の例では、人物の左肩が相対値座標で (x_6, y_6) であるから、背景画像と合成した合成画像において、人物の左肩の座標は $(x_{min} + x_6, y_{min} + y_6)$ のようになる（ばらつきがない場合）。

【0103】

ここでルーペなどをオクルージョン画像 305 に使う場合は、より表現を実態に合わせるため座標ずれを施すことが望ましい。図 15 の例では、オクルージョン画像 305 である「ルーペ」が、対象物画像 303 である「人物」に重なっている。合成処理のテンプレートにおいて、対象物画像 303 がオクルージョン画像 305 を基準とした所定範囲と重なる場合には、正解情報 304 の相対座標を背景画像 302 に固有の座標系に変える際に補正を行うように規定しておく。

【0104】

具体的には、「ルーペ」のオクルージョン画像 305 では、レンズの透明部分は切り抜かれているものとする。ここで合成条件のテンプレートでは、「オクルージョン画像がルーペの場合には、ルーペの枠内にある対象物画像の部位の座標は、相対値座標で (x_1, y_1) から $(x_1 + x_2, y_1 + y_2)$ に変更する」のように規定しておく。図 15 の例では、左目の位置の座標がルーペのレンズの屈折を再現するように補正されている。

【0105】

以上のように、背景画像 302、対象物画像 303 および、オクルージョン画像 305 の指定された箇所の画像を変形させたり、明度を変化させたり、ぼやけさせたりする特殊効果を追加することができる。

【0106】

このようなテンプレートで規定する特殊効果は、例えば透明や半透明のオクルージョンや、光の反射などを表現するのに有効である。

【実施例 5】

【0107】

以下では、画像解析でセグメンテーションを行う場合、前処理部 201 において、セグメンテーションした領域にカテゴリ分類を付して正解情報を得、画像合成部 202 で画像

10

20

30

40

50

を合成する場合を説明する。基本的に図 8 の実施例の構成と同様で良いが、本実施例特有の構成を以下説明する。

【0108】

前処理部 201 は、入出力インターフェース 102 を介してディスプレイに静止画像 2000 を表示し、適当な GUI を用いてユーザに静止画像 2000 からセグメンテーション領域を指定させカテゴリ分類を付与させる。

【0109】

図 16 は、本実施例の処理を説明する図である。本実施例の前処理では、セグメンテーションした領域をパーツとして対象物画像 303 に格納する。また、各パーツに紐づいたカテゴリ分類（テーブル、ベンチ、カッター、ループ、人、、、）を正解情報 304 に格納する。

10

【0110】

セグメンテーション処理では、セグメンテーションした領域が対象物であると同時にオクルージョンになり得るので、図 1 の実施例のように対象物画像 303 とオクルージョン画像 305 を区別する必要はない。

【0111】

画像合成部 202 では、セグメンテーションした領域を合成する際に、実施例 1 ～ 3 で説明したようにセグメンテーション領域の張り付け位置を指定する他、合成条件として領域ごとにレイヤーを指定してテンプレート化し、合成条件データ 307 の一部として記憶しておく。例えば、テーブル（レイヤー 0）とスパナ（レイヤー 1）とカッター（レイヤー 2）のようにテンプレートで定義しておく。定義したテンプレートを、パーツと関連付けておくのは、すでに説明した実施例と同様である。

20

【0112】

画像合成部 202 では、各パーツから合成画像を生成する際に、パーツと関連付けられた合成条件から、事前にテンプレートで定義したカテゴリと描画順序が合致するように合成画像を作る。

【0113】

例えば、上述のようにレイヤーを指定しておけば、テーブルの上にスパナとカッターが配置されるように、合成画像が生成される。

【実施例 6】

30

【0114】

以下の実施例では、より具体的な作業環境に応じて環境情報を調整する例を説明する。例えば、運搬業における段ボールの荷受け・仕分け作業における行動追跡や、逸脱作業で物体検出をする場合を考える。この場合、画像解析の対象物は段ボールや作業員（人）などである。

【0115】

この場合の課題として、非定型作業のため、同じ状況が少なく、正解情報 304 の再利用が困難となる。その結果、以下に示す環境変動で誤検知（検知対象外の物体を対象物と間違える）を引き起こすことが考えられる。

【0116】

40

環境変動の例としては、作業場の広さ、段ボールの置き場所、作業員の人数、作業員の服装、照明の位置、作業員や段ボールの状態の変化などの変動がある。また、上記に加えて、画像に写りこむ対象外の物体（例えば、脚立や台車、カラーコーン（登録商標）など）の存在が考えられる。

【0117】

図 17 は上記のような環境変動に対応できる学習用データの画像合成の概念を示す説明図である。本実施例の対象物画像 303 やオクルージョン画像 305 は、異なる形態の画像に全部または一部が同一の正解情報（ラベル）を付す。例えば、画像認識において、静止している人の対象物画像と走っている人の対象物画像に同じく「作業員」のラベルを付してシステム上同一に扱う。また、閉じている段ボール箱の対象物画像と開いている段ボ

50

ール箱の対象物画像に同じく「段ボール箱」のラベルを付してシステム上同一に扱う。このように同一物でも状況により形態が異なる対象物画像は、例えば動画像データの異なるタイミングから静止画像を切り出して、対象物画像を抽出することで得ることができる。

【0118】

画像合成に際しては、合成条件のテンプレートに従って、同一のラベルを持つ複数種の対象物画像からランダムに対象物画像を選択して合成する。これにより、環境変動を反映した合成画像を生成することができる。例えば、合成条件のテンプレートで「作業員」の対象物画像の一つ選択するとされる場合、静止している人の対象物画像と走っている人の対象物画像がランダムに選択されて画像が合成される。

【0119】

また、本実施例では、オクルージョン画像305にも正解情報304を付与する。すなわち、環境に紐づく対象物以外の物体にも正解値を付けて、画像合成を実施して学習用データを生成する。このようにすることで、誤検知を低減することが可能である。

【0120】

図18は、本実施例のオクルージョン画像305にも正解情報304を付けた場合の作用効果を説明する概念図である。オクルージョン画像305への正解情報304の付け方は、対象物画像303への正解情報304の付け方と同様である。

【0121】

画像解析の対象となる静止画像1801を、本実施例の機械学習モデル1802で解析した場合を説明する。本実施例の機械学習モデル1802は、オクルージョンである「カラーコーン」の物体検出が可能である。

【0122】

いま静止画像1801には、対象物である「作業員」と「段ボール」が含まれ、さらに対象物以外の「カラーコーン」が含まれている。機械学習モデル1802が対象物「段ボール」と「作業員」しか認識できない場合、「カラーコーン」を「段ボール」と誤認識する可能性があるが、「カラーコーン」も認識できる本実施例の機械学習モデルでは、誤認識の低減が可能である。

【0123】

上位のアプリケーションでは、検出対象物である「作業員」と「段ボール」について処理を行うので、機械学習モデルの出力1803からは「カラーコーン」の情報をフィルタする(S1804)。フィルタリングされたデータを上位のアプリケーションで処理して行動追跡などの処理を行う(S1805)。「カラーコーン」の正解値を付与せずに「作業員」と「段ボール」のみの正解値付与されたデータセットの場合、推論時に「カラーコーン」を学習時に記憶した「作業員」もしくは「段ボール」の特徴量との類似度を評価し、どちらか一方であると認識してしまう。しかし、「カラーコーン」「作業員」「段ボール」の正解値を付与したデータセットを用いる本実施例の場合、推論時には「カラーコーン」は学習時に記憶した3種の特徴量の内、「カラーコーン」のそれにもっとも近くなるため、誤認識の確率が減る効果が期待できる。そのため上記で説明したフィルタリングが有効となる。

【0124】

以上説明した実施例によれば、環境に応じた照度や光の種類、サイズや距離など、物体そのものは同じだが違うものに見える可能性のある環境条件を反映することができる。たとえば、類似の業務を行う工場で、作業員は似たような作業服を着ているが、機器や加工材料、配置が異なる場合である。また、ある作業員はカメラの真正面にまっすぐ向き合っ立つが、ある作業員は斜に傾いて立ったり、あるいは端の方に立ったりするなどの、環境条件の差異を学習用データに反映することができる。

【0125】

上記実施例によれば、少ない処理量で学習用データの生成が可能となるため、消費エネルギーが少なく、炭素排出量を減らし、地球温暖化を防止、持続可能な社会の実現に寄与することができる。

10

20

30

40

50

【符号の説明】

【 0 1 2 6 】

機械学習システム 1、中央処理装置 1 0 1、入出力インターフェース 1 0 2、メモリ 2 0 0、データベース 3 0 0

10

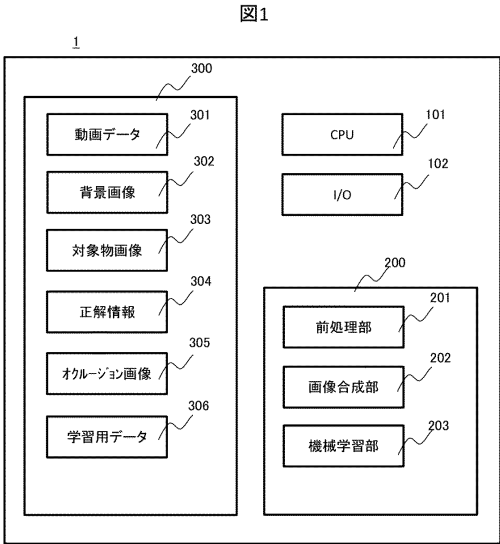
20

30

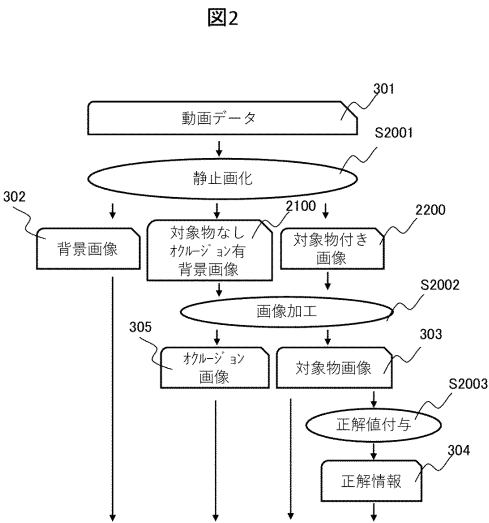
40

50

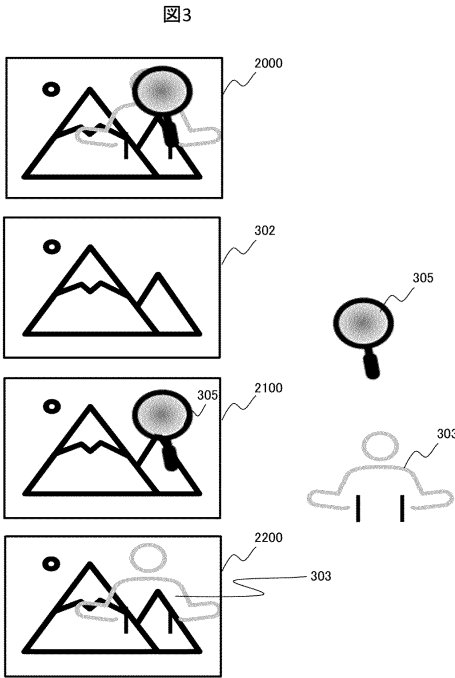
【図面】
【図 1】



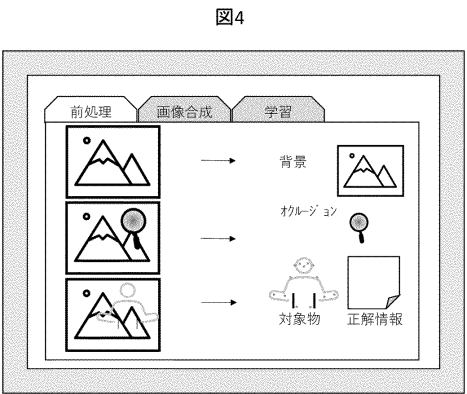
【図 2】



【図 3】



【図 4】



10

20

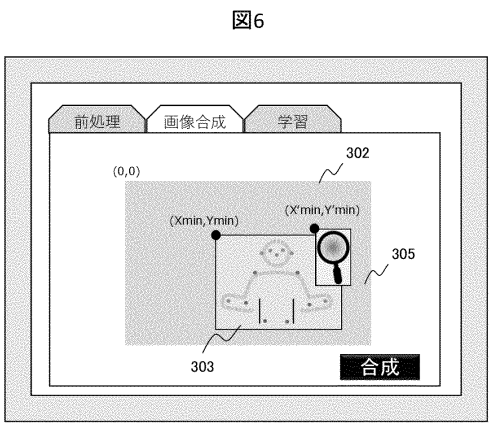
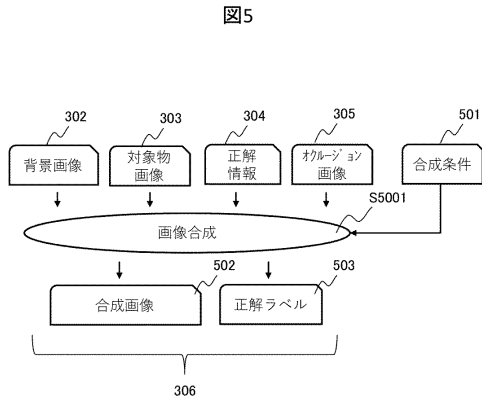
30

40

50

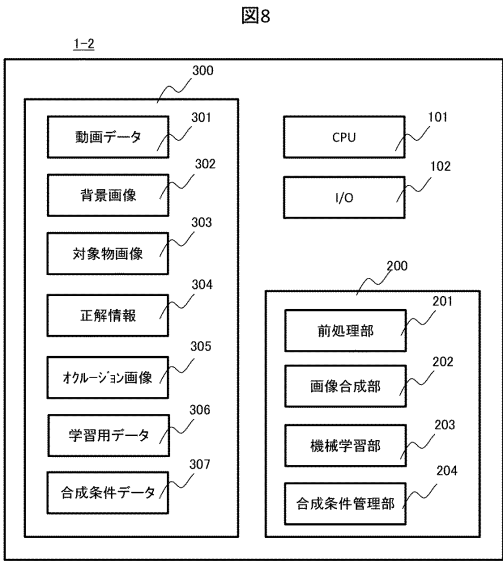
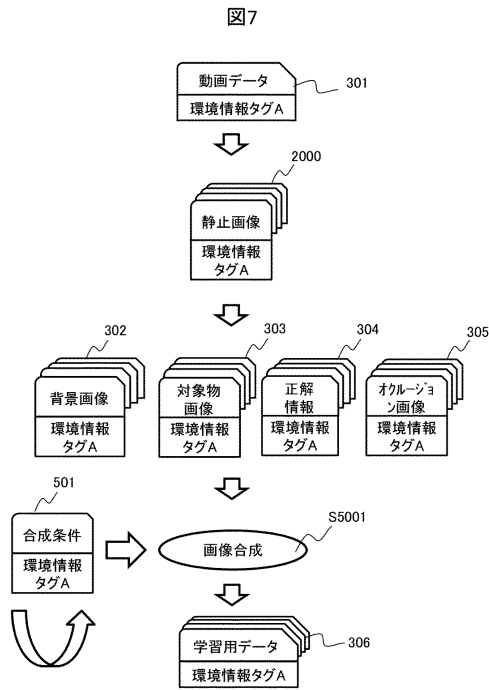
【図 5】

【図 6】



【図 7】

【図 8】



10

20

30

40

50

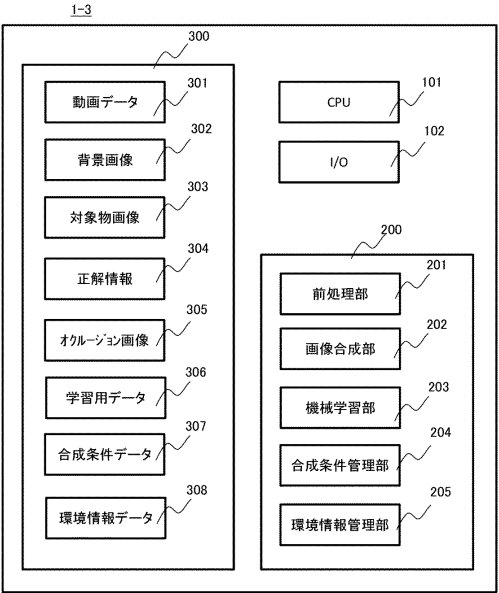
【図 9】

図9

大分類	中分類	小分類	合成条件
工場	前工程室	エッチング	条件1
工場	後工程室	ダイシング	条件2
道路	交差点	日本橋交差点	条件3
道路	高速道路	東北道上り車線	条件4

【図 1 0】

図10

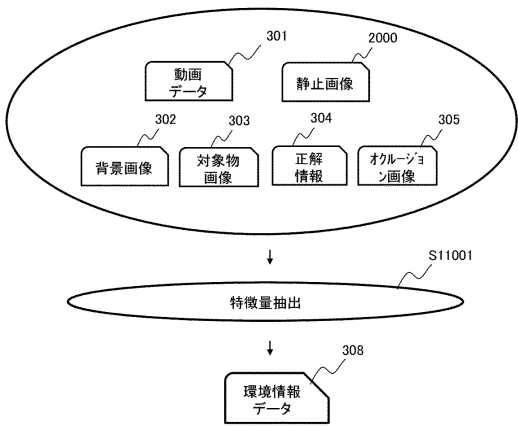


10

20

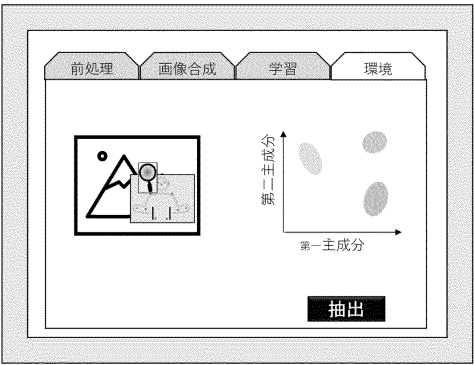
【図 1 1】

図11



【図 1 2】

図12



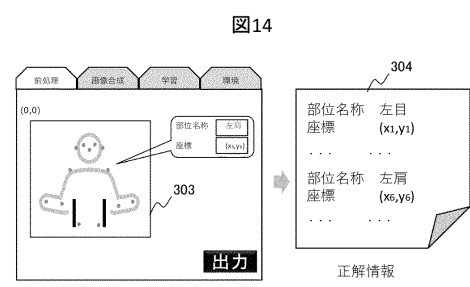
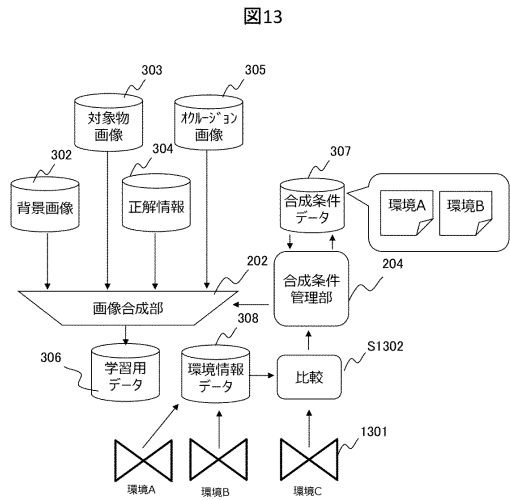
30

40

50

【図 1 3】

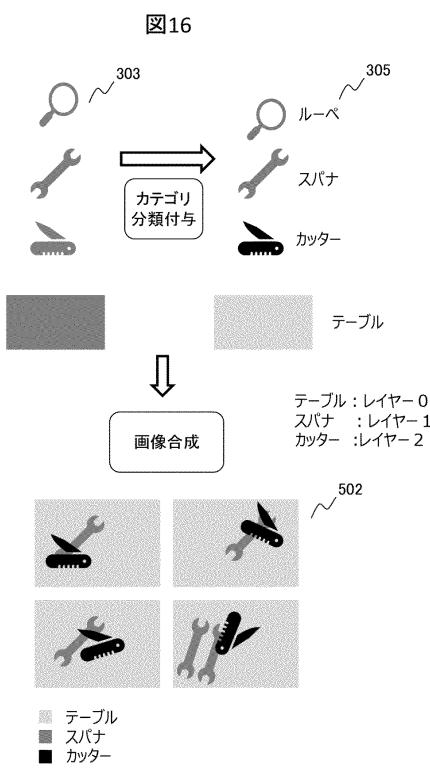
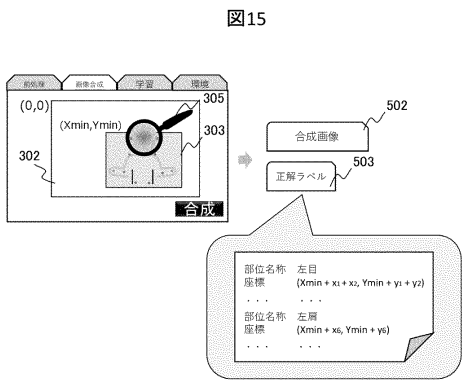
【図 1 4】



10

【図 1 5】

【図 1 6】



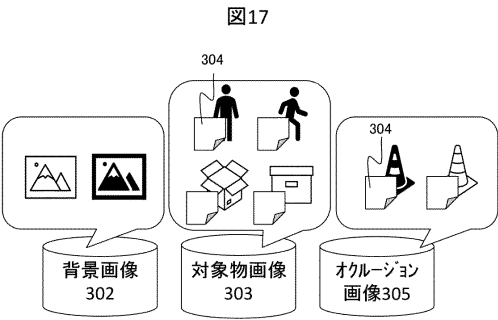
20

30

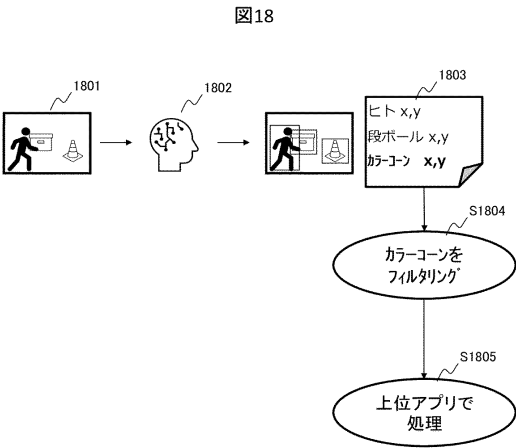
40

50

【図 1 7】



【図 1 8】



10

20

30

40

50

フロントページの続き

- (56)参考文献 特開 2 0 2 1 - 0 3 3 7 0 7 (J P , A)
国際公開第 2 0 2 0 / 1 0 2 7 6 7 (W O , A 1)
吉田英史,外5名,生成型学習法を用いた傘をさした歩行者の高精度な検出に関する検討,電子情報通信学会技術研究報告,日本,一般社団法人電子情報通信学会,2012年08月26日, Vol.112, No.198, p.191-196
- (58)調査した分野 (Int.Cl., D B 名)
- | | |
|---------|-------------|
| G 0 6 T | 7 / 0 0 |
| G 0 6 T | 5 / 5 0 |
| G 0 6 T | 3 / 4 0 4 6 |
| G 0 6 N | 2 0 / 0 0 |