

(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局

(43) 国際公開日
2025年1月23日(23.01.2025)



(10) 国際公開番号

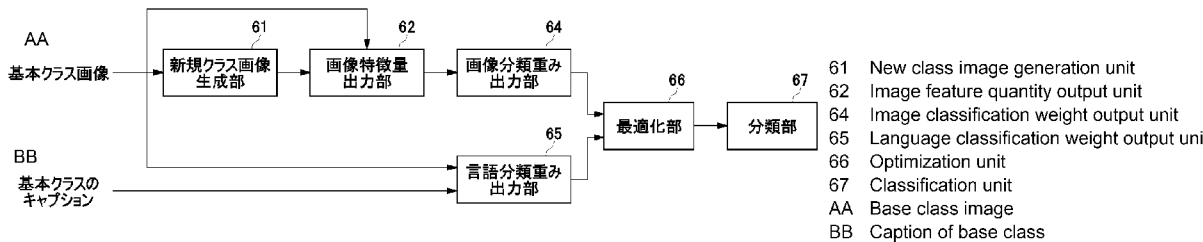
WO 2025/018010 A1

- (51) 国際特許分類:
G06V 10/764 (2022.01) G06T 7/00 (2017.01)
G06N 20/00 (2019.01)
- (21) 国際出願番号: PCT/JP2024/017617
- (22) 国際出願日: 2024年5月13日(13.05.2024)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (30) 優先権データ:
特願 2023-117326 2023年7月19日(19.07.2023) JP
- (71) 出願人: 株式会社 J V C ケンウッド (JVCKENWOOD CORPORATION) [JP/JP];
〒2210022 神奈川県横浜市神奈川区守屋町
3丁目12番地 Kanagawa (JP).
- (72) 発明者: 木田 晋吾 (KIDA Shingo); 〒2210022
神奈川県横浜市神奈川区守屋町3丁目
12番地株式会社 J V C ケンウッド 知
的財産部内 Kanagawa (JP).
- (74) 代理人: 森下 賢樹 (MORISHITA Sakaki);
〒1530061 東京都目黒区中目黒1-8-1 V
O R T 中目黒13階 Tokyo (JP).
- (81) 指定国(表示のない限り、全ての種類の国内保
護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA,
BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN,
CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC,
EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR,
HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, KE, KG, KH,
KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY,
MA, MD, MG, MK, MN, MU, MW, MX, MY, MZ,
NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT,
QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL,
ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG,
US, UZ, VC, VN, WS, ZA, ZM, ZW.
- (84) 指定国(表示のない限り、全ての種類の広域保
護が可能): ARIPO (BW, CV, GH, GM, KE, LR, LS,
MW, MZ, NA, RW, SC, SD, SL, ST, SZ, TZ, UG,
ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, RU,
TJ, TM), ヨーロッパ (AL, AT, BE, BG, CH, CY, CZ,
DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS,
IT, LT, LU, LV, MC, ME, MK, MT, NL, NO, PL, PT,

(54) Title: MACHINE LEARNING DEVICE, MACHINE LEARNING METHOD, AND MACHINE LEARNING PROGRAM

(54) 発明の名称: 機械学習装置、機械学習方法、および機械学習プログラム

【図7】



(57) Abstract: A new class image generation unit (61) processes an image of a base class and generates an image of a new class. An image feature quantity output unit (62) has been pre-trained by the image of the base class and uses the image of the base class or the image of the new class as input to output an image feature quantity. A language classification weight output unit (65) has been pre-trained by the image of the base class and a sentence describing the image of the base class, and uses the sentence describing the image of the base class as input to output a language classification weight. An image classification weight output unit (64) uses the image feature quantity as input to calculate the average value of the image feature quantity for each class and outputs the average value as an image classification weight for each class. An optimization unit (66) uses the image classification weight and the language classification weight as input to optimize the image classification weight and outputs a reconstruction classification weight.

[続葉有]

RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF,
CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE,
SN, TD, TG).

添付公開書類：

一 国際調査報告（条約第21条(3)）

(57) 要約：新規クラス画像生成部（61）は、基本クラスの画像を加工して、新規クラスの画像を生成する。画像特徴量出力部（62）は、基本クラスの画像によって事前学習されており、基本クラスの画像または新規クラスの画像を入力として画像特徴量を出力する。言語分類重み出力部（65）は、基本クラスの画像と、基本クラスの画像を説明する文章とによって事前学習されており、基本クラスの画像を説明する文章を入力として言語分類重みを出力する。画像分類重み出力部（64）は、画像特徴量を入力とし、クラス毎に画像特徴量の平均値を算出し、クラス毎の画像分類重みとして出力する。最適化部（66）は、画像分類重みと言語分類重みとを入力とし、画像分類重みを最適化し、再構築分類重みを出力する。

明 細 書

発明の名称：

機械学習装置、機械学習方法、および機械学習プログラム

技術分野

[0001] 本発明は、機械学習技術に関する。

背景技術

[0002] 人間は長期にわたる経験を通して新しい知識を学習することができ、昔の知識を忘れないように維持することができる。一方、畳み込みニューラルネットワーク (Convolutional Neural Network(CNN)) の知識は学習に使用したデータセットに依存しており、データ分布の変化に適応するためにはデータセット全体に対してCNNのパラメータの再学習が必要となる。

[0003] より効率的で実用的な手法として、既に獲得した知識を再利用しつつ、新たなタスクを学習する継続学習 (incremental learningまたはcontinual learning) という手法がある。特に、クラス分類タスクにおける継続学習は基本クラス (過去に学習したクラス) が分類できる状態から新しいクラス (新規クラス) を学習して分類できるようにする手法である。

[0004] 一方、深層学習では、過去に獲得した知識を大幅に喪失し、タスクの能力が大幅に低下する致命的忘却(catastrophic forgetting)という現象があるが、継続学習ではこれが特に問題となる。クラス分類タスクにおける継続学習では、致命的忘却を抑制して基本クラスの分類性能を維持しつつ、新規クラスの分類性能を獲得することが最大の課題となる。

[0005] 他方、新しいタスクは数少ないサンプルデータしか利用できないことが多いため、少ない教師データで効率的に学習する手法として、少数ショット学習 (few-shot learning) が提案されている。通常であれば学習には数千以上のサンプルが必要であるが、少数ショット学習では少数サンプル (例えば数個のサンプル) で学習が行われる。

[0006] また、基本 (ベース) クラスを学習済みのモデルに対して追加で学習させ

、新たなクラス（新規クラス）の分類を可能にするクラスインクリメンタル学習（Class Incremental Learning : CIL、継続学習）が提案されている。CILでは、クラス分類の学習済みモデルに対してタスクが継続的に追加され、新規タスクでは新規のクラス及び過去のクラスの分類性能が求められる。なお、通常、新規タスクの学習データはビッグデータである。

[0007] 基本（ベース）クラスの学習結果に対して致命的忘却を伴わずに新規クラスを学習する継続学習と、基本クラスに比べて少数しかない新規クラスを学習する少数ショット学習とを組み合わせたインクリメンタル少数ショット学習（Few-Shot Class Incremental Learning (FSCIL)）と呼ばれる手法が提案されている（非特許文献1）。継続少数ショット学習は、基本クラスについては大規模なデータセットから学習し、新規クラスについては少数のサンプルデータから学習することができる。FSCILはCILに近いクラス分類のインクリメンタル学習シナリオだが、新規タスクの学習データが少数（スモールデータ）である点が大きく異なる。

[0008] 継続少数ショット学習手法としてCEC (Continually Evolved Classifiers)が提案されている（非特許文献1）。CECは、元の画像を回転させた基本クラスの画像を疑似的に新しい新規クラスの画像とすることにより、疑似継続学習タスクを構築し、グラフアテンションネットワーク（graph attention network (GAT)）を学習する。

先行技術文献

非特許文献

[0009] 非特許文献1 : Zhang, C., Song, N., Lin, G., Zheng, Y., Pan, P., & Xu, Y. (2021). Few-shot incremental learning with continually evolved classifiers. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 12455-12464).

非特許文献2 : Nishida, K., Nishida, K., & Nishioka, S. (2022). Improving Few-Shot Image Classification Using Machine-and User-Generated Natural Language Descriptions. arXiv preprint arXiv:2207.03133.

非特許文献3: Chen, T., Kornblith, S., Norouzi, M., & Hinton, G. (2020, November). A simple framework for contrastive learning of visual representations. In International conference on machine learning (pp. 1597-1607). PMLR.

発明の概要

- [0010] 非特許文献1に記載の手法では、基本クラスの画像を分類するための特徴表現は既に学習済みであるため、回転させただけの学習済み画像を使用するだけでは、グラフモデルの学習が不十分になる可能性がある。そのため、十分な分類精度を得ることができない場合があるという課題があった。
- [0011] 本実施形態はこうした状況に鑑みてなされたものであり、その目的は、基本クラスの視覚的概念を含む言語情報を用いて特徴表現を最適化することによって、分類精度の向上を図ることができる機械学習技術を提供することにある。
- [0012] 上記課題を解決するために、本実施形態のある態様の機械学習装置は、基本クラスの画像に比べて少数の新規クラスの画像を継続学習する機械学習装置であって、前記基本クラスの画像を加工して、新規クラスの画像を生成する新規クラス画像生成部と、前記基本クラスの画像によって事前学習された画像特徴量出力部であって、前記基本クラスの画像または前記新規クラスの画像を入力として画像特徴量を出力する画像特徴量出力部と、前記基本クラスの画像と、前記基本クラスの画像を説明する文章とによって事前学習された言語分類重み出力部であって、前記基本クラスの画像を説明する文章を入力として言語分類重みを出力する言語分類重み出力部と、前記画像特徴量を入力とし、クラス毎に前記画像特徴量の平均値を算出し、クラス毎の画像分類重みとして出力する画像分類重み出力部と、前記画像分類重みと前記言語分類重みとを入力とし、前記画像分類重みを最適化し、再構築分類重みを出力する最適化部と、前記再構築分類重みをクラス分類の重みとして用い、前記画像特徴量出力部が出力した画像特徴量と前記クラス分類の重みとから、クラス分類を出力する分類部とを備える。

[0013] 本実施形態の別の態様は、機械学習方法である。この方法は、基本クラスの画像に比べて少数の新規クラスの画像を継続学習する機械学習方法であって、前記基本クラスの画像を加工して、新規クラスの画像を生成する新規クラス画像生成ステップと、前記基本クラスの画像によって事前学習された画像特徴量出力モジュールを用いて、前記基本クラスの画像または前記新規クラスの画像を入力として画像特徴量を出力する画像特徴量出力ステップと、前記基本クラスの画像と、前記基本クラスの画像を説明する文章とによって事前学習された言語分類重み出力モジュールを用いて、前記基本クラスの画像を説明する文章を入力として言語分類重みを出力する言語分類重み出力ステップと、前記画像特徴量を入力とし、クラス毎に前記画像特徴量の平均値を算出し、クラス毎の画像分類重みとして出力する画像分類重み出力ステップと、前記画像分類重みと前記言語分類重みとを入力とし、前記画像分類重みを最適化し、再構築分類重みを出力する最適化ステップと、前記再構築分類重みをクラス分類の重みとして用い、前記画像特徴量出力ステップが出力した画像特徴量と前記クラス分類の重みとから、クラス分類を出力する分類ステップとを備える。

[0014] なお、以上の構成要素の任意の組合せ、本実施形態の表現を方法、装置、システム、記録媒体、コンピュータプログラムなどの間で変換したものもまた、本実施形態の態様として有効である。

[0015] 本実施形態によれば、基本クラスの視覚的概念を含む言語情報を用いて特徴表現を最適化することによって、分類精度の向上を図ることができる機械学習技術を提供することができる。

図面の簡単な説明

[0016] [図1]従来のCECの手法を説明するための図である。

[図2]図2(a)、図2(b)は、従来のCECにおいて、GATの学習に使用する基本クラスの画像の回転角度に対する分類精度を示す図である。

[図3]従来の疑似継続学習におけるGATに入力されるクラス分類重みの例を示す図である。

[図4]本実施の形態の疑似継続学習におけるG A Tに入力されるクラス分類重みと言語分類重みの例を示す図である。

[図5]図5（a）、図5（b）は、疑似継続学習におけるG A Tの出力について従来技術と本実施の形態を比較した図である。

[図6]C E Cを利用する従来の機械学習装置の疑似継続学習モジュールの構成を説明するための機能ブロック図である。

[図7]C E Cを利用する本実施の形態の機械学習装置の疑似継続学習モジュールの構成を説明するための機能ブロック図である。

[図8]C E Cを利用する本実施の形態の機械学習装置の疑似継続学習モジュールの別の構成を説明するための機能ブロック図である。

発明を実施するための形態

[0017] C E CはF S C I Lの課題である基本クラスの忘却、新規クラスへのオーバーフィッティングを特徴抽出器と分類器の分離、グラフモデルによる分類器間の文脈情報の伝播によって改善する手法である。

[0018] 図1は、従来のC E Cの手法を説明するための図である。図1に示すように、C E Cはステージ1～3で構成される。従来の機械学習装置100は、ステージ1で用いられる事前トレーニングモジュール30と、ステージ2で用いられる疑似継続学習モジュール40と、ステージ3で用いられる新規クラス学習モジュール50と、を含む。

[0019] ステージ1は、事前学習ステージである。ステージ1では、大量の基本クラスのデータセット（以下、基本データセット）10を使用して、事前トレーニングモジュール30において標準的な教師あり学習で事前トレーニングモジュール30のバックボーンC N N 32の重みが事前学習される。基本データセット10はN個のデータサンプルを含む。データサンプルの一例は画像データであるが、これに限定されない。例えば、C I F A R 100データセットの場合、基本データセット10は60クラス×500枚の画像データを含む。基本データセット10は、複数の異なるクラスのデータセットを含んでもよい。バックボーンC N N 32は、基本データセット10を事前学習

した畳み込みニューラルネットワークである。バックボーンCNN32は、特徴抽出器Rの重みと基本クラスの分類器の重みベクトルである基本クラス分類重み W_0 とを有する。基本クラス分類重み W_0 は、基本データセット10のデータサンプルの平均特徴量を示す。事前学習したバックボーンCNN32の特徴抽出器Rのパラメータを以降のステージで固定することにより、基本クラスの忘却が抑制される。

[0020] ステージ2は、疑似継続学習ステージである。ステージ2では、疑似継続学習モジュール40において、各クラスの文脈情報を伝播させて全てのクラスに適応した分類器を生成するためにGAT44の重みが学習される。GAT44の学習は、基本データセット10の画像を回転させることによって生成した回転画像のデータセットから疑似継続学習タスクを構築することにより、エピソード形式で行う。以下、疑似継続学習ステージにおいて基本データセット10に基づいて生成されたデータセットを疑似データセット15という。

[0021] ステージ2では、基本クラスの別のデータセットである疑似データセット15をステージ1で事前学習したバックボーンCNN32の特徴抽出器Rに入力することにより生成された特徴ベクトルに基づいて、基本クラス分類重みが学習される。ステージ1で学習された基本クラス分類重み W_0 とステージ2で学習された基本クラス分類重みとを疑似継続学習モジュールのGAT44に入力することにより、GAT44においてこれらの基本クラス分類重みが適応されて再構築されることにより、 W'_0 の再構築分類重み45が出力される。以下、GATから出力される再構築された分類重みを再構築分類重みという。

[0022] エピソード形式について説明する。各エピソードは、サポートセットとクエリセットから構成される。疑似継続学習ステージにおいて、サポートセット及びクエリセットはいずれも基本データセット10と疑似データセット15とで構成される。ステージ2では、各エピソードにおいて、与えられたサポートセットのサポートサンプルに基づいて、クエリセットに含まれる基本

データセット 10 と疑似データセット 15 の両方のクエリサンプルをクラス分類し、クラス分類の損失を最小化するように GAT 44 のパラメータを更新する。

[0023] ここで、疑似継続学習タスクにおいて基本クラスの回転画像が使用されるのは、バックボーン CNN 32 はステージ 1 で既に基本クラスの画像をうまく分類するための特徴表現を学習していることから、仮に基本クラスの画像をそのまま利用すると GAT 44 がうまく学習されないためである。学習後の GAT 44 のパラメータは、以降のステージでは固定される。

[0024] ステージ 3 は、分類器学習及び適応ステージである。ステージ 3 では、新規クラス学習モジュール 50 において、セッション毎に与えられた少数の新規クラスのデータセット（以下、新規データセット）20 を使用して分類器を学習し、現在のセッションと以前のセッションで学習された全ての分類器を新規クラス学習モジュール 50 の GAT 53 に入力することにより、全ての分類器が適応される。新規クラス学習モジュール 50 の GAT 53 は、疑似継続学習ステージで学習した GAT である。GAT 53 によって適応された分類器によってクエリの推論が行われる。新規データセット 20 は、基本データセット 10 よりも少数の k 個のデータサンプルを含む。新規データセット 20 は、複数の異なるクラスのデータセットを含んでもよい。

[0025] ステージ 3 では、新規データセット 20 をステージ 1 で事前学習したバックボーン CNN 32 の特徴抽出器 R に入力することにより生成された特徴ベクトルに基づいて、セッション毎に新規クラス分類重みが学習される。ステージ 1 で学習された基本クラス分類重み W_0 と、ステージ 3 の i 番目のセッションまでの各セッションで学習された全ての新規クラス分類重み $\{W_1, \dots, W_i\}$ とを新規クラス学習モジュール 50 の GAT 53 に入力することにより、GAT 53 に入力された全てのクラスの分類重みが適応されて再構築され、GAT 53 から $\{W'_0, W'_1, \dots, W'_i\}$ の再構築分類重み 54 が出力される。

[0026] 本実施の形態の機械学習装置 200 では、ステージ 2 の疑似継続学習に着目し、疑似継続学習モジュール 40 に改良を加える。それ以外の構成につい

ては従来の機械学習装置100と同じである。

[0027] 図2(a)、図2(b)は、従来のCECにおいて、GATの学習に使用する基本クラスの画像の回転角度に対する分類精度を示す図である。図2(a)は回転角度に対する平均分類精度を示し、図2(b)は回転角度に対する初期セッションから最終セッションまでの平均分類精度の低下率を示す。図2(a)から、回転角度が 90° 、 180° 、 270° の場合に分類精度が高いことが確認できる。また、図2(b)から、回転角度が 90° 、 180° 、 270° の場合に初期セッションから最終セッションまでの平均分類精度の低下率が小さく、基本クラスの忘却が抑制されることがわかる。このことから疑似継続学習では、基本クラスの画像から視覚的にかけ離れた画像を使用することが分類精度向上の観点から望ましいと考えられる。

[0028] 図3は、従来の疑似継続学習におけるGATに入力されるクラス分類重みの例を示す図である。ここでは、特徴空間における各クラスの分類重みを2次元空間に視覚化したものを図示している。クラス分類重みは「プロトタイプ」とも呼ばれる。

[0029] GATにはランダムに選択された基本クラスの5クラスのプロトタイプ(B1~B5)と、その基本クラスの画像を回転させて疑似的に生成した新規クラスの5クラスのプロトタイプ(N1~N5)が入力される。新規クラスN1、N2、N3、N4、N5はそれぞれ基本クラスB1、B2、B3、B4、B5を回転させたものである。

[0030] 図3に示されるように、基本クラスの画像特徴を平均したプロトタイプとその基本クラスの回転画像の特徴を平均した新規クラスのプロトタイプでは、視覚的に大きな差がなく特徴空間上で近い位置となるため、GATの学習が不十分となる可能性がある。さらに、GATのパラメータの学習は、クロスエントロピー損失のみで最適化されているためプロトタイプの調整が限定的になる可能性がある。

[0031] 図4は、本実施の形態の疑似継続学習におけるGATに入力されるプロトタイプと言語分類重みの例を示す図である。ここでは、特徴空間における各

クラスのプロトタイプと言語分類重みを2次元空間に視覚化したものを図示している。

[0032] GATにはランダムに選択された基本クラスの5クラスのプロトタイプ（B1～B5）と、その基本クラスの画像を回転させて疑似的に生成した新規クラスの5クラスのプロトタイプ（N1～N5）と、基本クラスの5クラス（B1～B5）の言語分類重み（T1～T5）が入力される。新規クラスN1、N2、N3、N4、N5はそれぞれ基本クラスB1、B2、B3、B4、B5を回転させたものである。言語分類重みT1、T2、T3、T4、T5はそれぞれ基本クラスB1、B2、B3、B4、B5の言語分類重みである。

[0033] ここで、言語分類重みは、一例として非特許文献2に記載されるように、画像と言語をペアにしたビッグデータで十分に学習した学習済Text Encoderモデルによって生成される基本クラスの視覚的概念を含む特徴表現を用いる。具体的には、Text Encoderを用いて基本クラスの画像を説明する文章から基本クラスの言語分類重みを生成する。なお、言語分類重みは非特許文献2ではtext representation、text featureとされているものである。

[0034] GATのパラメータの学習には従来のクロスエントロピー損失に加えて、非特許文献3に記載されるコントラストティブ（対照的）損失（Contrastive Loss）を使用する。具体的には、基本クラスのプロトタイプは、基本クラスの視覚的概念を含む言語分類重みに近づけ、新規クラスのプロトタイプは、回転元の基本クラスの言語分類重みから遠ざけるようにGATのパラメータを疑似継続学習させる。

[0035] 図5（a）、図5（b）は、疑似継続学習におけるGATの出力について従来技術と本実施の形態を比較した図である。GATの出力は、疑似継続学習によって調整された後のプロトタイプである。

[0036] 図5（a）に示すように、従来技術ではプロトタイプの移動量が小さく、そのため分離精度が限定的である。

[0037] 図5(b)に示すように、本実施の形態では、基本クラスの調整後のプロトタイプは、基本クラスの言語分類重みに近づき、新規クラスの調整後のプロトタイプは、回転元の基本クラスの言語分類重みから遠ざかる。このようにして言語分類重みを用いた最適化によりGATを効果的に学習することができるため、プロトタイプの移動量が増し、各クラスの分類精度が向上する。

[0038] 図6は、CECを利用する従来の機械学習装置100の疑似継続学習モジュール40の構成を説明するための機能ブロック図である。疑似継続学習モジュール40は、新規クラス画像生成部61、画像特徴量出力部62、画像分類重み出力部64、最適化部66、および分類部67を含む。

[0039] 新規クラス画像生成部61は、事前トレーニングモジュール30で用いられた基本データセット10の基本クラスの画像を回転させることによって新規クラス画像の疑似データセット15を生成し、画像特徴量出力部62に供給する。

[0040] 画像特徴量出力部62は、新規クラス画像の疑似データセット15を入力として、新規クラス画像の疑似データセット15の特徴ベクトルを抽出し、抽出した画像特徴量を画像分類重み出力部64に供給する。画像特徴量出力部62は、ステージ1で基本クラス分類重みを事前学習したバックボーンCNN32の特徴抽出器Rに対応する。

[0041] 画像分類重み出力部64は、新規クラス画像の疑似データセット15の特徴ベクトルをクラス毎に平均化することにより、新規クラス画像の疑似データセット15の画像分類重みを算出し、最適化部66に供給する。

[0042] 最適化部66は、GAT44に対応し、ステージ1で基本クラスの分類重みを事前学習済みのバックボーンCNN32の基本クラス分類重み W_0 と、画像分類重み出力部64から供給される疑似データセット15の基本クラス分類重みとを入力とし、基本データセット10と疑似データセット15との間の依存関係をメタ学習して、入力した全てのクラス分類重みを適応させることにより、再構築分類重みを出力する。疑似継続学習モジュール40では、

メタモジュールであるGATをエピソード形式でメタ学習する。基本データセット10と疑似データセット15で構成されるクエリセットを用いて、エピソード毎に、最適化部66のパラメータが最適化されて更新される。最適化部66の最適化方法は非特許文献1に記載されるクロスエントロピー損失を最小化する方法を用いる。最適化部66はこのようにして得られた再構築分類重みを分類部67に供給する。

[0043] 分類部67は、再構築分類重みをクラス分類の重みとして用い、画像特徴量出力部62が出力した画像特徴量とクラス分類の重みとからクラス分類を出力する。

[0044] 図7は、CECを利用する本実施の形態の機械学習装置200の疑似継続学習モジュール40の構成を説明するための機能ブロック図である。疑似継続学習モジュール40は、新規クラス画像生成部61、画像特徴量出力部62、画像分類重み出力部64、言語分類重み出力部65、最適化部66、および分類部67を含む。ここでは、図6の従来の機械学習装置100の疑似継続学習モジュール40の機能ブロックと共通する構成と動作については適宜省略して、異なる構成と動作について説明する。

[0045] 言語分類重み出力部65は、基本クラス画像と基本クラスの画像を説明する文章（「キャプション」と呼ぶ）とによって事前学習されており、基本クラスの画像のキャプションを入力として基本クラスの画像の言語特徴量である言語分類重みを生成し、最適化部66に供給する。

[0046] 最適化部66は、画像分類重みと言語分類重みとを入力として、画像分類重みを最適化して再構築分類重みを算出し、分類部67に供給する。具体的には、最適化部66は、基本クラスの画像分類重みは言語分類重みに近づけ、新規クラスの画像分類重みは言語分類重みから遠ざけるようにコンストラスティブ損失を最小化することにより、再構築分類重みを算出する。

[0047] 本実施の形態の機械学習装置200の疑似継続学習モジュール40では、言語分類重みを基準とするコンストラスティブ損失を最小化することにより、従来技術よりもクラス間の距離が離れるため、分類精度が向上する。

[0048] 図8は、CECを利用する本実施の形態の機械学習装置200の疑似継続学習モジュール40の別の構成を説明するための機能ブロック図である。疑似継続学習モジュール40は、新規クラス画像生成部61、画像特徴量出力部62、言語特徴量出力部63、画像分類重み出力部64、言語分類重み出力部65、最適化部66、および分類部67を含む。ここでは、図7の機械学習装置200の疑似継続学習モジュール40の機能ブロックとは、言語特徴量出力部63がさらに設けられる点異なる。

[0049] 基本クラスの画像1つに対して複数のキャプションがある場合、言語特徴量出力部63は複数のキャプションのそれぞれから言語特徴量を抽出し、言語分類重み出力部65に供給する。言語分類重み出力部65は、言語特徴量を平均化することにより言語分類重みを算出し、最適化部66に供給する。それ以外の構成と動作は図7と同一である。

[0050] 以上説明した機械学習装置200の各種の処理は、CPUやメモリ等のハードウェアを用いた装置として実現することができるのは勿論のこと、ROM（リード・オンリ・メモリ）やフラッシュメモリ等に記憶されているファームウェアや、コンピュータ等のソフトウェアによっても実現することができる。そのファームウェアプログラム、ソフトウェアプログラムをコンピュータ等で読み取り可能な記録媒体に記録して提供することも、有線あるいは無線のネットワークを通してサーバと送受信することも、地上波あるいは衛星デジタル放送のデータ放送として送受信することも可能である。

[0051] 以上、本発明を実施の形態をもとに説明した。実施の形態は例示であり、それらの各構成要素や各処理プロセスの組合せにいろいろな変形例が可能なこと、またそうした変形例も本発明の範囲にあることは当業者に理解されるところである。

[0052] 基本クラスの画像から新規クラスの画像を生成する方法として、基本クラスの画像を回転させる以外の方法で基本クラスの画像を加工して新規クラスの画像を生成してもよい。たとえば、基本クラスの画像を複数の領域に分割し、分割した領域を互いに入れ替えることで新規クラスの画像を生成しても

よい。

産業上の利用可能性

[0053] 本発明は、機械学習技術に利用できる。

符号の説明

[0054] 10 基本データセット、 15 疑似データセット、 20 新規データセット、 30 事前トレーニングモジュール、 32 バックボーンCNN、 40 疑似継続学習モジュール、 44、53 GAT、 45、54 再構築分類重み、 50 新規クラス学習モジュール、 61 新規クラス画像生成部、 62 画像特徴量出力部、 63 言語特徴量出力部、 64 画像分類重み出力部、 65 言語分類重み出力部、 66 最適化部、 67 分類部、 100、200 機械学習装置。

請求の範囲

- [請求項1] 基本クラスの画像に比べて少数の新規クラスの画像を継続学習する機械学習装置であって、
- 前記基本クラスの画像を加工して、新規クラスの画像を生成する新規クラス画像生成部と、
- 前記基本クラスの画像によって事前学習された画像特徴量出力部であって、前記基本クラスの画像または前記新規クラスの画像を入力として画像特徴量を出力する画像特徴量出力部と、
- 前記基本クラスの画像と、前記基本クラスの画像を説明する文章とによって事前学習された言語分類重み出力部であって、前記基本クラスの画像を説明する文章を入力として言語分類重みを出力する言語分類重み出力部と、
- 前記画像特徴量を入力とし、クラス毎に前記画像特徴量の平均値を算出し、クラス毎の画像分類重みとして出力する画像分類重み出力部と、
- 前記画像分類重みと前記言語分類重みとを入力とし、前記画像分類重みを最適化し、再構築分類重みを出力する最適化部と、
- 前記再構築分類重みをクラス分類の重みとして用い、前記画像特徴量出力部が出力した画像特徴量と前記クラス分類の重みとから、クラス分類を出力する分類部とを備えることを特徴とする機械学習装置。
- [請求項2] 前記最適化部は、
- 前記基本クラスの画像分類重みは前記言語分類重みに近づけ、前記新規クラスの画像分類重みは前記言語分類重みから遠ざけるように学習することを特徴とする請求項1に記載の機械学習装置。
- [請求項3] 基本クラスの画像に比べて少数の新規クラスの画像を継続学習する機械学習方法であって、
- 前記基本クラスの画像を加工して、新規クラスの画像を生成する新規クラス画像生成ステップと、

前記基本クラスの画像によって事前学習された画像特徴量出力モジュールを用いて、前記基本クラスの画像または前記新規クラスの画像を入力として画像特徴量を出力する画像特徴量出力ステップと、

前記基本クラスの画像と、前記基本クラスの画像を説明する文章とによって事前学習された言語分類重み出力モジュールを用いて、前記基本クラスの画像を説明する文章を入力として言語分類重みを出力する言語分類重み出力ステップと、

前記画像特徴量を入力とし、クラス毎に前記画像特徴量の平均値を算出し、クラス毎の画像分類重みとして出力する画像分類重み出力ステップと、

前記画像分類重みと前記言語分類重みとを入力とし、前記画像分類重みを最適化し、再構築分類重みを出力する最適化ステップと、

前記再構築分類重みをクラス分類の重みとして用い、前記画像特徴量出力ステップが出力した画像特徴量と前記クラス分類の重みとから、クラス分類を出力する分類ステップとを備えることを特徴とする機械学習方法。

[請求項4] 基本クラスの画像に比べて少数の新規クラスの画像を継続学習する機械学習プログラムであって、

前記基本クラスの画像を加工して、新規クラスの画像を生成する新規クラス画像生成ステップと、

前記基本クラスの画像によって事前学習された画像特徴量出力モジュールを用いて、前記基本クラスの画像または前記新規クラスの画像を入力として画像特徴量を出力する画像特徴量出力ステップと、

前記基本クラスの画像と、前記基本クラスの画像を説明する文章とによって事前学習された言語分類重み出力モジュールを用いて、前記基本クラスの画像を説明する文章を入力として言語分類重みを出力する言語分類重み出力ステップと、

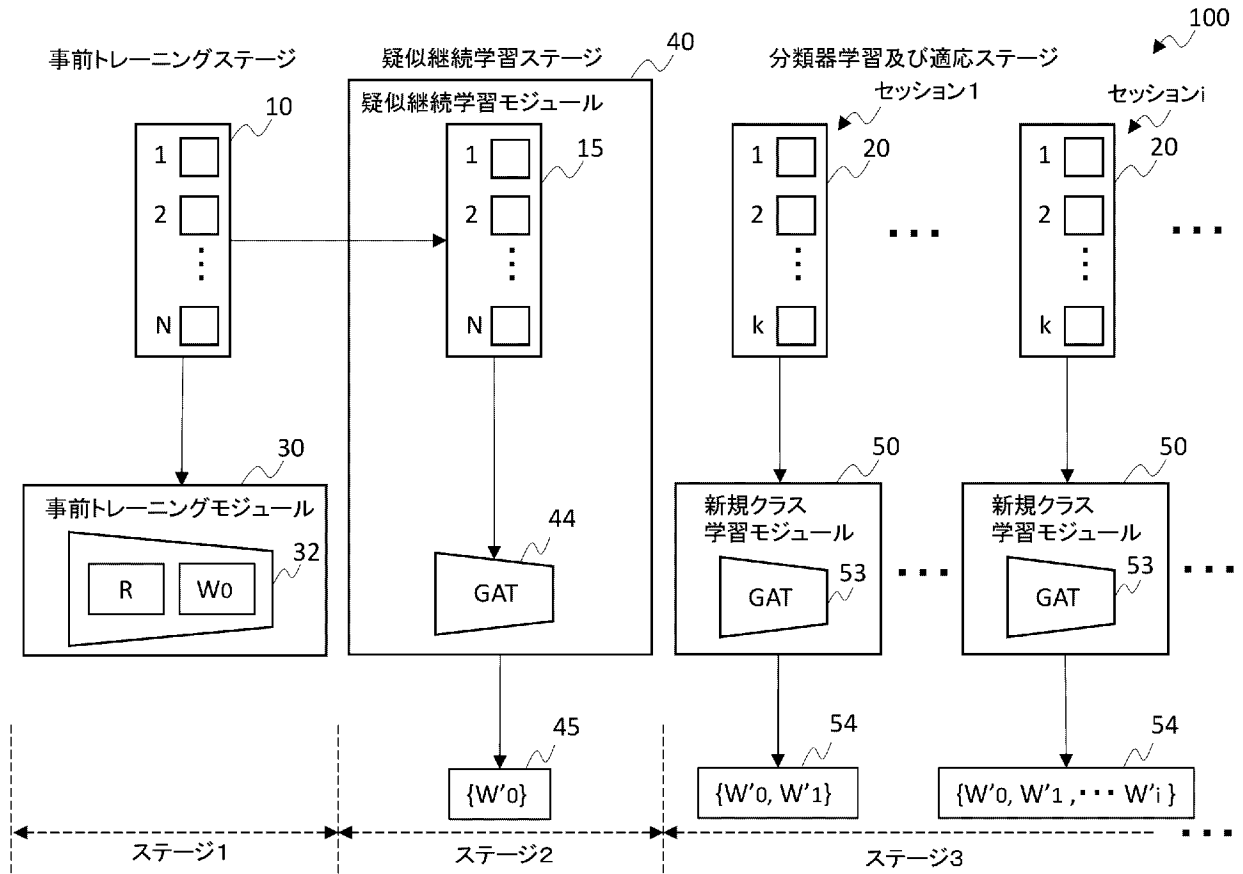
前記画像特徴量を入力とし、クラス毎に前記画像特徴量の平均値を

算出し、クラス毎の画像分類重みとして出力する画像分類重み出力ステップと、

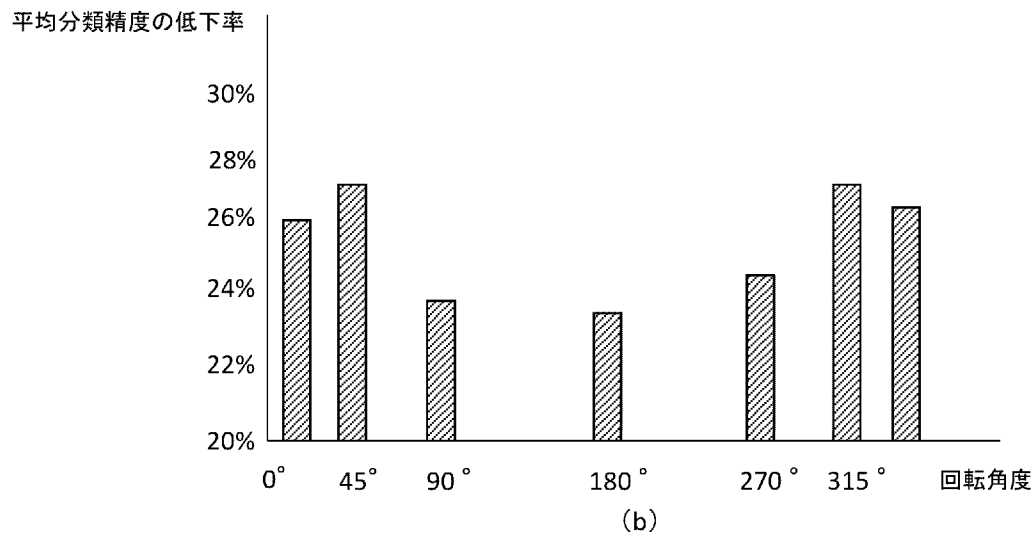
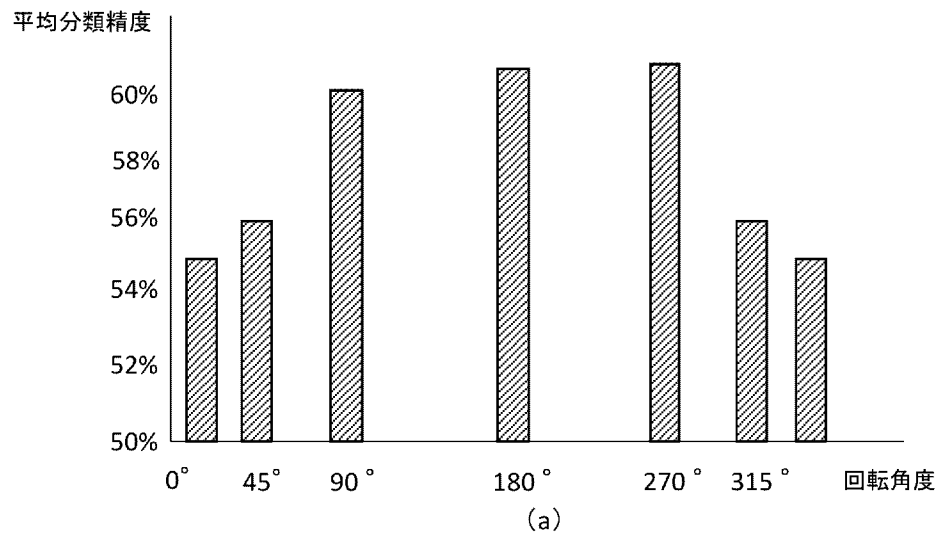
前記画像分類重みと前記言語分類重みとを入力とし、前記画像分類重みを最適化し、再構築分類重みを出力する最適化ステップと、

前記再構築分類重みをクラス分類の重みとして用い、前記画像特徴量出力ステップが出力した画像特徴量と前記クラス分類の重みとから、クラス分類を出力する分類ステップとをコンピュータに実行させることを特徴とする機械学習プログラム。

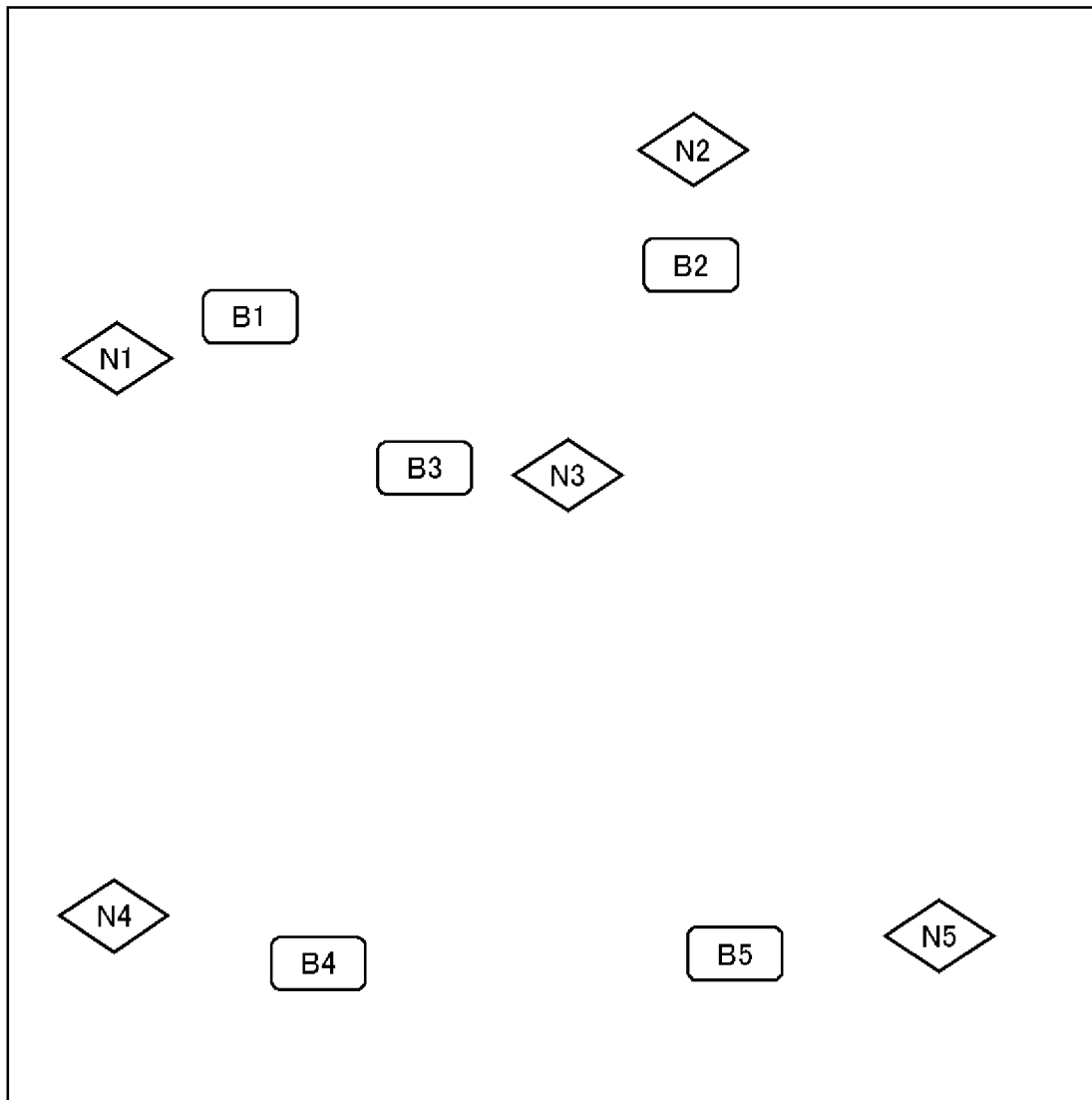
[図1]



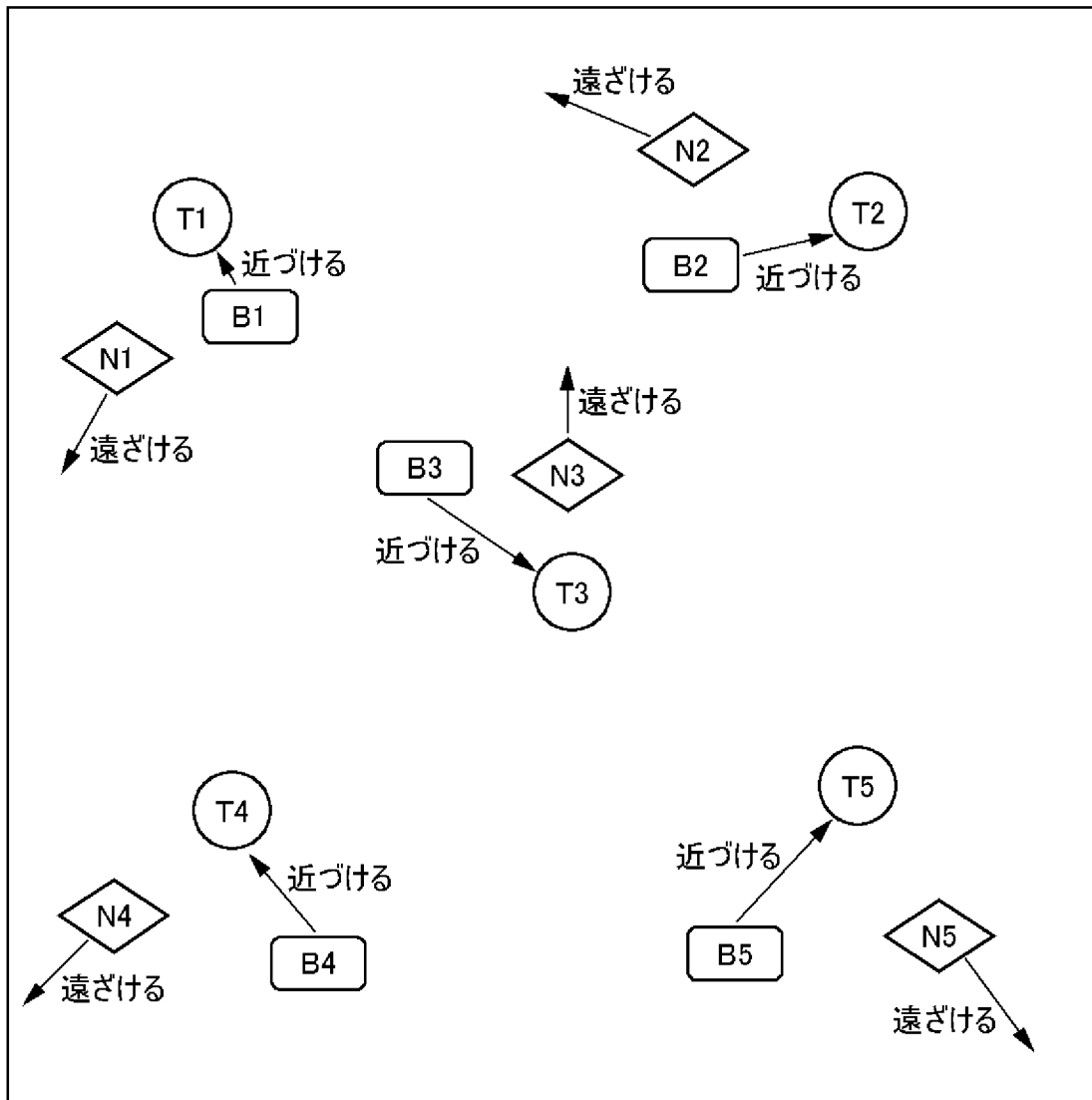
[図2]



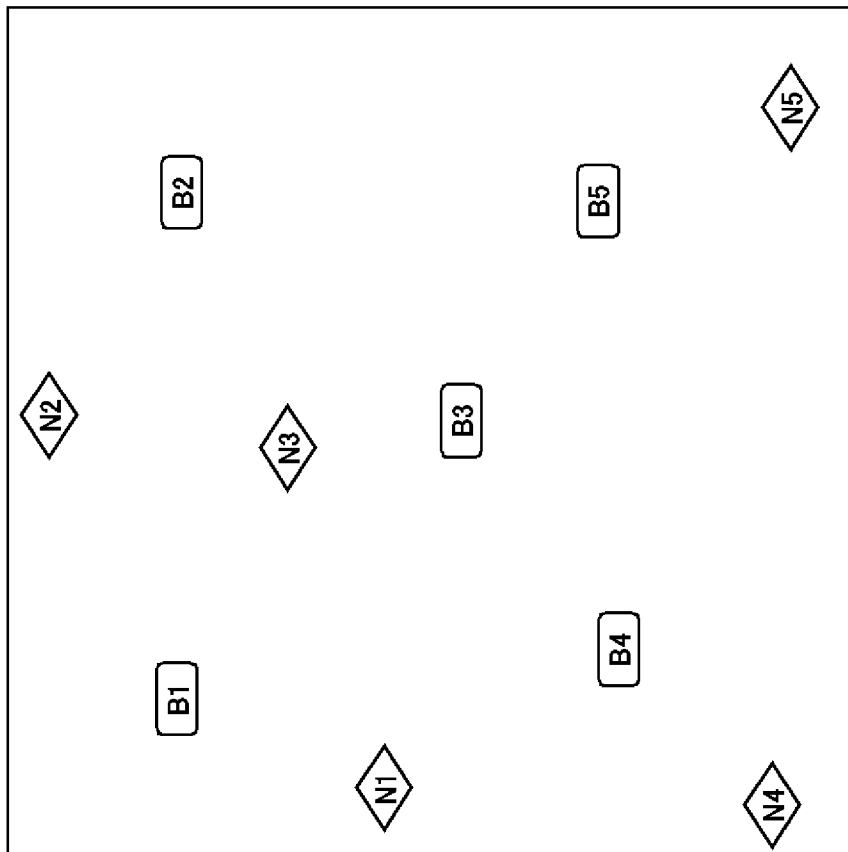
[図3]



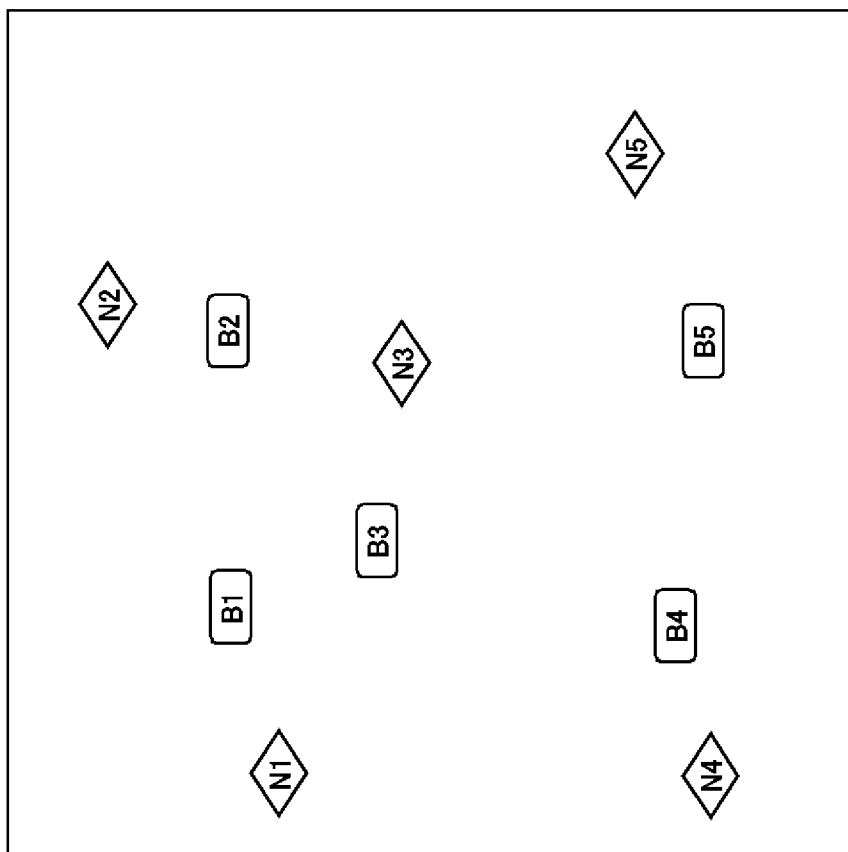
[図4]



[図5]

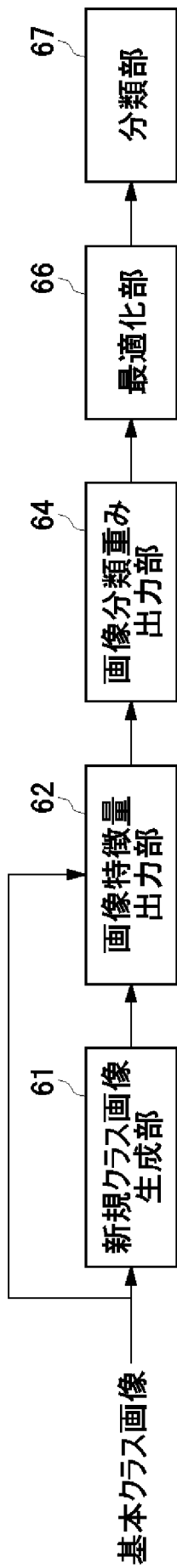


(b)

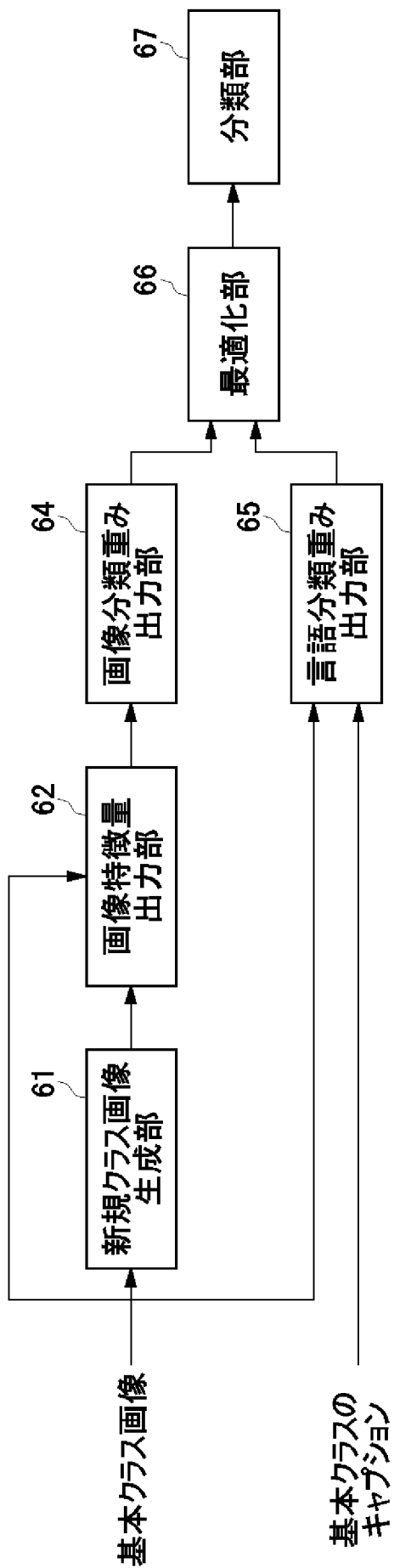


(a)

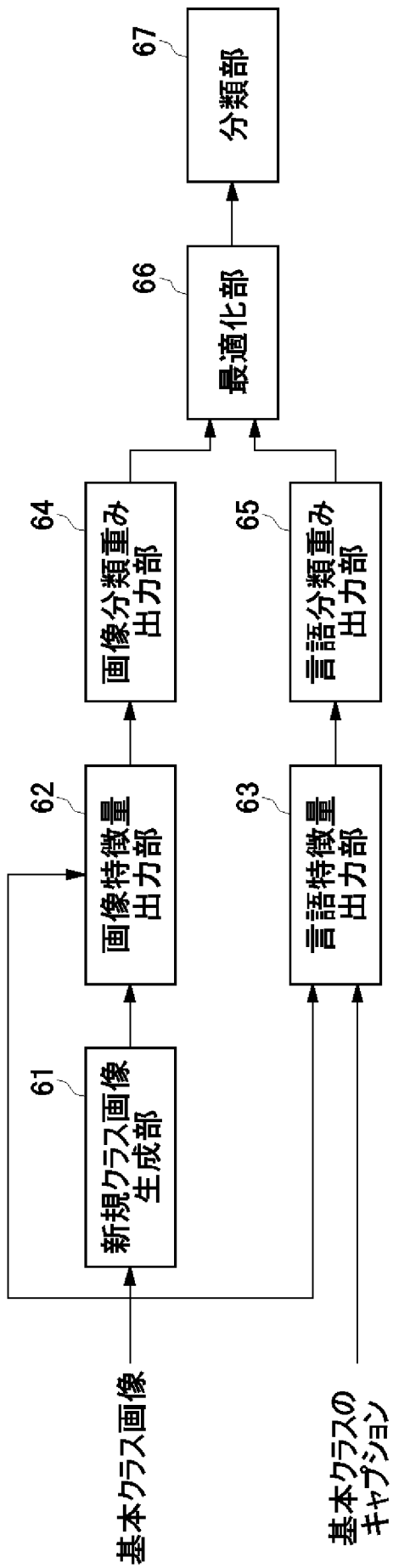
[図6]



[図7]



[図8]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2024/017617

A. CLASSIFICATION OF SUBJECT MATTER**G06V 10/764**(2022.01)i; **G06N 20/00**(2019.01)i; **G06T 7/00**(2017.01)i

FI: G06V10/764; G06N20/00; G06T7/00 350B

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06V10/00; G06N20/00; G06T7/00

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Published examined utility model applications of Japan 1922-1996

Published unexamined utility model applications of Japan 1971-2024

Registered utility model specifications of Japan 1996-2024

Published registered utility model applications of Japan 1994-2024

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 2023-48171 A (JVC KENWOOD CORP.) 07 April 2023 (2023-04-07) paragraphs [0002]-[0008], [0014]-[0088], fig. 1A-12	1-4
A	JP 2023-94215 A (JVC KENWOOD CORP.) 05 July 2023 (2023-07-05) paragraphs [0002]-[0009], [0016]-[0099], fig. 1A-11	1-4
P, A	JP 2024-45006 A (JVC KENWOOD CORP.) 02 April 2024 (2024-04-02) paragraphs [0002]-[0010], [0069]-[0074], fig. 11	1-4
P, A	JP 2024-33903 A (JVC KENWOOD CORP.) 13 March 2024 (2024-03-13) paragraphs [0002]-[0011], [0018]-[0076], fig. 1-11	1-4



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“D” document cited by the applicant in the international application

“E” earlier application or patent but published on or after the international filing date

“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

“O” document referring to an oral disclosure, use, exhibition or other means

“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

Date of the actual completion of the international search

12 June 2024

Date of mailing of the international search report

25 June 2024

Name and mailing address of the ISA/JP

Japan Patent Office (ISA/JP)
3-4-3 Kasumigaseki, Chiyoda-ku, Tokyo 100-8915
Japan

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/JP2024/017617

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)		Publication date (day/month/year)
JP	2023-48171	A	07 April 2023	WO	2023/053569	A1

JP	2023-94215	A	05 July 2023	WO	2023/119742	A1

JP	2024-45006	A	02 April 2024	(Family: none)		

JP	2024-33903	A	13 March 2024	WO	2024/047957	A1

A. 発明の属する分野の分類（国際特許分類（IPC））

G06V 10/764(2022.01)i; G06N 20/00(2019.01)i; G06T 7/00(2017.01)i

FI: G06V10/764; G06N20/00; G06T7/00 350B

B. 調査を行った分野

調査を行った最小限資料（国際特許分類（IPC））

G06V10/00; G06N20/00; G06T7/00

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報

1922-1996年

日本国公開実用新案公報

1971-2024年

日本国実用新案登録公報

1996-2024年

日本国登録実用新案公報

1994-2024年

国際調査で使用した電子データベース（データベースの名称、調査に使用した用語）

C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号
A	JP 2023-48171 A（株式会社 JVCケンウッド）07.04.2023（2023-04-07） 段落[0002]-[0008], [0014]-[0088], 図1A-12	1-4
A	JP 2023-94215 A（株式会社 JVCケンウッド）05.07.2023（2023-07-05） 段落[0002]-[0009], [0016]-[0099], 図1A-11	1-4
P, A	JP 2024-45006 A（株式会社 JVCケンウッド）02.04.2024（2024-04-02） 段落[0002]-[0010], [0069]-[0074], 図11	1-4
P, A	JP 2024-33903 A（株式会社 JVCケンウッド）13.03.2024（2024-03-13） 段落[0002]-[0011], [0018]-[0076], 図1-11	1-4

☐ C欄の続きにも文献が列挙されている。

☒ パテントファミリーに関する別紙を参照。

* 引用文献のカテゴリー

“A” 特に関連のある文献ではなく、一般的技術水準を示すもの

“D” 国際出願で出願人が先行技術文献として記載した文献

“E” 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの

“L” 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す）

“O” 口頭による開示、使用、展示等に言及する文献

“P” 国際出願日前で、かつ優先権の主張の基礎となる出願の日の後に公表された文献

“T” 国際出願日又は優先日後に公表された文献であって出願と抵触するものではなく、発明の原理又は理論の理解のために引用するもの

“X” 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの

“Y” 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの

“&” 同一パテントファミリー文献

国際調査を完了した日

12.06.2024

国際調査報告の発送日

25.06.2024

名称及びあて先

日本国特許庁(ISA/JP)

〒100-8915

日本国

東京都千代田区霞が関三丁目4番3号

権限のある職員（特許庁審査官）

長谷川 素直 5P 2948

電話番号 03-3581-1101 内線 3539

国際出願番号
PCT/JP2024/017617

様式 PCT/ISA/210 (パテントファミリー用別紙) (2022年7月)