

(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  
2020年3月5日(05.03.2020)



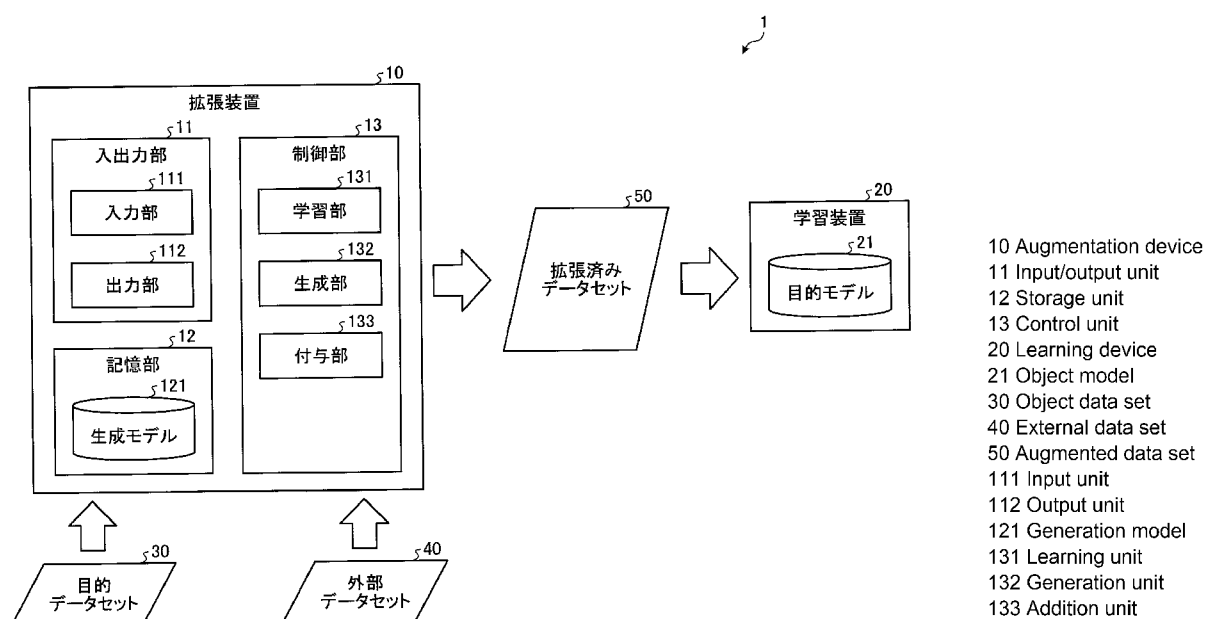
(10) 国際公開番号

WO 2020/045236 A1

- (51) 国際特許分類:  
G06N 20/00 (2019.01) G06T 7/00 (2017.01)
- (21) 国際出願番号: PCT/JP2019/032863
- (22) 国際出願日: 2019年8月22日(22.08.2019)
- (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 2018-158400 2018年8月27日(27.08.2018) JP
- (71) 出願人: 日本電信電話株式会社 (NIPPON TELEGRAPH AND TELEPHONE CORPORATION) [JP/JP]; 〒1008116 東京都千代田区大手町一丁目5番1号 Tokyo (JP).
- (72) 発明者: 山口 真弥 (YAMAGUCHI, Shinya); 〒1808585 東京都武蔵野市緑町3丁目9-1 NTT 知的財産センタ内 Tokyo (JP). 江田 毅晴(EDA, Takeharu); 〒1808585 東京都武蔵野市緑町3丁目9-1 NTT 知的財産センタ内 Tokyo (JP). 村松 沙那恵(MURAMATSU, Sanae); 〒1808585 東京都武蔵野市緑町3丁目9-1 NTT 知的財産センタ内 Tokyo (JP).
- (74) 代理人: 特許業務法人酒井国際特許事務所 (SAKAI INTERNATIONAL PATENT OFFICE); 〒1000013 東京都千代田区霞が関3丁目8番1号 虎の門三井ビルディング Tokyo (JP).

(54) Title: AUGMENTATION DEVICE, AUGMENTATION METHOD, AND AUGMENTATION PROGRAM

(54) 発明の名称: 拡張装置、拡張方法及び拡張プログラム



(57) Abstract: An augmentation device (10) causes a generation model for generating data from a label to learn second data and first data having the label added thereto. The augmentation device (10) generates augmentation data from the label added to the first data, by using the generation model which has learned the first and second data. The augmentation device (10) adds the label added to the first data, to augmented data in which the first data and the augmentation data have been integrated.

(57) 要約: 拡張装置 (10) は、ラベルからデータを生成する生成モデルに、ラベルが付与された第1のデータ及び第2のデータを学習させる。また、拡張装置 (10) は、第1のデータ及び第2のデータを学習した生成モデルを用いて、第1のデータに付与されたラベルから拡張用のデータを生成する。また、拡張装置 (10) は、第1のデータ及び拡張用のデータを統合した拡張済みデータに、第1のデータに付与されたラベルを付与する。

(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, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, 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, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, 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, GH, GM, KE, LR, LS, MW, MZ, NA, RW, 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, MK, MT, NL, NO, PL, PT, 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))

## 明 細 書

発明の名称： 拡張装置、拡張方法及び拡張プログラム

### 技術分野

[0001] 本発明は、拡張装置、拡張方法及び拡張プログラムに関する。

### 背景技術

[0002] 深層学習モデルにおける学習データの整備は、大きなコストを要する。学習データの整備には、学習データの収集だけでなく、学習データへのラベル等のアノテーションの付加が含まれる。

[0003] 従来、学習データの整備のコストを軽減するための技術として、ルールベースのデータ拡張 (Data Augmentation) が知られている。例えば、学習データとして用いられる画像に、反転、拡大縮小、ノイズ付加、回転等の特定のルールにしたがった変更を加えることで、別の学習データを生成する方法が知られている (例えば、非特許文献 1 又は 2 を参照)。また、学習データが音声やテキストである場合にも、同様のルールベースのデータ拡張が行われることがある。

### 先行技術文献

#### 非特許文献

[0004] 非特許文献1: Patrice Y. Simard, Dave Steinkraus, and John C. Platt. Best practices for convolutional neural networks applied to visual document analysis. In Proceedings of the Seventh International Conference on Document Analysis and Recognition - Volume 2, ICDAR '03, pp.958-, Washington, DC, USA, 2003. IEEE Computer Society.

非特許文献2: Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. Imagenet classification with deep convolutional neural networks. In Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1, NIPS

'12, pp. 1097-1105, USA, 2012. Curran Associates Inc.

非特許文献3: C. Szegedy, Wei Liu, Yangqing Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabino vich. Going deeper with convolutions. In 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1-9, June 2015.

非特許文献4: Tom Ko, Vijayaditya Peddinti, Daniel Povey, and S anjeev Khudanpur. Audio augmentation for speech recognition. I n INTERSPEECH, pp. 3586-3589. ISCA, 2015.

非特許文献5: Z. Xie, S. I. Wang, J. Li, D. Levy, A. Nie, D . Jurafsky, and A. Y. Ng. Data noising as smoothing in neu ral network language models. In International Conference on L earning Representations (ICLR), 2017.

非特許文献6: Mehdi Mirza, Simon Osindero: Conditional Generative Adversarial Nets. CoRR abs/1411.1784 (2014)

非特許文献7: D. Cheng, Y. Gong, S. Zhou, J. Wang and N. Zhe ng, "Person Re-identification by Multi-Channel Parts-Based CNN with Improved Triplet Loss Function," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, 2016, pp. 1335-1344. doi: 10.1109/CVPR.2016.149

## 発明の概要

### 発明が解決しようとする課題

[0005] しかしながら、従来の技術には、データ拡張により得られる学習データのバリエーションが少なく、モデルの精度を向上させられない場合があるという問題がある。具体的には、従来のルールベースのデータ拡張では、学習データの属性のバリエーションを増加させることが難しく、そのことがモデルの精度向上に限界を生じさせている。例えば、非特許文献1及び2に記載のルールベースのデータ拡張では、窓際にいる正面を向いた猫の画像の「窓際

」、「猫」及び「正面」をいった属性を変更した画像を生成することは困難である。

### 課題を解決するための手段

[0006] 上述した課題を解決し、目的を達成するために、拡張装置は、ラベルからデータを生成する生成モデルに、ラベルが付与された第1のデータ及び第2のデータを学習させる学習部と、前記第1のデータ及び前記第2のデータを学習した前記生成モデルを用いて、前記第1のデータに付与されたラベルから拡張用のデータを生成する生成部と、前記第1のデータ及び前記拡張用のデータを統合した拡張済みデータに、前記第1のデータに付与されたラベルを付与する付与部と、を有することを特徴とする。

### 発明の効果

[0007] 本発明によれば、データ拡張により得られる学習データのバリエーションを増加させ、モデルの精度を向上させることができる。

### 図面の簡単な説明

[0008] [図1]図1は、第1の実施形態に係る拡張装置の構成の一例を示す図である。

[図2]図2は、第1の実施形態に係る生成モデルの一例を示す図である。

[図3]図3は、第1の実施形態に係る生成モデルの学習処理を説明するための図である。

[図4]図4は、第1の実施形態に係る拡張画像の生成処理を説明するための図である。

[図5]図5は、第1の実施形態に係る付与処理を説明するための図である。

[図6]図6は、第1の実施形態に係る目的モデルの学習処理を説明するための図である。

[図7]図7は、第1の実施形態に係る拡張装置によって生成される拡張済みデータセットの一例を示す図である。

[図8]図8は、第1の実施形態に係る拡張装置の処理の流れを示すフローチャートである。

[図9]図9は、第1の実施形態の効果を示す図である。

[図10]図10は、拡張プログラムを実行するコンピュータの一例を示す図である。

### 発明を実施するための形態

[0009] 以下に、本願に係る拡張装置、拡張方法及び拡張プログラムの実施形態を図面に基づいて詳細に説明する。なお、本発明は、以下に説明する実施形態により限定されるものではない。

[0010] [第1の実施形態の構成]

まず、図1を用いて、第1の実施形態に係る拡張装置の構成について説明する。図1は、第1の実施形態に係る拡張装置の構成の一例を示す図である。図1に示すように、学習システム1は、拡張装置10及び学習装置20を有する。

[0011] 拡張装置10は、外部データセット40を用いて、目的データセット30のデータ拡張を行い、拡張済みデータセット50を出力する。また、学習装置20は、拡張済みデータセット50を用いて目的モデル21の学習を行う。目的モデル21は、機械学習を行う既知のモデルであってよい。例えば、目的モデル21は、非特許文献7に記載のMCCNN with Triplet lossである。

[0012] また、図1の各データセットは、目的モデル21で用いられるラベル付きのデータである。つまり、各データセットは、データとラベルの組み合わせである。例えば、目的モデル21が画像認識のためのモデルである場合、各データセットは、画像データとラベルの組み合わせである。また、目的モデル21は、音声認識モデルであってもよいし、自然言語認識モデルであってもよい。その場合、各データセットは、ラベル付きの音声データやラベル付きのテキストデータである。

[0013] ここでは、主に、各データセットが画像データとラベルの組み合わせである場合の例を説明する。また、以降の説明では、画像をコンピュータで処理可能な形式で表したデータを、画像データ又は単に画像と呼ぶ。

[0014] 図1に示すように、拡張装置10は、入出力部11、記憶部12及び制御

部 1 3 を有する。入出力部 1 1 は、入力部 1 1 1 及び出力部 1 1 2 を有する。入力部 1 1 1 は、ユーザからのデータの入力を受け付ける。入力部 1 1 1 は、例えば、マウスやキーボード等の入力装置である。出力部 1 1 2 は、画面の表示等により、データを出力する。出力部 1 1 2 は、例えば、ディスプレイ等の表示装置である。また、入出力部 1 1 は、通信によりデータの入出力を行う N I C (Network Interface Card) 等の通信インタフェースであってもよい。

[0015] 記憶部 1 2 は、H D D (Hard Disk Drive)、S S D (Solid State Drive)、光ディスク等の記憶装置である。なお、記憶部 1 2 は、R A M (Random Access Memory)、フラッシュメモリ、N V S R A M (Non Volatile Static Random Access Memory) 等のデータを書き換え可能な半導体メモリであってもよい。記憶部 1 2 は、拡張装置 1 0 で実行される O S (Operating System) や各種プログラムを記憶する。さらに、記憶部 1 2 は、プログラムの実行で用いられる各種情報を記憶する。また、記憶部 1 2 は、生成モデル 1 2 1 を記憶する。

[0016] 具体的には、記憶部 1 2 は、生成モデル 1 2 1 による各処理で用いられるパラメータを記憶する。本実施形態では、生成モデル 1 2 1 は、非特許文献 6 に記載の C G A N (Conditional Generative Adversarial Networks) であるものとする。ここで、図 2 を用いて、生成モデル 1 2 1 について説明する。図 2 は、第 1 の実施形態に係る生成モデルの一例を示す図である。

[0017] 図 2 に示すように、生成モデル 1 2 1 は、生成器 1 2 1 a 及び識別器 1 2 1 b を有する。例えば、生成器 1 2 1 a 及び識別器 1 2 1 b は、いずれもニューラルネットワークである。ここで、生成モデル 1 2 1 には、正解データセットが入力される。正解データセットは、正解データと、正解データに付与された正解ラベルの組み合わせである。例えば、正解データが特定の人物の画像である場合、正解ラベルは当該人物を識別する I D である。

[0018] 生成器 1 2 1 a は、所定のノイズとともに入力された正解ラベルから、生成データを生成する。また、識別器 1 2 1 b は、2 値判定誤差として、生成

データと正解データとの間の乖離の度合いを計算する。そして、生成モデル 121 の学習においては、生成器 121 a のパラメータは誤差が小さくなる方向に更新される。一方、識別器 121 b のパラメータは誤差が大きくなる方向に更新される。なお、学習における各パラメータの更新は、誤差逆伝播法 (Backpropagation) によって行われる。

[0019] つまり、生成器 121 a は、学習により、識別器 121 b によって正解データと同じものと識別されるような生成データを生成できるようになっていく。一方、識別器 121 b は、学習により、生成データを生成データと認識し、正解データを正解データと認識できるようになっていく。

[0020] 制御部 13 は、拡張装置 10 全体を制御する。制御部 13 は、例えば、CPU (Central Processing Unit)、MPU (Micro Processing Unit) 等の電子回路や、ASIC (Application Specific Integrated Circuit)、FPGA (Field Programmable Gate Array) 等の集積回路である。また、制御部 13 は、各種の処理手順を規定したプログラムや制御データを格納するための内部メモリを有し、内部メモリを用いて各処理を実行する。また、制御部 13 は、各種のプログラムが動作することにより各種の処理部として機能する。例えば、制御部 13 は、学習部 131、生成部 132 及び付与部 133 を有する。

[0021] 学習部 131 は、ラベルからデータを生成する生成モデル 121 に、ラベルが付与された第 1 のデータ及び第 2 のデータを学習させる。目的データセット 30 は、第 1 のデータ及び第 1 のデータに付与されたラベルの組み合わせの一例である。また、外部データセット 40 は、第 2 のデータ及び第 2 のデータに付与されたラベルの組み合わせの一例である。

[0022] ここで、目的データセット 30 は、目的データと目的データに付与された目的ラベルとの組み合わせであるものとする。また、外部データセット 40 は、外部データと外部データに付与された外部ラベルとの組み合わせであるものとする。

[0023] 目的ラベルは、目的モデル 21 の学習の対象のラベルである。例えば、目



的モデル21が画像中の人物を認識するためのモデルである場合、目的ラベルは、目的データの画像に映っている人物を識別するIDである。また、例えば、目的モデル21が音声からテキストを認識するモデルである場合、目的ラベルは、目的データの音声を書き起こしたテキストである。

[0024] 外部データセット40は、目的データセット30を拡張するためのデータセットである。外部データセット40は、目的データセット30と異なるドメインのデータセットであってもよい。ここで、ドメインとは、データセットに固有の特徴であって、データ、ラベル及び生成分布によって表される。例えば、データが $X_0$ 、ラベルが $Y_0$ であるデータセットのドメインは、 $(X_0, Y_0, P(X_0, Y_0))$ のように表される。

[0025] ここで、例として、目的モデル21が画像認識モデルであって、学習装置20は、画像からIDが「0002」である人物の画像を認識できるように目的モデル21の学習を行うものとする。この場合、目的データセット30は、ラベル「ID: 0002」と、当該人物が映っていることが既知の画像との組み合わせである。また、外部データセット40は、「0002」以外のIDを示すラベルと、当該IDに対応する人物が映っていることが既知の画像との組み合わせである。

[0026] また、外部データセット40は、必ずしも正確なラベルを有していなくてもよい。つまり、外部データセット40のラベルは、目的データセット30のラベルとの区別が付くものであればよく、例えば、未設定を意味するものであってもよい。

[0027] 拡張装置10は、目的データセット30のデータが有しない属性を外部データセット40から取り入れた拡張済みデータセット50を出力する。これにより、目的データセット30からだけでは得ることができなかったバリエーションのデータを得ることができる。例えば、拡張装置10によれば、目的データセット30に、ある人物の背面が映った画像のみが含まれている場合であっても、当該人物の正面が映った画像を得ることが可能になる。

[0028] 図3を用いて、学習部131による学習処理について説明する。図3は、

第1の実施形態に係る生成モデルの学習処理を説明するための図である。図3に示すように、データセット $S_{target}$ は、目的データセット30である。また、 $X_{target}$ 及び $Y_{target}$ は、それぞれデータセット $S_{target}$ のデータ及びラベルである。また、データセット $S_{outer}$ は、外部データセット40である。また、 $X_{outer}$ 及び $Y_{outer}$ は、それぞれデータセット $S_{outer}$ のデータ及びラベルである。

[0029] このとき、目的データセット30のドメインは、 $(X_{target}, Y_{target}, P(X_{target}, Y_{target}))$ のように表される。また、外部データセット40のドメインは、 $(X_{outer}, Y_{outer}, P(X_{outer}, Y_{outer}))$ のように表される。

[0030] 学習部131は、まず、各データに前処理を施す。例えば、学習部131は、前処理として、画像のサイズを一律の大きさ（例えば、 $128 \times 128$  pixel）に変更する。そして、学習部131は、データセット $S_{target}$ 及び $S_{outer}$ を結合し、データセット $S_{t+o}$ を生成する。例えば、 $S_{t+o}$ は、各データセットのデータ及びラベルを、それぞれ同じ配列に格納したものである。

[0031] そして、学習部131は、生成したデータセット $S_{t+o}$ を正解データセットとして生成モデル121に学習させる。具体的な学習方法は前述の通りである。つまり、学習部131は、生成モデル121の生成器121aが、第1のデータ及び第2のデータに近いデータを生成できるように、かつ、生成モデル121の識別器121bが、生成器121aが生成したデータと第1のデータ及び第2のデータとの違いを識別できるように学習を行う。

[0032] また、図3の $X'$ は、データセット $S_{t+o}$ のラベルから生成器121aが生成する生成データである。学習部131は、画像 $X'$ を基に、誤差逆伝播法により生成モデル121のパラメータを更新する。

[0033] 生成部132は、第1のデータ及び第2のデータを学習した生成モデル121を用いて、第1のデータに付与されたラベルから拡張用のデータを生成する。 $Y_{target}$ は、第1のデータに付与されたラベルの一例である。

- [0034] 図4を用いて、生成部132による生成処理について説明する。図4は、第1の実施形態に係る拡張画像の生成処理を説明するための図である。図4に示すように、生成部132は、ラベル $Y_{target}$ をノイズ $Z$ とともに生成モデル121に入力し、生成データ $X_{gen}$ を生成する。ここで、生成データ $X_{gen}$ は、生成器121aによって生成される。また、生成部132は、あらかじめ設定された分布に従ってノイズ $Z$ をランダムに発生させ、複数の生成データ $X_{gen}$ を生成することができる。ここでは、ノイズ $Z$ の分布は $N(0, 1)$ の正規分布であるとする。
- [0035] 付与部133は、第1のデータ及び拡張用のデータを統合した拡張済みデータに、第1のデータに付与されたラベルを付与する。付与部133は、生成部132によって生成された生成データ $X_{gen}$ にラベルを付与することで、学習装置20で利用可能なデータセット $S'_{target}$ を生成する。また、 $S'_{target}$ は、拡張済みデータセット50の一例である。
- [0036] 図5を用いて、付与部133による付与処理について説明する。図5に示すように、付与部133は、 $X_{target}$ と $X_{gen}$ を統合したデータに、ラベルとして $Y_{target}$ を付与する。このとき、目的データセット30のドメインは、 $(X_{target} + X_{gen}, Y_{target}, P(X_{target} + X_{gen}, Y_{target}))$ のように表される。
- [0037] その後、図6に示すように、学習装置20は、データセット $S'_{target}$ を用いて目的モデル21の学習を行う。図6は、第1の実施形態に係る目的モデルの学習処理を説明するための図である。
- [0038] 図7を用いて、拡張済みデータセット50の具体的な例について説明する。図7は、第1の実施形態に係る拡張装置によって生成される拡張済みデータセットの一例を示す図である。
- [0039] 図7に示すように、目的データセット30aは、画像301a及び「ID: 0002」というラベルを含む。また、外部データセット40aは、画像401a及び「ID: 0050」というラベルを含む。ここで、ラベルに含まれるIDは、画像中の人物を識別するものである。また、目的データセッ

ト 3 0 a 及び外部データセット 4 0 a には、図示のもの以外の画像が含まれていてもよい。

[0040] 画像 3 0 1 a には、黒髪で、赤 T シャツ及び短 G パンを着用し、背面を向いた黄色人種の人物が映っているものとする。このとき、画像 3 0 1 a には、「背面」、「黒髪」、「赤 T シャツ」、「黄色人種」、「短 G パン」といった属性が含まれる。

[0041] 画像 4 0 1 a には、バッグを肩にかけ、白 T シャツ、黒短パン及び靴を着用し、正面を向いた人物が映っているものとする。このとき、画像 4 0 1 a には、「正面」、「バッグ」、「白 T シャツ」、「黒短パン」、「靴」といった属性が含まれる。

[0042] なお、ここでの属性とは、目的モデル 2 1 が画像認識の際に利用する情報である。ただし、これらの属性は説明のために例として定義したものであり、画像認識処理においては、必ずしも明示的に個別の情報として扱われているわけではない。そのため、目的データセット 3 0 a 及び外部データセット 4 0 a は、どのような属性が含まれるかが未知のものであってもよい。

[0043] 拡張装置 1 0 は、目的データセット 3 0 a 及び外部データセット 4 0 a を入力とし、拡張済みデータセット 5 0 a を出力する。拡張用画像 5 0 1 a は、拡張装置 1 0 が生成した画像の 1 つである。拡張済みデータセット 5 0 a は、目的データセット 3 0 a と、ラベル「ID : 0 0 0 2」が付与された拡張用画像 5 0 1 a を統合したデータセットである。

[0044] 拡張用画像 5 0 1 a には、黒髪で、赤 T シャツ及び短 G パンを着用し、正面を向いた黄色人種の人物が映っているものとする。このとき、拡張用画像 5 0 1 a には、「正面」、「黒髪」、「赤 T シャツ」、「黄色人種」、「短 G パン」といった属性が含まれる。

[0045] ここで、「正面」という属性は、目的データセット 3 0 a からのみでは得ることができなかった属性である。このように、拡張装置 1 0 は、外部データセット 4 0 a から得られた属性を、目的データセット 3 0 a の属性と組み合わせた画像を生成することができる。

[0046] [第1の実施形態の処理]

図8を用いて、拡張装置10の処理の流れについて説明する。図8は、第1の実施形態に係る拡張装置の処理の流れを示すフローチャートである。ここでは、目的モデル21は画像認識を行うモデルであり、各データセットに含まれるデータは画像であるものとする。

[0047] 図8に示すように、まず、拡張装置10は、目的データセット30及び外部データセット40の入力を受け付ける（ステップS101）。次に、拡張装置10は、生成モデル121を用いて、目的データセット30及び外部データセット40から画像を生成する（ステップS102）。そして、拡張装置10は、生成した画像を基に生成モデル121のパラメータを更新する（ステップS103）。つまり、拡張装置10は、ステップS102及びステップS103により、生成モデル121の学習を行う。また、拡張装置10は、所定の条件が満たされるまで、ステップS102及びステップS103を繰り返し実行してもよい。

[0048] ここで、拡張装置10は、生成モデル121に、目的データセット30のラベルを指定し（ステップS104）、指定したラベルを基に拡張用画像を生成する（ステップS105）。次に、拡張装置10は、目的データセット30の画像と拡張用画像を統合し、統合したデータに目的データセット30のラベルを付与する（ステップS106）。

[0049] 拡張装置10は、ステップS106でラベルを付与したデータを拡張済みデータセット50として出力する（ステップS107）。学習装置20は、拡張済みデータセット50を用いて目的モデル21の学習を行う。

[0050] [第1の実施形態の効果]

これまで説明してきたように、拡張装置10は、ラベルからデータを生成する生成モデルに、ラベルが付与された第1のデータ及び第2のデータを学習させる。また、拡張装置10は、第1のデータ及び第2のデータを学習した生成モデルを用いて、第1のデータに付与されたラベルから拡張用のデータを生成する。また、拡張装置10は、第1のデータ及び拡張用のデータを

統合した拡張済みデータに、第1のデータに付与されたラベルを付与する。このように、本実施形態の拡張装置10は、データ拡張により、目的データセットに含まれない属性を持った学習データを生成することができる。このため、本実施形態によれば、データ拡張により得られる学習データのバリエーションを増加させ、モデルの精度を向上させることができる。

[0051] 拡張装置10は、生成モデルの生成器が、第1のデータ及び第2のデータに近いデータを生成できるように、かつ、生成モデルの識別器が、生成器が生成したデータと第1のデータ及び第2のデータとの違いを識別できるように学習を行う。これにより、生成モデルを用いて生成するデータを、目的データと似せることが可能になる。

[0052] [実験結果]

ここで、従来の技術と実施形態を比較するために行った実験について説明する。実験において、目的モデル21は、画像認識により画像から特定の人物を探すタスクを行うMCCNN with Triplet lossである。また、各手法の比較は、拡張前のデータ、すなわち目的データセット30を目的モデル21に入力した場合の認識精度により行った。生成モデル121は、CGANである。

[0053] また、目的データセット30は、人物再照合用のデータセットである「Market-1501」である。また、外部データセット40は、同じく人物再照合用のデータセットである「CHUK03」である。また、拡張するデータの量は、元データ量の3倍である。

[0054] 実験の結果を図9に示す。図9は、第1の実施形態の効果を示す図である。横軸は、目的データセット30のサイズを割合で示したものである。また、縦軸は、精度を示している。図9に示すように、また、各折れ線は、データ拡張をしなかった場合、実施形態の手法でデータ拡張を行った場合、及び従来のルールベースのデータ拡張を行った場合の結果を示している。

[0055] 図9に示すように、データサイズにかかわらず、実施形態の手法でデータ拡張を行った場合に最も精度が高くなった。特に、データサイズが20%程

度の場合、実施形態の手法の精度は、従来の手法の精度と比べて20%程度向上した。また、データサイズが33%程度の場合、実施形態の手法の精度が、データサイズが100%の場合の従来の手法の精度と同等であった。また、データサイズが100%であっても、実施形態の手法の精度は、従来の手法の精度と比べて10%程度向上した。これより、本実施形態によるデータ拡張は、従来の手法と比べて目的モデル21の認識精度をより向上させているといえる。

[0056] [その他の実施形態]

上記の実施形態では、目的モデル21の学習機能は、拡張装置10とは異なる学習装置20に備えられていた。一方で、拡張装置10に、拡張済みデータセット50を目的モデル21に学習させる目的モデル学習部が備えられていてもよい。これにより、拡張装置10は、装置間のデータ転送によるリソースの消費を抑え、データ拡張及び目的モデルの学習を、一連の処理として効率良く実行することができる。

[0057] [システム構成等]

また、図示した各装置の各構成要素は機能概念的なものであり、必ずしも物理的に図示のように構成されていることを要しない。すなわち、各装置の分散及び統合の具体的形態は図示のものに限られず、その全部又は一部を、各種の負荷や使用状況等に応じて、任意の単位で機能的又は物理的に分散又は統合して構成することができる。さらに、各装置にて行われる各処理機能は、その全部又は任意の一部が、CPU及び当該CPUにて解析実行されるプログラムにて実現され、あるいは、ワイヤードロジックによるハードウェアとして実現され得る。

[0058] また、本実施形態において説明した各処理のうち、自動的に行われるものとして説明した処理の全部又は一部を手動的に行うこともでき、あるいは、手動的に行われるものとして説明した処理の全部又は一部を公知の方法で自動的に行うこともできる。この他、上記文書中や図面中で示した処理手順、制御手順、具体的名称、各種のデータやパラメータを含む情報については、

特記する場合を除いて任意に変更することができる。

[0059] [プログラム]

一実施形態として、拡張装置10は、パッケージソフトウェアやオンラインソフトウェアとして上記のデータ拡張を実行する拡張プログラムを所望のコンピュータにインストールさせることによって実装できる。例えば、上記の拡張プログラムを情報処理装置に実行させることにより、情報処理装置を拡張装置10として機能させることができる。ここで言う情報処理装置には、デスクトップ型又はノート型のパーソナルコンピュータが含まれる。また、その他にも、情報処理装置にはスマートフォン、携帯電話機やPHS (Personal Handyphone System) 等の移動体通信端末、さらには、PDA (Personal Digital Assistant) 等のスレート端末等がその範疇に含まれる。

[0060] また、拡張装置10は、ユーザが使用する端末装置をクライアントとし、当該クライアントに上記のデータ拡張に関するサービスを提供する拡張サーバ装置として実装することもできる。例えば、拡張サーバ装置は、目的データを入力とし、拡張済みデータを出力とする拡張サービスを提供するサーバ装置として実装される。この場合、拡張サーバ装置は、Webサーバとして実装することとしてもよいし、アウトソーシングによって上記のデータ拡張に関するサービスを提供するクラウドとして実装することとしてもかまわない。

[0061] 図10は、拡張プログラムを実行するコンピュータの一例を示す図である。コンピュータ1000は、例えば、メモリ1010、CPU1020を有する。また、コンピュータ1000は、ハードディスクドライバインタフェース1030、ディスクドライバインタフェース1040、シリアルポートインタフェース1050、ビデオアダプタ1060、ネットワークインタフェース1070を有する。これらの各部は、バス1080によって接続される。

[0062] メモリ1010は、ROM (Read Only Memory) 1011及びRAM 1012を含む。ROM 1011は、例えば、BIOS (Basic Input Output



t System) 等のブートプログラムを記憶する。ハードディスクドライブインタフェース1030は、ハードディスクドライブ1090に接続される。ディスクドライブインタフェース1040は、ディスクドライブ1100に接続される。例えば磁気ディスクや光ディスク等の着脱可能な記憶媒体が、ディスクドライブ1100に挿入される。シリアルポートインタフェース1050は、例えばマウス1110、キーボード1120に接続される。ビデオアダプタ1060は、例えばディスプレイ1130に接続される。

[0063] ハードディスクドライブ1090は、例えば、OS1091、アプリケーションプログラム1092、プログラムモジュール1093、プログラムデータ1094を記憶する。すなわち、拡張装置10の各処理を規定するプログラムは、コンピュータにより実行可能なコードが記述されたプログラムモジュール1093として実装される。プログラムモジュール1093は、例えばハードディスクドライブ1090に記憶される。例えば、拡張装置10における機能構成と同様の処理を実行するためのプログラムモジュール1093が、ハードディスクドライブ1090に記憶される。なお、ハードディスクドライブ1090は、SSDにより代替されてもよい。

[0064] また、上述した実施形態の処理で用いられる設定データは、プログラムデータ1094として、例えばメモリ1010やハードディスクドライブ1090に記憶される。そして、CPU1020は、メモリ1010やハードディスクドライブ1090に記憶されたプログラムモジュール1093やプログラムデータ1094を必要に応じてRAM1012に読み出して、上述した実施形態の処理を実行する。

[0065] なお、プログラムモジュール1093やプログラムデータ1094は、ハードディスクドライブ1090に記憶される場合に限らず、例えば着脱可能な記憶媒体に記憶され、ディスクドライブ1100等を介してCPU1020によって読み出されてもよい。あるいは、プログラムモジュール1093及びプログラムデータ1094は、ネットワーク（LAN（Local Area Network）、WAN（Wide Area Network）等）を介して接続された他のコン

コンピュータに記憶されてもよい。そして、プログラムモジュール1093及びプログラムデータ1094は、他のコンピュータから、ネットワークインタフェース1070を介してCPU1020によって読み出されてもよい。

### 符号の説明

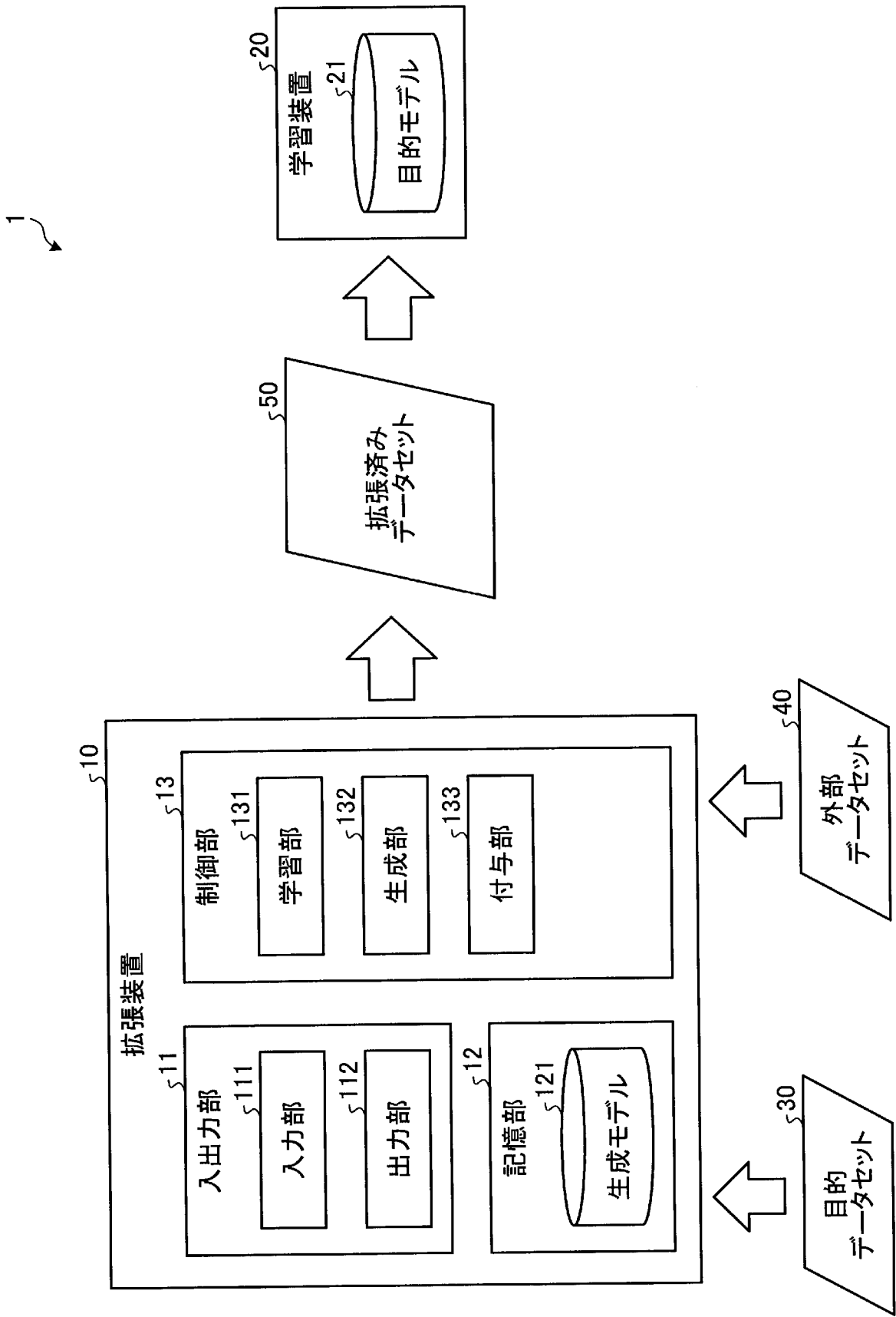
- [0066]    10   拡張装置
- 11   入出力部
- 12   記憶部
- 13   制御部
- 20   学習装置
- 21   目的モデル
- 30、30a   目的データセット
- 40、40a   外部データセット
- 50、50a   拡張済みデータセット
- 111   入力部
- 112   出力部
- 121   生成モデル
- 121a   生成器
- 121b   識別器
- 131   学習部
- 132   生成部
- 133   付与部
- 301a、401a   画像
- 501a   拡張用画像

## 請求の範囲

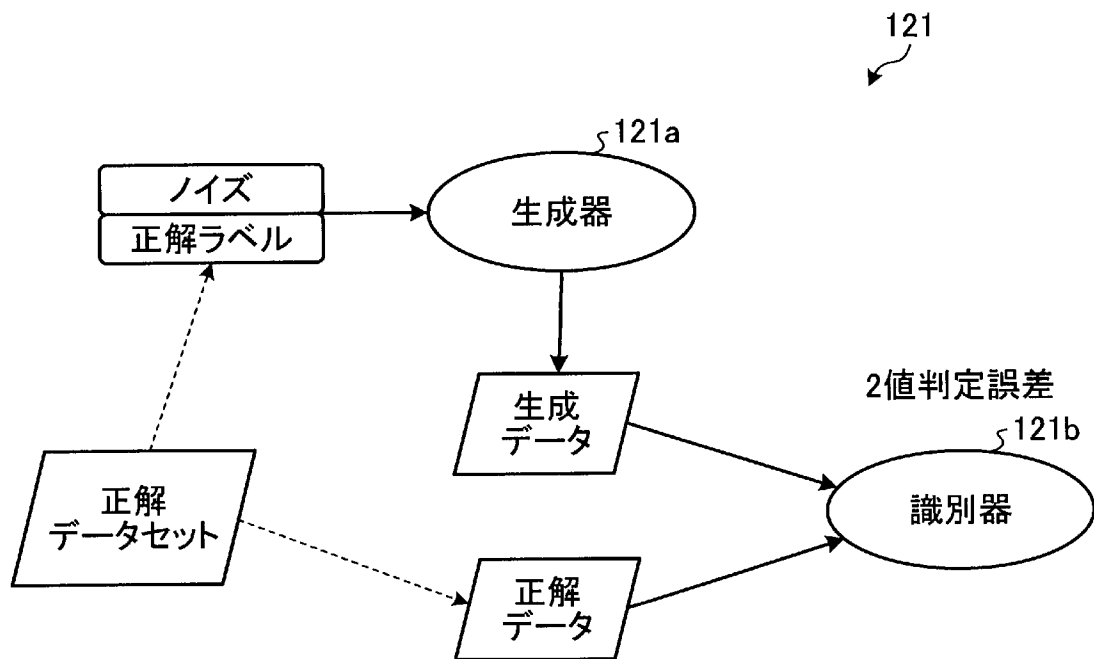
- [請求項1]           ラベルからデータを生成する生成モデルに、ラベルが付与された第1のデータ及び第2のデータを学習させる学習部と、
- 前記第1のデータ及び前記第2のデータを学習した前記生成モデルを用いて、前記第1のデータに付与されたラベルから拡張用のデータを生成する生成部と、
- 前記第1のデータ及び前記拡張用のデータを統合した拡張済みデータに、前記第1のデータに付与されたラベルを付与する付与部と、
- を有することを特徴とする拡張装置。
- [請求項2]           前記学習部は、前記生成モデルの生成器が、前記第1のデータ及び前記第2のデータに近いデータを生成できるように、かつ、前記生成モデルの識別器が、前記生成器が生成したデータと前記第1のデータ及び前記第2のデータとの違いを識別できるように学習を行い、
- 前記生成部は、前記生成器を用いて拡張用のデータを生成することを特徴とする請求項1に記載の拡張装置。
- [請求項3]           前記付与部によってラベルを付与された拡張済みデータを、目的モデルに学習させる目的モデル学習部をさらに有することを特徴とする請求項1又は2に記載の拡張装置。
- [請求項4]           コンピュータが実行する拡張方法であって、
- ラベルからデータを生成する生成モデルに、ラベルが付与された第1のデータ及び第2のデータを学習させる学習工程と、
- 前記第1のデータ及び前記第2のデータを学習した前記生成モデルを用いて、前記第1のデータに付与されたラベルから拡張用のデータを生成する生成工程と、
- 前記第1のデータ及び前記拡張用のデータを統合した拡張済みデータに、前記第1のデータに付与されたラベルを付与する付与工程と、
- を含むことを特徴とする拡張方法。
- [請求項5]           コンピュータを、請求項1から3のいずれか1項に記載の拡張装置

として機能させるための拡張プログラム。

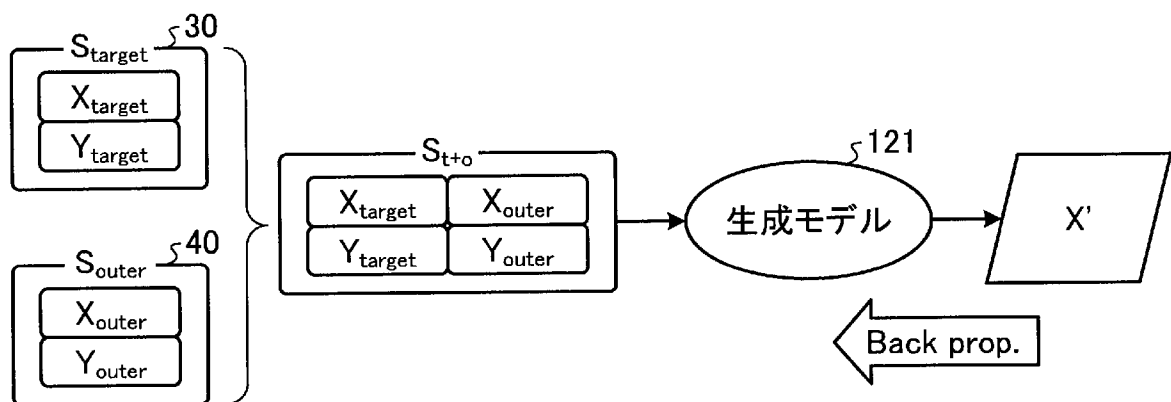
[図1]



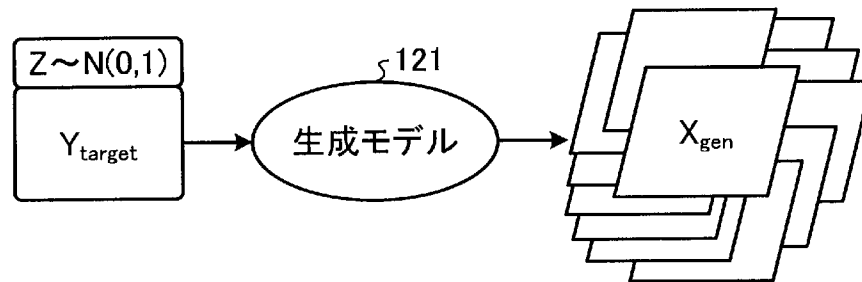
[図2]



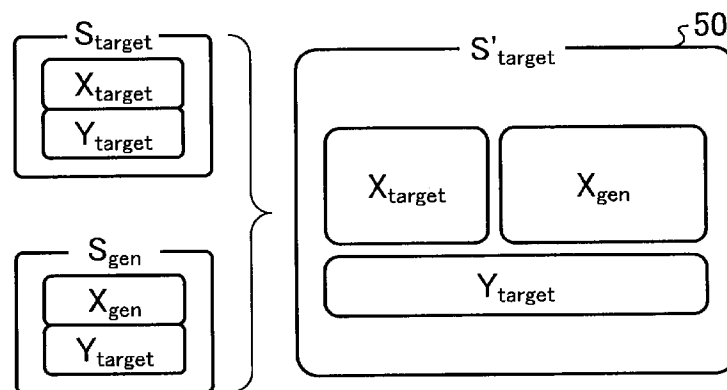
[図3]



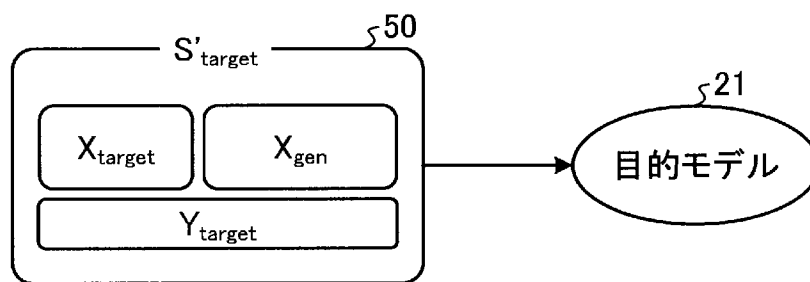
[図4]



[図5]

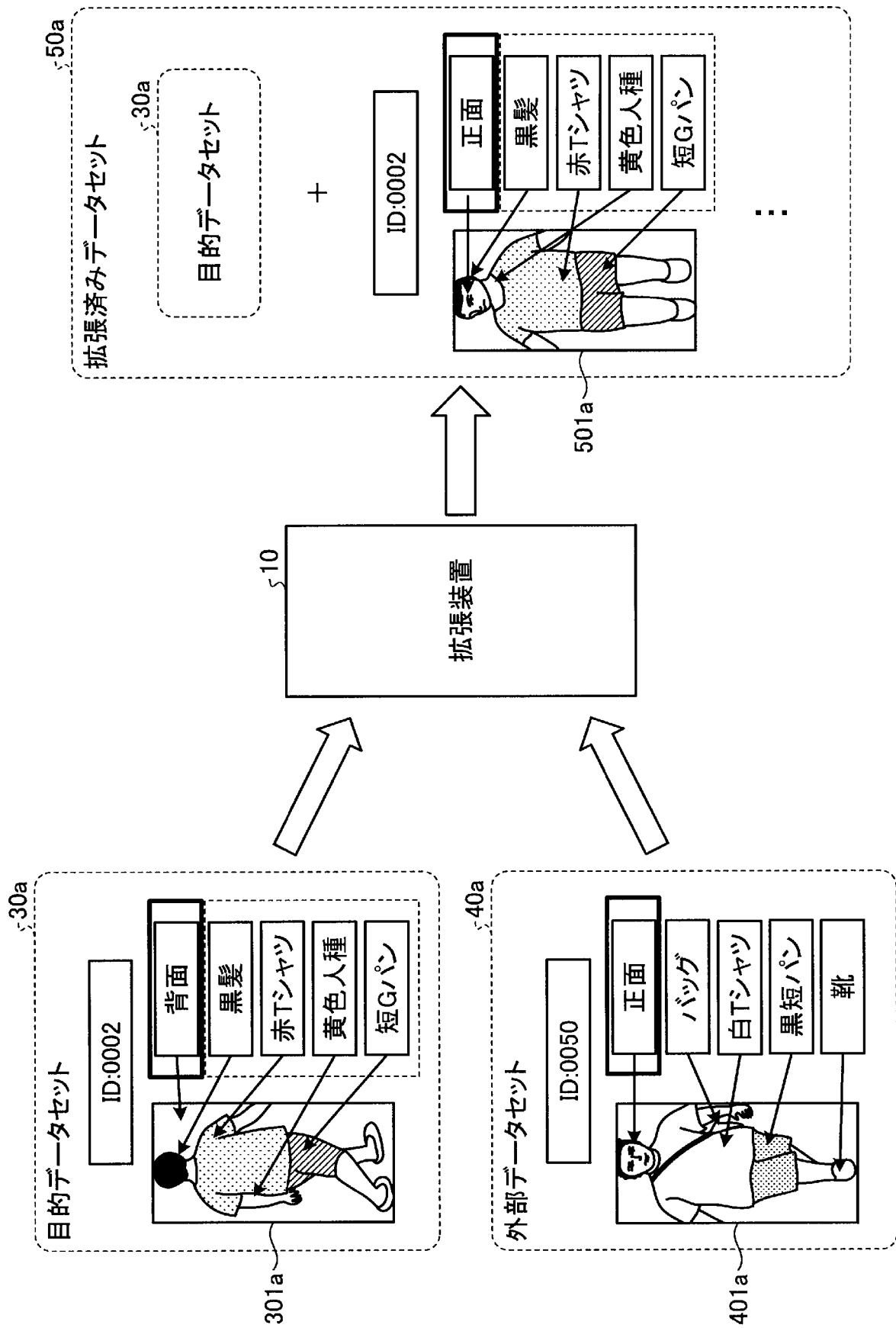


[図6]

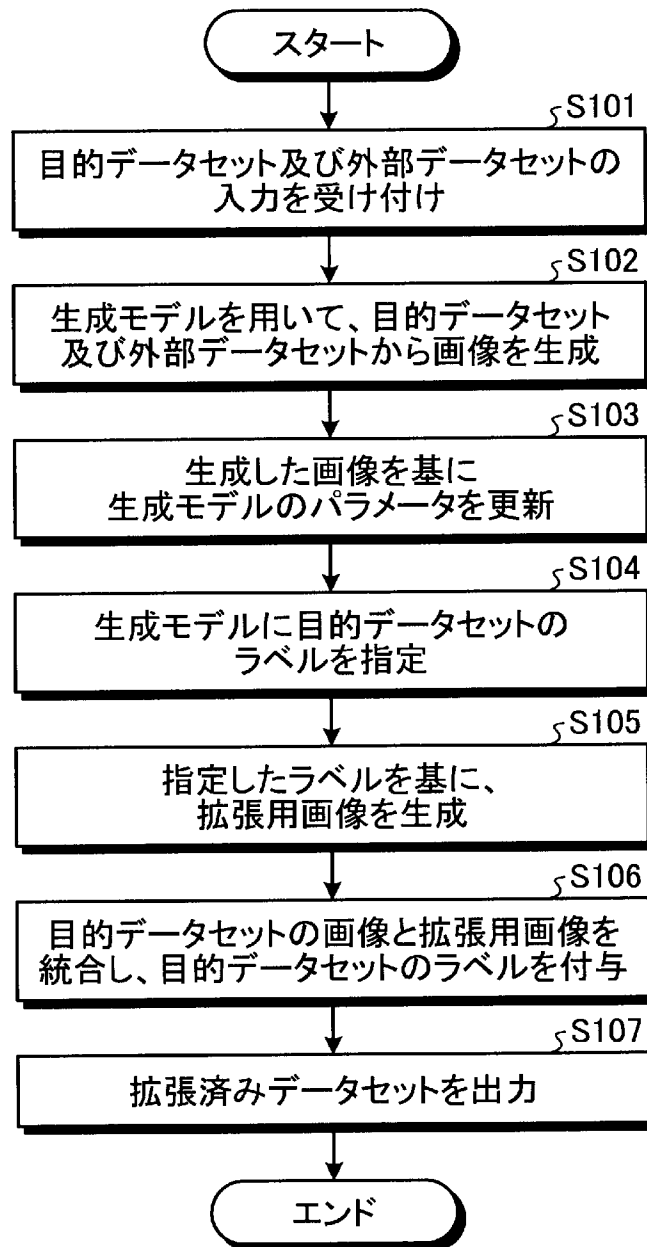




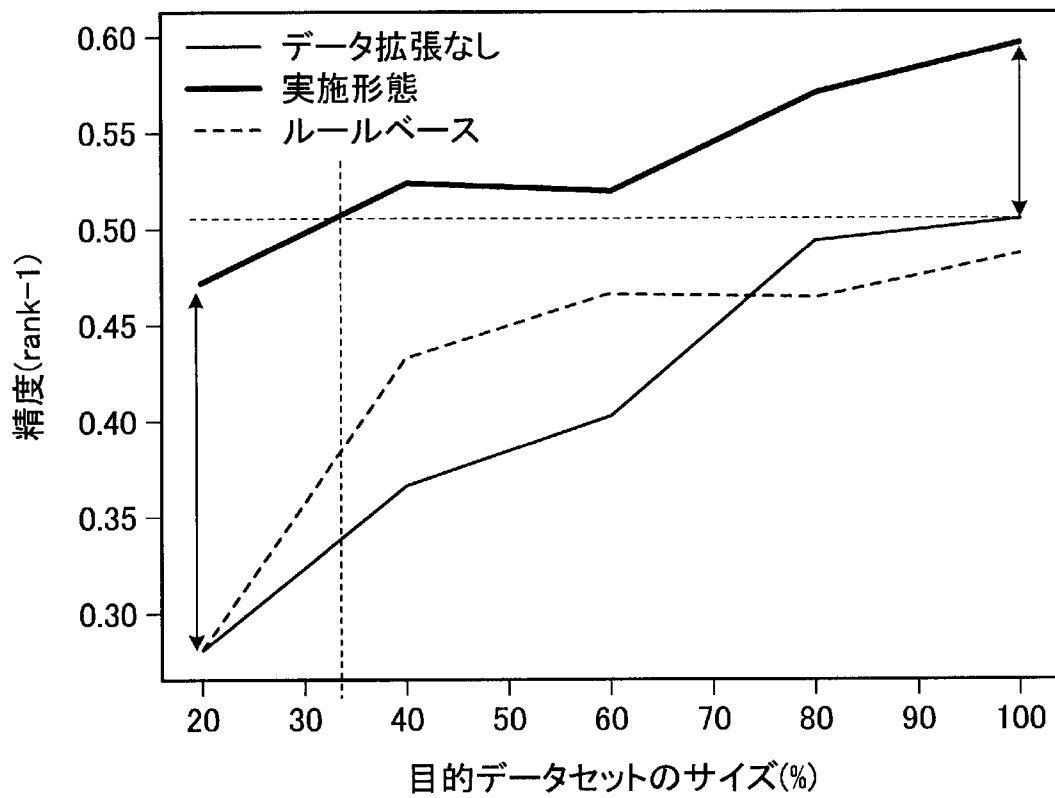
[図7]



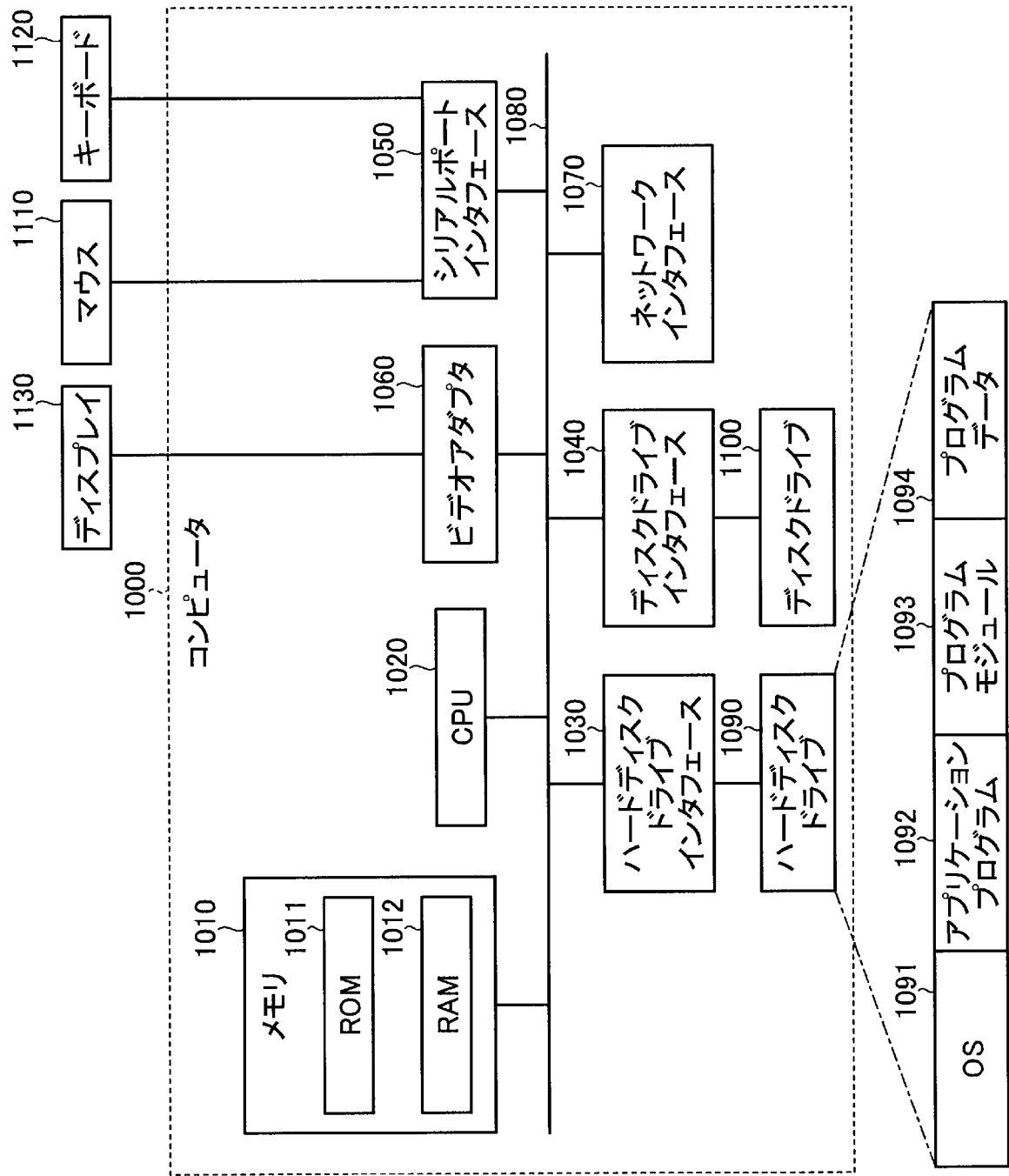
[図8]



[図9]



[図10]



## INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2019/032863

## A. CLASSIFICATION OF SUBJECT MATTER

Int.Cl. G06N20/00(2019.01)i, G06T7/00(2017.01)i

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)

Int.Cl. 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 1971-2019

Registered utility model specifications of Japan 1996-2019

Published registered utility model applications of Japan 1994-2019

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         | 堀田大地 ほか, 大量の Twitter 画像を用いた Conditional Cycle GAN による食事写真カテゴリ変換, 人工知能学会全国大会論文集 [online], 30 July 2018, [retrieved on 17 September 2019], pp. 1-4, retrieved from the Internet:<br><URL:https://www.jstage.jst.go.jp/article/pjsai/JS-AI2018/0/JSAI2018_4Pin110/_pdf/-char/ja>, (HORITA, Daichi et al., Conditional cycle GAN based transformation of food image category, Proceedings of the Annual Conference of JSAI) | 1-5                   |



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

“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

“&amp;” document member of the same patent family

Date of the actual completion of the international search

18.09.2019

Date of mailing of the international search report

08.10.2019

Name and mailing address of the ISA/

Japan Patent Office

3-4-3, Kasumigaseki, Chiyoda-ku,  
Tokyo 100-8915, Japan

Authorized officer

Telephone No.

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2019/032863

## C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages                                                                                                                                                               | Relevant to claim No. |
|-----------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------|
| A         | 渡部宏樹 ほか, DCGAN を用いたデータオーギュメンテーションによる猫の品種識別について, 映像情報メディア学会 2016 年年次大会講演予稿集, 17 August 2016, ISSN 1880-6953, (WATABE, Hiroki et al., On cat breed identification by data augmentation using DCGAN, Proceedings of the 2016 ITE annual convention) | 1-5                   |

## A. 発明の属する分野の分類（国際特許分類（IPC））

Int.Cl. G06N20/00(2019.01)i, G06T7/00(2017.01)i

## 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））

Int.Cl. 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 | 1971-2019年 |
| 日本国実用新案登録公報 | 1996-2019年 |
| 日本国登録実用新案公報 | 1994-2019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<br>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<br>請求項の番号 |
|-----------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------|
| A               | 堀田大地 ほか,<br>大量の Twitter 画像を用いた Conditional Cycle GAN による食事写真カテゴリー変換,<br>人工知能学会全国大会論文集[online], 2018.07.30<br>[retrieved on 2019.09.17], pp.1-4,<br>Retrieved from the Internet: <URL: <a href="https://www.jstage.jst.go.jp/article/pjsai/JSAI2018/0/JSAI2018_4Pin110/_pdf/-char/ja">https://www.jstage.jst.go.jp/article/pjsai/JSAI2018/0/JSAI2018_4Pin110/_pdf/-char/ja</a> > | 1-5            |

☒ 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  
「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矛盾するものではなく、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

18.09.2019

国際調査報告の発送日

08.10.2019

国際調査機関の名称及びあて先

日本国特許庁（ISA/J P）

郵便番号100-8915

東京都千代田区霞が関三丁目4番3号

特許庁審査官（権限のある職員）

北元 健太

5 B

1157

電話番号 03-3581-1101 内線 3545

| C（続き）． 関連すると認められる文献 |                                                                                                                      |                |
|---------------------|----------------------------------------------------------------------------------------------------------------------|----------------|
| 引用文献の<br>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<br>請求項の番号 |
| A                   | 渡部宏樹 ほか，<br>DCGAN を用いたデータオーギュメンテーションによる猫の品種識別<br>について，<br>映像情報メディア学会 2016 年年次大会講演予稿集，2016. 08. 17，<br>ISSN 1880-6953 | 1-5            |