(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
2022年12月1日(01.12.2022)

WIPO | PCT

(10) 国際公開番号
WO 2022/249285 A1

(51) 国際特許分類:
G06N 3/08 (2006.01) *G06N 20/00* (2019.01)

(21) 国際出願番号： PCT/JP2021/019817

(22) 国際出願日： 2021年5月25日(25.05.2021)

(25) 国際出願の言語： 日本語

(26) 国際公開の言語： 日本語

(71) 出願人: 富士通株式会社 (**FUJITSU LIMITED**) [JP/JP]; 〒2118588 神奈川県川崎市中原区上小田中４丁目１番１号 Kanagawa (JP).

(72) 発明者: 金月 寛彰 (**KINGETSU, Hiroaki**); 〒2118588 神奈川県川崎市中原区上小田中４丁目１番１号 富士通株式会社内 Kanagawa (JP).
小林 健一 (**KOBAYASHI, Kenichi**); 〒2118588 神奈川県川崎市中原区上小田中４丁目１番１号 富士通株式会社内 Kanagawa (JP).

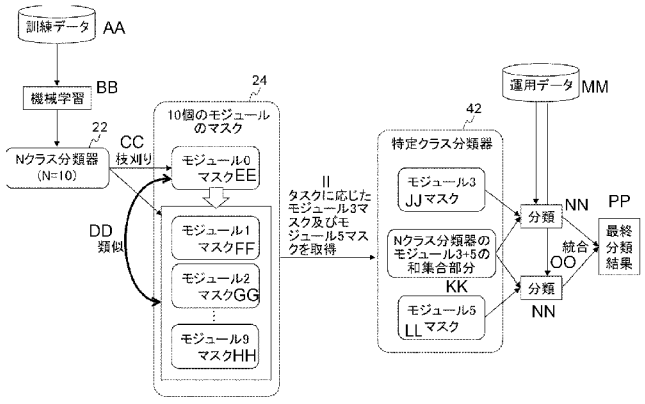
(74) 代理人: 中島 淳, 外 (**NAKAJIMA, Jun** et al.); 〒1600022 東京都新宿区新宿４丁目３番１７号 ＨＫ新宿ビル７階 太陽国際特許事務所 Tokyo (JP).

(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, 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, IT, JO, JP, 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, 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, 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,

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

(54) 発明の名称： 機械学習プログラム、装置、及び方法



22 N-class classifier
24 Masks of 10 modules
42 Specific class classifier
AA Training data
BB Machine learning
CC Pruning
DD Similar
EE Mask of module 0
FF Mask of module 1
GG Mask of module 2
HH Mask of module 9
II Acquire mask of module 3 and mask of module 5 corresponding to task
JJ Mask of module 3
KK Union part of modules 3+5 of N-class classifier
LL Mask of module 5
MM Operation data
NN Classification
OO Integration
PP Definitive classification result

(57) **Abstract:** This machine learning device generates an N-class classifier (22) by machine learning using training data, generates a base mask for generating a base module by machine learning using the N-class classifier (22) and training data associated with a correct label indicating one class, generates a mask for generating each of other modules such that each of other modules becomes similar to the base module, and stores the generated masks as a mask set (24). A classification device stores the masks of modules corresponding to a task and a union part thereof as a specific class classifier 42, and integrates classification results, which are obtained by inputting operation data into each module generated by applying each mask to the union part, to output a definitive classification result.

WO 2022/249285 A1

［続葉有］

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)）

(57) 要約：機械学習装置は、訓練データを用いた機械学習によりＮクラス分類器（２２）を生成し、１つのクラスを示す正解ラベルが対応付けられた訓練データとＮクラス分類器（２２）とを用いた機械学習により、ベースモジュールを生成するためのベースマスクを生成し、他のモジュールがベースモジュールに類似するように、他のモジュールの各々を生成するためのマスクを生成し、生成したマスクをマスク集合（２４）として記憶する。分類装置は、タスクに応じたモジュールのマスクと、その和集合部分とを特定クラス分類器４２として記憶し、和集合部分に各マスクを適用して生成した各モジュールに運用データを入力して得られる分類結果を統合して、最終的な分類結果を出力する。

明　細　書

発明の名称 ： 機械学習プログラム、装置、及び方法

技術分野

[0001] 開示の技術は、機械学習プログラム、機械学習装置、及び機械学習方法に関する。

背景技術

[0002] 従来、複数のクラスへ分類する機械学習モデルについて、枝刈りによって一部のクラスのみを分類する個別の機械学習モデルを複数生成し、それらを結合させることで特定の複数のクラスを分類する機械学習モデルを生成する技術が存在する。例えば、機械学習済みのニューラルネットワークから、スーパーマスクを用いてサブネットワークを切り出す技術が提案されている。この技術では、ニューラルネットワークのエッジの重みに対応したスコア行列を用意し、ニューラルネットワークの重みに、スコア行列の上位ｋ％の要素を１、その他の要素を０としたスーパーマスクを適用して、機械学習のフォワード処理が実行される。また、この技術では、バックワード処理時には、ニューラルネットワークのエッジの重みは固定した上で、スコア行列の各スコアについて、通常の勾配法により機械学習を実行する。

先行技術文献

非特許文献

[0003] 非特許文献1：Vivek Ramanujan, Mitchell Wortsman, Aniruddha Kembhavi, Ali Farhadi, and Mohammad Rastegari, "What's Hidden in a Randomly Weighted Neural Network?", CVPR, 31 Mar 2020.

発明の概要

発明が解決しようとする課題

[0004] 上述のように、枝刈りによって生成した個別の機械学習モデルのうち、運用時のタスクに応じた個別の機械学習モデルを選択して結合することで、元の機械学習モデルに比べて、機械学習モデルのパラメータ数を少なくできる

という効果がある。しかし、個別の機械学習モデルの構造によっては、結合により生成する機械学習モデルにおいて、枝刈りの効果が薄まる場合がある。

[0005] 一つの側面として、開示の技術は、元の機械学習モデルから枝刈りによって生成した一部のクラスのみを分類する個別の機械学習モデルを結合して機械学習モデルを生成する場合において、枝刈りの効果が薄まることを抑制する。

課題を解決するための手段

[0006] 一つの態様として、開示の技術は、複数のクラスへ分類する機械学習モデルのパラメータと、前記複数のクラスのうち第１のクラスを正解ラベルに含む訓練データとに基づいて、第１の枝刈り処理に関するパラメータを生成する。前記第１の枝刈り処理に関するパラメータは、前記機械学習モデルに第１の枝刈り処理を実行することによって前記第１のクラスを分類する第１の機械学習モデルを生成するためのパラメータである。また、開示の技術は、前記機械学習モデルのパラメータと、前記複数のクラスのうち第２のクラスを正解ラベルに含む訓練データと、前記第１の枝刈り処理に関するパラメータを含む損失関数とに基づいて、第２の枝刈り処理に関するパラメータを生成する。第２の枝刈り処理に関するパラメータは、前記機械学習モデルに第２の枝刈り処理を実行することによって前記第２のクラスを分類する第２の機械学習モデルを生成するためのパラメータである。

発明の効果

[0007] 一つの側面として、元の機械学習モデルから枝刈りによって生成した一部のクラスのみを分類する個別の機械学習モデルを結合して機械学習モデルを生成する場合において、枝刈りの効果が薄まることを抑制することができる、という効果を有する。

図面の簡単な説明

[0008] [図1]機械学習装置の機能ブロック図である。

[図2]Ｎクラス分類器の一例を示す図である。

[図3]Ｎクラス分類器のモジュール化を説明するための図である。

[図4]Ｎクラス分類器のモジュール化を説明するための図である。

[図5]ベースマスクの生成を説明するための図である。

[図6]ベースマスク以外のマスクの生成を説明するための図である。

[図7]分類装置の機能ブロック図である。

[図8]機械学習装置及び分類装置の処理の一例を説明するための図である。

[図9]機械学習装置として機能するコンピュータの概略構成を示すブロック図である。

[図10]分類装置として機能するコンピュータの概略構成を示すブロック図である。

[図11]機械学習処理の一例を示すフローチャートである。

[図12]分類処理の一例を示すフローチャートである。

[図13]本実施形態におけるモジュール間でのパラメータの共有割合の実験結果の一例を説明するための図である。

発明を実施するための形態

[0009] 以下、図面を参照して、開示の技術に係る実施形態の一例を説明する。本実施形態に係る機械学習システムは、機械学習装置と、分類装置とを含む。機械学習装置は、Ｎクラス分類器をモジュール化するためのマスクを生成する。分類器は、生成されたマスクを用いてタスクに応じた特定クラス分類器を生成し、運用データを分類する。以下、機械学習装置及び分類装置の各々について詳述する。

[0010] まず、機械学習装置について説明する。図１に示すように、機械学習装置１０は、機能的には、Ｎクラス分類器生成部１２と、第１マスク生成部１４と、第２マスク生成部１６とを含む。また、機械学習装置１０の所定の記憶領域には、Ｎクラス分類器２２と、マスク集合２４とが記憶される。

[0011] Ｎクラス分類器生成部１２は、入力されたデータを、Ｎ個のクラスのいずれかに分類するための分類器であるＮクラス分類器２２を生成する。例えば、Ｎクラス分類器２２は、図２に示すようなニューラルネットワーク（Neura

l Network：ＮＮ）としてよい。図２の例では、丸印が各ニューロンであり、各層のニューロン間はエッジで接続されている。出力層（図２中の破線部）は、各クラスに対応したＮ個のニューロンを持ち、出力層の各ニューロンからは、入力されたデータが各クラスに分類される確率が出力される。

[0012] Ｎクラス分類器生成部１２は、機械学習装置１０に入力された訓練データを取得し、取得した訓練データを用いた機械学習により、Ｎクラス分類器２２のパラメータであるエッジの重みを算出する。訓練データは、画像データ、音声データ、センサデータ等のデータに、そのデータが属するクラスを示す正解ラベルが対応付けられたデータである。

[0013] 本実施形態では、上記のようなＮクラス分類器２２から、一部のクラスのみを分類可能な部分ネットワークをモジュールとして切り出し、モジュール同士を適宜結合することにより、任意のマルチクラスを分類可能な分類器を生成する。モジュール化することの利点は、ネットワークが小さくなることで、リスク検証が容易になること、及び、モジュール同士の結合で生成される分類器のパラメータ数を少なくすることができる点である。分類器のパラメータ数が少ない方が、計算量が低減されるため、運用時に用いる分類器として実用性が高まる。

[0014] ここで、ＮＮのモジュール化について説明する。例えば、図３の上図に示すように、オリジナルのＮＮが、入力されたデータを、クラス「猫」、「犬」、及び「鳥」の３クラスのいずれかに分類するものであるとする。オリジナルのＮＮには、クラス「猫」を表現する部分ネットワークと、クラス「犬」を表現する部分ネットワークと、クラス「鳥」を表現する部分ネットワークとが含まれる。機械学習装置は、クラス「猫」を表現する部分ネットワークをモジュールＡ、クラス「犬」を表現する部分ネットワークをモジュールＢ、及びクラス「鳥」を表現する部分ネットワークをモジュールＣとして切り出す。そして、図３の下図に示すように、運用時のタスクが、入力されたデータが示すものが、猫か犬かを分類するタスクである場合、機械学習装置は、モジュールＡとモジュールＢとを結合した分類器を生成する。そして、

機械学習装置は、入力されたデータを、モジュールＡとモジュールＢとを結合した分類器を用いて、クラス「猫」又はクラス「犬」に分類する。同様に、運用時のタスクが、入力されたデータが示すものが、猫か鳥かを分類するタスクである場合、機械学習装置１０は、モジュールＡとモジュールＣとを結合した分類器を生成し、入力されたデータをクラス「猫」又はクラス「鳥」に分類する。

[0015] ＮＮをモジュール化する手法として、上述した非特許文献１に示すように、スーパーマスクを適用する手法が存在する。オリジナルのＮＮが手書き数字の画像データセットであるＭＮＩＳＴの１０個の数字に対応する１０クラス分類のＮＮである場合を例に、この手法によるモジュール化ついて説明する。図４に示すように、正解ラベル「０」の訓練データに基づく機械学習により生成した「マスク０」によりオリジナルのＮＮを枝刈りした際の、残りの部分ネットワークがモジュール０となる。モジュール０は、入力された画像データが示す数字が、「０」かそれ以外（１～９）かを分類するシングルクラス分類器である。同様に、正解ラベル「１」の訓練データに基づく機械学習により生成した「マスク１」によりオリジナルのＮＮを枝刈りした際の、残りの部分ネットワークがモジュール１となる。モジュール１は、入力された画像データが示す数字が、「１」かそれ以外（０、２～９）かを分類するシングルクラス分類器である。モジュール１とモジュール２とを結合することにより、「０」か「１」かを分類する２クラス分類の分類器が生成される。

[0016] 上述したように、モジュールを結合して生成される分類器のパラメータ数を少なくすることができる点が、オリジナルのＮＮの枝刈りによるモジュール化の効果である。しかし、生成されたモジュール間で、モジュールに含まれるエッジにばらつきが生じる場合がある。このような場合には、モジュール間で共有するエッジが少ないため、すなわち、モジュール間で同じパラメータを共有している割合が低いため、上記の枝刈りによるモジュール化の効果が薄まる。特に、オリジナルのＮＮに含まれるエッジ数が多い場合には、

枝切りのバリエーションが多くなるため、枝切りにより生成されるモジュール間で、モジュールに含まれるエッジのばらつきが大きくなる。そこで、本実施形態では、モジュール間でのエッジのばらつきを抑制するように、各モジュールを生成する。以下、モジュールを生成するためのマスクを生成する第１マスク生成部１４及び第２マスク生成部１６について詳述する。

[0017] 第１マスク生成部１４は、複数のクラスへ分類する機械学習モデルのパラメータと、複数のクラスのうち第１のクラスを正解ラベルに含む訓練データとに基づいて、第１の枝刈り処理に関するパラメータを生成する。第１の枝刈り処理に関するパラメータとは、機械学習モデルに第１の枝刈り処理を実行することによって第１のクラスを分類する第１の機械学習モデルを生成するためのパラメータである。

[0018] 具体的には、図５に示すように、第１マスク生成部１４は、オリジナルのＮＮのエッジの各々に対応したスコアを要素とするスコア行列を用意する。図５右図において、$s_{\{i,j\}}$は、オリジナルのＮＮのｉ番目の層のｊ番目のエッジに対応するスコアである。スコア行列の初期値はランダムに設定してよい。第１マスク生成部１４は、スコア行列において、スコアが高い順に上位ｋ％の要素を１、その他の要素を０としたマスクを生成する。また、第１マスク生成部１４は、スコア行列において、スコアが所定値以上の要素を１、所定値未満の要素を０とするマスクを生成してもよい。第１マスク生成部１４は、オリジナルのＮＮにおいて、生成したマスクの値が１の要素に対応するエッジ及びそのエッジで接続されたニューロンを含む部分ネットワークを用いて、機械学習のフォワード処理を実行する。図５の左図では、オリジナルのＮＮのうち、マスクの値が１の要素に対応する部分を実線、マスクの値が０の要素に対応する部分を点線で表している。より具体的には、第１マスク生成部１４は、マスクの値が１に対応するオリジナルのＮＮの部分に対して、訓練データを入力して順方向に伝播させる。この際、第１マスク生成部１４は、訓練データの正解ラベルｙを、特定のクラスについては正例を示す値、それ以外のクラスについては負例を示す値とする。「特定のクラス」は

、オリジナルのＮＮにより分類される複数のクラスのいずれか１つのクラスであり、複数のクラスのいずれであってもよい。そして、第１マスク生成部１４は、分類結果ｙ＾（後述する数式内では、「ｙ」の上に「＾（ハット）」）を得る。

[0019] また、第１マスク生成部１４は、ＮＮのエッジの重みではなく、スコア行列の各スコアに対して、機械学習のバックワード処理を実行する。具体的には、第１マスク生成部１４は、分類結果ｙ＾が正解ラベルｙに近づくように、誤差逆伝播法によりスコア行列の各スコアを更新する。例えば、第１マスク生成部１４は、下記（１）式に示す損失関数を最小化するように、スコア行列の各スコアを更新する。

[0020] [数1]

$$L = CE(y, \hat{y}) \qquad (1)$$

[0021] なお、ＣＥ（ｙ，ｙ＾）は、ｙとｙ＾とのクロスエントロピーである。第１マスク生成部１４は、上記のフォワード処理及びバックワード処理を機械学習の終了判定を満たすまで繰り返す。これにより、スコア行列において、オリジナルのＮＮに含まれるエッジのうち、特定のクラスを分類するためのモジュールのエッジとして残す適正度が高いエッジに対応するスコアほど大きな値となる。第１マスク生成部１４は、機械学習が終了した際のスコア行列から生成されるマスクを、特定のクラスを分類するためのモジュールを生成するためのマスクとする。以下では、第１マスク生成部１４により生成されるマスクを「ベースマスク」といい、ベースマスクにより生成されるモジュールを「ベースモジュール」という。第１マスク生成部１４は、生成したベースマスクをマスク集合２４に追加して記憶すると共に、ベースマスクに対応するスコア行列を第２マスク生成部１６へ受け渡す。

[0022] 第２マスク生成部１６は、機械学習モデルのパラメータと、複数のクラスのうち第２のクラスを正解ラベルに含む訓練データと、第１の枝刈り処理に関するパラメータを含む損失関数とに基づいて、第２の枝刈り処理に関するパラメータを生成する。第２の枝刈り処理に関するパラメータは、複数のク

ラスへ分類する機械学習モデルに第２の枝刈り処理を実行することによって第２のクラスを分類する第２の機械学習モデルを生成するためのパラメータである。

[0023] 具体的には、第２マスク生成部１６は、特定のクラス以外のクラスを示す正解ラベルの各々が対応付けられた訓練データに基づく機械学習により、第１マスク生成部１４と同様に、各クラスを分類するためのモジュールを生成するためのマスクを生成する。以下では、第２マスク生成部１６により生成されるマスクを「訓練対象マスク」といい、訓練対象マスクにより生成されるモジュールを「訓練対象モジュール」という。この機械学習の際、第２マスク生成部１６は、図６に示すように、訓練対象マスクに対応するスコア行列が、ベースマスクに対応するスコア行列と類似するように、訓練対象マスクに対応するスコア行列の各スコアを更新する。例えば、第２マスク生成部１６は、訓練対象マスク毎に、下記（２）式に示すような損失関数Ｌを最小化するように、スコア行列の各スコアを更新する。

[0024] [数2]

$$L = CE(y, \hat{y}) + \lambda \sum_{i} \sum_{j} \left| s_{\{i,j\}} - s^{*}_{\{i,j\}} \right| \qquad (2)$$

[0025] なお、$s_{\{i,j\}}$は、訓練対象マスクに対応するスコア行列の各要素の値、$s_{\{i,j\}}{}^{*}$は、ベースマスクに対応するスコア行列の各要素の値、λはハイパーパラメータである。すなわち、（２）式に示す損失関数は、（１）式に示す損失関数に、モジュール間のスコア行列の差をペナルティとした正則化項を追加したものである。これにより、ベースモジュールと類似した枝刈り処理を実行するための訓練対象マスクが生成され、ベースモジュールと類似した訓練対象モジュールを生成することができる。第２マスク生成部１６は、生成した訓練対象マスクをマスク集合２４に追加して記憶する。

[0026] 次に、分類装置について説明する。図７に示すように、分類装置３０は、機能的には、特定クラス分類器生成部３２と、分類部３４とを含む。また、分類装置３０の所定の記憶領域には、特定クラス分類器４２が記憶される。

[0027] 特定クラス分類器生成部３２は、運用時のタスクに関するタスク情報を受け付ける。タスク情報には、タスクに応じたモジュールを指定する情報が含まれる。特定クラス分類器生成部３２は、受け付けたタスク情報に基づいて、指定されたモジュールに対応するマスクをマスク集合２４から取得する。また、特定クラス分類器生成部３２は、Ｎクラス分類器２２であるオリジナルのＮＮも取得する。特定クラス分類器生成部３２は、取得したオリジナルのＮＮから、マスク集合２４から取得したマスクの各々に対応する部分の和集合（以下、「和集合部分」という）を切り出す。和集合部分に、取得したマスクを適用することにより、そのマスクに対応するモジュールが生成される。特定クラス分類器生成部３２は、切り出した和集合部分と、取得したマスクとを、特定クラス分類器４２として記憶する。

[0028] 分類部３４は、運用データを取得し、運用データを特定クラス分類器４２に入力し、タスクに応じたクラスのうち、運用データが属するクラスを分類結果として出力する。運用データは、正解ラベルが未知である点を除いて、訓練データと同様である。具体的には、分類部３４は、特定クラス分類器４２に含まれるマスクの各々を和集合部分に適用して、タスクに応じたクラスの各々を分類するためのモジュールの各々を生成する。分類部３４は、各モジュールに運用データを入力して分類結果を得て、分類結果の各々を統合して、最終的な分類結果を決定する。分類部３４は、各モジュールから出力される分類結果、すなわち、運用データが各モジュールに対応するクラスに属する確率のうち、最も高い確率を示すモジュールから出力された分類結果を最終的な分類結果として決定してよい。分類部３４は、決定した最終的な分類結果を出力する。

[0029] ここで、図８を参照して、ＭＮＩＳＴの０～９の１０個の数字に対応する１０クラス分類のＮＮをモジュール化して、運用データをクラス「３」又はクラス「５」に分類するタスクを例に、機械学習装置１０及び分類装置３０の処理の一例を説明する。

[0030] まず、Ｎクラス分類器生成部１２は、訓練データを用いた機械学習により

、Ｎ（Ｎ＝１０）クラス分類器２２を生成する。第１マスク生成部１４は、１つのクラスを示す正解ラベルが対応付けられた訓練データとＮクラス分類器２２とに基づく機械学習により、ベースマスクを生成する。図８では、正解ラベル「０」が対応付けられた訓練データに基づき、入力された画像データが示す数字が、０か０以外かを分類するシングルクラス分類器であるモジュール０を生成するためのモジュール０マスクをベースマスクとする例を示している。第２マスク生成部１４は、モジュール１～モジュール９の各々を生成するためのモジュール１マスク～モジュール９マスクの各々を生成する。この際、第２マスク生成部１６は、モジュール１マスク～モジュール９マスクの各々に対応するスコア行列の各々が、ベースマスクに対応するスコア行列と類似するように、モジュール１マスク～モジュール９マスクの各々を生成する。これにより、モジュール０～９の各々に対応する１０個のマスクが生成され、マスク集合２４として記憶される。

[0031] 特定クラス分類器生成部３２は、モジュール３及びモジュール５を指定するタスク情報を受け付け、マスク集合２４から、モジュール３マスク及びモジュール５マスクを取得する。特定クラス分類器生成部３２は、Ｎクラス分類器２２のうち、モジュール３マスクに対応する部分と、モジュール５マスクに対応する部分との和集合部分を切り出し、モジュール３マスク及びモジュール５マスクと共に、特定クラス分類器４２として記憶する。特定クラス分類器生成部３２は、和集合部分にモジュール３マスクを適用してモジュール３を生成し、和集合部分にモジュール５マスクを適用してモジュール５を生成する。分類部３４は、運用データとして、手書き数字の画像データを取得し、運用データをモジュール３に入力することにより、運用データが示す数字が３である確率を示す分類結果を得る。同様に、分類部３４は、運用データをモジュール５に入力することにより、運用データが示す数字が５である確率を示す分類結果を得る。分類部３４は、確率が高い方の分類結果を最終的な分類結果として出力する。例えば、分類部３４は、モジュール３の分類結果が９０％、モジュール５の分類結果が１０％の場合、運用データが示

す数字は「３」であるとする分類結果を出力する。

[0032] 機械学習装置１０は、例えば図９に示すコンピュータ５０で実現されてよい。コンピュータ５０は、ＣＰＵ（Central Processing Unit）５１と、一時記憶領域としてのメモリ５２と、不揮発性の記憶部５３とを備える。また、コンピュータ５０は、入力部、表示部等の入出力装置５４と、記憶媒体５９に対するデータの読み込み及び書き込みを制御するＲ／Ｗ（Read/Write）部５５とを備える。また、コンピュータ５０は、インターネット等のネットワークに接続される通信Ｉ／Ｆ（Interface）５６を備える。ＣＰＵ５１、メモリ５２、記憶部５３、入出力装置５４、Ｒ／Ｗ部５５、及び通信Ｉ／Ｆ５６は、バス５７を介して互いに接続される。

[0033] 記憶部５３は、ＨＤＤ（Hard Disk Drive）、ＳＳＤ（Solid State Drive）、フラッシュメモリ等によって実現してよい。記憶媒体としての記憶部５３には、コンピュータ５０を、機械学習装置１０として機能させるための機械学習プログラム６０が記憶される。機械学習プログラム６０は、Ｎクラス分類器生成プロセス６２と、第１マスク生成プロセス６４と、第２マスク生成プロセス６６とを有する。また、記憶部５３は、Ｎクラス分類器２２及びマスク集合２４の各々を構成する情報が記憶される情報記憶領域７０を有する。

[0034] ＣＰＵ５１は、機械学習プログラム６０を記憶部５３から読み出してメモリ５２に展開し、機械学習プログラム６０が有するプロセスを順次実行する。ＣＰＵ５１は、Ｎクラス分類器生成プロセス６２を実行することで、図１に示すＮクラス分類器生成部１２として動作する。また、ＣＰＵ５１は、第１マスク生成プロセス６４を実行することで、図１に示す第１マスク生成部１４として動作する。また、ＣＰＵ５１は、第２マスク生成プロセス６６を実行することで、図１に示す第２マスク生成部１６として動作する。また、ＣＰＵ５１は、情報記憶領域７０から情報を読み出して、Ｎクラス分類器２２及びマスク集合２４の各々をメモリ５２に展開する。これにより、機械学習プログラム６０を実行したコンピュータ５０が、機械学習装置１０として

機能することになる。なお、プログラムを実行するＣＰＵ５１はハードウェアである。

[0035] 分類装置３０は、例えば図１０に示すコンピュータ８０で実現されてよい。コンピュータ８０は、ＣＰＵ８１と、一時記憶領域としてのメモリ８２と、不揮発性の記憶部８３とを備える。また、コンピュータ８０は、入力部、表示部等の入出力装置８４と、記憶媒体８９に対するデータの読み込み及び書き込みを制御するＲ／W部８５と、通信Ｉ／Ｆ８６とを備える。ＣＰＵ８１、メモリ８２、記憶部８３、入出力装置８４、Ｒ／W部８５、及び通信Ｉ／Ｆ８６は、バス８７を介して互いに接続される。

[0036] 記憶部８３は、ＨＤＤ、ＳＳＤ、フラッシュメモリ等によって実現してよい。記憶媒体としての記憶部８３には、コンピュータ８０を、分類装置３０として機能させるための分類プログラム９０が記憶される。分類プログラム９０は、特定クラス分類器生成プロセス９２と、分類プロセス９４とを有する。また、記憶部８３は、特定クラス分類器４２を構成する情報が記憶される情報記憶領域１００を有する。

[0037] ＣＰＵ８１は、分類プログラム９０を記憶部８３から読み出してメモリ８２に展開し、分類プログラム９０が有するプロセスを順次実行する。ＣＰＵ８１は、特定クラス分類器生成プロセス９２を実行することで、図７に示す特定クラス分類器生成部３２として動作する。また、ＣＰＵ８１は、分類プロセス９４を実行することで、図７に示す分類部３４として動作する。また、ＣＰＵ８１は、情報記憶領域１００から情報を読み出して、特定クラス分類器４２をメモリ８２に展開する。これにより、分類プログラム９０を実行したコンピュータ８０が、分類装置３０として機能することになる。なお、プログラムを実行するＣＰＵ８１はハードウェアである。

[0038] なお、機械学習プログラム６０及び分類プログラム９０の各々により実現される機能は、例えば半導体集積回路、より詳しくはＡＳＩＣ（Application Specific Integrated Circuit）等で実現することも可能である。

[0039] 次に、本実施形態に係る機械学習システムの作用について説明する。機械

学習装置１０に訓練データが入力され、モジュール化のためのマスクの生成が指示されると、機械学習装置１０において、図１１に示す機械学習処理が実行される。また、分類装置３０にタスク情報及び運用データが入力され、運用データの分類が指示されると、分類装置３０において、図１２に示す分類処理が実行される。以下、機械学習処理及び分類処理の各々について説明する。なお、機械学習処理及び分類処理は、開示の技術の機械学習方法の一例である。

[0040] まず、図１１を参照して、機械学習処理について説明する。

ステップＳ１０で、Ｎクラス分類器生成部１２が、機械学習装置１０に入力された訓練データを取得する。次に、ステップＳ１２で、Ｎクラス分類器生成部１２が、取得した訓練データを用いた機械学習により、Ｎクラス分類器２２を生成する。

[0041] 次に、ステップＳ１４で、第１マスク生成部１４が、Ｎクラス分類器２２のエッジの各々に対応したスコアを要素とするスコア行列を用意する。また、第１マスク生成部１４が、取得した訓練データの正解ラベルｙを、特定のクラスについては正例、それ以外のクラスについては負例とする。そして、第１マスク生成部１４が、スコア行列において、スコアが高い順に上位ｋ％の要素を１、その他の要素を０としたマスクをＮクラス分類器２２に適用した部分に対して、訓練データを入力して順方向に伝播させ、分類結果ｙ＾を得る。さらに、第１マスク生成部１４が、分類結果ｙ＾が正解ラベルｙに近づくように、誤差逆伝播法によりスコア行列の各スコアを更新し、機械学習が終了した際のスコア行列からベースマスクを生成する。

[0042] 次に、ステップＳ１６で、上記ステップＳ１４と同様に、第２マスク生成部１６が、特定のクラス以外のクラスを分類するための訓練対象モジュールの各々を生成するための訓練対象マスクを生成する。この際、第２マスク生成部１６は、スコア行列がベースマスクのスコア行列と類似するように、訓練対象マスクの各々を生成する。次に、ステップＳ１８で、第１マスク生成部１４及び第２マスク生成部１６の各々が、生成したマスクをマスク集合２

４として記憶し、機械学習処理は終了する。

[0043] 次に、図１２を参照して、分類処理について説明する。

ステップＳ２０で、特定クラス分類器生成部３２が、タスクに応じたモジュールを指定する情報が含まれるタスク情報を取得する。次に、ステップＳ２２で、特定クラス分類器生成部３２が、タスク情報で指定されたモジュールに対応するマスクをマスク集合２４から取得する。また、特定クラス分類器生成部３２が、Ｎクラス分類器２２から、取得したマスクの各々に対応する部分の和集合部分を切り出し、取得したマスクと共に、特定クラス分類器４２として記憶する。

[0044] 次に、ステップＳ２４で、分類部３４が、分類装置３０に入力された運用データを取得する。次に、ステップＳ２６で、分類部３４が、特定クラス分類器４２に含まれるマスクの各々を和集合部分に適用して、タスクに応じたクラスの各々を分類するためのモジュールの各々を生成する。そして、分類部３４が、各モジュールに運用データを入力して分類結果を得て、分類結果の各々を統合して、最終的な分類結果を決定し、出力する。そして、分類処理は終了する。

[0045] 以上説明したように、本実施形態に係る機械学習システムによれば、機械学習装置が、Ｎクラス分類器と、Ｎクラスのうちの第１のクラスを正解ラベルに含む訓練データとに基づいて、ベースモジュールを生成するための枝刈り処理に用いるベースマスクを生成する。ベースマスクは、Ｎクラス分類器のエッジの各々に対応したスコアを要素とするスコア行列をバイナリ化することにより生成される。また、機械学習装置が、Ｎクラス分類器と、Ｎクラスのうち第２のクラスを正解ラベルに含む訓練データと、ベースマスクに対応するスコア行列の値を含む損失関数とに基づいて、他のモジュールを生成するための枝刈り処理に用いるマスクを生成する。このように生成されたマスクをＮクラス分類器に適用して各モジュールを生成することにより、モジュール間で同じパラメータを共有する割合が高まる。したがって、モジュールを結合して生成される分類器のパラメータ数を少なくすることができると

いう枝刈りの効果が薄まることを抑制することができる。

[0046] 図１３を参照して、本実施形態におけるモジュール間でのパラメータの共有割合の実験結果の一例について説明する。この実験では、図８で説明したＭＮＩＳＴの例と同様に、１０クラス分類器（オリジナルのＮＮ）から、０～９の各クラスを分類するための１０個のモジュールを生成した。なお、各モジュールは、オリジナルのＮＮの５％のパラメータ（各エッジの重み）数となるように枝切り処理を行うことにより生成した。図１３に示す比較例１及び比較例２は、マスク作成のためのスコア行列の更新時に用いる損失関数に、モジュール間のスコア行列の差をペナルティとした正則化項を含まない手法である。また、比較例１は、スコア行列の初期値をモジュール間で共有しない場合、比較例２は、共有する場合である。また、本手法１及び本手法２は、上記実施形態のように、損失関数に、モジュール間のスコア行列の差をペナルティとした正則化項を含む手法である。また、本手法１では、正則化項に対するハイパーパラメータλをλ＝１、本手法２では、λ＝１０とした。また、本手法１及び本手法２とも、スコア行列の初期値をモジュール間で共有した。

[0047] また、図１３において、「オリジナルのＮＮに対する結合時のパラメータ数の割合」は、オリジナルのＮＮのパラメータ数に対する、２つのモジュールを結合した場合のパラメータ数の割合である。また、この値は、１０個のモジュールのうちの２つのモジュールの全組み合わせ４５ペアの平均である。上述したように、各モジュールは、オリジナルのＮＮの５％のパラメータ数となるように生成されているため、この値は０．０５～０．１となり、この値が小さいほど２つのモジュールが類似していることを表す。また、図１３において、「共有パラメータの割合」は、結合する２つのモジュールで共有するパラメータ数の割合である。

[0048] 上述したように、モジュール同士が類似する場合は、モジュールを結合した場合のモデルのパラメータ数を削減することができるという効果がある。この効果を得るために、本手法とは異なり、スコア行列の初期値をモジュー

ル間で共有させることも考えられる。しかし、図１３に示すように、比較例１と比較例２とを比較すると、スコア行列の初期値をモジュール間で共有することにより、共有パラメータの割合が多少向上してはいる。これに対し、本手法１及び本手法２では、共有パラメータの割合が大きく向上している。すなわち、比較例１及び比較例２に比べ、本手法１及び本手法２は、よりモデルのデータ数を削減できる効果が高い。また、本手法２のように、損失関数に含まれる正則化項の影響をより大きくした場合には、その効果がより顕著になっている。

[0049] なお、上記実施形態では、１クラス単位でモジュール化する例について説明したが、これに限定されない。例えば、１０クラス分類器から、２クラスを分類するためのモジュールや、３クラスを分類するためのモジュールを生成するなど、元の機械学習モデルで分類可能な複数のクラスの一部を分類可能な部分ネットワークをモジュール化するものであればよい。

[0050] また、上記実施形態では、機械学習装置と分類装置とをそれぞれ別のコンピュータで実現する場合について説明したが、機械学習装置と分類装置とを１つのコンピュータで実現してもよい。

[0051] また、上記実施形態では、機械学習プログラム及び分類プログラムの各々が記憶部に予め記憶（インストール）されている態様を説明したが、これに限定されない。開示の技術に係るプログラムは、ＣＤ－ＲＯＭ、ＤＶＤ－ＲＯＭ、ＵＳＢメモリ等の記憶媒体に記憶された形態で提供することも可能である。

符号の説明

[0052] １０ 機械学習装置

１２ Ｎクラス分類器生成部

１４ 第１マスク生成部

１６ 第２マスク生成部

２２ Ｎクラス分類器

２４ マスク集合

３０　分類装置

３２　特定クラス分類器生成部

３４　分類部

４２　特定クラス分類器

５０、８０　コンピュータ

５１、８１　ＣＰＵ

５２、８２　メモリ

５３、８３　記憶部

５９、８９　記憶媒体

６０　機械学習プログラム

９０　分類プログラム

請求の範囲

[請求項1] 複数のクラスへ分類する機械学習モデルのパラメータと、前記複数のクラスのうち第１のクラスを正解ラベルに含む訓練データとに基づいて、前記機械学習モデルに第１の枝刈り処理を実行することによって前記第１のクラスを分類する第１の機械学習モデルを生成するための前記第１の枝刈り処理に関するパラメータを生成し、

前記機械学習モデルのパラメータと、前記複数のクラスのうち第２のクラスを正解ラベルに含む訓練データと、前記第１の枝刈り処理に関するパラメータを含む損失関数とに基づいて、前記機械学習モデルに第２の枝刈り処理を実行することによって前記第２のクラスを分類する第２の機械学習モデルを生成するための前記第２の枝刈り処理に関するパラメータを生成する、

処理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。

[請求項2] 前記第２の枝刈り処理に関するパラメータを生成する処理は、前記第２の機械学習モデルによる前記第２のクラスの分類結果と前記正解ラベルとの差を表す項と、前記第１の枝刈り処理に関するパラメータと前記第２の枝刈り処理に関するパラメータとの差を表す項とを含む前記損失関数を最小化することを含む、

請求項１に記載の機械学習プログラム。

[請求項3] 前記機械学習モデルはニューラルネットワークであり、

前記第１の枝刈り処理に関するパラメータは、前記複数のクラスへ分類する機械学習モデルに含まれるエッジの各々に対応し、前記第１の機械学習モデルのエッジとして残す適正度が高いエッジほど大きくなるスコアであり、

前記第２の枝刈り処理に関するパラメータは、前記複数のクラスへ分類する機械学習モデルに含まれるエッジの各々に対応し、前記第２の機械学習モデルのエッジとして残す適正度が高いエッジほど大きく

なるスコアである、

ことを特徴とする請求項１又は請求項２に記載の機械学習プログラム。

[請求項4] 前記複数のクラスへ分類する機械学習モデルに含まれるエッジのうち、対応する前記第１の枝刈り処理に関するパラメータが大きい順に上位所定割合の数のエッジ、又は対応する前記第１の枝刈り処理に関するパラメータが所定値以上のエッジを、前記第１の機械学習モデルに残すことを表す第１のマスクを生成し、

前記複数のクラスへ分類する機械学習モデルに含まれるエッジのうち、対応する前記第２の枝刈り処理に関するパラメータが大きい順に上位所定割合の数のエッジ、又は対応する前記第２の枝刈り処理に関するパラメータが所定値以上のエッジを、前記第２の機械学習モデルに残すことを表す第２のマスクを生成する、

処理を前記コンピュータに実行させることを特徴とする請求項３に記載の機械学習プログラム。

[請求項5] 前記第１のマスク及び前記第２のマスクから選択された１以上のマスクの和集合に対応する前記複数のクラスへ分類する機械学習モデルの部分に、前記選択された１以上のマスクの各々を適用することにより、前記選択された１以上のマスクに対応するクラスを分類する第３の機械学習モデルを生成する、

処理を前記コンピュータに実行させることを特徴とする請求項４に記載の機械学習プログラム。

[請求項6] 分類されるクラスが未知のデータを前記第３の機械学習モデルに入力し、前記データを前記選択された１以上のマスクに対応するクラスに分類する、

処理を前記コンピュータに実行させることを特徴とする請求項５に記載の機械学習プログラム。

[請求項7] 複数のクラスへ分類する機械学習モデルのパラメータと、前記複数

のクラスのうち第１のクラスを正解ラベルに含む訓練データとに基づいて、前記機械学習モデルに第１の枝刈り処理を実行することによって前記第１のクラスを分類する第１の機械学習モデルを生成するための前記第１の枝刈り処理に関するパラメータを生成し、

前記機械学習モデルのパラメータと、前記複数のクラスのうち第２のクラスを正解ラベルに含む訓練データと、前記第１の枝刈り処理に関するパラメータを含む損失関数とに基づいて、前記機械学習モデルに第２の枝刈り処理を実行することによって前記第２のクラスを分類する第２の機械学習モデルを生成するための前記第２の枝刈り処理に関するパラメータを生成する、

処理を実行する制御部を含むことを特徴とする機械学習装置。

[請求項8] 前記第２の枝刈り処理に関するパラメータを生成する処理は、前記第２の機械学習モデルによる前記第２のクラスの分類結果と前記正解ラベルとの差を表す項と、前記第１の枝刈り処理に関するパラメータと前記第２の枝刈り処理に関するパラメータとの差を表す項とを含む前記損失関数を最小化することを含む、

請求項７に記載の機械学習装置。

[請求項9] 前記機械学習モデルはニューラルネットワークであり、

前記第１の枝刈り処理に関するパラメータは、前記複数のクラスへ分類する機械学習モデルに含まれるエッジの各々に対応し、前記第１の機械学習モデルのエッジとして残す適正度が高いエッジほど大きくなるスコアであり、

前記第２の枝刈り処理に関するパラメータは、前記複数のクラスへ分類する機械学習モデルに含まれるエッジの各々に対応し、前記第２の機械学習モデルのエッジとして残す適正度が高いエッジほど大きくなるスコアである、

ことを特徴とする請求項７又は請求項８に記載の機械学習装置。

[請求項10] 前記制御部は、

前記複数のクラスへ分類する機械学習モデルに含まれるエッジのうち、対応する前記第１の枝刈り処理に関するパラメータが大きい順に上位所定割合の数のエッジ、又は対応する前記第１の枝刈り処理に関するパラメータが所定値以上のエッジを、前記第１の機械学習モデルに残すことを表す第１のマスクを生成し、

前記複数のクラスへ分類する機械学習モデルに含まれるエッジのうち、対応する前記第２の枝刈り処理に関するパラメータが大きい順に上位所定割合の数のエッジ、又は対応する前記第２の枝刈り処理に関するパラメータが所定値以上のエッジを、前記第２の機械学習モデルに残すことを表す第２のマスクを生成する、

処理をさらに実行することを特徴とする請求項９に記載の機械学習装置。

[請求項11] 前記制御部は、前記第１のマスク及び前記第２のマスクから選択された１以上のマスクの和集合に対応する前記複数のクラスへ分類する機械学習モデルの部分に、前記選択された１以上のマスクの各々を適用することにより、前記選択された１以上のマスクに対応するクラスを分類する第３の機械学習モデルを生成する、

処理をさらに実行することを特徴とする請求項１０に記載の機械学習装置。

[請求項12] 前記制御部は、分類されるクラスが未知のデータを前記第３の機械学習モデルに入力し、前記データを前記選択された１以上のマスクに対応するクラスに分類する、

処理をさらに実行することを特徴とする請求項１１に記載の機械学習装置。

[請求項13] 複数のクラスへ分類する機械学習モデルのパラメータと、前記複数のクラスのうち第１のクラスを正解ラベルに含む訓練データとに基づいて、前記機械学習モデルに第１の枝刈り処理を実行することによって前記第１のクラスを分類する第１の機械学習モデルを生成するため

の前記第１の枝刈り処理に関するパラメータを生成し、

前記機械学習モデルのパラメータと、前記複数のクラスのうち第２のクラスを正解ラベルに含む訓練データと、前記第１の枝刈り処理に関するパラメータを含む損失関数とに基づいて、前記機械学習モデルに第２の枝刈り処理を実行することによって前記第２のクラスを分類する第２の機械学習モデルを生成するための前記第２の枝刈り処理に関するパラメータを生成する、

処理をコンピュータが実行することを特徴とする機械学習方法。

[請求項14] 前記第２の枝刈り処理に関するパラメータを生成する処理は、前記第２の機械学習モデルによる前記第２のクラスの分類結果と前記正解ラベルとの差を表す項と、前記第１の枝刈り処理に関するパラメータと前記第２の枝刈り処理に関するパラメータとの差を表す項とを含む前記損失関数を最小化することを含む、

請求項１３に記載の機械学習方法。

[請求項15] 前記機械学習モデルはニューラルネットワークであり、

前記第１の枝刈り処理に関するパラメータは、前記複数のクラスへ分類する機械学習モデルに含まれるエッジの各々に対応し、前記第１の機械学習モデルのエッジとして残す適正度が高いエッジほど大きくなるスコアであり、

前記第２の枝刈り処理に関するパラメータは、前記複数のクラスへ分類する機械学習モデルに含まれるエッジの各々に対応し、前記第２の機械学習モデルのエッジとして残す適正度が高いエッジほど大きくなるスコアである、

ことを特徴とする請求項１３又は請求項１４に記載の機械学習方法。

[請求項16] 前記複数のクラスへ分類する機械学習モデルに含まれるエッジのうち、対応する前記第１の枝刈り処理に関するパラメータが大きい順に上位所定割合の数のエッジ、又は対応する前記第１の枝刈り処理に関

するパラメータが所定値以上のエッジを、前記第１の機械学習モデルに残すことを表す第１のマスクを生成し、

前記複数のクラスへ分類する機械学習モデルに含まれるエッジのうち、対応する前記第２の枝刈り処理に関するパラメータが大きい順に上位所定割合の数のエッジ、又は対応する前記第２の枝刈り処理に関するパラメータが所定値以上のエッジを、前記第２の機械学習モデルに残すことを表す第２のマスクを生成する、

処理を前記コンピュータが実行することを特徴とする請求項１５に記載の機械学習方法。

[請求項17] 前記第１のマスク及び前記第２のマスクから選択された１以上のマスクの和集合に対応する前記複数のクラスへ分類する機械学習モデルの部分に、前記選択された１以上のマスクの各々を適用することにより、前記選択された１以上のマスクに対応するクラスを分類する第３の機械学習モデルを生成する、

処理を前記コンピュータが実行することを特徴とする請求項１６に記載の機械学習方法。

[請求項18] 分類されるクラスが未知のデータを前記第３の機械学習モデルに入力し、前記データを前記選択された１以上のマスクに対応するクラスに分類する、

処理を前記コンピュータが実行することを特徴とする請求項１７に記載の機械学習方法。

[請求項19] 複数のクラスへ分類する機械学習モデルのパラメータと、前記複数のクラスのうち第１のクラスを正解ラベルに含む訓練データとに基づいて、前記機械学習モデルに第１の枝刈り処理を実行することによって前記第１のクラスを分類する第１の機械学習モデルを生成するための前記第１の枝刈り処理に関するパラメータを生成し、

前記機械学習モデルのパラメータと、前記複数のクラスのうち第２のクラスを正解ラベルに含む訓練データと、前記第１の枝刈り処理に

関するパラメータを含む損失関数とに基づいて、前記機械学習モデルに第２の枝刈り処理を実行することによって前記第２のクラスを分類する第２の機械学習モデルを生成するための前記第２の枝刈り処理に関するパラメータを生成する、

処理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を記憶した記憶媒体。

[請求項20] 機械学習装置と分類装置とを含む機械学習システムであって、

前記機械学習装置は、

複数のクラスへ分類する機械学習モデルのパラメータと、前記複数のクラスのうち第１のクラスを正解ラベルに含む訓練データとに基づいて、前記機械学習モデルに第１の枝刈り処理を実行することによって前記第１のクラスを分類する第１の機械学習モデルを生成するための前記第１の枝刈り処理に関するパラメータを生成し、

前記機械学習モデルのパラメータと、前記複数のクラスのうち第２のクラスを正解ラベルに含む訓練データと、前記第１の枝刈り処理に関するパラメータを含む損失関数とに基づいて、前記機械学習モデルに第２の枝刈り処理を実行することによって前記第２のクラスを分類する第２の機械学習モデルを生成するための前記第２の枝刈り処理に関するパラメータを生成し、

処理を実行する第１の制御部を含み、

前記分類装置は、

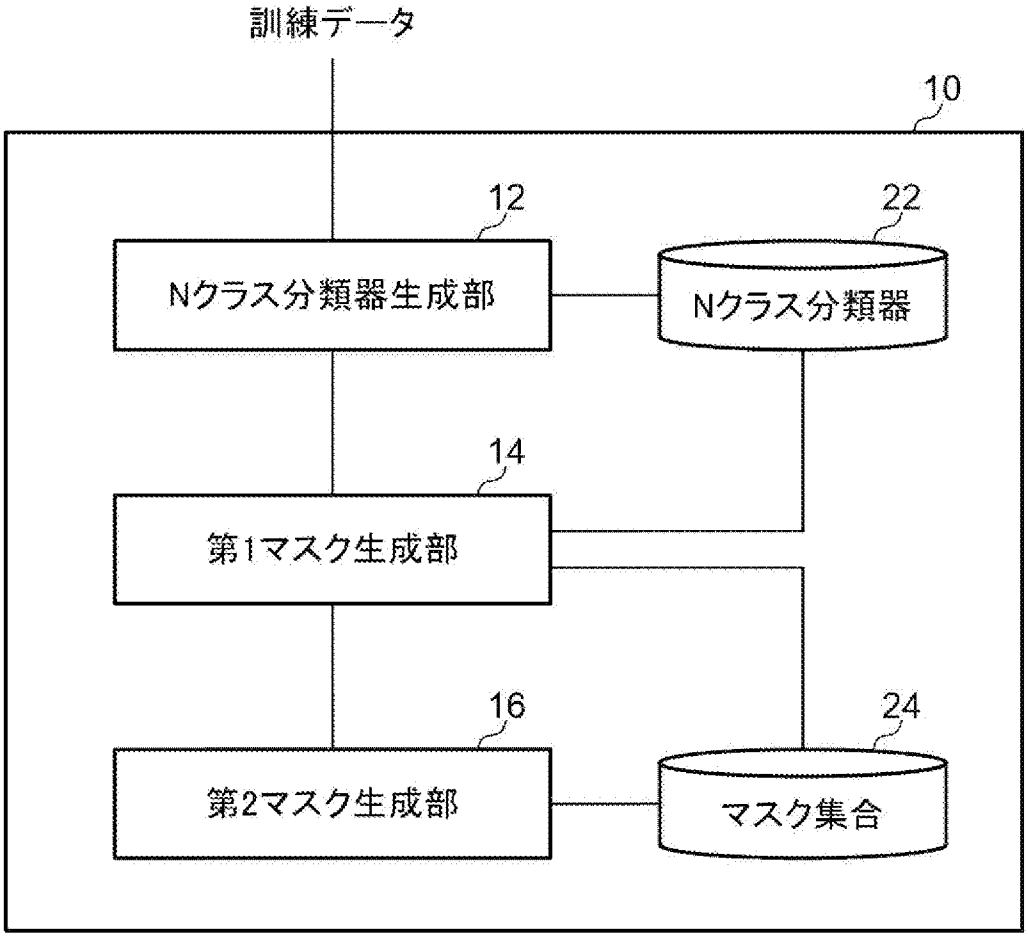
前記第１の機械学習モデル及び前記第２の機械学習モデルから選択された１以上の機械学習モデルを結合し、前記選択された１以上の機械学習モデルに対応するクラスを分類する第３の機械学習モデルを生成し、

分類されるクラスが未知のデータを前記第３の機械学習モデルに入力し、前記データを前記選択された１以上の機械学習モデルに対応するクラスに分類する、

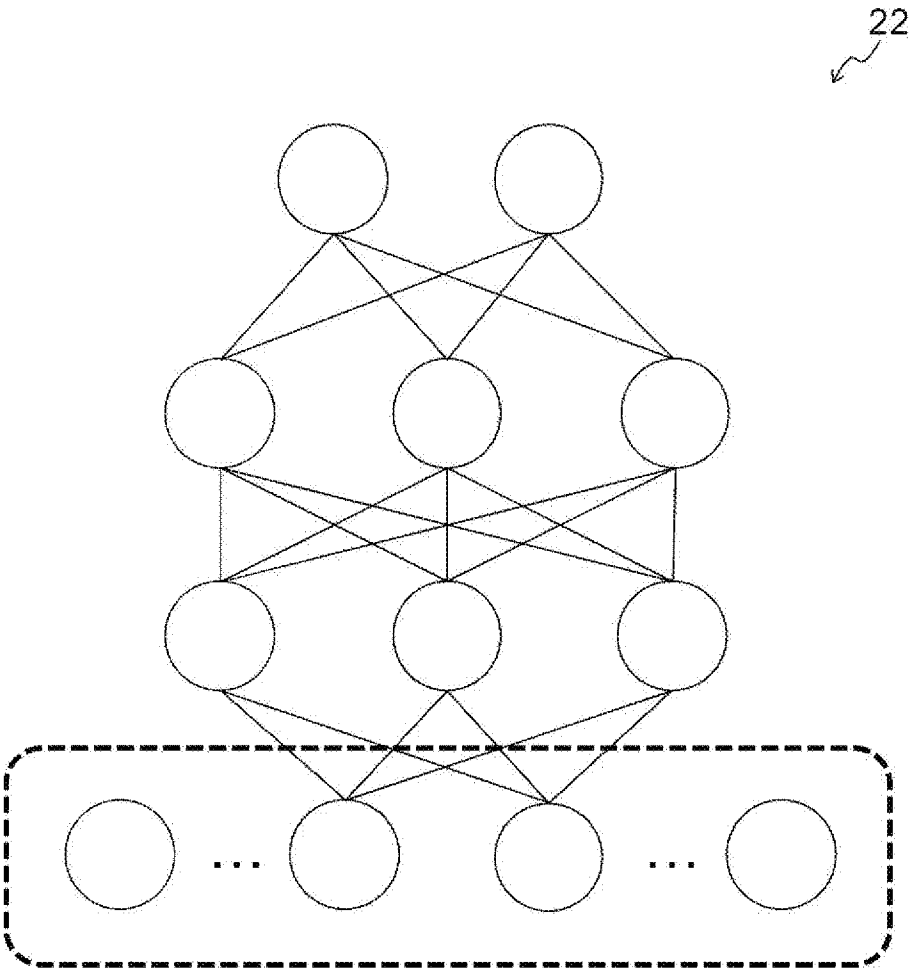
処理を実行する第２の制御部を含む、

ことを特徴とする機械学習システム。

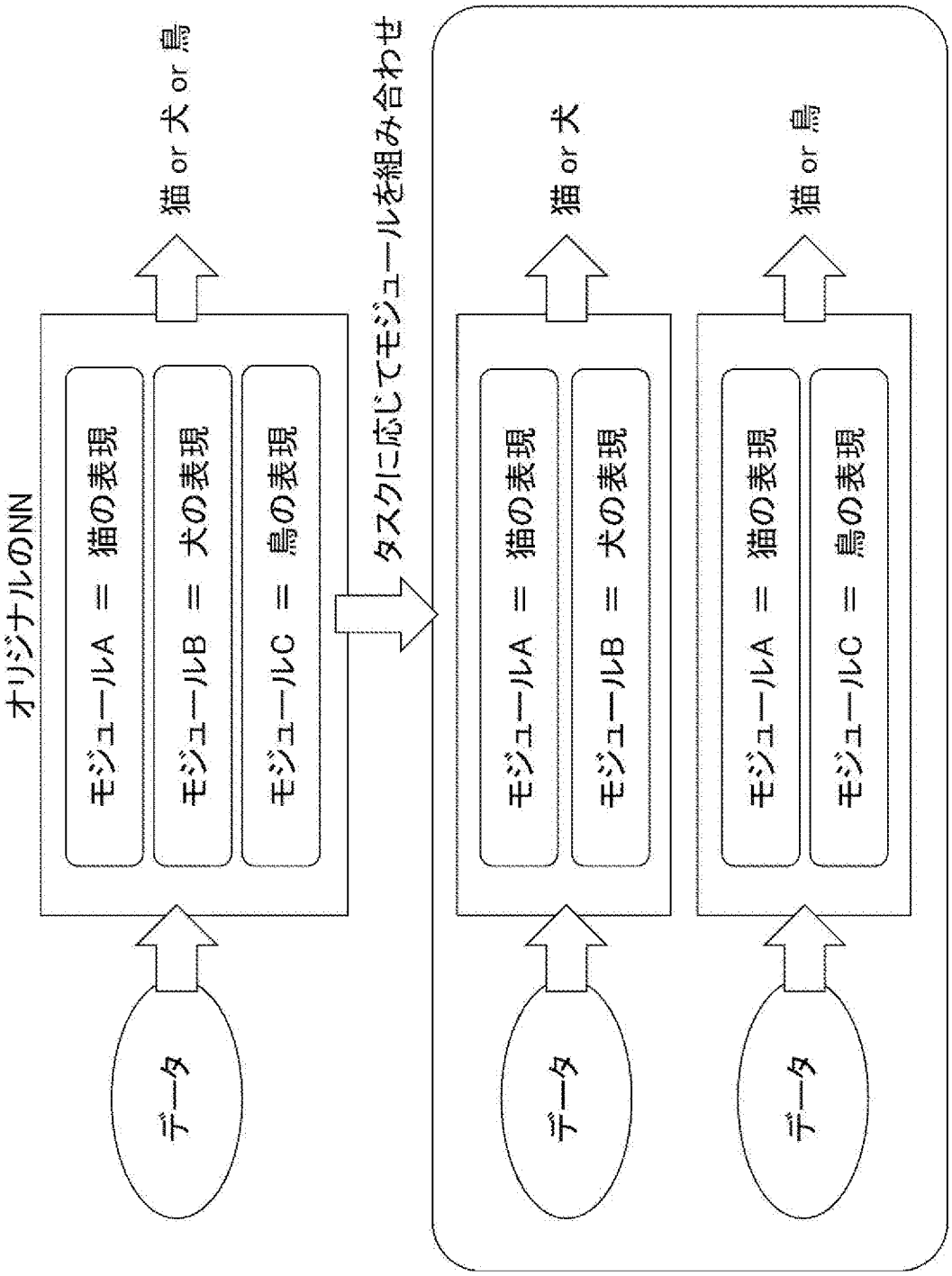
[図1]


訓練データ
10
12
Nクラス分類器生成部
22
Nクラス分類器
14
第1マスク生成部
16
第2マスク生成部
24
マスク集合

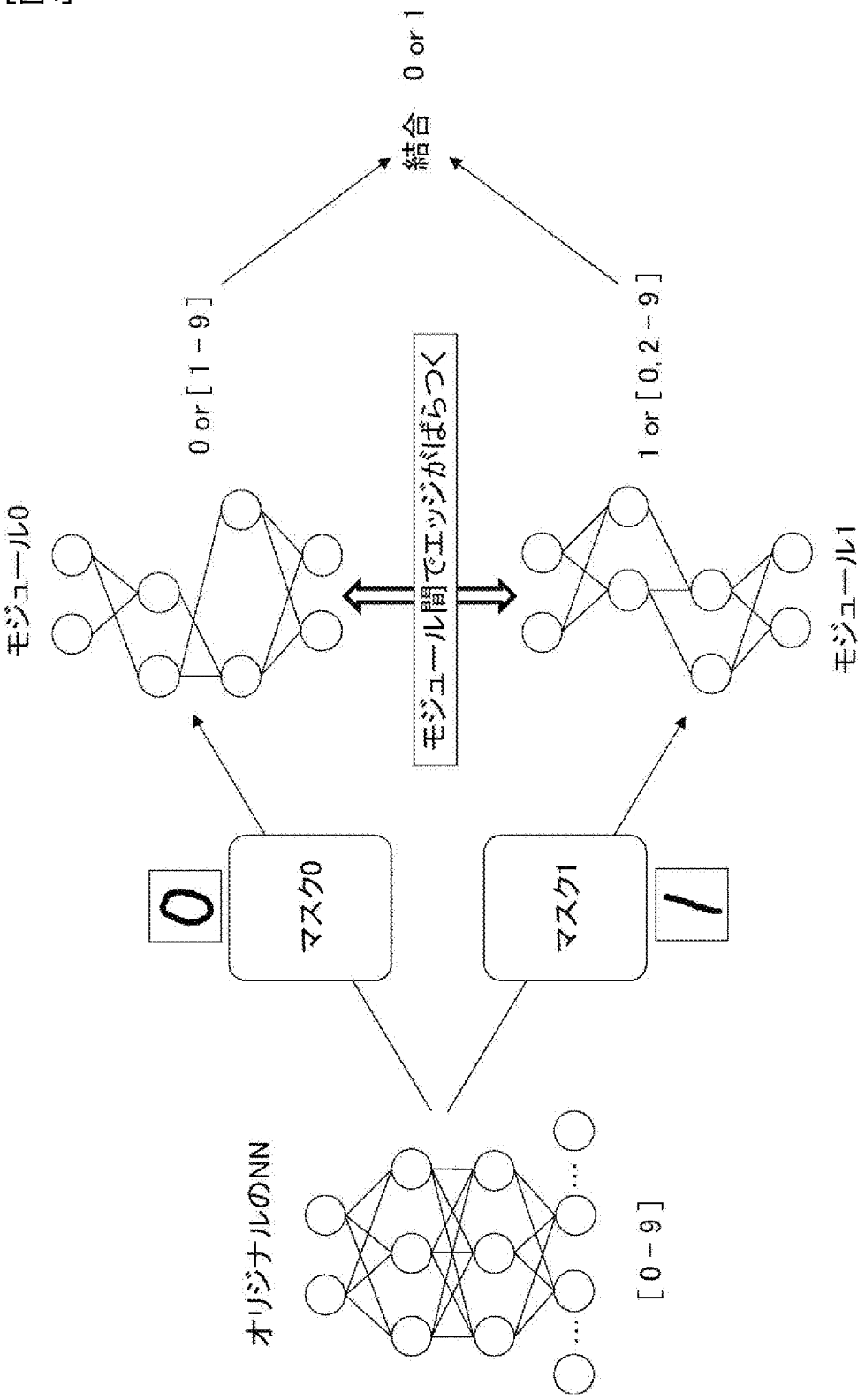
[図2]


22

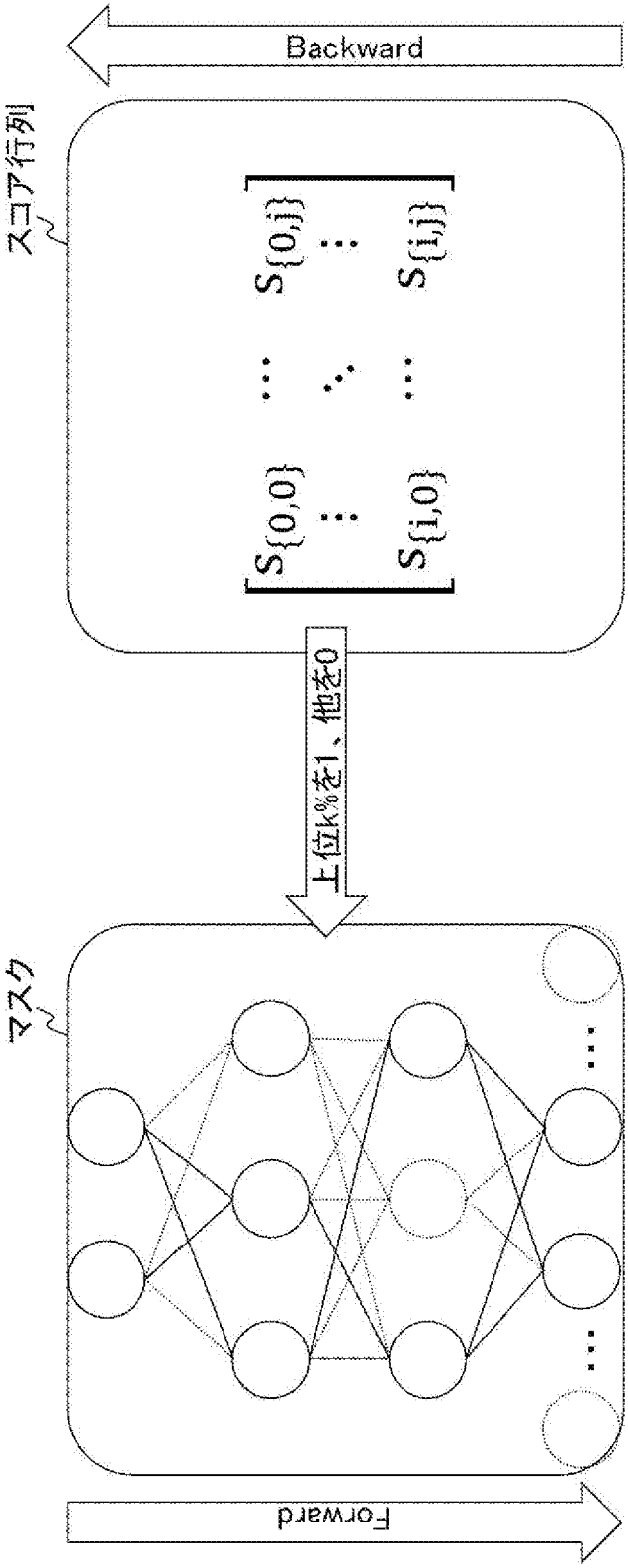
[図3]


オリジナルのNN
モジュールA ＝ 猫の表現
モジュールB ＝ 犬の表現
モジュールC ＝ 鳥の表現
猫 or 犬 or 鳥
データ
タスクに応じてモジュールを組み合わせ
モジュールA ＝ 猫の表現
モジュールB ＝ 犬の表現
猫 or 犬
データ
モジュールA ＝ 猫の表現
モジュールC ＝ 鳥の表現
猫 or 鳥
データ

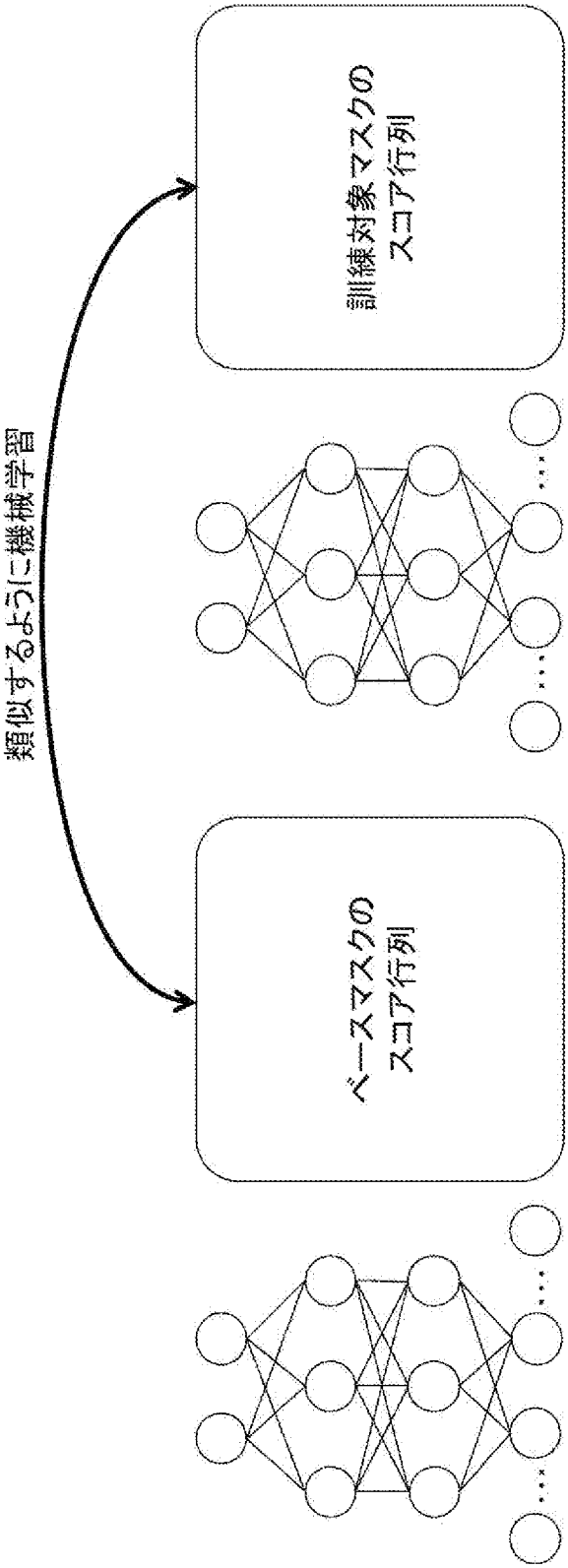
[図4]


0 or 1
結合
0 or [1 – 9]
1 or [0, 2 – 9]
モジュール0
モジュール1
モジュール間でエッジがばらつく
0
マスク0
マスク1
1
オリジナルのNN
[0 – 9]

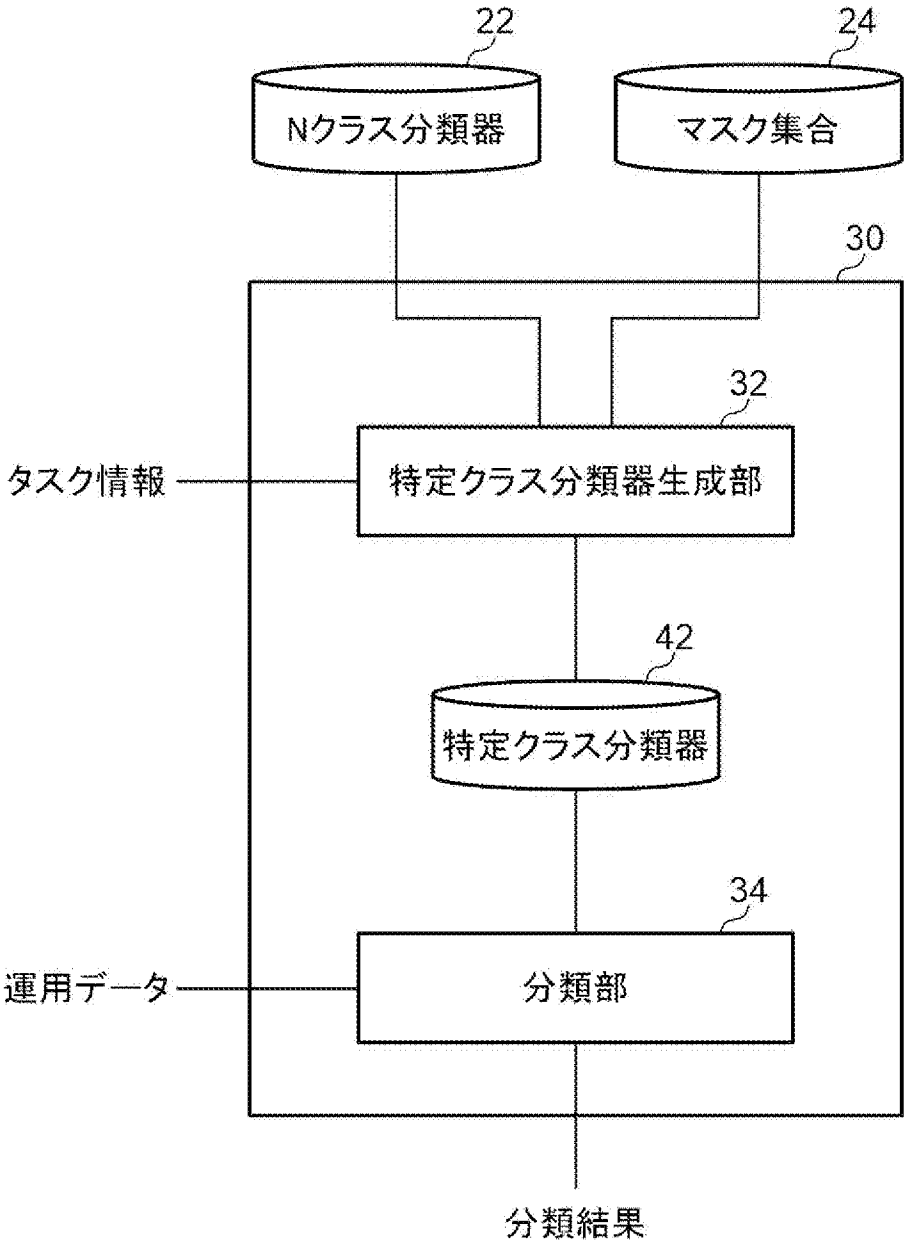
[図5]



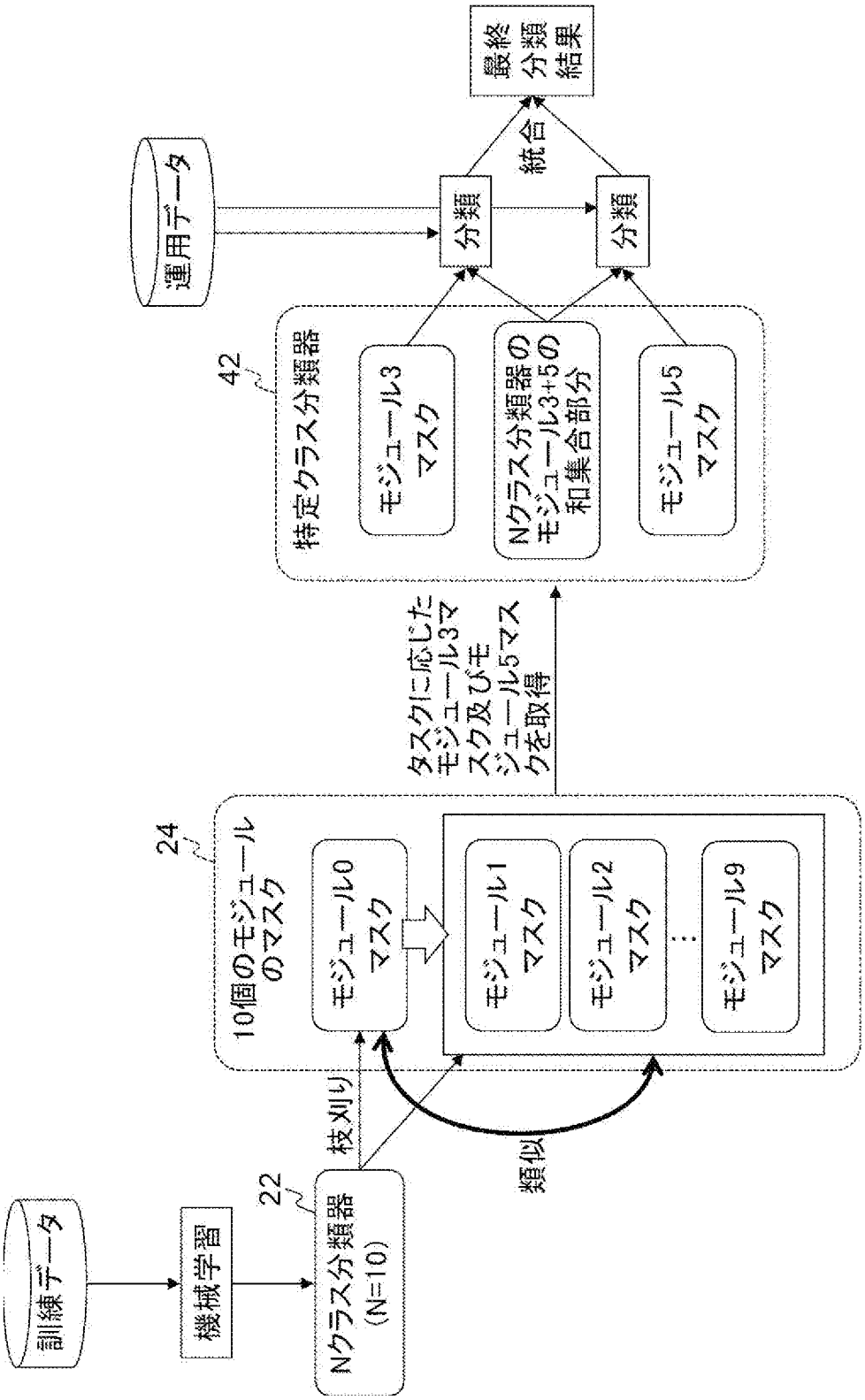
[図6]



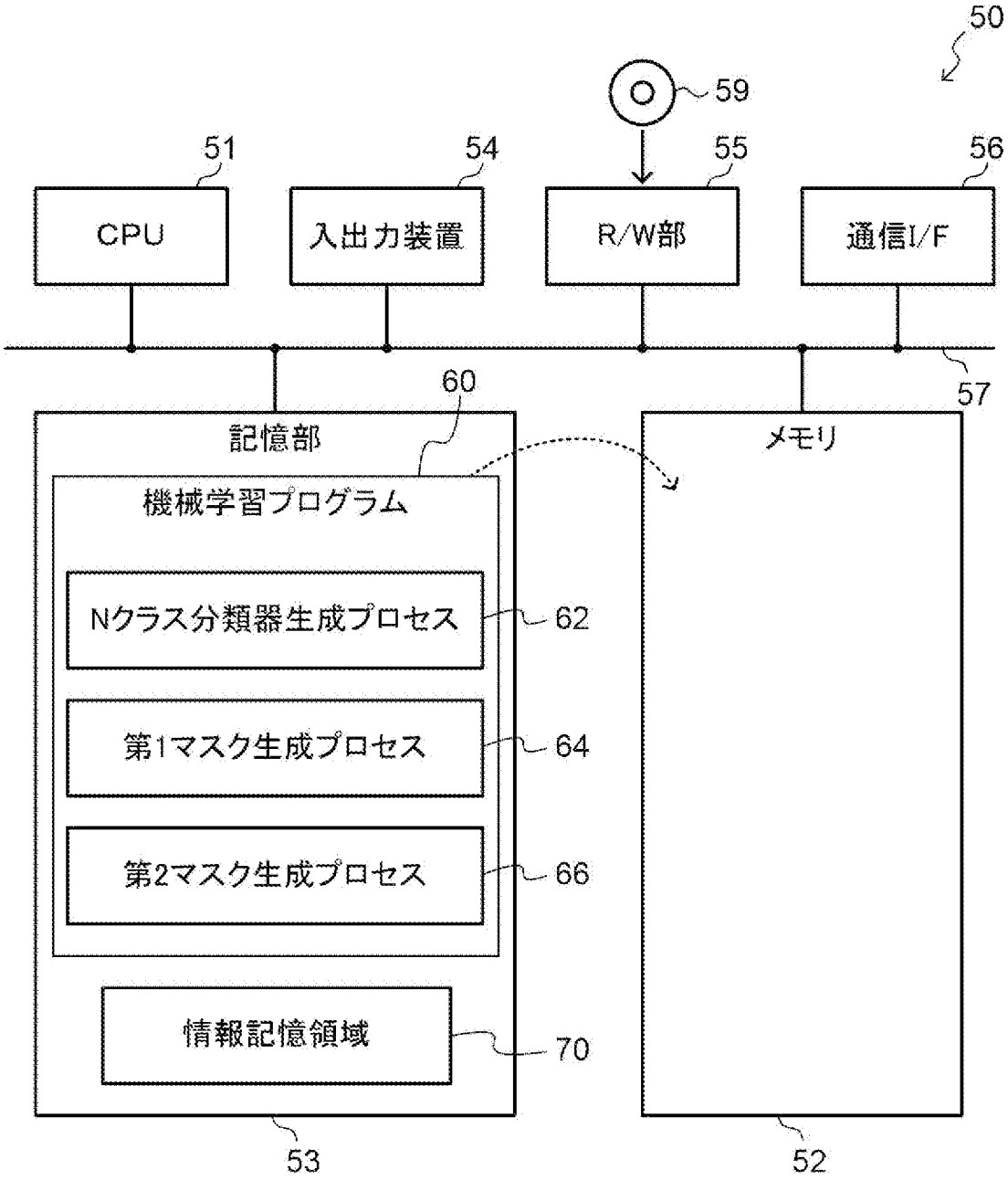
[図7]


22
24
Nクラス分類器
マスク集合
30
32
タスク情報
特定クラス分類器生成部
42
特定クラス分類器
34
運用データ
分類部
分類結果

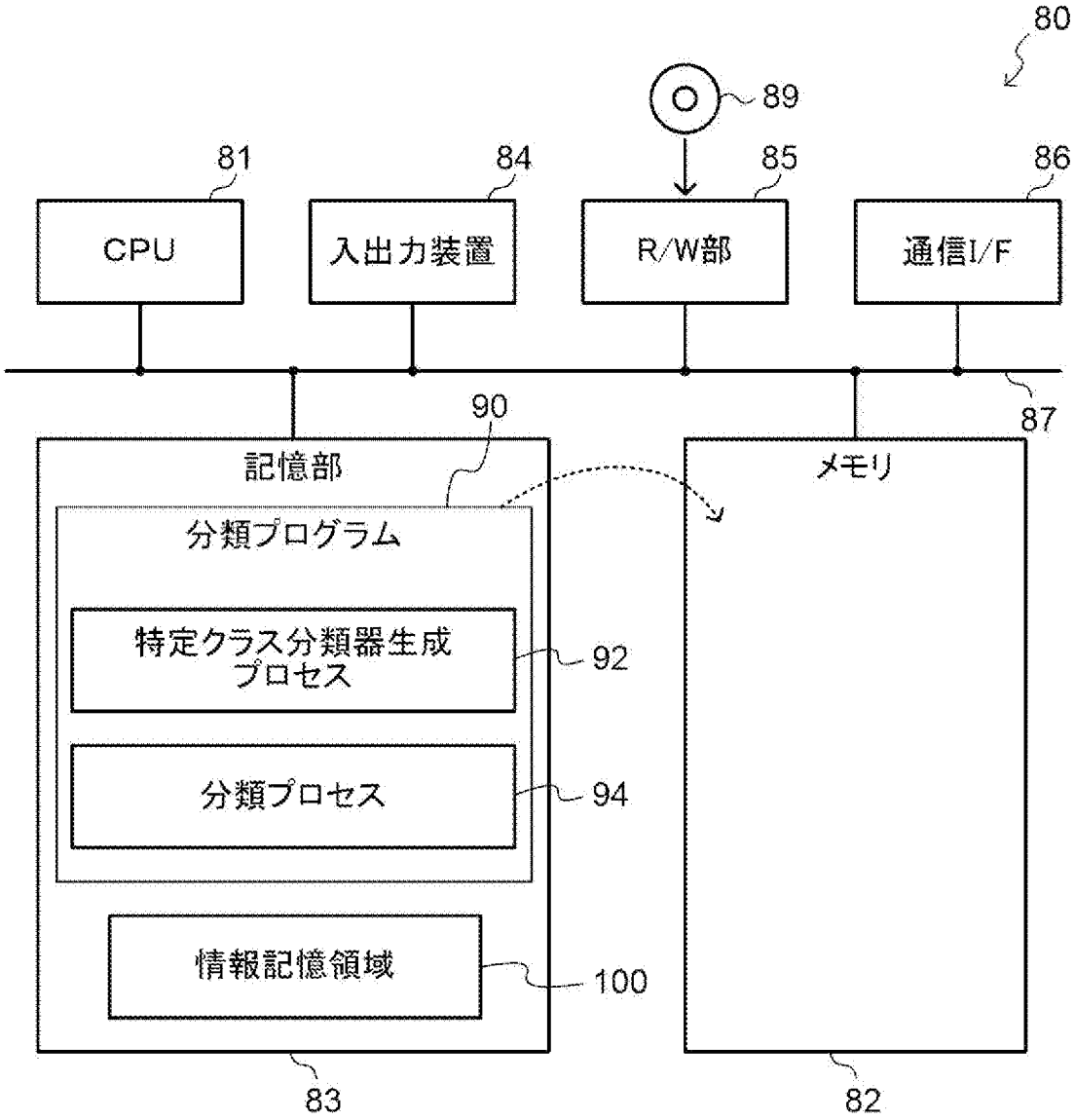
[図8]


訓練データ
機械学習
22
Nクラス分類器
(N=10)
枝刈り
類似
24
10個のモジュール
のマスク
モジュール0
マスク
モジュール1
マスク
モジュール2
マスク
モジュール9
マスク
タスクに応じた
モジュール3マ
スク及びモ
ジュール5マス
クを取得
42
特定クラス分類器
モジュール3
マスク
Nクラス分類器の
モジュール3+5の
和集合部分
モジュール5
マスク
運用データ
分類
分類
統合
最終
分類
結果

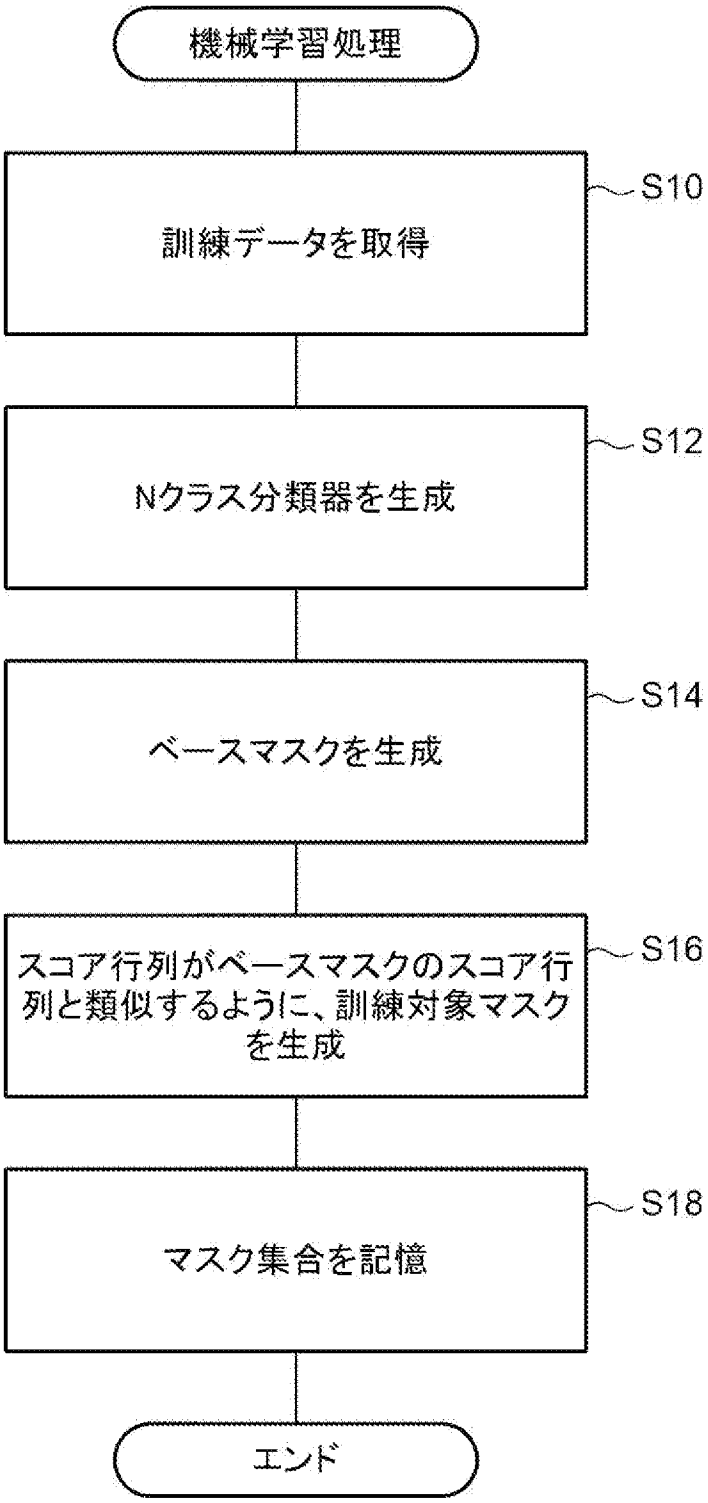
[図9]


50
59
51
54
55
56
CPU
入出力装置
R/W部
通信I/F
57
60
記憶部
メモリ
機械学習プログラム
Nクラス分類器生成プロセス
62
第1マスク生成プロセス
64
第2マスク生成プロセス
66
情報記憶領域
70
53
52

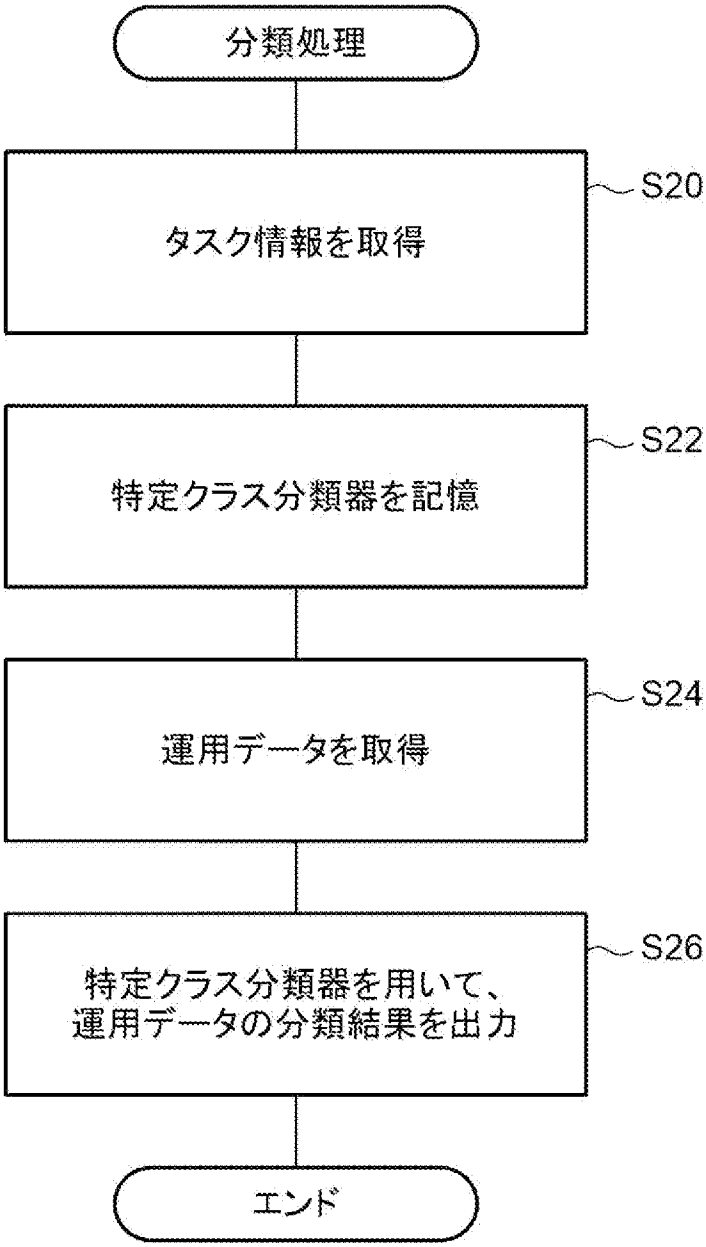
[図10]


80
89
81
84
85
86
CPU
入出力装置
R/W部
通信I/F
87
90
記憶部
メモリ
分類プログラム
特定クラス分類器生成
プロセス
92
分類プロセス
94
情報記憶領域
100
83
82

[図11]


機械学習処理
訓練データを取得
S10
Nクラス分類器を生成
S12
ベースマスクを生成
S14
スコア行列がベースマスクのスコア行列と類似するように、訓練対象マスクを生成
S16
マスク集合を記憶
S18
エンド

[図12]


分類処理
タスク情報を取得
S20
特定クラス分類器を記憶
S22
運用データを取得
S24
特定クラス分類器を用いて、
運用データの分類結果を出力
S26
エンド

[図13]

手法	スコア行列の初期値共有	損失関数	オリジナルのNNに対する結合時のパラメータ数の割合	共有パラメータの割合
比較例1	共有なし	CEのみ	0.0750	50%
比較例2	共有あり	CEのみ	0.0710	58%
本手法1	共有あり	CE+L1	0.0671	65.8%
本手法2	共有あり	CE+ λ * L1	0.0543	91.4%

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2021/019817

A. CLASSIFICATION OF SUBJECT MATTER
G06N 3/08(2006.01)i; G06N 20/00(2019.01)i
FI: G06N3/08 120; G06N20/00 130
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)
G06N3/00-99/00

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	1971-2021
Registered utility model specifications of Japan	1996-2021
Published registered utility model applications of Japan	1994-2021

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

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	HU, Xiaolong et al. DRNet: Dissect and Reconstruct the Convolutional Neural Network via Interpretable Manners. arXiv:1911.08691v2., 26 February 2020, version v2, [online]. arXiv (Cornell University)., pp. 1-8, [retrieved on 19 July 2021], Retrieved from the Internet: <URL:https://arxiv.org/abs/1911.08691v2>., abstract; sections 1.1, 3.1.	1-20

☐ 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
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
21 July 2021 (21.07.2021)	03 August 2021 (03.08.2021)
Name and mailing address of the ISA/ Japan Patent Office 3-4-3, Kasumigaseki, Chiyoda-ku, Tokyo 100-8915, Japan	Authorized officer Telephone No.

Form PCT/ISA/210 (second sheet) (January 2015)

国際調査報告

国際出願番号
PCT/JP2021/019817

A. 発明の属する分野の分類（国際特許分類（ＩＰＣ））

G06N 3/08(2006.01)i; G06N 20/00(2019.01)i
FI: G06N3/08 120; G06N20/00 130

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））

G06N3/00-99/00

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	１９７１－２０２１年
日本国実用新案登録公報	１９９６－２０２１年
日本国登録実用新案公報	１９９４－２０２１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）

IEEE Xplore

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	HU, Xiaolong, et al., DRNet: Dissect and Reconstruct the Convolutional Neural Network via Interpretable Manners, arXiv:1911.08691v2, 2020.02.26, version v2, [online], arXiv (Cornell University), Pages 1-8, [retrieved on 2021.07.19], Retrieved from the Internet: <URL: https://arxiv.org/abs/1911.08691v2>. Abstract; Sections 1.1 and 3.1.	1-20

☐ 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	国際調査報告の発送日
21.07.2021	03.08.2021

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）
久保　光宏 5C 9189
電話番号 03-3581-1101 内線 3539