(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10月13日(13.10.2022)

WIPO | PCT

(10) 国際公開番号
WO 2022/215559 A1

(51) 国際特許分類:
G06T 7/00 (2017.01) *G06N 20/00* (2019.01)

(21) 国際出願番号： PCT/JP2022/014692

(22) 国際出願日： 2022年3月25日(25.03.2022)

(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：
63/171,016 2021年4月5日(05.04.2021) US

(71) 出願人：パナソニックＩＰマネジメント株式会社(**PANASONIC INTELLECTUAL PROPERTY MANAGEMENT CO., LTD.**) [JP/JP]；〒5406207 大阪府大阪市中央区城見２丁目１番６１号 Osaka (JP).

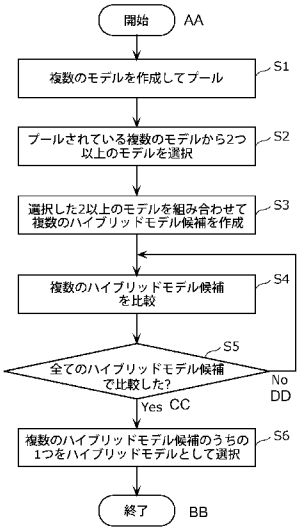
(72) 発明者：ズウオウ ヤオ(**ZHOU, Yao**). マテェウ アテュル(**MATHEW, Athul M.**). ベック アリエル(**BECK, Ariel**). ウィジャヤ チャンドラ スワンディ(**WIJAYA, Chandra Suwandi**). アウン グンウェイ ンウェイ(**AUNG, Nway Nway**). ケック カイ ジュン(**KEK, Khai Jun**). 菅澤 裕也(**SUGASAWA, Yuya**). フェルナンド ジェッフリー(**FERNANDO, Jeffry**). 佐藤 吉宣(**SATOU, Yoshinori**). 村田 久治(**MURATA, Hisaji**).

(74) 代理人：新 居 広 守，外 (**NII, Hiromori** et al.)；〒5320011 大阪府大阪市淀川区西中島５丁目３番１０号タナカ・イトーピア新大阪ビル６階新居国際特許事務所内 Osaka (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, JM, 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,

(54) **Title:** HYBRID MODEL CREATION METHOD, HYBRID MODEL CREATION DEVICE, AND PROGRAM

(54) 発明の名称： ハイブリッドモデル作成方法、ハイブリッドモデル作成装置、及び、プログラム



S1 Create and pool plurality of models
S2 Select at least two models from among plurality of pooled models
S3 Create plurality of hybrid model candidates by combining at least two selected models
S4 Make comparison between plurality of hybrid model candidates
S5 Has comparison made between all hybrid model candidates?
S6 Select one of plurality of hybrid model candidates as hybrid model
AA Start
BB End
CC Yes
DD No

(57) **Abstract:** The present invention involves, first, pooling a plurality of models for inferring a category of data being inputted (S1). At least one of the plurality of models is a model obtained by machine learning. The present invention next involves: creating a plurality of hybrid model candidates for determining a category by selecting and combining at least two models from among the plurality of pooled models (S2); making a comparison between the plurality of hybrid model candidates (S4); and selecting, as a hybrid model, one of the plurality of hybrid model candidates (S6).

(57) 要約：まず、入力されるデータのカテゴリを推定する複数のモデルをプールし（Ｓ１）、複数のモデルの少なくとも一つのモデルは、機械学習されたモデルである。次に、プールされている複数のモデルから２つ以上のモデルを選択して組み合わせることで、カテゴリを判定するハイブリッドモデル候補を複数作成し（Ｓ２）、複数のハイブリッドモデル候補を比較することで（Ｓ４）、複数のハイブリッドモデル候補のうちの１つをハイブリッドモデルとして選択する（Ｓ６）。

WO 2022/215559 A1

［続葉有］

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, 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 ：

ハイブリッドモデル作成方法、ハイブリッドモデル作成装置、及び、プログラム

技術分野

[0001]　本開示は、ハイブリッドモデル作成方法、ハイブリッドモデル作成装置、及び、プログラムに関する。

背景技術

[0002]　ＡＩ技術を使用した外観検査システムが一般的になりつつある。ＡＩモデルの種類によって利点と不利点とが異なることから、複数のＡＩモデルを組み合わせてそれぞれの利点を相補的に取得することで精度を高める技術が提案されている（例えば特許文献１参照）。特許文献１には、装置が有する複数のモデルすべてを使用して得た結果を統合することで、最終判定結果を得ることが開示されている。

先行技術文献

特許文献

[0003] 特許文献1：国際公開第２０１８／０７９８４０号

発明の概要

発明が解決しようとする課題

[0004]　しかしながら、上記特許文献１に開示される技術では、装置が有する複数のモデルすべてを使用するので、他のモデルと相補的でない冗長なモデルが組み合わされて使用されてしまうという課題がある。

[0005]　本開示は、上述の事情を鑑みてなされたもので、より精度が高いハイブリッドモデルを作成することができるハイブリッドモデル作成方法等を提供することを目的とする。

課題を解決するための手段

[0006] 上記目的を達成するために、本開示の一形態に係るハイブリッドモデル作成方法は、入力されるデータのカテゴリを推定する複数のモデルをプールし、前記複数のモデルの少なくとも一つのモデルは、機械学習されたモデルであり、プールされている複数のモデルから２つ以上のモデルを選択して組み合わせることで、前記カテゴリを判定するハイブリッドモデル候補を複数作成し、複数の前記ハイブリッドモデル候補を比較することで、前記複数のハイブリッドモデル候補のうちの１つをハイブリッドモデルとして選択する。

[0007] これにより、複数のモデルを用いてより精度が高いハイブリッドモデルを作成することができる。

[0008] なお、これらの全般的または具体的な態様は、装置、方法、集積回路、コンピュータプログラムまたはコンピュータで読み取り可能なＣＤ－ＲＯＭなどの記録媒体で実現されてもよく、システム、方法、集積回路、コンピュータプログラム及び記録媒体の任意な組み合わせで実現されてもよい。

発明の効果

[0009] 本開示により、複数のモデルを用いてより精度が高いハイブリッドモデルを作成することができるハイブリッドモデル作成方法などを提供できる。

図面の簡単な説明

[0010] [図1]図１は、実施の形態に係るハイブリッドモデル作成装置の機能構成を示すブロック図である。

[図2]図２は、実施の形態に係るハイブリッドモデル作成方法が実行される際の処理を概念的に説明するための図である。

[図3]図３は、実施の形態に係るハイブリッドモデル作成装置の動作概要を示すフローチャートである。

[図4]図４は、実施例１に係るステップＳ１の詳細処理の一例を示すフローチャートである。

[図5]図５は、実施例２に係るステップＳ１の詳細処理の一例を示すフローチャートである。

[図6]図６は、実施例３に係るステップＳ３の詳細処理の一例を示すフローチ

ャートである。

[図7A]図７Ａは、実施例３に係る相関の強い３つのモデルを組み合わせたときのハイブリッドモデル候補の精度を説明するための図である。

[図7B]図７Ｂは、実施例３に係る相関の弱い３つのモデルを組み合わせたときのハイブリッドモデル候補の精度を説明するための図である。

[図8]図８は、実施例４に係るハイブリッドモデル候補作成処理の詳細の一例を概念的に説明するための図である。

[図9]図９は、実施例４に係るハイブリッドモデル作成装置の処理の一例を示すフローチャートである。

[図10]図１０は、実施例５に係るステップＳ３の詳細処理の一例を示すフローチャートである。

[図11]図１１は、実施例６に係るモデル１とモデル２とで組み合わせて作成されるハイブリッドモデル候補の一例を概念的に示す図である。

[図12]図１２は、実施例６に係るステップＳ３の詳細処理の一例を示すフローチャートである。

[図13]図１３は、実施例７に係るモデル１とモデル２との出力と不良品画像に対応する出力の分布の凸包とを概念的に示す図である。

[図14]図１４は、図１３に示す凸包の頂点を除く不良品画像に対応する出力を除去したモデル１とモデル２との出力から作成されるハイブリッドモデル候補の一例を概念的に示す図である。

[図15]図１５は、実施例７に係るステップＳ３の詳細処理の一例を示すフローチャートである。

[図16]図１６は、実施例７に係るモデル１とモデル２との出力と除外領域とを概念的に示す図である。

[図17]図１７は、図１６に示す除外領域に含まれる不良品画像に対応する出力を除去したモデル１とモデル２との出力から作成されるハイブリッドモデル候補の一例を概念的に示す図である。

[図18]図１８は、実施例８に係るＦＡＲ曲線をモデル１に対して算出する方

法を説明するための図である。

[図19]図１９は、実施例８に係るモデル１のＦＡＲ表の一例を示す図である。

[図20]図２０は、実施例８に係る２つのモデルそれぞれの第１ＦＡＲ値と、２つのモデルを組み合わせて作成されるハイブリッドモデル候補の第２ＦＡＲ値とを概念的に示す図である。

[図21]図２１は、その他の実施の形態に係るハイブリッドモデル作成方法の一例を示す図である。

[図22]図２２は、その他の実施の形態に係るハイブリッドモデル作成方法の他の一例を示す図である。

[図23A]図２３Ａは、その他の実施の形態に係る混同行列の表の一例を示す図である。

[図23B]図２３Ｂは、その他の実施の形態に係る混同行列の表の一例を示す図である。

発明を実施するための形態

[0011] 以下、本開示の実施の形態について、図面を用いて詳細に説明する。なお、以下で説明する実施の形態は、いずれも本開示の一具体例を示す。以下の実施の形態で示される数値、形状、材料、規格、構成要素、構成要素の配置位置及び接続形態、ステップ、ステップの順序等は、一例であり、本開示を限定する主旨ではない。また、以下の実施の形態における構成要素のうち、本開示の独立請求項に記載されていない構成要素については、任意の構成要素として説明される。また、各図は、必ずしも厳密に図示したものではない。各図において、実質的に同一の構成については同一の符号を付し、重複する説明は省略又は簡略化する場合がある。

[0012] （実施の形態）

まず、本実施の形態に係るハイブリッドモデル作成装置及びハイブリッドモデル作成方法の概要について説明する。

[0013] ［１．ハイブリッドモデル作成装置１０の概要］

以下、本実施の形態に係るハイブリッドモデル作成装置１０の構成等の概要について説明する。

[0014] 図１は、本実施の形態に係るハイブリッドモデル作成装置１０の機能構成を示すブロック図である。図２は、本実施の形態に係るハイブリッドモデル作成方法が実行される際の処理を概念的に説明するための図である。

[0015] ハイブリッドモデル作成装置１０は、コンピュータ等で実現され、複数のモデルを用いて、より精度が高いハイブリッドモデルを作成することができる装置である。

[0016] 本実施の形態では、図１に示されるように、ハイブリッドモデル作成装置１０は、モデルプール部１１と、モデル選択部１２と、ハイブリッドモデル候補作成部１３と、ハイブリッドモデル選択部１４と、判定閾値決定部１５とを備える。なお、判定閾値決定部１５は、ハイブリッドモデル作成装置１０とは、別の装置に備えられてもよい。

[0017] ［１－１．モデルプール部１１］

モデルプール部１１は、ＨＤＤ（Hard Disk Drive）またはメモリ等で構成され、入力されるデータのカテゴリを推定する複数のモデルをプール（記憶）している。本実施の形態では、モデルプール部１１は、図２に示すように、例えばモデル１、モデル２、モデル３及びモデル４などの予め作成された複数のモデル１１ａをプールしている。ここで、複数のモデル１１ａの少なくとも一つのモデルは、機械学習されたモデルである。複数のモデル１１ａのそれぞれは、ＡＩモデルとも称することができる。本実施の形態では、入力されるデータは、製造品の検査画像であるとして説明する。複数のモデル１１ａのうちの少なくとも１つのモデルは、深層学習により学習されたＡＩモデルである。複数のモデル１１ａには、人手により特徴量が設計されたＡＩモデルが含まれていてもよい。例えば、複数のモデル１１ａのそれぞれは、製造品の検査画像を入力とし、検査画像に映る製造品が不良である確率を推定して出力する。なお、複数のモデル１１ａのそれぞれは、検査画像に映る製造品が不良であるか否かの２値の推定結果を出力してもよい。

[0018] ［１－２．モデル選択部１２］

モデル選択部１２は、モデルプール部１１にプールされている複数のモデルから２つ以上のモデルを選択する。本実施の形態では、モデル選択部１２は、モデルプール部１１にプールされている複数のモデルのうち所定のモデルを除外した上で２つ以上のモデルを選択する。図２に示す例では、モデル選択部１２は、例えばモデル１、モデル２、モデル３及びモデル４のうち、モデル４を所定のモデルとして除外した上で、２つ以上のモデルを選択するモデル選択処理１２ａを行う。モデル選択部１２は、モデルプール部１１にプールされている複数のモデルから所定のモデルを除外してから、２つ以上のモデルを選択してもよいし、モデルプール部１１にプールされている複数のモデルのうち、２つ以上のモデルを所定のモデルを除外した上で選択してもよい。また、所定のモデルは、例えば推定精度の低いモデルであってもよいし、他のモデルとの相関が強いモデルであってもよい。このような所定のモデルを除外する方法等の詳細は、後述する実施例１及び実施例２で説明するのでここでの説明は省略する。

[0019] ［１－３．ハイブリッドモデル候補作成部１３］

ハイブリッドモデル候補作成部１３は、モデル選択部１２により選択された２つ以上のモデルを組み合わせることで、カテゴリを判定するハイブリッドモデル候補を複数作成する。なお、ハイブリッドモデル候補作成部１３は、閾値より強い相関があるモデルの組み合わせを含めないように、モデル選択部１２により選択された２つ以上のモデルを組み合わせることで、ハイブリッドモデル候補を複数作成してもよい。ハイブリッドモデル候補は、モデル選択部１２により選択された２つ以上のモデルを単純に連結（カスケード）することで組み合わせてもよいし、後述するようにロジスティック回帰などを用いて組み合わせてもよい。

[0020] 図２に示す例では、ハイブリッドモデル候補作成部１３は、モデル選択部１２により選択されたモデル１、モデル２及びモデル３を組み合わせることでハイブリッドモデル候補を作成するハイブリッドモデル候補作成処理１３

ａを行う。より具体的には、ハイブリッドモデル候補作成部１３は、例えばモデル１とモデル２とを組み合わせたハイブリッドモデル候補１と、例えばモデル２とモデル３とを組み合わせたハイブリッドモデル候補２を作成する。また、ハイブリッドモデル候補作成部１３は、例えばモデル１とモデル３とを組み合わせたハイブリッドモデル候補３と、例えばモデル１とモデル２とモデル３とを組み合わせたハイブリッドモデル候補４を作成する。本実施の形態では、判定されるカテゴリとしては、検査画像に映る製造品が良品または不良品であるかである。つまり、ハイブリッドモデル候補３は、検査画像に映る製造品が良品または不良品であるかを判定する。なお、ハイブリッドモデル候補１～３は、検査画像に映る製造品が不良であるかを確率で判定（推定）した判定結果を出力してもよい。

[0021] ハイブリッドモデル候補を作成する方法等の詳細については、後述する実施例３～実施例６で説明するのでここでの説明は省略する。

[0022] また、ハイブリッドモデル候補作成部１３は、作成したハイブリッドモデル候補を比較する。

[0023] 図２に示す例では、ハイブリッドモデル候補作成部１３は、作成したハイブリッドモデル候補１～４を比較する比較処理を行う。ハイブリッドモデル候補１～４の比較方法としては、例えばハイブリッドモデル候補１～４のそれぞれの判定結果の精度を比較する方法、当該それぞれの判定結果から算出できる構成される２つ以上のモデルのそれぞれの重要度（寄与度とも称される）を比較する方法などが挙げられる。

[0024] なお、複数のハイブリッドモデル候補の比較方法等の詳細については、実施例２等で後述するのでここでの説明は省略する。

[0025] ［１－４．ハイブリッドモデル選択部１４］

ハイブリッドモデル選択部１４は、複数のハイブリッドモデル候補の比較結果に基づき、複数のハイブリッドモデル候補のうちの１つをハイブリッドモデルとして選択する。

[0026] 図２に示す例では、ハイブリッドモデル選択部１４は、ハイブリッドモデ

ル候補１～４の比較結果から、ハイブリッドモデル候補１～４のうちの１つをハイブリッドモデルとして選択するハイブリッドモデル選択処理１４ａを行う。

[0027] ハイブリッドモデル選択処理１４ａでは、ハイブリッドモデル候補１～４の比較結果に基づき、判定結果の精度のうちで一番精度の高い、または、重要度が高いモデルの組み合わせからなるハイブリッドモデル候補が、ハイブリッドモデルとして選択される。

[0028] なお、ハイブリッドモデルの選択方法の詳細については、後述するのでここでの説明は省略する。

[0029] ［１－５．判定閾値決定部１５］

判定閾値決定部１５は、例えば製造品の検査画像などの検証用データセットを用いて、ハイブリッドモデル選択部１４により選択されたハイブリッドモデルの感度を調整し、誤判定を抑制するために許容できる過検出率の閾値を決定する。判定閾値決定部１５は、例えば製造品の検査画像などの検証用データセットを入力して当該製造品が良品または不良品であるかを判定させた判定結果を取得する。判定閾値決定部１５は、取得した判定結果から混同行列を生成し、誤判定を抑制するために許容できる過検出率の閾値（判定閾値）を決定する。なお、図２に示す判定閾値決定処理１５ａにおいて示されるCascading Modelは、ハイブリッドモデル選択部１４により選択されたハイブリッドモデルを意味し、判定閾値が最適化されている。

[0030] ［２．ハイブリッドモデル作成装置１０の動作概要］

以上のように構成されたハイブリッドモデル作成装置１０の動作概要について以下説明する。

[0031] 図３は、本実施の形態に係るハイブリッドモデル作成装置１０の動作概要を示すフローチャートである。

[0032] まず、ハイブリッドモデル作成装置１０は、入力されるデータのカテゴリを推定する複数のモデルをプールする（Ｓ１）。本実施の形態では、複数のモデルの少なくとも一つのモデルは、機械学習されたモデルである。また、

例えば、複数のモデルのそれぞれは、製造品の検査画像を入力とし、検査画像に映る製造品が不良である確率を推定して出力する。

[0033] 次に、ハイブリッドモデル作成装置１０は、プールされている複数のモデルから２つ以上のモデルを選択する（Ｓ２）。本実施の形態では、ハイブリッドモデル作成装置１０は、プールされている複数のモデルから、一部（所定のモデル）を除いて２つ以上のモデルを選択する。

[0034] 次に、ハイブリッドモデル作成装置１０は、ステップＳ２で選択された２つ以上のモデルを組み合わせることで、カテゴリを判定するハイブリッドモデル候補を複数作成する（Ｓ３）。本実施の形態では、ハイブリッドモデル作成装置１０は、ステップＳ２で選択された２つ以上のモデルをシーケンシャルにカスケードして組み合わせてもよいし、ロジスティック回帰を用いて組み合わせてもよい。

[0035] 次に、ハイブリッドモデル作成装置１０は、ステップＳ３で作成した複数のハイブリッドモデル候補を比較する（Ｓ４）。本実施の形態では、ハイブリッドモデル作成装置１０は、例えばハイブリッドモデル候補のそれぞれの判定結果の精度を比較したり、ハイブリッドモデル候補のそれぞれの判定結果から算出できる構成される２つ以上のモデルのそれぞれの重要度を比較したりすることができる。

[0036] 次に、ハイブリッドモデル作成装置１０は、全てのハイブリッドモデル候補で比較したかを判定する（Ｓ５）。ステップＳ５において、全てのハイブリッドモデル候補で比較していない場合（Ｓ５でＮｏ）、ステップＳ４に戻る。

[0037] 一方、ステップＳ５において、全てのハイブリッドモデル候補で比較済みの場合（Ｓ５でＹｅｓ）、複数のハイブリッドモデル候補のうちの１つをハイブリッドモデルとして選択する（Ｓ６）。本実施の形態では、ハイブリッドモデル作成装置１０は、ハイブリッドモデル候補のそれぞれの判定結果のうちで一番精度の高い、または、重要度が高いモデルの組み合わせからなるハイブリッドモデル候補をハイブリッドモデルとして選択することができる

。

[0038] このように、本実施の形態のハイブリッドモデル作成方法によれば、プールされている複数のモデルの全部を用いずに、複数のハイブリッドモデル候補を作成し、例えば判定精度などを用いて複数のハイブリッドモデル候補を比較する。これにより、例えば判定結果のうちで一番精度の高いハイブリッドモデル候補をハイブリッドモデルとして選択することができる。つまり、複数のモデルを用いてより精度が高いハイブリッドモデルを作成することができる。

[0039] （実施例１）

図３に示すステップＳ１において、プールされている複数のモデルから、推定精度の低いモデルを所定のモデルとして除いてもよい。すなわち、プールされている複数のモデルのうち、推定精度の低いモデルをハイブリッドモデル候補から除外してもよい。以下、この場合の具体例を実施例１として説明する。なお、推定精度は、正解率に限らず、適合率、再現率、適合率及び再現率の調和平均により算出されるＦ値、ＲＯＣ（Receiver Operating Characteristic）曲線のＡＵＣ(Area Under Curve)並びに、正解率のうちの少なくとも一の組み合わせであればよい。

[0040] 図４は、実施例１に係るステップＳ１の詳細処理の一例を示すフローチャートである。

[0041] ステップＳ１において、まず、ハイブリッドモデル作成装置１０は、入力されるデータのカテゴリを推定する複数のモデルをプールする（Ｓ１１１）。

[0042] 次に、ハイブリッドモデル作成装置１０は、検証用データセットを用いて、複数のモデルそれぞれの推定精度を取得する（Ｓ１１２）。より具体的には、モデル選択部１２は、２つ以上のモデルを選択する前に、モデルプール部１１にプールされている複数のモデルそれぞれに、複数の検証用データセットを入力してカテゴリを推定させることで当該複数のモデルそれぞれの推定精度を取得する。

[0043] なお、プールされている複数のモデルそれぞれの推定精度は、予め用意された検証用データセットをすべて用いて算出してもよいが、これに限らない。すべての検証用データセットのうち、モデルによって推定結果が異なる検証用データセットを用いてもよい。例えば、プールされている複数のモデルがモデル１、モデル２、モデル３及びモデル４である場合、モデル１の推定結果と、モデル２、モデル３及びモデル４の推定結果とが異なる検証用データセットを用いる。

[0044] 次に、ハイブリッドモデル作成装置１０は、推定精度が閾値以下のモデルを除外する（Ｓ１１３）。より具体的には、モデル選択部１２は、推定精度が閾値以下のモデルを、モデルプール部１１にプールされている複数のモデルの中から除外する。そして、モデル選択部１２は、閾値以下のモデルが除外された複数のモデルから、２つ以上のモデルを選択する。なお、閾値は、事前にユーザにより設定される。

[0045] 例えば、プールされている複数のモデルがモデル１、モデル２、モデル３及びモデル４であり、モデル４の推定精度のみが閾値以下の場合、モデル選択部１２は、モデルプール部１１にプールされているモデル１～４からモデル４を除外する。そして、モデル選択部１２は、モデルプール部１１にプールされているモデル１、モデル２及びモデル３から、２つ以上のモデルを選択する。

[0046] このようにして、ハイブリッドモデル作成装置１０は、プールされている複数のモデルのうち、推定精度が閾値以下のモデルをハイブリッドモデル候補から除外することができる。

[0047] （実施例２）

図３に示すステップＳ１において、プールされている複数のモデルから、他のすべてのモデルとの相関が強いモデルを所定のモデルとして除いてもよい。すなわち、プールされている複数のモデルのうち、他のすべてのモデルとの相関が強いモデルをハイブリッドモデル候補から除外してもよい。以下、この場合の具体例を実施例２として説明する。

[0048] 図５は、実施例２に係るステップＳ１の詳細処理の一例を示すフローチャートである。

[0049] ステップＳ１において、まず、ハイブリッドモデル作成装置１０は、入力されるデータのカテゴリを推定する複数のモデルをプールする（Ｓ１２１）。

[0050] 次に、ハイブリッドモデル作成装置１０は、検証用データセットを用いて、複数のモデルそれぞれの推定結果を取得する（Ｓ１２２）。より具体的には、モデル選択部１２は、２つ以上のモデルを選択する前に、モデルプール部１１にプールされている複数のモデルそれぞれに、複数の検証用データセットを入力してカテゴリを推定させることで当該複数のモデルそれぞれの推定結果を取得する。ここで、推定結果は、モデルの最終出力結果でもよいし、モデルの中間量であってもよい。例えば、深層学習されたモデルでは、推定結果は、深層学習されたモデルの中間層または最終層の出力結果である。

[0051] 次に、ハイブリッドモデル作成装置１０は、ステップＳ１２２で取得した推定結果を用いて、プールされている複数のモデルすべての相関を算出する（Ｓ１２３）。より具体的には、モデル選択部１２は、モデルプール部１１にプールされている複数のモデルのすべてについて２つのモデルの相関を算出する。

[0052] ここで、相関の算出方法について説明する。

[0053] 検証用データセットに対するｊ（ｊは自然数）番目のモデルの推定結果をc_jとする。例えば検証用データセットのうちのｉ（ｉは自然数）番目の検証用データに対する推定結果を$c_{j,i}$とする。また、推定結果は、モデルの最終出力結果またはスカラーの中間量であるとする。

[0054] この場合、ｊ番目とｋ（ｋは自然数、かつｊ≠ｋ）番目のモデルとの相関は、（式１）または、（式２）、（式３）もしくは（式４）を用いて算出することができる。なお、（式１）は、推定結果の一致率（Jcacard係数）を算出する式であり、推定結果が０または１の２値の場合に用いることができる。（式１）においてδは、クロネッカーのδである。

[0055] [数1]

$$\frac{1}{N}\sum_{i=1}^{N}\delta_{c_{j,i},c_{k,i}} \qquad (式１)$$

[0056] [数2]

$$cov(c_j, c_k) = E[c_j c_k] - E[c_j]E[c_k] \qquad (式２)$$

[0057] [数3]

$$\rho(c_j, c_k) = \frac{cov(c_j,c_k)}{\sqrt{V[c_j]V[c_k]}} \qquad (式３)$$

[0058] [数4]

$$cos(c_j, c_k) = \frac{\boldsymbol{c_j}\cdot\boldsymbol{c_k}}{|\boldsymbol{c_j}||\boldsymbol{c_k}|} \qquad (式４)$$

[0059] 一方で、（式２）～（式４）は、推定結果が２値である場合に限らず連続値の場合にも用いることができる。（式２）は、共分散を算出する式であり、Ｅ［Ｘ］はＸの平均を示す。（式３）のＶ［Ｘ］はＸの分散を示す。また、（式３）は相関係数を算出する式であり、（式４）はコサイン類似度を算出する式であり、c_jは$c_{j,k}$をｉに対して並べて作ったベクトルである。

[0060] 続いて、推定結果がベクトルの中間量である場合の相関の算出方法について説明する。

[0061] この場合、ｊ番目とｋ番目のモデルとの相関は、（式５）または、（式６）を用いて、検証用データごとの中間量類似度sim_iを算出することができる。なお、$f_{j,i}$は、複数値からなるベクトルの中間量である。そして、中央値または（式７）で示される平均値などの統計量を算出する。これにより、推定結果がベクトルの中間量であっても算出した相関を比較することができる。

[0062]

[数5]

$$\mathrm{sim_i} = \frac{\mathbf{f_{j,i}} \cdot \mathbf{f_{k,i}}}{|\mathbf{f_{j,i}}||\mathbf{f_{k,i}}|} \quad (式５)$$

[0063] [数6]

$$\mathrm{sim_i} = ||\mathbf{f_{j,i}} - \mathbf{f_{k,i}}|| \quad (式６)$$

[0064] [数7]

$$\mathrm{average(sim)} = \frac{1}{N}\sum_{i=1}^{N} \mathrm{sim_i} \quad (式７)$$

[0065] 以下、図５に戻って説明を続ける。

[0066] 次に、ハイブリッドモデル作成装置１０は、ステップＳ１２３で算出された相関に基づき、他のすべてのモデルとの相関が閾値より強いモデルを除外する（Ｓ１２４）。より具体的には、モデル選択部１２は、他のすべてのモデルとの相関係数の平均または中央値が閾値より強いモデルを、モデルプール部１１にプールされている複数のモデルの中から除外する。そして、モデル選択部１２は、閾値以下のモデルが除外された複数のモデルから、２つ以上のモデルを選択する。なお、閾値は、事前にユーザにより設定される。

[0067] 例えば、プールされている複数のモデルがモデル１、モデル２、モデル３及びモデル４であり、モデル４と他のモデル１、２または３の相関が閾値より強い場合、モデル選択部１２は、モデルプール部１１にプールされているモデル１～４からモデル４を除外する。そして、モデル選択部１２は、モデルプール部１１にプールされているモデル１、モデル２及びモデル３から、２つ以上のモデルを選択する。

[0068] このようにして、ハイブリッドモデル作成装置１０は、プールされている複数のモデルのうち、他のすべてのモデルとの相関が閾値より強いモデルをハイブリッドモデル候補から除外することができる。

[0069] （実施例３）

実施例２では、図３に示すステップＳ１において、プールされている複数のモデルから、他のすべてのモデルとの相関が強いモデルを所定のモデルとして除いた場合について説明したが、これに限らない。図３に示すステップ３において、相関の強いモデルの組み合わせを含めないようにしてハイブリッドモデル候補を作成してもよい。以下、この場合の具体例を実施例３として説明する。

[0070] 図６は、実施例３に係るステップＳ３の詳細処理の一例を示すフローチャートである。

[0071] ステップＳ３において、まず、ハイブリッドモデル作成装置１０は、検証用データセットを用いて、複数のモデルそれぞれの推定結果を取得する（Ｓ３１１）。より具体的には、ハイブリッドモデル候補作成部１３は、ハイブリッドモデル候補を複数作成する前に、モデルプール部１１にプールされている複数のモデルそれぞれに、複数の検証用データセットを入力してカテゴリを推定させることで当該複数のモデルそれぞれの推定結果を取得する。なお、ハイブリッドモデル候補作成部１３は、モデル選択部１２に選択された複数のモデルそれぞれに、複数の検証用データセットを入力してカテゴリを推定させることで当該複数のモデルそれぞれの推定結果を取得してもよい。ここで、推定結果は、実施例２で説明したのと同様に、モデルの最終出力結果でもよいし、モデルの中間量であってもよい。例えば、深層学習されたモデルでは、推定結果は、深層学習されたモデルの中間層または最終層の出力結果である。

[0072] 次に、ハイブリッドモデル作成装置１０は、ステップＳ３１１で取得した推定結果を用いて、プールされているまたは選択された複数のモデルすべての相関を算出する（Ｓ３１２）。より具体的には、ハイブリッドモデル候補作成部１３は、モデルプール部１１にプールされているまたはモデル選択部１２に選択された複数のモデルのすべてについて２つのモデルの相関を算出する。なお、相関の算出方法については実施例２で説明したのでここでの説

明を省略する。

[0073] 次に、ハイブリッドモデル作成装置１０は、プールされている複数のモデルから、閾値より強い相関がある２つのモデルの組み合わせを含めないように、２つ以上のモデルを選択する（Ｓ３１３）。より具体的には、ハイブリッドモデル候補作成部１３は、閾値より強い相関がある２つのモデルの組み合わせを含めないように、モデル選択部１２により選択された２つ以上のモデルを組み合わせることで、ハイブリッドモデル候補を複数作成する。

[0074] このようにして、ハイブリッドモデル作成装置１０は、選択された複数のモデルから、相関が弱いモデルを組み合わせたハイブリッドモデル候補を作成することができる。

[0075] ここで、相関が弱いモデルを組み合わせたハイブリッドモデル候補を作成する理由について説明する。

[0076] 図７Ａは、実施例３に係る相関の強い３つのモデルを組み合わせたときのハイブリッドモデル候補の精度を説明するための図である。図７Ｂは、実施例３に係る相関の弱い３つのモデルを組み合わせたときのハイブリッドモデル候補の精度を説明するための図である。図７Ａ及び図７Ｂに示されるハイブリッドモデル候補は、ロジスティック回帰などを用いて３つのモデルの推定結果を組み合わせる。なお、説明を簡単にするため、図７Ａ及び図７Ｂにおいてハイブリッドモデル候補は、３つのモデルの推定結果の多数決を出力するものとして説明する。

[0077] 図７Ａには、相関の強いモデル１、モデル２及びモデル３それぞれと、ハイブリッドモデル候補とに、検証用データセットのうちの１０個の検証用データを用いたときの２値の推定結果及び判定結果と、１０個の検証用データの真の値とが示されている。図７Ａに示されているように、モデル１、モデル２及びモデル３の精度（推定精度）は、８０％、７０％及び８０％であり、モデル１、モデル２及びモデル３を組み合わせたハイブリッドモデル候補の精度（判定精度）は、８０％となっている。

[0078] 図７Ｂには、相関の弱いモデル１、モデル２及びモデル３それぞれと、ハ

イブリッドモデル候補とに、検証用データセットのうちの１０個の検証用データを用いたときの２値の推定結果及び判定結果と、１０個の検証用データの真の値とが示されている。図７Ｂに示されているように、モデル１、モデル２及びモデル３の精度（推定精度）は８０％、６０％及び５０％であり、モデル１、モデル２及びモデル３を組み合わせたハイブリッドモデル候補の精度（判定精度）は９０％となっている。

[0079] つまり、相関の強い３つのモデルを組み合わせてもハイブリッドモデル候補の精度は改善しない。一方、相関の弱い３つのモデルの精度が高くなくても、相関の弱い３つのモデルを組み合わせたハイブリッドモデル候補の精度は改善することが可能である。

[0080] 以上のように、実施例３によれば、ハイブリッドモデル作成装置１０は、相関が弱いモデルを組み合わせたハイブリッドモデル候補を作成することができる。そして、ハイブリッドモデル作成装置１０は、このようなハイブリッドモデル候補から１つのハイブリッドモデルを選べるので、より精度が高いハイブリッドモデルを作成することができる。

[0081] （実施例４）

実施例４では、ロジスティック回帰などを用いてハイブリッドモデル候補を作成する場合の具体例について説明する。

[0082] 本実施例では、ハイブリッドモデル候補作成部１３は、モデル選択部１２により選択された２つ以上のモデルを、ロジスティック回帰などを用いて組み合わせることで、カテゴリを判定するハイブリッドモデル候補を複数作成する。なお、組み合わせるモデルの数の最大数は、予め設定されているが、ハイブリッドモデル候補を作成する際に都度設定されてもよい。ハイブリッドモデル候補作成部１３は、複数のハイブリッドモデル候補のそれぞれを機械学習モデルとして作成する。機械学習モデルは、当該ハイブリッドモデル候補を構成するために選択された２つ以上のモデルのそれぞれに検証用データセットを入力してカテゴリを推定させることで得た２つ以上の出力結果を入力とし、検証用データセットのカテゴリを判定した判定結果を出力させる

モデルである。

[0083] また、ハイブリッドモデル候補作成部１３は、作成した複数のハイブリッドモデル候補に出力させた判定結果を比較する。より具体的には、ハイブリッドモデル候補作成部１３は、作成した複数のハイブリッドモデル候補を機械学習させた後に出力させた判定結果を比較する。

[0084] 図８は、実施例４に係るハイブリッドモデル候補作成処理１３ａの詳細の一例を概念的に説明するための図である。図８に示すハイブリッドモデル候補作成処理１３ａは、図２に示されるハイブリッドモデル候補作成処理１３ａの詳細の一例である。

[0085] 図８に示す例では、ハイブリッドモデル候補作成部１３は、モデル選択部１２により選択されたモデル１、モデル２及びモデル３を組み合わせることでハイブリッドモデル候補を作成するハイブリッドモデル候補作成処理１３ａを行う。より具体的には、ハイブリッドモデル候補作成部１３は、例えばモデル１とモデル２とをロジスティック回帰を用いて組み合わせた機械学習モデル１＆２（ハイブリッドモデル候補１）を作成する。また、ハイブリッドモデル候補作成部１３は、例えばモデル２とモデル３とをロジスティック回帰を用いて組み合わせた機械学習モデル２＆３（ハイブリッドモデル候補２）を作成する。また、ハイブリッドモデル候補作成部１３は、例えばモデル１とモデル３とをロジスティック回帰を用いて組み合わせた機械学習モデル１＆３（ハイブリッドモデル候補３）を作成する。なお、図８に示す例では、組み合わせるモデルの数の最大数は２であるとして、総当たりに組み合わせた機械学習モデルが作成されている。

[0086] 図８に示す例では、ハイブリッドモデル候補作成部１３は、機械学習モデル１＆２、機械学習モデル２＆３及び機械学習モデル１＆３を、検証用データセットを用いて学習させた後に得られる出力結果（判定結果）を取得する。ハイブリッドモデル候補作成部１３は、機械学習モデル１＆２、機械学習モデル２＆３及び機械学習モデル１＆３の出力結果を比較する比較処理を行う。ハイブリッドモデル候補作成部１３は、比較処理の結果、例えば精度の

高い順にランキングする。図８に示す例では、機械学習モデル２＆３、機械学習モデル１＆３及び機械学習モデル１＆２の順にランキングされている。

[0087] ここで、ロジスティック回帰を用いて複数のモデルを組み合わせる方法について説明する。

[0088] ロジスティック回帰を用いて組み合わせることで得られる機械学習モデルは、下記の（式８）で示されるようなロジスティック関数（シグモイド関数）を用いて表すことができる。なお、（式８）では、２つのモデルを組み合わせているが、３つ以上のモデルを組み合わせる場合も同様である。

[0089] [数8]

$$p=\frac{b^{\beta_0+\beta_1 x_1+\beta_2 x_2}}{b^{\beta_0+\beta_1 x_1+\beta_2 x_2}+1}=\frac{1}{1+b^{-(\beta_0+\beta_1 x_1+\beta_2 x_2)}}=S_b(\beta_0+\beta_1 x_1+\beta_2 x_2) \quad \text{（式８）}$$

[0090] （式８）において、関数S_b（$\beta_0+\beta_1 x_1+\beta_2 x_2$）は、０～１までの出力を有するシグモイド関数であり、β_0は定数であり、β_1及びβ_2はx_1及びx_2の係数である。また、x_1及びx_2は、２つのモデルの出力を示す。

[0091] 本実施例では、x_1及びx_2は、２つのモデルそれぞれを学習させた後に得られる出力（推定結果）に該当し、確率で表現される。関数S_b（$\beta_0+\beta_1 x_1+\beta_2 x_2$）の出力は、２つのモデルを組み合わせた機械学習モデルを、検証用データセットを用いて係数を学習させた後に得られる出力（判定結果）に該当し、０～１の確率で表現される。

[0092] 例えば、ロジスティック回帰を用いて組み合わせることで得られる機械学習モデル１＆２は、モデル１の出力及びモデル２の出力を入力として、検証用データセットを用いて係数を学習させたロジスティック関数を作用させて判定結果を出力するハイブリッドモデル候補である。同様に、ロジスティック回帰を用いて組み合わせることで得られる機械学習モデル２＆３は、モデル２の出力及びモデル３の出力を入力として、検証用データセットを用いて係数を学習させたロジスティック関数を作用させて判定結果を出力するハイブリッドモデル候補である。ロジスティック回帰を用いて組み合わせることで得られる機械学習モデル１＆３は、モデル１の出力及びモデル３の出力を

入力として、検証用データセットを用いて係数を学習させたロジスティック関数を作用させて判定結果を出力するハイブリッドモデル候補である。

[0093] なお、複数のモデルを組み合わせる方法は、ロジスティック回帰を用いる方法に限らない。複数のモデルそれぞれを学習させた後に得られる出力（推定結果）を入力として機械学習することができれば、サポートベクトルマシン、ランダムフォレスト、勾配ブースティング法、ニューラルネットワークといった機械学習手法を適宜選択できる。

[0094] 次に、以上のように説明した実施例４に係るハイブリッドモデル作成装置１０の処理について説明する。

[0095] 図９は、実施例４に係るハイブリッドモデル作成装置１０の処理の一例を示すフローチャートである。なお、図９に示すステップＳ１、ステップＳ２、ステップＳ５及びステップＳ６は、図３で説明したステップＳ１、ステップＳ２、ステップＳ５及びステップＳ６と同様であるため説明を省略する。

[0096] ステップＳ３２１において、ハイブリッドモデル作成装置１０は、検証用データセットを用いて、複数のモデルそれぞれの推定結果を取得する。より具体的には、ハイブリッドモデル候補作成部１３は、モデルプール部１１にプールされているまたはモデル選択部１２により選択された複数のモデルそれぞれに、複数の検証用データセットを入力してカテゴリを推定させることで当該複数のモデルそれぞれの推定結果を取得する。推定結果は、上述したように、モデルの最終出力結果でもよいし、モデルの中間量であってもよい。なお、モデルプール部１１にプールされている複数のモデルそれぞれの推定結果を取得する場合、ステップＳ３２１は、ステップＳ２の前に実行されてもよい。

[0097] 次に、ハイブリッドモデル作成装置１０は、ステップＳ２で選択された２つ以上のモデルを組み合わせた複数のハイブリッドモデル候補を、機械学習モデルとして作成する（Ｓ３２２）。

[0098] ここで、複数のハイブリッドモデル候補のそれぞれは、組み合わせとして選ばれた２つ以上のモデルから出力された推定結果を入力として、検証用デ

ータセットのカテゴリを判定した判定結果を出力させる機械学習モデルである。この機械学習モデルは、典型的には、組み合わせとして選ばれた２つ以上のモデルを、ロジスティック回帰を用いて組み合わせることで得られるモデルである。また、機械学習モデルは、ユーザの指示に従いハイブリッドモデル候補作成部１３により作成される。

[0099] 次に、ハイブリッドモデル作成装置１０は、ステップＳ３２２で作成した複数のハイブリッドモデル候補それぞれに出力させた判定結果を比較する（Ｓ４１）。より具体的には、ハイブリッドモデル候補作成部１３は、例えばハイブリッドモデル候補のそれぞれに検証用データセットを入力して、出力させた判定結果の精度を比較する。

[0100] なお、ハイブリッドモデル候補作成部１３は、ハイブリッドモデル候補のそれぞれの判定結果から算出できる構成される２つ以上のモデルのそれぞれの重要度を比較してもよい。より具体的には、ハイブリッドモデル候補作成部１３は、複数のハイブリッドモデル候補を比較する際、複数のハイブリッドモデル候補それぞれに出力させた判定結果から、当該ハイブリッドモデル候補を構成するために選択された２つ以上のモデルのそれぞれの重要度を算出してもよい。そして、ハイブリッドモデル候補作成部１３は、算出された重要度のうち予め設定されていた閾値を下回った重要度のモデルを通知することで、ステップＳ４１の比較処理を行ってもよい。

[0101] また、ハイブリッドモデル候補作成部１３は、上記の通知をハイブリッドモデル選択部１４に対して行ってもよいし、閾値を下回った重要度のモデルをディスプレイなどに表示するなどで上記の通知を行ってもよい。これにより、ステップＳ６において、ハイブリッドモデル選択部１４は、予め設定されていた閾値を下回った重要度のモデルを有するハイブリッドモデル候補を除いた複数のハイブリッドモデル候補のうちの１つを、ハイブリッドモデルとして選択することができる。

[0102] 以下、重要度（寄与度）を比較する方法について説明する。

[0103] （式８）において、上述したように、関数$S_b(\beta_0+\beta_1 x_1+\beta_2 x_2)$の

出力は、２つのモデルを組み合わせた機械学習モデルを、検証用データセットを用いて係数を学習させた後に得られる出力（判定結果）である。係数β_1は、この機械学習モデルにおいてx_1を出力するモデルの重要度を示し、係数β_2は、この機械学習モデルにおいてx_2を出力するモデルの重要度を示す。つまり、係数β_iは、複数のモデルを組み合わせた機械学習モデルにおいてx_iを出力するモデルｉの重要度を示す。

[0104] ここで、係数β_iが０である、または、係数β_iが他の係数β_k（$i \neq k$）と比べて小さい場合には、モデルｉは、機械学習モデルにおいて判定結果に与える影響（貢献）が小さいと解析できる。

[0105] また、係数β_iが負の値である場合、モデルｉは、機械学習モデルの過学習の原因となっている可能性があると解析できる。機械学習モデルの判定結果と、機械学習モデルを構成する複数のモデルのそれぞれは、正の相関をもつべきと考えられるからである。

[0106] このように、複数のモデルを組み合わせた機械学習モデルを、検証用データセットを用いて係数を学習させ、係数を解析することで、組み合わせた複数のモデルそれぞれの重要度を解析することができる。

[0107] なお、係数β_iが０である、もしくは、係数βｉが他の係数β_kと比べて小さい場合、または、係数β_iが負の値である場合には、機械学習モデルを構成するモデルとしてモデルｉを用いないようにすればよい。

[0108] このように、ロジスティック回帰を用いて複数のモデルを組み合わせる場合、組み合わせられる複数のモデルのそれぞれの重要度の算出は容易である。一方、他の機械学習モデルでは、上述した係数解析の手法を用いることができない場合がある。しかし、推論結果が出たときの特徴量の寄与を解釈するためのツールであるＳＨＡＰ(SHapley Additive exPlanation)を利用すれば、組み合わせられる複数のモデルのそれぞれの重要度を算出することができる。

[0109] （実施例５）

複数のモデルには計算コストが高いモデルが含まれている場合がある。こ

のような場合、計算コストが高いモデルが組み合わされて作成されたハイブリッドモデル候補は、使用するハードウェアの要件または実行時間の要件を満たせない可能性がある。なお、実行時間が要件内であっても処理速度は速い方がよいと考えられる。

[0110] そこで、実施例５では、ロジスティック回帰を用いてハイブリッドモデル候補を機械学習により作成するときに、処理速度を加味して作成する。以下、その具体例について説明する。

[0111] 処理速度を加味する方法としては、ハイブリッドモデル候補を構成する複数のモデルのそれぞれの実行時間の合計を用いる方法と、機械学習する際の損失関数に正則化項を追加する方法とがある。

[0112] 処理速度を加味する１つ目の方法として、ハイブリッドモデル候補を構成する複数のモデルのそれぞれの実行時間の合計を用いる方法について説明する。

[0113] まず、ハイブリッドモデル候補作成部１３は、モデルプール部１１にプールされているまたはモデル選択部１２により選択された複数のモデルのそれぞれにおいて、検証用データセットが入力されて検証用データセットのカテゴリを推定するまでに要した処理時間を計測（取得）する。ここで、検証用データセットにはＸ個のサンプルデータが含まれているとする。

[0114] 次に、ハイブリッドモデル候補作成部１３は、計測した処理時間から、複数のモデルのそれぞれについて、１個のサンプルデータ当たりの処理時間である平均処理時間を算出する。

[0115] 次に、ハイブリッドモデル候補作成部１３は、モデル選択部１２で選択された２つ以上のモデルのうち、平均処理時間の合計が実行時間の要件を満たす組み合わせのみで、複数のハイブリッドモデル候補を作成する。なお、ロジスティック回帰を用いてハイブリッドモデル候補を機械学習により作成する方法については、実施例４で説明した通りであるので、ここでの説明は省略する。

[0116] 続いて、処理速度を加味する２つ目の方法として、機械学習する際の損失

関数に正則化項を追加する方法について説明する。

[0117] まず、ハイブリッドモデル候補作成部１３は、モデルプール部１１にプールされているまたはモデル選択部１２により選択された複数のモデルのそれぞれにおいて、検証用データセットが入力されて検証用データセットのカテゴリを推定するまでに要した処理時間を取得する。ここで、検証用データセットにはＸ個のサンプルデータが含まれているとする。

[0118] 次に、ハイブリッドモデル候補作成部１３は、取得した処理時間から、複数のモデルのそれぞれについて、１個のサンプルデータ当たりの処理時間である平均処理時間を算出する。ハイブリッドモデル候補作成部１３は、複数のモデルのすべての平均処理時間の和に対する複数のモデルのそれぞれの平均処理時間の値をハードウェアコストと定義する。

[0119] すなわち、ハードウェアコストC_mは、下記の（式９）のように定義できる。

c_m=avg(modelの処理速度)/sum(avg_all_modelsの処理速度) （式９）

[0120] なお、ロジスティック回帰を用いて、ハイブリッドモデル候補を機械学習により作成する方法については、実施例４で説明した通りであるので、ここでの説明は省略する。

[0121] 次に、ハイブリッドモデル候補作成部１３は、複数のハイブリッドモデル候補それぞれの機械学習モデルの損失関数に、当該ハイブリッドモデル候補を構成するために選択された２つ以上のモデルのそれぞれのハードウェアコストを加味した正則化項を追加する。

[0122] ハードウェアコストを加味した正則化項は、例えば、Ｌａｓｓｏ（Ｌ１ノルムまたはＬ１正則化）などの正則化項に、パラメータαとハードウェアコストC_mとを乗算したα・C_m・Ｌ１正則化項で表すことができる。ここで、パラメータαは、ハードウェアコストの重みを変えることができるハイパーパラメータである。詳細について後述する。

[0123] 次に、ハイブリッドモデル候補作成部１３は、ハードウェアコストを加味した正則化項を追加したうえで、ロジスティック回帰の機械学習を実行する

。これにより、計算コストが大きい割にハイブリッドモデル候補への貢献が小さいモデルの係数（重み）を小さくすることができるので、ハイブリッドモデル候補への貢献が小さいモデルを除外することができる。

[0124] ここで、ハードウェアコストを正則化項として追加する方法の詳細例について説明する。

[0125] 学習に用いるデータセットのデータ数をＮとし、データセットのｎ番目のデータの真値をｔ$_n$とし、ｎ番目のデータの説明変数セットをϕ_nとすると、ロジスティック回帰の損失関数Ｅ（ｗ）は、（式１０）のように表される。そして、機械学習を行う際には、損失関数Ｅ（ｗ）を最小化するような重み（係数）の組み合わせを得るように学習することになる。

[0126] [数9]

$$E(\boldsymbol{w}) = -\sum_{n=1}^{N} t_n \ln y_n + (1-t_n)\ln(1-y_n)$$

$$y_n = \sigma(\boldsymbol{w}^{\mathrm{T}}\boldsymbol{\phi}_n) = \frac{1}{1-\exp(-\mathbf{w}^{\mathrm{T}}\boldsymbol{\phi}_n)} \qquad \text{（式１０）}$$

[0127] ここで、説明変数の次元数をｍとすると、例えばＬ１正則化項を損失関数Ｅ（ｗ）に追加した損失関数Ｅ´（ｗ）は、（式１１）のように表される。（式１１）において、パラメータαは、ハイパーパラメータである。

[0128] [数10]

$$E'(\boldsymbol{w}) = E(\boldsymbol{w}) + \alpha\sum_{m=1}^{M}|w_m| \qquad \text{（式１１）}$$

[0129] 本実施例では、説明変数は、各モデルの出力値である。したがって、対応するモデルのハードウェアコストＣ$_m$を用いると、ハードウェアコストを加味した損失関数Ｅ´（ｗ）は、（式１２）のように表すことができる。

[0130] [数11]

$$E'(\boldsymbol{w}) = E(\boldsymbol{w}) + \alpha\sum_{m=1}^{M}c_m|w_m| \qquad \text{（式１２）}$$

[0131] なお、ロジスティック回帰の場合においてハードウェアコストを正則化項

として追加する方法について説明したがこれに限らない。一般の機械学習に対しても、同様に、損失関数Ｅ（ｗ）に（式１２）の右辺第二項を追加すれば、ハードウェアコスト$Ｃ_{m}$を加味した損失関数Ｅ´（ｗ）を作ることができる。

[0132] また、上記では、正則化項としてＬ１正則化項を用いる場合の例を説明したが、Ｌ２正則化項を用いてもよい。この場合でも、同様に、ハードウェアコスト$Ｃ_{m}$を加味した損失関数を定義できる。なお、Ｌ１正則化項は、重み（係数）の値を小さくできるだけでなく０にする効果が期待できる。このため、処理時間の長いモデルを除外した組み合わせによりハイブリッドモデル候補を作成するという目的には、Ｌ２正則化項よりもＬ１正則化項を用いた方がよい。

[0133] 次に、以上のように説明した実施例５に係るハイブリッドモデル候補の作成処理について説明する。

[0134] 図１０は、実施例５に係るステップＳ３の詳細処理の一例を示すフローチャートである。なお、図１０は、図９に示すステップＳ３２１及びステップＳ３２２の処理の別の例に該当する。

[0135] ステップＳ３において、ハイブリッドモデル作成装置１０は、検証用データセットを用いて、複数のモデルそれぞれにおける処理時間と推定結果とを取得する（Ｓ３３１）。より具体的には、ハイブリッドモデル候補作成部１３は、モデル選択部１２により選択された複数のモデルそれぞれに、複数の検証用データセットを入力してカテゴリを推定させる。ハイブリッドモデル候補作成部１３は、モデル選択部１２により選択された複数のモデルそれぞれにおいて、検証用データセットが入力されて検証用データセットのカテゴリを推定するまでに要した処理時間と推定結果とを取得する。推定結果は、上述したように、モデルの最終出力結果でもよいし、モデルの中間量であってもよい。なお、処理時間と推定結果とは、モデルプール部１１にプールされている複数のモデルそれぞれから取得してもよい。この場合、ステップＳ３３１は、ステップＳ２の前に実行されてもよい。

[0136] 次に、ハイブリッドモデル作成装置１０は、ステップＳ３３１で取得した処理時間に基づき、当該複数のモデルのすべての処理時間の和に対する当該複数のモデルのそれぞれの当該要する時間の値をハードウェアコストと定義する（Ｓ３３２）。ここで、ハードウェアコストを定義するために用いる処理時間は、平均処理時間である。

[0137] 次に、ハイブリッドモデル作成装置１０は、ステップＳ２で選択された２つ以上のモデルを組み合わせた複数のハイブリッドモデル候補を、機械学習モデルとして作成する。ここで、ハイブリッドモデル作成装置１０は、機械学習する際の複数のハイブリッドモデル候補それぞれの損失関数にハードウェアコストを加味した正則化項を追加する（Ｓ３３３）。より具体的には、複数のハイブリッドモデル候補のそれぞれの損失関数には、当該ハイブリッドモデル候補を構成するために選択された２つ以上のモデルのそれぞれのハードウェアコストが加味（乗算）された正則化項が追加される。

[0138] なお、続くステップＳ４において複数のハイブリッドモデル候補を比較する前に、ハイブリッドモデル候補作成部１３は、検証用データセットを用いて学習させた後に得られる出力（判定結果）から係数解析を行う。これにより、ハイブリッドモデル候補作成部１３は、処理時間の長いモデルを含むハイブリッドモデル候補を除外することができる。よって、ハイブリッドモデル候補作成部１３は、続くステップＳ４において、処理時間の長いモデルを含むハイブリッドモデル候補を除外した上で複数のハイブリッドモデル候補の比較処理を行えばよい。

[0139] （実施例６）

ハイブリッドモデル候補は、組み合わせとして選ばれた２つ以上のモデルから出力された推定結果を入力として、検証用データセットのカテゴリを判定した判定結果を出力させる機械学習モデルとして作成され、機械学習される。実施例６では、機械学習モデルを機械学習する際に、２つ以上のモデルから出力された推定結果が確実にＮＧを示しかつ、真の値（ラベル）もＮＧである明確なＮＧを示す出力を、除外して機械学習される場合について説明

する。

[0140] 図１１は、実施例６に係るモデル１とモデル２とで組み合わせて作成されるハイブリッドモデル候補の一例を概念的に示す図である。図１１に示すハイブリッドモデル候補は、機械学習により作成されるロジスティック回帰モデル（境界）である。縦軸は、検証用データセットをモデル２に入力したときに出力（推定）される出力値であり、確率で表現される。同様に、横軸は、検証用データセットをモデル１に入力したときに出力（推定）される出力値であり、確率で表現される。図１１において、検証用データセットに含まれるサンプルデータが製造品の検査画像であるとすると、黒丸は、サンプルデータの真の値が良品である検査画像であり、良品画像と称している。白丸は、サンプルデータの真の値が不良品である検査画像であり、不良品画像と称している。

[0141] ハイブリッドモデル候補を作成する際、真の値が不良品である検査画像を良品と判定するような見逃し判定を極力抑えるすなわち判定精度を底上げすることが必要になる。見逃し判定は、ＮＧ（真の値が不良品）をＯＫ（良品）と誤判定することである。判定精度を底上げする方法としては、上述したように、各モデルで推定結果が異なっているサンプルデータを用いて機械学習されることがある。また、図１１に示すロジスティック回帰モデル（境界）となるように機械学習させるために、境界に近い良品画像に対応するモデル１及びモデル２の出力の存在が重要である。一方、モデル１及びモデル２の出力値（確率）が共に大きい不良品画像の出力（明確なＮＧを示す出力と称する）は、図１１に示すロジスティック回帰モデル（境界）となるように機械学習させる場合には、相対的に重要度が低いことがわかる。

[0142] また、図１１に示す例では、円で囲まれた領域の出力は、明確なＮＧを示す出力である。円で囲まれた領域に含まれている明確なＮＧの数は多い。このため、図１１に示すようなモデル１及びモデル２の出力と、検証用データセットの真の値（ラベル）とでハイブリッドモデル候補を機械学習で作成する際、円で囲まれた領域にあるような、明確なＮＧを示す出力に強く影響を

受け、図１１に示す境界を得られない可能性がある。

[0143] そこで、ハイブリッドモデル候補を機械学習で作成する際に、円で囲まれた領域にある出力である明確なＮＧを示す出力を、除外して機械学習する。

[0144] 具体的には、ハイブリッドモデル候補作成部１３は、当該ハイブリッドモデル候補を構成するために選択された２つ以上のモデルのそれぞれに検証用データセットを入力してカテゴリを推定させることで得た複数の出力値から、閾値より高い値で不良品であると推定された出力値を除外する。次いで、ハイブリッドモデル候補作成部１３は、閾値より高い出力値が除外された複数の出力値を入力として用いて、かつ、当該複数の出力値に対応する検証用データセットの真の値を用いて機械学習を行うことで、複数のハイブリッドモデル候補を作成する。

[0145] 次に、以上のように説明した実施例６に係るハイブリッドモデル候補の作成処理について説明する。

[0146] 図１２は、実施例６に係るステップＳ３の詳細処理の一例を示すフローチャートである。なお、図１２は、図９に示すステップＳ３２１及びステップＳ３２２の処理の別の例に該当する。

[0147] ステップＳ３において、ハイブリッドモデル作成装置１０は、検証用データセットを用いて、ハイブリッドモデル候補を構成する２つ以上のモデルのそれぞれに推定させて複数の出力値を取得する（Ｓ３４１）。

[0148] 次に、ハイブリッドモデル作成装置１０は、ステップＳ３４１で取得した複数の出力値から、閾値より高い値で不良品であると推定された出力値を除外する（Ｓ３４２）。ここで、閾値より高い値で不良品であると推定された出力値は、図１１を用いて説明した明確なＮＧを示す出力値である。

[0149] 次に、ハイブリッドモデル作成装置１０は、ステップＳ３４２で閾値より高い出力値が除外された複数の出力値を入力として用いて、かつ、当該複数の出力値に対応する検証用データセットの真の値を用いて機械学習を行うことで、複数のハイブリッドモデル候補を作成する（Ｓ３４２）。

[0150] このように、ハイブリッドモデル作成装置１０は、明確なＮＧが集まる領

域に含まれる出力を除外して機械学習することで、判定精度の高い複数のハイブリッドモデル候補を作成することができる。

[0151] （実施例７）

実施例６では、ハイブリッドモデル候補を構成する２つ以上のモデルのそれぞれの出力のうち明確なＮＧが集まる領域に含まれる出力を除外して機械学習する場合について説明したが、これに限らない。実施例７では、明確なＮＧを示す出力を除外する別の方法として、凸包を用いる方法について説明する。なお、凸包とは、与えられた点をすべて包含する最小の凸多角形（凸多面体）のことを意味する。

[0152] 図１３は、実施例７に係るモデル１とモデル２との出力と不良品画像に対応する出力の分布の凸包とを概念的に示す図である。図１４は、図１３に示す凸包の頂点を除く不良品画像に対応する出力を除去した、モデル１とモデル２との出力から作成されるハイブリッドモデル候補の一例を概念的に示す図である。図１４の（ａ）には、図１３に示す凸包の頂点以外のＮＧを示す出力が除去された、モデル１とモデル２との出力が概念的に示されている。図１４の（ｂ）には、図１４の（ａ）に示されるモデル１とモデル２との出力から機械学習で作成されるハイブリッドモデル候補としてのロジスティック回帰モデル（境界）の一例が概念的に示されている。

[0153] 具体的には、ハイブリッドモデル候補作成部１３は、当該ハイブリッドモデル候補を構成するために選択された２つ以上のモデルのそれぞれに検証用データセットを入力してカテゴリを推定させることで得た複数の出力値うち、不良品であると推定された出力値をプロットしたときの凸包を算出する。次いで、ハイブリッドモデル候補作成部１３は、複数の出力値から、凸包の頂点を除き凸包に含まれる出力値を除外する。そして、ハイブリッドモデル候補作成部１３は、凸包の頂点を除き凸包に含まれる出力値が除外された複数の出力値を入力して用いて、かつ、当該複数の出力値に対応する検証用データセットの真の値を用いて機械学習を行うことで、複数のハイブリッドモデル候補を作成する。

[0154] これにより、見逃し（見逃し判定）が０となるような判定精度を有するハイブリッドモデル候補を作成することができる。

[0155] 次に、以上のように説明した実施例７に係るハイブリッドモデル候補の作成処理について説明する。

[0156] 図１５は、実施例７に係るステップＳ３の詳細処理の一例を示すフローチャートである。なお、図１５は、図９に示すステップＳ３２１及びステップＳ３２２の処理の別の例に該当する。

[0157] ステップＳ３において、ハイブリッドモデル作成装置１０は、検証用データセットを用いて、ハイブリッドモデル候補を構成する２つ以上のモデルのそれぞれに推定させて複数の出力値を取得する（Ｓ３５１）。

[0158] 次に、ハイブリッドモデル作成装置１０は、ステップＳ３５１で取得した複数の出力値のうち、不良品であると推定された出力値をプロットしたときの凸包を算出する（Ｓ３５２）。

[0159] 次に、ハイブリッドモデル作成装置１０は、ステップＳ３５１で取得した複数の出力値から、凸包の頂点を除く凸包に含まれる出力値を除外する（Ｓ３５３）。

[0160] 次に、ハイブリッドモデル作成装置１０は、凸包の頂点を除く凸包に含まれる出力値が除外された複数の出力値を入力として用いて、かつ、当該複数の出力値に対応する検証用データセットの真の値を用いて機械学習を行うことで、複数のハイブリッドモデル候補を作成する（Ｓ３５４）。

[0161] このように、凸包を用いて明確なＮＧを示す出力を除外することで、ハイブリッドモデル作成装置１０は、見逃し（見逃し判定）が０となるような判定精度の高い複数のハイブリッドモデル候補を作成することができる。

[0162] なお、ハイブリッドモデル候補を構成するために選択されたモデルの数（次元数）が例えば１０以上などの大きい数の場合には、凸包の頂点の数が膨大になったり、凸包の算出コストが大きくなったりするため、凸包を用いる方法が採用できない場合がある。

[0163] このような場合には、実施例６で説明したが、図１６に示されるような明

確なＮＧが集まる領域（除外領域）に含まれる出力を除外すればよい。

[0164] 図１６は、実施例７に係るモデル１とモデル２との出力と除外領域とを概念的に示す図である。図１７は、図１６に示す除外領域に含まれる不良品画像に対応する出力を除去したモデル１とモデル２との出力から作成されるハイブリッドモデル候補の一例を概念的に示す図である。図１７の（ａ）には、図１６に示す除外領域に含まれるＮＧを示す出力が除去されたモデル１とモデル２との出力が概念的に示されている。図１７の（ｂ）には、図１７の（ａ）に示されるモデル１とモデル２との出力から機械学習で作成されるハイブリッドモデル候補としてのロジスティック回帰モデル（境界）の一例が概念的に示されている。

[0165] 図１６及び図１７に示すように、モデル１の出力及びモデル２の出力において共に、ＮＧを示す出力値（確率）が大きく、真の値もＮＧである明確なＮＧを示す出力が集まる領域を除外領域として算出すればよい。このような算出方法は、凸包の算出の近似的な方法として用いることができる。そして、除外領域にある明確なＮＧを示す出力を、除外して機械学習すればよい。

[0166] なお、このような近似的手法は、次元数が小さい低次元の場合においても有効である。凸包を用いる手法を行う場合、次元数が小さい低次元のときには機械学習に用いることができるモデルの出力が少なくなりすぎ、機械学習が不安定になるからである。

[0167] （実施例８）

複数のハイブリッドモデル候補を比較する比較方法として、ハイブリッドモデル候補のそれぞれの判定結果を比較する方法がある。

[0168] ここで、通常、機械学習による判定結果は確率で出力される。しかし、判定結果として出力される確率は、判定結果として示されるカテゴリの実際の確率を表すわけではない。つまり、例えば、サンプルデータとして入力された検査画像に示される製造品が不良品であるか否かを判定した判定結果が０．９であっても、その製造品が不良品である確率は９０％であるとは限らず、実際の確率と判定結果との間には差異があることが知られている。

[0169] また、ＡＩ判定結果として示される確率を実際の確率に合わせこむ技術も知られており、Confidence Calibrationと呼ばれている。

[0170] 本実施例では、複数のハイブリッドモデル候補それぞれに出力させた判定結果として、ハイブリッドモデル候補の見逃し率を使用する。また、ハイブリッドモデル候補を作成するために選択された複数のモデルのＦＡＲ表を算出し、ハイブリッドモデル候補の見逃し率を調整するパラメータとして使用する。

[0171] 図１８は、実施例８に係るＦＡＲ曲線をモデル１に対して算出する方法を説明するための図である。図１９は、実施例８に係るモデル１のＦＡＲ表の一例を示す図である。モデル１は、ハイブリッドモデル候補を作成するために選択された複数のモデルのうちの一つである。

[0172] ここで、ＦＡＲは、False Acceptance Rateの略であり、ＮＧをＯＫと判定と誤判定する確率である。本実施例では、ＦＡＲ値を見逃し率と称する。また、ＦＡは、False Acceptの略であり、ＮＧをＯＫと判定と誤判定することである。本実施例では、ＦＡを見逃しまたは見逃し判定と称している。ＦＡＲ表は、所定のステップサイズで閾値を変動させたときの見逃し率（ＦＡＲ値）を表にしたものである。

[0173] 具体的には、まず、ハイブリッドモデル候補作成部１３は、機械学習の際、検証用データセットを用いて、ハイブリッドモデル候補を作成するために選択された複数のモデルそれぞれのＦＡＲ表を作成する。

[0174] 例えば図１８に示す例で説明すると、まず、モデル１に検証用データセットを入力してカテゴリを推定させることで得た出力値と頻度とを取得する。次いで、図１８の（ａ）に示すように、検証用データセットのうち良品を示す出力値（確率）及びその頻度で示される分布と、検証用データセットのうち不良品を示す出力値（確率）及びその頻度で示される分布とに層別する。次いで、図１８の（ａ）に示される不良品を示す出力値（確率）の分布において、閾値を徐々に増やしたときに良品と判定される出力値の面積割合を取得することで図１８の（ｂ）に示すＦＡＲ曲線を得ることができる。なお、

当該分布全体の面積を１としたときの面積割合が、見逃し率（ＦＡＲ値）に該当する。また、図１８の（ａ）に示される不良品を示す出力値（確率）の分布において、所定のステップサイズで見逃し率（ＦＡＲ値）を取得することで、図１９に示されるＦＡＲ表を取得できる。なお、図１９に示されるＦＡＲ表では、ステップサイズが０．００７８１２５と設定され、ステップサイズごとに振られたインデックスに対して０～１のＦＡＲ値が記載されている。

[0175] このように、ハイブリッドモデル候補作成部１３は、選択された複数のモデルのそれぞれにおいて、検証用データセットのうち不良品を示す複数のデータを入力してカテゴリを推定させることで得た出力値の分布からＦＡＲ表を作成できる。ハイブリッドモデル候補作成部１３は、取得した出力値の分布において閾値を変動させることで見逃し率を得ることができるので、見逃し率の表であるＦＡＲ表を作成することができる。

[0176] 次に、ハイブリッドモデル候補作成部１３は、ハイブリッドモデル候補を構成するために選択された２つ以上のモデルのそれぞれに、検証用データセットに含まれるデータサンプルを入力してカテゴリを推定させた推定結果を取得する。ハイブリッドモデル候補作成部１３は、取得した推定結果（出力値）と、予め作成しておいたＦＡＲ表とを照合することで、当該データサンプルに対する２つ以上のモデルのそれぞれのＦＡＲ値である第１ＦＡＲ値を取得する。

[0177] ここで、ハイブリッドモデル候補を構成する例えばモデル１に、真の値が不良品を示す検査画像であるサンプル画像を入力したときの推定結果として０．９９が得られたとする。この場合、図１９に示すモデル１のＦＡＲ表から、推定結果が０．９９であるときのＦＡＲ値を取得する。具体的には、Ｔａｂｌｅ［（１－推定結果）／ｓｔｅｐ＿ｓｉｚｅ］＝Ｔａｂｌｅ［（１－０．９９）／０．００７８１２５］＝Ｔａｂｌｅ［1］と算出できることから、図１９に示すＦＡＲ表において、Ｉｎｄｅｘ＝１のＦＡＲ値を取得する。これにより、ハイブリッドモデル候補作成部１３は、第１ＦＡＲ値として

ＦＡＲ値＝０．０００００２３を取得できる。

[0178] このようにして、本実施例では、検査画像が不良品である確率を、ＦＡＲ表を作成したときの不良品を示す出力値（確率）の分布に基づいて推定（調整）することができる。

[0179] 次に、ハイブリッドモデル候補作成部１３は、取得した２つ以上のモデルのそれぞれの第１ＦＡＲ値を乗算する。これにより、ハイブリッドモデル候補作成部１３は、当該２つ以上のモデルで組み合わされたハイブリッドモデル候補のＦＡＲ値である第２ＦＡＲ値を取得することができる。

[0180] 図２０は、実施例８に係る２つのモデルそれぞれの第１ＦＡＲ値と、２つのモデルを組み合わせて作成されるハイブリッドモデル候補の第２ＦＡＲ値とを概念的に示す図である。図２０に示すように、２つ以上のモデルのそれぞれの第１ＦＡＲ値を乗算することで、２つ以上のモデルのそれぞれの第１ＦＡＲ値よりも改善された第２ＦＡＲ値を得ることができる。

[0181] ここで、ハイブリッドモデル候補を構成する複数のモデルのＦＡＲ分布は独立であると仮定している。このため、独立な事象に対する確率の法則により、２つ以上のモデルのそれぞれの第１ＦＡＲ値を乗算することで、当該２つ以上のモデルで組み合わされたハイブリッドモデル候補の第２ＦＡＲ値を取得できる。

[0182] なお、ハイブリッドモデル候補を構成する複数のモデルのＦＡＲ分布が独立でない場合には、複数のモデルすべての相関係数を算出して、性能が良いものが支配的となるように第２ＦＡＲ値を補正すればよい。

[0183] 相関係数の算出方法は、例えば次の通りである。すなわち、まず、ハイブリッドモデル候補作成部１３は、複数のハイブリッドモデル候補を作成するために選択された複数のモデルのそれぞれにおいて、複数の検証用データセットを入力してカテゴリを推定させることで当該複数のモデルそれぞれの推定結果を取得する。次いで、ハイブリッドモデル候補作成部１３は、取得した推定結果を用いて、複数のモデルのうち２つのモデルの組み合わせすべての相関係数を算出すればよい。これにより、ハイブリッドモデル候補作成部

１３は、取得した２つ以上のモデルのそれぞれの第１ＦＡＲ値を乗算して、さらに、相関係数が大きいほど小さくなる係数を乗算することで、補正した第２ＦＡＲ値を取得することができる。

[0184] 次に、ハイブリッドモデル候補作成部１３は、第２ＦＡＲ値が、事前に設定された閾値（ＦＡＲ閾値）より小さい場合に、当該データサンプルが良品であると判定する。ハイブリッドモデル候補作成部１３は、この判定結果を、ハイブリッドモデル候補に当該データサンプル入力させたときの判定結果として取得することができる。これにより、ハイブリッドモデル候補作成部１３は、第２ＦＡＲ値と事前に設定された閾値とを用いて調整した判定結果を、複数のハイブリッドモデル候補にデータサンプルを入力したときの判定結果として取得できる。そして、ハイブリッドモデル候補作成部１３は、調整した判定結果を用いて、複数のハイブリッドモデル候補を比較できる。

[0185] なお、ＦＡＲ閾値は、事前にハイブリッドモデルを利用するユーザがどの程度の見逃し率を許容できるかに基づいて決定されればよい。

[0186] ここで、例えば真の値が不良品を示す検査画像のうち１ｐｐｍの見逃し率を許容することをユーザが決定し、事前に閾値（ＦＡＲ閾値）を１／１，０００，０００と設定したとする。また、ハイブリッドモデル候補を構成する複数のモデルがモデル１とモデル２である。この場合、上述のようにして、あるサンプルデータに対するモデル１とモデル２とのそれぞれの第１ＦＡＲ値を取得すると、第２ＦＡＲ値は、これらを乗算した値として取得できる。そして、第２ＦＡＲ値が、ＦＡＲ閾値である１／１，０００，０００より小さければサンプルデータはＮＧ（不良品を示す）、大きければＯＫ（良品を示す）と判定できる。

[0187] 以上の実施の形態及び実施例によれば、本開示に係るハイブリッドモデル作成装置１０及びハイブリッドモデル作成方法は、予め準備されプールされている複数のモデルを全部使用しないハイブリッドモデルを作成することができる。また、本開示に係るハイブリッドモデル作成装置１０及びハイブリッドモデル作成方法は、処理速度の観点から、計算コストが高く貢献のすく

ないモデルを除外したハイブリッドモデル候補を作成できるので、ハイブリッドモデルを軽量かつ効果的に作成できる。さらに、本開示に係るハイブリッドモデル作成装置１０及びハイブリッドモデル作成方法は、重要度を用いて精度向上に貢献しないモデルを除外したハイブリッドモデル候補を作成できるので、ハイブリッドモデルを軽量かつ効果的に作成できる。

[0188] 以上、本開示に係るハイブリッドモデル作成装置１０などについて、実施の形態及び各実施例に基づいて説明したが、本開示は、これら実施の形態等に限定されるものではない。本開示の主旨を逸脱しない限り、当業者が思いつく各種変形を実施の形態及び各実施例に施したものや、実施の形態及び各実施例における一部の構成要素を組み合わせて構築される別の形態も、本開示の範囲内に含まれる。

[0189] （その他の実施の形態）

（１）上記の実施の形態では、ハイブリッドモデル作成装置１０は、プールされている複数のモデルから選択した複数のモデルを、ロジスティック回帰などを用いて組み合わせたハイブリッドモデル候補を作成し、比較することで１つのハイブリッドモデルを選択したが、これに限らない。プールされている複数のモデルから選択した複数のモデルを組み合わせされて、組み合わせた順番に論理式で推定処理を行わせるハイブリッドモデル候補を作成して、精度を比較することで１つのハイブリッドモデルを選択してもよい。

[0190] 図２１は、その他の実施の形態に係るハイブリッドモデル作成方法の一例を示す図である。

[0191] 図２１では、プールされている複数のモデルから、モデル１、モデル２及びモデル３が選択された場合のハイブリッドモデル作成方法が示されている。図２１では、矢印で繋がれている異なる３つのモデルをこの順で組み合わせてハイブリッドモデル候補を作成する。ハイブリッドモデル候補は、モデル１、モデル２及びモデル３の組み合わせされた順番にそれぞれの精度の論理和または論理積を取った精度で比較される。図２１に示される例では、モデル３－モデル１－モデル２の順で組み合わせたハイブリッドモデル候補の

精度が９３％と一番高いため、ハイブリッドモデルとして選択されることが示されている。

[0192] 図２２は、その他の実施の形態に係るハイブリッドモデル作成方法の他の一例を示す図である。図２２では、プールされている複数のモデルから、モデル１、モデル２及びモデル３が選択されている場合に、モデル１、モデル２及びモデル３の少なくとも２つ以上を組み合わせたハイブリッドモデル候補を作成する方法が示されている。図２２では、モデル２とモデル１とがこの順で組み合わされたハイブリッドモデル候補の精度が９３％と一番高いため、ハイブリッドモデルとして選択される。

[0193] （２）上記の実施の形態では、ハイブリッドモデル作成装置１０を構成する判定閾値決定部１５は、混同行列を用いて判定閾値を決定すると説明したが、図２３Ａ及び図２３Ｂに示す混同行列の表を用いて、以下の２ステップで判定閾値を決定してもよい。

[0194] 図２３Ａ及び図２３Ｂは、その他の実施の形態に係る混同行列の表の一例を示す図である。

[0195] まず、ステップ１において、判定閾値決定部１５は、検証用データセットを用いてハイブリッドモデル選択部１４により選択されたハイブリッドモデルの判定結果（ＯＫまたはＮＧの２値の予測値）を取得する。判定閾値決定部１５は、閾値を０．５として、判定結果と真の値（ＯＫまたはＮＧの２値）との組み合わせから、例えば図２３Ａに示す混同行列にまとめた表を作成する。

[0196] 次に、ステップ２において、例えば過検出率を０．８６％など、所望の精度を入力して、上記の判定結果（ＯＫまたはＮＧの２値の予測値）を、真の値（ＯＫまたはＮＧの２値）のリストに並び替えて図２３Ｂに示す混同行列の表を作成する。ここで、０．８６%の過検出率が所望の精度である場合、図２３Ｂに示される閾値０．４２を最適な閾値（判定閾値）として選択することができる。

[0197] また、以下に示す形態も、本開示の一つ又は複数の態様の範囲内に含まれ

てもよい。

[0198] （３）上記のハイブリッドモデル作成装置１０を構成する構成要素の一部は、マイクロプロセッサ、ＲＯＭ、ＲＡＭ、ハードディスクユニット、ディスプレイユニット、キーボード、マウスなどから構成されるコンピュータシステムであってもよい。前記ＲＡＭ又はハードディスクユニットには、コンピュータプログラムが記憶されている。前記マイクロプロセッサが、前記コンピュータプログラムにしたがって動作することにより、その機能を達成する。ここでコンピュータプログラムは、所定の機能を達成するために、コンピュータに対する指令を示す命令コードが複数個組み合わされて構成されたものである。

[0199] （４）上記のハイブリッドモデル作成装置１０を構成する構成要素の一部は、１個のシステムＬＳＩ（Large Scale Integration：大規模集積回路）から構成されているとしてもよい。システムＬＳＩは、複数の構成部を１個のチップ上に集積して製造された超多機能ＬＳＩであり、具体的には、マイクロプロセッサ、ＲＯＭ、ＲＡＭなどを含んで構成されるコンピュータシステムである。前記ＲＡＭには、コンピュータプログラムが記憶されている。前記マイクロプロセッサが、前記コンピュータプログラムにしたがって動作することにより、システムＬＳＩは、その機能を達成する。

[0200] （５）上記のハイブリッドモデル作成装置１０を構成する構成要素の一部は、各装置に脱着可能なＩＣカード又は単体のモジュールから構成されているとしてもよい。前記ＩＣカード又は前記モジュールは、マイクロプロセッサ、ＲＯＭ、ＲＡＭなどから構成されるコンピュータシステムである。前記ＩＣカード又は前記モジュールは、上記の超多機能ＬＳＩを含むとしてもよい。マイクロプロセッサが、コンピュータプログラムにしたがって動作することにより、前記ＩＣカード又は前記モジュールは、その機能を達成する。このＩＣカード又はこのモジュールは、耐タンパ性を有するとしてもよい。

[0201] （６）また、上記のハイブリッドモデル作成装置１０を構成する構成要素の一部は、前記コンピュータプログラム又は前記デジタル信号をコンピュー

タで読み取り可能な記録媒体、例えば、フレキシブルディスク、ハードディスク、ＣＤ－ＲＯＭ、ＭＯ、ＤＶＤ、ＤＶＤ－ＲＯＭ、ＤＶＤ－ＲＡＭ、ＢＤ（Blu-ray（登録商標） Disc）、半導体メモリなどに記録したものとしてもよい。また、これらの記録媒体に記録されている前記デジタル信号であるとしてもよい。

[0202] また、上記のハイブリッドモデル作成装置１０を構成する構成要素の一部は、前記コンピュータプログラム又は前記デジタル信号を、電気通信回線、無線又は有線通信回線、インターネットを代表とするネットワーク、データ放送等を経由して伝送するものとしてもよい。

[0203] （７）本開示は、上記に示す方法であるとしてもよい。また、これらの方法をコンピュータにより実現するコンピュータプログラムであるとしてもよいし、前記コンピュータプログラムからなるデジタル信号であるとしてもよい。

[0204] （８）また、本開示は、マイクロプロセッサとメモリを備えたコンピュータシステムであって、前記メモリは、上記コンピュータプログラムを記憶しており、前記マイクロプロセッサは、前記コンピュータプログラムにしたがって動作するとしてもよい。

[0205] （９）また、前記プログラム又は前記デジタル信号を前記記録媒体に記録して移送することにより、又は前記プログラム又は前記デジタル信号を、前記ネットワーク等を経由して移送することにより、独立した他のコンピュータシステムにより実施するとしてもよい。

[0206] （１０）上記実施の形態及び上記変形例をそれぞれ組み合わせるとしてもよい。

産業上の利用可能性

[0207] 本開示は、検査工程における良品判定などを行うために機械学習のモデルを組み合わせたハイブリッドモデルを作成する方法、ハイブリッドモデル方法、ハイブリッドモデル作成装置、及び、プログラムなどに利用できる。

符号の説明

[0208] １０　ハイブリッドモデル作成装置

１１　モデルプール部

１１ａ　モデル

１２　モデル選択部

１２ａ　モデル選択処理

１３　ハイブリッドモデル候補作成部

１３ａ　ハイブリッドモデル候補作成処理

１４　ハイブリッドモデル選択部

１４ａ　ハイブリッドモデル選択処理

１５　判定閾値決定部

１５ａ　判定閾値決定処理

請求の範囲

[請求項1] 入力されるデータのカテゴリを推定する複数のモデルをプールし、

前記複数のモデルの少なくとも一つのモデルは、機械学習されたモデルであり、

プールされている複数のモデルから２つ以上のモデルを選択して組み合わせることで、前記カテゴリを判定するハイブリッドモデル候補を複数作成し、

複数の前記ハイブリッドモデル候補を比較することで、前記複数のハイブリッドモデル候補のうちの１つをハイブリッドモデルとして選択する、

ハイブリッドモデル作成方法。

[請求項2] 前記入力されるデータは、製造品の検査画像であり、

判定されるカテゴリは、前記製造品が良品または不良品であるかである、

請求項１に記載のハイブリッドモデル作成方法。

[請求項3] 前記複数のハイブリッドモデル候補を作成する際、

当該ハイブリッドモデル候補を構成するために選択された前記２つ以上のモデルのそれぞれに検証用データセットを入力してカテゴリを推定させることで得た複数の出力値から、閾値より高い値で不良品であると推定された出力値を除外し、

前記閾値より高い出力値が除外された前記複数の出力値を入力として用いて、かつ、前記複数の出力値に対応する検証用データセットの真の値を用いて機械学習を行うことで、前記複数のハイブリッドモデル候補を作成する、

請求項２に記載のハイブリッドモデル作成方法。

[請求項4] 前記複数のハイブリッドモデル候補を作成する際、

当該ハイブリッドモデル候補を構成するために選択された前記２つ以上のモデルのそれぞれに複数の検証用データセットを入力してカテ

ゴリを推定させることで得た複数の出力値のうち、不良品であると推定された出力値をプロットしたときの凸包を算出し、

前記複数の出力値から、前記凸包の頂点を除き前記凸包に含まれる出力値を除外し、

前記凸包の頂点を除き前記凸包に含まれる出力値が除外された前記複数の出力値を入力して用いて、かつ、前記複数の出力値に対応する検証用データセットの真の値を用いて機械学習を行うことで、前記複数のハイブリッドモデル候補を作成する、

請求項２に記載のハイブリッドモデル作成方法。

[請求項5] 前記２つ以上のモデルを選択する前に、プールされている複数のモデルそれぞれに、複数の検証用データセットを入力してカテゴリを推定させることでプールされている前記複数のモデルそれぞれの推定精度を取得し、

前記推定精度が閾値以下のモデルを、プールされている前記複数のモデルの中から除外し、

前記閾値以下のモデルが除外された前記複数のモデルから、前記２つ以上のモデルを選択する、

請求項１～４のいずれか１項に記載のハイブリッドモデル作成方法。

[請求項6] 前記２つ以上のモデルを選択する前に、プールされている複数のモデルそれぞれに、複数の検証用データセットを入力してカテゴリを推定させることでプールされている前記複数のモデルそれぞれの推定結果を取得し、

前記推定結果を用いて、プールされている前記複数のモデルのすべての相関を算出し、

他のすべてのモデルとの相関が閾値より強いモデルをプールされている前記複数のモデルの中から除外し、

前記閾値より強いモデルが除外された後において前記複数のモデル

から、前記２つ以上のモデルを選択する、

請求項１～４のいずれか１項に記載のハイブリッドモデル作成方法。

[請求項7] 前記ハイブリッドモデル候補を複数作成する前に、プールされているまたは選択された複数のモデルそれぞれに、複数の検証用データセットを入力してカテゴリを推定させることで当該複数のモデルそれぞれの推定結果を取得し、

前記推定結果を用いて、当該複数のモデルのすべての相関を算出し、

前記ハイブリッドモデル候補を複数作成する際に、閾値より強い相関がある２つのモデルの組み合わせを含めないように、選択された前記２つ以上のモデルを組み合わせることで、前記ハイブリッドモデル候補を複数作成する、

請求項１～４のいずれか１項に記載のハイブリッドモデル作成方法。

[請求項8] 深層学習されたモデルでは、前記推定結果は、前記深層学習されたモデルの中間層または最終層の出力結果である、

請求項６または７に記載のハイブリッドモデル作成方法。

[請求項9] 前記複数のハイブリッドモデル候補それぞれは、当該ハイブリッドモデル候補を構成するために選択された前記２つ以上のモデルのそれぞれに複数の検証用データセットを入力してカテゴリを推定させることで得た２つ以上の出力結果を入力として、前記複数の検証用データセットのカテゴリを判定した判定結果を出力させる機械学習モデルであり、

前記複数のハイブリッドモデル候補に出力させた判定結果を比較することで、前記複数のハイブリッドモデル候補のうちの１つをハイブリッドモデルとして選択する、

請求項１～８のいずれか１項に記載のハイブリッドモデル作成方法

。

[請求項10] 前記複数のハイブリッドモデル候補を比較する際、

前記複数のハイブリッドモデル候補それぞれに出力させた判定結果から、当該ハイブリッドモデル候補を構成するために選択された前記２つ以上のモデルのそれぞれの重要度を算出し、

算出された重要度のうち予め設定されていた閾値を下回った重要度のモデルを通知する、

請求項９に記載のハイブリッドモデル作成方法。

[請求項11] 前記複数のハイブリッドモデル候補を比較する際、

前記複数のハイブリッドモデル候補それぞれに出力された判定結果から、当該ハイブリッドモデル候補を構成するために選択された前記２つ以上のモデルのそれぞれの重要度を算出し、

前記ハイブリッドモデルとして選択する際、

算出された重要度のうち予め設定されていた閾値を下回った重要度のモデルを有するハイブリッドモデル候補を除いた前記複数のハイブリッドモデル候補のうちの１つを、ハイブリッドモデルとして選択する、

請求項９に記載のハイブリッドモデル作成方法。

[請求項12] さらに、前記複数のハイブリッドモデル候補を作成するために選択された複数のモデルのそれぞれにおいて、検証用データセットが入力されて前記検証用データセットのカテゴリを推定するまでに要した処理時間を取得し、

取得した前記処理時間に基づき、前記複数のモデルのすべての前記処理時間の和に対する前記複数のモデルのそれぞれの当該処理時間の値をハードウェアコストと定義し、

前記複数のハイブリッドモデル候補を作成する際、

前記複数のハイブリッドモデル候補それぞれの機械学習モデルの損失関数に、当該ハイブリッドモデル候補を構成するために選択された

前記２つ以上のモデルのそれぞれのハードウェアコストを加味した正則化項を追加する、

請求項９～１１のいずれか１項に記載のハイブリッドモデル作成方法。

[請求項13] さらに、前記複数のハイブリッドモデル候補を作成するために選択された複数のモデルのそれぞれにおいて、前記入力されるデータとして、検証用データセットのうち不良品を示す複数のデータを入力してカテゴリを推定させることで得た出力値の分布に基づいて、閾値を変動させたときの見逃し率の表であるＦＡＲ表を作成し、

前記複数のハイブリッドモデル候補を比較する際、

当該ハイブリッドモデル候補を構成するために選択された前記２つ以上のモデルのそれぞれに、検証用データセットに含まれるデータサンプルを入力してカテゴリを推定させることで得た出力値と、前記ＦＡＲ表とを照合することで、前記データサンプルに対する前記２つ以上のモデルのそれぞれの第１ＦＡＲ値を取得し、

取得した前記２つ以上のモデルのそれぞれの第１ＦＡＲ値を乗算することで、前記ハイブリッドモデル候補の第２ＦＡＲ値を取得し、

前記第２ＦＡＲ値が事前に設定された閾値より小さい場合に、前記データサンプルが良品であるとの判定結果を、複数のハイブリッドモデル候補それぞれに出力させた判定結果として取得して、比較する、

請求項２に記載のハイブリッドモデル作成方法。

[請求項14] さらに、前記複数のハイブリッドモデル候補を作成するために選択された複数のモデルのそれぞれにおいて、複数の検証用データセットを入力してカテゴリを推定させることで前記選択された複数のモデルそれぞれの推定結果を取得し、

前記推定結果を用いて、前記選択された複数のモデルのうち２つのモデルの組み合わせすべての相関係数を算出し、

前記第２ＦＡＲ値を取得する際、

取得した前記２つ以上のモデルのそれぞれの第１ＦＡＲ値を乗算して、さらに、前記相関係数が大きいほど小さくなる係数を乗算することで、前記第２ＦＡＲ値を取得する、

請求項１３に記載のハイブリッドモデル作成方法。

[請求項15] 入力されるデータのカテゴリを推定する複数のモデルをプールするモデルプール部と、

前記複数のモデルの少なくとも一つのモデルは、機械学習されたモデルであり、プールされている複数のモデルから２つ以上のモデルを選択するモデル選択部と、

選択された前記２つ以上のモデルを組み合わせることで、前記カテゴリを判定するハイブリッドモデル候補を複数作成し、複数の前記ハイブリッドモデル候補を比較するハイブリッドモデル候補作成部と、

前記複数のハイブリッドモデル候補のうちの１つをハイブリッドモデルとして選択するハイブリッド選択部とを備える、

ハイブリッドモデル作成装置。

[請求項16] 入力されるデータのカテゴリを推定する複数のモデルをプールし、

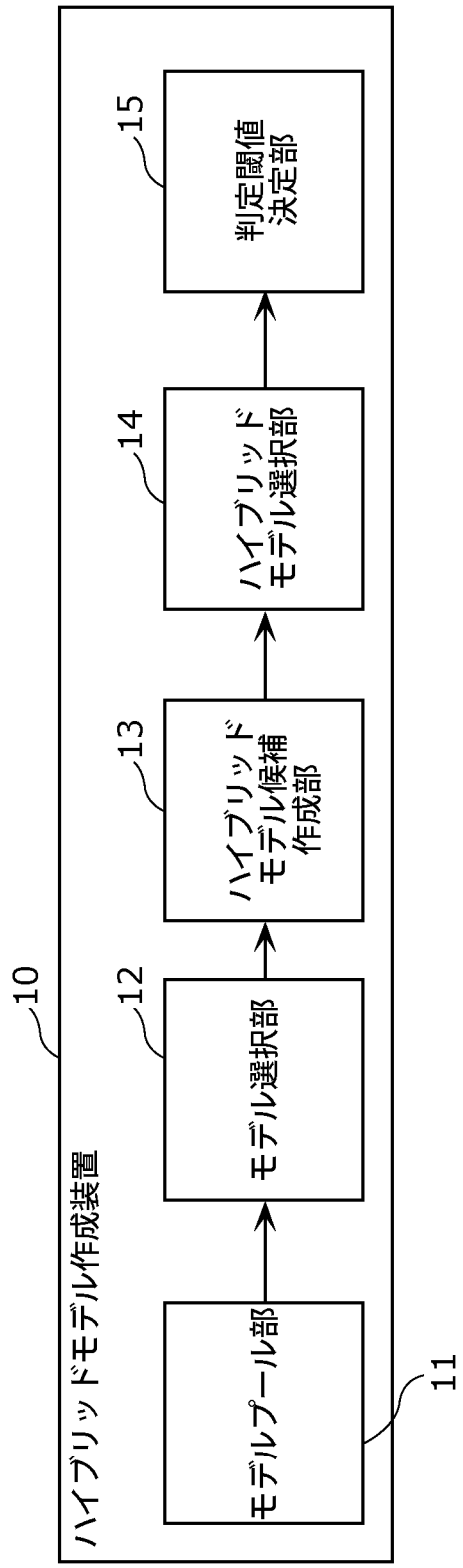
前記複数のモデルの少なくとも一つのモデルは、機械学習されたモデルであり、

プールされている複数のモデルから２つ以上のモデルを選択して組み合わせることで、前記カテゴリを判定するハイブリッドモデル候補を複数作成し、

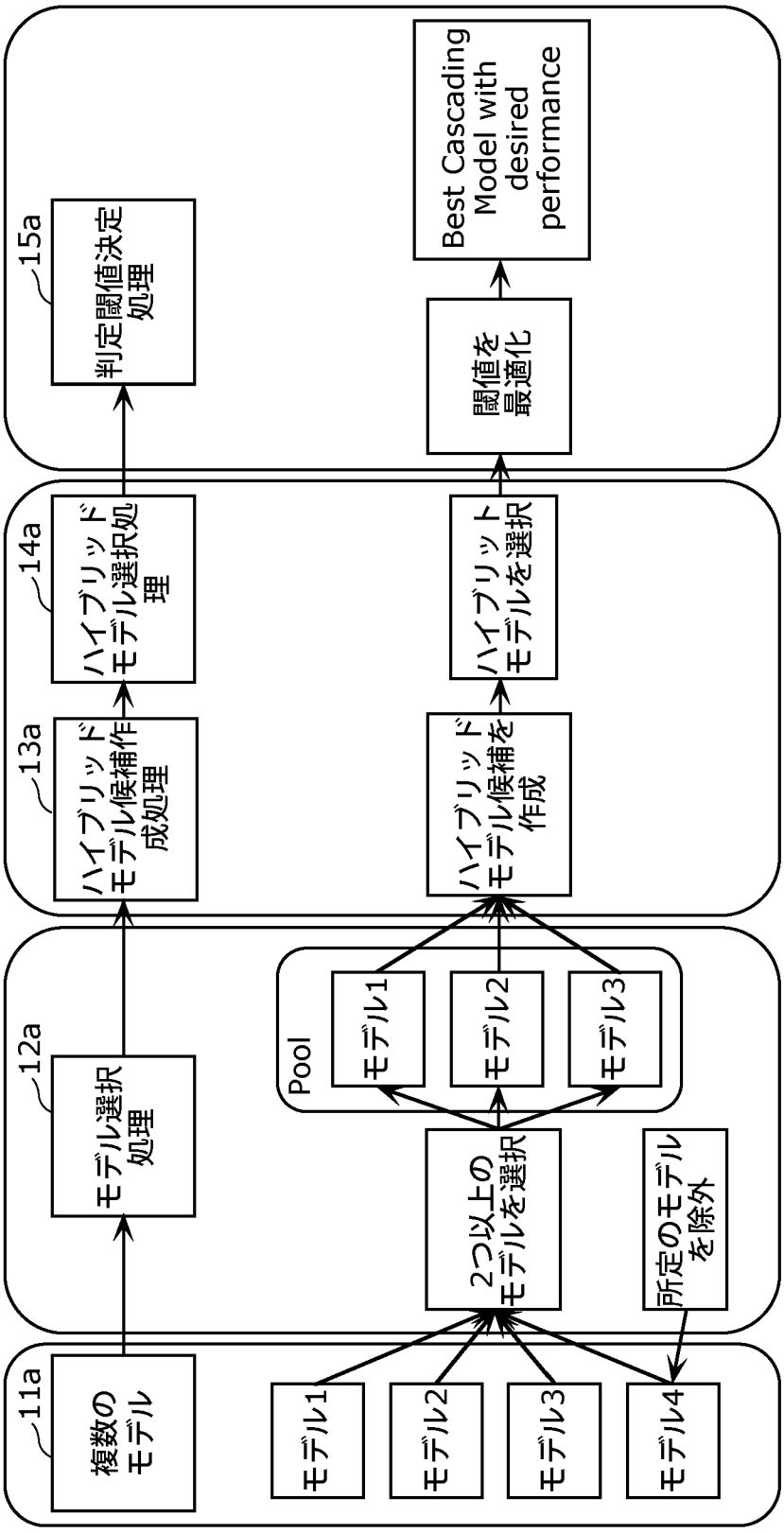
複数の前記ハイブリッドモデル候補を比較することで、前記複数のハイブリッドモデル候補のうちの１つをハイブリッドモデルとして選択することを、

コンピュータに実行させるプログラム。

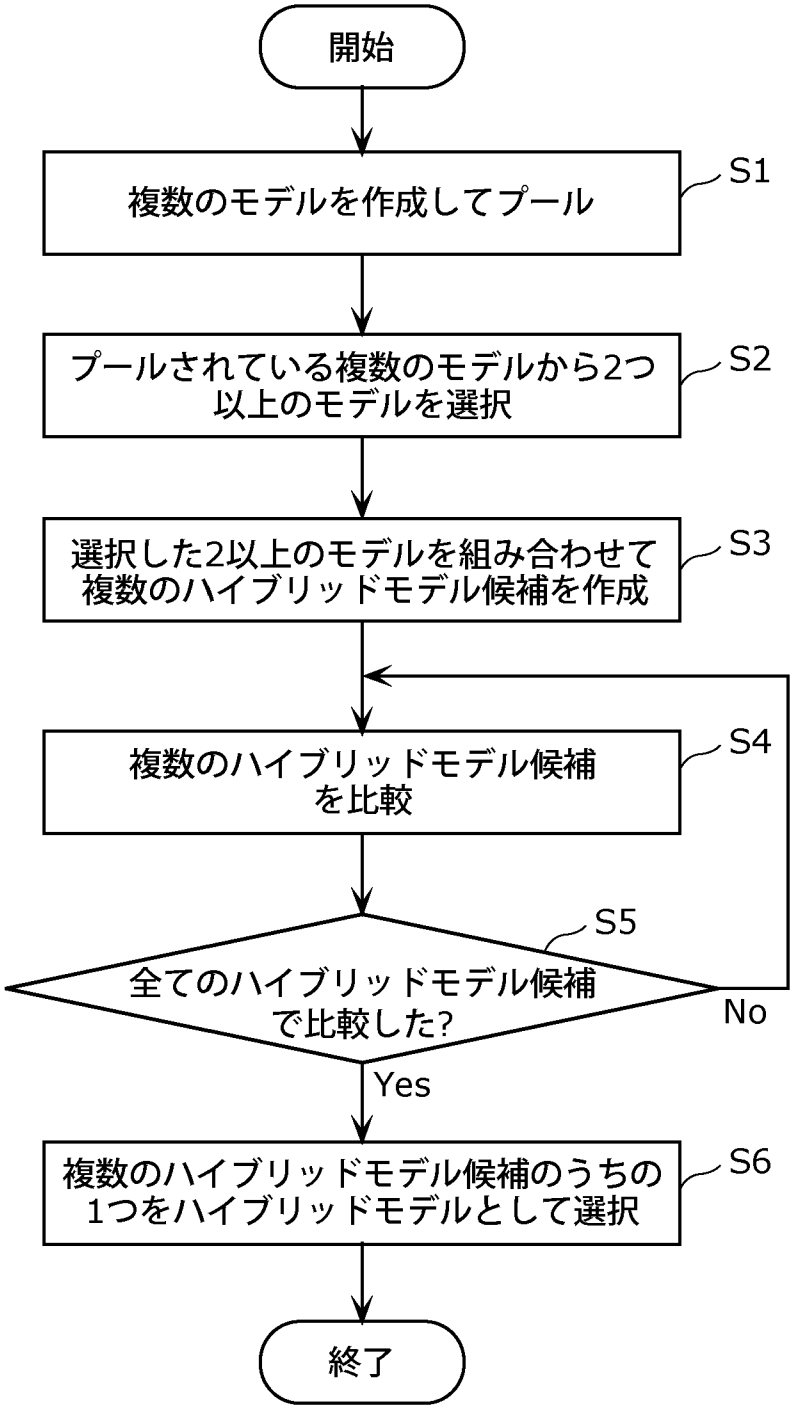
[図1]


10
ハイブリッドモデル作成装置
11
モデルプール部
12
モデル選択部
13
ハイブリッドモデル候補作成部
14
ハイブリッドモデル選択部
15
判定閾値決定部

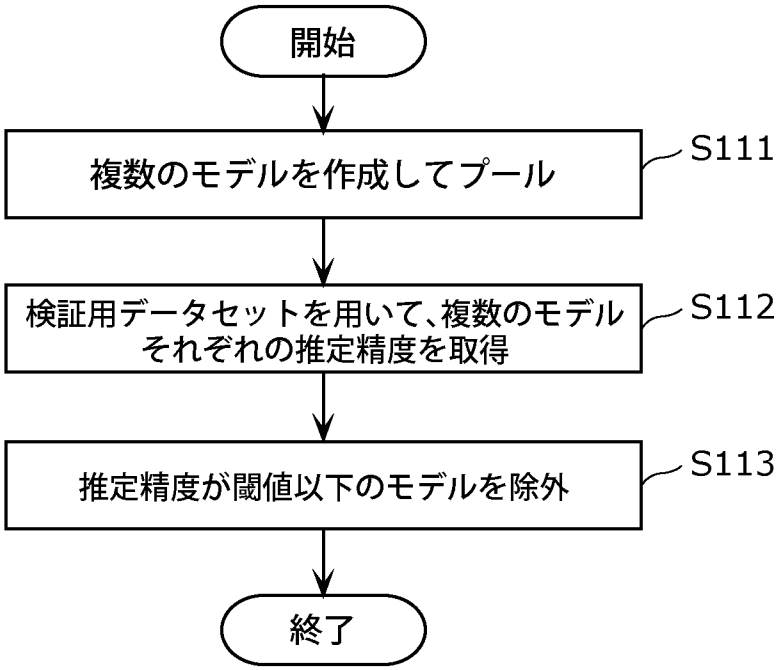
[図2]


11a
複数の
モデル
モデル1
モデル2
モデル3
モデル4
12a
モデル選択
処理
2つ以上の
モデルを選択
所定のモデル
を除外
Pool
モデル1
モデル2
モデル3
13a
ハイブリッド
モデル候補作
成処理
ハイブリッド
モデル候補を
作成
14a
ハイブリッド
モデル選択処
理
ハイブリッド
モデルを選択
15a
判定閾値決定
処理
閾値を
最適化
Best Cascading
Model with
desired
performance

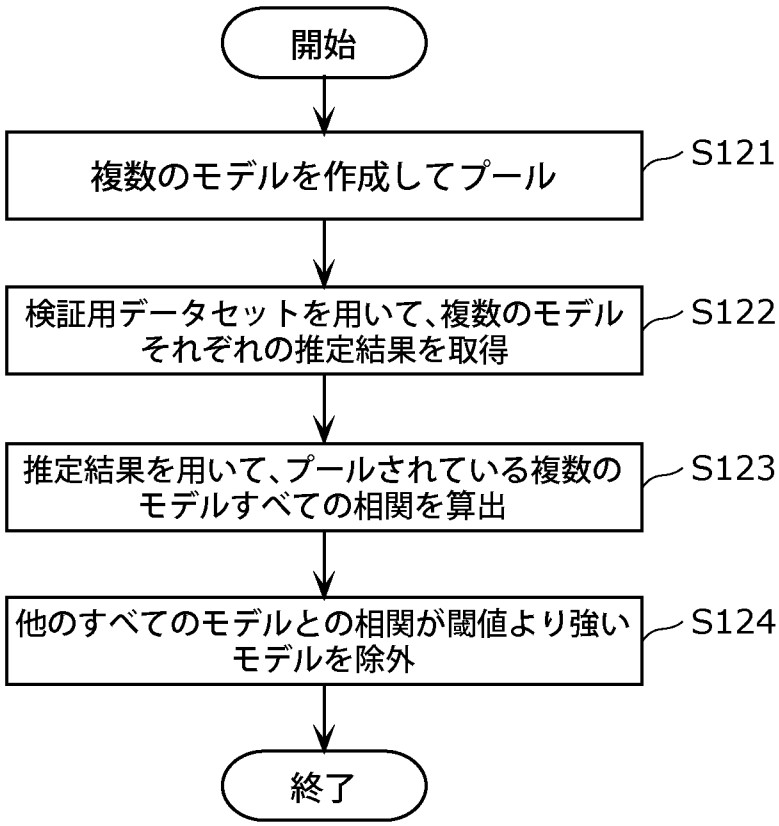
[図3]


開始
複数のモデルを作成してプール
S1
プールされている複数のモデルから2つ以上のモデルを選択
S2
選択した2以上のモデルを組み合わせて複数のハイブリッドモデル候補を作成
S3
複数のハイブリッドモデル候補を比較
S4
全てのハイブリッドモデル候補で比較した?
S5
No
Yes
複数のハイブリッドモデル候補のうちの1つをハイブリッドモデルとして選択
S6
終了

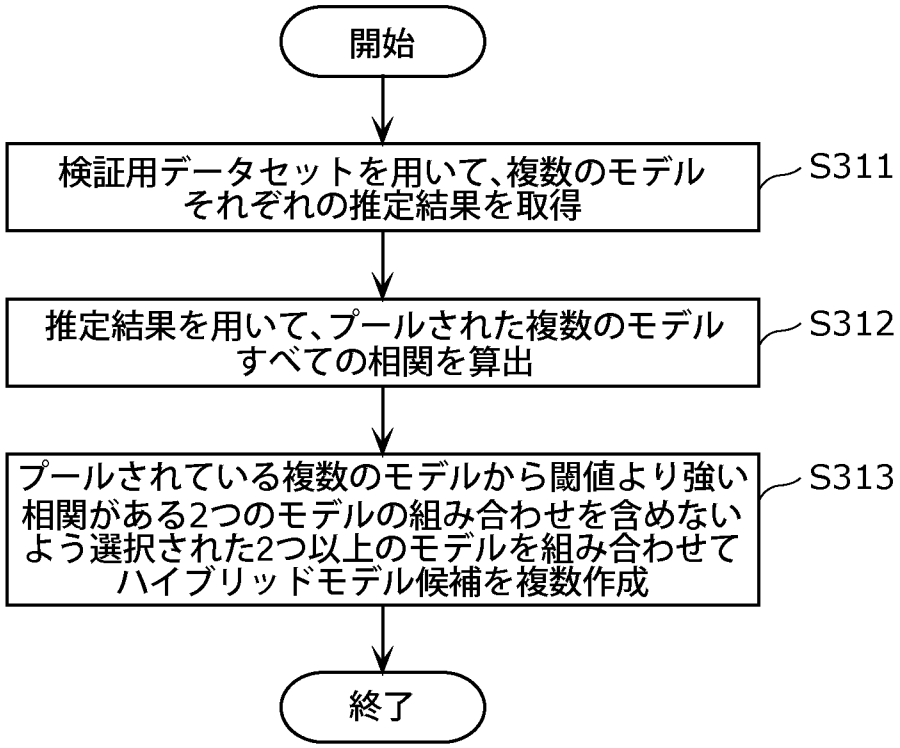
[図4]


開始
複数のモデルを作成してプール
S111
検証用データセットを用いて、複数のモデルそれぞれの推定精度を取得
S112
推定精度が閾値以下のモデルを除外
S113
終了

[図5]


開始
複数のモデルを作成してプール
S121
検証用データセットを用いて、複数のモデルそれぞれの推定結果を取得
S122
推定結果を用いて、プールされている複数のモデルすべての相関を算出
S123
他のすべてのモデルとの相関が閾値より強いモデルを除外
S124
終了

[図6]



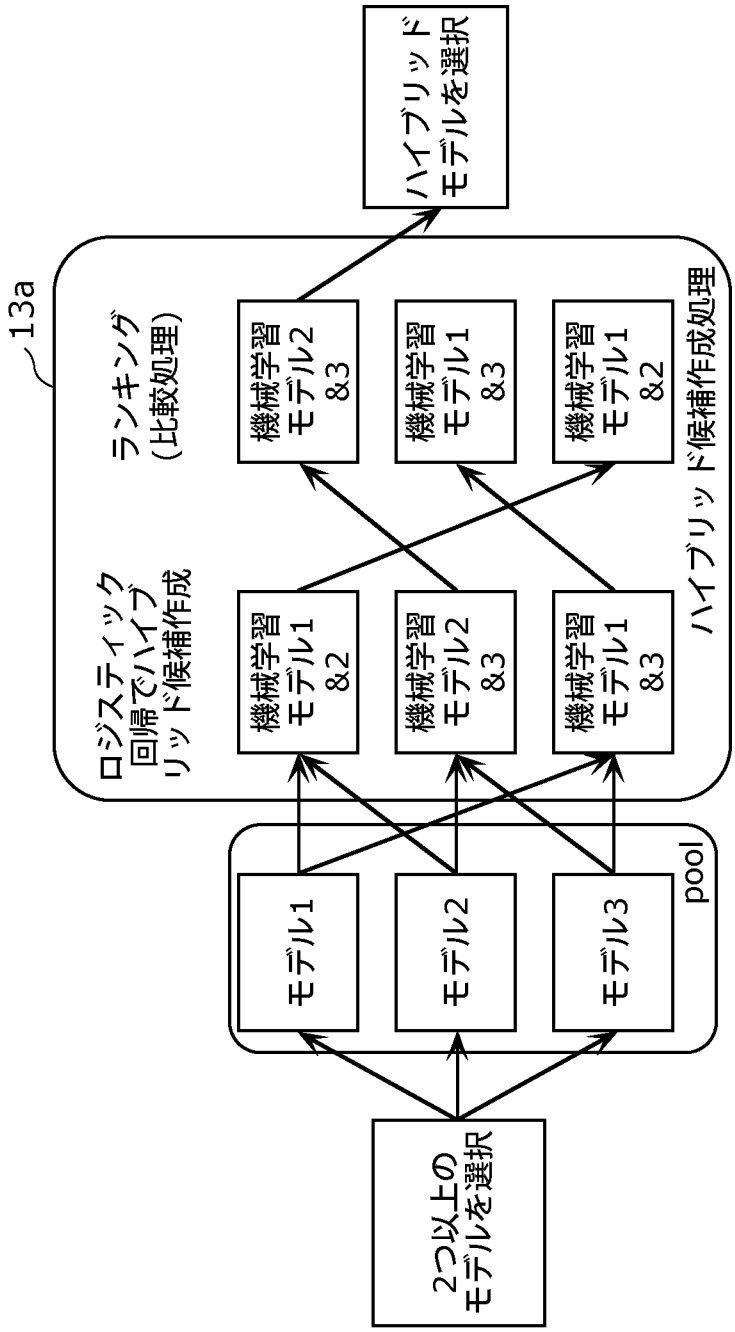
[図7A]

真の値	-> 1 1 1 1 1 1 1 1 1 1	
モデル1	-> 0 0 1 1 1 1 1 1 1 1	=> 80% accuracy
モデル2	-> 0 0 1 0 1 1 1 1 1 1	=> 70% accuracy
モデル3	-> 0 0 1 1 1 1 1 1 1 1	=> 80% accuracy
ハイブリッドモデル候補 (多数決の場合)	-> 0 0 1 1 1 1 1 1 1 1	=> 80% accuracy

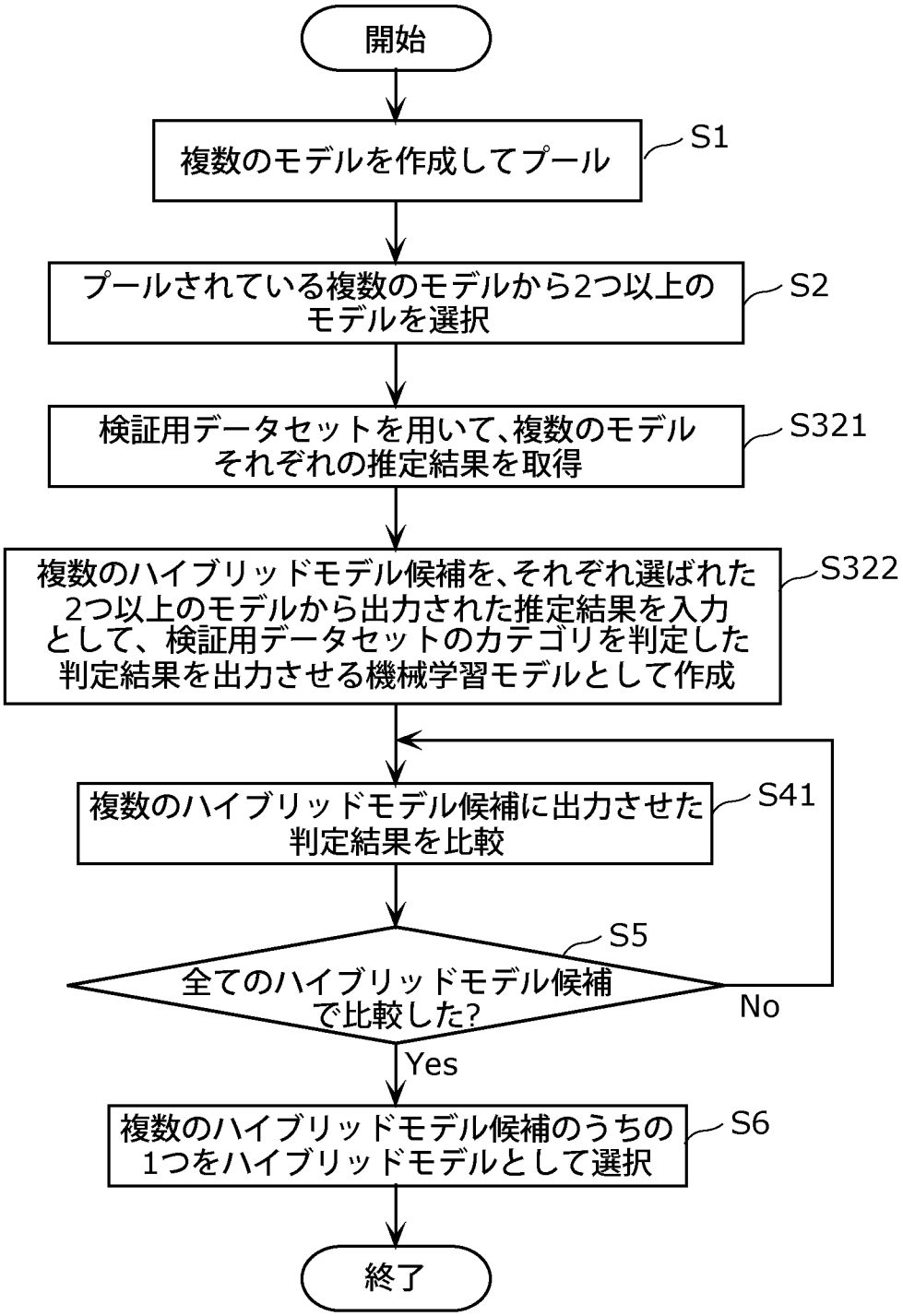
[図7B]

真の値	-> 1 1 1 1 1 1 1 1 1 1	
モデル1	-> 0 0 1 1 1 1 1 1 1 1	=> 80% accuracy
モデル2	-> 0 1 0 1 1 1 0 0 1 1	=> 60% accuracy
モデル3	-> 1 1 1 0 0 0 1 1 0 0	=> 50% accuracy
ハイブリッドモデル候補 (多数決の場合)	-> 0 1 1 1 1 1 1 1 1 1	=> 90% accuracy

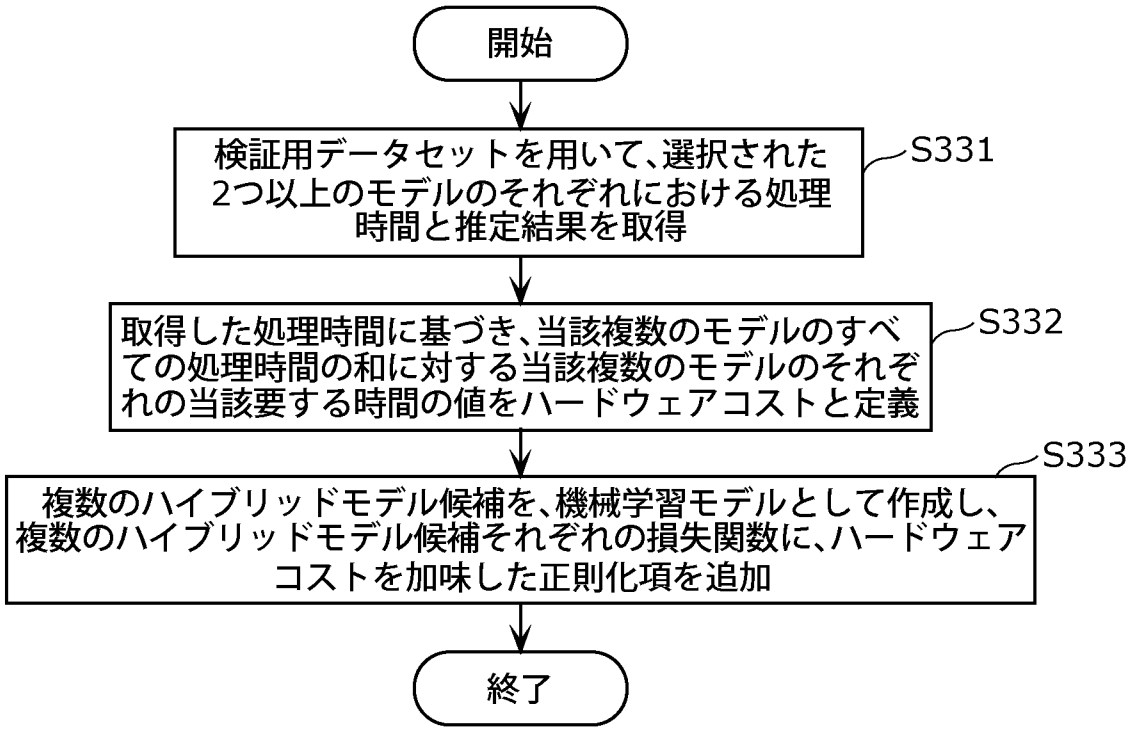
[図8]


2つ以上の
モデルを選択
pool
モデル1
モデル2
モデル3
13a
ロジスティック
回帰でハイブ
リッド候補作成
ランキング
(比較処理)
ハイブリッド候補作成処理
機械学習
モデル1
&2
機械学習
モデル2
&3
機械学習
モデル1
&3
機械学習
モデル2
&3
機械学習
モデル1
&3
機械学習
モデル1
&2
ハイブリッド
モデルを選択

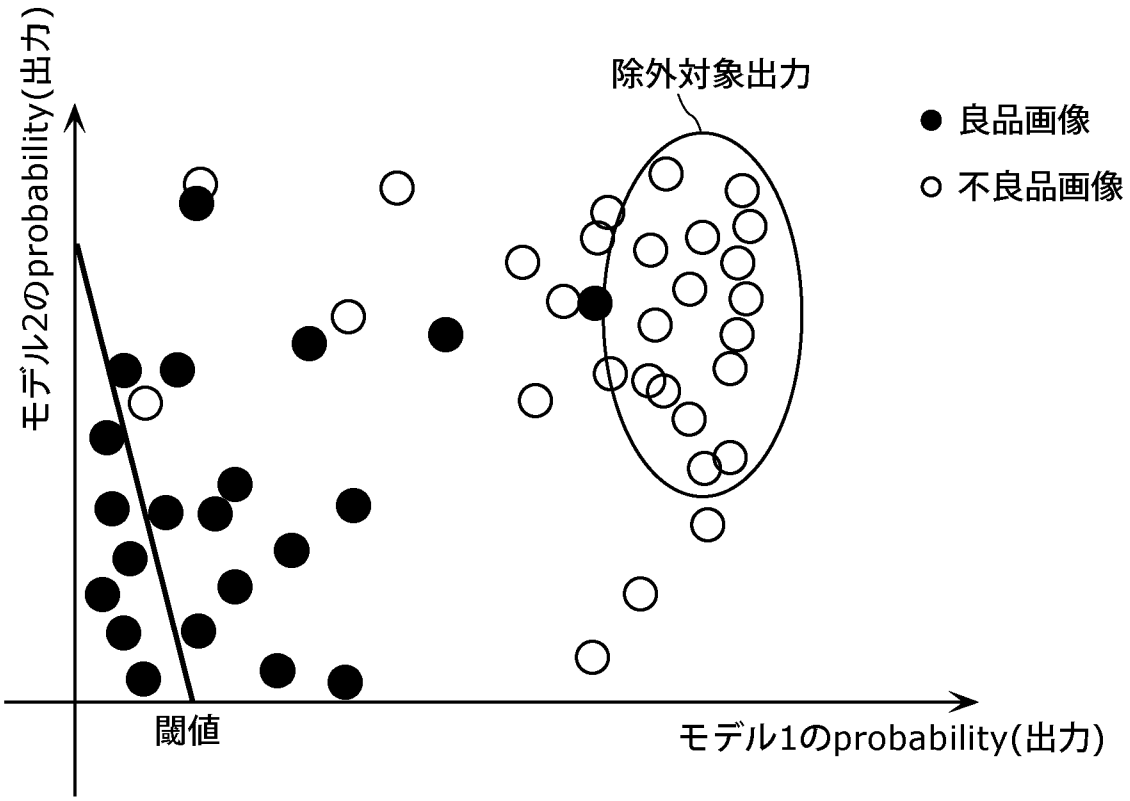
[図9]


開始
複数のモデルを作成してプール
S1
プールされている複数のモデルから2つ以上のモデルを選択
S2
検証用データセットを用いて、複数のモデルそれぞれの推定結果を取得
S321
複数のハイブリッドモデル候補を、それぞれ選ばれた2つ以上のモデルから出力された推定結果を入力として、検証用データセットのカテゴリを判定した判定結果を出力させる機械学習モデルとして作成
S322
複数のハイブリッドモデル候補に出力させた判定結果を比較
S41
全てのハイブリッドモデル候補で比較した?
S5
No
Yes
複数のハイブリッドモデル候補のうちの1つをハイブリッドモデルとして選択
S6
終了

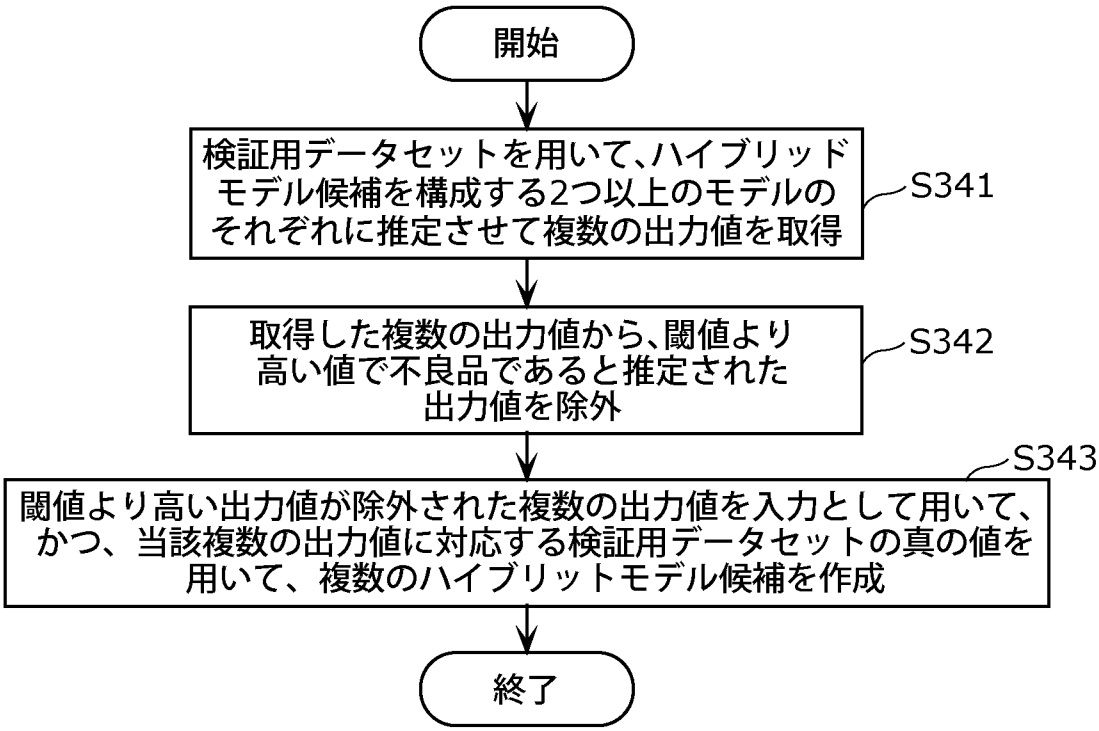
[図10]


開始
検証用データセットを用いて、選択された2つ以上のモデルのそれぞれにおける処理時間と推定結果を取得
S331
取得した処理時間に基づき、当該複数のモデルのすべての処理時間の和に対する当該複数のモデルのそれぞれの当該要する時間の値をハードウェアコストと定義
S332
複数のハイブリッドモデル候補を、機械学習モデルとして作成し、複数のハイブリッドモデル候補それぞれの損失関数に、ハードウェアコストを加味した正則化項を追加
S333
終了

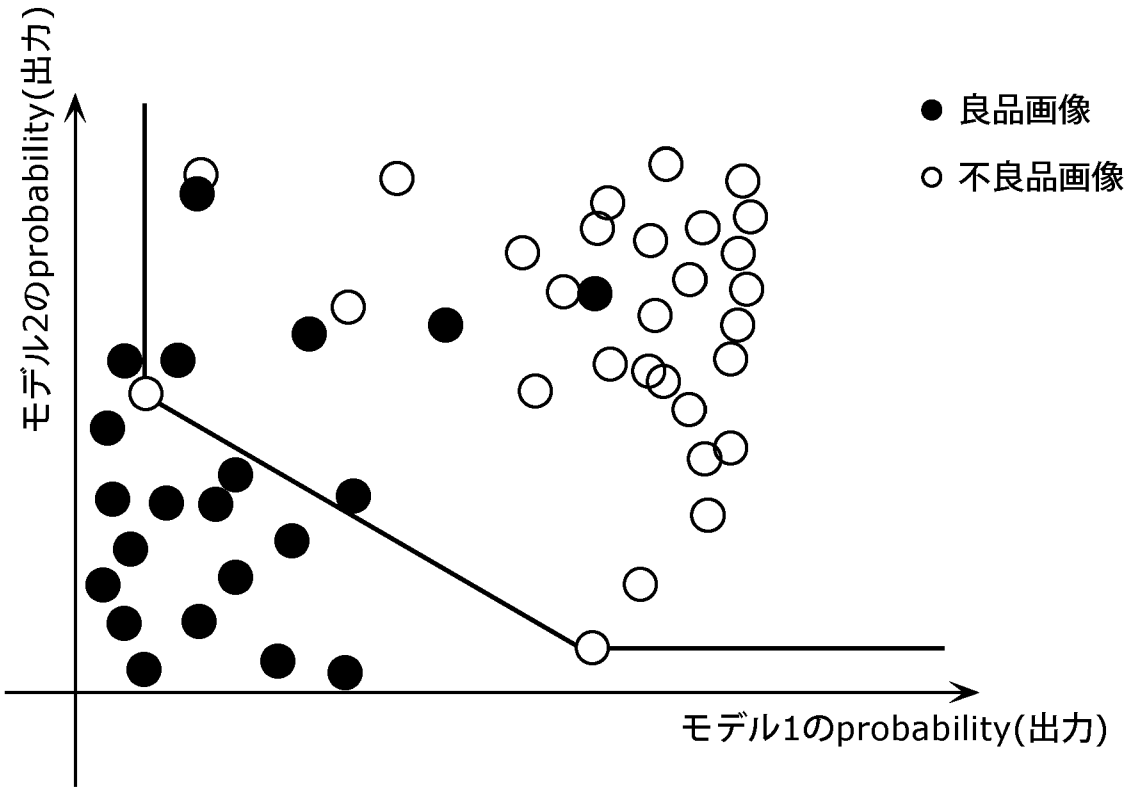
[図11]


モデル2のprobability(出力)
除外対象出力
良品画像
不良品画像
閾値
モデル1のprobability(出力)

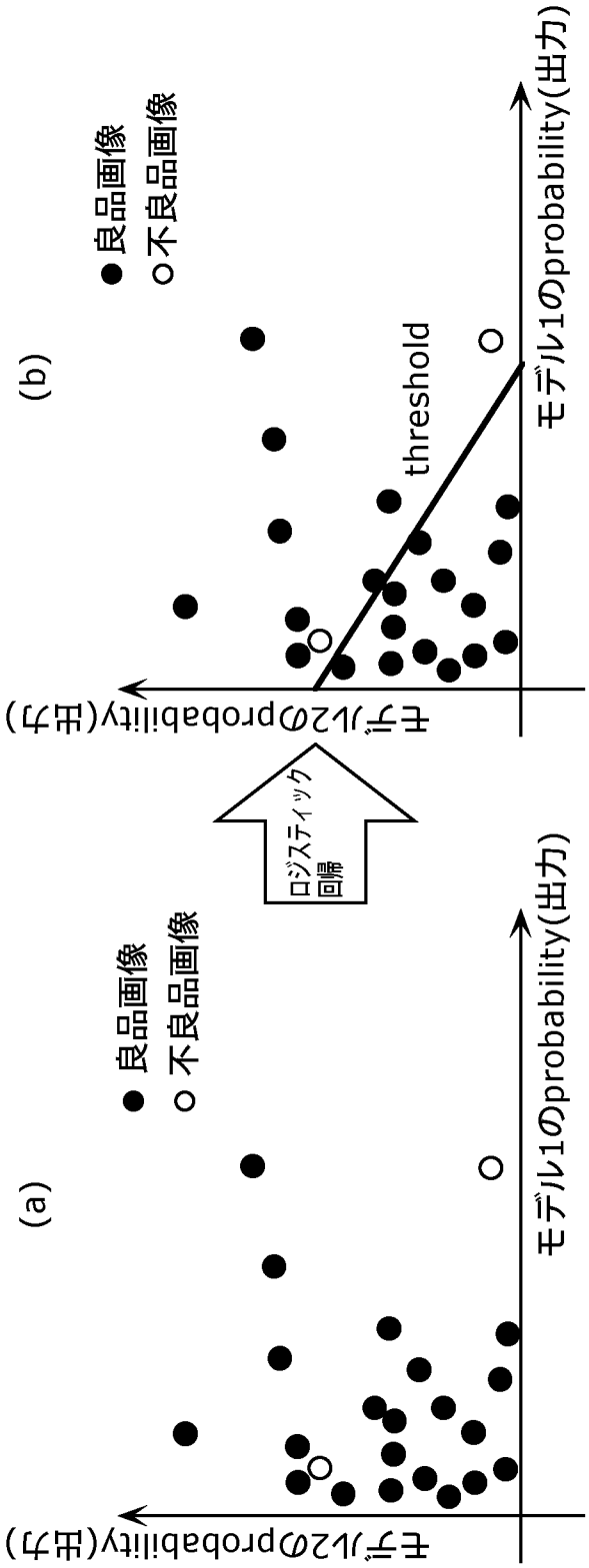
[図12]


開始
検証用データセットを用いて、ハイブリッドモデル候補を構成する2つ以上のモデルのそれぞれに推定させて複数の出力値を取得
S341
取得した複数の出力値から、閾値より高い値で不良品であると推定された出力値を除外
S342
S343
閾値より高い出力値が除外された複数の出力値を入力として用いて、かつ、当該複数の出力値に対応する検証用データセットの真の値を用いて、複数のハイブリットモデル候補を作成
終了

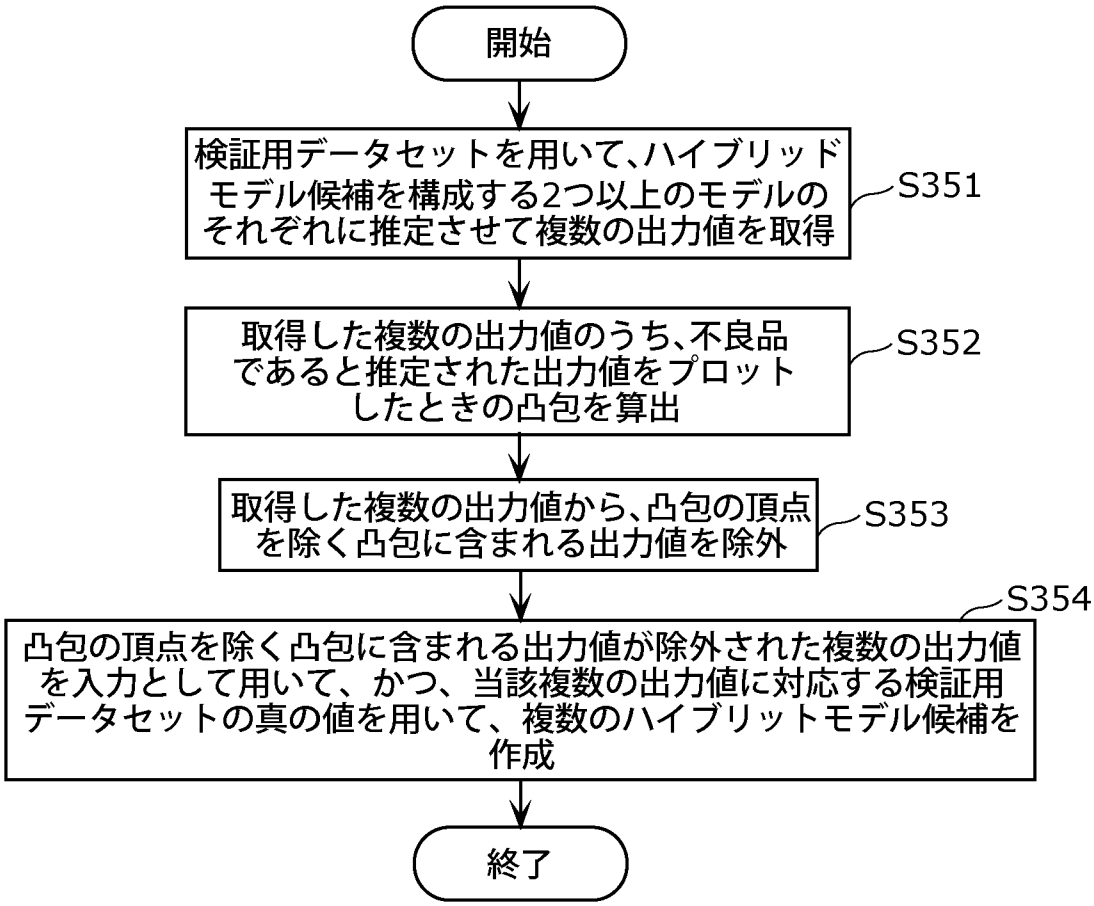
[図13]


モデル2のprobability(出力)
良品画像
不良品画像
モデル1のprobability(出力)

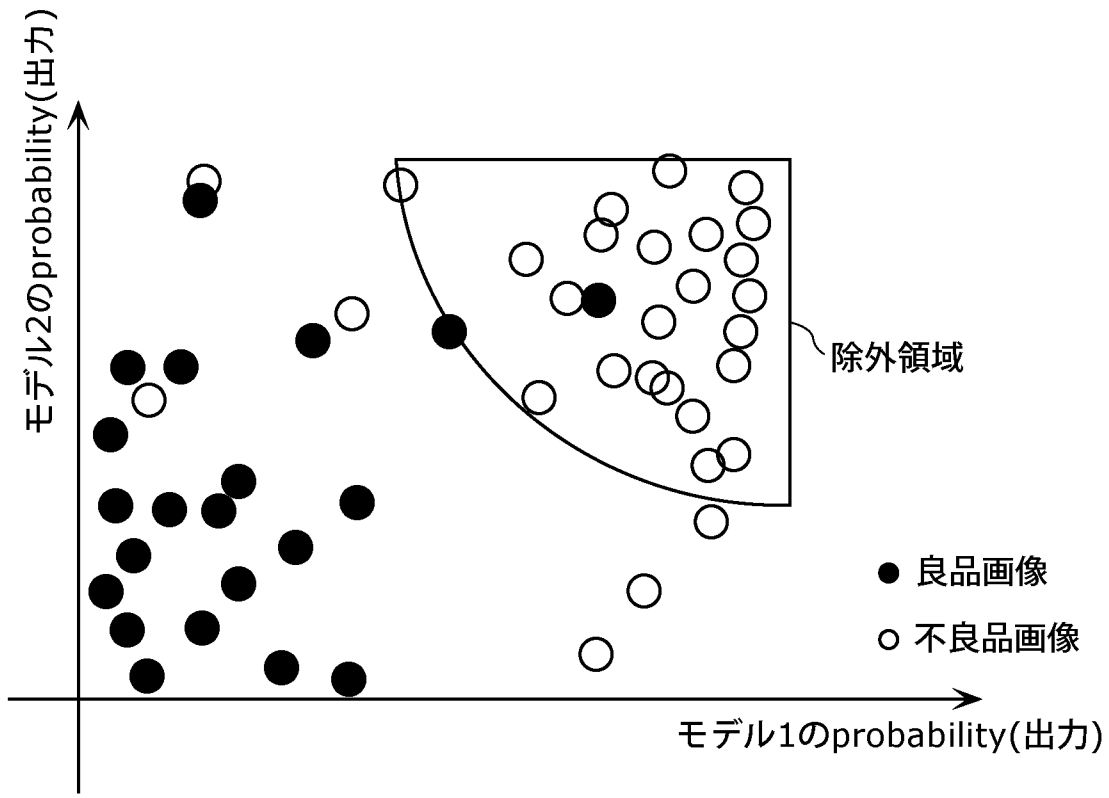
[図14]


(a)
● 良品画像
○ 不良品画像
モデル1のprobability(出力)
モデル2のprobability(出力)
ロジスティック回帰
(b)
● 良品画像
○ 不良品画像
threshold
モデル1のprobability(出力)
モデル2のprobability(出力)

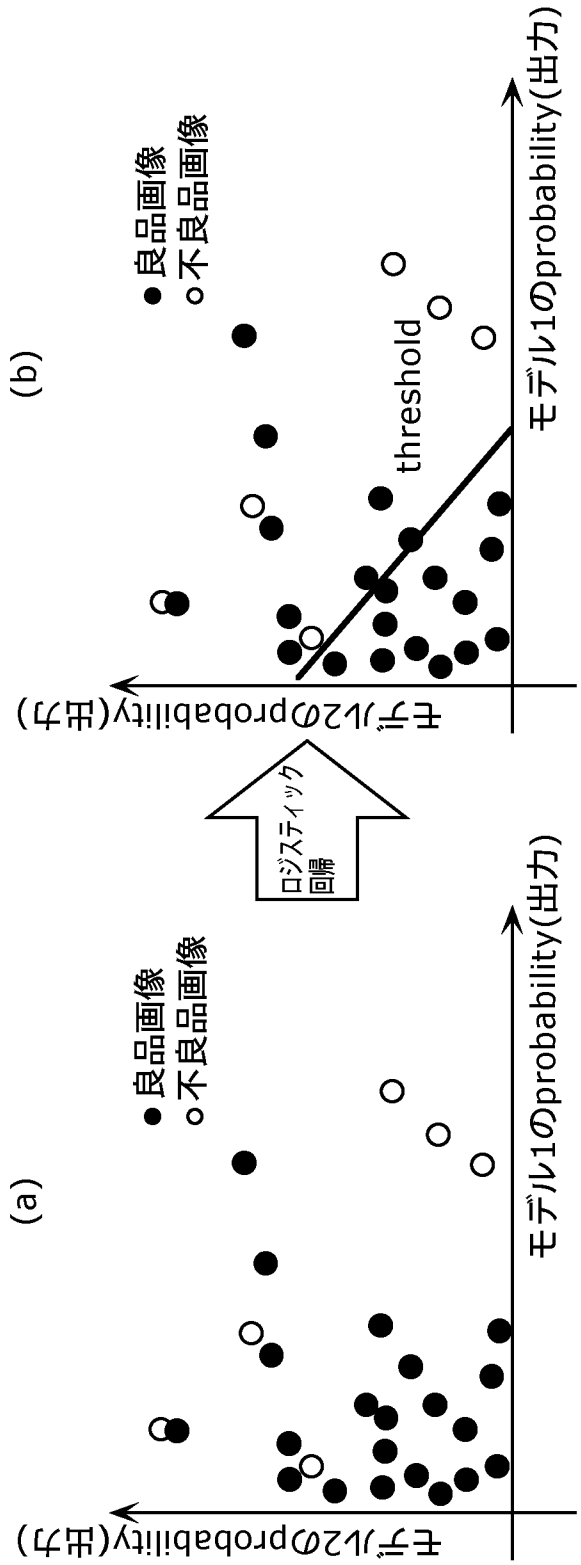
[図15]


開始
検証用データセットを用いて、ハイブリッドモデル候補を構成する2つ以上のモデルのそれぞれに推定させて複数の出力値を取得
S351
取得した複数の出力値のうち、不良品であると推定された出力値をプロットしたときの凸包を算出
S352
取得した複数の出力値から、凸包の頂点を除く凸包に含まれる出力値を除外
S353
S354
凸包の頂点を除く凸包に含まれる出力値が除外された複数の出力値を入力として用いて、かつ、当該複数の出力値に対応する検証用データセットの真の値を用いて、複数のハイブリットモデル候補を作成
終了

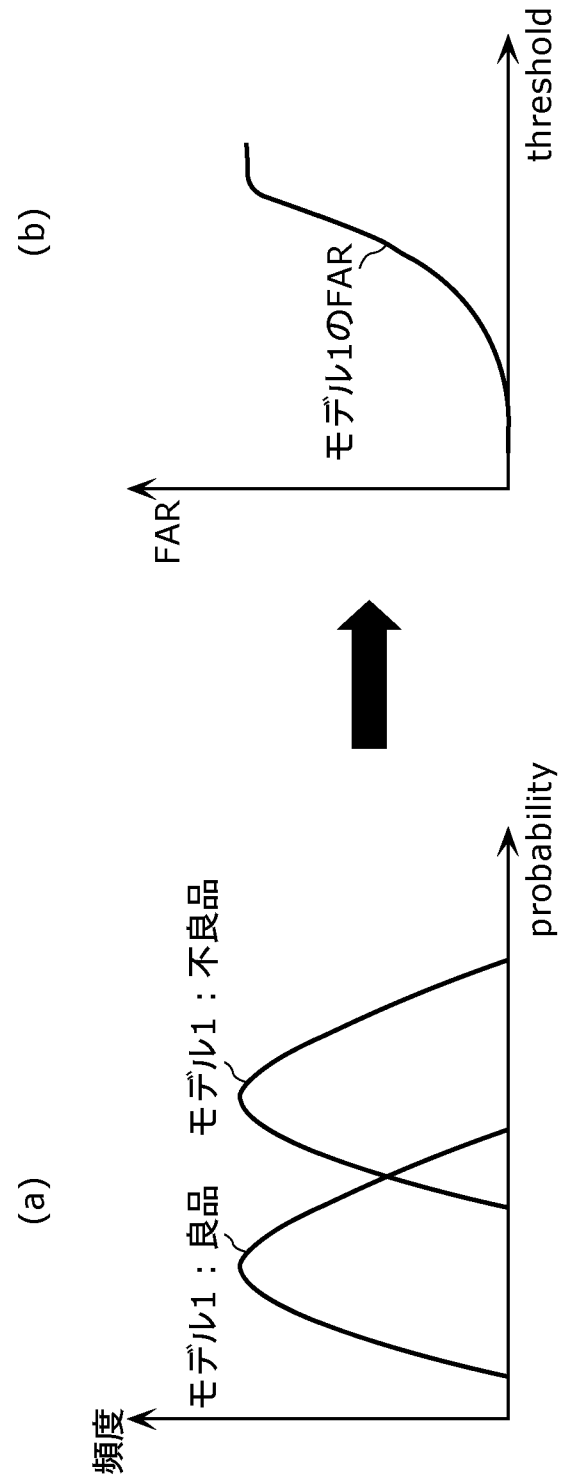
[図16]


モデル2のprobability(出力)
除外領域
良品画像
不良品画像
モデル1のprobability(出力)

[図17]


(a)
●良品画像
○不良品画像
モデル1のprobability(出力)
モデル2のprobability(出力)
ロジスティック回帰
(b)
●良品画像
○不良品画像
threshold
モデル1のprobability(出力)
モデル2のprobability(出力)

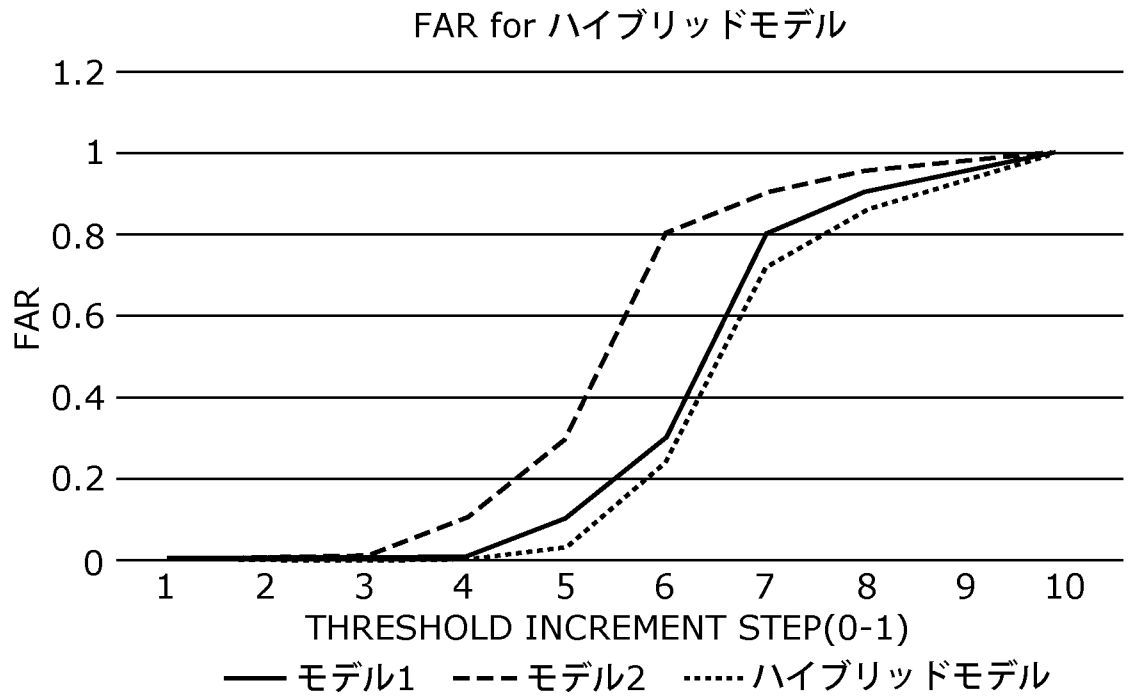
[図18]



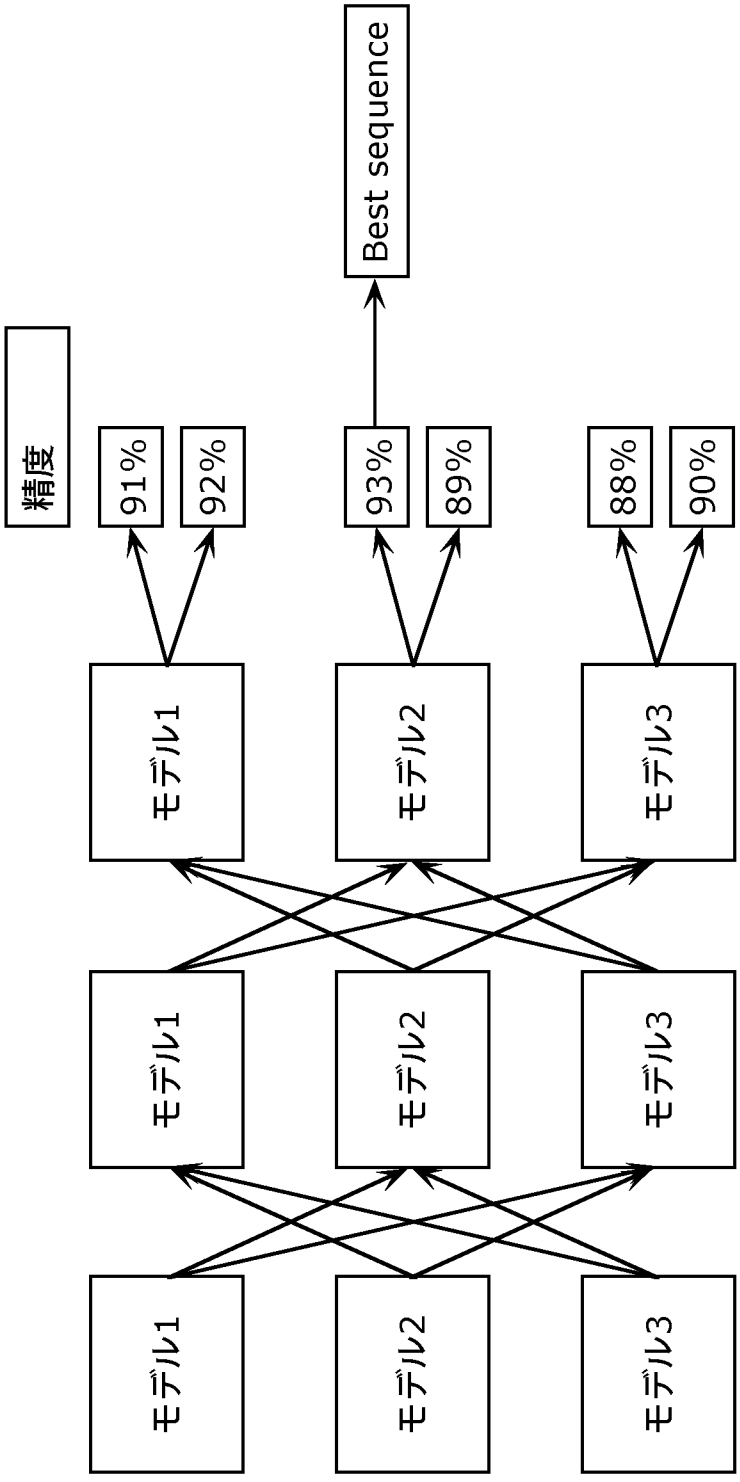
[図19]

Index	Threshold	FAR
1	0	0.000001
2	0.0078125	0.000023
3	0.015625	0.0005
---	---	---
64	0.5	0.5
---	---	---
128	1.0	1.0

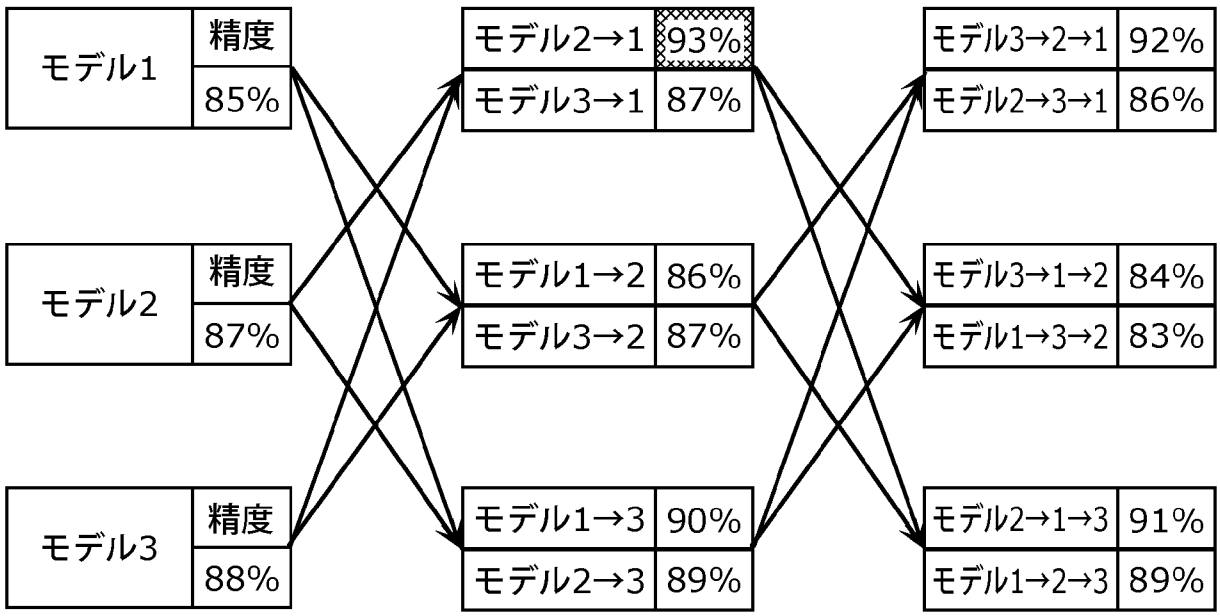
[図20]



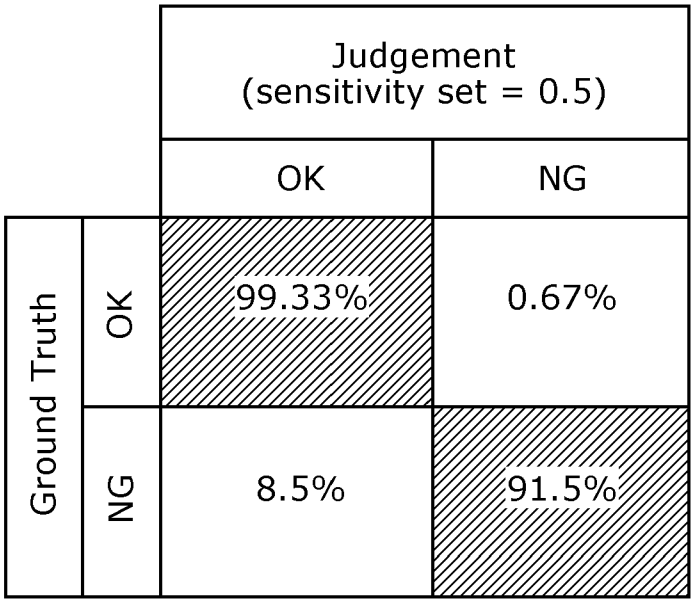
[図21]


Best sequence
精度
91%
92%
93%
89%
88%
90%
モデル1
モデル2
モデル3
モデル1
モデル2
モデル3
モデル1
モデル2
モデル3

[図22]


モデル1
精度
85%
モデル2
精度
87%
モデル3
精度
88%
モデル2→1
93%
モデル3→1
87%
モデル1→2
86%
モデル3→2
87%
モデル1→3
90%
モデル2→3
89%
モデル3→2→1
92%
モデル2→3→1
86%
モデル3→1→2
84%
モデル1→3→2
83%
モデル2→1→3
91%
モデル1→2→3
89%

[図23A]


Judgement
(sensitivity set = 0.5)
OK
NG
Ground Truth
OK
NG
99.33%
0.67%
8.5%
91.5%

[図23B]

		Judgement (sensitivity set = 0.42)	
		OK	NG
Ground Truth	OK	99.15%	0.85%
	NG	7.07%	92.93%

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2022/014692

A. CLASSIFICATION OF SUBJECT MATTER

G06T 7/00(2017.01)i; ***G06N 20/00***(2019.01)i
FI: 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)

G06T7/00; G06N20/00

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 1971-2022
Registered utility model specifications of Japan 1996-2022
Published registered utility model applications of Japan 1994-2022

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.
X	JP 6801149 B1 (RAKUTEN INC) 16 December 2020 (2020-12-16) paragraphs [0030]-[0031], [0037]-[0039]	1, 15-16
Y		2
A		3-14
Y	JP 2019-15654 A (FANUC LTD) 31 January 2019 (2019-01-31) paragraphs [0002]-[0003]	2
A	US 2021/0081725 A1 (ADVANCED NEW TECHNOLOGIES CO., LTD.) 18 March 2021 (2021-03-18) entire text, all drawings	1-16
A	CN 112085205 A (THE FOURTH PARADIGM (BEIJING) TECH CO LTD) 15 December 2020 (2020-12-15) entire text, all drawings	1-16
A	KR 10-2211852 B1 (LUNIT INC) 03 February 2021 (2021-02-03) entire text, all drawings	1-16

☐ 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
14 June 2022	**28 June 2022**
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/JP2022/014692

Patent document cited in search report	Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
JP 6801149 B1	16 December 2020	EP 3876181 A1 paragraphs [0040]-[0043], [0053]-[0059] WO 2021/149118 A1	
JP 2019-15654 A	31 January 2019	US 2019/0012579 A1 paragraphs [0002]-[0003] DE 102018211212 A1 CN 109242820 A	
US 2021/0081725 A1	18 March 2021	WO 2020/029689 A1 EP 3779723 A1 CN 109165249 A TW 202018533 A	
CN 112085205 A	15 December 2020	WO 2020/249125 A1	
KR 10-2211852 B1	03 February 2021	(Family: none)	

国際調査報告

国際出願番号
PCT/JP2022/014692

A. 発明の属する分野の分類（国際特許分類（ＩＰＣ））
G06T 7/00(2017.01)i; G06N 20/00(2019.01)i
FI: 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））
G06T7/00; G06N20/00

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　１９７１－２０２２年
日本国実用新案登録公報　１９９６－２０２２年
日本国登録実用新案公報　１９９４－２０２２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
X	JP 6801149 B1 (楽天株式会社) 16.12.2020 (2020-12-16) [0030]-[0031], [0037]-[0039]	1, 15-16
Y		2
A		3-14
Y	JP 2019-15654 A (ファナック株式会社) 31.01.2019 (2019-01-31) [0002]-[0003]	2
A	US 2021/0081725 A1 (ADVANCED NEW TECHNOLOGIES CO., LTD.) 18.03.2021 (2021-03-18) 全文, 全図	1-16
A	CN 112085205 A (THE FOURTH PARADIGM (BEIJING) TECH CO LTD) 15.12.2020 (2020-12-15) 全文, 全図	1-16
A	KR 10-2211852 B1 (LUNIT INC) 03.02.2021 (2021-02-03) 全文, 全図	1-16

☐ 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
14.06.2022	28.06.2022

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）
金田　孝之 5B 3144
電話番号 03-3581-1101 内線 3545

国際調査報告
パテントファミリーに関する情報

国際出願番号
PCT/JP2022/014692

引用文献	公表日	パテントファミリー文献	公表日
JP 6801149 B1	16.12.2020	EP 3876181 A1 [0040]-[0043], [0053]-[0059] WO 2021/149118 A1	
JP 2019-15654 A	31.01.2019	US 2019/0012579 A1 [0002]-[0003] DE 102018211212 A1 CN 109242820 A	
US 2021/0081725 A1	18.03.2021	WO 2020/029689 A1 EP 3779723 A1 CN 109165249 A TW 202018533 A	
CN 112085205 A	15.12.2020	WO 2020/249125 A1	
KR 10-2211852 B1	03.02.2021	（ファミリーなし）	