

(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6月23日(23.06.2022)



(10) 国際公開番号

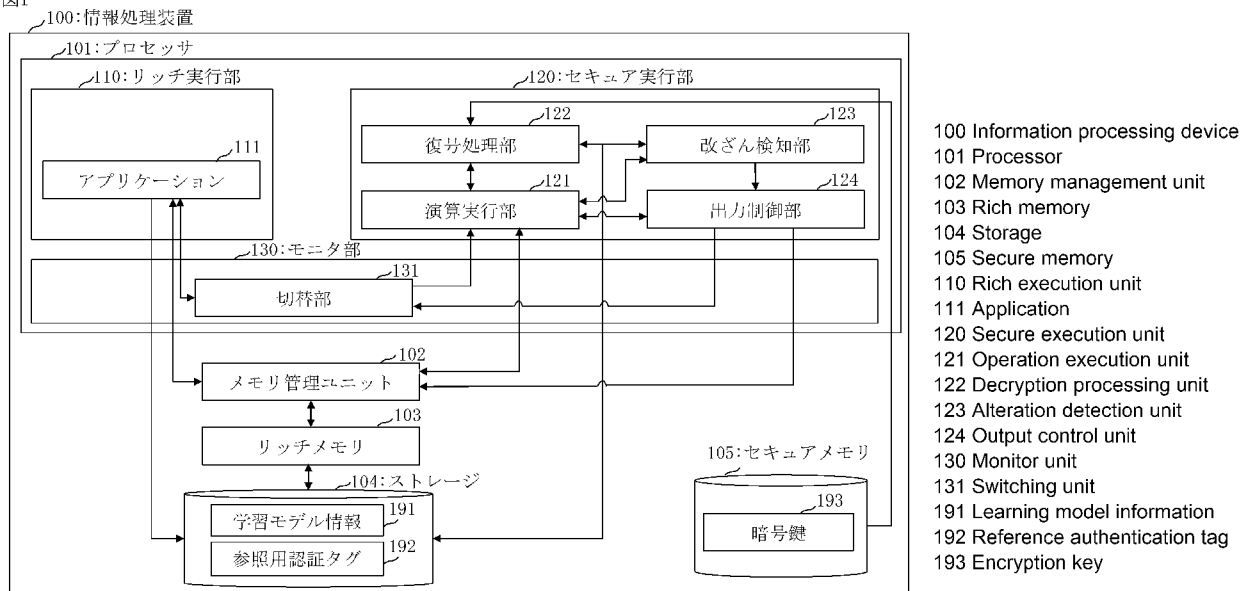
WO 2022/130558 A1

- (51) 国際特許分類:
G06F 21/74 (2013.01) G06N 20/00 (2019.01)
G06F 21/53 (2013.01) H04L 9/32 (2006.01)
G06F 21/64 (2013.01)
- (21) 国際出願番号: PCT/JP2020/047098
- (22) 国際出願日: 2020年12月17日(17.12.2020)
- (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: 三菱電機株式会社(MITSUBISHI ELECTRIC CORPORATION) [JP/JP]; 〒1008310 東京都千代田区丸の内二丁目7番3号 Tokyo (JP).
- (72) 発明者: 中井 綱人(NAKAI, Tsunato); 〒1008310 東京都千代田区丸の内二丁目7番3号 三菱電機株式会社内 Tokyo (JP).
- (74) 代理人: 溝井 国際特許業務法人(MIZOI INTERNATIONAL PATENT FIRM); 〒2470056 神奈川県鎌倉市大船二丁目17番10号3階 Kanagawa (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,

(54) Title: INFORMATION PROCESSING DEVICE, INFORMATION PROCESSING METHOD, AND INFORMATION PROCESSING PROGRAM

(54) 発明の名称: 情報処理装置、情報処理方法および情報処理プログラム

図1



(57) Abstract: A rich execution unit (110) executes an application (111) while executing a rich OS. A secure execution unit (120) executes a target operation, which is an operation based on a learning model, while running a secure OS. When the application requests execution of the target operation, a decryption processing unit (122) decrypts encrypted model information of the learning model into decrypted model information. An operation execution unit (121) uses this model information to execute the target operation. An alteration detection unit (123) determines whether the encrypted model information is altered information. If it is determined that the encrypted model information is not altered information, an output control unit (124) passes the execution result of the target operation to the application.

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:

— 国際調査報告 (条約第21条(3))

(57) 要約: リッチ実行部 (1 1 0) は、リッチOSを実行しながらアプリケーション (1 1 1) を実行する。セキュア実行部 (1 2 0) は、セキュアOSを実行しながら、学習モデルに基づく演算である対象演算を実行する。復号処理部 (1 2 2) は、アプリケーションから対象演算の実行が要求された場合に、学習モデルの暗号化されたモデル情報からモデル情報を復号する。演算実行部 (1 2 1) は、モデル情報を使って対象演算を実行する。改ざん検知部 (1 2 3) は、暗号化されたモデル情報が改ざん情報であるか判定する。出力制御部 (1 2 4) は、暗号化されたモデル情報が改ざん情報でないと判定された場合に、対象演算の実行結果をアプリケーションに渡す。

明 細 書

発明の名称：

情報処理装置、情報処理方法および情報処理プログラム

技術分野

[0001] 本開示は、学習モデルのモデル情報のための秘匿と改ざん検知に関するものである。

背景技術

[0002] 機械学習によって得られる学習モデルについて、アーキテクチャとパラメータといったモデル情報を秘匿することが望まれている。また、モデル情報が改ざんされた場合に改ざんを検知することが望まれている。

[0003] 非特許文献1は、学習モデル情報の秘匿を実現するための方法を開示している。この方法では、実行環境が仮想的に分離され、セキュアな実行環境（セキュアOS、トラステッドOS）でのみ、暗号化された学習モデル情報が復号され、機械学習の演算が実行される。

先行技術文献

非特許文献

[0004] 非特許文献1：Peter M. VanNostrandら著、「Confidential Deep Learning: Executing Proprietary Model on Untrusted Devices」arXiv、2019年8月28日発行

発明の概要

発明が解決しようとする課題

[0005] 非特許文献1の方法では、学習モデル情報の秘匿のみが実現され、改ざん検知が実現されない。

[0006] 本開示は、学習モデルのモデル情報の秘匿と学習モデル情報の改ざん検知とを実現することを目的とする。

課題を解決するための手段

[0007] 本開示の情報処理装置は、

リッチオペレーティングシステムを実行しながらアプリケーションを実行するリッチ実行部と、

セキュアオペレーティングシステムを実行しながら、学習モデルに基づく演算である対象演算を実行するセキュア実行部と、
を備える。

前記セキュア実行部は、

前記アプリケーションから前記対象演算の実行が要求された場合に、前記学習モデルの暗号化されたモデル情報からモデル情報を復号する復号処理部と、

前記モデル情報を使って前記対象演算を実行する演算実行部と、

前記暗号化されたモデル情報が改ざん情報であるか判定する改ざん検知部と、

前記暗号化されたモデル情報が前記改ざん情報でないと判定された場合に、前記対象演算の実行結果を前記アプリケーションに渡す出力制御部と、
を備える。

発明の効果

[0008] 本開示によれば、学習モデルのモデル情報の秘匿と学習モデル情報の改ざん検知とを実現することができる。

図面の簡単な説明

[0009] [図1]実施の形態1における情報処理装置100の構成図。

[図2]実施の形態1における情報処理方法のフローチャート。

[図3]実施の形態1における情報処理方法のフローチャート。

[図4]実施の形態2における情報処理装置100の構成図。

[図5]実施の形態2における情報処理方法のフローチャート。

[図6]実施の形態2における情報処理方法のフローチャート。

[図7]実施の形態2におけるGCMの適用例を示す図。

[図8]実施の形態2における3層ニューラルネットワークに対するGCMの適用例を示す図。

[図9]実施の形態における情報処理装置100のハードウェア構成図。

発明を実施するための形態

[0010] 実施の形態および図面において、同じ要素または対応する要素には同じ符号を付している。説明した要素と同じ符号が付された要素の説明は適宜に省略または簡略化する。図中の矢印はデータの流れ又は処理の流れを主に示している。

[0011] 実施の形態1.

情報処理装置100について、図1から図3に基づいて説明する。

[0012] ***構成の説明***

図1に基づいて、情報処理装置100の構成を説明する。

情報処理装置100は、プロセッサ101とメモリ管理ユニット102とリッチメモリ103とストレージ104とセキュアメモリ105といったハードウェアを備えるコンピュータである。これらのハードウェアは、信号線を介して互いに接続されている。

[0013] プロセッサ101は、演算処理を行うICであり、他のハードウェアを制御する。例えば、プロセッサ101は、CPU、DSPまたはGPUである。

ICは、Integrated Circuitの略称である。

CPUは、Central Processing Unitの略称である。

DSPは、Digital Signal Processorの略称である。

GPUは、Graphics Processing Unitの略称である。

[0014] メモリ管理ユニット102は、リッチメモリ103へのアクセスを制御する。

[0015] リッチメモリ１０３は、汎用のメモリであり、主記憶装置またはメインメモリともいう。

メモリは、揮発性または不揮発性の記憶装置である。例えば、メモリはRAMである。メモリに記憶されたデータは必要に応じてストレージ１０４に保存される。

RAMは、Random Access Memoryの略称である。

[0016] ストレージ１０４は、不揮発性の記憶装置であり、補助記憶装置ともいう。

例えば、ストレージ１０４は、ROM、HDDまたはフラッシュメモリである。ストレージ１０４に記憶されたデータは必要に応じてリッチメモリ１０３にロードされる。

ROMは、Read Only Memoryの略称である。

HDDは、Hard Disk Driveの略称である。

例えば、ストレージ１０４には、学習モデル情報１９１および参照用認証タグ１９２などが記憶される。

[0017] セキュアメモリ１０５は、セキュリティの機能を有してセキュリティが強化されたメモリである。

例えば、セキュアメモリ１０５には、暗号鍵１９３などが記憶される。

[0018] 情報処理装置１００は、リッチ実行部１１０とセキュア実行部１２０とモニタ部１３０といった要素を備える。これらの要素はソフトウェアで実現される。

[0019] ストレージ１０４には、リッチ実行部１１０とセキュア実行部１２０とモニタ部１３０としてコンピュータを機能させるための情報処理プログラムが記憶されている。情報処理プログラムは、リッチメモリ１０３にロードされて、プロセッサ１０１によって実行される。

ストレージ１０４には、さらに、OSが記憶されている。OSの少なくとも一部は、リッチメモリ１０３にロードされて、プロセッサ１０１によって実行される。

プロセッサ１０１は、ＯＳを実行しながら、情報処理プログラムを実行する。

ＯＳは、Operating Systemの略称である。

[0020] 情報処理装置１００は、プロセッサ１０１を代替する複数のプロセッサを備えてもよい。

[0021] 情報処理プログラムは、光ディスクまたはフラッシュメモリ等の不揮発性の記録媒体にコンピュータ読み取り可能に記録（格納）することができる。

[0022] リッチ実行部１１０とセキュア実行部１２０とモニタ部１３０について説明する。

リッチ実行部１１０は、リッチＯＳを実行しながらアプリケーション１１１を実行する。

リッチＯＳは、汎用のＯＳである。

アプリケーション１１１は、アプリケーションプログラムである。

[0023] セキュア実行部１２０は、セキュアＯＳを実行しながら、対象演算を実行する。

セキュアＯＳは、セキュリティの機能を有してセキュリティが強化されたＯＳである。

対象演算は、学習モデルに基づく演算である。具体的には、対象演算は、機械学習によって学習モデルを更新する演算、又は、学習モデルを利用して入力データを解析する演算である。機械学習の具体例はディープラーニングである。但し、機械学習は、ディープラーニングに限定されず、回帰法、決定木学習、ベイズ方またはクラスタリングなどであってもよい。

[0024] セキュア実行部１２０は、演算実行部１２１と復号処理部１２２と改ざん検知部１２３と出力制御部１２４といった要素を備える。これら要素については後述する。

[0025] モニタ部１３０は、仮想モニタとして機能し、切替部１３１を備える。

切替部１３１は、情報処理装置１００の実行環境をリッチＯＳによる環境またはセキュアＯＳによる環境に切り替える。

但し、実行環境の切り替えは他の方法によって実現されてもよい。例えば、情報処理装置 100 がモニタ部 130 を備えず、メモリ管理ユニット 102 が切替部 131 を備えてもよい。

[0026] セキュア実行部 120 の要素について説明する。

復号処理部 122 は、アプリケーション 111 から対象演算の実行が要求された場合に、学習モデル情報 191 からモデル情報を復号する。学習モデル情報 191 は、暗号化モデル情報を含む。暗号化モデル情報は、学習モデルの暗号化されたモデル情報である。

演算実行部 121 は、モデル情報を使って対象演算を実行する。

改ざん検知部 123 は、暗号化モデル情報が改ざん情報であるか判定する。改ざん情報は改ざんされた情報を意味する。

出力制御部 124 は、暗号化モデル情報が改ざん情報でないと判定された場合に、対象演算の実行結果をアプリケーション 111 に渡す。

[0027] 演算実行部 121 は、暗号化モデル情報が改ざん情報でないと判定された場合に、モデル情報を使って対象演算を実行する。

[0028] 出力制御部 124 は、暗号化モデル情報が改ざん情報であると判定された場合に、改ざん通知をアプリケーション 111 に渡す。改ざん通知は、暗号化モデル情報の改ざんを知らせるための通知であり、改ざん検知アラートともいう。

[0029] ***動作の説明***

情報処理装置 100 の動作の手順は情報処理方法に相当する。また、情報処理装置 100 の動作の手順は情報処理プログラムによる処理の手順に相当する。

[0030] 図 2 および図 3 に基づいて、情報処理方法を説明する。

ステップ S101 において、リッチ実行部 110 は、アプリケーション 111 の実行を開始する。

そして、リッチ実行部 110 は、アプリケーション 111 からの演算要求を切替部 131 に送る。演算要求は、対象演算の実行の要求である。

切替部 131 は、演算要求を演算実行部 121 に送る。

[0031] ステップ S111 において、演算実行部 121 は、メモリ管理ユニット 102 によって、学習モデル情報 191 を取得する。メモリ管理ユニット 102 は、学習モデル情報 191 をストレージ 104 からリッチメモリ 103 の中の演算実行部 121 用のメモリ領域に読み込む。

学習モデル情報 191 は、参照情報と暗号化モデル情報とを含む。

暗号化モデル情報は、学習モデルの暗号化されたモデル情報である。

モデル情報は、モデル構成情報とパラメータとを含む。モデル構成情報は学習モデルの構成を示す。

参照情報は、暗号化されたモデル構成情報と暗号化パラメータとのそれぞれの参照先を示す。参照先は、学習モデル情報 191 の中の位置を指す。

[0032] ステップ S112 において、演算実行部 121 は、学習モデル情報 191 の中の参照情報を参照し、暗号化されたモデル構成情報の参照先を復号処理部 122 に通知する。

復号処理部 122 は、メモリ管理ユニット 102 によって、通知された参照先の暗号化されたモデル構成情報を取得する。メモリ管理ユニット 102 は、暗号化されたモデル構成情報をストレージ 104 からリッチメモリ 103 の中の復号処理部 122 用のメモリ領域に読み込む。

そして、復号処理部 122 は、セキュアメモリ 105 に記憶されている暗号鍵 193 を使って、暗号化されたモデル構成情報からモデル構成情報を復号する。

[0033] ステップ S113 において、演算実行部 121 は、復号処理部 122 からモデル構成情報を取得し、モデル構成情報を対象演算に設定する。

[0034] ステップ S114 において、演算実行部 121 は、学習モデル情報 191 の中の参照情報を参照し、暗号化されたパラメータの参照先を復号処理部 122 に通知する。

復号処理部 122 は、メモリ管理ユニット 102 によって、通知された参照先の暗号化されたパラメータを取得する。メモリ管理ユニット 102 は、

暗号化されたパラメータをストレージ１０４からリッチメモリ１０３の中の復号処理部１２２用のメモリ領域に読み込む。

そして、復号処理部１２２は、セキュアメモリ１０５に記憶されている暗号鍵１９３を使って、暗号化されたパラメータからパラメータを復号する。

[0035] ステップＳ１１５において、演算実行部１２１は、復号処理部１２２からパラメータを取得し、パラメータを対象演算に設定する。

[0036] ステップＳ１２１において、演算実行部１２１は、学習モデル情報１９１を改ざん検知部１２３に渡す。

改ざん検知部１２３は、学習モデル情報１９１の認証タグを生成する。認証タグは、特定の計算によって算出される。認証タグはメッセージ認証コードともいう。

生成される認証タグを「検証用認証タグ」と称する。

[0037] ステップＳ１２２において、改ざん検知部１２３は、メモリ管理ユニット１０２によって、参照用認証タグ１９２を取得する。メモリ管理ユニット１０２は、参照用認証タグ１９２をストレージ１０４からリッチメモリ１０３の中の改ざん検知部１２３用のメモリ領域に読み込む。

参照用認証タグ１９２は、当初の学習モデル情報１９１の認証タグ、つまり、改ざんされていない学習モデル情報１９１の認証タグである。

[0038] ステップＳ１２３において、改ざん検知部１２３は、検証用認証タグを参照用認証タグ１９２と照合する。

そして、改ざん検知部１２３は、照合結果に基づいて、暗号化されたモデル情報が改ざん情報であるか判定する。

検証用認証タグが参照用認証タグ１９２と一致する場合、改ざん検知部１２３は、暗号化されたモデル情報が改ざん情報でないと判定する。そして、処理はステップＳ１３１に進む。

検証用認証タグが参照用認証タグ１９２と一致しない場合、改ざん検知部１２３は、暗号化されたモデル情報が改ざん情報であると判定する。そして、処理はステップＳ１４１に進む。

[0039] ステップS 1 3 1において、改ざん検知部 1 2 3は、暗号化されたモデル情報が改ざん情報でないことを示す判定結果を出力制御部 1 2 4に通知する。

出力制御部 1 2 4は、実行指示を演算実行部 1 2 1に通知する。

演算実行部 1 2 1は、メモリ管理ユニット 1 0 2によって、アプリケーション 1 1 1から入力データを取得する。メモリ管理ユニット 1 0 2は、入力データをリッチメモリ 1 0 3の中のアプリケーション 1 1 1用のメモリ領域からリッチメモリ 1 0 3の中の演算実行部 1 2 1用のメモリ領域に読み込む。

[0040] ステップS 1 3 2において、演算実行部 1 2 1は、入力データに対して対象演算を実行する。

[0041] ステップS 1 3 3において、演算実行部 1 2 1は、演算結果を出力制御部 1 2 4に渡す。演算結果は、対象演算の実行結果、すなわち、対象演算を実行して得られた結果である。

出力制御部 1 2 4は、メモリ管理ユニット 1 0 2によって、演算結果をアプリケーション 1 1 1に渡す。メモリ管理ユニット 1 0 2は、演算結果をリッチメモリ 1 0 3の中のアプリケーション 1 1 1用のメモリ領域に書き込む。

ステップS 1 3 3の後、処理はステップS 1 5 1に進む。

[0042] ステップS 1 4 1において、改ざん検知部 1 2 3は、暗号化されたモデル情報が改ざん情報であることを示す判定結果を出力制御部 1 2 4に通知する。

出力制御部 1 2 4は、メモリ管理ユニット 1 0 2によって、改ざん通知をアプリケーション 1 1 1に渡す。メモリ管理ユニット 1 0 2は、改ざん通知をリッチメモリ 1 0 3の中のアプリケーション 1 1 1用のメモリ領域に書き込む。

ステップS 1 4 1の後、処理はステップS 1 5 1に進む。

[0043] ステップS 1 5 1において、出力制御部 1 2 4は、完了応答を切替部 1 3

1 に送る。完了応答は演算要求に対する応答である。

切替部 131 は、完了応答をリッチ実行部 110 に送る。

リッチ実行部 110 は、アプリケーション 111 により、演算結果が得られた場合の処理または改ざん通知が得られた場合の処理を実行する。

[0044] ***実施の形態 1 の効果***

セキュアな実行環境での機械学習演算において、情報処理装置 100 は、暗号化されたモデル構成情報と暗号化されたパラメータとを復号する。また、情報処理装置 100 は、改ざん検知を実施する。そして、情報処理装置 100 は、改ざんの有無に基づいて、演算を実行して結果を出力するか、または、改ざん通知を出力するかを選択する。

これにより、セキュアな実行環境において、モデル情報の秘匿とモデル情報の改ざん検知とを両立することができる。

[0045] 実施の形態 2.

メモリリソースおよび計算時間のオーバーヘッドを抑える形態について、主に実施の形態 1 と異なる点を図 4 から図 8 に基づいて説明する。

[0046] ***構成の説明***

図 4 に基づいて、情報処理装置 100 の構成を説明する。

情報処理装置 100 は、さらに、ストレージ参照部 125 を備える。ストレージ参照部 125 は、ソフトウェアによって実現される。

情報処理プログラムは、さらに、ストレージ参照部 125 としてコンピュータを機能させる。

[0047] セキュア実行部 120 の要素について説明する。

改ざん検知部 123 は、暗号化されたモデル情報が改ざん情報であるか判定するための計算処理を対象演算と並列に又は対象演算と並行に実行する。

[0048] 改ざん検知部 123 は、上記計算処理を以下のように実行する。

対象演算は、実行順序が決められた複数の演算関数に対する演算を含む。

暗号化されたモデル情報は、複数の演算関数のそれぞれの暗号化されたパラメータを含む。

復号処理部 1 2 2 は、各演算関数の暗号化されたパラメータから各演算関数のパラメータを復号する。

演算実行部 1 2 1 は、実行順序にしたがって、各演算関数のパラメータを使って各演算関数を演算する。

改ざん検知部 1 2 3 は、実行順序にしたがって、各演算関数の演算と並列に又は各演算関数の演算と並行に、各演算関数の暗号化されたパラメータを使って上記計算処理を実行する。

[0049] 具体的には、改ざん検知部 1 2 3 は、上記計算処理を以下のように実行する。

暗号化されたモデル情報は、暗号化されたモデル構成情報を含む。

まず、改ざん検知部 1 2 3 は、暗号化されたモデル構成情報の認証タグを暫定認証タグとして生成する。

次に、改ざん検知部 1 2 3 は、実行順序にしたがって各演算関数の暗号化されたパラメータを使って上記計算処理を実行して暫定認証タグの更新を行う。

次に、改ざん検知部 1 2 3 は、更新後の暫定認証タグを用いて検証用認証タグを生成する。

そして、改ざん検知部 1 2 3 は、検証用認証タグを参照用認証タグ 1 9 2 と照合して、暗号化されたモデル情報が改ざん情報であるか判定する。

[0050] 出力制御部 1 2 4 は、暗号化されたモデル情報が改ざん情報であると判定された場合に、改ざん通知をアプリケーション 1 1 1 に渡す。

[0051] ***動作の説明***

図 5 および図 6 に基づいて、情報処理方法を説明する。

ステップ S 2 0 1 において、リッチ実行部 1 1 0 は、アプリケーション 1 1 1 からの演算要求を切替部 1 3 1 に送る。

ステップ S 2 0 1 は、実施の形態 1 におけるステップ S 1 0 1 と同じである。

[0052] ステップ S 2 1 1 において、演算実行部 1 2 1 は、メモリ管理ユニット 1

02によって、学習モデル情報191を取得する。

ステップS211は、実施の形態1におけるステップS111と同じである。

[0053] ステップS212において、演算実行部121は、学習モデル情報191をストレージ参照部125に渡す。

ストレージ参照部125は、暗号化されたモデル構成情報を学習モデル情報191から取得し、暗号化されたモデル構成情報を復号処理部122に渡す。

復号処理部122は、セキュアメモリ105に記憶されている暗号鍵193を使って、暗号化されたモデル構成情報からモデル構成情報を復号する。

[0054] ステップS213において、演算実行部121は、復号処理部122からモデル構成情報を取得し、モデル構成情報を対象演算に設定する。

ステップS213は、実施の形態1におけるステップS113と同じである。

[0055] ステップS221において、ストレージ参照部125は、暗号化されたモデル構成情報を改ざん検知部123に渡す。

改ざん検知部123は、暗号化されたモデル構成情報の認証タグを生成する。

生成される認証タグを「暫定認証タグ」と称する。

[0056] ステップS222において、演算実行部121は、メモリ管理ユニット102により、アプリケーション111から入力データを取得する。メモリ管理ユニット102は、入力データをリッチメモリ103の中のアプリケーション111用のメモリ領域からリッチメモリ103の中の演算実行部121用のメモリ領域に読み込む。

[0057] 対象演算は、実行順序が決められた複数の演算関数に対する演算を含む。

ステップS231からステップS234は、複数の演算関数の実行順序にしたがって各演算関数に対して実行される。つまり、ステップS231からステップS234は、演算関数の数だけ繰り返される。

対象となる演算関数を「対象関数」と称する。

[0058] ステップS 2 3 1において、ストレージ参照部 1 2 5は、対象関数の暗号化されたパラメータを学習モデル情報 1 9 1から取得し、対象関数の暗号化されたパラメータを復号処理部 1 2 2に渡す。

復号処理部 1 2 2は、セキュアメモリ 1 0 5に記憶されている暗号鍵 1 9 3を使って、対象関数の暗号化されたパラメータからパラメータを復号する。

[0059] ステップS 2 3 2において、演算実行部 1 2 1は、復号処理部 1 2 2からパラメータを取得し、パラメータを対象関数に設定する。

[0060] ステップS 2 3 3において、ストレージ参照部 1 2 5は、対象関数の暗号化されたパラメータを改ざん検知部 1 2 3に渡す。

改ざん検知部 1 2 3は、対象関数の暗号化されたパラメータを使って計算処理を実行して暫定認証タグを更新する。

[0061] ステップS 2 3 4において、演算実行部 1 2 1は、対象関数を演算する。

例えば、対象関数が1番目の演算関数である場合、演算実行部 1 2 1は、入力データに対して対象関数を実行する。また、対象関数が2番目以降の演算関数である場合、演算実行部 1 2 1は、前回の対象関数の演算結果に対して対象関数を実行する。

[0062] ステップS 2 3 5において、演算実行部 1 2 1は、全ての演算関数の演算が終了したか判定する。

全ての演算関数の演算が終了していない場合、処理はステップS 2 3 1に進む。

全ての演算関数の演算が終了した場合、処理はステップS 2 4 1に進む。

[0063] ステップS 2 4 1において、改ざん検知部 1 2 3は、暫定認証タグを用いて検証用認証タグを生成する。

[0064] ステップS 2 4 2において、ストレージ参照部 1 2 5は、メモリ管理ユニット 1 0 2によって、参照用認証タグ 1 9 2を取得する。メモリ管理ユニット 1 0 2は、参照用認証タグ 1 9 2をストレージ 1 0 4からリッチメモリ 1

03の中のストレージ参照部125用のメモリ領域に読み込む。

ストレージ参照部125は、参照用認証タグ192を改ざん検知部123に渡す。

改ざん検知部123は、参照用認証タグ192をストレージ参照部125から取得する。

[0065] ステップS243において、改ざん検知部123は、検証用認証タグを参照用認証タグ192と照合する。

そして、改ざん検知部123は、照合結果に基づいて、暗号化されたモデル情報が改ざん情報であるか判定する。

検証用認証タグが参照用認証タグ192と一致する場合、改ざん検知部123は、暗号化されたモデル情報が改ざん情報でないと判定する。そして、処理はステップS251に進む。

検証用認証タグが参照用認証タグ192と一致しない場合、改ざん検知部123は、暗号化されたモデル情報が改ざん情報であると判定する。そして、処理はステップS261に進む。

[0066] ステップS251において、演算実行部121は、演算結果を出力制御部124に渡す。

ステップS251は、実施の形態1におけるステップS133と同じである。但し、出力制御部124には、最後の演算関数の実行結果が渡される。

[0067] ステップS261において、出力制御部124は、メモリ管理ユニット102によって、改ざん通知をアプリケーション111に渡す。

ステップS261は、実施の形態1におけるステップS141と同じである。

[0068] ステップS271において、出力制御部124は、完了応答を切替部131に送る。

ステップS271は、実施の形態1におけるステップS151と同じである。

[0069] ***実施の形態2の補足***

モデル情報の暗号化とモデル情報の改ざん検知のためのアルゴリズムとして、G a l o i s / C o u n t e r M o d e (G C M) による認証付き暗号を利用することができる。

図7に基づいて、G C Mによる認証付き暗号処理を実施の形態2に適用した例を説明する。

機械学習アルゴリズムは、機械学習済みの複数の演算関数 (M L) で構成され、入力データ (I n p u t D a t a) に対する演算結果 (O u t p u t D a t a) を返す。

機械学習演算に用いるパラメータ (P l a i n t e x t) は暗号化されている。

演算関数の実行時には、暗号化されたパラメータ (C i p h e r t e x t) が復号され、パラメータが演算関数に用いられる。また、暫定認証タグ (A u t h D a t a) も逐次的に更新される。

なお、逐次的に展開されるパラメータおよび演算関数によって得られる中間結果は、演算関数の実行ごとに削除または上書きされる。これにより、リッチメモリ 103 に展開されるデータ量が削減される。

全ての演算関数が実行されると、検証用認証タグ (T a g) が生成され、ストレージ 104 に格納された参照用認証タグ 192 (A u t h T a g) と検証用タグが比較され、改ざんの有無が確認される。

改ざんが無い場合には、演算結果 (O u t p u t D a t a) が出力される。

改ざんが有る場合には、演算結果の代わりに改ざん通知 (O u t p u t D a t a) が出力される。

認証付き暗号処理と機械学習演算処理は、並列で実行されてもよいし、並行で処理されてもよい。

[0070] 図8に基づいて、学習モデルが3層のニューラルネットワークである場合について、G C Mによる認証付き暗号処理を実施の形態2に適用した例を説明する。

リッチ実行部 110 により、入力データと学習モデル情報 191 と参照用認証タグ 192 が管理される。

学習モデル情報 191 は、暗号化されたモデル構成情報を含む。セキュア実行部 120 により、暗号化されたモデル構成情報からモデル構成情報が復号される。モデル構成情報には、3 層のニューラルネットワークの複数の演算関数が設定されている。

参照用認証タグ 192 は、予め生成されている。

まず、入力データがリッチ実行部 110 からセキュア実行部 120 に渡される。

次に、第 1 層の演算関数（畳込み演算）が実行される。このとき、第 1 層のパラメータが復号されてセキュア実行部 120 用のメモリ領域に展開される。また、暫定認証タグが更新される。第 1 層の演算関数の終了後、演算に用いられたデータ（パラメータなど）はメモリ領域から削除される。

次に、第 2 層の演算関数（畳込み演算）が実行される。このとき、第 2 層のパラメータが復号されてセキュア実行部 120 用のメモリ領域に展開される。また、暫定認証タグが更新される。第 2 層の演算関数の終了後、演算に用いられたデータ（パラメータなど）はメモリ領域から削除される。

次に、第 3 層の第 1 演算関数（全結合乗算）が実行される。このとき、第 3 層のパラメータ（重み）が復号されてセキュア実行部 120 用のメモリ領域に展開される。また、暫定認証タグが更新される。第 3 層の第 1 演算関数の終了後、演算に用いられたデータ（パラメータなど）はメモリ領域から削除される。

次に、第 3 層の第 2 演算関数（全結合加算）が実行される。このとき、第 3 層のパラメータ（バイアス）が復号されてセキュア実行部 120 用のメモリ領域に展開される。また、暫定認証タグが更新される。第 3 層の第 2 演算関数の終了後、演算に用いられたデータ（パラメータなど）はメモリ領域から削除される。

そして、暫定認証タグを用いて検証用認証タグが生成され、検証用認証タ

グが参照用認証タグ 192 と比較され、改ざんの有無が判定される。

改ざんが無い場合、演算結果（出力データ）がセキュア実行部 120 からリッチ実行部 110 に渡される。

改ざんが有る場合、改ざん通知（出力データ）がセキュア実行部 120 からリッチ実行部 110 に渡される。

1つの層に複数の演算がある場合、第3層のように複数の演算を分離して実行してもよい。また、複数の演算を分離せずにまとめて実行してもよい。

また、2つ以上の層をまとめてもよい。つまり、異なる層の2つ以上の演算をまとめて実行してもよい。さらに、入力データを一定のデータサイズに分割しても各演算を実行してもよい。

[0071] ***実施の形態2の効果***

情報処理装置 100 は、機械学習演算を演算関数単位で並列または並行に処理する。これにより、メモリリソースおよび計算時間のオーバーヘッドを抑えて、セキュアな実行環境においてモデル情報の秘匿とモデル情報の改ざん検知とを両立することができる。

[0072] ステップ S231 からステップ S234 において、全てのパラメータが復号して展開されるのではなく、演算関数ごとにパラメータが復号して展開され演算関数が演算される。これにより、演算時のメモリ使用量を抑えることができる。

[0073] ***実施の形態の補足***

図9に基づいて、情報処理装置 100 のハードウェア構成を説明する。記憶装置はリッチメモリ 103、ストレージ 104 およびセキュアメモリ 105 を表している。

情報処理装置 100 は処理回路 109 を備える。

処理回路 109 は、リッチ実行部 110 とセキュア実行部 120 とモニタ部 130 とを実現するハードウェアである。

処理回路 109 は、専用のハードウェアであってもよいし、記憶装置に格納されるプログラムを実行するプロセッサ 101 であってもよい。

[0074] 処理回路１０９が専用のハードウェアである場合、処理回路１０９は、例えば、単回路、複合回路、プログラム化したプロセッサ、並列プログラム化したプロセッサ、ＡＳＩＣ、ＦＰＧＡまたはこれらの組み合わせである。

ＡＳＩＣは、Ａｐｐｌｉｃａｔｉｏｎ Ｓｐｅｃｉｆｉｃ Ｉｎｔｅｇｒａｔｅｄ Ｃｉｒｃｕｉｔの略称である。

ＦＰＧＡは、Ｆｉｅｌｄ Ｐｒｏｇｒａｍｍａｂｌｅ Ｇａｔｅ Ａｒｒａｙの略称である。

[0075] 情報処理装置１００は、処理回路１０９を代替する複数の処理回路を備えてもよい。

[0076] 処理回路１０９において、一部の機能が専用のハードウェアで実現されて、残りの機能がソフトウェアまたはファームウェアで実現されてもよい。

[0077] このように、情報処理装置１００の機能はハードウェア、ソフトウェア、ファームウェアまたはこれらの組み合わせで実現することができる。

[0078] 各実施の形態は、好ましい形態の例示であり、本開示の技術的範囲を制限することを意図するものではない。各実施の形態は、部分的に実施してもよいし、他の形態と組み合わせて実施してもよい。フローチャート等を用いて説明した手順は、適宜に変更してもよい。

[0079] 情報処理装置１００の要素である「部」は、「処理」、「工程」、「回路」または「サーキットリ」と読み替えてもよい。

符号の説明

[0080] １００ 情報処理装置、１０１ プロセッサ、１０２ メモリ管理ユニット、１０３ リッチメモリ、１０４ ストレージ、１０５ セキュアメモリ、１０９ 処理回路、１１０ リッチ実行部、１１１ アプリケーション、１２０ セキュア実行部、１２１ 演算実行部、１２２ 復号処理部、１２３ 改ざん検知部、１２４ 出力制御部、１２５ ストレージ参照部、１３０ モニタ部、１３１ 切替部、１９１ 学習モデル情報、１９２ 参照用認証タグ、１９３ 暗号鍵。

請求の範囲

- [請求項1] リッチオペレーティングシステムを実行しながらアプリケーションを実行するリッチ実行部と、
- セキュアオペレーティングシステムを実行しながら、学習モデルに基づく演算である対象演算を実行するセキュア実行部と、
- を備え、
- 前記セキュア実行部は、
- 前記アプリケーションから前記対象演算の実行が要求された場合に、前記学習モデルの暗号化されたモデル情報からモデル情報を復号する復号処理部と、
- 前記モデル情報を使って前記対象演算を実行する演算実行部と、
- 前記暗号化されたモデル情報が改ざん情報であるか判定する改ざん検知部と、
- 前記暗号化されたモデル情報が前記改ざん情報でないと判定された場合に、前記対象演算の実行結果を前記アプリケーションに渡す出力制御部と、
- を備える情報処理装置。
- [請求項2] 前記演算実行部は、前記暗号化されたモデル情報が前記改ざん情報でないと判定された場合に、前記モデル情報を使って前記対象演算を実行する
- 請求項1に記載の情報処理装置。
- [請求項3] 前記出力制御部は、前記暗号化されたモデル情報が前記改ざん情報であると判定された場合に、改ざん通知を前記アプリケーションに渡す
- 請求項1または請求項2に記載の情報処理装置。
- [請求項4] 前記改ざん検知部は、前記暗号化されたモデル情報が前記改ざん情報であるか判定するための計算処理を前記対象演算と並列に又は前記対象演算と並行に実行する

請求項 1 に記載の情報処理装置。

[請求項5] 前記対象演算は、実行順序が決められた複数の演算関数に対する演算を含み、

前記暗号化されたモデル情報は、前記複数の演算関数のそれぞれの暗号化されたパラメータを含み、

前記復号処理部は、各演算関数の暗号化されたパラメータから各演算関数のパラメータを復号し、

前記演算実行部は、前記実行順序にしたがって、各演算関数のパラメータを使って各演算関数を演算し、

前記改ざん検知部は、前記実行順序にしたがって、各演算関数の演算と並列に又は各演算関数の演算と並行に、各演算関数の暗号化されたパラメータを使って前記計算処理を実行する

請求項 4 に記載の情報処理装置。

[請求項6] 前記暗号化されたモデル情報は、暗号化されたモデル構成情報を含み、

前記改ざん検知部は、前記暗号化されたモデル構成情報の認証タグを暫定認証タグとして生成し、前記実行順序にしたがって各演算関数の暗号化されたパラメータを使って前記計算処理を実行して前記暫定認証タグの更新を行い、更新後の前記暫定認証タグを用いて検証用認証タグを生成し、前記検証用認証タグを参照用認証タグと照合して前記暗号化されたモデル情報が前記改ざん情報であるか判定する

請求項 5 に記載の情報処理装置。

[請求項7] 前記出力制御部は、前記暗号化されたモデル情報が前記改ざん情報であると判定された場合に、改ざん通知を前記アプリケーションに渡す

請求項 4 から請求項 6 のいずれか 1 項に記載の情報処理装置。

[請求項8] リッチ実行部が、リッチオペレーティングシステムを実行しながらアプリケーションを実行し、

セキュア実行部が、セキュアオペレーティングシステムを実行しながら、学習モデルに基づく演算である対象演算を実行する情報処理方法であって、

前記セキュア実行部において、

復号処理部が、前記アプリケーションから前記対象演算の実行が要求された場合に、前記学習モデルの暗号化されたモデル情報からモデル情報を復号し、

演算実行部が、前記モデル情報を使って前記対象演算を実行し、

改ざん検知部が、前記暗号化されたモデル情報が改ざん情報であるか判定し、

出力制御部が、前記暗号化されたモデル情報が前記改ざん情報でないと判定された場合に、前記対象演算の実行結果を前記アプリケーションに渡す

情報処理方法。

[請求項9]

リッチオペレーティングシステムを実行しながらアプリケーションを実行するリッチ実行処理と、

セキュアオペレーティングシステムを実行しながら、学習モデルに基づく演算である対象演算を実行するセキュア実行処理と、
をコンピュータに実行させるための情報処理プログラムであって、

前記セキュア実行処理は、

前記アプリケーションから前記対象演算の実行が要求された場合に、前記学習モデルの暗号化されたモデル情報からモデル情報を復号する復号処理と、

前記モデル情報を使って前記対象演算を実行する演算実行処理と、

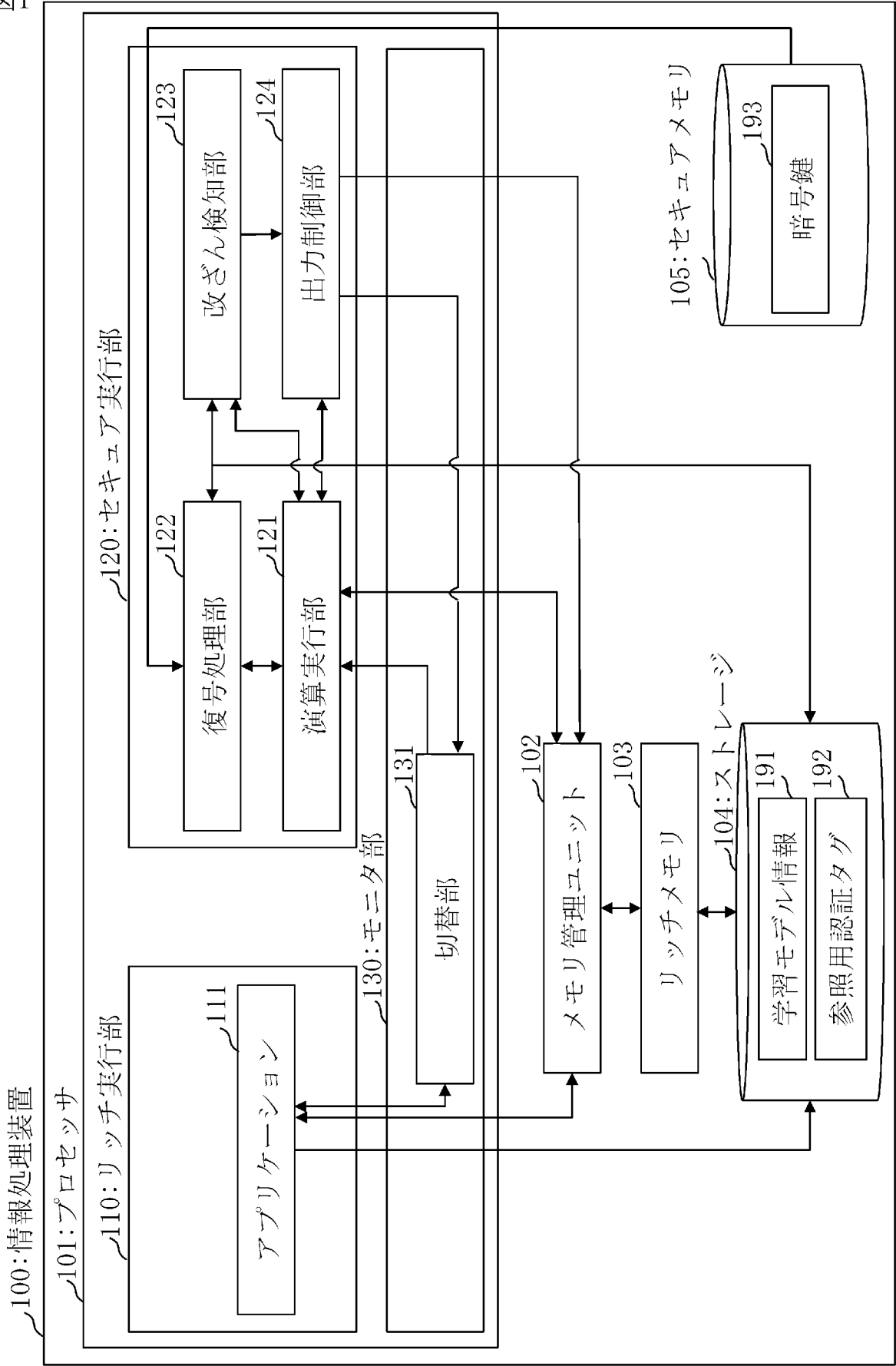
前記暗号化されたモデル情報が改ざん情報であるか判定する改ざん検知処理と、

前記暗号化されたモデル情報が前記改ざん情報でないと判定された場合に、前記対象演算の実行結果を前記アプリケーションに渡す出力

制御処理と、を含む
情報処理プログラム。

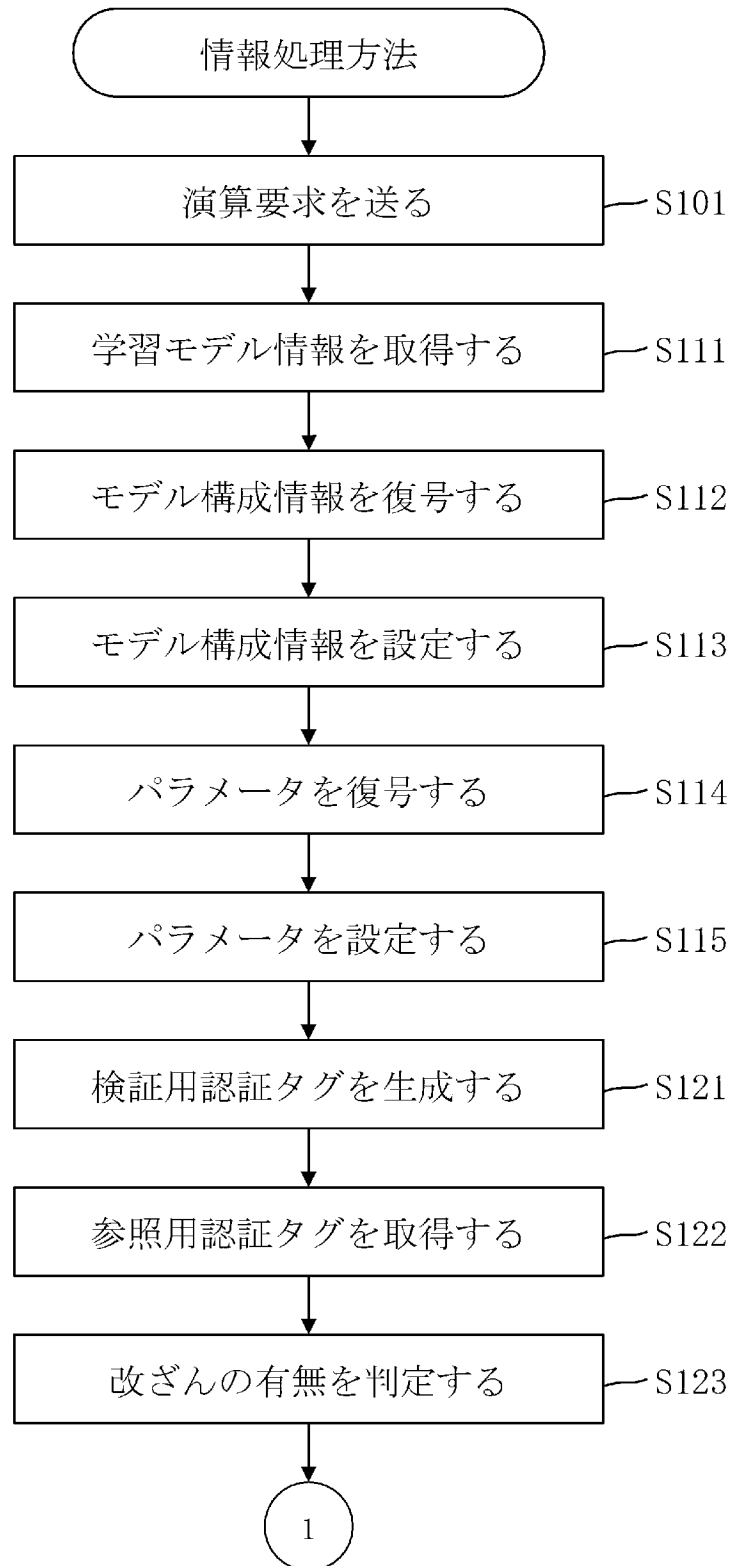
[図1]

図1



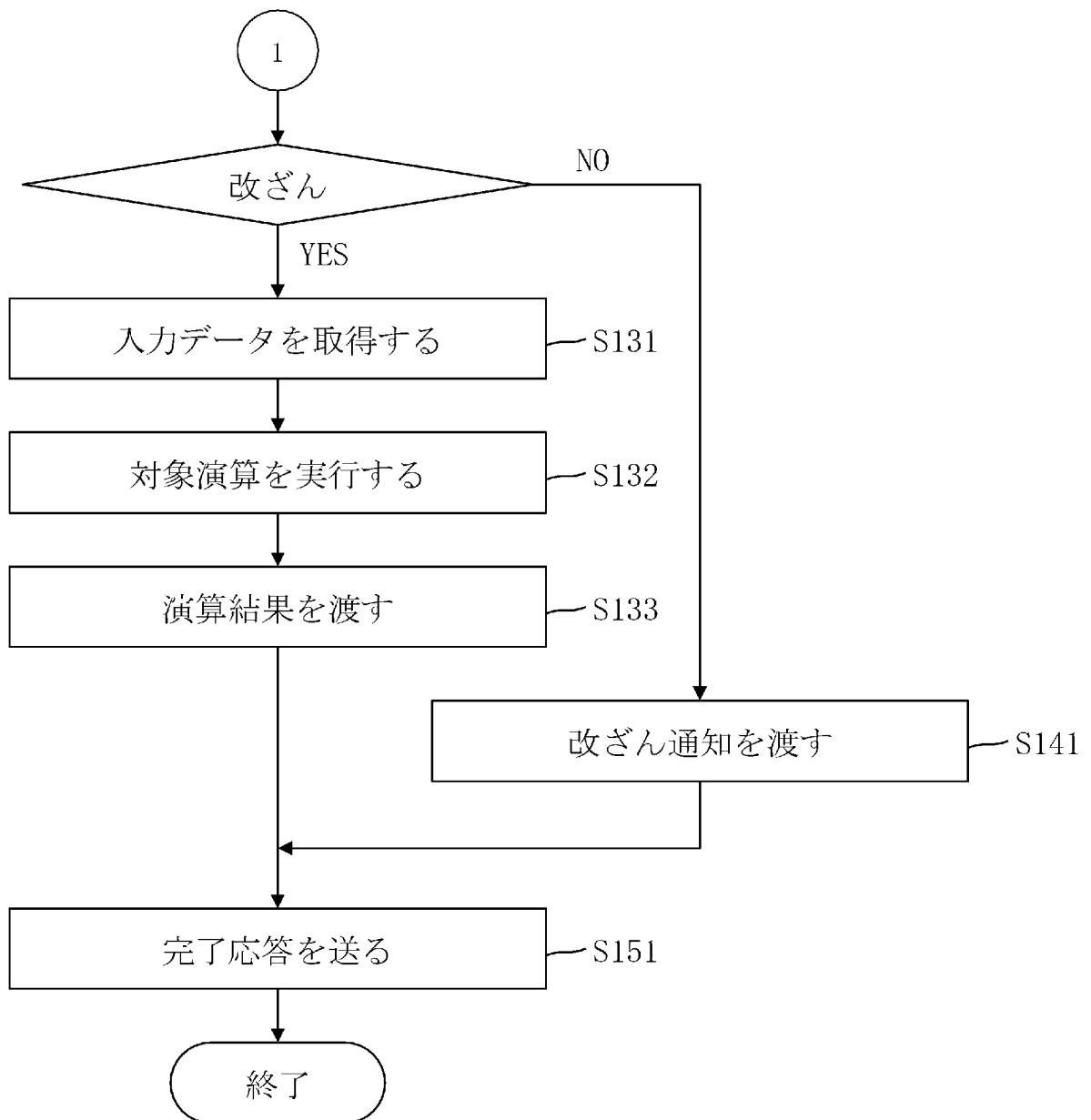
[図2]

図2



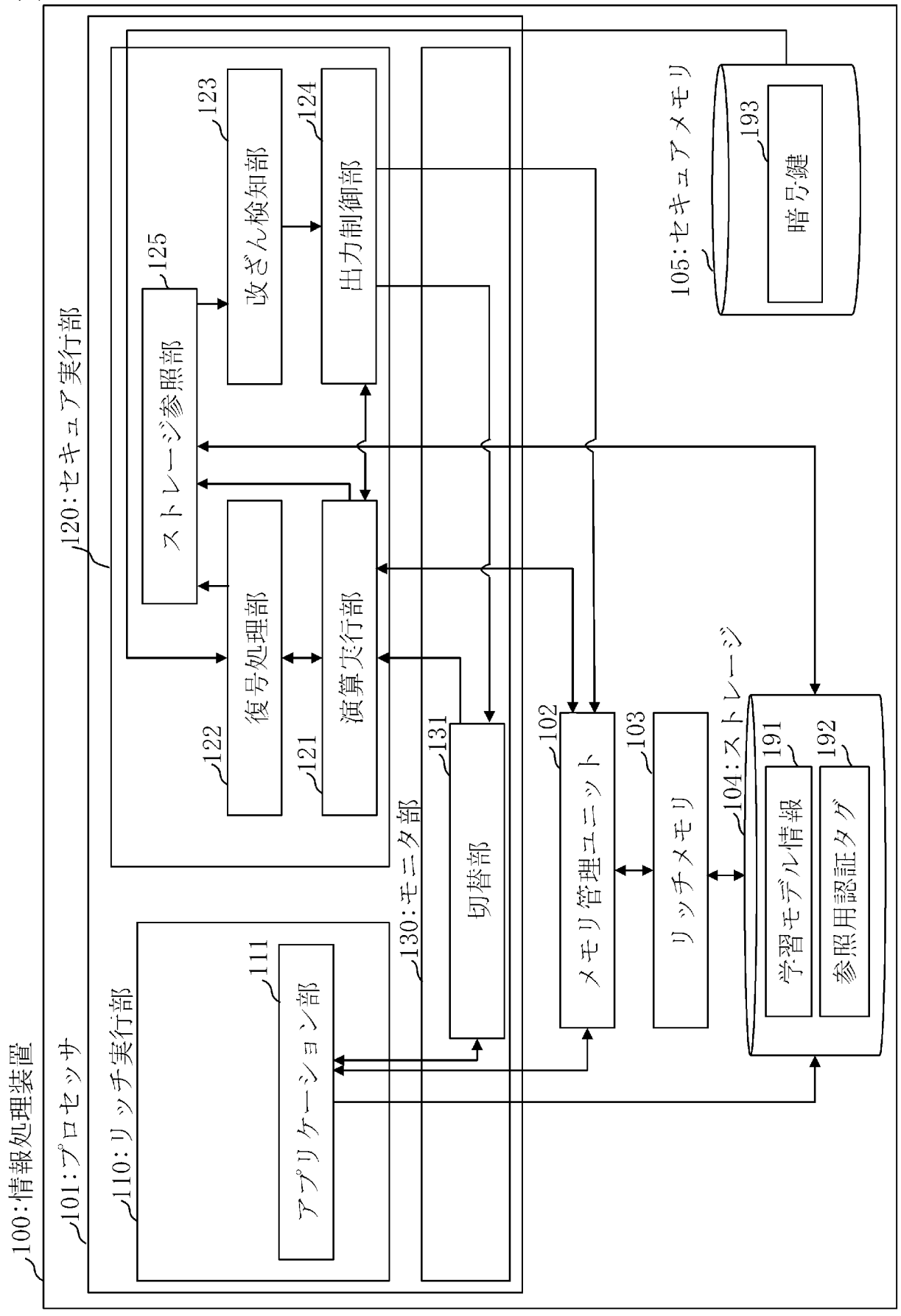
[図3]

図3



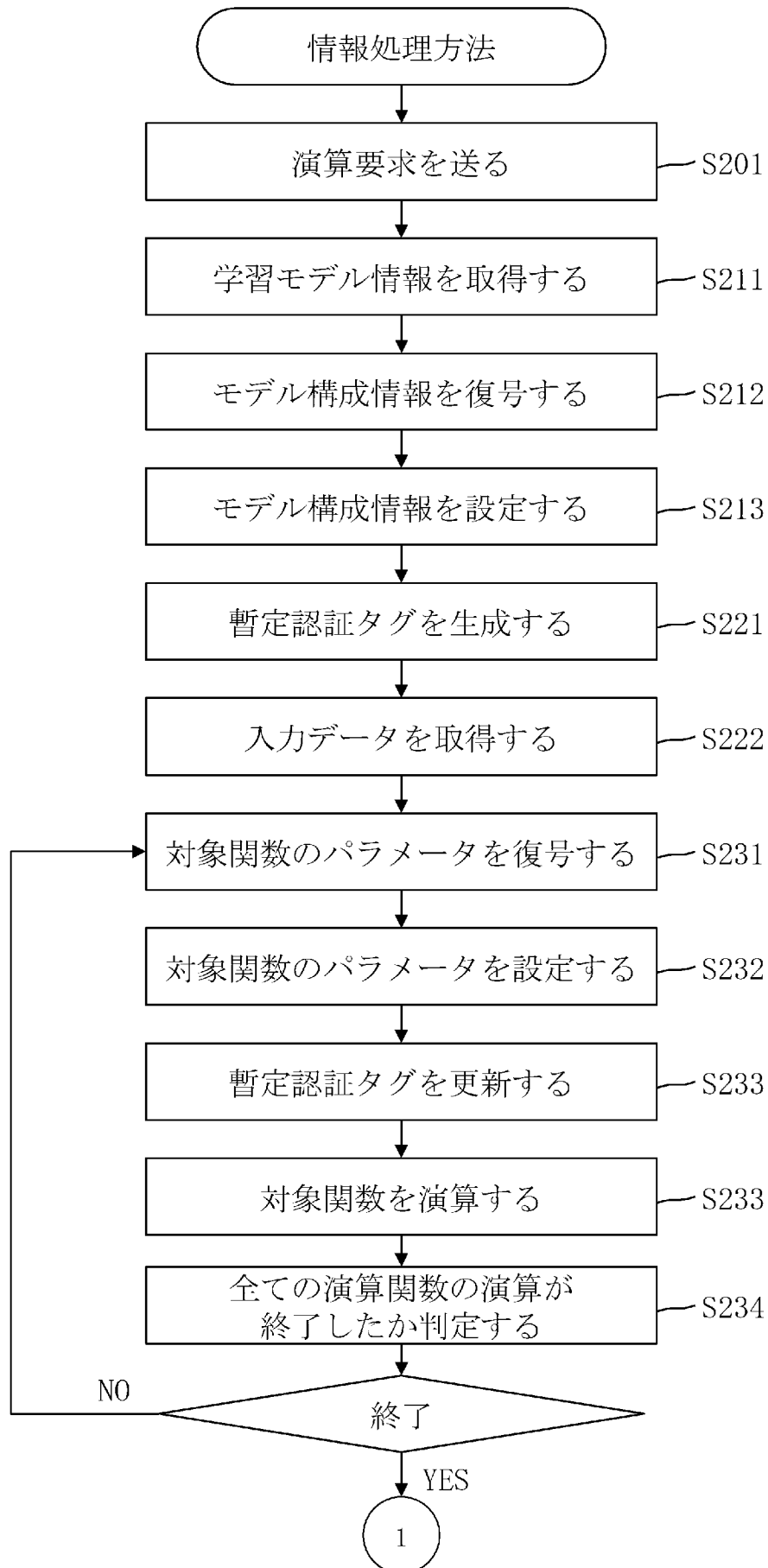
[図4]

図4



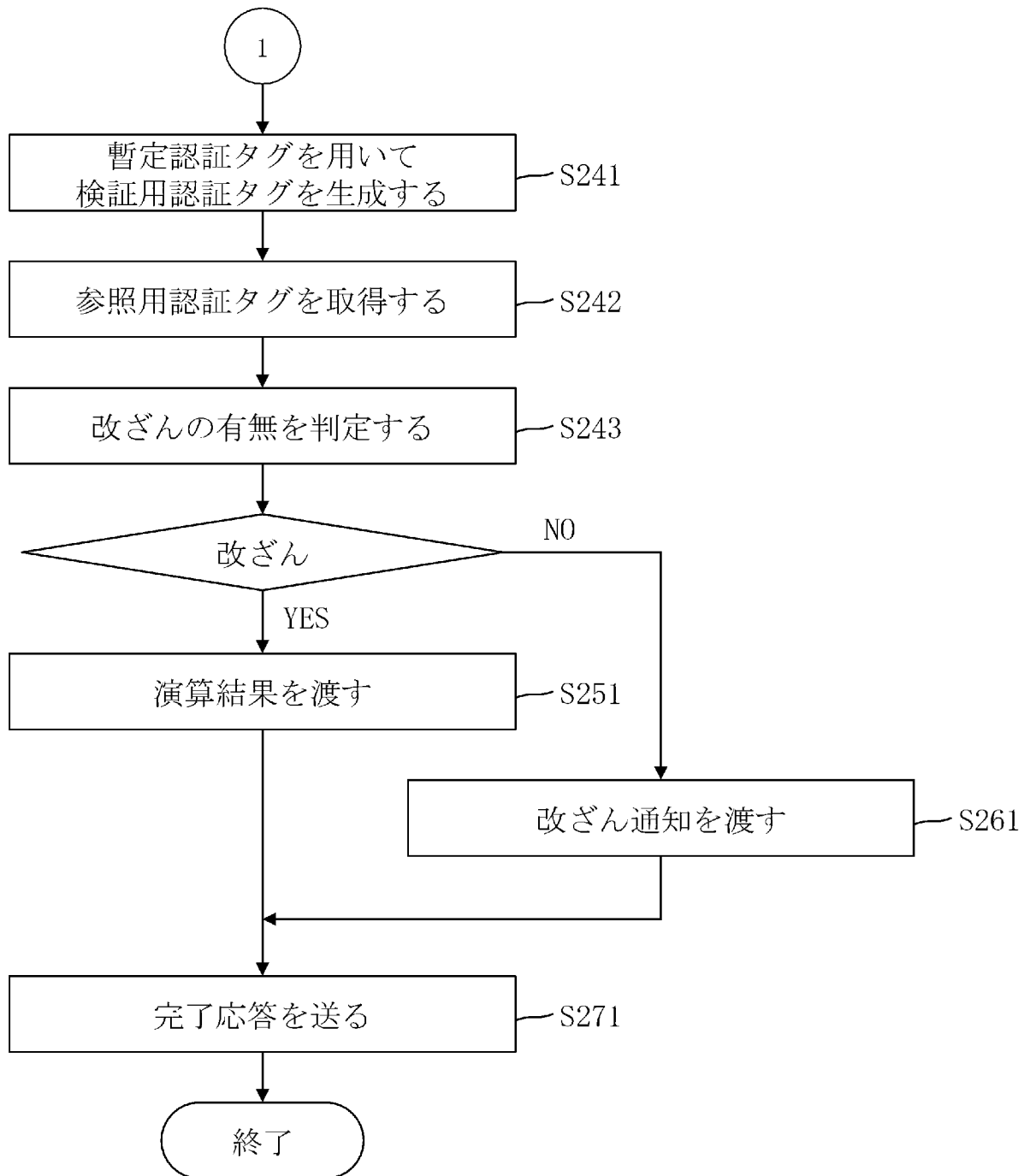
[図5]

図5

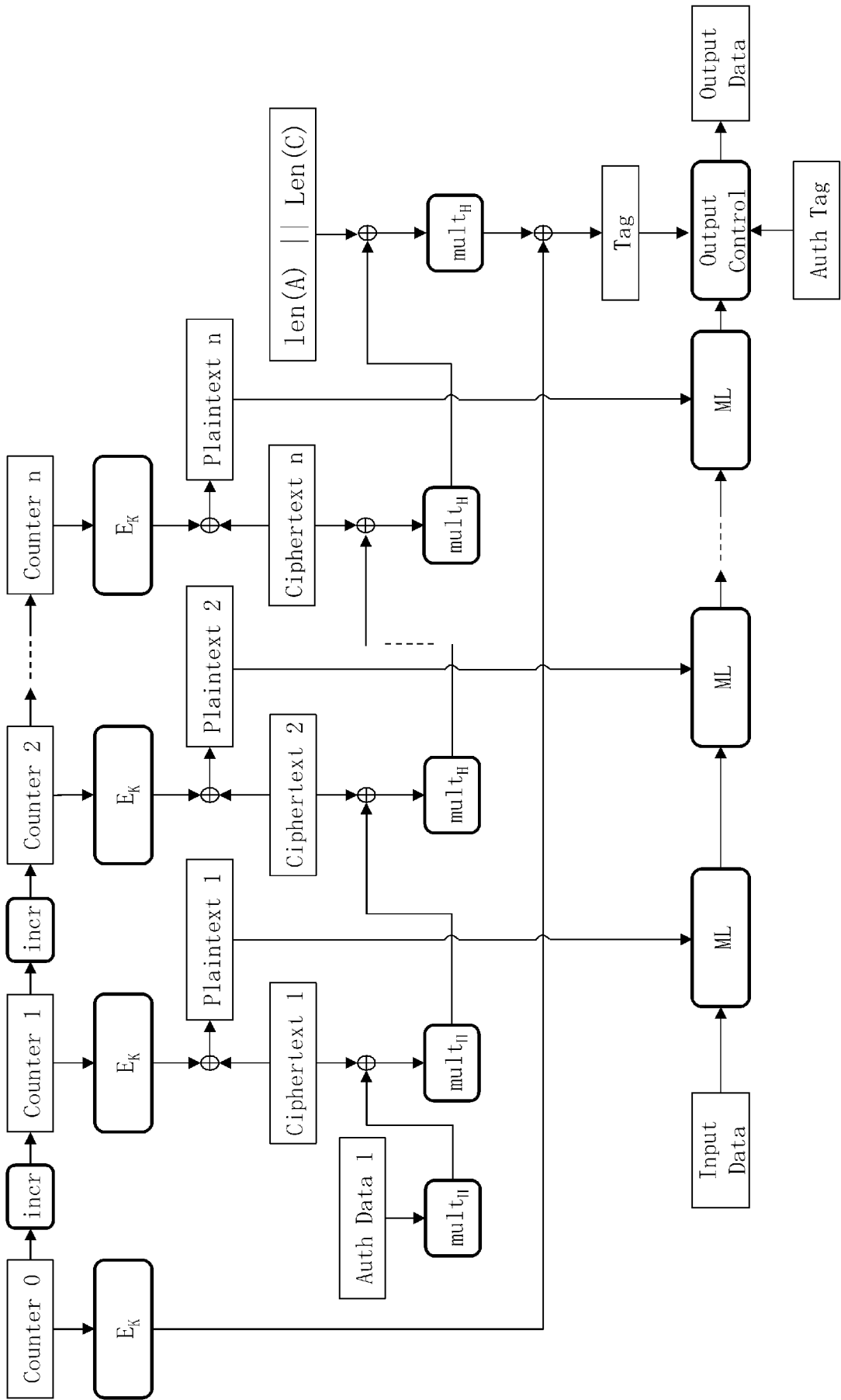


[図6]

図6

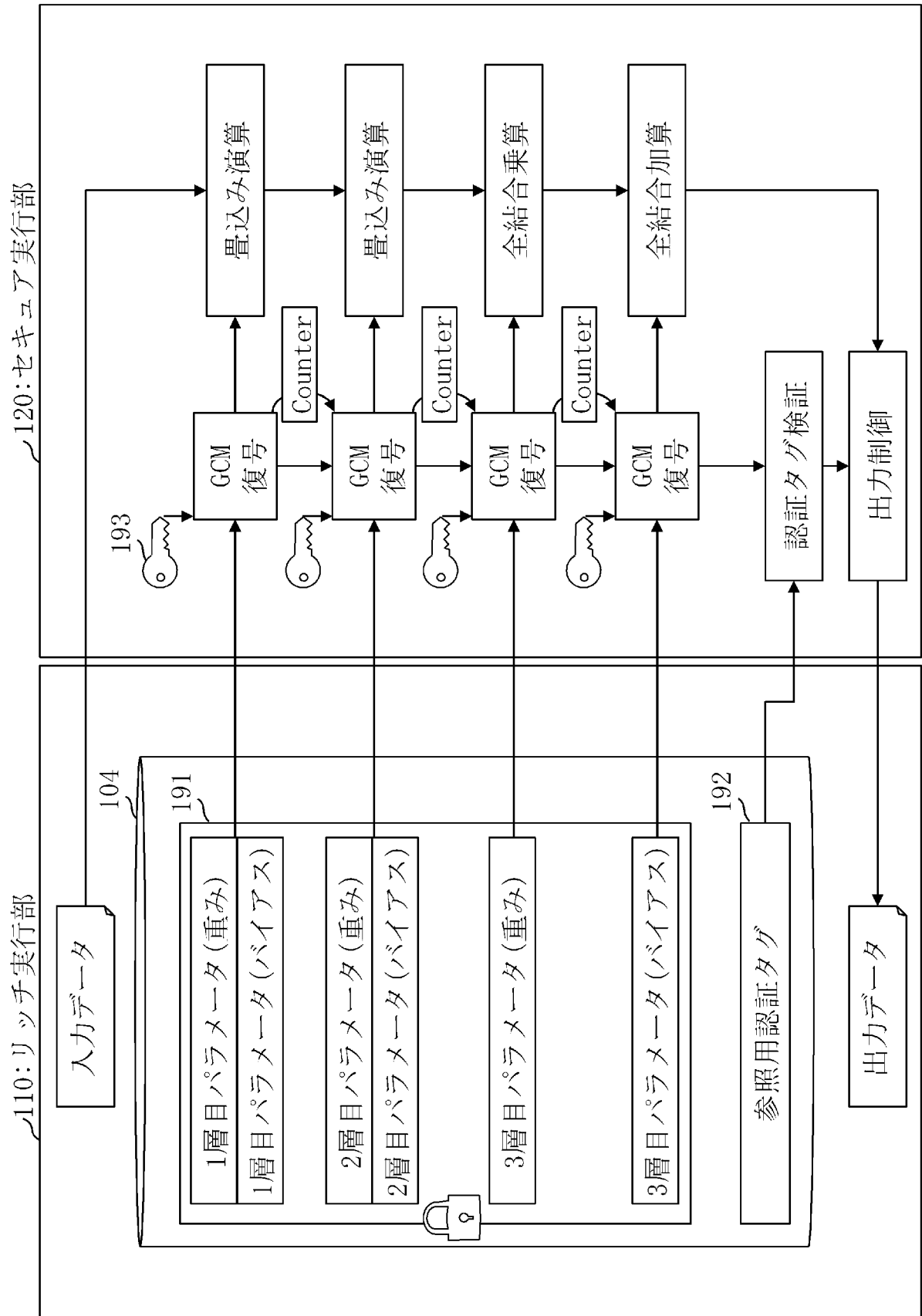


[図7]
図7



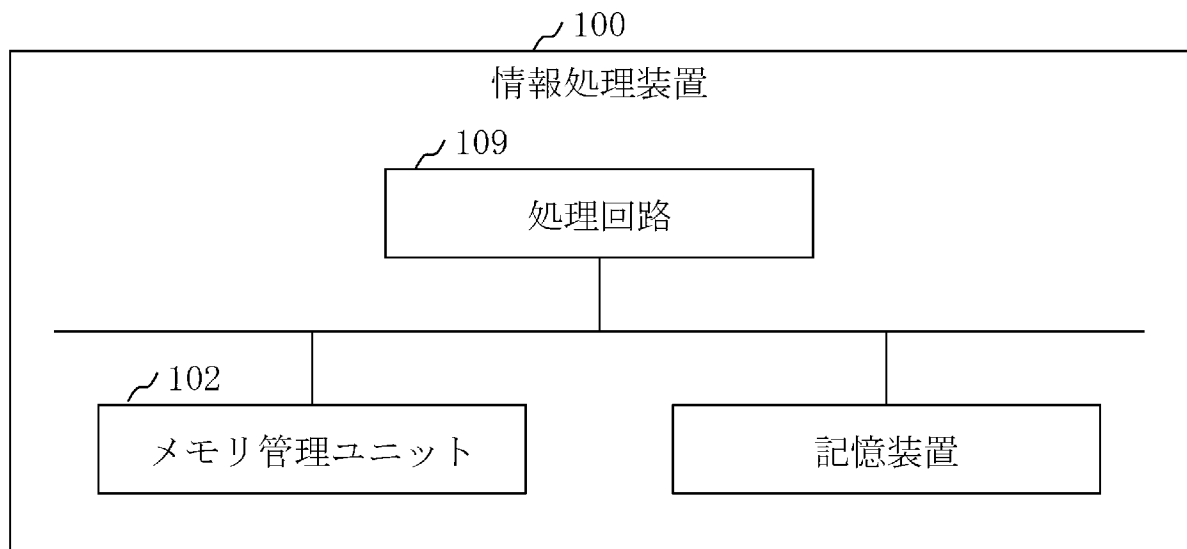
[図8]

図8



[図9]

図9



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2020/047098

A. CLASSIFICATION OF SUBJECT MATTER

G06F 21/74(2013.01)i; G06F 21/53(2013.01)i; G06F 21/64(2013.01)i; G06N 20/00(2019.01)i; H04L 9/32(2006.01)i

FI: G06F21/74; G06F21/64; G06N20/00; H04L9/00 675A; G06F21/53

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)

G06F21/74; G06F21/53; G06F21/64; G06N20/00; H04L9/32

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)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	JP 2019-109680 A (DAINIPPON PRINTING CO., LTD.) 04 July 2019 (2019-07-04) paragraphs [0017]-[0088], [0102]-[0113] (Family: none)	1-5, 7-9
Y A	VANNOSTRAND, Peter M. et al., "Confidential Deep Learning: Executing Proprietary Models on Untrusted Devices", arXiv.org, 28 August 2019, arXiv:1908.10730v1, pp. 1-6, [retrieved on 10 March 2021], Retrieved from the Internet: <URL:https://arxiv.org/abs/1908.10730v1>, in particular, [2. Background], [4. Model Partitioning to Support Confidential Deep Learning]	1-5, 7-9 6
Y	JP 2019-121141 A (DAINIPPON PRINTING CO., LTD.) 22 July 2019 (2019-07-22) paragraphs [0058]-[0072], [0089]-[0097], [0103], [0138]-[0151] (Family: none)	1-5, 7-9

☐

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

"I" 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
08 March 2021 (08.03.2021)Date of mailing of the international search report
23 March 2021 (23.03.2021)Name and mailing address of the ISA/
Japan Patent Office
3-4-3, Kasumigaseki, Chiyoda-ku,
Tokyo 100-8915, Japan

Authorized officer

Telephone No.

A. 発明の属する分野の分類（国際特許分類（IPC）） G06F 21/74(2013.01)i; G06F 21/53(2013.01)i; G06F 21/64(2013.01)i; G06N 20/00(2019.01)i; H04L 9/32(2006.01)i FI: G06F21/74; G06F21/64; G06N20/00; H04L9/00 675A; G06F21/53		
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）） G06F21/74; G06F21/53; G06F21/64; G06N20/00; H04L9/32 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）		
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
Y	JP 2019-109680 A（大日本印刷株式会社）04.07.2019（2019-07-04） 段落0017-0088, 0102-0113 （ファミリーなし）	1-5, 7-9
Y	VANNOSTRAND, Peter M. et al., Confidential Deep Learning: Executing Proprietary Models on Untrusted Devices, arXiv.org, 2019.08.28, arXiv:1908.10730v1, pp. 1-6, [retrieved on 2021.03.10], Retrieved from the Internet: <URL: https://arxiv.org/abs/1908.10730v1> 特に「2. Background」及び「4. Model Partitioning to Support Confidential Deep Learning」	1-5, 7-9
A		6
Y	JP 2019-121141 A（大日本印刷株式会社）22.07.2019（2019-07-22） 段落0058-0072, 0089-0097, 0103, 0138-0151 （ファミリーなし）	1-5, 7-9
☐ 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、一般的な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 08.03.2021		国際調査報告の発送日 23.03.2021
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） 金沢 史明 5S 4538 電話番号 03-3581-1101 内線 3546