

(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局

(43) 国際公開日
2020年3月5日(05.03.2020)



(10) 国際公開番号

WO 2020/044567 A1

(51) 国際特許分類:
G06N 3/08 (2006.01)

(21) 国際出願番号: PCT/JP2018/032484

(22) 国際出願日: 2018年8月31日(31.08.2018)

(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: オリンパス株式会社 (**OLYMPUS CORPORATION**) [JP/JP]; 〒1928507 東京都八王子市石川町2951番地 Tokyo (JP).

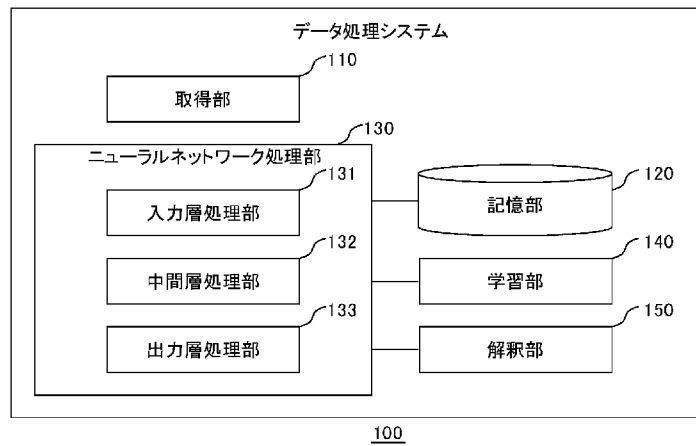
(72) 発明者: 矢口 陽一 (**YAGUCHI Yoichi**); 〒1928507 東京都八王子市石川町2951番地オリンパス株式会社内 Tokyo (JP).

(74) 代理人: 森下 賢樹 (**MORISHITA Sakaki**); 〒1500021 東京都渋谷区恵比寿西2-11-12 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, 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,

(54) **Title:** DATA PROCESSING SYSTEM AND DATA PROCESSING METHOD

(54) 発明の名称: データ処理システムおよびデータ処理方法



- 100 Data processing system
- 110 Acquisition unit
- 120 Storage unit
- 130 Neural network processing unit
- 131 Input layer processing unit
- 132 Intermediate layer processing unit
- 133 Output layer processing unit
- 140 Learning unit
- 150 Interpretation unit

(57) **Abstract:** A data processing system 100 is provided with: a neural network processing unit 130 that executes processing in accordance with a neural network including an input layer, one or more intermediate layers, and an output layer; and a learning unit 140 that causes the neural network to carry out learning by optimizing an optimization target parameter for the neural network on the basis of a comparison between output data outputted as a result of execution of the processing on learning data, by the network processing unit 130, in accordance with the neural network and ideal output data for the

SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA,
UG, US, UZ, VC, VN, ZA, ZM, ZW.

- (84) 指定国(表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, RU, TJ, TM), ヨーロッパ (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

添付公開書類:

- 一 国際調査報告 (条約第21条(3))

learning data. In the learning, the neural network processing unit 130 executes, on intermediate data representing input data to an intermediate layer element constituting an M-th (M is an integer of 1 or greater) intermediate layer or output data from the intermediate layer element, coefficient processing of performing multiplication by a coefficient the absolute value of which monotonically increases according to the degree of progress of learning.

(57) 要約: データ処理システム 100 は、入力層、1 以上の中間層および出力層を含むニューラルネットワークにしたがった処理を実行するニューラルネットワーク処理部 130 と、ニューラルネットワーク処理部 130 が学習データに対してニューラルネットワークにしたがった処理を実行することにより出力される出力データと、その学習データに対する理想的な出力データとの比較に基づいて、ニューラルネットワークの最適化対象パラメータを最適化することにより、ニューラルネットワークを学習させる学習部 140 と、を備える。ニューラルネットワーク処理部 130 は、学習において、第M層 (Mは1 以上の整数) の中間層を構成する中間層要素への入力データまたは当該中間層要素からの出力データを表す中間データに対して、学習の進行度に応じて絶対値が単調増加する係数を乗じる係数処理を実行する。

明 細 書

発明の名称：データ処理システムおよびデータ処理方法

技術分野

[0001] 本発明は、データ処理技術に関し、特に、学習された深層ニューラルネットワークを用いたデータ処理技術に関する。

背景技術

[0002] ニューラルネットワークは、1以上の非線形ユニットを含む数学的モデルであり、入力に対応する出力を予測する機械学習モデルである。多くのニューラルネットワークは、入力層と出力層の他に、1以上の中間層（隠れ層）をもつ。各中間層の出力は次の層（中間層または出力層）の入力となる。ニューラルネットワークの各層は、入力および自身のパラメータに応じて出力を生成する。

先行技術文献

非特許文献

[0003] 非特許文献1：Alex Krizhevsky, Ilya Sutskever, Geoffrey E. Hinton, 「ImageNet Classification with Deep Convolutional Neural Networks」、NIPS2012_4824

非特許文献2：Sergey Ioffe, Christian Szegedy, 「Batch normalization: Accelerating deep network training by reducing internal covariate shift」、ICML 2015 448?456

発明の概要

発明が解決しようとする課題

[0004] 一般に、ネットワーク全体の入出力の関係が大きく変化すると、学習が困難になる。非特許文献2では、入力ミニバッチの統計量を利用して次の層への入力を正規化して入出力の関係が大きく変化するのを抑えることで、学習の困難さを解決している。しかしながら、過剰な正規化はネットワークの表現力の低下にもつながる。一方、ネットワーク全体の入出力の関係が大きく

変化する問題は、中間層のパラメータの更新量が大きい学習初期に顕著となる。

[0005] 本発明はこうした状況に鑑みなされたものであり、その目的は、ニューラルネットワークの学習を容易にする技術を提供することにある。

課題を解決するための手段

[0006] 上記課題を解決するために、本発明のある態様のデータ処理システムは、入力層、1以上の中間層および出力層を含むニューラルネットワークにしたがった処理を実行するニューラルネットワーク処理部と、ニューラルネットワーク処理部が学習データに対して処理を実行することにより出力される出力データと、学習データに対する理想的な出力データとの比較に基づいて、ニューラルネットワークの最適化対象パラメータを最適化することにより、ニューラルネットワークを学習させる学習部と、を備える。ニューラルネットワーク処理部は、学習において、第M層（Mは1以上の整数）の中間層を構成する中間層要素への入力データまたは当該中間層要素からの出力データを表す中間データに対して、学習の進行度に応じて絶対値が単調増加する係数を乗じる係数処理を実行する。

[0007] 本発明の別の態様もまた、データ処理システムである。このデータ処理システムは、入力層、1以上の中間層および出力層を含むニューラルネットワークにしたがった処理を実行するニューラルネットワーク処理部を備える。ニューラルネットワーク処理部は、学習データに対して処理を実行することにより出力される出力データと、学習データに対する理想的な出力データとの比較に基づいて、ニューラルネットワークの最適化対象パラメータを最適化することにより学習されており、ニューラルネットワーク処理部は、学習において、第M層（Mは1以上の整数）の中間層を構成する中間層要素への入力データまたは当該中間層要素からの出力データを表す中間データに対して、学習の進行度に応じて絶対値が単調増加する係数を乗じる係数処理を実行する。

[0008] 本発明のさらに別の態様は、データ処理方法である。この方法は、学習デ

ータに対して、入力層、1以上の中間層および出力層を含むニューラルネットワークにしたがった処理を実行することにより、学習データに対応する出力データを出力するステップと、学習データに対応する出力データと、学習データに対する理想的な出力データとの比較に基づいて、ニューラルネットワークの最適化対象パラメータを最適化するステップと、を備える。最適化対象パラメータを最適化するステップでは、第M層（Mは1以上の整数）の中間層を構成する中間層要素への入力データまたは当該中間層要素からの出力データを表す中間データに対して、学習の進行度に応じて絶対値が単調増加する係数を乗じる係数処理を実行する。

[0009] 本発明のさらに別の態様もまた、データ処理方法である。この方法は、入力層、1以上の中間層および出力層を含むニューラルネットワークにしたがった処理を実行するステップを備える。ニューラルネットワークは、学習データに対して処理を実行することにより出力される出力データと、学習データに対する理想的な出力データとの比較に基づいて、最適化対象パラメータが最適化されており、学習において、第M層（Mは1以上の整数）の中間層を構成する中間層要素への入力データまたは当該中間層要素からの出力データを表す中間データに対して、学習の進行度に応じて絶対値が単調増加する係数を乗じる係数処理が実行されている。

[0010] なお、以上の構成要素の任意の組み合わせ、本発明の表現を方法、装置、システム、記録媒体、コンピュータプログラムなどの間で変換したものもまた、本発明の態様として有効である。

発明の効果

[0011] 本発明によれば、ニューラルネットワークの学習を容易にする技術を提供できる。

図面の簡単な説明

[0012] [図1]実施の形態に係るデータ処理システムの機能および構成を示すブロック図である。

[図2]ニューラルネットワークの構成の一例を模式的に示す図である。

[図3]データ処理システムによる学習処理のフローチャートを示す図である。

[図4]データ処理システムによる適用処理のフローチャートを示す図である。

[図5]ニューラルネットワークの構成の他の一例を模式的に示す図である。

[0013] 以下、本発明を好適な実施の形態をもとに図面を参照しながら説明する。

[0014] 以下ではデータ処理装置を画像処理に適用する場合を例に説明するが、当業者によれば、データ処理装置を音声認識処理、自然言語処理、その他の処理にも適用可能であることが理解されよう。

[0015] 図1は、実施の形態に係るデータ処理システム100の機能および構成を示すブロック図である。ここに示す各ブロックは、ハードウェア的には、コンピュータのCPU (central processing unit) をはじめとする素子や機械装置で実現でき、ソフトウェア的にはコンピュータプログラム等によって実現されるが、ここでは、それらの連携によって実現される機能ブロックを描いている。したがって、これらの機能ブロックはハードウェア、ソフトウェアの組合せによっていろいろなかたちで実現できることは、当業者には理解されるところである。

[0016] データ処理システム100は、学習用の画像（学習データ）と、その画像に対する理想的な出力データである正解値とに基づいてニューラルネットワークの学習を行う「学習処理」と、学習済みのニューラルネットワークを未知の画像（未知データ）に適用し、画像分類、物体検出または画像セグメンテーションなどの画像処理を行う「適用処理」と、を実行する。

[0017] 学習処理では、データ処理システム100は、学習用の画像に対してニューラルネットワークにしたがった処理を実行し、学習用の画像に対する出力データを出力する。そしてデータ処理システム100は、出力データが正解値に近づく方向にニューラルネットワークの最適化（学習）対象のパラメータ（以下、「最適化対象パラメータ」と呼ぶ）を更新する。これを繰り返すことにより最適化対象パラメータが最適化される。

[0018] 適用処理では、データ処理システム100は、学習処理において最適化された最適化対象パラメータを用いて、未知の画像に対してニューラルネット

ワークにしたがった処理を実行し、その画像に対する出力データを出力する。データ処理システム１００は、出力データを解釈して、画像を画像分類したり、画像から物体検出したり、画像に対して画像セグメンテーションを行ったりする。

[0019] データ処理システム１００は、取得部１１０と、記憶部１２０と、ニューラルネットワーク処理部１３０と、学習部１４０と、解釈部１５０と、を備える。主にニューラルネットワーク処理部１３０と学習部１４０により学習処理の機能が実現され、主にニューラルネットワーク処理部１３０と解釈部１５０により適用処理の機能が実現される。

[0020] 取得部１１０は、学習処理においては、一度に複数の学習用の画像と、それら複数の学習用の画像のそれぞれに対応する正解値とを取得する。また取得部１１０は、適用処理においては、処理対象の未知の画像を取得する。なお、画像は、チャンネル数は特に問わず、例えばRGB画像であっても、また例えばグレースケール画像であってもよい。

[0021] 記憶部１２０は、取得部１１０が取得した画像を記憶する他、ニューラルネットワーク処理部１３０、学習部１４０および解釈部１５０のワーク領域や、ニューラルネットワークのパラメータの記憶領域となる。

[0022] ニューラルネットワーク処理部１３０は、ニューラルネットワークにしたがった処理を実行する。ニューラルネットワーク処理部１３０は、ニューラルネットワークの入力層に対応する処理を実行する入力層処理部１３１と、中間層に対応する処理を実行する中間層処理部１３２と、出力層に対応する処理を実行する出力層処理部１３３と、を含む。

[0023] 図２は、ニューラルネットワークの構成の一例を模式的に示す図である。この例では、ニューラルネットワークは２つの中間層を含み、各中間層は畳み込み処理を行う中間層要素とプーリング処理を行う中間層要素とを含んで構成されている。なお、中間層の数は特に限定されず、例えば中間層の数が１であっても、３以上であってもよい。図示の例の場合、中間層処理部１３２は、各中間層の各中間層要素の処理を実行する。

[0024] また、本実施の形態では、ニューラルネットワークは、少なくとも1つの係数要素を含む。図示の例では、ニューラルネットワークは各中間層の前後に係数要素を含んでいる。中間層処理部132は、この係数要素に対応する処理も実行する。

[0025] 中間層処理部132は、学習処理時は、係数要素に対応する処理として係数処理を実行する。係数処理とは、中間層要素への入力データまたは中間層要素からの出力データを表す中間データに対して、学習の進行度に応じて絶対値が単調増加（広義単調増加）する係数を乗じる処理をいう。本実施の形態の係数処理では、当該中間データに対して、学習の進行度に応じて絶対値が0以上1以下の範囲で単調増加する係数を乗じる。なお、本実施の形態では、学習の進行度は学習の繰り返し回数であるとする。

[0026] 係数処理は、一例として以下の式（1）により与えられる。

[数1]

$$y = (1 - \alpha^t)x \cdots (1)$$

x : 入力

y : 出力

α : 係数の増幅速度を規定するハイパーパラメータ

t : 学習の繰り返し回数

ここで、 α には、0より大きく1より小さい値（例えば0.999）が設定される。したがって、 α^t は、0より大きく1より小さい範囲で、学習が進むにつれて徐々に小さくなる。したがって、係数（ $1 - \alpha^t$ ）は、0より大きく1より小さい範囲で、学習が進むにつれて単調増加する。係数（ $1 - \alpha^t$ ）は特に、学習が進むにつれて1に近づく。この場合、中間データは、学習の初期は比較的小さい値に変換され、学習が進むにつれて徐々に変換度合いが小さくなり、学習の後期では、1に近い値を乗じることから明らかなように、実質的に変換がないものと見なせる程度に変換される。

[0027] また、中間層処理部132は、適用処理時は、係数処理として以下の式（2）により与えられる処理を実行する。つまり、入力をそのまま出力する処

理を実行する。なお別の捉え方をすると、中間層処理部132は、適用処理時は、係数処理として1を乗じる処理を実行するとも言える。いずれにしろ、本発明を利用しない場合と同程度の処理時間で適用処理を実行できる。

[数2]

$$y = x \cdots (2)$$

[0028] 学習部140は、ニューラルネットワークの最適化対象パラメータを最適化することにより、ニューラルネットワークを学習させる。学習部140は、学習用の画像をニューラルネットワーク処理部130に入力することにより得られた出力と、その画像に対応する正解値とを比較する目的関数（誤差関数）により、誤差を算出する。学習部140は、算出された誤差に基づいて、勾配逆伝搬法等によりパラメータについての勾配を計算し、モーメンタム法に基づいてニューラルネットワークの最適化対象パラメータを更新する。

[0029] 取得部110による学習用の画像の取得と、ニューラルネットワーク処理部130による学習用画像に対するニューラルネットワークにしたがった処理と、学習部140による最適化対象パラメータの更新とを繰り返すことにより、最適化対象パラメータが最適化される。

[0030] また、学習部140は、学習を終了すべきか否かを判定する。学習を終了すべき終了条件は、例えば学習が所定回数行われたことや、外部から終了の指示を受けたことや、最適化対象パラメータの更新量の平均値が所定値に達したことや、算出された誤差が所定の範囲内に収まったことである。学習部140は、終了条件が満たされる場合、学習処理を終了させる。学習部140は、終了条件が満たされない場合、処理をニューラルネットワーク処理部130に戻す。

[0031] 解釈部150は、出力層処理部133からの出力を解釈して、画像分類、物体検出または画像セグメンテーションを実施する。

[0032] 実施の形態に係るデータ処理システム100の動作を説明する。

図3は、データ処理システム100による学習処理のフローチャートを示

す。取得部110は、複数枚の学習用の画像を取得する(S10)。ニューラルネットワーク処理部130は、取得部110が取得した複数枚の学習用の画像のそれぞれに対して、ニューラルネットワークにしたがった処理を実行し、それぞれについての出力データを出力する(S12)。学習部140は、複数枚の学習用の画像のそれぞれについての出力データと、それぞれについての正解値とに基づいて、パラメータを更新する(S14)。学習部140は、終了条件が満たされるか否かを判定する(S16)。終了条件が満たされない場合(S16のN)、処理はS10に戻される。終了条件が満たされる場合(S16のY)、処理は終了する。

[0033] 図4は、データ処理システム100による適用処理のフローチャートを示す。取得部110は、適用処理の対象の画像を取得する(S20)。ニューラルネットワーク処理部130は、取得部110が取得した画像に対して、最適化対象パラメータが最適化されたすなわち学習済みのニューラルネットワークにしたがった処理を実行し、出力データを出力する(S22)。解釈部150は、出力データを解釈し、対象の画像を画像分類したり、対象の画像から物体検出したり、対象の画像に対して画像セグメンテーションを行ったりする(S24)。

[0034] 以上説明した実施の形態に係るデータ処理システム100によると、係数処理として、中間層要素への入力データまたは中間層要素からの出力データを表す中間データに対して、学習の進行度に応じて絶対値が0以上1以下の範囲で単調増加する係数を乗じる処理が実行される。これにより、学習初期において、ニューラルネットワーク全体の入出力の関係が大きく変化するのを抑えられ、その結果として学習が容易となる。また、係数処理の出力が係数処理への入力よりも大きくならないため、学習の発散を抑制できる。

[0035] 以上、本発明を実施の形態をもとに説明した。この実施の形態は例示であり、それらの各構成要素や各処理プロセスの組合せにいろいろな変形例が可能なこと、またそうした変形例も本発明の範囲にあることは当業者に理解されるところである。

[0036] (変形例1)

図5は、ニューラルネットワークの構成の他の一例を模式的に示す図である。この例では、第M（Mは1以上の整数）層の中間層は、1以上の中間層要素を含む。ニューラルネットワーク処理部130は、第M層の処理において、その中間層要素への入力データを表す中間データおよびその中間層要素からの出力データを表す中間データの少なくとも一方に対して係数処理を実行する。図示の例では、ニューラルネットワーク処理部130は、第M層の中間層を構成する1以上の中間層要素のうちの最初の中間層要素への入力データを表す中間データと、最後の中間層要素からの出力データを表す中間データに対して係数処理を実行する。

[0037] また、ニューラルネットワーク処理部130は、第M層の中間層に入力されるべき中間データと、当該中間データを第M層の中間層に入力することにより出力される中間データとを統合する統合処理を実行する。例えば、ニューラルネットワーク処理部130は、統合処理として、第M層の中間層に入力されるべき中間データと、当該中間データを第M層の中間層に入力することにより出力される中間データとを足し合わせてもよい。この場合のニューラルネットワークは、Residual networksに係数要素を含めたものに相当する。また例えば、ニューラルネットワーク処理部130は、統合処理として、第M層の中間層に入力されるべき中間データと当該中間データを第M層の中間層に入力することにより出力される中間データとをチャンネル連結してもよい。この場合のニューラルネットワークは、Densely connected networksに係数要素を含めたものに相当する。

[0038] 本変形例によれば、ニューラルネットワーク全体の入出力の関係が恒等写像に近くなるため学習が容易になる。なお具体的には、第M層の中間層を構成する1以上のうちの中間層要素のうちの最初の中間層要素への入力データを表す中間データに対して係数処理を実行すると、順伝搬が恒等写像に近くなり、最後の中間層要素からの出力データを表す中間データに対して係数処理を実行すると、逆伝搬が恒等写像に近くなる。

[0039] (変形例2)

係数処理において、係数が1に十分近づいたら、言い換えると、1と係数との差分値が所定の値以下となったら、係数を乗算しなくてもよい。具体的には例えば、係数処理は、以下の式(3)により与えられてもよい。

[数3]

$$y = \begin{cases} (1 - \alpha^t)x & \dots \alpha^t > \epsilon \\ x & \dots \alpha^t \leq \epsilon \end{cases} \dots (3)$$

ϵ : 係数乗算を無視する程度を規定するハイパーパラメータ

上述したように、 α^t は、0より大きく1より小さい範囲で、学習が進むにつれて徐々に小さくなる。係数 $(1 - \alpha^t)$ は、0より大きく1より小さい範囲で、学習が進むにつれて1に近づく。そして、この例では、係数 $(1 - \alpha^t)$ が1にある程度以上近づくと、すなわち1と係数 $(1 - \alpha^t)$ との差分値が ϵ よりも小さくなると、係数を乗算せずに、入力をそのまま出力する処理を実行する。本発明によれば、学習の途中からは、本発明を利用しない場合と同程度の処理時間で学習処理を実行できる。

[0040] (変形例3)

実施の形態では、学習の進行度は学習の繰り返し回数である場合について説明したが、これに限定されない。例えば、学習の収束具合を学習の進行度としてもよい。この場合、進行度は例えば、学習データをニューラルネットワークに入力して得られた出力と、その学習データに対する理想的な出力データである正解値との差に対して単調減少する関数に基づく値であってもよい。具体的には例えば、以下の式4に基づく値であってもよい。

[数4]

$$t = \frac{1}{L} \dots (4)$$

L : 学習用の画像をニューラルネットワーク処理部130に

入力することにより得られた出力と、その画像に対応する正解値とを比較する目的関数(誤差関数)によって算出された誤差の値

[0041] 実施の形態および変形例において、データ処理システムは、プロセッサと

、メモリ等のストレージを含んでもよい。ここでのプロセッサは、例えば各部の機能が個別のハードウェアで実現されてもよいし、あるいは各部の機能が一体のハードウェアで実現されてもよい。例えば、プロセッサはハードウェアを含み、そのハードウェアは、デジタル信号を処理する回路およびアナログ信号を処理する回路の少なくとも一方を含むことができる。例えば、プロセッサは、回路基板に実装された１又は複数の回路装置（例えばＩＣ等）や、１又は複数の回路素子（例えば抵抗、キャパシター等）で構成することができる。プロセッサは、例えばＣＰＵ（Central Processing Unit）であってもよい。ただし、プロセッサはＣＰＵに限定されるものではなく、ＧＰＵ（Graphics Processing Unit）、あるいはＤＳＰ（Digital Signal Processor）等、各種のプロセッサを用いることが可能である。またプロセッサはＡＳＩＣ（application specific integrated circuit）又はＦＰＧＡ（field-programmable gate array）によるハードウェア回路でもよい。またプロセッサは、アナログ信号を処理するアンプ回路やフィルター回路等を含んでもよい。メモリは、ＳＲＡＭ、ＤＲＡＭなどの半導体メモリであってもよいし、レジスタであってもよいし、ハードディスク装置等の磁気記憶装置であってもよいし、光学ディスク装置等の光学式記憶装置であってもよい。例えば、メモリはコンピュータにより読み取り可能な命令を格納しており、当該命令がプロセッサにより実行されることで、データ処理システムの各部の機能が実現されることになる。ここでの命令は、プログラムを構成する命令セットの命令でもよいし、プロセッサのハードウェア回路に対して動作を指示する命令であってもよい。

符号の説明

[0042] １００ データ処理システム、 １３０ ニューラルネットワーク処理部、
 １４０ 学習部。

産業上の利用可能性

[0043] 本発明は、データ処理システムおよびデータ処理方法に利用できる。

請求の範囲

- [請求項1] 入力層、1 以上の中間層および出力層を含むニューラルネットワークにしたがった処理を実行するニューラルネットワーク処理部と、
- 前記ニューラルネットワーク処理部が学習データに対して前記処理を実行することにより出力される出力データと、前記学習データに対する理想的な出力データとの比較に基づいて、前記ニューラルネットワークの最適化対象パラメータを最適化することにより、前記ニューラルネットワークを学習させる学習部と、を備え、
- 前記ニューラルネットワーク処理部は、前記学習において、第M層（Mは1 以上の整数）の中間層を構成する中間層要素への入力データまたは当該中間層要素からの出力データを表す中間データに対して、学習の進行度に応じて絶対値が単調増加する係数を乗じる係数処理を実行することを特徴とするデータ処理システム。
- [請求項2] 入力層、1 以上の中間層および出力層を含むニューラルネットワークにしたがった処理を実行するニューラルネットワーク処理部を備え、
- 、
- 前記ニューラルネットワークは、学習データに対して前記処理を実行することにより出力される出力データと、前記学習データに対する理想的な出力データとの比較に基づいて、最適化対象パラメータが最適化されており、
- 前記ニューラルネットワーク処理部は、前記学習において、第M層（Mは1 以上の整数）の中間層を構成する中間層要素への入力データまたは当該中間層要素からの出力データを表す中間データに対して、学習の進行度に応じて絶対値が単調増加する係数を乗じる係数処理を実行することを特徴とするデータ処理システム。
- [請求項3] 前記係数の絶対値は、0 以上1 以下であることを特徴とする請求項1 または2 に記載のデータ処理システム。
- [請求項4] 前記ニューラルネットワーク処理部は、1 と前記係数との差分値が

所定の値以下となった場合、係数処理として入力をそのまま出力する処理を実行することを特徴とする請求項 1 から 3 のいずれかに記載のデータ処理システム。

[請求項5] 前記ニューラルネットワーク処理部は、適用処理時は、係数処理として入力をそのまま出力する処理を実行することを特徴とする請求項 1 から 4 のいずれかに記載のデータ処理システム。

[請求項6] 前記第M層の中間層は、1 以上の中間層要素を含み、
前記ニューラルネットワーク処理部は、(i) 前記第M層の中間層の処理において、その中間層要素への入力データを表す中間データおよび当該中間層要素からの出力データを表す中間データの少なくとも一方に対して係数処理を実行し、かつ、(ii) 当該第M層の中間層に入力されるべき中間データと、当該中間データを当該第M層の中間層に入力することにより出力される中間データとを統合する統合処理を実行することを特徴とする請求項 1 から 5 のいずれかに記載のデータ処理システム。

[請求項7] 前記ニューラルネットワーク処理部は、前記第M層の中間層の最初の間層要素への入力データを表す中間データに対して係数処理を実行することを特徴とする請求項 6 に記載のデータ処理システム。

[請求項8] 前記ニューラルネットワーク処理部は、前記第M層の中間層の最後の間層要素からの出力データを表す中間データに対して係数処理を実行することを特徴とする請求項 6 に記載のデータ処理システム。

[請求項9] 前記ニューラルネットワーク処理部は、前記統合処理として、それぞれの中間データを足し合わせることを特徴とする請求項 6 から 8 のいずれかに記載のデータ処理システム。

[請求項10] 前記ニューラルネットワーク処理部は、前記統合処理として、それぞれの中間データをチャンネル連結することを特徴とする請求項 6 から 8 のいずれかに記載のデータ処理システム。

[請求項11] 学習の進行度は、学習の繰り返し回数であることを特徴とする請求

項 1 から 10 のいずれかに記載のデータ処理システム。

[請求項12] 学習の進行度は、学習データに対して前記処理を実行することにより出力される出力データとその学習データに対する理想的な出力データとの差に対して単調減少する関数に基づき決定されることを特徴とする請求項 1 から 10 のいずれかに記載のデータ処理システム。

[請求項13] 学習データに対して、入力層、1 以上の中間層および出力層を含むニューラルネットワークにしたがった処理を実行することにより、学習データに対応する出力データを出力するステップと、

学習データに対応する出力データと、前記学習データに対する理想的な出力データとの比較に基づいて、前記ニューラルネットワークの最適化対象パラメータを最適化するステップと、を備え、

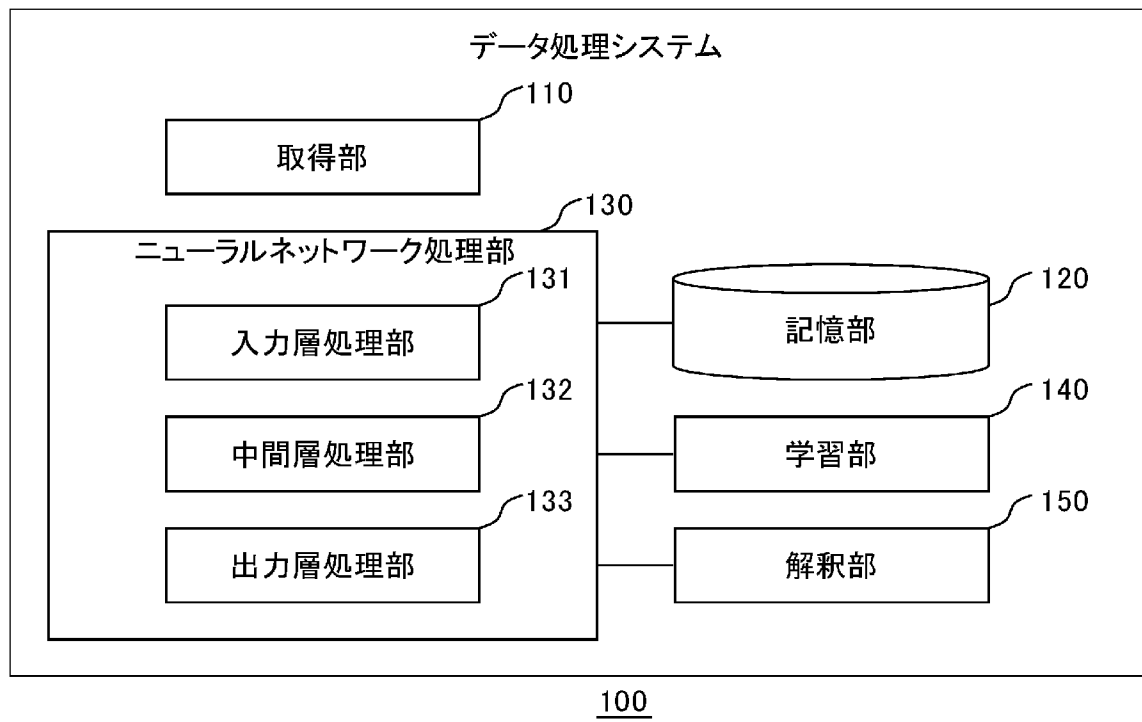
前記最適化対象パラメータを最適化するステップでは、第M層（Mは1以上の整数）の中間層を構成する中間層要素への入力データまたは当該中間層要素からの出力データを表す中間データに対して、学習の進行度に応じて絶対値が単調増加する係数を乗じる係数処理を実行することを特徴とするデータ処理方法。

[請求項14] 入力層、1 以上の中間層および出力層を含むニューラルネットワークにしたがった処理を実行するステップを備え、

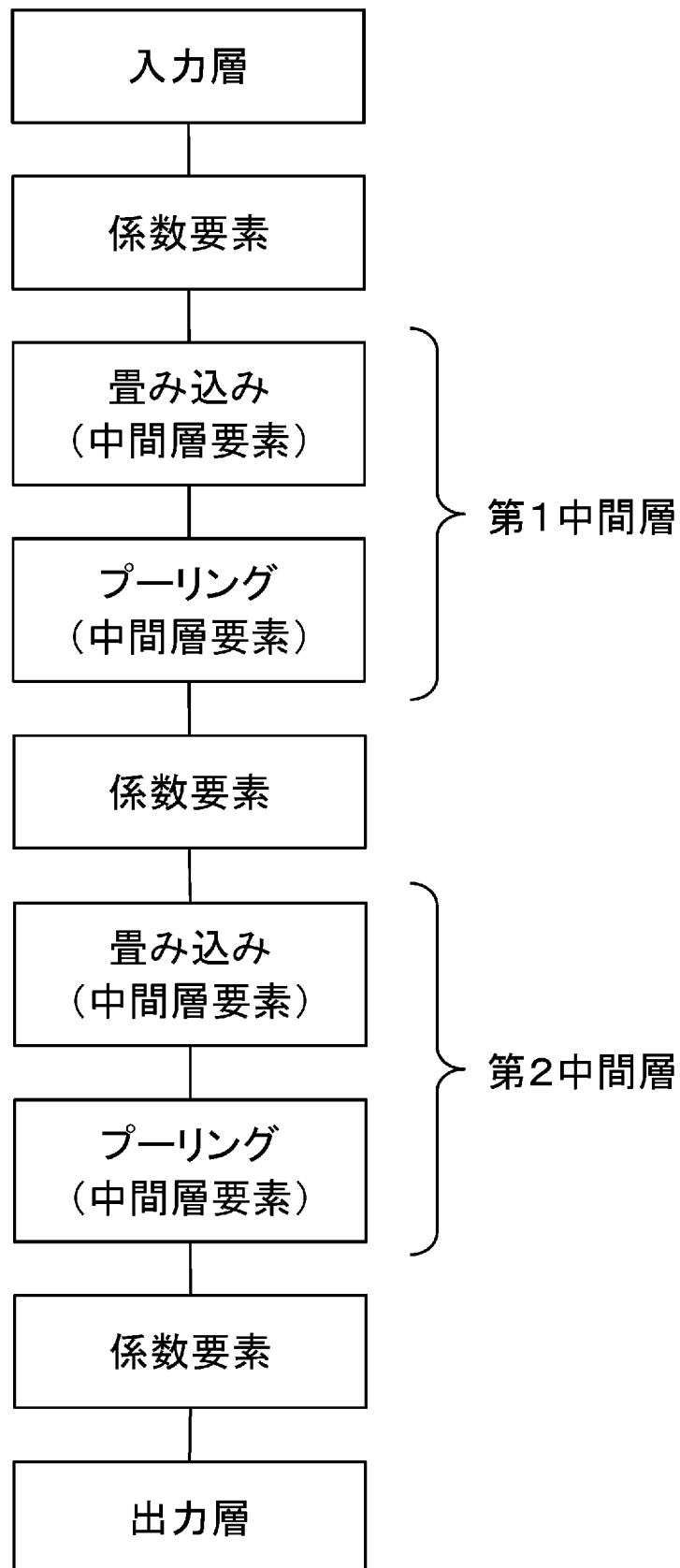
前記ニューラルネットワークは、学習データに対して前記処理を実行することにより出力される出力データと、前記学習データに対する理想的な出力データとの比較に基づいて、最適化対象パラメータが最適化されており、

前記学習において、第M層（Mは1以上の整数）の中間層を構成する中間層要素への入力データまたは当該中間層要素からの出力データを表す中間データに対して、学習の進行度に応じて絶対値が単調増加する係数を乗じる係数処理が実行されていることを特徴とするデータ処理方法。

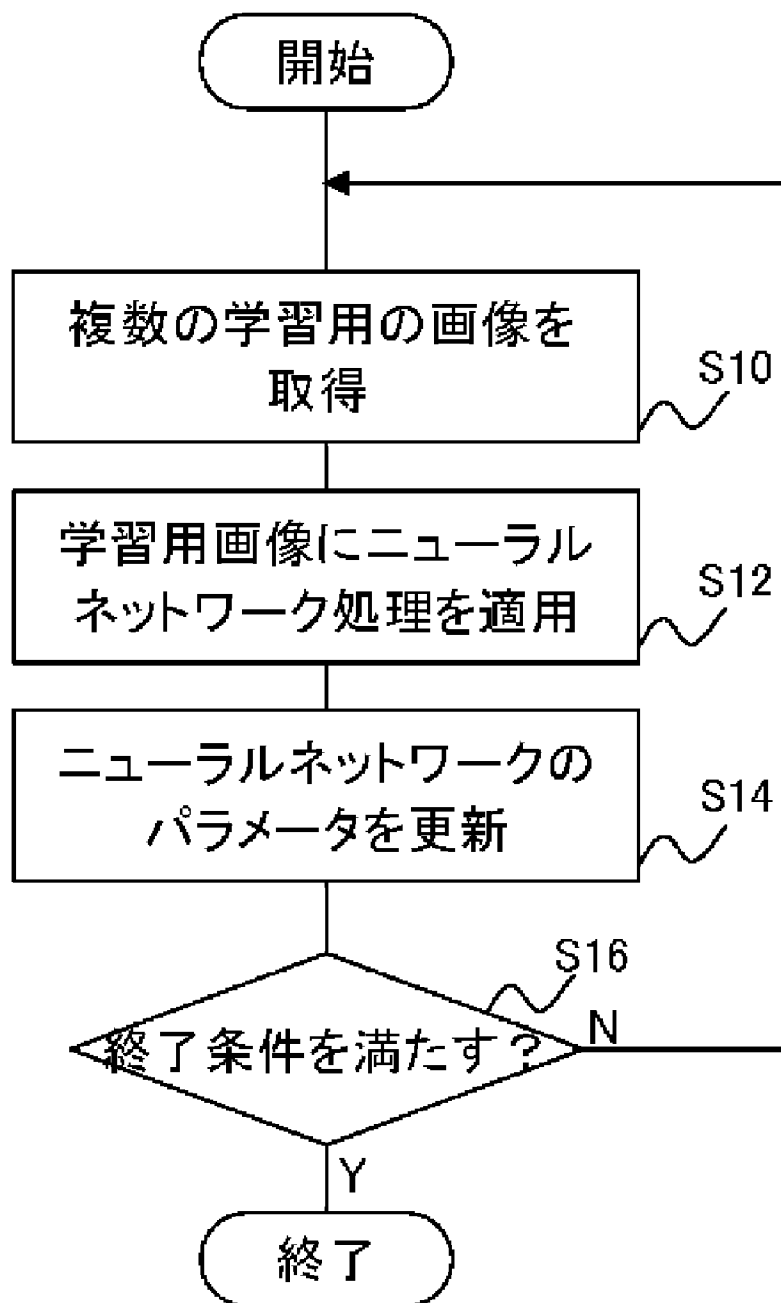
[図1]



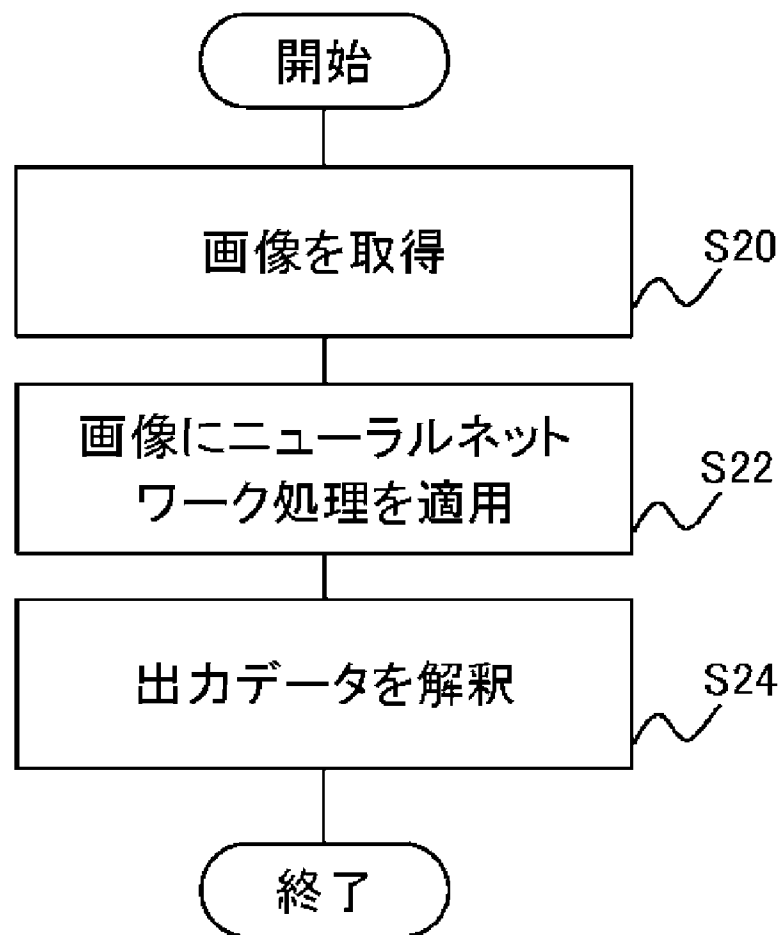
[図2]



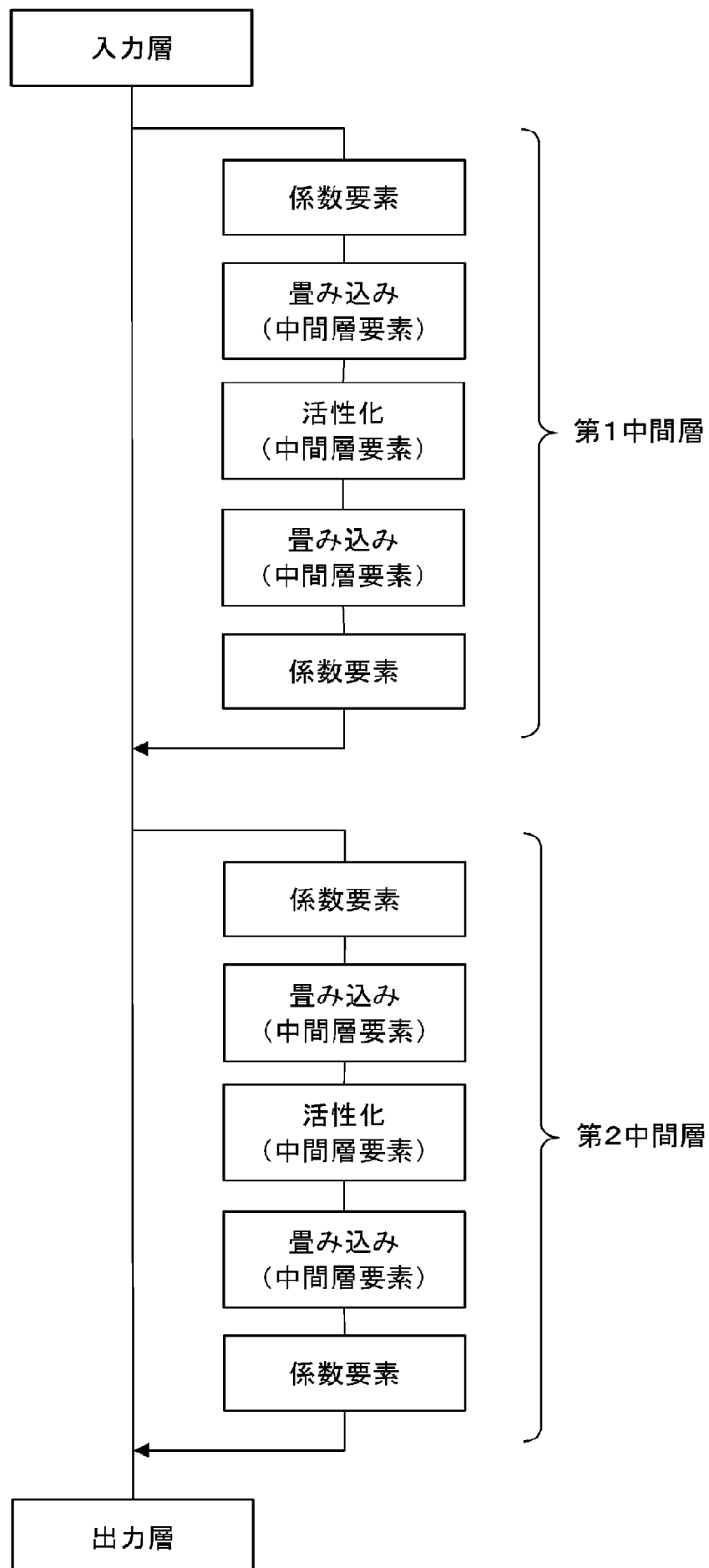
[図3]



[図4]



[図5]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2018/032484

A. CLASSIFICATION OF SUBJECT MATTER

Int. Cl. G06N3/08 (2006.01) i

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

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Int. Cl. G06N3/08

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 1971-2018

Registered utility model specifications of Japan 1996-2018

Published registered utility model applications of Japan 1994-2018

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

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 2017-211939 A (YAHOO JAPAN CORP.) 30 November 2017, paragraphs [0020]-[0038], fig. 1 & US 2017/0345196 A1, paragraphs [0030]-[0043]	1-14
A	JP 2017-134853 A (YAHOO JAPAN CORP.) 03 August 2017, paragraphs [0020]-[0032], fig. 1 (Family: none)	1-14

☐ 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
16.10.2018

Date of mailing of the international search report
30.10.2018

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））

Int.Cl. G06N3/08(2006.01)i

B. 調査を行った分野

調査を行った最小限資料（国際特許分類（IPC））

Int.Cl. G06N3/08

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	1971-2018年
日本国実用新案登録公報	1996-2018年
日本国登録実用新案公報	1994-2018年

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

C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号
A	JP 2017-211939 A（ヤフー株式会社）2017.11.30, 段落[0020]-[0038]、図1 & US 2017/0345196 A1（[0030]-[0043]）	1-14
A	JP 2017-134853 A（ヤフー株式会社）2017.08.03, 段落[0020]-[0032]、図1 （ファミリーなし）	1-14

C欄の続きにも文献が列挙されている。

パテントファミリーに関する別紙を参照。

* 引用文献のカテゴリー

「A」特に関連のある文献ではなく、一般的技術水準を示すもの
「E」国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの
「L」優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す）
「O」口頭による開示、使用、展示等に言及する文献
「P」国際出願日前で、かつ優先権の主張の基礎となる出願

の日の後に公表された文献

「T」国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの
「X」特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの
「Y」特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの
「&」同一パテントファミリー文献

国際調査を完了した日

16.10.2018

国際調査報告の発送日

30.10.2018

国際調査機関の名称及びあて先

日本国特許庁（ISA/J P）

郵便番号100-8915

東京都千代田区霞が関三丁目4番3号

特許庁審査官（権限のある職員）

渡部 博樹

5B

3868

電話番号 03-3581-1101 内線 3545