

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7704766号
(P7704766)

(45)発行日 令和7年7月8日(2025.7.8)

(24)登録日 令和7年6月30日(2025.6.30)

(51)国際特許分類

F I

G 0 6 N 3/0464(2023.01)

G 0 6 N 3/0464

G 1 6 B 30/00 (2019.01)

G 1 6 B 30/00

G 1 6 B 40/00 (2019.01)

G 1 6 B 40/00

請求項の数 20 (全88頁)

(21)出願番号	特願2022-549987(P2022-549987)	(73)特許権者	500358711
(86)(22)出願日	令和3年2月19日(2021.2.19)		イルミナ インコーポレイテッド
(65)公表番号	特表2023-515804(P2023-515804 A)		アメリカ合衆国 カリフォルニア州 9 2
(43)公表日	令和5年4月14日(2023.4.14)		1 2 2 サンディエゴ イルミナ ウェイ
(86)国際出願番号	PCT/US2021/018913	(74)代理人	5 2 0 0
(87)国際公開番号	WO2021/168356		100108453
(87)国際公開日	令和3年8月26日(2021.8.26)		弁理士 村山 靖彦
審査請求日	令和6年2月19日(2024.2.19)	(74)代理人	100110364
(31)優先権主張番号	62/979,399		弁理士 実広 信哉
(32)優先日	令和2年2月20日(2020.2.20)	(74)代理人	100133400
(33)優先権主張国・地域又は機関	米国(US)		弁理士 阿部 達彦
(31)優先権主張番号	62/979,411	(72)発明者	ジェリー・ベッセル
(32)優先日	令和2年2月20日(2020.2.20)		アメリカ合衆国・カリフォルニア・9 2
	最終頁に続く		1 2 2・サン・ディエゴ・イルミナ・ウ
			エイ・5 2 0 0
			最終頁に続く

(54)【発明の名称】 人工知能ベースのベースコールのためのデータ圧縮

(57)【特許請求の範囲】

【請求項1】

ベースコールの人工知能ベースの方法であって、前記方法が、

クラスタを示す画像のサイクルごとの分析物チャネルセットのシリーズにアクセスすることであって、前記画像が、シーケンシングランのシーケンシングサイクルの間に基材の上の関連する分析物のクラスタにおけるヌクレオチド取り込みの結果として生成された強度放射を示し、前記サイクルごとの分析物チャネルセットが、異なる照明および／またはフィルタ波長帯域を適用する結果として生じる、ことと、

畳み込みニューラルネットワークベースのベースコーラーの空間ネットワークを介して、前記シーケンシングランのシーケンシングサイクルの第1のウィンドウに対する、前記シリーズ内の前記クラスタを示す画像のサイクルごとの分析物チャネルセットの第1のウィンドウを処理し、1D、2D、または3Dの畳み込みを適用し、前記シーケンシングサイクルの第1のウィンドウにおけるそれぞれのシーケンシングサイクルに対する空間出力セットのそれぞれのシーケンスを生成することと、

前記畳み込みニューラルネットワークベースのベースコーラーの圧縮ネットワークを介して、前記空間出力セットのそれぞれのシーケンス内のそれぞれの最終空間出力セットを処理し、畳み込みフィルタを適用して前記それぞれの最終空間出力セット内の特徴マップの数を減少させ、前記シーケンシングサイクルの第1のウィンドウにおける前記それぞれのシーケンシングサイクルに対するそれぞれの圧縮空間出力セットを生成することと、

前記それぞれの圧縮空間出力セットに基づいて、前記シーケンシングサイクルの第1の

ウィンドウにおける1つ以上のシーケンシングサイクルに対するベースコール予測を生成することと、を含む、人工知能ベースの方法。

【請求項2】

前記それぞれの最終空間出力セットが、M個のチャネルを有し、前記それぞれの圧縮空間出力セットが、N個のチャネルを有し、MがNより大きい、請求項1に記載の人工知能ベースの方法。

【請求項3】

前記シーケンシングサイクルの第1のウィンドウと、前記空間ネットワークが以前に空間出力セットを生成した1つ以上の重なり合うシーケンシングサイクル、及び少なくとも1つの重なり合わないシーケンシングサイクルを共有する、前記シーケンシングランのシーケンシングサイクルの第2のウィンドウに対して、

10

前記空間ネットワークを介して、前記少なくとも1つの重なり合わないシーケンシングサイクルに対してのみサイクルごとの分析物チャネルセットを処理し、前記少なくとも1つの重なり合わないシーケンシングサイクルのための空間出力セットのシーケンスを生成することと、

前記圧縮ネットワークを介して、前記空間出力セットのシーケンス内の最終空間出力セットを処理し、前記少なくとも1つの重なり合わないシーケンシングサイクルのための圧縮空間出力セットを生成することであって、前記最終空間出力セットが、M個のチャネルを有し、前記圧縮空間出力セットの圧縮空間出力が、N個のチャネルを有し、MがNより大きい、ことと、

20

前記シーケンシングサイクルの第1のウィンドウに対して以前に生成された前記1つ以上の重なり合うシーケンシングサイクルのためのそれぞれの圧縮空間出力セット、及び前記圧縮空間出力セットに基づいて、前記シーケンシングサイクルの第2のウィンドウにおける1つ以上のシーケンシングサイクルに対するベースコール予測を生成することと、を更に含む、請求項1または2に記載の人工知能ベースの方法。

【請求項4】

前記シーケンシングサイクルの第1及び第2のウィンドウと、前記空間ネットワークが以前に空間出力セットを生成した1つ以上の重なり合うシーケンシングサイクル、及び少なくとも1つの重なり合わないシーケンシングサイクルを共有する前記シーケンシングランのシーケンシングサイクルの第3のウィンドウに対して、

30

前記空間ネットワークを介して、前記少なくとも1つの重なり合わないシーケンシングサイクルに対してのみサイクルごとの分析物チャネルセットを処理し、前記少なくとも1つの重なり合わないシーケンシングサイクルのための空間出力セットのシーケンスを生成することと、

前記圧縮ネットワークを介して、前記空間出力セットのシーケンス内の最終空間出力セットを処理し、前記少なくとも1つの重なり合わないシーケンシングサイクルのための圧縮空間出力セットを生成することであって、前記最終空間出力セットが、M個のチャネルを有し、前記圧縮空間出力が、N個のチャネルを有し、MがNより大きい、ことと、

前記シーケンシングサイクルの第1及び第2のウィンドウに対して以前に生成された前記1つ以上の重なり合うシーケンシングサイクルのためのそれぞれの圧縮空間出力セット、及び前記圧縮空間出力セットに基づいて、前記シーケンシングサイクルの第3のウィンドウにおける1つ以上のシーケンシングサイクルに対するベースコール予測を生成することと、を更に含む、請求項3に記載の人工知能ベースの方法。

40

【請求項5】

前記空間ネットワークが、前記シーケンシングランのシーケンシングサイクルの特定のウィンドウのための前記シリーズにおけるサイクルごとの分析物チャネルセットの特定のウィンドウにおける各サイクルごとの分析物チャネルセットを、別々に処理する空間畳み込み層のシーケンスを有し、前記シーケンシングサイクルの特定のウィンドウ内の各シーケンシングサイクルに対して、空間出力セットのシーケンスを生成し、前記シーケンシングサイクルの特定のウィンドウ内の異なるシーケンシングサイクルの、サイクルごとの分

50

析物チャンネルセット間ではなく、対象のシーケンシングサイクルのサイクルごとの分析物チャンネルセット内でのみ、強度を組み合わせる第1の空間畳み込み層から始めることと、前記シーケンシングサイクルの特定のウィンドウ内の前記異なるシーケンシングサイクル間ではなく、対象のシーケンシングサイクル内でのみ、先行する空間畳み込み層の空間出力を組み合わせる後続の空間畳み込み層により引き続き行うことと、を含む、請求項1から4のいずれか一項に記載の人工知能ベースの方法。

【請求項6】

前記空間畳み込み層のシーケンス内のそれぞれの空間畳み込み層が、異なる数の畳み込みフィルタを有し、前記空間畳み込み層のシーケンス内の最終空間畳み込み層が、M個の畳み込みフィルタを有し、Mが4より大きい整数である、請求項5に記載の人工知能ベースの方法。

10

【請求項7】

前記空間畳み込み層のシーケンス内のそれぞれの空間畳み込み層が、同じ数の畳み込みフィルタを有し、前記同じ数がMであり、Mが4より大きい整数である、請求項5に記載の人工知能ベースの方法。

【請求項8】

前記空間ネットワーク内の前記畳み込みフィルタが、2次元(2D)の畳み込みを使用する、請求項1から7のいずれか一項に記載の人工知能ベースの方法。

【請求項9】

前記空間ネットワーク内の前記畳み込みフィルタが、3次元(3D)の畳み込みを使用する、請求項1から7のいずれか一項に記載の人工知能ベースの方法。

20

【請求項10】

前記畳み込みニューラルネットワークベースのベースコーラーが、時間ネットワークを有し、前記時間ネットワークが、前記シーケンシングサイクルの特定のウィンドウ内の後続のシーケンシングサイクルのウィンドウに対するそれぞれの圧縮空間出力セットをグループ化して処理する時間畳み込み層のシーケンスを有し、前記シーケンシングサイクルの特定のウィンドウに対する時間出力セットのシーケンスを生成し、前記シーケンシングサイクルの特定のウィンドウ内の前記異なるシーケンシングサイクル間で圧縮空間出力セットを組み合わせる第1の時間畳み込み層から開始することと、先行する時間畳み込み層の後続する時間出力を組み合わせる後続の時間畳み込み層により引き続き行うことと、を含む、請求項5から9のいずれか一項に記載の人工知能ベースの方法。

30

【請求項11】

前記シーケンシングサイクルの第1のウィンドウに対して、

前記時間ネットワークの前記時間畳み込み層のシーケンスにおける第1の時間畳み込み層を介して、前記シーケンシングサイクルの第1のウィンドウ内の後続のシーケンシングサイクルのウィンドウに対する前記それぞれの圧縮空間出力セットを処理し、前記シーケンシングサイクルの第1のウィンドウに対する複数の時間出力セットを生成することと、

前記圧縮ネットワークを介して、前記複数の時間出力セットを処理し、前記複数の時間出力セット内のそれぞれの時間出力セットのためのそれぞれの圧縮時間出力セットを生成することとあって、前記それぞれの時間出力セットが、M個のチャンネルを有し、前記それぞれの圧縮時間出力セットが、N個のチャンネルを有し、MがNより大きい、ことと、

40

前記時間ネットワークの前記時間畳み込み層のシーケンスにおける最終時間畳み込み層を介して、前記それぞれの圧縮時間出力セットを処理し、前記シーケンシングサイクルの第1のウィンドウに対する最終時間出力セットを生成することと、

前記最終時間出力セットに基づいて、前記シーケンシングサイクルの第1のウィンドウにおける1つ以上のシーケンシングサイクルに対する前記ベースコール予測を生成することと、を更に含み、

出力層が、前記最終時間出力セットを処理し、前記シーケンシングサイクルの第1のウィンドウに対する最終出力を生成し、前記ベースコール予測が、前記最終出力に基づいて生成される、請求項10に記載の人工知能ベースの方法。

50

【請求項 1 2】

前記シーケンシングサイクルの第 1 のウィンドウと、前記第 1 の時間畳み込み層が以前に時間出力セットを生成した後続のシーケンシングサイクルの 1 つ以上の重なり合うウィンドウ、及び前記第 1 の時間畳み込み層がまだ時間出力セットを生成していない後続のシーケンシングサイクルの少なくとも 1 つの重なり合わないウィンドウを共有するシーケンシングサイクルの第 2 のウィンドウに対して、

前記第 1 の時間畳み込み層を介して、前記後続のシーケンシングサイクルの少なくとも 1 つの重なり合わないウィンドウ内のそれぞれのシーケンシングサイクルに対してのみ、それぞれの圧縮空間出力セットを処理し、前記後続のシーケンシングサイクルの少なくとも 1 つの重なり合わないウィンドウに対する時間出力セットを生成することと、

10

前記圧縮ネットワークを介して、前記時間出力セットを処理し、前記後続のシーケンシングサイクルの少なくとも 1 つの重なり合わないウィンドウに対する圧縮時間出力セットを生成することであって、前記時間出力セットが、M 個のチャンネルを有し、前記圧縮時間出力セットが、N 個のチャンネルを有し、M が N より大きい、ことと、

前記最終時間畳み込み層を介して、前記シーケンシングサイクルの第 1 のウィンドウ及び前記圧縮時間出力セットに対して以前に生成された前記後続のシーケンシングサイクルの重なり合うウィンドウに対するそれぞれの圧縮時間出力セットを処理し、前記シーケンシングサイクルの第 2 のウィンドウに対する最終時間出力セットを生成することと、

前記最終時間出力セットに基づいて、前記シーケンシングサイクルの第 2 のウィンドウにおける 1 つ以上のシーケンシングサイクルに対する前記ベースコール予測を生成することと、を更に含み、

20

出力層が、前記最終時間出力セットを処理し、前記シーケンシングサイクルの第 2 のウィンドウに対する最終出力を生成し、前記ベースコール予測が、前記最終出力に基づいて生成される、請求項 1 1 に記載の人工知能ベースの方法。

【請求項 1 3】

前記シーケンシングサイクルの第 1 及び第 2 のウィンドウと、前記第 1 の時間畳み込み層が以前に時間出力セットを生成した後続のシーケンシングサイクルの 1 つ以上の重なり合うウィンドウ、及び後続のシーケンシングサイクルの少なくとも 1 つの重なり合わないウィンドウを共有するシーケンシングサイクルの第 3 のウィンドウに対して、

前記第 1 の時間畳み込み層を介して、前記後続のシーケンシングサイクルの少なくとも 1 つの重なり合わないウィンドウ内のそれぞれのシーケンシングサイクルに対してのみ、それぞれの圧縮空間出力セットを処理し、前記後続のシーケンシングサイクルの少なくとも 1 つの重なり合わないウィンドウに対する時間出力セットを生成することと、

30

前記圧縮ネットワークを介して、前記時間出力セットを処理し、前記後続のシーケンシングサイクルの少なくとも 1 つの重なり合わないウィンドウに対する圧縮時間出力セットを生成することであって、前記時間出力セットが、M 個のチャンネルを有し、前記圧縮時間出力セットが、N 個のチャンネルを有し、M が N より大きい、ことと、

前記最終時間畳み込み層を介して、前記シーケンシングサイクルの第 1 及び第 2 のウィンドウ並びに前記圧縮時間出力セットに対して以前に生成された前記後続のシーケンシングサイクルの重なり合うウィンドウに対するそれぞれの圧縮時間出力セットを処理し、前記シーケンシングサイクルの第 3 のウィンドウの最終時間出力セットを生成することと、

40

前記最終時間出力セットに基づいて、前記シーケンシングサイクルの第 3 のウィンドウにおける 1 つ以上のシーケンシングサイクルに対する前記ベースコール予測を生成することと、を更に含み、

出力層が、前記最終時間出力セットを処理し、前記シーケンシングサイクルの第 3 のウィンドウに対する最終出力を生成し、前記ベースコール予測が、前記最終出力に基づいて生成される、請求項 1 2 に記載の人工知能ベースの方法。

【請求項 1 4】

前記時間ネットワークの前記時間畳み込み層のシーケンス内のそれぞれの時間畳み込み層が、異なる数の畳み込みフィルタを有し、前記第 1 の時間畳み込み層が、M 個の畳み込

50

みフィルタを有し、Mが4より大きい整数である、請求項10から13のいずれか一項に記載の人工知能ベースの方法。

【請求項15】

前記時間ネットワークの前記時間畳み込み層のシーケンス内のそれぞれの時間畳み込み層が、同じ数の畳み込みフィルタを有し、前記同じ数がMであり、Mが4より大きい整数である、請求項10から13のいずれか一項に記載の人工知能ベースの方法。

【請求項16】

前記時間ネットワーク内の畳み込みフィルタが、1次元(1D)の畳み込みを使用する、請求項14または15に記載の人工知能ベースの方法。

【請求項17】

前記圧縮ネットワークが、 1×1 の畳み込みを使用して、圧縮空間出力セット内の圧縮空間出力の数を制御し、前記圧縮ネットワークが、N個の畳み込みフィルタを有し、Nは4以下の整数である、請求項1から16のいずれか一項に記載の人工知能ベースの方法。

【請求項18】

信頼できない分析物に対応する圧縮空間出力セット内の圧縮空間出力の部分を除去するために前記信頼できない分析物を識別するデータを使用することを更に含み、フィルタリングされた圧縮空間出力セットを生成して、前記圧縮空間出力セットを置換し、前記信頼できない分析物ではないそれらの分析物に対してのみベースコール予測を生成する、請求項1から17のいずれか一項に記載の人工知能ベースの方法。

【請求項19】

前記1つ以上のシーケンシングサイクルに対する前記ベースコール予測に基づいて前記画像のサイクルごとの分析物チャネルセットの前記第1のウィンドウの前記画像に示される前記クラスタのターゲットクラスタに対するベースコール生成することをさらに含む、請求項1に記載の人工知能ベースの方法。

【請求項20】

前記画像のサイクルごとの分析物チャネルセットのシリーズの画像が、シーケンシング機器の撮像装置によって捕捉される、請求項1に記載の人工知能ベースの方法。

【発明の詳細な説明】

【技術分野】

【0001】

開示される技術は、人工知能タイプコンピュータ及びデジタルデータ処理システム、並びに知能のエミュレーションのための対応するデータ処理方法及び製品（すなわち、知識ベースシステム、推論システム、及び知識獲得システム）に関し、開示される技術は、不確実性下で推論するためのシステム（例えば、ファジィ論理システム）、適応型システム、機械学習システム、及び人工ニューラルネットワークを含む。具体的には、開示される技術は、データを分析するための深層畳み込みニューラルネットワークなどの深層ニューラルネットワークを使用することに関する。

【0002】

（優先権出願）

本PCT出願は、2020年2月20日に出願された「DATA COMPRESSION FOR ARTIFICIAL INTELLIGENCE-BASED BASE CALLING」と題する米国仮特許出願第62/979,411号（代理人整理番号ILLM 1029-1/IP-1964-PRV）、及び2021年2月18日に出願された「DATA COMPRESSION FOR ARTIFICIAL INTELLIGENCE-BASED BASE CALLING」と題する米国特許出願第17/179,395号（代理人整理番号ILLM 1029-2/IP-1964-US）の優先権及び利益を主張する。優先権出願は、本明細書に完全に記載されているかのように、全ての目的のために参照により本明細書に組み込まれる。

【0003】

本PCT出願は、2020年2月20日に出願された「SQUEEZING LAYER

10

20

30

40

50

FOR ARTIFICIAL INTELLIGENCE-BASED BASE CALLING」と題する米国仮特許出願第62/979,399号（代理人整理番号、ILLM1030-1/IP-1982-PRV）、及び2021年2月19日に出願された「SPLIT ARCHITECTURE FOR ARTIFICIAL INTELLIGENCE-BASED BASE CALLER」と題する米国特許出願第17/180,480号（代理人整理番号ILLM1030-2/IP-1982-US）の優先権及び利益を主張する。優先権出願は、本明細書に完全に記載されているかのように、全ての目的のために参照により本明細書に組み込まれる。

【0004】

本PCT出願は、2021年2月19日に出願された「BUS NETWORK FOR ARTIFICIAL INTELLIGENCE-BASED BASE CALLER」と題する米国特許出願第17/180,513号（代理人整理番号ILLM1031-2/IP-1965-US）の優先権及び利益を主張する。優先権出願は、本明細書に完全に記載されているかのように、全ての目的のために参照により本明細書に組み込まれる。

【0005】

（組み込み）

以下の文献は、本明細書に完全に記載されているかのように、参照により組み込まれる：

2020年2月20日に出願された「ARTIFICIAL INTELLIGENCE-BASED BASE CALLING OF INDEX SEQUENCES」と題する米国仮特許出願第62/979,384号（代理人整理番号ILLM 1015-1/IP-1857-PRV）、

2020年2月20日に出願された「ARTIFICIAL INTELLIGENCE-BASED MANY-TO-MANY BASE CALLING」と題する米国仮特許出願第62/979,414号（代理人整理番号ILLM 1016-1/IP-1858-PRV）、

2020年2月20日に出願された「KNOWLEDGE DISTILLATION-BASED COMPRESSION OF ARTIFICIAL INTELLIGENCE-BASED BASE CALLER」と題する米国仮特許出願第62/979,385号（代理人整理番号ILLM 1017-1/IP-1859-PRV）、

2020年8月28日に出願された「DETECTING AND FILTERING CLUSTERS BASED ON ARTIFICIAL INTELLIGENCE-PREDICTED BASE CALLS」と題する米国仮特許出願第63/072,032号（代理人整理番号ILLM 1018-1/IP-1860-PRV）、

2020年2月20日に出願された「MULTI-CYCLE CLUSTER BASED REAL TIME ANALYSIS SYSTEM」と題する米国仮特許出願第62/979,412号（代理人整理番号ILLM 1020-1/IP-1866-PRV）、

2020年3月20日に出願された「TRAINING DATA GENERATION FOR ARTIFICIAL INTELLIGENCE-BASED SEQUENCING」と題する米国非仮特許出願第16/825,987号（代理人整理番号ILLM 1008-16/IP-1693-US）、

2020年3月20日に出願された「ARTIFICIAL INTELLIGENCE-BASED GENERATION OF SEQUENCING METADATA」と題する米国非仮特許出願第16/825,991号（代理人整理番号ILLM 1008-17/IP-1741-US）、

2020年3月20日に出願された「ARTIFICIAL INTELLIGENCE-BASED BASE CALLING」と題する米国非仮特許出願第16/826,126号（代理人整理番号ILLM 1008-18/IP-1744-US）、

2020年3月20日に出願された「ARTIFICIAL INTELLIGENCE-BASED QUALITY SCORING」と題する米国非仮特許出願第16/8

10

20

30

40

50

26, 134号(代理人整理番号 ILLM 1008-19/IP-1747-US)、及び

2020年3月21日に出願された「ARTIFICIAL INTELLIGENCE-BASED SEQUENCING」と題する米国非仮特許出願第16/826, 168号(代理人整理番号 ILLM 1008-20/IP-1752-PRV-US)。

【背景技術】

【0006】

このセクションで考察される主題は、単にこのセクションにおける言及の結果として、先行技術であると想定されるべきではない。同様に、このセクションで言及した問題、又は背景として提供された主題と関連付けられた問題は、先行技術において以前に認識されていると想定されるべきではない。このセクションの主題は、単に、異なるアプローチを表し、それ自体はまた、特許請求される技術の実施態様に対応し得る。

【0007】

計算能力の急速な改善により、近年、多くのコンピュータビジョンタスクにおいて、深層畳み込みニューラルネットワーク(convolution neural network、CNN)が、著しく改善された精度で、大成功を収めることが可能となった。推論段階の間、多くのアプリケーションは、厳密な電力消費要件を伴う、1つの画像の低遅延の処理を必要とし、これにより、グラフィックス処理ユニット(graphics processing unit、GPU)及び他の汎用プラットフォームの効率が低下し、そのことは、特定のアクセレーションハードウェア、例えば、フィールドプログラマブルゲートアレイ(Field Programmable Gate Array、FPGA)にとっては、深層学習アルゴリズムの推論に特に効果的となるようにデジタル回路をカスタマイズすることによって、好機をもたらすこととなる。しかしながら、携帯型及び埋め込み式システムにCNNを配備することは、大きなデータ量、集中的な計算、様々なアルゴリズム構造、及び頻繁なメモリアクセスのために依然として困難である。

【0008】

畳み込みが、CNNにおけるほとんどの演算を提供するので、畳み込みアクセレーションスキームが、ハードウェアCNNアクセラレータの効率及び性能に大きく影響することになる。畳み込みは、カーネル及び特徴マップに沿ってスライドする4つのレベルのループを伴う、積和(multiply and accumulate、MAC)演算を含む。第1のループレベルは、1つのカーネルウィンドウ内のピクセルのMACを計算する。第2のループレベルは、様々な異なる入力特徴マップにわたるMACの積の和を累積する。第1及び第2のループレベルを完了した後、バイアスを追加することにより、最終出力ピクセルが得られる。第3のループレベルは、入力特徴マップ内で、カーネルウィンドウをスライドさせる。第4のループレベルは、様々な異なる出力特徴マップを生成する。

【0009】

FPGAは、特に、推論タスクを加速化するために、より多くの関心を集め、より広く普及してきた。それは、FPGAが、(1)再構成可能性が高く、(2)CNNの急速な進化にキャッチアップするために必要な開発時間の速さという点で、特定用途向け集積回路(application specific integrated circuit、ASIC)と比較して優れており、(3)良好な性能を有し、(4)GPUと比較して、エネルギー効率が優れている、ということに起因する。FPGAの高い性能及び高い効率性は、特定の計算のためにカスタマイズされた回路を合成して、カスタマイズされたメモリシステムで数十億回の演算を直接処理することによって実現することができる。例えば、現代のFPGAにおける数百から数千のデジタル信号処理(digital signal processing、DSP)ブロックは、コア畳み込み演算、例えば、高度の並列処理を伴う積和演算をサポートする。外部オンチップメモリとオンチッププロセッシングエンジン(processing engine、PE)と間の専用データバッファは、フィールドプログラマブルゲートアレイ(FPGA)チップ上に、数十メガバイトのオンチップブロックランダムアクセスメモリ(block random access memory、BRAM)を構成することにより、優先データフローを実現するように設計することがで

10

20

30

40

50

きる。

【0010】

高性能を達成するためにリソース利用を最大化しながら、データ通信を最小限に抑えるために、CNNアクセレレーションの効率的なデータフロー及びハードウェアアーキテクチャが望まれている。アクセレレーションハードウェア上の様々なCNNアルゴリズムの推論プロセスを加速化し、高い性能、高い効率、及び高い柔軟性を実現するための方法論及びフレームワークを設計するための好機が生じることとなる。

【発明の概要】

【課題を解決するための手段】

【0011】

ベースコールの人工知能ベースの方法であって、前記方法が、
シーケンシングランのシーケンシングサイクルに対して生成されたサイクルごとの分析物チャンネルセットのシリーズにアクセスすることと、

ニューラルネットワークベースのベースコーラーの空間ネットワークを介して、前記シーケンシングランのシーケンシングサイクルの第1のウィンドウに対する、前記シリーズ内のサイクルごとの分析物チャンネルセットの第1のウィンドウを処理し、前記シーケンシングサイクルの第1のウィンドウにおけるそれぞれのシーケンシングサイクルに対する空間出力セットのそれぞれのシーケンスを生成することと、

前記ニューラルネットワークベースのベースコーラーの圧縮ネットワークを介して、前記空間出力セットのそれぞれのシーケンス内のそれぞれの最終空間出力セットを処理し、前記シーケンシングサイクルの第1のウィンドウにおける前記それぞれのシーケンシングサイクルに対するそれぞれの圧縮空間出力セットを生成することと、

前記それぞれの圧縮空間出力セットに基づいて、前記シーケンシングサイクルの第1のウィンドウにおける1つ以上のシーケンシングサイクルに対するベースコール予測を生成することと、を含む。

【0012】

図面では、同様の参照文字は、概して、異なる図全体を通して同様の部分を指す。また、図面は必ずしも縮尺どおりではなく、その代わりに、開示された技術の原理を例示することを強調している。以下の説明では、開示される技術の様々な実施態様が、以下の図面を参照して説明される。

【図面の簡単な説明】

【0013】

【図1A】ベースコールの第1の反復のための圧縮空間マップセットを生成する、開示された圧縮ロジックの一実施態様を示す。

【図1B】開示された時間ロジックによって圧縮空間マップセットを処理して、時間マップセットを生成する一実施態様を示す。

【図1C】開示された出力ロジックによって時間マップセットを処理して、ベースコール分類データを生成する一実施態様を示す。

【図1D】対象シーケンシングサイクルに対するサイクルごとの画像パッチの処理に応答して、空間畳み込み層のカスケードによって連続的に生成された特徴マップボリュームのシーケンスの一例を示す。

【図1E】 1×1 の畳み込みが特徴マップをどのように圧縮するかを示す一例を示す。

【図1F】開示された圧縮ロジックによって達成される圧縮比が、開示された圧縮ロジックによって適用される圧縮フィルタの数の関数であることを示す。

【図1G】例示的なソフトマックス関数を示す。

【図1H】開示された技術によって生成されたクラスごと、サイクルごとの確率四分位数の例を示す。

【図2A】ベースコールの2回目の反復中に、空間マップと対応する圧縮空間マップが、重なり合わないシーケンシングサイクル6に対してのみ生成されることを示す。

【図2B】ベースコールの第1の反復中に生成された圧縮空間マップセットが、ベースコ

10

20

30

40

50

ールの第2の反復中に生成された圧縮空間マップセットと併せて使用されて、中央シーケンシングサイクル4に対するベースコールを生成することを示す。

【図2C】出力層が、ベースコールの第2の反復中に生成された最終時間マップセットを処理し、中央シーケンシングサイクル4のベースコールを生成することを示す。

【図3A】ベースコールの3回目の反復中に、空間マップと対応する圧縮空間マップが、重なり合わないシーケンシングサイクル7に対してのみ生成されることを示す。

【図3B】ベースコールの第1及び第2の反復中に生成された圧縮空間マップセットが、ベースコールの第3の反復中に生成された圧縮空間マップセットと併せて使用されて、中央シーケンシングサイクル5のベースコールを生成することを示す。

【図3C】出力層が、ベースコールの第3の反復中に生成された最終時間マップセットを処理し、中央シーケンシングサイクル5のベースコールを生成することを示す。

10

【図4A】中央シーケンシングサイクル16をベースコールするためのベースコールの第14の反復を示す。

【図4B】シーケンシングサイクル1～29に対して以前に生成された圧縮空間マップを使用して、中央シーケンシングサイクル16をベースコールするための最終時間マップセットを生成することを示す。

【図4C】出力層が、ベースコールの第14の反復中に生成された最終時間マップセットを処理し、中央シーケンシングサイクル16のベースコールを生成することを示す。

【図5A】フィルタリングロジックを使用して、それぞれのシーケンシングサイクル1、2、3、4、及び5に対する圧縮空間マップセットをフィルタリングして、ベースコールの第1の反復中に、それぞれのフィルタリングされた圧縮空間マップを生成する一実施態様を示す。

20

【図5B】出力層が、ベースコールの第1の反復中に生成された、フィルタリングされた最終時間マップセットを処理し、中央シーケンシングサイクル3のベースコールを生成することを示す。

【図6A】フィルタリングロジックを使用して、それぞれのシーケンシングサイクル2、3、4、5、及び6の圧縮空間マップセットをフィルタリングして、ベースコールの第2の反復中に、それぞれのフィルタリングされた圧縮空間マップを生成する一実施態様を示す。

【図6B】出力層が、ベースコールの第2の反復中に生成された、フィルタリングされた最終時間マップセットを処理し、中央シーケンシングサイクル4のベースコールを生成することを示す。

30

【図7A】フィルタリングロジックを使用して、それぞれのシーケンシングサイクル3、4、5、6、及び7の圧縮空間マップセットをフィルタリングして、ベースコールの第3の反復中に、それぞれのフィルタリングされた圧縮空間マップを生成する一実施態様を示す。

【図7B】出力層が、ベースコールの第3の反復中に生成された、フィルタリングされた最終時間マップセットを処理し、中央シーケンシングサイクル5のベースコールを生成することを示す。

【図8A】圧縮ロジックによって、ベースコールの第1の反復中に生成された時間特徴マップのセットを処理して、圧縮時間特徴マップのそれぞれのセットを生成する一実施態様を示す。

40

【図8B】出力層が、ベースコールの第1の反復中に生成された、最終圧縮時間マップセットを処理し、中央シーケンシングサイクル3のベースコールを生成することを示す。

【図9A】第1のベースコール反復中に生成された圧縮時間マップセットが、ベースコールの第2の反復中に生成された圧縮時間マップセットと併せて使用されて、中央シーケンシングサイクル4のベースコールを生成することを示す。

【図9B】出力層が、ベースコールの第2の反復中に生成された、最終圧縮時間マップセットを処理し、中央シーケンシングサイクル4のベースコールを生成することを示す。

【図10A】第1及び第2のベースコール反復中に生成された圧縮時間マップセットが、

50

ベースコールの第3の反復中に生成された圧縮時間マップセットと併せて使用されて、中央シーケンシングサイクル5のベースコールを生成することを示す。

【図10B】出力層が、ベースコールの第3の反復中に生成された、最終圧縮時間マップセットを処理し、中央シーケンシングサイクル5のベースコールを生成することを示す。

【図11A】圧縮ロジックによって、ベースコールの第1の反復中に生成された、フィルタリングされた時間特徴マップのセットを処理して、フィルタリングされた圧縮時間特徴マップのそれぞれのセットを生成する一実施態様を示す。

【図11B】出力層が、ベースコールの第1の反復中に生成された、フィルタリングされた最終圧縮時間マップセットを処理し、中央シーケンシングサイクル3のベースコールを生成することを示す。

10

【図12A】第1のベースコールの反復中に生成された、フィルタリングされた圧縮時間マップセットが、ベースコールの第2の反復中に生成されたフィルタリングされた圧縮時間マップセットと併せて使用されて、中央シーケンシングサイクル4のベースコールを生成することを示す。

【図12B】出力層が、ベースコールの第2の反復中に生成された、フィルタリングされた最終圧縮時間マップセットを処理し、中央シーケンシングサイクル4のベースコールを生成することを示す。

【図13A】第1及び第2のベースコールの反復中に生成された、フィルタリングされた圧縮時間マップセットが、ベースコールの第3の反復中に生成された、フィルタリングされた圧縮時間マップセットと併せて使用されて、中央シーケンシングサイクル5のベースコールを生成することを示す。

20

【図13B】出力層が、ベースコールの第3の反復中に生成された、フィルタリングされた最終圧縮時間マップセットを処理し、中央シーケンシングサイクル5のベースコールを生成することを示す。

【図14】本明細書に開示されるニューラルネットワークベースのベースコーラーの、第1の例示的なアーキテクチャを示す。

【図15】本明細書に開示されるニューラルネットワークベースのベースコーラーの、第2の例示的なアーキテクチャを示す。

【図16】本明細書に開示されるニューラルネットワークベースのベースコーラーの、第3の例示的なアーキテクチャを示す。

30

【図17】本明細書に開示されるニューラルネットワークベースのベースコーラーの、第4の例示的なアーキテクチャを示す。

【図18】入力データ内のチャンネルの数に応じて、圧縮層内の畳み込みフィルタの数（又は大きな数）を構成する、フィルタ構成ロジックの一実施態様を示す。

【図19A】シーケンシングシステムの一実施態様を示す。

【図19B】シーケンシングシステムの一実施態様を示す。シーケンシングシステムは、構成可能なプロセッサを含む。

【図19C】ベースコールセンサ出力などの、シーケンシングシステムからのセンサデータの、分析のためのシステムの簡略ブロック図である。

【図20A】ホストプロセッサによって実行される実行時プログラムの機能を含む、ベースコール動作の態様を示す簡略図である。

40

【図20B】構成可能なプロセッサの構成の簡略図である。

【図21】オフチップメモリ（例えば、オフチップのDRAM、ホストRAM、ホスト高帯域幅メモリ（high bandwidth memory、HBM））から、第2のベースコール反復の間に利用可能な、第1のベースコール反復中に生成された圧縮空間マップを作成する、開示されたデータフローロジックの別の実施態様を示している。

【図22】オンチップメモリ（例えば、オンチップのDRAM、オンチップのSRAM、オンチップのBRAM、相互接続を介してプロセッサに取り付けられたDRAM）から、第2のベースコール反復中に利用可能な、第1のベースコール反復中に生成された圧縮空間マップを作成する、開示されたデータフローロジックの一実施態様を示す。

50

【図 2 3】開示されたニューラルネットワークベースのベースコーラーの、いわゆる分割アーキテクチャの一実施態様を示す。

【図 2 4 A】特徴マップ付加を介して、以前の情報を下流に再度注入する、残留接続を示す。

【図 2 4 B】残留ブロック及びスキップ接続の一実施態様を示す。

【図 2 4 C】空間畳み込み層がスキップ接続で残留ブロックにグループ化される、ニューラルネットワークベースのベースコーラーの残留アーキテクチャを示す。

【図 2 5 A】本明細書に記載のニューラルネットワークベースのベースコーラーの開示されたバスネットワークの詳細を示す。

【図 2 5 B】開示されたバスネットワークの例示的な動作を示す。

【図 2 5 C】開示されたバスネットワークの次元適合性ロジックの一実施態様を示す。

【図 2 6】開示されたバスネットワークの別の例を示す。

【図 2 7】開示されたバスネットワークの更に別の例を示す。

【図 2 8】開示されたバスネットワークのスケーリングロジックの一実施態様を示す。

【図 2 9】時間ネットワークの時間畳み込み層間のスキップ接続の一実施態様を示す。

【図 3 0】圧縮ロジックで構成されたネットワークネットワークベースのベースコーラー（s q z 2 ベースコーラー）によるベースコール性能を、圧縮ロジックを有しないネットワークネットワークベースのベースコーラーと、I l l u m i n a の非ニューラルネットワークベースのベースコーラーリアルタイム分析（Real-Time Analysis、R T A）ソフトウェアとに対して、比較するグラフである。

【図 3 1】開示された圧縮ロジックの使用によってもたらされる R A M 及び D R A M の使用の節約を示す。

【図 3 2】分割及びスキップアーキテクチャ（s p l i t _ r e s）で構成されたネットワークネットワークベースのベースコーラーによるベースコール性能を、R T A ベースコーラーに対して、かつ分割及びスキップアーキテクチャを伴わない（ディスティルド）ネットワークネットワークベースのベースコーラーの別のバージョンに対して比較するグラフである。

【図 3 3】開示される技術を実装するように使用され得るコンピュータシステムである。

【発明を実施するための形態】

【0 0 1 4】

以下の考察は、開示された技術を当業者が作製及び使用することを可能にするために提示され、特定の用途及びその要件に関連して提供される。開示される実施態様に対する様々な修正は、当業者には容易に明らかとなり、本明細書で定義される一般原理は、開示される技術の趣旨及び範囲から逸脱することなく、他の実施態様及び用途に適用され得る。したがって、開示される技術は、示される実施態様に限定されることを意図するものではなく、本明細書に開示される原理及び特徴と一致する最も広い範囲を与えられるものである。

【0 0 1 5】

シーケンス画像

ベースコールは、シーケンスのヌクレオチド組成を決定するプロセスである。ベースコールは、I l l u m i n a の i S e q、H i S e q X、H i S e q 3 0 0 0、H i S e q 4 0 0 0、H i S e q 2 5 0 0、N o v a S e q 6 0 0 0、N e x t S e q 5 5 0、N e x t S e q 1 0 0 0、N e x t S e q 2 0 0 0、N e x t S e q D x、M i S e q、及び M i S e q D x などのシーケンシング機器によって行われる、シーケンシングラン（又はシーケンシング反応）中に生成された画像データ、すなわちシーケンス画像の分析を伴う。

【0 0 1 6】

以下の説明は、一実施態様に従って、シーケンス画像がどのように生成されるか、及びそれらの画像が何を描いているのかを概説する。

【0 0 1 7】

ベースコールは、シーケンス画像にコード化された強度データを、ヌクレオチドシーケ

10

20

30

40

50

ンスにデコードする。一実施態様では、I l l u m i n a のシーケンシングプラットフォームは、ベースコールのための循環可逆的ターミネーション (Cyclic Reversible Termination、C R T) 化学を採用する。このプロセスは、新たに添加された各ヌクレオチドの放出信号を追跡しながら、蛍光標識されたヌクレオチドを有するテンプレート鎖に相補的な、成長しつつある新生ストランドに依存する。蛍光標識されたヌクレオチドは、ヌクレオチド型のフルオロフォアシグナルをアンカーする、3' の取り外し可能なブロックを有する。

【0018】

シーケンシングは、反復サイクルで行われ、反復サイクルの各々は、以下の3つのステップを含む：(a) 蛍光標識ヌクレオチドを添加することによる新生鎖の伸長、(b) シーケンシング機器の光学システムの1つ以上のレーザを使用したフルオロフォアの励起と、光学システムの異なるフィルタを通した画像化による、シーケンス画像の生成、並びに (c) フルオロフォアの切断及び次のシーケンシングサイクルの準備における、3' ブロックの除去。組み込み及び撮像サイクルを、最大で指定された数のシーケンシングサイクルで繰り返し、読み取り長さを規定する。このアプローチを使用して、各サイクルはテンプレートストランドに沿って新しい位置を問い合わせる。

【0019】

I l l u m i n a シーケンサの途方もないパワーは、C R T 反応を受ける数百万のクラスタ又は更には数十億のクラスタ (分析物) を同時に実行及び感知する能力に起因する。クラスタどうしはサイズ及び形状が異なるが、1つのクラスタには、テンプレートストランドの約千個の同一のコピーが含まれる。クラスタは、シーケンシングラン前に、入力ライブラリのブリッジ増幅又は排除増幅によって、テンプレートストランドから成長させる。増幅及びクラスタ成長の目的は、撮像装置が単一のストランドのフルオロフォアシグナルを確実に感知するということができないため、放出されるシグナルの強度を増大させることにある。しかしながら、クラスタ内のストランドどうしの物理的距離は小さいため、撮像装置は、単一のスポットとしてストランドのクラスタを知覚する。

【0020】

シーケンシングは、フローセル (又はバイオセンサ)、すなわち、入力ストランドを保持する小さなガラススライドにおいて行われる。フローセルは、光学システムに接続され、その光学システムは、顕微鏡撮像装置、励起レーザ装置、及び蛍光フィルタを含む。フローセルは、レーンと呼ばれる複数のチャンバを含む。レーンどうしは、互いに物理的に分離され、異なるタグ付けされたシーケンシングライブラリを含んでもよく、試料の交差汚染なしで区別可能である。いくつかの実施態様では、フローセルは、パターン化された表面を含む。「パターン化された表面」は、固体支持体の露出層内又はその上における、異なる領域の配置を指す。

【0021】

撮像装置 (例えば、電荷結合素子 (Charge-Coupled Device、C C D) 又は相補的金属酸化物半導体 (Complementary Metal-Oxide-Semiconductor、C M O S) センサなどのソリッドステート撮像素子) は、レーンに沿った複数の場所で、タイルと呼ばれる一連の重なり合わない領域において、スナップショットを取る。例えば、レーン当たり64個又は96個のタイルが存在し得る。タイルは数十万〜数百万個のクラスタを保持する。

【0022】

シーケンシングランの出力は、シーケンス画像である。シーケンス画像は、ピクセル化ユニット (例えば、ピクセル、スーパーピクセル、サブピクセル) のグリッド (又はアレイ) を使用するクラスタ及びそれらの周囲の背景の強度放射を示す。強度放射は、ピクセル化ユニットの強度値として記憶される。シーケンス画像は、ピクセル化ユニットのグリッドの、w 個 × h 個の寸法を有し、ここで、w (幅) 及び h (高さ) は、1 ~ 100, 000 の範囲の任意の数 (例えば、115 × 115、200 × 200、1800 × 2000、2200 × 25000、2800 × 3600、4000 × 400) である。いくつかの実施態様では、w 及び h は同じである。他の実施態様では、w 及び h は異なる。シーケン

10

20

30

40

50

ス画像は、シーケンシングラン中にヌクレオチドシーケンスにヌクレオチドを組み込む結果として生成される強度放射を示す。強度放射は、関連するクラスタ及びそれらの周囲の背景に由来するものである。

【0023】

ニューラルネットワークベースのベースコール

以下の考察は、本明細書に記載のニューラルネットワークベースのベースコーラー100に焦点を当てている。まず、一実施態様による、ニューラルネットワークベースのベースコーラー100への入力の説明される。次いで、ニューラルネットワークベースのベースコーラー100の構造及び形態の例が提供される。最後に、一実施態様による、ニューラルネットワークベースのベースコーラー100の出力の説明される。

【0024】

データフローロジックは、ベースコールのために、ニューラルネットワークベースのベースコーラー100にシーケンス画像を提供する。ニューラルネットワークベースのベースコーラー100は、パッチごとに（又はタイルごとに）シーケンス画像にアクセスする。パッチのそれぞれは、シーケンス画像を形成するピクセル化されたユニットのグリッド内のピクセル化ユニットのサブグリッド（又はサブアレイ）である。パッチは、ピクセル化ユニットのサブグリッドの q 個 $\times$ r 個の寸法を有し、 q （幅）及び r （高さ）は、1～10000の範囲の任意の数である（例えば、 3×3 、 5×5 、 7×7 、 10×10 、 15×15 、 25×25 、 64×64 、 78×78 、 115×115 ）。いくつかの実施態様では、 q 及び r は同じである。他の実施態様では、 q 及び r は互いに異なる。いくつかの実施態様では、1つのシーケンス画像から抽出されたパッチは、同じサイズである。他の実施態様では、パッチは、異なるサイズのものである。いくつかの実施態様では、パッチどうしは、（例えば、エッジの上で）重なり合うピクセル化ユニットを有することができる。

【0025】

シーケンシングにより、対応する m 個の画像チャンネルに対して、シーケンシングサイクルごとに、 m 個のシーケンス画像を生成する。すなわち、シーケンス画像の各々は、1つ以上の画像（又は強度）チャンネル（カラー画像の赤、緑、青（RGB）のチャンネルと類似）を有する。一実施態様では、各画像チャンネルは、複数のフィルタ波長帯域のうちの1つに対応する。別の実施態様では、各画像チャンネルは、1つのシーケンシングサイクルで複数の撮像イベントのうちの1つに対応する。更に別の実施態様では、各画像チャンネルは、特定のレーザを用いた照明と特定の光学フィルタを通した撮像との組み合わせに対応する。画像パッチは、特定のシーケンシングサイクルについて、 m 個の画像チャンネルの各々からタイリング（又はアクセス）される。4-、2-、及び1-チャンネル化学などの異なる実施態様では、 m は、4又は2である。他の実施態様では、 m は、1、3、又は4よりも大きい。

【0026】

例えば、シーケンシングランが、2つの異なる画像化チャンネル、すなわち、青色チャンネルと緑色チャンネルとを使用して実施され则认为てみる。すると、各シーケンシングサイクルにおいて、シーケンシングランは、青色画像及び緑色画像を生成する。このようにして、シーケンシングランの k 個のシーケンシングサイクルのシリーズについて、 k 対の青色画像及び緑色画像のシーケンスが出力として生成され、シーケンス画像として記憶される。したがって、 k 対の青色及び緑色画像パッチのシーケンスが、ニューラルネットワークベースのベースコーラー100によるパッチレベル処理に対して生成される。

【0027】

ベースコールの反復1回（又はフォワードパス又は単一の順方向横断の1つのインスタンス）に対する、ニューラルネットワークベースのベースコーラー100への入力画像データは、複数のシーケンシングサイクルを含む1つのスライドウィンドウに対するデータを含む。スライドウィンドウは、例えば、現在のシーケンシングサイクル、1つ以上の先行するシーケンシングサイクル、及び1つ以上の後続のシーケンシングサイクルを含むこ

10

20

30

40

50

とができる。

【0028】

一実施態様では、入力画像データは、3回のシーケンシングサイクルのデータを含み、ベースコールされる現在の（時間 t ）シーケンシングサイクルのデータは、（ i ）左側の隣接／コンテキスト／以前の／先行する／前の（時間 $t-1$ ）シーケンシングサイクルのデータ、及び（ $i\ i$ ）右側の隣接／コンテキスト／次の／後続の／その後の（時間 $t+1$ ）シーケンシングサイクルのデータを伴う。

【0029】

別の実施態様では、入力画像データは、5回のシーケンシングサイクルのデータを含み、ベースコールされる現在の（時間 t ）シーケンシングサイクルのデータは、（ i ）第1の左側の隣接／コンテキスト／以前の／先行する／前の（時間 $t-1$ ）シーケンシングサイクルについてのデータと、（ $i\ i$ ）第2の左側の隣接／コンテキスト／以前の／先行する／前の（時間 $t-2$ ）シーケンシングサイクルについてのデータと、（ $i\ i\ i$ ）第1の右側の隣接／コンテキスト／次の／後続の／その後の（時間 $t+1$ ）シーケンシングサイクルについてのデータと、（ $i\ v$ ）第2の右側の隣接／コンテキスト／次の／後続の／その後の（時間 $t+2$ ）シーケンシングサイクルについてのデータと、を伴う。

【0030】

更に別の実施態様では、入力画像データは、7回のシーケンシングサイクルのデータを含み、ベースコールされる現在の（時間 t ）シーケンシングサイクルのデータは、（ i ）第1の左側の隣接／コンテキスト／以前の／先行する／前の（時間 $t-1$ ）シーケンシングサイクルについてのデータと、（ $i\ i$ ）第2の左側の隣接／コンテキスト／以前の／先行する／前の（時間 $t-2$ ）シーケンシングサイクルについてのデータと、（ $i\ i\ i$ ）第3の左側の隣接／コンテキスト／以前の／先行する／前の（時間 $t-3$ ）シーケンシングサイクルについてのデータと、（ $i\ v$ ）第1の右側の隣接／コンテキスト／次の／後続の／その後の（時間 $t+1$ ）シーケンシングサイクルについてのデータと、（ v ）第2の右側の隣接／コンテキスト／次の／後続の／その後の（時間 $t+2$ ）シーケンシングサイクルについてのデータと、（ $v\ i$ ）第3の右側の隣接／コンテキスト／次の／後続の／その後の（時間 $t+3$ ）シーケンシングサイクルについてのデータと、を伴う。他の実施態様では、入力画像データは、1回のシーケンシングサイクルのデータを含む。更に他の実施態様では、入力画像データは、10回、15回、20回、30回、58回、75回、92回、130回、168回、175回、209回、225回、230回、275回、318回、325回、330回、525回、又は625回のシーケンシングサイクルのデータを含む。

【0031】

1つの実施態様によれば、ニューラルネットワークベースのベースコーラー100は、その畳み込み層を介して画像パッチを処理し、代替表現を生成する。次いで、その代替表現は、出力層（例えば、ソフトマックス層）によって使用され、現在のシーケンシングサイクル（時間 t ）、又はシーケンシングサイクルの各々（すなわち、現在のシーケンシングサイクル（時間 t ）、第1及び第2の先行するシーケンシングサイクル（時間 $t-1$ 、時間 $t-2$ ）、及び第1及び第2の後続のシーケンシングサイクル（時間 $t+1$ 、時間 $t+2$ ））のいずれかに対するベースコールを生成する。得られたベースコールは、シーケンシングリードを形成する。

【0032】

一実施態様では、ニューラルネットワークベースのベースコーラー100は、特定のシーケンシングサイクルに対する単一のターゲットクラスタに対するベースコールを出力する。別の実施態様では、ニューラルネットワークベースのベースコーラー100は、特定のシーケンシングサイクルで複数のターゲットクラスタ内の各ターゲットクラスタに対するベースコールを出力する。更に別の実施態様では、ニューラルネットワークベースのベースコーラー100は、複数のシーケンシングサイクル内の各シーケンシングサイクルで、複数のターゲットクラスタ内の各ターゲットクラスタに対するベースコールを出力し、

それによって、各ターゲットクラスタに対するベースコールシーケンスを生成する。

【0033】

一実施態様では、ニューラルネットワークベースのベースコーラー100は、多層パーセプトロン（Multilayer Perceptron、MLP）である。別の実施態様では、ニューラルネットワークベースのベースコーラー100は、フィードフォワードニューラルネットワークである。更に別の実施態様では、ニューラルネットワークベースのベースコーラー100は、完全に接続されたニューラルネットワークである。更なる実施態様では、ニューラルネットワークベースのベースコーラー100は、完全畳み込みニューラルネットワークである。なお更なる実施態様では、ニューラルネットワークベースのベースコーラー100は、セマンティックセグメンテーションニューラルネットワークである。なお別の更なる実施態様では、ニューラルネットワークベースのベースコーラー100は、敵対的生成ネットワークニューラルネットワーク（generative adversarial network、GAN）である。

10

【0034】

一実施態様では、ニューラルネットワークベースのベースコーラー100は、複数の畳み込み層を有する畳み込みニューラルネットワーク（CNN）である。別の実施態様では、ニューラルネットワークベースのベースコーラー100は、長い短期メモリネットワーク（long short-term memory network、LSTM）、双方向LSTM（bi-directional LSTM、Bi-LSTM）、又はゲートされた反復単位（gated recurrent unit、GRU）などの反復ニューラルネットワーク（recurrent neural network、RNN）である。更に別の実施態様では、ニューラルネットワークベースのベースコーラー100は、CNN及びRNNの両方を含む。

20

【0035】

更に他の実施態様では、ニューラルネットワークベースのベースコーラー100は、1D畳み込み、2D畳み込み、3D畳み込み、4D畳み込み、5D畳み込み、拡張又は膨張畳み込み、転置畳み込み、奥行分離可能な畳み込み、点ごとの畳み込み、 1×1 の畳み込み、グループ畳み込み、扁平畳み込み、空間及びクロスチャネルの畳み込み、シャッフルされたグループ化畳み込み、空間的な分離可能な畳み込み、並びに逆畳み込みを使用することができる。ニューラルネットワークベースのベースコーラー100は、ロジスティック回帰／対数損失、多クラスクロスエントロピー／ソフトマックス損失、二値クロスエントロピー損失、平均二乗誤差損失、L1損失、L2損失、平滑L1損失、及びHuber損失などの、1つ以上の損失機能を使用することができる。ニューラルネットワークベースのベースコーラー100は、TFRecord、圧縮符号化（例えば、PNG）、シャープ化、マップ変換に対する平行コール、バッチング、プリフェッチ、モデル並列、データ並列、及び同期／非同期確率的勾配降下法（stochastic gradient descent、SDG）のような、任意の並列、効率、及び圧縮方式を使用することができる。ニューラルネットワークベースのベースコーラー100は、アップサンプリング層、ダウンサンプリング層、反復接続、ゲート及びゲートされたメモリユニット（LSTM又はGRUなど）、残留ブロック、残留接続、ハイウェイ接続、スキップ接続、ピープホール接続、活性化関数（例えば、非線形変換関数（正規化線形ユニット（rectifying linear unit、ReLU）、リーキーReLU、指数関数的ライナーユニット（exponential liner unit、ELU）、シグモイド、及び双曲線正接（tanh））など）、バッチ正規化層、規則化層、ドロップアウト、プール層（例えば、最大又は平均プール）、グローバル平均プール層、及び注意機構のような非線形変換機能を含み得る。

30

40

【0036】

ニューラルネットワークベースのベースコーラー100は、逆伝播ベースの勾配更新技術を使用して学習する。ニューラルネットワークベースのベースコーラー100が学習するために使用することができる例示的な勾配降下技術としては、確率的勾配降下、バッチ勾配降下、及びミニバッチ勾配降下が挙げられる。ニューラルネットワークベースのベースコーラー100が学習するために使用され得る勾配降下最適化アルゴリズムのいくつか

50

の例としては、Momentum、Nestor v 加速勾配、Adagrad、Adadelta、RMSprop、Adam、AdaMax、Nadam、及びAMSGradが挙げられる。

【0037】

一実施態様では、ニューラルネットワークベースのベースコーラー100は、異なるシーケンシングサイクルに対するデータの処理を分離するために、専用のアーキテクチャを使用する。上記の専用のアーキテクチャを使用する動機をまず説明する。上述のように、ニューラルネットワークベースのベースコーラー100は、現在のシーケンシングサイクル、1つ以上の先行するシーケンシングサイクル、及び1つ以上の後続のシーケンシングサイクルに対する画像パッチを処理する。追加のシーケンシングサイクルに対するデータは、シーケンスごとに固有のコンテキストを提供する。ニューラルネットワークベースのベースコーラー100は、学習中にシーケンスごとに固有のコンテキストを学習し、それらをベースコールする。更に、事前及び事後シーケンシングサイクルに対するデータは、プレフェージング及びフェージング信号の2次の寄与を現在のシーケンシングサイクルに提供する。

10

【0038】

しかしながら、異なるシーケンシングサイクルで、かつ異なる画像チャネル内に捕捉される画像は、位置合わせ不良であり、互いに残留位置合わせ誤差を有する。この位置合わせ不良を考慮するために、専用アーキテクチャは、シーケンシングサイクルどうしの間では情報を混合せず、同一のシーケンシングサイクル内でのみ情報を混合する、空間畳み込み層を含む。

20

【0039】

空間畳み込み層（又は空間ロジック）は、畳み込みの「専用の非共有」シーケンスを介して複数のシーケンシングサイクルの各々に対して独立してデータを処理することによって分離を操作する、いわゆる「分離された畳み込み」を使用する。分離された畳み込みは、任意の他のシーケンシングサイクルのデータ及び得られた特徴マップ上で畳み込むことなく、所与のシーケンシングサイクル、すなわち、サイクル内のみのデータ及び得られた特徴マップ上で畳み込む。

【0040】

例えば、入力画像データが、(i) ベースコールされる現在の（時間 t ）シーケンシングサイクルに対する現在の画像パッチ、(ii) 以前の（時間 $t-1$ ）シーケンシングサイクルに対する以前の画像パッチ、及び(iii) 次の（時間 $t+1$ ）シーケンシングサイクルに対する次の画像パッチと、を含むと考える。次いで、専用アーキテクチャは、3つの別個の畳み込みパイプライン、すなわち、現在の畳み込みパイプライン、以前の畳み込みパイプライン、及び次の畳み込みパイプラインを開始する。現在のデータ処理パイプラインは、現在の（時間 t ）シーケンシングサイクルに対する現在の画像パッチを入力として受信し、複数の空間畳み込み層を介して独立してそれを処理して、最終空間畳み込み層の出力としていわゆる「現在の空間畳み込み表現」を生成する。以前の畳み込みパイプラインは、以前の（時間 $t-1$ ）シーケンシングサイクルに対する以前の画像パッチを入力として受信し、複数の空間畳み込み層を介して独立してそれを処理して、最終空間畳み込み層の出力としていわゆる「以前の空間畳み込み表現」を生成する。次の畳み込みパイプラインは、次の（時間 $t+1$ ）シーケンシングサイクルに対する次のデータを入力として受信し、複数の空間畳み込み層を介して独立してそれを処理して、最終空間畳み込み層の出力としていわゆる「次の空間畳み込み表現」を生成する。

30

40

【0041】

いくつかの実施態様では、現在、以前、及び次の畳み込みパイプラインは、並行して実行される。いくつかの実施態様では、空間畳み込み層は、専用アーキテクチャ内の空間畳み込みネットワーク（又はサブネットワーク）の一部である。

【0042】

ニューラルネットワークベースのベースコーラー100は、シーケンシングサイクル間

50

、すなわち、サイクル間で情報を混合する時間畳み込み層（又は時間ロジック）を更に含む。時間畳み込み層は、空間畳み込みネットワークからそれらの入力を受信し、それぞれのデータ処理パイプラインに対して最終空間畳み込み層によって生成される空間畳み込み表現で動作する。

【0043】

時間畳み込み層のサイクル間動作性自由度は、空間畳み込みネットワークへの入力として供給される画像データ内に存在する位置合わせ不良特性が、空間畳み込み層のシーケンスによって実行される、分離された畳み込みのスタック又はカスケードによって空間畳み込み表現からパージされるという事実から生じる。

【0044】

時間畳み込み層は、スライドウィンドウベースでの後続の入力で入力チャンネル上でグループごとに畳み込む、いわゆる「組み合わせ畳み込み」を使用する。一実施態様では、後続の入力は、以前の空間畳み込み層又は以前の時間畳み込み層によって生成される後続の出力である。

【0045】

いくつかの実施態様では、時間畳み込み層は、専用アーキテクチャ内の時間畳み込みネットワーク（又はサブネットワーク）の一部である。時間畳み込みネットワークは、空間畳み込みネットワークからその入力を受信する。一実施態様では、時間畳み込みネットワークの第1の時間畳み込み層は、シーケンシングサイクル間の空間畳み込み表現をグループごとに組み合わせる。別の実施態様では、時間畳み込みネットワークの後続の時間畳み込み層は、以前の時間畳み込み層の後続する出力を組み合わせる。最終時間畳み込み層の出力は、出力を生成する出力層に供給される。出力は、1つ以上のシーケンシングサイクルで1つ以上のクラスタをベースコールするために使用される。

【0046】

ニューラルネットワークベースのベースコーラー100に関する追加の詳細は、参照により本明細書に組み込まれる、2019年3月21日に出版された「ARTIFICIAL INTELLIGENCE-BASED SEQUENCING」と題する米国仮特許出願第62/821,766号（代理人整理番号ILLM1008-9/IP-1752-PRV）に見出すことができる。

【0047】

圧縮ネットワーク

上述のように、ニューラルネットワークベースのベースコーラー100の専用アーキテクチャは、対応するシーケンシングサイクルの画像パッチのスライドウィンドウを処理する。後続のスライドウィンドウのシーケンシングサイクルどうしの間には重複が存在する。これにより、ニューラルネットワークベースのベースコーラー100は、重なり合うシーケンシングサイクルの画像パッチを冗長に処理することになる。これにより、計算リソースの無駄が生じる。例えば、一実施態様では、ニューラルネットワークベースのベースコーラー100の各空間畳み込み層は、ほぼ1億回の乗算操作を有する。次いで、5つのシーケンシングサイクル及び7つの空間畳み込み層のカスケード（又はシーケンス）のウィンドウに対して、空間畳み込みニューラルネットワークは、約6億2千万回の乗算操作を実行する。更に、時間畳み込みニューラルネットワークは、約1千万回の乗算操作を実行する。

【0048】

現在のスライドウィンドウ（又はベースコールの現在の反復）におけるサイクルN-1に対する画像データは、以前のスライドウィンドウ（又はベースコールの以前の反復）におけるサイクルNとして処理されるため、現在のスライドウィンドウで行われた処理の中間結果と、後続のウィンドウで行われた処理の中間結果を記憶して、それによって、後続のスライドウィンドウどうしの間で重なり合うシーケンシングサイクルに対する、入力画像データの冗長処理（又は再処理）を回避する（又は不要にする）という機会がもたらされる。

10

20

30

40

50

【0049】

しかしながら、中間結果は、実用的とは言えない容量のストレージを必要とする数テラバイトのデータである。この技術的問題を克服するために、開示された技術は、中間結果を、その中間結果が最初にニューラルネットワークベースのベースコーラー100によって生成される際に圧縮し、その後のスライドウィンドウでその圧縮された中間結果を再目的化して利用して冗長な計算を回避し、それによって、中間結果を再生成しない（又は、一度だけ生成する）ということを提案する。いくつかの実施態様では、開示される技術は、ニューラルネットワークベースのベースコーラー100の空間ネットワークにおいて約80%の畳み込みを節約する。一実施態様では、80%の節約は、その後のシーケンシングサイクルにおいて、圧縮ロジック及び圧縮特徴マップの再目的化しての利用が5回のシーケンシングサイクルの入力ウィンドウ（例えば、サイクルN、サイクルN+1、サイクルN-1、サイクルN+2、及びサイクルN-2）に対して使用されるときに、空間畳み込みにおいて観察される。別の実施態様では、その後のシーケンシングサイクルにおける圧縮ロジック及び圧縮特徴マップの再目的化しての利用が、10回のシーケンシングサイクルの入力ウィンドウ（例えば、サイクルN、サイクルN+1、サイクルN-1、サイクルN+2、サイクルN-2、サイクルN+3、及びサイクルN-3）に使用される場合、空間畳み込みにおいて90%の節約が観察される。すなわち、ウィンドウサイズが大きいほど、圧縮ロジックの使用及び圧縮特徴マップの再目的化からの節約が大きくなり、ウィンドウサイズが大きいほど、追加の隣接サイクルからのより大きなコンテキストの組み込みによるベースコール性能が良好である。より大きなウィンドウに対するより大きな節約は、所与の計算能力に対する全体的な性能を改善する。

【0050】

圧縮ロジックによってもたらされる計算効率及びコンパクトな計算フットプリントは、中央演算処理装置（Central Processing Unit、CPU）、グラフィック処理ユニット（Graphics Processing Unit、GPU）、フィールドプログラマブルゲートアレイ（FPGA）、粗粒度再構成可能アーキテクチャ（Coarse-Grained Reconfigurable Architecture、CGRA）、特定用途向け集積回路（ASIC）、特定用途向け命令セットプロセッサ（Application Specific Instruction-set Processor、ASIP）、及びデジタル信号プロセッサ（DSP）のようなリソース制約プロセッサ上のニューラルネットワークベースのベースコーラー100のハードウェア実装を容易にする。

【0051】

圧縮ロジックによって節約された計算は、ニューラルネットワークベースのベースコーラー100により多くの畳み込み演算子を組み込むことを可能にする。例としては、空間畳み込み層及び時間畳み込み層により多くの畳み込みフィルタを追加すること、畳み込みフィルタのサイズを増加させること、空間畳み込み層及び時間畳み込み層の数を増加させることが挙げられる。追加の畳み込み演算は、ニューラルネットワークベースのベースコーラー100の強度パターン検出及び全体的なベースコール精度を改善する。

【0052】

圧縮ロジックによって節約された計算はまた、対象スライドウィンドウの入力画像データを拡張して、より多くのシーケンシングサイクルを含むようにすることを可能にする。拡張されたスライドウィンドウは、追加の隣接するシーケンシングサイクルから余剰画像パッチをもたらすことによって、ベースコールコンテキストを広げる。

【0053】

更に、元の間中間結果とは対照的に、圧縮された中間結果の使用によって生じ得る精度のいかなる低下も、追加の畳み込み演算子の組み込み及びスライドウィンドウの拡張によって補償される。

【0054】

図1Aは、ベースコールの第1の反復のための圧縮空間マップセットを生成する、開示された圧縮ロジックの一実施態様を示す。図示の例では、シーケンシングサイクルの第1のウィンドウは、シーケンシングサイクル1、2、3、4、及び5を含む。それぞれのシ

ーケンシングサイクル1、2、3、4、及び5に対するそれぞれの画像パッチ102、112、122、132、及び142（又はサイクルごとの分析物チャネルセット）は、空間ロジック104（又は空間ネットワーク若しくは空間サブネットワーク若しくは空間畳み込みニューラルネットワーク）によって別々に処理されて、それぞれのシーケンシングサイクル1、2、3、4、及び5のそれぞれの空間マップ106、116、126、136、及び146（又は中間結果若しくは空間出力セット若しくは空間特徴マップセット）を生成する。空間畳み込みネットワーク104は、1D、2D、又は3D畳み込みを使用することができる。

【0055】

空間ロジック104は、空間畳み込み層のシーケンス（又はカスケード）を含む。各空間畳み込み層は、分離された畳み込みを実装する複数の空間畳み込みフィルタを有するフィルタバンクを有する。したがって、各空間畳み込み層は、複数の空間特徴マップを出力として生成する。ある対象のサブジェクト空間畳み込み層によって生成される空間特徴マップの数は、その対象の空間畳み込み層に構成された空間畳み込みフィルタの数の関数である。例えば、対象の空間畳み込み層が14個の空間畳み込みフィルタを有する場合、対象の空間畳み込み層は14個の空間特徴マップを生成する。集合的観点から、14個の空間特徴マップは、14個のチャネル（又は深さ次元＝14）との空間特徴マップの体積（又はテンソル）と見なすことができる。

【0056】

更に、対象の空間畳み込み層に続く次の空間畳み込み層も、14個の空間畳み込みフィルタで構成することができる。そのような場合、次の空間畳み込み層は、入力として、14個の空間特徴マップを処理して、対象の空間畳み込み層を生成し、その対象の空間畳み込み層自体が、14個の新しい空間特徴マップを出力として生成する。図1Aは、それぞれのシーケンシングサイクル1、2、3、4、及び5に対して、空間ネットワーク104の最終空間畳み込み層によって生成された、5つの空間特徴マップセット106、116、126、136、及び146を示す。図示の例では、5つの空間特徴マップセット106、116、126、136、及び146の各々は、14個の特徴マップを有する。

【0057】

図1Dは、空間ネットワーク104の7つの空間畳み込み層のカスケードによって生成された7つの空間特徴マップセット196a、196b、196c、196d、196e、196f、及び196gのシーケンスを示す。対象のシーケンシングサイクルiに対する、サイクルごとの入力パッチデータ194は、空間次元数が 115×115 であり、（元のシーケンス画像内の2つの画像チャネルによる）深さ次元数が2である。一実施態様では、7つの空間畳み込み層の各々は、後続の空間特徴マップボリュームの空間次元数を2だけ、例えば、 10×10 から 8×8 に低減する、 3×3 の畳み込みを使用する。

【0058】

第1の空間特徴マップボリューム196aは、 113×113 の空間次元（すなわち、第1の空間畳み込み層の 3×3 の畳み込みによって 115×115 から減少したもの）と、14の深さ次元（すなわち、第1の空間畳み込み層内の14個の空間畳み込みフィルタによる、14個の特徴マップ又は14個のチャネル）と、を有する。第2の空間特徴マップボリューム196bは、 111×111 の空間次元（すなわち、第2の空間畳み込み層の 3×3 の畳み込みによって 113×113 から減少したもの）と、14の深さ次元（すなわち、第2の空間畳み込み層内の14個の空間畳み込みフィルタによる14個の特徴マップ又は14個のチャネル）と、を有する。第3の空間特徴マップボリューム196cは、 109×109 の空間次元（すなわち、第3の空間畳み込み層の 3×3 の畳み込みによって 111×111 から減少したもの）と、14の深さ次元（すなわち、第3の空間畳み込み層内の14個の空間畳み込みフィルタによる14個の特徴マップ又は14個のチャネル）と、を有する。第4の空間特徴マップボリューム196dは、 107×107 の空間次元（すなわち、第4の空間畳み込み層の 3×3 の畳み込みによって 109×109 から減少したもの）と、14の深さ次元（すなわち、第4の空間畳み込み層内の14個の空間

畳み込みフィルタによる14個の特徴マップ又は14個のチャンネル)と、を有する。第5の空間特徴マップボリューム196eは、 105×105 の空間次元(すなわち、第5の空間畳み込み層の 3×3 の畳み込みによって 107×107 から減少したもの)と、14の深さ次元(すなわち、第5の空間畳み込み層内の14個の空間畳み込みフィルタによる14個の特徴マップ又は14個のチャンネル)と、を有する。第6の空間特徴マップボリューム196fは、 103×103 の空間次元(すなわち、第6の空間畳み込み層の 3×3 の畳み込みによって 105×105 から減少したもの)と、14の深さ次元(すなわち、第6の空間畳み込み層内の14個の空間畳み込みフィルタによる14個の特徴マップ又は14個のチャンネル)と、を有する。第7の空間特徴マップボリューム196gは、 101×101 の空間次元(すなわち、第7の空間畳み込み層の 3×3 の畳み込みによって 103×103 から減少したもの)と、14の深さ次元(すなわち、第7の空間畳み込み層内の14個の空間畳み込みフィルタによる14個の特徴マップ又は14個のチャンネル)と、を有する。

【0059】

図1Aに示されるマルチサイクルの例に類似して、5回のシーケンシングサイクル1、2、3、4、及び5と、サイクルごとの5つの画像パッチ102、112、122、132、及び142とに対して、空間ロジック104は、7つの空間特徴マップボリューム196a、196b、196c、196d、196e、196f、及び196gの、5つのそれぞれのシーケンスを別々に生成する。なお、図1Aの空間マップ106、116、126、136、及び146が、図1Dの最終空間特徴マップボリューム196gの5つの別個のインスタンスに同等である。

【0060】

圧縮ロジック108(又は圧縮ネットワーク若しくは圧縮サブネットワーク若しくは圧縮層若しくはスクイーズ層)は、空間ロジック104の出力を処理し、出力の圧縮表現を生成する。一実施態様では、圧縮ネットワーク108は、空間ネットワーク104によって生成された特徴マップの深さ次元数を低減する圧縮畳み込み層を備える。

【0061】

例えば、図1Aでは、空間マップ106、116、126、136、及び146の深さ次元数は14である(すなわち、空間出力当たり14個の特徴マップ又は14個のチャンネル)。圧縮ネットワーク108は、空間マップ106、116、126、136、及び146を、それぞれのシーケンシングサイクル1、2、3、4、及び5に対するそれぞれの圧縮空間マップセット110、120、130、140、及び150に減縮させる。圧縮空間マップセット110、120、130、140、及び150のそれぞれは、深さ次元数(すなわち、圧縮空間出力当たり2つの特徴マップ、又は2つのチャンネル)が2である。他の実施態様では、圧縮空間マップセット110、120、130、140、及び150は、深さ次元数(すなわち、圧縮空間出力当たり、3つ若しくは4つの特徴マップ、又は3つ若しくは4つのチャンネル)が3又は4であってもよい。更に他の実施態様では、圧縮空間マップセット110、120、130、140、及び150は、深さ次元数(すなわち、圧縮空間出力あたり1つの特徴マップ、又は1つのチャンネル)が1であってもよい。一実施態様では、圧縮層108は、ReLUのような活性化関数を含まない。他の実施態様では、圧縮層108は、活性化関数を含むことができる。他の実施態様では、圧縮ロジック108は、対応する圧縮空間マップセットを、各々が4つより多くの特徴マップを有するように構成することができる。

【0062】

ここで、圧縮ロジック108が圧縮出力をどのように生成するかを論じる。

【0063】

一実施態様では、圧縮ロジック108は、非線形性を導入しながら、特徴マップの数(すなわち、深さ次元、又はチャンネルの数)を減少させるために 1×1 の畳み込みを使用する。 1×1 の畳み込みは、1のカーネルサイズを有する。 1×1 の畳み込みは、空間次元を変化させることなく、体積深さを別のスクイーズされた又は拡張された表現に変換する

ことができる。1×1の畳み込みは、入力チャンネルにわたる、完全接続線形層のように動作する。これは、多くのチャンネルを有する特徴マップから、より少数の特徴マップにマッピングするのに有用である。図1Eでは、単一の1×1の畳み込みが、2つの特徴マップを有する入力テンソルに適用される。1×1の畳み込みは、2チャンネル入力を単一チャンネルの出力に圧縮する。

【0064】

圧縮層108によって生成された圧縮出力（又は圧縮特徴マップ若しくは圧縮空間マップ若しくは圧縮時間マップ）の数は、圧縮層108内に構成された1×1の畳み込みフィルタ（又は圧縮畳み込みフィルタ若しくは圧縮フィルタ）の数の関数である。図1Fでは、圧縮層108は、2つの1×1の畳み込みフィルタ198a及び198bを有する。第1の1×1の畳み込みフィルタ198aは、14個の特徴マップを有する空間特徴ボリューム196gを処理し、101×101の空間次元数を保存しながら第1の特徴マップ199aを生成する。第2の1×1の畳み込みフィルタ198bはまた、14個の特徴マップを有する空間特徴ボリューム196gを処理し、101×101の空間次元数を保存しながら第2の特徴マップ199bを生成する。したがって、圧縮層108は、14個の特徴マップを有する空間特徴ボリューム196gを、2つの空間特徴マップ199a及び199bを有する圧縮出力（すなわち、圧縮比＝7）に減縮する。

【0065】

時系列の観点から、シーケンシングサイクル5は、中央シーケンシングサイクル（N）であり、シーケンシングサイクル1及び2は、左側隣接シーケンシングサイクル（N－1、N－2）であり、シーケンシングサイクル4及び5は、左側隣接シーケンシングサイクル（N＋1、N＋2）である。したがって、中央の圧縮出力130は、中央シーケンシングサイクル（N）に対して生成され、左側隣接圧縮出力120は、左側隣接シーケンシングサイクル（N－1）に対して生成され、更なる左側隣接圧縮出力110は、更なる左側隣接シーケンシングサイクル（N－2）に対して生成され、右側隣接圧縮出力140は、右側隣接シーケンシングサイクル（N＋1）に対して生成され、更なる右側隣接圧縮出力150は、更なる右側隣接シーケンシングサイクル（N＋2）に対して生成される。

【0066】

パイプラインの観点から、ニューラルネットワークベースのベースコーラー100は、（例えば、データ並列処理に基づくマルチスレッド実行又はマルチクラスタ化実行として）空間ロジック104及び圧縮ロジック108によって、画像パッチ102、112、122、132、及び142をそれぞれ処理する、5つの並列で独立のパイプラインを実行する。したがって、5つの圧縮出力110、120、130、140、及び150は、別々に、同時に、かつ独立して、ニューラルネットワークベースのベースコーラー100によって生成される。

【0067】

いくつかの実施態様では、圧縮層108は、空間ネットワーク104の最終空間畳み込み層と見なすことができる。他の実施態様では、圧縮ネットワーク108は、ニューラルネットワークベースのベースコーラー100の専用アーキテクチャ内又はその外側の、別個のネットワークと見なすことができる。

【0068】

図1Bは、ニューラルネットワークベースのベースコーラー100の時間ロジック160（又は時間ネットワーク若しくは時間サブネットワーク若しくは時間畳み込みニューラルネットワーク）を介して圧縮空間マップセット110、120、130、140、及び150を処理する一実施態様を示す。時間ロジック160は、スライドウィンドウベースで、連続した圧縮空間マップセットのグループを処理する。例えば、図1Bでは、時間ロジック160は、シーケンシングサイクル1、2、及び3にそれぞれ対する圧縮空間マップセット110、120、及び130の第1のグループ／ウィンドウを処理し、出力として、時間マップ172（又は時間マップセット若しくは時間特徴マップ若しくは時間特徴マップセット）を生成する。時間ロジック160は、それぞれのシーケンシングサイクル

2、3、及び4に対する圧縮空間マップセット120、130、及び140の第2のグループ／ウィンドウを処理し、出力として、時間マップ174を生成する。時間ロジック160は、それぞれのシーケンシングサイクル3、4、及び5に対する圧縮空間マップセット130、140、及び150の第3のグループ／ウィンドウを処理し、出力として、時間マップ176を生成する。時間畳み込みネットワーク160は、1D、2D、又は3D畳み込みを使用することができる。

【0069】

図1Bに示される時間ロジック160の3つのインスタンスは、時間ネットワーク160の第1の時間畳み込み層の3つのフィルタバンクを表す。第1のフィルタバンクは、圧縮空間マップ110、120、及び130の第1のグループに、時間畳み込みフィルタの第1のセットを適用し、時間マップの第1のセット172を生成する。第2のフィルタバンクは、圧縮空間マップ120、130、及び140の第2のグループに、時間畳み込みフィルタの第2のセットを適用し、時間マップの第2のセット174を生成する。第3のフィルタバンクは、圧縮空間マップ130、140、及び150の第3のグループに、時間畳み込みフィルタの第3のセットを適用し、時間マップの第3のセット176を生成する。

【0070】

時間マップの第1のセット172、第2のセット174、及び第3のセット176は、時間ロジック160によってグループとして処理されて、時間マップ182を生成する。図1Bに示される時間ロジック160の第4のインスタンスは、時間ネットワーク160の第2の時間畳み込み層を表し、この第2の時間畳み込み層は、全てのシーケンシングサイクル1、2、3、4、及び5に対する出力を生成し、この全てのシーケンシングサイクル1、2、3、4、及び5に対して、サイクルごとの画像パッチ対が、図1Aのニューラルネットワークベースのベースコーラー100への入力として供給された。

【0071】

時間ネットワーク160は、時間畳み込み層のカスケード（例えば、シーケンスに配置された2つ、3つ、4つ、5つ以上の時間畳み込み層）を有する。時間畳み込み層のカスケードは、異なるレベルのグループ化を有する階層形式でデータを処理する。すなわち、所与のレベルでは、スライドウィンドウアプローチが、その所与のレベルでの入力をグループ式で処理して、出力を生成し、この出力は、その後、スライドウィンドウ様式の次のレベルで、グループ式で処理されるものである。

【0072】

時間畳み込み層は、組み合わせ畳み込みを実装する時間畳み込みフィルタを用いて構成される。組み合わせ畳み込みは、複数のシーケンシングサイクルにわたる特徴マップどうしの中で、情報を混合する。組み合わせ畳み込みは、時間ネットワーク160内の現在のレベルの、対象のグループ／ウィンドウにおける後続のシーケンシングサイクルどうしの中で、データを組み合わせる。例えば、第1の時間畳み込み層は、第1のグループのシーケンシングサイクル1、2、及び3に対する、第1のグループの圧縮空間マップ110、120、及び130を組み合わせ、時間マップの第1のセット172を生成し、第1の時間畳み込み層は更に、第2のグループのシーケンシングサイクル2、3、及び4に対する、第2のグループの圧縮空間マップ120、130、及び140を組み合わせ、時間マップの第2のセット174を生成し、更に、第1の時間畳み込み層は、第3のグループのシーケンシングサイクル3、4、及び5に対する、第3のグループの圧縮空間マップ130、140、及び150を組み合わせ、時間マップの第3のセット176を生成する。

【0073】

組み合わせ畳み込みはまた、時間ネットワーク160内の現在のレベルの、対象のグループ／ウィンドウにおける、シーケンシングサイクルの連続するグループどうしの中で、データを組み合わせる。例えば、第2の時間畳み込み層は、時間マップの第1のセット172、第2のセット174、及び第3のセット176を組み合わせ、時間マップの最終的なセット182にする。レベル2では、レベル1からのシーケンシングサイクルの第1

10

20

30

40

50

、第2、及び第3のグループ／ウィンドウが、シーケンシングサイクル1、2、3、4、及び5の第1のグループ／ウィンドウにグループ化される。

【0074】

組み合わせ畳み込みは、組み合わせられる入力の数と同じ数のカーネルを有して構成されている（すなわち、時間畳み込みフィルタの深さカラム又はファイバは、現在のレベルでの対象のグループ／ウィンドウにおける入力の数と一致する）。例えば、時間畳み込み層が3つの圧縮空間マップを組み合わせる場合、その時間畳み込み層は、3つの圧縮空間マップの深さ全体を通して要素ごとの乗算及び和算を実行する3つのカーネルを各々有する、複数の時間畳み込みフィルタを使用する。

【0075】

時間マップの最終セット182は、時間ネットワーク160の最終的な（又は最後の）時間畳み込み層によって生成される。図1Cは、開示された出力ロジック190（又は出力層若しくは出力ネットワーク若しくは出力サブネットワーク）を介して、最終時間マップセット182を処理して、ベースコール分類データを生成する一実施態様を示す。一実施態様では、複数のクラスタが、1つ以上のシーケンシングサイクルに対して、同時にベースコールされる。図1Cに示される例では、ベースコール192は、中央シーケンシングサイクル3に対してのみ多くのクラスタについて生成される。他の実施態様では、開示された技術により、出力ロジック190は、所与の入力ウィンドウに対して、中央のシーケンシングサイクルだけでなく、隣接するシーケンシングサイクル（オプションの点線で示される）に対しても、1つの実施態様によってベースコールを生成する。例えば、一実施態様では、開示される技術は、所与の入力ウィンドウに対して、サイクルN、サイクルN+1、サイクルN-1、サイクルN+2、サイクルN-2などに対するベースコールを、同時に生成する。すなわち、ニューラルネットワークベースのベースコーラー102の、1回の順方向伝播／トラバース／ベースコール反復では、シーケンシングサイクルの入力ウィンドウにおける複数のシーケンシングサイクルに対するベースコールが生成され、これは本明細書では「多対多のベースコール」と呼ばれる。

【0076】

出力層190の例としては、ソフトマックス関数、 $\log\text{-softmax}$ 関数、アンサンブル出力平均関数、多層パーセプトロン不確実関数、ベイズガウス分布関数、及びクラスタ強度関数が挙げられる。一実施態様では、出力層190は、各クラスタに対して、かつ各シーケンシングサイクルに対して、クラスタごと、サイクルごとの確率四分位数を生成する。

【0077】

以下の考察は、一例としてソフトマックス関数を使用して、クラスタごと、サイクルごとの確率四分位数に焦点を当てている。最初に、ソフトマックス関数、次いでクラスタごと、サイクルごとの確率四分位数を説明する。

【0078】

ソフトマックス関数は、マルチクラス分類のための好ましい関数である。ソフトマックス関数は、全ての可能な標的クラスにわたって各標的クラスの確率を計算する。ソフトマックス関数の出力範囲はゼロと1との間であり、全ての確率の合計は1に等しい。ソフトマックス関数は、所与の入力値の指数及び全ての入力値の指数値の合計を計算する。入力値の指数と指数値の合計との比は、本明細書で「指数正規化」と称されるソフトマックス関数の出力である。

【0079】

形式的には、いわゆるソフトマックス分類子を学習することは、クラスではなくむしろ各クラスの確率の信頼性予測を返すため、真の分類子よりもむしろクラス確率への回帰である。ソフトマックス関数は、ある種類の値を取り、合計で1になる確率にそれらを変換する。ソフトマックス関数は、任意の実数値のn次元ベクトルを0～1の範囲内の実数値のn次元ベクトルに押し込む。したがって、ソフトマックス関数を使用することは、出力が有効で、指数関数的に正規化された確率質量関数（非負及び合計1になる）であること

10

20

30

40

50

を保証する。

【0080】

直感的に、ソフトマックス関数は、最大関数の「ソフト」バージョンである。「ソフト」という用語は、ソフトマックス関数が連続的であり微分可能であるという事実に由来する。1つの最大要素を選択する代わりに、最大入力要素が比例してより大きな値を取得し、他方が値の割合が少なくなるように、ベクトルを全体の部分に分解する。確率分布を出力する特性は、分類タスクにおける確率的解釈に適したソフトマックス関数をもたらす。

【0081】

ソフトマックス層への入力のベクトルとして、 z を考慮するとしよう。ソフトマックス層ユニットは、ソフトマックス層内のノードの数であり、したがって、 z ベクトルの長さは、ソフトマックス層内のユニットの数である（10個の出力ユニットを有する場合、10個の z 要素がある）。

【0082】

n -次元ベクトル $Z = [z_1, z_2, \dots, z_n]$ の場合、ソフトマックス関数は、指数正規化（exp）を使用して、その和が1になる、範囲 $[0, 1]$ の範囲内の正規化された値を有する別の n -次元ベクトル $p(Z)$ を生成する。

【0083】

【数1】

$$Z = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix} \text{ 及び } p(Z) \rightarrow \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_n \end{bmatrix}$$

$$p_j = \frac{\exp^{z_j}}{\sum_{k=1}^n \exp^{z_k}} \forall j \in 1, 2, \dots, n$$

【0084】

図1Gは、例示的なソフトマックス関数を示す。ソフトマックス関数は、

【0085】

【数2】

$$z \mapsto \text{ソフトマックス} \left(\left[z; \frac{z}{10}; -2z \right] \right)$$

として3つのクラスに適用される。3つの出力は常に、合計1になることに留意されたい。したがって、それらは、離散確率質量関数を定義する。

【0086】

特定のクラスごと、サイクルごとの確率四分位数は、特定のシーケンシングサイクルで特定のクラスに組み込まれた塩基である、A、C、T、及びGの確率を識別する。ニューラルネットワークベースのベースコーラー100の出力層がソフトマックス関数を使用する場合、クラスごと、サイクルごとの確率四分位数は、合計すると1になる、指数的に正規化された分類スコアである。図1Hは、クラス1（121、茶色で示される）に対する、かつシーケンシングサイクル1～S（122）に対するソフトマックス関数によって生成された、クラスごと、サイクルごとの確率四分位数123の例を示す。言い換えれば、シーケンシングサイクルの第1のサブセットは、S個のシーケンシング

サイクルを含む。

【0087】

信頼できないクラスタ識別子125は、クラスタ、サイクルごとの確率四分位数から、フィルタ値を生成することに基づいて、信頼できないクラスタを識別する。本出願では、クラスタごと、サイクルごとの確率四分位数は、ベースコール分類スコア又は正規化ベースコール分類スコア又は初期ベースコール分類スコア又は正規化された初期ベースコール分類スコア又は初期ベースコールとも呼ばれる。

【0088】

フィルタ計算機127は、それが識別する確率に基づいて、クラスタごと、サイクルごとの確率四分位数に対するフィルタ値を決定し、それによって各クラスタのフィルタ値のシーケンスを生成する。フィルタ値のシーケンスは、フィルタ値124として記憶される。

【0089】

クラスタごと、サイクルごとの確率四分位数に対するフィルタ値は、確率のうちの1つ以上を含む計算に基づいて決定される。一実施態様では、フィルタ計算機127によって使用される計算は減算である。例えば、図1Hに示す実施態様では、クラスタごと、サイクルごとの確率四分位数に対するフィルタ値は、確率のうちの最も高い確率（マゼンタ色で示される）から確率の2番目に高い確率（青色に示される）を差し引くことによって決定される。

【0090】

別の一実施態様では、フィルタ計算機116によって使用される計算は除算である。例えば、クラスタごと、サイクルごとの確率四分位数に対するフィルタ値は、確率のうちの最も高い確率（マゼンタ色に示されている）の確率のうちの2番目に高い確率（青色で示される）に対する比として決定される。更に別の一実施態様では、フィルタ計算機127によって使用される計算は、加算である。なお更なる実施態様では、フィルタ計算機127によって使用される計算は、乗算である。

【0091】

一実施態様では、フィルタ計算機127は、フィルタリング関数を使用してフィルタ値124を生成する。一例では、フィルタリング関数は、最も明るいベース強度を、最も明るいベース強度と2番目に明るいベース強度との合計で割った比として、チャスティティを定義するチャスティフィルタである。別の例では、フィルタリング関数は、最大対数確率関数、最小二乗誤差関数、平均信号対雑音比（signal-to-noise ratio、SNR）、及び最小絶対誤差関数のうちの少なくとも1つである。

【0092】

信頼できないクラスタ識別子125は、フィルタ値124を使用して、複数のクラスタ内のいくつかのクラスタを、信頼できないクラスタ128として識別する。信頼できないクラスタ128を識別するデータは、コンピュータ可読形式又は媒体にあり得る。信頼できないクラスタは、器具ID、機器上のラン番号、フローセルID、レーン番号、タイル番号、クラスタのX座標、クラスタのY座標、及び固有の分子識別子（unique molecular identifier、UMI）によって識別することができる。信頼できないクラスタ識別子125は、複数のクラスタ内の、フィルタ値のシーケンスが閾値「H」を下回る「G」数のフィルタ値を含むクラスタを、信頼できないクラスタとして、識別する。一実施態様では、「G」は1～5の範囲である。別の一実施態様では、「H」は0.5～0.99の範囲である。一実施態様では、信頼できないクラスタ128は、信頼できないクラスタに対応する（すなわちその強度放射を描く）ピクセルを識別する。このようなピクセルは、本出願において後述するように、フィルタリングロジック502によってフィルタリングされ取り除かれる。

【0093】

信頼できないクラスタは、所望の信号を、バックグラウンド信号と比較して有意ではない量しか発しない、低品質クラスタである。信頼できないクラスタの信号対雑音比は、実質的に低く、例えば、1未満である。いくつかの実施態様では、信頼できないクラスタは

10

20

30

40

50

、所望の信号を全く生成しない場合がある。他の実施態様では、信頼できないクラスタは、バックグラウンドと比較して非常に少ない量の信号しか生成しない場合があり得る。一実施態様では、信号は、光信号であり、例えば、蛍光、発光、散乱、又は吸収信号を含むことを意図する。信号レベルとは、所望又は所定の特性を有する検出されたエネルギー又は符号化された情報の量を意味する。例えば、光信号は、強度、波長、エネルギー、周波数、電力、輝度などのうちの1つ以上によって定量化することができる。他の信号は、電圧、電流、電界強度、磁場強度、周波数、電力、温度などの特性に従って定量化することができる。信頼できないクラスタにおける信号の不在は、ゼロの信号レベル、又はノイズとは有意に区別されない信号レベルであると理解される。

【0094】

信頼できないクラスタの不十分な品質の信号には、多くの潜在的な理由がある。信頼できないクラスタ中の約1000個の分子のうちの相当大きな割合が、特定の位置で異なる塩基を含むように、コロニー増幅におけるポリメラーゼ連鎖反応 (polymerase chain reaction、PCR) エラーが存在する場合、2つの塩基に対する信号を観察し得るが、これは、不十分な品質のしるしとして解釈され、フェーズエラーと称される。フェーズエラーは、信頼できないクラスタ内の個々の分子が、(例えば、フェージングと呼ばれる、3'ターミネーターの不完全な除去を原因として)いくつかのサイクルでヌクレオチドを組み込まず、他の分子よりも遅れてしまう場合、又は(例えば、プリフェージングと呼ばれる、効果的な3'ブロックなしでヌクレオチドの組み込みを原因として)個々の分子が単一のサイクルで2つ以上のヌクレオチドを組み込んでいる場合に生じる。これにより、シーケンスコピーの読み出しにおける、同期の喪失がもたらされる。信頼できないクラスタにおける、フェージング及びプリフェージングによって影響を受けるシーケンスの割合は、サイクル数の増加と共に増加し、読み取りの品質が高いサイクル数で低下する傾向がある主な理由である。

【0095】

信頼できないクラスタはまた、フェーディングの結果として生じる。フェーディングは、サイクル数の関数としての信頼できないクラスタの信号強度における指数関数的減衰である。シーケンシングランが進行するにつれて、信頼できないクラスタのストランドが過度に洗浄され、反応種を作成するレーザ放出に曝露され、過酷な環境条件に置かれる。これらの全ては、信頼できないクラスタにおいて断片が徐々に失われる結果を招き、それらの信号強度を低下させる。

【0096】

信頼できないクラスタはまた、発育の十分ではないコロニー、すなわち、パターン形成されたフローセル上に空又は部分的にしか充填されないウェルを生成してしまう、信頼できないクラスタの小さなクラスタサイズにも起因する。すなわち、いくつかの実施態様では、信頼できないクラスタは、パターン化されたフローセル上の、空のウェル、多クローン性ウェル、及び曖昧なウェルを示す。信頼できないクラスタはまた、非排他的な増幅によって引き起こされる、重なり合うコロニーから生じる。信頼できないクラスタはまた、例えば、フローセルの縁部に位置することに起因する、照明不足又は不均一な照明から生じる。信頼できないクラスタはまた、放出された信号を不明確化するフローセル上の不純物から生じる。信頼できないクラスタはまた、複数のクラスタが同一のウェルに堆積される場合に生じる、多クローン性クラスタも含む。

【0097】

シーケンシングサイクルの第1のウィンドウは、シーケンシングサイクル1、2、3、4、及び5を含み、ベースコールの第1の反復は、中央シーケンシングサイクル3に対するベースコール192を生成する。シーケンシングサイクルの第2のウィンドウは、シーケンシングサイクル2、3、4、5、及び6を含み、ベースコールの第2の反復は、中央シーケンシングサイクル4に対するベースコール292を生成する。したがって、シーケンシングサイクル2、3、4、及び5は、第1のウィンドウと第2のウィンドウとの間の、又はベースコールの第2の反復と第3の反復との間の、重なり合うシーケンシングサイ

10

20

30

40

50

クルである。

【0098】

開示されたベースコーリングシステム及び技術は、シーケンシングサイクル2、3、4、及び5のベースコールの第1の反復中に生成される圧縮空間マップセット120、130、140、及び150を、メモリ（例えば、オンチップDRAM、オンチップSRAM又はBRAM、オフチップDRAM）に格納する。ベースコールの第2の反復中、開示されたベースコールシステム及び技術は、重なり合うサイクル2、3、4、及び5に対するそれぞれの入力画像パッチ112、122、132、及び142を、空間ネットワーク104を介して再処理しない。その代わりに、開示されたベースコールシステム及び技術は、ベースコールの第2の反復中に、それぞれの入力画像パッチ112、122、132、及び142に替えて、以前に生成された圧縮空間マップセット120、130、140、及び150を再利用する。

10

【0099】

圧縮ロジックは、圧縮空間マップセット120、130、140、及び150、並びにそれぞれの入力画像パッチ112、122、132、及び142が、同じ数のサイクル特徴マップ／チャンネルを有することを求めるように、更に設定されている。これにより、圧縮空間マップセット120、130、140、及び150が、それぞれの入力画像パッチ112、122、132、及び142の損失のない表現であることを保証する。すなわち、それぞれの入力画像パッチ112、122、132、及び142が、各々2つの特徴マップ／チャンネルを有する場合、圧縮ロジック108は、圧縮空間マップセット120、130、140、及び150も、2つの特徴マップ／チャンネルを有するように設定する。同様に、それぞれの入力画像パッチ112、122、132、及び142が、各々3つの特徴マップ／チャンネルを有する場合、圧縮ロジック108は、圧縮空間マップセット120、130、140、及び150も、3つの特徴マップ／チャンネルを有するように設定する。同じ様に、それぞれの入力画像パッチ112、122、132、及び142が、各々4つの特徴マップ／チャンネルを有する場合、圧縮ロジック108は、圧縮空間マップセット120、130、140、及び150も、4つの特徴マップ／チャンネルを有するように設定する。

20

【0100】

図2Aは、ベースコールの第2の反復中に、空間ロジック104及び圧縮ロジック108によって入力画像データ222を処理することによって、重なり合わないシーケンシングサイクル6に対してのみ、空間マップ226及び対応する圧縮空間マップ230が生成されることを示す。したがって、重なり合うサイクル2、3、4、及び5に対する入力画像パッチ112、122、132、及び142（凡例中の灰色の網掛けで強調表示される）は、重なり合う畳み込みを回避するため、再度処理されることはない。

30

【0101】

図2Bは、ベースコールの第1の反復中に生成された圧縮空間マップセット120、130、140、及び150が、ベースコールの第2の反復中に生成された圧縮空間マップセット230と併せて使用されて、中央シーケンシングサイクル4に対するベースコール292を生成することを示す。図2Bでは、時間マップセット174、176、及び278が、図1Bに関して上で論じたものと同様の方法で、時間ネットワーク160の第1の時間畳み込み層によって生成される。時間マップセット282は、図1Bに関して上述したものと同様の方法で、時間ネットワーク160の第2のかつ最後の時間畳み込み層によって生成される。図2Cは、出力層190が、ベースコールの第2の反復中に生成された最終時間マップセット282を処理し、中央シーケンシングサイクル4に対するベースコール292を生成することを示す。

40

【0102】

シーケンシングサイクルの第3のウィンドウは、シーケンシングサイクル3、4、5、6、及び7を含み、ベースコールの第3の反復は、中央シーケンシングサイクル5に対するベースコール392を生成する。したがって、シーケンシングサイクル3、4、5、及

50

び6は、第2のウィンドウと第3のウィンドウとの間の、又はベースコールの第2の反復と第3の反復との間の、重なり合うシーケンシングサイクルである。

【0103】

開示されたベースコーリングシステム及び技術は、それぞれのシーケンシングサイクル3、4、及び5に対するベースコールの第1の反復中に生成される圧縮空間マップセット130、140、及び150と、シーケンシングサイクル6に対するベースコールの第2の反復中に生成される圧縮空間マップセット230とを、メモリ（例えば、オンチップDRAM、オンチップSRAM又はBRAM、オフチップDRAM）に格納する。ベースコールの第3の反復中、開示されたベースコーリングシステム及び技術は、重なり合うサイクル3、4、5、及び6に対するそれぞれの入力画像パッチ122、132、142、及び222を、空間ネットワーク104を介して再処理しない。その代わりに、開示されたベースコーリングシステム及び技術は、ベースコールの第3の反復中に、それぞれの入力画像パッチ122、132、142、及び222に替えて、以前に生成された圧縮空間マップセット130、140、150、及び230を再利用する。

10

【0104】

図3Aは、ベースコールの第3の反復中に、空間ロジック104及び圧縮ロジック108によって入力画像データ322を処理することによって、重なり合わないシーケンシングサイクル7に対してのみ、空間マップ326及び対応する圧縮空間マップ330が生成されることを示す。したがって、重なり合うサイクル3、4、5、及び6に対する入力画像パッチ122、132、142、及び222（凡例中の灰色の網掛けで強調表示される）は、重なり合う畳み込みを回避するため、再度処理されることはない。

20

【0105】

図3Bは、ベースコールの第1及び第2の反復中に生成された圧縮空間マップセット130、140、150、及び230が、ベースコールの第3の反復中に生成された圧縮空間マップセット330と併せて使用されて、中央シーケンシングサイクル5に対するベースコール392を生成することを示す。図3Bでは、時間マップセット176、278、及び378は、図1Bに関して上述したものと同様の方法で、時間ネットワーク160の第1の時間畳み込み層によって生成される。時間マップセット382は、図1Bに関して上述したものと同様の方法で、時間ネットワーク160の第2のかつ最後の時間畳み込み層によって生成される。図3Cは、出力層190が、ベースコールの第3の反復中に生成された最終時間マップセット382を処理し、中央シーケンシングサイクル5に対するベースコール392を生成することを示す。

30

【0106】

ひとたび圧縮空間マップセットが、所与のシーケンシングサイクルに対して生成されると、その圧縮空間マップセットは、任意の後続のシーケンシングサイクルをベースコールするために再度利用され得る。図4Aは、中央シーケンシングサイクル16をベースコールするためのベースコールの第14の反復を示す。図4Bは、シーケンシングサイクル1〜29に対して以前に生成された圧縮空間マップを使用して、中央シーケンシングサイクル16をベースコールするための最終時間マップセット482を生成することを示す。図4Cは、出力層190が、ベースコールの第14の反復中に生成された最終時間マップセット482を処理し、中央シーケンシングサイクル16のベースコール492を生成することを示す。

40

【0107】

図5Aは、フィルタリングロジック502を使用して、それぞれのシーケンシングサイクル1、2、3、4、及び5に対する圧縮空間マップセット110、120、130、140、及び150をフィルタリングして、ベースコールの第1の反復中に、それぞれのフィルタリングされた圧縮空間マップ510、520、530、540、及び550（信頼性の高いクラスタのみを示す）を生成する一実施態様を示す。上で論じたように、信頼性の低いクラスタデータ128は、空間マップ及び圧縮された空間マップのうちの、信頼性のないクラスタに対応するそれらの部分（例えば、ピクセル）を識別する。そのようなピ

50

クセルは、例えば、信頼できないクラスタの位置座標に基づいて識別することができる。

【0108】

フィルタリングロジック502は、信頼できないクラスタを識別するデータ128を使用して、圧縮空間マップセット110、120、130、140、及び150から、信頼できないクラスタに対応する（すなわち、その強度放射を描く）ピクセルをフィルタリングして取り除く（又は廃棄する若しくは除去する）。いくつかの実施態様では、これは、圧縮空間マップセットから破棄されたピクセルの75%をもたらし、それによって多くの非生産的な量込み込みを防止する。

【0109】

図5Aでは、中央シーケンシングサイクル3をベースコールするために、フィルタリング処理された時間マップセット572、574、及び576（信頼性の高いクラスタのみを示す）が、フィルタリングされた圧縮空間マップ510、520、530、540、及び550から生成される。フィルタリングされた時間マップセット572、574、及び576（信頼性の高いクラスタのみを示す）は、図1Bに関して上述したものと同様の方法で、時間ネットワーク160の第1の時間量込み込み層によって生成される。フィルタリングされた時間マップセット582（信頼性の高いクラスタのみを示す）は、図1Bに関して上述したものと同様の方法で、時間ネットワーク160の第2のかつ最後の時間量込み込み層によって生成される。図5Bは、出力層190が、ベースコールの第1の反復中に生成された、フィルタリングされた最終時間マップセット582を処理し、中央シーケンシングサイクル3のベースコール592を生成することを示す。

【0110】

図6Aは、フィルタリングロジック502を使用して、それぞれのシーケンシングサイクル2、3、4、5、及び6に対する圧縮空間マップセット120、130、140、150、及び230をフィルタリングして、ベースコールの第2の反復中に、それぞれのフィルタリングされた圧縮空間マップ520、530、540、550、及び650（信頼性の高いクラスタのみを示す）を生成する一実施態様を示す。フィルタリングロジック502は、信頼できないクラスタを識別するデータ128を使用して、圧縮空間マップセット120、130、140、150、及び230から、信頼できないクラスタに対応する（すなわち、その強度放射を描く）ピクセルをフィルタリングして取り除く（又は廃棄する若しくは除去する）。

【0111】

図6Aでは、中央シーケンシングサイクル4をベースコールするために、フィルタリング処理された時間マップセット574、576、及び676（信頼性の高いクラスタのみを示す）が、フィルタリングされた圧縮空間マップ520、530、540、550、及び650から生成される。フィルタリングされた時間マップセット574、576、及び676（信頼性の高いクラスタのみを示す）は、図1Bに関して上述したものと同様の方法で、時間ネットワーク160の第1の時間量込み込み層によって生成される。フィルタリングされた時間マップセット682（信頼性の高いクラスタのみを示す）は、図1Bに関して上述したものと同様の方法で、時間ネットワーク160の第2のかつ最後の時間量込み込み層によって生成される。図6Bは、出力層190が、ベースコールの第2の反復中に生成された、フィルタリングされた最終時間マップセット682を処理し、中央シーケンシングサイクル4のベースコール692を生成することを示す。

【0112】

図7Aは、フィルタリングロジック502を使用して、それぞれのシーケンシングサイクル3、4、5、6、及び7に対する圧縮空間マップセット130、140、150、230、及び330をフィルタリングして、ベースコールの第3の反復中に、それぞれのフィルタリングされた圧縮空間マップ530、540、550、650、及び750（信頼性の高いクラスタのみを示す）を生成する一実施態様を示す。フィルタリングロジック502は、信頼できないクラスタを識別するデータ128を使用して、圧縮空間マップセット130、140、150、230、及び330から、信頼できないクラスタに対応する

（すなわち、その強度放射を描く）ピクセルをフィルタリングして取り除く（又は廃棄する若しくは除去する）。

【0113】

図7Aでは、中央シーケンシングサイクル5をベースコールするために、フィルタリング処理された時間マップセット576、676、及び776（信頼性の高いクラスタのみを示す）が、フィルタリングされた圧縮空間マップ530、540、550、650、及び750から生成される。フィルタリングされた時間マップセット576、676、及び776（信頼性の高いクラスタのみを示す）は、図1Bに関して上述したものと同様の方法で、時間ネットワーク160の第1の時間畳み込み層によって生成される。フィルタリングされた時間マップセット782（信頼性の高いクラスタのみを示す）は、図1Bに関して上述したものと同様の方法で、時間ネットワーク160の第2のかつ最後の時間畳み込み層によって生成される。図7Bは、出力層190が、ベースコールの第3の反復中に生成された、フィルタリングされた最終時間マップセット782を処理し、中央シーケンシングサイクル5のベースコール792を生成することを示す。

【0114】

他の実施態様では、圧縮ロジック108は、対応する圧縮時間マップセットを、各々が4つより多くの特徴マップを有するように設定することができる。

【0115】

空間特徴マップに関して上述した圧縮ロジック108は、時間ロジック160によって生成された時間特徴マップの圧縮に同等に適用される。ひとたび生成された圧縮空間特徴マップを、後続のシーケンシングサイクルにおいて再使用することはまた、ひとたび生成された圧縮された時間特徴マップを、その後のシーケンシングサイクルにおいて再使用することにも同等に適用される。

【0116】

いくつかの実施態様では、圧縮された時間特徴マップを再使用することは、その結果、圧縮された時間特徴マップが、処理パイプラインのより後の段階で、圧縮空間特徴マップから生成されるため、圧縮空間特徴マップを再使用するよりも、100倍以上の効率性及びコンピューティングの節約をもたらす。更なる処理エンジン（すなわち、時間ネットワーク160）から中間結果の再目的化することにより、スキップされ得る、先行する処理ステップの数を増加させる。すなわち、圧縮空間特徴マップを再使用することは、空間ネットワーク104を介した、元の画像データの冗長処理を排除するが、時間ネットワーク160を介した、圧縮空間特徴マップを冗長に処理することを含む可能性がある。対照的に、圧縮時間特徴マップを再使用することは、空間ネットワーク104を介した、元の画像データの冗長処理と、時間ネットワーク160を介した、圧縮空間特徴マップの冗長処理と、の両方を排除する。

【0117】

図8Aは、圧縮ロジック108によって、ベースコールの第1の反復中に生成された時間特徴マップのセット172、174、及び176を処理して、圧縮時間特徴マップのそれぞれのセット802、804、及び806を生成する一実施態様を示す。圧縮時間特徴マップセット802、804、及び806は、図1E及び図1Fに関して上述したものと同様の方法で、圧縮ロジック108によって生成される。すなわち、例えば、第1の時間特徴マップセット172が、例えば21個の特徴マップを有する（又はチャンネル若しくは深さが21である）場合、圧縮ロジック108は、対応する圧縮時間特徴マップセット802を、1つ、2つ、3つ、又は4つの特徴マップを有するように設定することができる。圧縮時間特徴マップのセット802、804、及び806は、図1Bに関して上述したものと同様の方法で、時間ネットワーク160の第2の時間畳み込み層によって処理されて、最終圧縮時間特徴マップセット814を生成する。図8Bは、出力層190が、ベースコールの第1の反復中に生成された最終圧縮時間特徴マップセット814を処理し、中央シーケンシングサイクル3のベースコール892を生成することを示す。

【0118】

図9 Aは、第2のベースコール反復において、第1のベースコール反復で生成された圧縮時間マップを再利用する一実施態様を示す。すなわち、圧縮時間マップの第1及び第2のセット804及び806は、第1のベースコール反復のために図8 Aで生成され、現在、図9 A及び図9 Bに示される第2のベースコール反復では再目的化される。

【0119】

圧縮時間マップの第1及び第2のセット804及び806は、時間マップの第1及び第2のセット172及び174から図8 Aで生成されたということに留意されたい。更に、時間マップの第1及び第2のセット172及び174は、図1 Bでは、圧縮空間マップ110、120、130、及び140から生成され、圧縮空間マップ110、120、130、及び140は、図1 Aでは、それぞれ対応する空間マップ106、116、126、及び136から生成され、空間マップ106、116、126、及び136は、図1 Aでは、それぞれ対応する画像パッチ102、112、122、及び132から生成されたものである。

【0120】

重なり合う時間マップ172、174、及び176を、冗長に生成する図1 B、図2 B、及び図3 Bとは異なり、図9 Aでは、重なり合う時間マップ174及び176（点線及び薄い文字で図9 Aに示される）は、図8 A（第1のベースコール反復）から図9 A（第2のベースコール反復）まで冗長に生成されることはない。これが起こるのは、圧縮ロジック108が時間ネットワーク160に組み込まれて、第1のベースコール反復において、圧縮時間マップの第1及び第2のセット804及び806を生成し、これらが第2のベースコール反復において、重なり合う時間マップ174及び176を置き換えるためである。圧縮時間マップは、メモリ（例えば、オンチップDRAM、オンチップSRAM、又はBRAM、オフチップDRAM）に格納することができる。

【0121】

図9 Aはまた、圧縮ロジック108によって、重なり合わない時間マップ278（すなわち、第1のベースコール反復と第2のベースコール反復との間で重なり合わない）を処理し、圧縮ロジックマップ906を生成することを示す。圧縮時間マップセット804、806、及び906は、図1 Bに関して上述したものと同様の方法で、時間ネットワーク160の第2の時間畳み込み層によって処理されて、最終圧縮時間特徴マップセット914を生成する。図9 Bは、出力層190が、ベースコールの第2の反復中に生成された最終圧縮時間特徴マップセット914を処理し、中央シーケンシングサイクル4のベースコール992を生成することを示す。

【0122】

重なり合う時間マップ174、176、及び278を、冗長に生成する図1 B、図2 B、及び図3 Bとは異なり、図10 Aでは、重なり合う時間マップ176及び278（点線及び薄い文字で図10 Aに示される）は、図9 A（第2のベースコール反復）から図10 A（第3のベースコール反復）まで冗長に生成されることはない。これが起こるのは、圧縮ロジック108が時間ネットワーク160に組み込まれて、第1及び第2のベースコール反復において、圧縮時間マップの第1及び第2のセット806及び906を生成し、これらが第3のベースコール反復において、重なり合う時間マップ176及び278を置き換えるためである。圧縮時間マップは、メモリ（例えば、オンチップDRAM、オンチップSRAM、又はBRAM、オフチップDRAM）に格納することができる。

【0123】

図10 Aはまた、圧縮ロジック108によって、重なり合わない時間マップ378（すなわち、第2のベースコール反復と第3のベースコール反復との間で重なり合わない）を処理し、圧縮ロジックマップ1006を生成することを示す。圧縮時間マップセット806、906、及び1006は、図1 Bに関して上述したものと同様の方法で、時間ネットワーク160の第2の時間畳み込み層によって処理されて、最終圧縮時間特徴マップセット1014を生成する。図10 Bは、出力層190が、ベースコールの第3の反復中に生成された最終圧縮時間特徴マップセット1014を処理し、中央シーケンシングサイクル

5のベースコール1092を生成することを示す。

【0124】

図11Aは、圧縮ロジック108によって、ベースコールの第1の反復中に生成されたフィルタリングされた時間特徴マップのセット572、574、及び576を処理して、フィルタリングされた圧縮時間特徴マップのそれぞれのセット1102、1104、及び1306を生成する一実施態様を示す。フィルタリングされた圧縮時間特徴マップセット1102、1104、及び1106（信頼性の高いクラスタのみを示す）は、図1E及び図1Fに関して上述したものと同様の方法で、圧縮ロジック108によって生成される。すなわち、例えば、第1のフィルタリングされた時間特徴マップセット572が、例えば21個の特徴マップを有する（又はチャンネル又は深さが21である）場合、圧縮ロジック108は、対応する、フィルタリングされた圧縮時間特徴マップセット1102を、1つ、2つ、3つ、又は4つの特徴マップを有するように構成することができる。フィルタリングされた圧縮時間特徴マップのセット1102、1104、及び1106は、図1Bに関して上述したものと同様の方法で、フィルタリングされた時間ネットワーク160の第2のフィルタリングされた時間畳み込み層によって処理されて、フィルタリングされた最終圧縮時間特徴マップセット1114を生成する。図8Bは、出力層190が、ベースコールの第1の反復中に生成された、フィルタリングされた最終圧縮時間特徴マップセット1114を処理し、中央シーケンシングサイクル3のベースコール1192を生成することを示す。

10

【0125】

他の実施態様では、圧縮ロジック108は、対応する圧縮特徴マップセットを、各々が4つより多くの特徴マップを有するように設定することができる。

20

【0126】

図12Aは、第2のベースコール反復において、第1のベースコール反復で生成された、フィルタリングされた圧縮時間マップを再利用する一実施態様を示す。すなわち、フィルタリングされた圧縮時間マップの第1及び第2のセット1104及び1106は、第1のベースコール反復のために図11Aで生成され、現在、図12A及び図2Bに示される第2のベースコール反復では再目的化される。

【0127】

フィルタリングされた圧縮時間マップの第1及び第2のセット1104及び1106は、フィルタリングされた時間マップの第1及び第2のセット572及び574から図11Aで生成されたということに留意されたい。更に、フィルタリングされた時間マップの第1及び第2のセット572及び574は、図5Aでは、フィルタリングされた圧縮空間マップ510、520、530及び540から生成され、フィルタリングされた圧縮空間マップ510、520、530及び540は、図5Aでは、それぞれ対応する圧縮空間マップ110、120、130、及び140から生成され、圧縮空間マップ110、120、130、及び140は、図1Aでは、それぞれ対応する空間マップ106、116、126、及び136から生成され、空間マップ106、116、126、及び136は、図1Aでは、それぞれ対応する画像パッチ102、112、122、及び132から生成されたものである。

30

【0128】

フィルタリングされた、重なり合う時間マップ572、574、及び576を、冗長に生成する図5A、6A、及び7Aとは異なり、図12Aでは、フィルタリングされた、重なり合う時間マップ574及び576（点線及び薄い文字で図12Aに示される）は、図11A（第1のベースコール反復）から図12A（第2のベースコール反復）へは、重複して生成されることはない。これが起こるのは、圧縮ロジック108がフィルタリングされた時間ネットワーク160に組み込まれて、第1のベースコール反復において、フィルタリングされた圧縮時間マップの第1及び第2のセット1104及び1106を生成し、これらが第2のベースコール反復において、フィルタリングされた、重なり合う時間マップ574及び576を置き換えるためである。フィルタリングされた圧縮時間マップは、

40

50

メモリ（例えば、サイト、オンチップ D R M、オンチップ S R A M、又は B R A M、オフチップ D R A M）に格納することができる。

【0129】

図12Aはまた、圧縮ロジック108によって、フィルタリングされた重なり合わない時間マップ676（すなわち、第1のベースコール反復と第2のベースコール反復との間で重なり合わない）を処理し、フィルタリングされた圧縮ロジックマップ1206（信頼性の高いクラスタのみを示す）を生成することを示す。フィルタリングされた圧縮時間マップのセット1104、1106、及び1206は、図1Bに関して上述したものと同様の方法で、時間ネットワーク160の第2のフィルタリングされた時間畳み込み層によって処理されて、フィルタリングされた最終圧縮時間特徴マップセット1214を生成する。図12Bは、出力層190が、ベースコールの第2の反復中に生成された、フィルタリングされた最終圧縮時間特徴マップセット1214を処理し、中央シーケンシングサイクル4のベースコール1292を生成することを示す。

10

【0130】

フィルタリングされた、重なり合う時間マップ574、576、及び676を、冗長に生成する図5A、6A、及び7Aとは異なり、図13Aでは、フィルタリングされた、重なり合う時間マップ576及び676（点線及び薄い文字で図13Aに示される）は、図12A（第2のベースコール反復）から図132A（第3のベースコール反復）へは、重複して生成されることはない。これが起こるのは、圧縮ロジック108がフィルタリングされた時間ネットワーク160に組み込まれて、第1及び第2のベースコール反復において、フィルタリングされた圧縮時間マップの第1及び第2のセット1106及び1206を生成し、これらが第3のベースコール反復において、フィルタリングされた、重なり合う時間マップ576及び676を置き換えるためである。フィルタリングされた圧縮時間マップは、メモリ（例えば、サイト、オンチップ D R M、オンチップ S R A M、又は B R A M、オフチップ D R A M）に格納することができる。

20

【0131】

図13Aはまた、圧縮ロジック108によって、フィルタリングされた重なり合わない時間マップ776（すなわち、第2のベースコール反復と第3のベースコール反復との間で重なり合わない）を処理し、フィルタリングされた圧縮ロジックマップ1306（信頼性の高いクラスタのみを示す）を生成することを示す。フィルタリングされた圧縮時間マップのセット1106、1206、及び1306は、図1Bに関して上述したものと同様の方法で、時間ネットワーク160の第2のフィルタリングされた時間畳み込み層によって処理されて、フィルタリングされた最終圧縮時間特徴マップセット1314を生成する。図13Bは、出力層190が、ベースコールの第3の反復中に生成された、フィルタリングされた最終圧縮時間特徴マップセット1314を処理し、中央シーケンシングサイクル5のベースコール1392を生成することを示す。

30

【0132】

図14は、ニューラルネットワークベースのベースコーラー100の第1の例示的なアーキテクチャを示す。図示の実施態様では、ニューラルネットワークベースのベースコーラー100は、空間ネットワーク104、圧縮ネットワーク108、及び時間ネットワーク160を含む。空間ネットワーク104は、7つの空間畳み込み層を含む。圧縮ネットワーク108は、圧縮層を備える。時間ネットワーク160は、2つの時間畳み込み層を含む。

40

【0133】

7つの空間畳み込み層の各々は、同じ数の畳み込みフィルタを有することができるか、又は異なる数の畳み込みフィルタを有することができる。第1の空間畳み込み層は、S1という数のフィルタを有することができ、S1は、例えば、7、14、21、64、128、又は254であり得る。第2の空間畳み込み層は、S2という数のフィルタを有することができ、S2は、例えば、7、14、21、64、128、又は254であり得る。第3の空間畳み込み層は、S3という数のフィルタを有することができ、S3は、例えば

50

、7、14、21、64、128、又は254であり得る。第4の空間畳み込み層は、S4という数のフィルタを有することができ、S4は、例えば、7、14、21、64、128、又は254であり得る。第5の空間畳み込み層は、S5という数のフィルタを有することができ、S5は、例えば、7、14、21、64、128、又は254であり得る。第6の空間畳み込み層は、S6という数のフィルタを有することができ、S6は、例えば、7、14、21、64、128、又は254であり得る。第7の空間畳み込み層は、S7という数のフィルタを有することができ、S7は、例えば、7、14、21、64、128、又は254であり得る。

【0134】

圧縮層は、C1という数のフィルタを有することができ、C1は、例えば、1、2、3、4、又はそれ以上であり得る。

【0135】

2つの時間畳み込み層の各々は、同じ数の畳み込みフィルタを有することができるか、又は異なる数の畳み込みフィルタを有することができる。第1の時間畳み込み層は、T1という数のフィルタを有することができ、T1は、例えば、7、14、21、64、128、又は254であり得る。第2の時間畳み込み層は、T2という数のフィルタを有することができ、T2は、例えば、7、14、21、64、128、又は254であり得る。図14はまた、ニューラルネットワークベースのベースコーラー100の層の各々によって生成された特徴マップ1412を示す。

【0136】

図15は、ニューラルネットワークベースのベースコーラー100の第2の例示的なアーキテクチャを示す。図15は、ニューラルネットワークベースのベースコーラー100の一部としてのフィルタリングロジック502を示す。他の実施態様では、フィルタリングロジック502は、ニューラルネットワークベースのベースコーラー100の一部ではない。圧縮特徴マップC1は、 $P1 \times P2$ の空間次元数を有する。フィルタリングロジック502は、圧縮特徴マップC1内の、信頼できないクラスタに対応するピクセルをフィルタリングして除外し、 $P3 \times P4$ の空間次元数を有する、フィルタリングされた圧縮特徴マップF1を生成する。フィルタリングされた圧縮特徴マップF1は、信頼できるクラスタのみを示す。一実施態様では、フィルタリングロジック502は、圧縮特徴マップC1内のピクセルの75%を廃棄するため、P3は、P1の25%であり、P4は、P2の25%である。図15はまた、ニューラルネットワークベースのベースコーラー100の層の各々によって生成された特徴マップ1512を示す。

【0137】

図16は、ニューラルネットワークベースのベースコーラー100の第3の例示的なアーキテクチャを示す。図16は、圧縮ネットワーク108が、空間ネットワーク104の出力及び時間ネットワーク160の出力を圧縮するために使用されることを示す。図16はまた、ニューラルネットワークベースのベースコーラー100の層の各々によって生成された特徴マップ1612を示す。

【0138】

図17は、ニューラルネットワークベースのベースコーラー100の第4の例示的なアーキテクチャを示す。図17は、フィルタリングロジック502が、空間ネットワーク104の圧縮出力に適用されて、時間ネットワーク160からの圧縮及びフィルタリングされた時間出力を生成することを示す。図17はまた、ニューラルネットワークベースのベースコーラー100の層の各々によって生成された特徴マップ1712を示す。

【0139】

図18は入力データ内のチャンネルの数に応じて、圧縮層108内の畳み込みフィルタの数（又は大きな数）を構成する、フィルタ構成ロジック1804の一実施態様を示す。これにより、圧縮特徴マップは、入力データの損失のない表現であることを可能にする。いくつかの実施態様では、入力データは、その後のシーケンシングサイクルで再利用するために、対応する圧縮表現を用いてメモリに上書きすることができる。

【0140】

一実施態様では、サイクルごとの入力の各々（例えば、1つの画像チャネルのみ）に、1つのチャネル1812のみを含む入力データの場合、フィルタ構成ロジック1804は、圧縮層108を、シーケンシングサイクルごとに1つの圧縮特徴マップ1818のみを生成する1つの畳み込みフィルタ1816のみを有するように構成する。別の実施態様では、サイクルごとの入力の各々（例えば、青色及び緑色レーザに対応するシーケンス画像内の青色及び緑色の画像チャネルのような2つの画像チャネル）に、2つのチャネル1822を含む入力データの場合、フィルタ構成ロジック1804は、圧縮層108を、シーケンシングサイクルごとに2つの圧縮特徴マップ1828を生成する2つの畳み込みフィルタ1826を用いて構成する。更に別の実施態様では、サイクルごとの入力の各々（例えば、3つの画像チャネルのみ）に、3つのチャネル1832のみを含む入力データの場合、フィルタ構成ロジック1804は、圧縮層108を、シーケンシングサイクルごとに3つの圧縮特徴マップ1838を生成する3つの畳み込みフィルタ1836を有するように構成する。なお更なる実施態様では、サイクルごとの入力の各々（例えば、ヌクレオチドA、C、T、及びGに対応する、シーケンス画像内のA、C、T、及びGチャネルのような4つの画像チャネル）に、4つのチャネル1842を含む入力データの場合、フィルタ構成ロジック1804は、圧縮層108を、シーケンシングサイクルごとに4つの圧縮特徴マップ1848を生成する4つの畳み込みフィルタ1846を用いて構成する。他の実施態様では、圧縮ロジック108は、対応する圧縮特徴マップセットを、各々が4つより多くの特徴マップを有し、それゆえに圧縮層108に対して4つより多くのフィルタを選択するように構成することができる。

【0141】

図19A及び図19Bは、シーケンシングシステム1900Aの一実施態様を示す。シーケンシングシステム1900Aは、構成可能プロセッサ1946を含む。構成可能プロセッサ1946は、本明細書に開示されるベースコール技術を実施する。シーケンシングシステムは、「シーケンサ」とも称される。

【0142】

シーケンシングシステム1900Aは、生物学的物質又は化学物質のうちの少なくとも1つに関連する任意の情報又はデータを得ることができる。いくつかの実施態様では、シーケンシングシステム1900Aは、ベンチトップデバイス又はデスクトップコンピュータと同様であり得るワークステーションである。例えば、所望の反応を実施するためのシステム及び構成要素の大部分（又は全て）は、共通のハウジング1902内にあってもよい。

【0143】

特定の実施態様では、シーケンシングシステム1900Aは、De Novoシーケンシング、全ゲノム又は標的ゲノム領域の再配列、及びメタゲノミクスを含むがこれらに限定されない、様々な用途のために構成された核酸シーケンシングシステムである。シーケンサはまた、DNA又はRNA分析に使用されてもよい。いくつかの実施態様では、シーケンシングシステム1900Aはまた、バイオセンサ内に反応部位を生成するように構成されてもよい。例えば、シーケンシングシステム1900Aは、サンプルを受容し、サンプル由来のクロノウイルス増幅核酸の表面結合クラスタを生成するように構成され得る。各クラスタは、バイオセンサ内の反応部位を構成するか、又はその一部であってもよい。

【0144】

例示的なシーケンシングシステム1900Aは、バイオセンサ1912と相互作用して、バイオセンサ1912内で所望の反応を行うように構成されたシステム受け部又はインターフェース1910を含んでもよい。図19Aに関して以下の説明では、バイオセンサ1912はシステム受け部1910内に装填される。しかしながら、バイオセンサ1912を含むカートリッジは、システム受け部1910に挿入されてもよく、一部の状態では、カートリッジは一時的又は永久的に除去され得ることが理解される。上述のように、カートリッジは、とりわけ、流体制御及び流体貯蔵構成要素を含んでもよい。

【0145】

特定の実施態様では、シーケンシングシステム1900Aは、バイオセンサ1912内で多数の平行反応を行うように構成されている。バイオセンサ1912は、所望の反応が生じ得る1つ以上の反応部位を含む。反応部位は、例えば、バイオセンサの固体表面に固定化されてもよく、又はバイオセンサの対応する反応チャンバ内に位置するビーズ（又は他の可動基材）に固定化されてもよい。反応部位は、例えば、クローン的に増幅された核酸のクラスタを含むことができる。バイオセンサ1912は、固体撮像装置（例えば、CCD又はCMOSイメージャ）及びそれに取り付けられたフローセルを含んでもよい。フローセルは、シーケンシングシステム1900Aから溶液を受容し、溶液を反応部位に向けて方向付ける1つ以上の流路を含んでもよい。任意選択的に、バイオセンサ1912は、熱エネルギーを流路の内外に伝達するための熱要素と係合するように構成することができる。

10

【0146】

シーケンシングシステム1900Aは、相互に相互作用して、生物学的又は化学的分析のための所定の方法又はアッセイプロトコルを実行する、様々な構成要素、アセンブリ、及びシステム（又はサブシステム）を含んでもよい。例えば、シーケンシングシステム1900Aは、システムコントローラ1906を含み、そのシステムコントローラ1906は、シーケンシングシステム1900Aの様々な構成要素、アセンブリ、及びサブシステムと、またバイオセンサ1912とも通信し得る。例えば、システム受け部1910に加えて、シーケンシングシステム1900Aはまた、シーケンシングシステム1900A及びバイオセンサ1912の流体ネットワーク全体にわたる流体の流れを制御するための流体制御システム1908と、バイオアッセイシステムによって使用され得る全ての流体（例えば、気体又は液体）を保持するように構成されている流体貯蔵システム1914と、流体ネットワーク、流体貯蔵システム1914、及び／又はバイオセンサ1912内の流体の温度を調整し得る温度制御システム1904と、バイオセンサ1912を照明するように構成されている照明システム1916と、を備え得る。上述のように、バイオセンサ1912を有するカートリッジがシステム容器1910内に装填される場合、カートリッジはまた、流体制御及び流体貯蔵構成要素を含んでもよい。

20

【0147】

また、シーケンシングシステム1900Aは、ユーザーと対話するユーザーインターフェース1918を含んでもよい。例えば、ユーザーインターフェース1918は、ユーザーから情報を表示又は要求するディスプレイ1920と、ユーザー入力を受け取るためのユーザー入力デバイス1922とを含むことができる。いくつかの実施態様では、ディスプレイ1920及びユーザー入力デバイス1922は、同一のデバイスである。例えば、ユーザーインターフェース1918は、個々のタッチの存在を検出し、またディスプレイ上のタッチの場所を識別するように構成されたタッチ感知ディスプレイを含んでもよい。しかしながら、マウス、タッチパッド、キーボード、キーパッド、ハンドヘルドスキャナー、音声認識システム、動き認識システムなどの他のユーザー入力デバイス1922が使用されてもよい。以下でより詳細に説明するように、シーケンシングシステム1900Aは、所望の反応を実施するために、バイオセンサ1912（例えば、カートリッジの形態）を含む様々な構成要素と通信してもよい。シーケンシングシステム1900Aはまた、バイオセンサから得られたデータを分析して、ユーザーに所望の情報を提供するように構成されてもよい。

30

40

【0148】

システムコントローラ1906は、マイクロコントローラ、縮小命令セットコンピュータ（Reduced Instruction Set Computer、RISC）、特定用途向け集積回路（Application Specific Integrated Circuit、ASIC）、フィールドプログラマブルゲートアレイ（Field Programmable Gate Array、FPGA）、粗粒度再構成可能構造（Coarse-Grained Reconfigurable Architecture、CGRA）、論理回路、及び本明細書に記載される機能を実行することができる任意の他の回路又はプロセッサと、を備える。上記

50

の実施例は、例示的なものに過ぎず、したがって、システムコントローラという用語の定義及び／又は意味を制限することを意図するものではない。例示的实施態様では、システムコントローラ1906は、検出データを取得し分析する少なくとも1つのために、1つ以上の記憶要素、メモリ、又はモジュール内に記憶された命令のセットを実行する。検出データは、ピクセル信号の複数のシーケンスを含むことができ、それにより、数百万個のセンサ（又はピクセル）のそれぞれからのピクセル信号のシーケンスを、多くのベースサイクルにわたって検出することができる。記憶要素は、シーケンシングシステム1900A内の情報源又は物理メモリ要素の形態であってもよい。

【0149】

命令セットは、本明細書に記載される様々な実施態様の方法及びプロセスなどの特定の動作を実行するようにシーケンシングシステム1900A又はバイオセンサ1912に指示する様々なコマンドを含んでもよい。命令のセットは、有形の非一時的コンピュータ可読媒体又は媒体の一部を形成し得るソフトウェアプログラムの形態であってもよい。本明細書で使用する時、用語「ソフトウェア」及び「ファームウェア」は互換可能であり、RAMメモリ、ROMメモリ、EPROMメモリ、EEPROMメモリ、及び不揮発性RAM（NVRAM）メモリを含むコンピュータによって実行されるメモリに記憶された任意のコンピュータプログラムを含む。上記メモリタイプは、例示的なものに過ぎず、したがって、コンピュータプログラムの記憶に使用可能なメモリの種類に限定されない。

【0150】

ソフトウェアは、システムソフトウェア又はアプリケーションソフトウェアなどの様々な形態であってもよい。更に、ソフトウェアは、別個のプログラムの集合、又はより大きいプログラム内のプログラムモジュール若しくはプログラムモジュールの一部分の形態であってもよい。ソフトウェアはまた、オブジェクト指向プログラミングの形態のモジュール式プログラミングを含んでもよい。検出データを取得した後、検出データは、ユーザー入力に応じて処理されたシーケンシングシステム1900Aによって自動的に処理されてもよく、又は別の処理マシン（例えば、通信リンクを介したリモート要求）によって行われる要求に応じて処理されてもよい。図示の別の実施態様では、システムコントローラ1906は分析モジュール1944を含む。他の別の実施態様では、システムコントローラ1906は分析モジュール1944を含まず、代わりに分析モジュール1944へのアクセスを有する（例えば、分析モジュール1944は、クラウド上で別個にホスティングされ得る）。

【0151】

システムコントローラ1906は、通信リンクを介して、バイオセンサ1912及びシーケンシングシステム1900Aの他の構成要素に接続されてもよい。システムコントローラ1906はまた、オフサイトシステム又はサーバに通信可能に接続されてもよい。通信リンクは、配線、コード、又は無線であってもよい。システムコントローラ1906は、ユーザーインターフェース1918及びユーザー入力デバイス1922からユーザー入力又はコマンドを受信してもよい。

【0152】

流体制御システム1908は、流体ネットワークを含み、流体ネットワークを通る1つ以上の流体の流れを方向付けるように構成されている。流体ネットワークは、バイオセンサ1912及び流体貯蔵システム1914と流体連通していてもよい。例えば、流体貯蔵システム1914から流体を選択し、制御された方法でバイオセンサ1912に向けてもよく、又は流体は、バイオセンサ1912から引き出され、例えば、流体貯蔵システム1914内の廃棄物リザーバに向けられてもよい。図示されていないが、流体制御システム1908は、流体ネットワーク内の流体の流量又は圧力を検出する流量センサを含んでもよい。センサは、システムコントローラ1906と通信してもよい。

【0153】

温度制御システム1904は、流体ネットワーク、流体貯蔵システム1914及び／又はバイオセンサ1912の異なる領域における流体の温度を調節するように構成されてい

10

20

30

40

50

る。例えば、温度制御システム1904は、バイオセンサ1912と相互作用し、バイオセンサ1912内の反応部位に沿って流れる流体の温度を制御する熱循環器を含んでもよい。温度制御システム1904はまた、シーケンシングシステム1900A又はバイオセンサ1912の固体要素又は構成要素の温度を調節してもよい。図示されていないが、温度制御システム1904は、流体又は他の構成要素の温度を検出するためのセンサを含んでもよい。センサは、システムコントローラ1906と通信してもよい。

【0154】

流体貯蔵システム1914は、バイオセンサ1912と流体連通しており、所望の反応を行うために使用される様々な反応成分又は反応物質を貯蔵してもよい。流体貯蔵システム1914はまた、流体ネットワーク及びバイオセンサ1912を洗浄又はクリーニングし、反応物質を希釈するための流体を貯蔵してもよい。例えば、流体貯蔵システム1914は、試料、試薬、酵素、他の生体分子、緩衝液、水性、及び非極性溶液などを保存するための様々なリザーバを含んでもよい。更に、流体貯蔵システム1914はまた、バイオセンサ1912から廃棄物を受容するための廃棄物リザーバを含んでもよい。カートリッジを含む実施態様形態では、カートリッジは、流体貯蔵システム、流体制御システム、又は温度制御システムのうちの1つ以上を含み得る。したがって、これらのシステムに関する本明細書に記載される構成要素のうちの1つ以上は、カートリッジハウジング内に収容され得る。例えば、カートリッジは、サンプル、試薬、酵素、他の生体分子、緩衝液、水性及び非極性溶液、廃棄物などを保存するための様々なリザーバを有し得る。したがって、流体貯蔵システム、流体制御システム、又は温度制御システムのうちの1つ以上は、カートリッジ又は他のバイオセンサを介してバイオアッセイシステムと取り外し可能に係合され得る。

【0155】

照明システム1916は、バイオセンサを照明するための光源（例えば、1つ以上の発光ダイオード（Light-Emitting Diode、LED））及び複数の光学構成要素を含んでもよい。光源の例としては、レーザ、アークランプ、LED、又はレーザダイオードを挙げることができる。光学構成要素は、例えば、反射器、偏光板、ビームスプリッタ、コリメータ、レンズ、フィルタ、ウェッジ、プリズム、鏡、検出器などであってもよい。照明システムを使用する実施態様では、照明システム1916は、励起光を反応部位に向けるように構成されてもよい。一例として、蛍光団は、緑色の光の波長によって励起されてもよく、そのため、励起光の波長は約1932nmであり得る。一実施態様では、照明システム1916は、バイオセンサ1912の表面の表面法線に平行な照明を生成するように構成されている。別の実施態様では、照明システム1916は、バイオセンサ1912の表面の表面法線に対してオフアングルである照明を生成するように構成されている。更に別の実施態様では、照明システム1916は、いくつかの平行照明及びある程度のオフアングル照明を含む複数の角度を有する照明を生成するように構成されている。

【0156】

システム受け部又はインターフェース1910は、機械的、電氣的、及び流体的な方法のうちの少なくとも1つにおいてバイオセンサ1912と係合するように構成される。システム受け部1910は、バイオセンサ1912を所望の配向に保持して、バイオセンサ1912を通る流体の流れを容易にすることができる。システム受け部1910はまた、バイオセンサ1912と係合するように構成された電気接点を含んでもよく、それにより、シーケンシングシステム1900Aは、バイオセンサ1912と通信してもよく、及び／又はバイオセンサ1912に電力を供給することができる。更に、システム受け部1910は、バイオセンサ1912と係合するように構成された流体ポート（例えば、ノズル）を含んでもよい。いくつかの実施態様では、バイオセンサ1912は、機械的に、電氣的に、また流体方式で、システム受け部1910に取り外し可能に連結される。

【0157】

加えて、シーケンシングシステム1900Aは、他のシステム若しくはネットワークと、又は他のバイオアッセイシステム1900Aと、遠隔で通信してもよい。バイオアッセ

10

20

30

40

50

イシステム 1900A によって得られた検出データは、リモートデータベースに記憶されてもよい。

【0158】

図 19B は、図 19A のシステムで使用することができるシステムコントローラ 1906 のブロック図である。一実施態様では、システムコントローラ 1906 は、互いに通信することができる 1 つ以上のプロセッサ又はモジュールを含む。プロセッサ又はモジュールのそれぞれは、特定のプロセスを実行するためのアルゴリズム（例えば、有形及び／又は非一時的コンピュータ可読記憶媒体上に記憶された命令）又はサブアルゴリズムを含んでもよい。システムコントローラ 1906 は、モジュールの集合として概念的に例示されるが、専用ハードウェアボード、DSP、プロセッサなどの任意の組み合わせを利用して実装されてもよい。あるいは、システムコントローラ 1906 は、単一のプロセッサ又は複数のプロセッサを備えた既製の PC を使用して実装されてもよく、機能動作はプロセッサ間に分散される。更なる選択肢として、以下に記載されるモジュールは、特定のモジュール式機能が専用ハードウェアを利用して実施されるハイブリッド構成を利用して実装されてもよく、残りのモジュール式機能は、既製の PC などを利用して実施される。モジュールはまた、処理ユニット内のソフトウェアモジュールとして実装されてもよい。

【0159】

動作中、通信ポート 1950 は、バイオセンサ 1912（図 19A）及び／又はサブシステム 1908、1914、1904（図 19A）に情報（例えば、コマンド）を送信してもよく、あるいはそれらから情報（例えば、データ）を受信してもよい。あるいくつかの実施態様では、通信ポート 1950 は、ピクセル信号の複数のシーケンスを出力することができる。通信リンク 1934 は、ユーザーインターフェース 1918 からユーザー入力を受信してもよく（図 19A）、ユーザーインターフェース 1918 にデータ又は情報を送信してもよい。バイオセンサ 1912 又はサブシステム 1908、1914、1904 からのデータは、バイオアッセイセッション中に、システムコントローラ 1906 によってリアルタイムで処理されてもよい。追加的に又は代替的に、データは、バイオアッセイセッション中にシステムメモリ内に一時的に記憶され、リアルタイムよりも遅いタイミングで、又はオフライン操作で処理されてもよい。

【0160】

図 19B に示すように、システムコントローラ 1906 は、中央演算処理装置（Central Processing Unit、CPU）1952 と共に、主制御モジュール 1924 と通信する複数のモジュール 1926～1948 を含んでもよい。主制御モジュール 1924 は、ユーザーインターフェース 1918 と通信してもよい（図 19A）。モジュール 1926～1948 は、主制御モジュール 1924 と直接通信するものとして示されているが、モジュール 1926～1948 はまた、互いに、ユーザーインターフェース 1918 と、及びバイオセンサ 1912 と直接通信してもよい。また、モジュール 1926～1948 は、他のモジュールを介して主制御モジュール 1924 と通信してもよい。

【0161】

複数のモジュール 1926～1948 は、サブシステム 1908、1914、1904 及び 1916 とそれぞれ通信するシステムモジュール 1928～1932、1926 を含む。流体制御モジュール 1928 は、流体ネットワークを通る 1 つ以上の流体の流れを制御するために、流体制御システム 1908 と通信して、流体ネットワークの弁及び流量センサを制御してもよい。流体貯蔵モジュール 1930 は、流体の貯蔵量が低い場合、又は廃棄物リザーバが容量一杯にある又はその容量一杯近くにあるときにユーザーに通知することができる。流体貯蔵モジュール 1930 はまた、流体が所望の温度で貯蔵され得るように、温度制御モジュール 1932 と通信してもよい。照明モジュール 1926 は、所望の反応（例えば、結合事象）が生じた後など、プロトコル中に指定された時間で反応部位を照明するために、照明システム 1916 と通信してもよい。いくつかの実施態様では、照明モジュール 1926 は、照明システム 1916 と通信して、指定された角度で反応部位を照明することができる。

【0162】

複数のモジュール1926～1948はまた、バイオセンサ1912と通信する装置モジュール1936と、バイオセンサ1912に関連する識別情報を判定する識別モジュール1938とを含んでもよい。装置モジュール1936は、例えば、システム受け部1910と通信して、バイオセンサがシーケンシングシステム1900Aとの電氣的及び流体的接続を確立したことを確認することができる。識別モジュール1938は、バイオセンサ1912を識別する信号を受信してもよい。識別モジュール1938は、バイオセンサ1912の識別情報を使用して、他の情報をユーザーに提供してもよい。例えば、識別モジュール1938は、ロット番号、製造日、又はバイオセンサ1912で実行することが推奨されるプロトコルを決定し、その後表示してもよい。

10

【0163】

複数のモジュール1926～1948はまた、バイオセンサ1912から信号データ（例えば、画像データ）を受信及び分析する分析モジュール1944（信号処理モジュール又は信号プロセッサとも呼ばれる）も含む。分析モジュール1944は、検出／画像データを記憶するためのメモリ（例えば、RAM又はフラッシュ）を含む。検出データは、ピクセル信号の複数のシーケンスを含むことができ、それにより、数百万個のセンサ（又はピクセル）のそれぞれからのピクセル信号のシーケンスを、多くのベースコールサイクルにわたって検出することができる。信号データは、その後の分析のために記憶されてもよく、又はユーザーインターフェース1918に送信されて、所望の情報をユーザーに表示することができる。いくつかの実施態様では、信号データは、分析モジュール1944が信号データを受信する前に、固体撮像素子（例えば、CMOS画像センサ）によって処理され得る。

20

【0164】

分析モジュール1944は、複数のシーケンシングサイクルのそれぞれにおいて、光検出器から画像データを取得するように構成される。画像データは、光検出器によって検出される発光信号から導出され、ニューラルネットワークベースのベースコーラー100を介して、複数のシーケンシングサイクルの各々に対して画像データを処理し、複数のシーケンシングサイクルの各々で分析物のうちの少なくともいくつかに対するベースコールを生成する。光検出器は、1つ以上のオーバーヘッドカメラ（例えば、バイオセンサ1912上のクラスタの画像を上から撮影するIlluminaのGAIIxのCCDカメラ）の一部であってもよく、又はバイオセンサ1912自体の一部（例えば、バイオセンサ1912上のクラスタの下にあり、底部からのクラスタの画像を取るIlluminaのiSeqのCMOS画像センサ）であってもよい。

30

【0165】

光検出器の出力は、それぞれクラスタの強度放射及びそれらの周囲背景を示すシーケンス画像である。シーケンス画像は、シーケンシング中にシーケンスにヌクレオチドを組み込む結果として生成される強度放射を示す。強度放射は、関連する分析物及びそれらの周囲の背景からのものである。シーケンス画像は、メモリ1948に記憶される。

【0166】

プロトコルモジュール1940及び1942は、メイン制御モジュール1924と通信して、所定のアッセイプロトコルを実施する際にサブシステム1908、1914及び1904の動作を制御する。プロトコルモジュール1940及び1942は、所定のプロトコルに従って特定の動作を実行するようにシーケンシングシステム1900Aに命令するための命令セットを含み得る。図示のように、プロトコルモジュールは、シーケンシングごとの合成プロセスを実行するための様々なコマンドを発行するように構成された、合成によるシーケンシング（Sequencing-By-Synthesis、SBS）モジュール1940であってもよい。SBSにおいて、核酸テンプレートに沿った核酸プライマーの伸長を監視して、テンプレート中のヌクレオチドシーケンスを決定する。基礎をなす化学プロセスは、重合（例えば、ポリメラーゼ酵素により触媒される）又はライゲーション（例えば、リガーゼ酵素により触媒される）であり得る。特定のポリマー系SBSの実施態様では、プラ

40

50

イマーに付加されるヌクレオチドの順序及び種類の検出を使用してテンプレートのシーケンスを決定することができるように、蛍光標識ヌクレオチドをテンプレート依存様式でプライマー（それによってプライマーを伸長させる）に添加する。例えば、第1のSBSサイクルを開始するために、1つ以上の標識されたヌクレオチド、DNAポリメラーゼなどを、核酸テンプレートのアレイを収容するフローセル内に／それを介して送達することができる。核酸テンプレートは、対応する反応部位に位置してもよい。プライマー伸長が、組み込まれる標識ヌクレオチドを、撮像事象を通して検出することができる、これらの反応部位が検出され得る。撮像イベントの間、照明システム1916は、反応部位に励起光を提供することができる。任意選択的に、ヌクレオチドは、ヌクレオチドがプライマーに付加されると、更なるプライマー伸長を終結する可逆的終結特性を更に含むことができる。例えば、可逆的ターミネーター部分を有するヌクレオチド類似体をプライマーに付加して、デブロッキング作用因子が送達されてその部分を除去するまで、その後の伸長が起こらないようにすることができる。したがって、可逆終端を使用する別の実施態様では、フローセルにデブロッキング試薬を、（検出前又は検出後に）送達するために、コマンドを与えることができる。1つ以上のコマンドは、様々な送達工程間の洗浄（複数可）をもたらすために与えられ得る。次いで、サイクルをn回繰り返してプライマーをn個のヌクレオチド分だけ延長し、それによって長さnのシーケンスを検出することができる。例示的なシーケンシング技術は、例えば、Bentley et al., Nature 419 6:193-199 (20019)、国際特許出願公開第04/0119497号、米国特許第7,0197,026号、国際特許出願公開第91/066719号、国際特許出願公開第07/123744号、米国特許第7,329,492号、米国特許第7,211,414号、米国特許第7,3119,019号、米国特許第7,4019,2191号、及び米国特許出願公開第20019/01470190192号に記載されており、これらの各々は、参照により本明細書に組み込まれる。

【0167】

SBSサイクルのヌクレオチド送達工程では、単一の種類のヌクレオチドのいずれかを一度に送達することができ、又は複数の異なるヌクレオチドタイプ（例えば、A、C、T、及びG）を送達することができる。一度に単一の種類のヌクレオチドのみが存在するヌクレオチド送達構成では、異なるヌクレオチドは、個別化された送達に固有の時間的分離に基づいて区別することができるため、異なるヌクレオチドは別個の標識を有する必要はない。したがって、シーケンシング方法又は装置は、単一の色検出を使用することができる。例えば、励起源は、単一の波長又は単一の波長範囲の励起のみを提供する必要がある。ある時点で、送達がフローセル内に存在する複数の異なるヌクレオチドをもたらすヌクレオチド送達構成では、異なるヌクレオチドタイプを組み込む部位は、混合物中のそれぞれのヌクレオチドタイプに結合された異なる蛍光標識に基づいて区別することができる。例えば、4つの異なる蛍光団のうちの1つをそれぞれ有する4つの異なるヌクレオチドを使用することができる。一実施態様では、4つの異なるフルオロフォアは、スペクトルの4つの異なる領域における励起を使用して区別することができる。例えば、4つの異なる励起放射線源を使用することができる。あるいは、4つ未満の異なる励起源を使用することができるが、単一源からの励起放射線の光学的濾過を使用して、フローセルにおいて異なる励起放射線の範囲を生成することができる。

【0168】

いくつかの実施態様では、4つ未満の異なる色を、4つの異なるヌクレオチドを有する混合物中で検出することができる。例えば、ヌクレオチドの対は、同じ波長で検出することができるが、対のうちの1つのメンバーに対する強度の差に基づいて、又は、対の他の部材について検出された信号と比較して明らかなシグナルを出現又は消失させる、対の1つのメンバーへの変化（例えば、化学修飾、光化学修飾、又は物理的改質を行うことを介して）に基づいて区別され得る。4個未満の色の検出を使用して4個の異なるヌクレオチドを区別するための例示的な装置及び方法が、例えば、米国特許出願第61/19319,294号及び同第61/619,19719号に記載されており、それらの全体が参照

により本明細書に組み込まれる。2012年9月21日に出願された米国特許出願第13／624,200号もまた、その全体が参照により組み込まれる。

【0169】

複数のプロトコルモジュールはまた、バイオセンサ1912内の製品を増幅するための流体制御システム1908及び温度制御システム1904に命令を発行するように構成された試料調製（又は生成）モジュール1942を含んでもよい。例えば、バイオセンサ1912は、シーケンシングシステム1900Aに係合されてもよい。増幅モジュール1942は、バイオセンサ1912内の反応チャンバに必要な増幅成分を送達するために、流体制御システム1908に命令を発行することができる。他の実施態様では、反応部位は、テンプレートDNA及び／又はプライマーなどの増幅のためのいくつかの成分を既に含有していてもよい。増幅成分を反応チャンバに送達した後、増幅モジュール1942は、既知の増幅プロトコルに従って異なる温度段階を通して温度制御システム1904にサイクルするように指示し得る。いくつかの実施態様では、増幅及び／又はヌクレオチドの取り込みは、等温的に実施される。

10

【0170】

SBSモジュール1940は、クローン性アンプリコンのクラスタがフローセルのチャネル内の局所領域上に形成されるブリッジPCRを実行するコマンドを発行することができる。ブリッジPCRを介してアンプリコンを生成した後、アンプリコンを「線形化」して、一本鎖テンプレートDNAを作製してもよく、sstDNA及びシーケンシングプライマーは、関心領域に隣接する普遍シーケンスにハイブリダイズされてもよい。例えば、合成方法による可逆的ターミネーター系シーケンシングを、上記のように又は以下のように使用することができる。

20

【0171】

各ベースコール又はシーケンシングサイクルは、例えば、修飾DNAポリメラーゼ及び4種類のヌクレオチドの混合物を使用することによって達成することができる単一の塩基によってsstDNAを延長することができる。異なる種類のヌクレオチドは、固有の蛍光標識を有することができ、各ヌクレオチドは、各サイクルにおいて単一塩基の組み込みのみが生じることを可能にする可逆的ターミネーターを更に有し得る。単一の塩基をsstDNAに添加した後、励起光が反応部位に入射し、蛍光発光を検出することができる。検出後、蛍光標識及びターミネーターは、sstDNAから化学的に切断され得る。別の同様のベースコール又はシーケンシングサイクルが、後に続いてよい。そのようなシーケンシングプロトコルでは、SBSモジュール1940は、バイオセンサ1912を通る試薬及び酵素溶液の流れを方向付けるように流体制御システム1908に指示することができる。本明細書に記載される装置及び方法と共に利用することができる例示的な可逆性ターミネーターベースのSBS方法は、米国特許出願公開第2007／01667019（A1）号、米国特許出願公開第2006／01196＊3901（A1）号、米国特許第7,0197,026号、米国特許出願公開第2006／0240439（A1）号、米国特許出願公開第2006／021914714709（A1）号、国際特許出願公開第019／0619514号、米国特許出願公開第20019／014700900（A1）号、国際特許出願公開第06／019B199号及び国際特許出願公開第07／014702191号（それぞれ参照によりその全体が本明細書に組み込まれる）に記載されている。可逆的ターミネーター系SBSの例示的な試薬は、米国特許第7,1941,444号、米国特許第7,0197,026号、米国特許第7,414,14716号、米国特許第7,427,673号、米国特許第7,1966,1937号、米国特許第7,1992,4319号、国際特許出願公開第07／14193193619号に記載され、これらの文書のそれぞれ全体が、参照により本明細書に組み込まれる。

30

40

【0172】

いくつかの実施態様では、増幅及びSBSモジュールは、単一のアッセイプロトコルで動作してもよく、例えば、テンプレート核酸は増幅され、続いて同じカートリッジ内でシーケンスされる。

50

【0173】

シーケンシングシステム1900Aはまた、ユーザーがアッセイプロトコルを再構成することを可能にし得る。例えば、シーケンシングシステム1900Aは、決定されたプロトコルを修正するために、ユーザーインターフェース1918を通じてユーザーにオプションを提供することができる。例えば、バイオセンサ1912が増幅のために使用されると判定された場合、シーケンシングシステム1900Aは、アニーリングサイクルの温度を要求し得る。更に、シーケンシングシステム1900Aは、選択されたアッセイプロトコルに対して一般的に許容されないユーザー入力をユーザーが提供した場合に、ユーザーに警告を発行し得る。

【0174】

実施態様形態では、バイオセンサ1912は、数百万個のセンサ（又はピクセル）を含み、それらのそれぞれは、後続のベースコールサイクルにわたって複数のピクセル信号のシーケンスを生成する。分析モジュール1944は、センサのアレイ上のセンサの行方向及び／又は列方向の位置に従って、ピクセル信号の複数のシーケンスを検出し、それらを対応するセンサ（又はピクセル）に属させる。

【0175】

図19Cは、ベースコールセンサ出力などのシーケンシングシステム1900Aからのセンサデータの分析のためのシステムの簡略ブロック図である。図19Cの例では、システムは、構成可能プロセッサ1946を含む。構成可能プロセッサ1946は、中央演算処理装置（CPU）1952（すなわち、ホストプロセッサ）によって実行される実行時プログラム／ロジック1980と協調して、ベースコーラー（例えば、ニューラルネットワークベースのベースコーラー100）を実行することができる。シーケンシングシステム1900Aは、バイオセンサ1912及びフローセルを含む。フローセルは、遺伝物質のクラスタが、遺伝物質中の塩基を識別するためにクラスタ内の反応を引き起こすために使用される一連の分析物フローに曝露される1つ以上のタイルを含み得る。センサは、タイルデータを提供するために、フローセルの各タイルにおけるシーケンスの各サイクルの反応を検知する。遺伝的シーケンシングはデータ集約的操作であり、このデータ集約的操作は、ベースコールセンサデータを、ベースコール動作中に感知された各遺伝物質群のベースコールのシーケンスに変換する。

【0176】

本実施例のシステムは、ベースコール動作を調整するための実行時プログラム／ロジック1980を実行するCPU1952と、タイルデータのアレイのシーケンスを記憶するメモリ1948Bと、ベースコール動作によって生成されたベースコールリードと、ベースコール動作で使用される他の情報とを記憶する。また、この図では、システムは、構成ファイル（又は複数の構成ファイル）、例えば、フィールドプログラマブルゲートアレイ（FPGA）ビットファイル、及び構成可能プロセッサ1946を構成及び再構成し、ニューラルネットワークを実行するために使用されるモデルパラメータを記憶するためのメモリ1948Aを含む。シーケンシングシステム1900Aは、構成可能プロセッサを構成するためのプログラムを含むことができ、いくつかの実施態様では、ニューラルネットワークを実行する再構成可能なプロセッサを含み得る。

【0177】

シーケンシングシステム1900Aは、バス1989によって構成可能プロセッサ1946に結合される。バス1989は、一例では、PCI-SIG規格（PCI Special Interest Group）によって現在維持及び開発されているPCIe規格（Peripheral Component Interconnect Express）と互換性のあるバス技術などの高スループット技術を使用して実装することができる。また、この例では、メモリ1948Aは、バス1993によって構成可能プロセッサ1946に結合される。メモリ1948Aは、構成可能プロセッサ1946を有する回路基板上に配置されたオンボードメモリであってもよい。メモリ1948Aは、ベースコール動作で使用される作業データの構成可能プロセッサ1946による高速アクセスに使用される。バス1993はまた、PCIe規格と互換性のあるバス

10

20

30

40

50

技術などの高スループット技術を使用して実装することもできる。

【0178】

フィールドプログラマブルゲートアレイ（FPGA）、粗粒度再構成可能アレイCGRA、並びに他の構成可能かつ再構成可能なデバイスを含む構成可能なプロセッサは、コンピュータプログラムを実行する汎用プロセッサを使用して達成され得るよりも、より効率的に又はより高速に様々な機能を実装するように構成することができる。構成可能なプロセッサの構成は、時にはビットストリーム又はビットファイルと称される構成ファイルを生成するために機能的な説明を編集することと、構成ファイルをプロセッサ上の構成可能要素に配布することと、を含む。構成ファイルは、データフローパターンを設定するように回路を構成することにより、分散メモリ及び他のオンチップメモリリソースの使用、ルックアップテーブルコンテンツ、構成可能な論理ブロックの動作、及び構成可能な論理ブロックの動作、及び構成可能なアレイの構成可能な相互接続及び他の要素のような構成可能な実行ユニットとを含む。構成ファイルがフィールド内で変更され得る場合、ロードされた構成ファイルを変更することによって構成ファイルを変更することができる場合に再構成可能である。例えば、構成ファイルは、揮発性SRAM要素内に、不揮発性読み書きメモリ素子内に記憶されてもよく、構成可能又は再構成可能なプロセッサ上の構成可能要素のアレイ間に分散されたものであってもよい。様々な市販の構成可能なプロセッサは、本明細書に記載されるようなベースコール動作において使用するのに好適である。例としては、Googleのテンソル処理ユニット（TPU）（商標）、GX4 Rackmount Series（商標）、GX9 Rackmount Series（商標）、NVIDIA DGX-1（商標）、Microsoft's Stratix V FPGA（商標）、GraphcoreのIntelligent Processor Unit（IPU）（商標）、QualcommのZeroth Platform（商標）（Snapdragon processors（商標）、NVIDIA Volta（商標）、NVIDIAのドライブPX（商標）、NVIDIAのJETSON TX1/TX2 MODULE（商標）、Intel's Nirvana™、Movidius VPU（商標）、Fujitsu DPI（商標）、アームDynamicIQ（商標）、IBM TrueNorth（商標）、Lambda GPU Server with Testa V100s（商標）、Xilinx Alveo（商標）U200、Xilinx Alveo（商標）U2190、Xilinx Alveo（商標）U280、Intel/Altera Stratix（商標）GX2800、Intel/Altera Stratix（商標）GX2800、及びIntel Stratix（商標）GX10M、が含まれる。いくつかの実施例では、ホストCPUは、構成可能プロセッサと同じ集積回路上に実装することができる。

【0179】

本明細書に記載の実施態様は、構成可能なプロセッサ1946を使用して、ニューラルネットワークベースのベースコーラー100を実装する。構成可能プロセッサ1946の構成ファイルは、高レベルの記述言語HDL又はレジスタ転送レベルRTL言語仕様を使用して実行される論理機能を指定することによって実装することができる。本明細書は、選択された構成可能プロセッサが構成ファイルを生成するように設計されたリソースを使用してコンパイルすることができる。構成可能なプロセッサではない場合がある特定用途向け集積回路の設計を生成する目的で、同じ又は類似の仕様をコンパイルすることができる。

【0180】

したがって、本明細書に記載される全ての実施態様における構成可能プロセッサ構成可能プロセッサ1946の代替例は、特定用途向けASIC又は専用集積回路又は集積回路のセットを含む構成されたプロセッサを含み、又はシステムオンチップSOCデバイス、又は本明細書に記載されるようなニューラルネットワークベースのベースコール動作を実行するように構成された、システムオンチップSOCデバイス、又はグラフィック処理ユニット（GPU）プロセッサ若しくは粗粒度再構成可能アーキテクチャ（Coarse-Grained）

d Reconfigurable Architecture、C G R A) プロセッサである。

【0181】

一般に、ニューラルネットワークの動作を実行するように構成された、本明細書に記載の構成可能なプロセッサ及び構成されたプロセッサは、本明細書ではニューラルネットワークプロセッサと称される。

【0182】

構成可能プロセッサ1946は、この例では、CPU1952によって実行されるプログラムを使用して、又は構成可能要素1991のアレイを構成する他のソースによってロードされた構成ファイルによって構成される（例えば、構成論理ブロック（Configuration Logic Block、CLB）、例えばルックアップテーブル（Look Up Table、LUT）、フリップフロップ、演算処理ユニット（PMU）、及び計算メモリユニット（Compute Memory Unit、CMU）、構成可能なI/Oブロック、プログラマブル相互接続）を計算して、ベースコール機能を実行する。この例では、構成は、バス1989及び1993に結合され、ベースコール動作で使用される要素間でデータ及び制御パラメータを分散する機能を実行するデータフローロジック1997を含む。

【0183】

また、構成可能なプロセッサ1946は、データフローロジック1997を用いて構成されて、ニューラルネットワークベースのベースコーラー100を実行する。ロジック1997は、マルチサイクル実行クラスタ（例えば、1979）を含み、この実施例では、実行クラスタXを介した実行クラスタ1を含む。多重サイクル実行クラスタの数は、動作の所望のスループットを伴うトレードオフ、及び構成可能プロセッサ1946上の利用可能なリソースに従って選択することができる。

【0184】

多重サイクル実行クラスタは、構成可能なプロセッサ1946上の構成可能な相互接続及びメモリリソースを使用して実装されるデータ流路1999によってデータフローロジック1997に結合される。また、マルチサイクル実行クラスタは、構成可能な相互接続及びメモリリソースを使用して、例えば構成可能プロセッサ1946上で実施される制御経路1995によってデータフローロジック1997に結合されている。それは、利用可能な実行クラスタを示す制御信号を提供し、ニューラルネットワークベースのベースコーラー100の動作の実行のための入力ユニットを利用可能な実行クラスタに提供する準備ができており、ニューラルネットワークベースのベースコーラー100の学習されたパラメータを提供する準備ができており、ベースコール分類データの出力パッチ、並びにニューラルネットワークベースのベースコーラー100の実行に使用される他の制御データを提供する準備ができています。

【0185】

構成可能プロセッサ1946は、学習されたパラメータを使用してニューラルネットワークベースのベースコーラー100のランを実行して、ベースコール動作の検知サイクルに関する分類データを生成するように構成されている。ニューラルネットワークベースのベースコーラー100のランを実行して、ベースコール動作の被験者検知サイクルの分類データを生成する。ニューラルネットワークベースのベースコーラー100のランは、N個の検知サイクルのそれぞれの検知サイクルからのタイルデータのアレイの数Nを含むシーケンスで動作し、N個の検知サイクルは、本明細書に記載される例での時間シーケンスにおける動作ごとの1つの塩基位置に対する異なるベースコール動作のためのセンサデータを提供する。任意選択的に、N個の検知サイクルのうちのいくつかは、実行される特定のニューラルネットワークモデルに従って必要に応じて、シーケンスから出ることができる。数Nは、1を超える任意の数であり得る。本明細書に記載されるいくつかの実施例では、N個の検知サイクルの検知サイクルは、被験者の検知サイクルに先行する少なくとも1つの検知サイクル、及び被験者サイクルの後の少なくとも1回の検知サイクルについての検知サイクルのセットを表す。本明細書では、数Nが5以上の整数である、実施例が記載される。

【0186】

データフローロジック1997は、N個のアレイの空間的に位置合わせされたパッチのタイルデータを含む所与の動作のための入力ユニットを使用して、ニューラルネットワークベースのベースコーラー100のランのために、メモリ1948Aから構成可能プロセッサ1946に、タイルデータ及びモデルパラメータの少なくともいくつかの学習されたパラメータを移動させるように構成されている。入力ユニットは、1回のDMA動作におけるダイレクトメモリアクセス動作によって、又は、配備されたニューラルネットワークの実行と協調して、利用可能なタイムスロットの間に移動するより小さいユニット内で移動させることができる。

【0187】

本明細書に記載される感知サイクルのタイルデータは、1つ以上の特徴を有するセンサデータのアレイを含むことができる。例えば、センサデータは、DNA、RNA、又は他の遺伝物質の遺伝的シーケンスにおける塩基位置で4塩基のうちの1つを識別するために分析される2つの画像を含むことができる。タイルデータはまた、画像及びセンサに関するメタデータを含むことができる。例えば、ベースコール動作の実施態様では、タイルデータは、タイル上の遺伝物質群の中心からのセンサデータのアレイ内の各ピクセルの距離を示す中心情報からの距離などの、クラスタとの画像の位置合わせに関する情報を含むことができる。

【0188】

以下に記載されるように、ニューラルネットワークベースのベースコーラー100の実行中に、タイルデータはまた、ニューラルネットワークベースのベースコーラー100の実行中に生成されるデータも含み得る。それは、ニューラルネットワークベースのベースコーラー100のラン中に再計算されるのではなく再使用することができる中間データと称される。例えば、ニューラルネットワークベースのベースコーラー100の実行中に、データフローロジック1997は、タイルデータのアレイの所与のパッチのセンサデータの代わりに、中間データをメモリ1948Aに書き込むことができる。このような実施態様は、以下により詳細に記載される。

【0189】

図示されているように、ベースコール動作の検知サイクルからタイルのセンサデータを含むタイルデータを記憶する実行時プログラム／ロジック1980によってアクセス可能なメモリ（例えば、1948A）を含む、ベースコールセンサ出力の分析のためのシステムが説明される。また、システムは、メモリへのアクセスを有する構成可能プロセッサ1946などのニューラルネットワークプロセッサを含む。ニューラルネットワークプロセッサは、学習されたパラメータを使用してニューラルネットワークのランを実行して、検知サイクルのための分類データを生成するように構成される。本明細書に記載されるように、ニューラルネットワークの動作は、被験者サイクルを含むN個の感知サイクルのそれぞれの感知サイクルからタイルデータのN個のアレイのシーケンスで動作して、被験者サイクルの分類データを生成する。データフローロジック1997は、N個の感知サイクルのそれぞれの感知サイクルからのN個のアレイの空間的に整合されたパッチのデータを含む入力ユニットを使用して、ニューラルネットワークの実行のために、メモリからニューラルネットワークプロセッサにタイルデータ及び学習されたパラメータを移動させるために提供される。

【0190】

また、ニューラルネットワークプロセッサがメモリへのアクセスを有し、複数の実行クラスタを含み、ニューラルネットワークを実行するように構成された複数の実行クラスタ内の実行クラスタを含むシステムも説明される。データフローロジック1997は、メモリへのアクセス、及び複数の実行クラスタ内のクラスタを実行して、複数の実行クラスタ内の利用可能な実行クラスタにタイルデータの入力ユニットを提供し、入力ユニットは、それぞれの感知サイクルからタイルデータのアレイの空間的に整列されたパッチの数Nを含む、入力ユニットと、被験者検知サイクルを含み、N個の空間的に整合されたパッチを

10

20

30

40

50

ニューラルネットワークに適用して、被験者検知サイクルの空間的に整合されたパッチの分類データの出力パッチを生成させるように、実行クラスタに、ニューラルネットワークにN個の空間的に位置合わせされたパッチを適用させることと、を含み、Nは1より大きい。

【0191】

図20Aは、ホストプロセッサによって実行される実行時プログラム（例えば、実行時ロジック1980）の機能を含む、ベースコール動作の態様を示す簡略図である。この図では、フローセルからの画像センサの出力は、ライン2000上で画像処理スレッド2001に提供され、画像処理スレッド2001は、個々のタイルのセンサデータのアレイ内の位置合わせ及び配置、及び画像の再サンプリングなどの画像上のプロセスを実行することができ、フローセル内の各タイルのタイルクラスタマスクを計算するプロセスによって使用することができ、フローセルの対応するタイル上の遺伝子材料のクラスタに対応するセンサデータのアレイ内のピクセルを識別するプロセスによって使用することができる。画像処理スレッド2001の出力は、CPU内のディスパッチロジック2003に、ライン2002を介して提供され、それは、高速バス2004上又は高速バス2006上のデータキャッシュ2005（例えば、SSD記憶装置）に、ベースコール動作の状態に従って、図19Cの構成可能プロセッサ1946などのニューラルネットワークプロセッサハードウェア2007にタイルデータのアレイを転送する。処理され、変換された画像は、以前に使用されたサイクルを検知するために、データキャッシュ2005上に記憶され得る。ハードウェア2007は、ニューラルネットワークによって出力された分類データをディスパッチロジック2003に返し、ディスパッチロジック2003は、情報をデータキャッシュ2005に、又はライン2008上でスレッド2009に渡し、分類データを使用してベースコール及び品質スコア計算を実行し、ベースコール読み取りのための標準フォーマットでデータを配置することができる。ベースコール及び品質スコア計算を実行するスレッド2009の出力は、ベースコールリードを集約するスレッド2011に、ライン2010上に提供され、データ圧縮などの他の動作を実行し、結果として得られるベースコール出力を顧客による利用のために指定された宛先に書き込む。

【0192】

いくつかの実施態様では、ホストは、ニューラルネットワークを支持するハードウェア2007の出力の最終処理を実行する、スレッド（図示せず）を含むことができる。例えば、ハードウェア2007は、マルチクラスタニューラルネットワークの最終層から分類データの出力を提供することができる。ホストプロセッサは、ベースコール及び品質スコアスレッド2002によって使用されるデータを設定するために、分類データを超えて、ソフトマックス関数などの出力起動機能を実行することができる。また、ホストプロセッサは、ハードウェア2007への入力前のタイルデータのバッチ正規化などの入力動作（図示せず）を実行することができる。

【0193】

図20Bは、図19Cの構成などの構成可能プロセッサ1946の構成の簡略図である。図20Bでは、構成可能プロセッサ1946は、複数の高速PCIeインターフェースを有するFPGAを含む。FPGAは、図19Cを参照して説明されるデータフローロジック1997を含むラッパー2090を用いて構成されている。ラッパー2090は、CPU通信リンク2077を介して、CPU内の実行時プログラムとのインターフェース及び調整を管理し、DRAM通信リンク2097を介してオンボードDRAM2099（例えば、メモリ1448A）との通信を管理する。ラッパー2090内のデータフローロジック1997は、数Nサイクルについて、オンボードDRAM2099上のタイルデータのアレイを横断することによって取得されたパッチデータをクラスタ2085に提供し、クラスタ2085からプロセスデータ2087を取得して、オンボードDRAM2099に配信する。ラッパー2090はまた、タイルデータの入力アレイ、及び分類データの出力パッチの両方について、オンボードDRAM2099とホストメモリとの間のデータの転送を管理する。ラッパーは、ライン2083上のパッチデータを、割り当てられたクラ

スタ2085に転送する。ラッパーは、オンボードDRAM2099から取得されたクラスタ2085にライン2081の重み及びバイアスなどの学習されたパラメータを提供する。ラッパーは、CPU通信リンク2077を介してホスト上の実行時プログラムから提供されるか、又はそれに応答して生成されるクラスタ2085に、ライン2079上の構成及び制御データを提供する。クラスタはまた、ホストからの制御信号と協働して使用されて、空間的に整列したパッチデータを提供し、クラスタ2085のリソースを使用して、パッチデータを介して多重サイクルニューラルネットワークをパッチデータの上で実行するために、ホストからの制御信号と協働して使用されるラッパー2090に、ライン2089上の状態信号を提供することができる。

【0194】

上述のように、タイルデータの複数のパッチのうちの対応するパッチ上で実行するように構成されたラッパー2090によって管理される単一の構成可能なプロセッサ上に複数のクラスタが存在し得る。各クラスタは、本明細書に記載される複数の感知サイクルのタイルデータを使用して、被験者検知サイクルにおけるベースコールの分類データを提供するように構成することができる。

【0195】

システムの例では、フィルタ重み及びバイアスのようなカーネルデータを含むモデルデータをホストCPUから構成可能プロセッサに送信することができ、その結果、モデルは、サイクル数の関数として更新され得る。ベースコール動作は、代表的な例では、数百の感知サイクルの順序で含むことができる。ベースコール動作は、いくつかの実施態様では、ペアリングされた端部読み取りを含むことができる。例えば、モデル学習されたパラメータは、20サイクル（又は他の数のサイクル）ごとに、又は特定のシステム及びニューラルネットワークモデルに実装される更新パターンに従って更新されてもよい。タイル上の遺伝的クラスタ内の所与のストリングのためのシーケンスが、ストリングの第1の端部から下方に（又は上方に）延在する第1の部分と、ストリングの第2の端部から上方に（又は下方に）延在する第2の部分とを含む、ペアリングされた端部リードを含むいくつかの実施態様では、学習されたパラメータは、第1の部分から第2の部分への遷移で更新され得る。

【0196】

いくつかの実施例では、タイルのための感知データの複数サイクルの画像データは、CPUからラッパー2090に送信され得る。ラッパー2090は、任意選択的に、感知データの一部の前処理及び変換を行い、その情報をオンボードDRAM2099に書き込むことができる。各感知サイクルの入力タイルデータは、タイル当たり4000×3000ピクセル／タイル以上を含むセンサデータのアレイを含むことができ、2つの特徴はタイルの2つの画像の色を表し、1ピクセル当たり1つ又は2つのバイトを含むセンサデータのアレイを含むことができる。数Nが、多重サイクルニューラルネットワークの各動作において使用される3回の検知サイクルである実施態様用では、多重サイクルニューラルネットワークの各動作のためのタイルデータのアレイは、タイル当たり数百メガバイトのオーダーで消費することができる。システムのいくつかの実施態様では、タイルデータはまた、タイルごとに1回記憶されたクラスタセンターからの距離（distance-from-cluster center、DFC）データのアレイ、又はセンサデータ及びタイルに関する他のタイプのメタデータも含む。

【0197】

動作中、多重サイクルクラスタが利用可能である場合、ラッパーは、パッチをクラスタに割り当てる。ラッパーはタイルの横断面にタイルデータの次のパッチをフェッチし、適切な制御及び構成情報と共に割り当てられたクラスタに送信する。クラスタは、構成可能プロセッサ上の十分なメモリを用いて構成されて、パッチを含むデータのパッチを、定位位置に処理されているいくつかのシステム内で複数サイクルから保持するのに十分なメモリを有するように構成することができ、様々な実施態様では、ピンポンバッファ技術又はラスタ走査技術を使用して処理される。

10

20

30

40

50

【0198】

割り当てられたクラスタが、現在のパッチのニューラルネットワークのその動作を完了し、出力パッチを生成すると、それはラッパーに信号を送る。ラッパーは、割り当てられたクラスタから出力パッチを読み出すか、あるいは割り当てられたクラスタは、データをラッパーにプッシュする。次いで、ラッパーは、DRAM2099内の処理されたタイルのための出力パッチを組み立てる。タイル全体の処理が完了し、データの出力パッチがDRAMに転送されると、ラッパーは、処理された出力アレイを、特定のフォーマットでホスト／CPUに返送する。いくつかの実施態様では、オンボードDRAM2099は、ラッパー2090内のメモリ管理論理によって管理される。実行時プログラムは、リアルタイム分析を提供するために連続フローで動作する全てのサイクルについての全てのタイルデータのアレイの分析を完了するために、シーケンシング動作を制御することができる。

10

【0199】

図21は、オフチップメモリ（例えば、オフチップのDRAM、ホストRAM、ホスト高帯域幅メモリ（high bandwidth memory、HBM））2116から、第2のベースコール反復の間に利用可能な、第1のベースコール反復中に生成された圧縮空間マップを作成する、開示されたデータフローロジックの別の実施態様を示している。

【0200】

一実施態様では、ホストプロセッサ（例えば、CPU1952）に取り付けられたホストメモリ（例えば、メモリ1948B）は、シーケンシングランが進行するにつれて、シーケンス画像2102の進行を受信するように構成されている。構成可能なプロセッサ（例えば、構成可能なプロセッサ1946）は、処理ユニットのアレイを有する。処理ユニットのアレイ内の処理ユニットは、ニューラルネットワークベースのベースコーラー100を実行して、ベースコール予測を生成するように構成される。データフローロジック1997は、ホストメモリ、ホストプロセッサ、及び構成可能なプロセッサへのアクセスを有する。第1のベースコール反復に対して、データフローロジック1997は、シーケンシングサイクルの第1のウィンドウのシーケンシングサイクル（例えば、図1Aのシーケンシングサイクル1～5）のシーケンス画像を、ホストメモリから構成可能なプロセッサ上にロードする。

20

【0201】

実行時ロジック1980は、構成可能なプロセッサの処理ユニットをサイクルごとにシーケンス画像2102上でニューラルネットワークベースのベースコーラー100の空間ネットワーク104を実行させ、シーケンシングサイクルの第1のウィンドウ内のシーケンシングサイクルの各々について空間特徴マップセット2106を生成させるように構成されている。一実施態様では、実行時ロジック1980は、並列に、シーケンス画像2102からタイリングされたパッチ2104上のニューラルネットワークベースのベースコーラー100の複数の処理クラスタを実行する。複数の処理クラスタは、パッチ2104上の空間ネットワーク104をパッチごと2105に適用する。

30

【0202】

実行時ロジック1980は、構成可能なプロセッサの処理ユニットが、サイクルごとに空間特徴マップセット2106上で、ニューラルネットワークベースのベースコーラー100の圧縮ネットワーク108を実行し、圧縮空間特徴マップセット2107を生成し、更に、圧縮空間特徴マップセット2107を時間ネットワーク160及び出力ネットワーク190を通じて処理し、シーケンシングサイクルの第1のウィンドウ内の1つ以上のシーケンシングサイクルに対してベースコール予測2111を生成するように構成されている。時間ネットワーク160は、時間特徴マップ2108を生成する。出力ネットワーク190は、ベースコール分類スコア2110を生成する（例えば、非正規化ベースワイズスコア）。一実施態様では、圧縮空間特徴マップセット2107は、オフチップメモリ2116上に記憶される。

40

【0203】

一実施態様では、データフローロジック1997は、圧縮空間特徴マップセット210

50

7をホストメモリ2116に移動させ、圧縮空間特徴マップセット2107でシーケンス画像2102の対応するものを上書きするように構成される。他の実施態様では、パッチ2104の対応するものは、圧縮空間特徴マップセット2107によって置き換えられる。
【0204】

シーケンシングサイクルの第2のウィンドウ（図2Aのシーケンシングサイクル2～6）は、シーケンシングサイクルの第1のウィンドウと、1つ以上の重なり合うシーケンシングサイクル（例えば、シーケンシングサイクル2～5）を共有し、かつ少なくとも1つの重なり合わないシーケンシングサイクル（例えば、試験シングサイクル6）を有するが、第2のベースコール反復に対し、かつシーケンシングサイクルの第2のウィンドウに対して、データフローロジック1997は、重なり合うシーケンシングサイクルのために圧縮空間特徴マップセット216と、重なり合わないシーケンシングサイクルのためにシーケンス画像2122（又はパッチ2124）とを、構成可能なプロセッサに、ホストメモリからロードするように構成されている。

10

【0205】

実行時ロジック1980は、構成可能なプロセッサの処理ユニットが、重なり合わないシーケンシングサイクルのシーケンス画像2122に対して空間ネットワーク104を実行し、重なり合わないシーケンシングサイクルのための空間特徴マップセット2126を生成するように構成されている。一実施態様では、複数の処理クラスタは、パッチ2124に空間ネットワーク104をパッチごと2125に適用する。

【0206】

実行時ロジック1980は、構成可能なプロセッサの処理ユニットが、空間特徴マップセット2126上で圧縮ネットワーク108を実行し、重なり合わないシーケンシングサイクルに対して圧縮空間特徴マップセット2127を生成し、重なり合うシーケンシングサイクルに対して、圧縮空間特徴マップセット2126を処理し、かつ重なり合わないシーケンシングサイクルに対して、時間ネットワーク160及び出力ネットワーク190を介して、圧縮空間特徴マップセット2127を処理し、かつシーケンシングサイクルの第2のウィンドウ中の1つ以上のシーケンシングサイクルに対してベースコール予測2131を生成するように構成されている。時間ネットワーク160は、時間特徴マップ2128を生成する。出力ネットワーク190は、ベースコール分類スコア2129を生成する（例えば、非正規化ベースワイズスコア）。一実施態様では、圧縮空間特徴マップセット2127は、オフチップメモリ2116上に記憶される。

20

【0207】

図22は、オンチップメモリ（例えば、オンチップのDRAM、オンチップSRAM、オンチップBRAM、相互接続を介してプロセッサに取り付けられたDRAMのようなプロセッサメモリ）2216から、第2のベースコール反復中に利用可能な、第1のベースコール反復中に生成された圧縮空間マップを作成する、開示されたデータフローロジックの一実施態様を示す。図22では、圧縮空間特徴マップセット2107及び圧縮空間特徴マップセット2127は、オンチップメモリ2216に記憶される。また、図22において、データフローロジック1997は、オンチップメモリ2216から構成可能なプロセッサ上に、重なり合うシーケンシングサイクルの圧縮空間特徴マップセット2126をロードするように構成される。

30

40

【0208】

分割アーキテクチャ

図23は、ニューラルネットワークベースのベースコーラー100の、いわゆる分割アーキテクチャの一実施態様を示す。上述のように、空間畳み込みネットワーク104は、シーケンシングランのシーケンシングサイクル（サイクル $N+2$ 、 $N+1$ 、 N 、 $N-1$ 、 $N-2$ 、）のシリーズに対して、サイクルごとのシーケンス画像セットのウィンドウを、サイクルごとに、空間畳み込み層のそれぞれのシーケンス2301、2302、2303、2304、及び2405を通じて、それぞれのサイクルごとのシーケンス画像セット別々に畳み込むことによって処理し、それぞれのシーケンシングサイクルに対して、それぞ

50

れのサイクルごとの空間特徴マップセットを生成するように構成されている。例えば、空間畳み込み層の5つのシーケンス2301、2302、2303、2304、及び2405の各々は、7つの空間畳み込み層（すなわち、図23の層L1～L7）を有する。

【0209】

空間畳み込み層のそれぞれのシーケンス2301、2302、2303、2304、及び2405は、空間畳み込みフィルタバンクのそれぞれ対応するシーケンスを有する（例えば、空間畳み込み層のシーケンス2301に対して、空間畳み込みフィルタバンク2310、2311、2312、2313、2314、2315、及び2316を含む）。一実施態様では、空間畳み込みフィルタバンクのそれぞれのシーケンスの空間畳み込みフィルタバンクにおける空間畳み込みフィルタの学習された係数（又は重み）は、空間畳み込み層のそれぞれのシーケンス内の空間畳み込み層のシーケンス間で変化する。

10

【0210】

例えば、空間畳み込み層シーケンス2301、2302、2303、2304、及び2405は、異なる学習された係数を有する畳み込みフィルタを用いて構成される。別の例では、対応するレベルの空間畳み込み層内の畳み込みフィルタは、異なる学習された係数（空間畳み込み層の5つのシーケンス2301、2302、2303、2304、及び2405の、それぞれの第3の空間畳み込み層内の畳み込みフィルタバンク2382、2383、2384、2385、及び2312）を有する。

【0211】

時間畳み込みネットワーク160は、サイクルごとの空間特徴マップセットにおけるサイクルごとの空間特徴マップセットのそれぞれの重なり合うグループ（例えばグループ2360、2361、2362）に畳み込むことによって、サイクルごとの空間特徴マップセットをグループごとに処理し、その際、第1の時間畳み込み層2320のそれぞれの時間畳み込みフィルタバンク2321、2322、及び2323を用いて、サイクルごとの空間特徴マップセットのそれぞれの重なり合うグループに対する、それぞれのグループごとの時間特徴マップセットを生成するように構成されている。一実施態様では、それぞれの時間畳み込みフィルタバンク内の時間畳み込みフィルタの学習された係数（又は重み）は、それぞれの時間畳み込みフィルタバンク内の時間畳み込みフィルタバンク2321、2322、及び2323どうしの間で変化する。

20

【0212】

スキップアーキテクチャ

図24Aは、特徴マップ付加を介して、以前の情報を下流に再度注入する、残留（又はスキップ）接続を示す。残留接続は、データ処理フローに沿った情報損失を防止するのに役立つ、過去の出力テンソルを後の出力テンソルに追加することによって、以前の表現をデータの下流の流れに再注入することを含む。残留接続は、任意の大規模な深層学習モデルを悩ませる2つの一般的な問題、勾配の消失と、表現上のボトルネックとに対処するためのものである。

30

【0213】

残留接続は、先行する層の出力を、後続の層に入力として利用可能にすることを含み、ショートカットを連続的なネットワーク内で効果的に作成するというものである。先行する出力は、後のアクティブ化に連結されるのではなく、後のアクティブ化と合計される。なお、この場合には、両方のアクティブ化は、同じサイズであると仮定する。それらが異なるサイズの場合、目標の形状への先行のアクティブ化を再成形するために、線形変換を使用することができる。

40

【0214】

図24Bは、残留ブロック及びスキップ接続の一実施態様を示す。残留ネットワークは、学習精度の劣化を軽減するために、いくつかの残留ユニットを積み重ねる。残留ブロックは、ディープニューラルネットワークにおける勾配の消失に対抗するための特別な追加スキップ接続を使用する。残留ブロックの開始時に、データフローは2つのストリームに分離される：第1のストリームは、ブロックの変化しない入力を担うが、第2は、重み及

50

び非線形性を適用する。ブロックの終わりに、2つのストリームは、要素ごとの合計を使用してマージされる。そのような構成の主な利点は、勾配がネットワークを介してより容易に流れることを可能にすることである。

【0215】

残留ネットワークで構成されている場合、いくつかの実施態様では、ニューラルネットワークベースのベースコーラー100は容易に学習され、画像分類及びオブジェクト検出のために精度が改善され得る。ニューラルネットワークベースのベースコーラー100は、第 l 層の出力を第 $(l+1)$ 層の入力として接続し、これにより、次の層の移行が可能になる： $x_l = H_l(x_{l-1})$ 。残留ブロックは、識別関数 $x_l = H_l(x_{l-1}) + x_{l-1}$ により非線形変換を迂回するスキップ接続を追加する。残留ブロックの利点は、勾配が、後の層から前の層まで、識別関数を直接流れることができることである（例えば、空間的及び時間畳み込み層）。識別関数及び H_l の出力は、合計（追加）によって組み合わせられる。

10

【0216】

図24Cは、空間畳み込み層がスキップ接続で残留ブロックにグループ化される、ニューラルネットワークベースのベースコーラー100の残留アーキテクチャを示す。他の実施態様では、ニューラルネットワークベースのベースコーラー100の時間畳み込み層は、スキップ接続で、残留ブロックにグループ化される。

【0217】

図24Cに示される実施態様では、第2及び第3の空間畳み込み層は、第1の残留ブロック2412にグループ化され、第4及び第5の空間畳み込み層が、第2の残留ブロック2422にグループ化され、第6及び第7の空間畳み込み層が、第3の残留ブロック2432にグループ化される。

20

【0218】

図25Aは、ニューラルネットワークベースのベースコーラー100のバスネットワークの詳細を示す。一実施態様では、バスネットワークの所与の残留ブロック2585は、空間畳み込み層2590及び2592のセットを含む。空間畳み込み層のセット内の第1の空間畳み込み層2590は、所与の残留ブロック2585（例えば、空間ネットワーク104内の第1の空間畳み込み層2590に先行するゼロ空間畳み込み層）の一部ではない前の空間畳み込み層によって生成された、先行する出力2586を、入力として受信する。第1の空間畳み込み層2590は、先行する出力2586を処理し、第1の出力2591を生成する。空間畳み込み層のセット内の、第1の空間畳み込み層2590の後ろに続く第2の空間畳み込み層2592は、第1の出力2591を受信し、第1の出力2591を処理し、第2の出力2593を生成する。一実施態様では、第1の空間畳み込み層2590は、第1の出力2591を生成するReLUのような非線形起動関数を有する。別の実施態様では、第2の空間畳み込み層2592は、非線形活性化関数を欠いている。

30

【0219】

スキップ接続2589は、先行する出力2586を、加算器2594に提供する。加算器2594はまた、第2の空間畳み込み層2592から第2の出力2593を受信する。加算器2594は、先行する出力2586と第2の出力2593とを組み合わせ、合計された出力2595を生成する。合計された出力2595は、ReLUのような非線形起動を通して更に処理されて、最終合計出力2587を生成する。次いで、いくつかの実施態様では、最終合計出力2587を後続の残留ブロックに入力として供給する。いくつかの実施態様では、先行する出力2586は、第2の出力2593と次元適合性であるように修正される。例えば、先行する出力2586内の特徴マップのエッジはトリミングされて、第2の出力2593内の特徴マップと同じ空間次元数を有する特徴マップを生成する。

40

【0220】

図25Bは、開示されたバスネットワークの例示的な動作を示す。一実施態様では、バスネットワークは、空間畳み込み層のそれぞれのシーケンス内の空間畳み込み層間でバス（例えば、2516、2526、2536、2602、2604、2702、及び271

50

2) を形成するように構成される。バスは、特定のシーケンシングサイクルに対する空間畳み込み層の特定のシーケンス内の2つ以上の空間畳み込み層によって生成されたそれぞれのサイクルの空間特徴マップセットを組み合わせて、組み合わせられた、サイクルごとの空間特徴マップセットを形成し、その組み合わせられた、サイクルごとの空間特徴マップセットを、空間畳み込み層の上記の特定のシーケンス内の別の空間畳み込み層への入力として提供するように構成されている。

【0221】

例えば、ここで第1の残留ブロック2412を取り上げる。ここで、2つ以上の空間畳み込み層は、第1の空間畳み込み層及び第3の空間畳み込み層を含む。第1の空間畳み込み層は、第1のサイクルごとの空間特徴マップセット2520を生成する。第1の空間畳み込み層は、第1のサイクルごとの空間特徴マップセット2520を、第2の空間畳み込み層への入力として提供する。第2の空間畳み込み層は、第1のサイクルごとの空間特徴マップセット2520を処理し、第2のサイクルごとの空間特徴マップセット2522を生成する。第2の空間畳み込み層は、第3の空間畳み込み層への入力として、第2のサイクルごとの空間特徴マップセット2522を提供する。第3の空間畳み込み層は、第2のサイクルごとの空間特徴マップセット2522を処理し、第3のサイクルごとの空間特徴マップセット2524を生成する。バス（例えば、スキップバス2519）は、第1の空間特徴マップセット2520及び第3のサイクルごとの空間特徴マップセット2524を結合して（例えば、コンバイナ2502によって合計又は連結され）、サイクルごとの空間特徴マップセット2518とするように更に構成されている。次いで、別の空間畳み込み層が、空間畳み込み層の特定のシーケンス内の第3の空間畳み込み層の直後に続く、第4の空間畳み込み層である。第4の空間畳み込み層は、入力として、上記の結合された、サイクルごとの空間特徴マップセット2518を処理する。同じ考え方が、第2及び第3の残留ブロック2422及び2432にも適用されるが、2526及び2536は、スキップバス2516のようなスキップバスであり、それぞれのコンバイナ2512及び2532に、サイクルごとの空間特徴マップセット2528及び2538を生成させるスキップバスである。

【0222】

図25Cは、結合前に、スキップバスによって供給される入力特徴マップが修正され（例えば、トリミングされ）、入力される側の特徴マップが、バスネットワークのコンバイナによって結合される受信側の特徴マップと同じ空間次元を有するようされていることを保証する、次元適合性ロジック2532の一実施態様を示す。

【0223】

図26は、スキップバス2602が第1の空間畳み込み層の出力を第1の残留ブロックの出力と組み合わせる、開示されたバスネットワークの別の例を示す。図26はまた、スキップバス2604が、特徴マップを複数の残留ブロックを横断的に結合させることができる、及び非後続的な層を横断的に結合させて（例えば、層1から層5）、結合された表現を生成することができ、この結合された表現は、別の層（例えば、層6）によって処理され得ることを示す。

【0224】

図27は、開示されたバスネットワークの更に別の例を示し、この例では、複数の後続の層及び／又は非後続の層からの入力及び結合表現の表現（例えば、層1、コンバイナ2502、及びコンバイナ2512からのもの）が、例示的なスキップバス2702、2604、及びコンバイナ2712によって組み合わせられて、別の層（例えば、層6）によって処理され得る結合表現を生成することができることを示す。

【0225】

図28は、スキップバスによって供給される入来側特徴マップを、それらがバスネットワークのコンバイナによって結合される対象の受け入れ側の特徴マップと結合される前に、スケーリング（調整）を行うという、スケーリングロジック2832の一実施態様を示す。スケーリングロジック2832によって使用される値は、例えば、0と1を含む、

0～1の間のいずれかの範囲であり得る。スケーリングロジックは、例えば、入来側特徴マップの強度／大きさ／値（例えば、その特徴量（例えば、浮動小数点の値））を減衰又は増幅するために使用することができる。

【0226】

図29は、時間ネットワーク160の時間的な畳み込み層2902、2912、2922、2932、2942、2952、2962、及び2972の間のスキップ接続の一実施態様を示す。例えば、スキップ接続2922は、第1の時間畳み込み層2902から第3の時間畳み込み層2932に時間特徴マップを供給する。

【0227】

図30は、圧縮ロジック108で構成されたネットワークネットワークベースのベースコーラー100（sqz2ベースコーラー）によるベースコール性能を、圧縮ロジック108を有しないネットワークネットワークベースのベースコーラー100（ニューラルモデルのベースラインとして使用）と、Illuminaの非ニューラルネットワークベースのベースコーラーリアルタイム分析（Real-Time Analysis、RTA）ソフトウェア（従来型画像処理モデルのベースライン）とに対して、比較するグラフである。図30のチャートに見られるように、sqz2ベースコーラー（紫色フィットライン）は、RTAベースコーラー（黒フィットライン）よりも低いベースコールエラー率（Y軸上の「エラー%」）を有し、また、圧縮ロジック108（赤色及びシアンフィットライン）なしのネットワークネットワークベースのベースコーラー100の2つの事例よりも低い率を有する。

【0228】

図31は、開示された圧縮ロジック108の使用によってもたらされるRAM及びDRAMの使用の節約を示す。

【0229】

図32は、分割及びスキップアーキテクチャ（split_res）で構成されたネットワークネットワークベースのベースコーラー100によるベースコール性能を、RTAベースコーラーに対して、及び分割及びスキップアーキテクチャを伴わない（ディスティルド）ネットワークネットワークベースのベースコーラー100の別のバージョンに対して比較するグラフである。図32のチャートに見られるように、split_resベースコーラー（オレンジフィットライン）は、RTAベースコーラー（青色フィットライン）よりも低いベースコールエラー率（「Y軸上のエラーカウント」）を有する。

【0230】

本明細書で使用される場合、「ロジック（論理）」（例えば、データフローロジック）は、本明細書に記載の方法ステップを実行するためのコンピュータ使用可能なプログラムコードを備えた非一時的コンピュータ可読記憶媒体を含むコンピュータ製品の形態で実装することができる。「ロジック」は、メモリと、メモリに結合され、例示的な方法ステップを実行するように動作する少なくとも1つのプロセッサとを含む装置の形態で実装することができる。「ロジック」は、本明細書に記載の方法ステップの1つ又は複数を実行するための手段の形で実施することができる。手段は、（i）ハードウェアモジュール（複数可）、（ii）1つ以上のハードウェアプロセッサ上で実行されるソフトウェアモジュール、又は（iii）ハードウェア及びソフトウェアモジュールの組み合わせを含むことができ、（i）～（iii）のいずれかは、本明細書に記載の特定の技術を実装し、ソフトウェアモジュールは、コンピュータ可読記憶媒体（又は複数のそのような媒体）に記憶される。一実施態様では、ロジックは、データ処理機能を実装する。ロジックは、機能を指定するコンピュータプログラムを伴う、汎用、単一コア又はマルチコアプロセッサであってよく、コンピュータプログラムを備えたデジタル信号プロセッサ、構成ファイルで構成可能なFPGAなどのロジック、状態機械などの専用回路、又はこれらの任意の組み合わせであってよい。また、コンピュータプログラム製品は、ロジックのコンピュータプログラム及び構成ファイル部分を具現化することができる。

【0231】

図33は、本明細書に開示される技術を実施するためにシーケンシングシステム190

0 Aによって使用され得るコンピュータシステム3300である。コンピュータシステム3300は、バスサブシステム3355を介して多数の周囲デバイスと通信する、少なくとも1つの中央演算処理装置(CPU)3372を含む。これらの周囲デバイスは、記憶サブシステム3358を含むことができ、この記憶サブシステム3358は、例えば、メモリデバイス及びファイルストレージサブシステム3336、ユーザーインターフェース入力デバイス3338、ユーザーインターフェース出力デバイス3376、及びネットワークインターフェースサブシステム3374を含む。入力及び出力デバイスは、コンピュータシステム3300とのユーザー対話を可能にする。ネットワークインターフェースサブシステム3374は、他のコンピュータシステム内の対応するインターフェースデバイスへのインターフェースを含む外部ネットワークへのインターフェースを提供する。

10

【0232】

一実施態様では、システムコントローラ1906は、記憶サブシステム3310及びユーザーインターフェース入力デバイス3338に通信可能にリンクされている。

【0233】

ユーザーインターフェース入力デバイス3338は、キーボードと、マウス、トラックボール、タッチパッド、又はグラフィックスタブレットなどのポインティングデバイスと、スキャナと、ディスプレイに組み込まれたタッチスクリーンと、音声認識システム及びマイクロフォンなどのオーディオ入力デバイス、並びに他のタイプの入力デバイスと、を含んでもよい。一般に、用語「入力デバイス」の使用は、コンピュータシステム3300に情報を入力するための全ての可能なタイプのデバイス及び方式を含むことを意図している。

20

【0234】

ユーザーインターフェース出力デバイス3376は、ディスプレイサブシステム、プリンタ、ファックス装置、又はオーディオ出力デバイスなどの非視覚ディスプレイを含むことができる。ディスプレイサブシステムは、LEDディスプレイ、陰極線管(Cathode Ray Tube、CRT)、液晶ディスプレイ(Liquid Crystal Display、LCD)などのフラットパネルデバイス、投影デバイス、又は可視画像を作成するための何らかの他の機構を含むことができる。ディスプレイサブシステムはまた、音声出力デバイスなどの非視覚ディスプレイを提供することができる。一般に、用語「出力デバイス」の使用は、コンピュータシステム3300からユーザー又は別のマシン若しくはコンピュータシステムに情報を出力するための、全ての可能なタイプのデバイス及び方法を含むことを意図している。

30

【0235】

記憶サブシステム3358は、本明細書に記載されるモジュール及び方法のうちのいくつか又は全ての機能を提供するプログラミング及びデータ構築物を記憶する。これらのソフトウェアモジュールは、概して、深層学習プロセッサ3378によって実行される。

【0236】

深層学習プロセッサ3378は、グラフィック処理ユニット(GPU)、フィールドプログラマブルゲートアレイ(FPGA)、特定用途向け集積回路(ASIC)、及び／又は粗粒度再構成可能構造(CGRA)であり得る。深層学習プロセッサ3378は、Google Cloud Platform(商標)、Xilinx(商標)及びCirrascal(商標)などの深層学習クラウドプラットフォームによってホスティングすることができる。深層学習プロセッサ3378の例は、GoogleのTensor Processing Unit(TPU)(商標)、GX4 Rackmount Series(商標)、GX33 Rackmount Series(商標)のようなラックマウントソリューション、NVIDIA DGX-1(商標)、Microsoft's Stratix V FPGA(商標)、GraphcoreのIntelligent Processor Unit(IPU)(商標)、Snapdragon processors(商標)を有するQualcommのZeroth Platform(商標)、NVIDIAのVolta(商標)、NVIDIAのDRIVE PX(商標)、NVIDIAのJETSON TX1/TX2 MODULE(商標)、IntelのNirvana(

40

50

商標)、Movidius VPU (商標)、Fujitsu DPI (商標)、ARMのDynamicIQ (商標)、IBM TrueNorth (商標)、Testa V100s (商標)を有するLambda GPU Server、SambaNova's Reconfigurable Dataflow Unit (RDU) (商標)、及び他のものを含む。

【0237】

記憶サブシステム3358で使用されるメモリサブシステム3322は、プログラム実行中に命令及びデータを記憶するためのメインランダムアクセスメモリ(RAM)3332と、固定命令が記憶された読み取り専用メモリ(ROM)3334とを含む多数のメモリを含むことができる。ファイル記憶サブシステム3336は、プログラム及びデータファイルのための永続的な記憶装置を提供することができ、ハードディスクドライブ、関連する取り外し可能な媒体、CD-ROMドライブ、光学ドライブ、又は取り外し可能な媒体カートリッジを含むことができる。特定の実施態様の機能を実装するモジュールは、ファイル記憶サブシステム3336によって、ストレージサブシステム3358内に、又はプロセッサによってアクセス可能な他のマシン内に記憶され得る。

【0238】

バスサブシステム3355は、コンピュータシステム3300の様々な構成要素及びサブシステムを、意図されるように互いに通信させるための機構を提供する。バスサブシステム3355は、単一のバスとして概略的に示されているが、バスサブシステムの代替実施態様は、複数のバスを使用することができる。

【0239】

コンピュータシステム3300自体は、パーソナルコンピュータ、ポータブルコンピュータ、ワークステーション、コンピュータ端末、ネットワークコンピュータ、テレビ、メインフレーム、サーバファーム、緩く分散した一組の緩くネットワーク化されたコンピュータ、又は任意の他のデータ処理システム若しくはユーザーデバイスを含む様々なタイプのものであり得る。コンピュータ及びネットワークは絶え間なく変化する性質のものであるため、図33に示されるコンピュータシステム3300の説明は、本発明の好ましい実施態様を例示する目的のための特定の例としてのみ意図される。コンピュータシステム3300の多くの他の構成は、図33に示されるコンピュータシステムよりも多く又は少ない構成要素を有することができる。

【0240】

項目

本発明者らは、以下の項目を開示する。

圧縮(スクイーズ)

1. ベースコールの人工知能ベースの方法であって、方法が、シーケンシングランのシーケンシングサイクルに対して生成された、サイクルごとの分析物チャンネルセットのシリーズにアクセスすることと、

ニューラルネットワークベースのベースコーラーの空間ネットワークを介して、シーケンシングランのシーケンシングサイクルの第1のウィンドウに対するシリーズ内のサイクルごとの分析物チャンネルセットの第1のウィンドウを処理し、かつシーケンシングサイクルの第1のウィンドウでのそれぞれのシーケンシングサイクルに対するそれぞれの空間出力セットそれぞれのシーケンスを生成することと、

ニューラルネットワークベースのベースコーラーの圧縮ネットワークを介して、空間出力セットのそれぞれのシーケンス内のそれぞれの最終空間出力セットを処理し、かつシーケンシングサイクルの第1のウィンドウ内のそれぞれのシーケンシングサイクルに対するそれぞれの圧縮空間出力セットを生成することと、

それぞれの圧縮空間出力セットに基づいて、シーケンシングサイクルの第1のウィンドウ内の1つ以上のシーケンシングサイクルに対するベースコール予測を生成することと、を含む、人工知能ベースの方法。

2. それぞれの最終空間出力セットが、M個のチャンネル(特徴マップ)を有し、それぞ

れの圧縮空間出力セットが、N個のチャンネル（特徴マップ）を有し、かつMがNより大きい、項目1に記載の人工知能ベースの方法。

3. シーケンシングサイクルの第1のウィンドウと、空間ネットワークが、以前に空間出力セットを生成した1つ以上の重なり合うシーケンシングサイクルと、空間ネットワークが、まだ空間出力セットを生成していない少なくとも1つの重なり合わないシーケンシングサイクルと、を共有するシーケンシングランのシーケンシングサイクルの第2のウィンドウに対して、

空間ネットワークを介して、重なり合わないシーケンシングサイクルのみに対するサイクルごとの分析物チャンネルセットを処理し、重なり合わないシーケンシングサイクルのための空間出力セットのシーケンスを生成し、それによって、空間ネットワークを介して重なり合うシーケンシングサイクルに対するそれぞれのサイクルごとの分析物チャンネルセットの再処理を回避することと、

10

圧縮ネットワークを介して、空間出力セットのシーケンス内の最終空間出力セットを処理し、重なり合わないシーケンシングサイクルのための圧縮空間出力セットを生成することであって、最終空間出力が、M個のチャンネル（特徴マップ）を有し、圧縮空間出力が、N個のチャンネル（特徴マップ）を有し、MがNより大きい、ことと、

シーケンシングサイクルの第1のウィンドウ及び圧縮空間出力セットに対して以前に生成された重なり合うシーケンシングサイクルのためのそれぞれの圧縮空間出力セットに基づいて、シーケンシングサイクルの第2のウィンドウ内の1つ以上のシーケンシングサイクルについてのベースコール予測を生成し、それによって、重なり合うシーケンシングサイクルに対するそれぞれの圧縮された空間出力セットを、重なり合うシーケンシングサイクルに対するそれぞれのサイクルごとの分析物チャンネルセットに置換することと、を更に含む、項目1に記載の人工知能ベースの方法。

20

4. シーケンシングサイクルの第1及び第2のウィンドウと、空間ネットワークが、以前に空間出力セットを生成した、1つ以上の重なり合うシーケンシングサイクルと、空間ネットワークが、まだ空間出力セットを生成していない少なくとも1つの重なり合わないシーケンシングサイクルと、を共有するシーケンシングランのシーケンシングサイクルの第3のウィンドウに対して、

空間ネットワークを介して、重なり合わないシーケンシングサイクルのみに対するサイクルごとの分析物チャンネルセットを処理し、重なり合わないシーケンシングサイクルのための空間出力セットのシーケンスを生成し、それによって、空間ネットワークを介して重なり合うシーケンシングサイクルに対するそれぞれのサイクルごとの分析物チャンネルセットの再処理を回避することと、

30

圧縮ネットワークを介して、空間出力セットのシーケンス内の最終空間出力セットを処理し、重なり合わないシーケンシングサイクルのための圧縮空間出力セットを生成することであって、最終空間出力が、M個のチャンネル（特徴マップ）を有し、圧縮空間出力が、N個のチャンネル（特徴マップ）を有し、MがNより大きい、ことと、

シーケンシングサイクルの第1及び第2のウィンドウ及び圧縮空間出力セットに対して以前に生成された重なり合うシーケンシングサイクルのためのそれぞれの圧縮空間出力セットに基づいて、シーケンシングサイクルの第3のウィンドウ内の1つ以上のシーケンシングサイクルについてのベースコール予測を生成し、それによって、重なり合うシーケンシングサイクルに対するそれぞれの圧縮された空間出力セットを、重なり合うシーケンシングサイクルに対するそれぞれのサイクルごとの分析物チャンネルセットに置換することと、を更に含む、項目3に記載の人工知能ベースの方法。

40

5. シリーズ内の各サイクルごとの分析物チャンネルセットが、シーケンシングランにおける対応するシーケンシングサイクルでの分析物へのヌクレオチド取り込みに応答して登録された強度を示す、項目1に記載の人工知能ベースの方法。

6. 空間ネットワークが、シーケンシングランのシーケンシングサイクルの特定のウィンドウのためのシリーズにおけるサイクルごとの分析物チャンネルセットの特定のウィンドウにおける各サイクルごとの分析物チャンネルセットを、別々に処理する、空間畳み込み層

50

のシーケンスを有し、シーケンシングサイクルの特定のウィンドウ内の各シーケンシングサイクルに対して、空間出力セットのシーケンスを生成し、シーケンシングサイクルの特定のウィンドウ内の異なるシーケンシングサイクルの、サイクルごとの分析物チャネルセットどうしの間ではなく、対象のシーケンシングサイクルの、1つの、サイクルごとの分析物チャネルセット内でのみ、強度を組み合わせる、第1空間畳み込み層から開始することと、先行する空間畳み込み層の空間出力を、シーケンシングサイクルの特定のウィンドウ内の異なるシーケンシングサイクル間ではなく、対象のシーケンシングサイクル内でのみ組み合わせる、後続の空間畳み込み層により引き続き行うことと、を含む、項目5に記載の人工知能ベースの方法。

7. 空間畳み込み層のシーケンス内のそれぞれの空間畳み込み層が、異なる数の畳み込みフィルタを有し、空間畳み込み層のシーケンス内の最終空間畳み込み層が、M個の畳み込みフィルタを有し、Mが、4より大きい整数である、項目6に記載の人工知能ベースの方法。

10

8. 空間畳み込み層のシーケンス内のそれぞれの空間畳み込み層が、同じ数の畳み込みフィルタを有し、その同じ数がMであり、Mが、4より大きい整数である、項目7に記載の人工知能ベースの方法。

9. 空間ネットワーク内の畳み込みフィルタが、2次元 (two-dimensional、2D) の畳み込みを使用する、項目8に記載の人工知能ベースの方法。

10. 空間ネットワーク内の畳み込みフィルタが、3次元 (three-dimensional、3D) の畳み込みを使用する、項目8に記載の人工知能ベースの方法。

20

11. ニューラルネットワークベースのベースコーラーが、時間ネットワークを有し、時間ネットワークが、シーケンシングサイクルの特定のウィンドウ内の後続のシーケンシングサイクルのウィンドウに対するそれぞれの圧縮空間出力セットをグループ化して処理する時間畳み込み層のシーケンスを有し、シーケンシングサイクルの特定のウィンドウに対する時間出力セットのシーケンスを生成し、シーケンシングサイクルの特定のウィンドウ内の異なるシーケンシングサイクルどうしの間で圧縮空間出力セットを組み合わせる第1の時間畳み込み層から開始することと、先行する時間畳み込み層の後続する時間出力を組み合わせる後続の時間畳み込み層により引き続き行うことと、を含む、項目6に記載の人工知能ベースの方法。

12. シーケンシングサイクルの第1のウィンドウに対して、

30

時間ネットワークの時間畳み込み層のシーケンスで第1の時間畳み込み層を介して、シーケンシングサイクルの第1のウィンドウ内の後続のシーケンシングサイクルのウィンドウに対するそれぞれの圧縮空間出力セットを処理し、シーケンシングサイクルの第1のウィンドウに対する複数の時間出力セットを生成することと、

圧縮ネットワークを介して、複数の時間出力セットを処理し、複数の時間出力セット内のそれぞれの時間出力セットのためのそれぞれの圧縮時間出力セットを生成することであって、それぞれの時間出力セットが、M個のチャネル (特徴マップ) を有し、それぞれの圧縮時間出力セットが、N個のチャネル (特徴マップ) を有し、MがNより大きい、ことと、

時間ネットワークの時間畳み込み層のシーケンスにおける最終時間畳み込み層を介して、それぞれの圧縮時間出力セットを処理し、シーケンシングサイクルの第1のウィンドウに対する最終時間出力セットを生成することと、

40

最終時間出力セットに基づいて、シーケンシングサイクルの第1のウィンドウ内の1つ以上のシーケンシングサイクルに対するベースコール予測を生成することと、を更に含み、出力層が、最終時間出力セットを処理し、シーケンシングサイクルの第1のウィンドウに対する最終出力を生成し、ベースコール予測は、最終出力に基づいて生成される、項目11に記載の人工知能ベースの方法。

13. シーケンシングサイクルの第1のウィンドウと、第1の時間畳み込み層が以前に時間出力セットを生成した後続のシーケンシングサイクルの1つ以上の重なり合うウィンドウと、第1の時間畳み込み層がまだ時間出力セットを生成していない後続のシーケンシ

50

ングサイクルの少なくとも1つの重なり合わないウィンドウを共有するシーケンシングサイクルの第2のウィンドウに対して、

第1の時間量み込み層を介して、後続のシーケンシングサイクルの重なり合わないウィンドウにおけるそれぞれのシーケンシングサイクルに対してのみ、それぞれの圧縮空間出力セットを処理し、後続のシーケンシングサイクルの重なり合わないウィンドウに対する時間出力セットを生成し、それによって、後続のシーケンシングサイクルの重なり合うウィンドウ内のそれぞれのシーケンシングサイクルに対する、第1の時間量み込み層を通じた、それぞれの圧縮空間出力セットの再処理を回避することと、

圧縮ネットワークを介して、時間出力セットを処理し、後続のシーケンシングサイクルの重なり合わないウィンドウに対する圧縮時間出力セットを生成することであって、時間出力セットが、M個のチャンネル（特徴マップ）を有し、圧縮時間出力が、N個のチャンネル（特徴マップ）を有し、MがNより大きい、ことと、

最終時間量み込み層を介して、シーケンシングサイクルの第1のウィンドウ及び圧縮時間出力セットに対して以前に生成された、後続のシーケンシングサイクルの重なり合うウィンドウに対するそれぞれの圧縮時間出力セットを処理し、シーケンシングサイクルの第2のウィンドウのための最終時間出力セットを生成し、それによって、後続のシーケンシングサイクルの重なり合うウィンドウに対するそれぞれの圧縮時間出力セットを、後続のシーケンシングサイクルの重なり合うウィンドウに対するそれぞれのサイクルごとの分析物チャンネルセットに置換することと、

最終時間出力セットに基づいて、シーケンシングサイクルの第2のウィンドウ内の1つ以上のシーケンシングサイクルに対するベースコール予測を生成することと、を更に含み、出力層が、最終時間出力セットを処理し、シーケンシングサイクルの第2のウィンドウに対する最終出力を生成し、ベースコール予測は、最終出力に基づいて生成される、項目12に記載の人工知能ベースの方法。

14. シーケンシングサイクルの第1及び第2のウィンドウと、第1の時間量み込み層が以前に時間出力セットを生成した後続のシーケンシングサイクルの1つ以上の重なり合うウィンドウ、及び第1の時間量み込み層がまだ時間出力セットを生成していない後続のシーケンシングサイクルの少なくとも1つの重なり合わないウィンドウを共有するシーケンシングサイクルの第3のウィンドウに対して、

第1の時間量み込み層を介して、後続のシーケンシングサイクルの重なり合わないウィンドウにおけるそれぞれのシーケンシングサイクルに対してのみ、それぞれの圧縮空間出力セットを処理し、後続のシーケンシングサイクルの重なり合わないウィンドウに対する時間出力セットを生成し、それによって、後続のシーケンシングサイクルの重なり合うウィンドウ内のそれぞれのシーケンシングサイクルに対する、第1の時間量み込み層を介する、それぞれの圧縮空間出力セットの再処理を回避することと、

圧縮ネットワークを介して、時間出力セットを処理し、後続のシーケンシングサイクルの重なり合わないウィンドウに対する圧縮時間出力セットを生成することであって、時間出力セットが、M個のチャンネル（特徴マップ）を有し、圧縮時間出力が、N個のチャンネル（特徴マップ）を有し、MがNより大きい、ことと、

最終時間量み込み層を介して、シーケンシングサイクルの第1及び第2のウィンドウ及び圧縮時間出力セットに対して以前に生成された、後続のシーケンシングサイクルの重なり合うウィンドウに対するそれぞれの圧縮時間出力セットを処理し、シーケンシングサイクルの第3のウィンドウのための最終時間出力セットを生成し、それによって、後続のシーケンシングサイクルの重なり合うウィンドウに対するそれぞれの圧縮時間出力セットを、後続のシーケンシングサイクルの重なり合うウィンドウに対するそれぞれのサイクルごとの分析物チャンネルセットに置換することと、

最終時間出力セットに基づいて、シーケンシングサイクルの第3のウィンドウ内の1つ以上のシーケンシングサイクルに対するベースコール予測を生成することと、を更に含み、出力層が、最終時間出力セットを処理し、シーケンシングサイクルの第3のウィンドウに対する最終出力を生成し、ベースコール予測は、最終出力に基づいて生成される、項目

10

20

30

40

50

13に記載の人工知能ベースの方法。

15. 時間ネットワークの時間畳み込み層のシーケンス内のそれぞれの時間畳み込み層が、異なる数の畳み込みフィルタを有し、第1の時間畳み込み層が、M個の畳み込みフィルタを有し、Mが4より大きい整数である、項目11に記載の人工知能ベースの方法。

16. 時間ネットワークの時間畳み込み層のシーケンス内のそれぞれの時間畳み込み層が、同じ数の畳み込みフィルタを有し、その同じ数がMであり、Mが4より大きい整数である、項目11に記載の人工知能ベースの方法。

17. 時間ネットワーク内の畳み込みフィルタが1次元(1D)畳み込みを使用する、項目16に記載の人工知能ベースの方法。

18. 圧縮ネットワークが、 1×1 の畳み込みを使用して、圧縮空間出力セット内の圧縮空間出力の数を制御し、圧縮ネットワークはN個の畳み込みフィルタを有し、Nは4以下の整数である、項目1に記載の人工知能ベースの方法。

19. 信頼できない分析物を識別するデータを使用して、信頼できない分析物に対応する圧縮空間出力セット内の圧縮空間出力の部分を除き、フィルタリングされた圧縮空間出力セットを生成して、圧縮空間出力セットを置換し、信頼できない分析物ではないそれらの分析物に対してのみベースコール予測を生成することを更に含む、項目1に記載の人工知能ベースの方法。

20. フィルタリングされた圧縮空間出力セットを、対応する圧縮空間出力セットの代わりに、時間ネットワークを介して処理することを更に含む、項目19に記載の人工知能ベースの方法。

21. フィルタリングされた圧縮空間出力セットから圧縮時間出力セットを生成することを更に含む、項目20に記載の人工知能ベースの方法。

22. 信頼できない分析物を識別するデータが、信頼できないクラスタの強度を示すピクセルを識別する、項目19に記載の人工知能ベースの方法。

23. 信頼できない分析物を識別するデータが、いかなる強度も示さないピクセルを識別する、項目19に記載の人工知能ベースの方法。

24. 圧縮空間出力セットが、対応するフィルタリングされた圧縮空間出力セットの合計ピクセル数の4~9倍の数の合計ピクセル数を有する、項目20に記載の人工知能ベースの方法。

25. フィルタリングされた圧縮空間出力セットにより、時間ネットワークが75%少ないピクセルで動作し、それにより、時間ネットワークの計算操作、メモリアクセス、及びメモリ占有を75%削減する、項目24に記載の人工知能ベースの方法。

26. 空間ネットワークを介した再処理を回避することが、時間ネットワークに対する計算動作、メモリアクセス、及びメモリ占有を80%低減させる、項目5に記載の人工知能ベースの方法。

27. 時間ネットワークを介した再処理を回避することが、時間ネットワークの計算動作、メモリアクセス、及びメモリ占有を低減する、項目14に記載の人工知能ベースの方法。

28. 空間ネットワーク及び時間ネットワークにおける補足的な畳み込みフィルタの追加に、圧縮ネットワークによって利用可能とされた計算リソースを再度割り当てることを更に含む、項目27に記載の人工知能ベースの方法。

29. 特定のシーケンシングサイクルに対するベースコール予測を生成するために使用されるサイクルごとの分析物チャネルセットの各ウィンドウ内の補足的なサイクルごとの分析物チャネルセットの追加に、圧縮ネットワークによって利用可能とされた計算リソースを再度割り当てることを更に含む、項目27に記載の人工知能ベースの方法。

30. 圧縮ネットワークによって利用可能とされた計算リソースを、空間ネットワーク内の補足的な空間畳み込み層の追加に再度割り当てることを更に含む、項目27に記載の人工知能ベースの方法。

31. 圧縮ネットワークによって利用可能とされた計算リソースを、時間ネットワーク内の補足的な時間畳み込み層の追加に再度割り当てることを更に含む、項目27に記載の

10

20

30

40

50

人工知能ベースの方法。

32. シーケンシングサイクルの現在のウィンドウに対して生成された1つ以上の圧縮空間出力セットと併せて、シーケンスサイクルの1つ以上の先行するウィンドウに対して生成された1つ以上の圧縮空間出力セットを使用して、シーケンシングサイクルの現在のウィンドウにおける1つ以上のシーケンシングサイクルに対するベースコール予測を生成することを更に含む、項目1に記載の人工知能ベースの方法。

33. シーケンスサイクルの現在のウィンドウに対して生成された1つ以上の圧縮空間出力セットと併せて、シーケンスサイクルの1つ以上の後続ウィンドウに対して生成された1つ以上の圧縮空間出力セットを使用して、シーケンシングサイクルの現在のウィンドウにおける1つ以上のシーケンシングサイクルに対するベースコール予測を生成することを更に含む、項目1に記載の人工知能ベースの方法。

10

34. シーケンスサイクルの現在のウィンドウに対して生成された1つ以上の圧縮時間出力セットと併せて、シーケンスサイクルの1つ以上の先行するウィンドウに対して生成された1つ以上の圧縮時間出力セットを使用して、シーケンシングサイクルの現在のウィンドウにおける1つ以上のシーケンシングサイクルに対するベースコール予測を生成することを更に含む、項目1に記載の人工知能ベースの方法。

35. シーケンスサイクルの現在のウィンドウに対して生成された1つ以上の圧縮時間出力セットと併せて、シーケンスサイクルの1つ以上の後続のウィンドウに対して生成された1つ以上の圧縮時間出力セットを使用して、シーケンシングサイクルの現在のウィンドウにおける1つ以上のシーケンシングサイクルに対するベースコール予測を生成することを更に含む、項目1に記載の人工知能ベースの方法。

20

36. サイクルごとの分析物チャネルセットが、シーケンシングラン中にシーケンシングされた分析物の分析物データをコード化する、項目1に記載の人工知能ベースの方法。

37. 分析物データが、分析物から収集された強度放射を識別する画像データである、項目36に記載の人工知能ベースの方法。

38. 画像データが、複数の画像チャネル（画像）を有する、項目37に記載の人工知能ベースの方法。

39. 画像チャネル（画像）が、(i) 特定のレーザでの照明及び(ii) 特定の光学フィルタを介した撮像の組み合わせによって生成される、項目38に記載の人工知能ベースの方法。

30

40. 分析物データが、分析物活性に基づいて検出された電流及び／又は電圧データである、項目36に記載の人工知能ベースの方法。

41. 分析物データが、分析物活性に基づいて検出されたpHスケールデータである、項目36に記載の人工知能ベースの方法。

42. 直列に設定された各サイクルごとの分析物チャネル内のチャネルの数が、圧縮ネットワーク内の畳み込みフィルタの数、したがって圧縮空間出力セット及び圧縮時間出力セット内のチャネルの数を決定する、項目1に記載の人工知能ベースの方法。

43. 圧縮空間出力セット、圧縮フィルタ空間出力セット、及び圧縮時間出力セットが、量子化された形態で格納される、項目1に記載の人工知能ベースの方法。

44. システムであって、

40

ホストプロセッサに取り付けられ、シーケンシングランの進行としてシーケンス画像の進行を受信するように構成されたホストメモリと、

処理ユニットのアレイを有する構成可能なプロセッサであって、処理ユニットのアレイ内の処理ユニットが、ニューラルネットワークベースのベースコーラーを実行してベースコール予測を生成するように構成されている、構成可能なプロセッサと、

ホストメモリ、ホストプロセッサ、及び構成可能なプロセッサへのアクセスを有し、ホストメモリから構成可能なプロセッサ上のシーケンシングサイクルの第1のウィンドウ内のシーケンシングサイクルに対してシーケンス画像をロードするように構成されたデータフローロジックと、

処理ユニットに、ニューラルネットワークベースのベースコーラーの空間ネットワーク

50

を、サイクルごとに、シーケンシングサイクルの第1のウィンドウにおけるシーケンシングサイクルのシーケンス画像に対して実行させ、シーケンシングサイクルの第1のウィンドウ内のシーケンシングサイクルの各々に対して空間特徴マップセットを生成するように構成された実行時ロジックであって、

処理ユニットに、ニューラルネットワークベースのベースコーラーの圧縮ネットワークを、サイクルごとに空間特徴マップセットに対して実行させ、圧縮空間特徴マップセットを生成し、圧縮空間特徴マップセットを、時間ネットワーク及び出力ネットワークを介して処理して、シーケンシングサイクルの第1のウィンドウ内の1つ以上のシーケンシングサイクルのためのベースコール予測を生成させるように構成された、実行時ロジックと、を備え、

10

データフローロジックが、圧縮空間特徴マップセットをホストメモリに移動させ、圧縮空間特徴マップセットでシーケンス画像を上書きするように構成され、

1つ以上の重なり合うシーケンシングサイクルを、シーケンシングサイクルの第1のウィンドウと共有し、少なくとも1つの重なり合わないシーケンシングサイクルを有する、シーケンシングサイクルの第2のウィンドウに対して、データフローロジックが構成可能なプロセッサ上にホストメモリから、重なり合うシーケンシングサイクルに対して圧縮空間特徴マップセットをロードし、重なり合わないシーケンシングサイクルに対してシーケンス画像をロードするように構成され、

実行時ロジックが、重なり合わないシーケンシングサイクルのシーケンス画像上で処理ユニットに空間ネットワークを実行させ、重なり合わないシーケンシングサイクルのための空間特徴マップセットを生成させるように構成され、

20

実行時ロジックが、処理ユニットが、空間特徴マップセット上で圧縮ネットワークを実行し、重なり合わないシーケンシングサイクルに対して圧縮空間特徴マップセットを生成し、重なり合うシーケンシングサイクルに対して、圧縮空間特徴マップセットを処理し、かつ重なり合わないシーケンシングサイクルに対して、時間ネットワーク及び出力ネットワークを介して、圧縮空間特徴マップセットを処理し、かつシーケンシングサイクルの第2のウィンドウ中の1つ以上のシーケンシングサイクルに対してベースコール予測を生成するように構成されている、システム。

45. システムであって、

ホストプロセッサに取り付けられ、シーケンシングランの進行としてシーケンス画像の進行を受信するように構成されたホストメモリと、

30

プロセッサメモリに取り付けられた処理ユニットのアレイを有する構成可能なプロセッサであって、処理ユニットのアレイ内の処理ユニットが、ニューラルネットワークベースのベースコーラーを実行して、ベースコール予測を生成するように構成された、構成可能なプロセッサと、

ホストメモリ、ホストプロセッサ、構成可能なプロセッサ、及びプロセッサメモリへのアクセスを有し、ホストメモリから構成可能なプロセッサ上のシーケンシングサイクルの第1のウィンドウ内のシーケンシングサイクルに対してシーケンス画像をロードするように構成されたデータフローロジックと、

処理ユニットに、ニューラルネットワークベースのベースコーラーの空間ネットワークを、サイクルごとに、シーケンシングサイクルの第1のウィンドウにおけるシーケンシングサイクルのシーケンス画像に対して実行させ、シーケンシングサイクルの第1のウィンドウ内のシーケンシングサイクルの各々に対して空間特徴マップセットを生成させるように構成された実行時ロジックであって、

40

処理ユニットに、ニューラルネットワークベースのベースコーラーの圧縮ネットワークを、サイクルごとに空間特徴マップセットに対して実行させ、圧縮空間特徴マップセットを生成し、圧縮空間特徴マップセットを、時間ネットワーク及び出力ネットワークを介して処理して、シーケンシングサイクルの第1のウィンドウ内の1つ以上のシーケンシングサイクルのためのベースコール予測を生成させるように構成された、実行時ロジックと、を備え、

50

データフローロジックが、圧縮空間特徴マップセットをプロセッサメモリに移動させるように構成され、

1つ以上の重なり合うシーケンシングサイクルをシーケンシングサイクルの第1のウィンドウと共有し、少なくとも1つの重なり合わないシーケンシングサイクルを有する、シーケンシングサイクルの第2のウィンドウに対して、データフローロジックが構成可能なプロセッサ上にプロセッサメモリから、重なり合うシーケンシングサイクルに対して圧縮空間特徴マップセットをロードし、重なり合わないシーケンシングサイクルに対してシーケンス画像をホストメモリからロードするように構成され、

実行時ロジックが、重なり合わないシーケンシングサイクルのシーケンス画像上で処理ユニットに空間ネットワークを実行させ、重なり合わないシーケンシングサイクルのための空間特徴マップセットを生成させるように構成され、

10

実行時ロジックが、処理ユニットが、空間特徴マップセット上で圧縮ネットワークを実行し、重なり合わないシーケンシングサイクルに対して圧縮空間特徴マップセットを生成し、重なり合うシーケンシングサイクルに対して、圧縮空間特徴マップセットを処理し、かつ重なり合わないシーケンシングサイクルに対して、時間ネットワーク及び出力ネットワークを介して、圧縮空間特徴マップセットを処理し、かつシーケンシングサイクルの第2のウィンドウ中の1つ以上のシーケンシングサイクルに対してベースコール予測を生成するように構成されている、システム。

46. システムであって、

ニューラルネットワークグラフの第1の横断を実行して、第1の処理ロジックを介して第1の入力セット内のそれぞれの入力を独立して処理し、第1の入力セット内のそれぞれの入力どうしの間で情報を混合することなく、第1の入力セット内のそれぞれの入力のそれぞれの代替表現を生成し、第1の入力セット内のそれぞれの入力のそれぞれの代替表現に基づいて、第1の横断の出力を生成するように構成されたニューラルネットワークロジックであって、

20

ニューラルネットワークグラフの第2の横断を実行して、第1の処理ロジックを介して第2の入力セット内のそれぞれの入力を独立して処理し、第2の入力セット内のそれぞれの入力どうしの間で情報を混合することなく、第2の入力セット内のそれぞれの入力のそれぞれの代替表現を生成し、第2の入力セット内のそれぞれの入力のそれぞれの代替表現に基づいて、第2の横断の出力を生成するように構成され、第1及び第2の入力セットが、1つ以上の重なり合う入力、及び少なくとも1つの重なり合わない入力を有する、ニューラルネットワークロジックと、

30

ニューラルネットワークロジックとともに、第1の横断を実行して、第1の入力セット内のそれぞれの入力のそれぞれの代替表現を生成して、第1の入力セット内のそれぞれの入力のそれぞれの代替表現を、メモリに、圧縮された形式で記憶し、第1の入力セット内のそれぞれの入力の圧縮された形式のそれぞれの代替表現に基づいて、第1の横断に対する出力を生成するように構成された実行時ロジックであって、

第2の横断を実行して、第1の処理ロジックを介して重なり合わない入力のみを処理し、重なり合わない入力の代替表現を生成して、その重なり合わない入力の代替表現を、圧縮された形式でメモリに記憶して、第1の横断中に生成された重なり合う入力の、圧縮した形式のそれぞれの代替表現を読み出して、第2の横断における重なる入力のそれぞれの代替表現の冗長な生成を回避する補償とし、かつ重なり合う入力の圧縮された形式のそれぞれの代替表現及び重なり合わない入力の、圧縮された形式の代替表現に基づいて、第2の横断に対する出力を生成するように構成された、実行時ロジックと、を備える、システム。

40

47. メモリが、オンチップメモリである、項目46に記載のシステム。

48. メモリが、オフチップメモリである、項目46に記載のシステム。

49. 圧縮形態のチャネルの数が、第1及び第2の入力セット内の入力のチャネルの数に対応する、項目46に記載のシステム。

50. ベースコールの人工知能ベースの方法であって、方法が、

50

シーケンシングランのシーケンシングサイクルのために生成された、サイクルごとの分析物チャンネルセットのシリーズにアクセスすることであって、対象のサイクルごとの分析物チャンネルセットが、シーケンシングランの対象のシーケンシングサイクルにおいて、分析物に対して検知された分析物データをコード化する、ことと、

ニューラルネットワークの第1の処理モジュールを介して対象のサイクルごとの分析物チャンネルセットを処理し、M個の特徴マップを有する、対象のサイクルごとの分析物チャンネルセットの中間表現を生成することと、

ニューラルネットワークの第2の処理モジュールを介して中間表現を処理し、N個（ただしNはMより小さい）の特徴マップを有する、対象のサイクルごとの分析物チャンネルセットの中間表現をより少ない数で生成することと、

対象のサイクルごとの分析物チャンネルセットの、より少ない数の中間表現を使用して、対象のシーケンシングサイクル及び／又はシーケンシングランの他のシーケンシングサイクルでの、分析物に対するベースコール予測を生成することと、を含む、人工知能ベースの方法。

51. 第1の処理モジュールが、M個の畳み込みフィルタを備えた畳み込み層である、項目50に記載の人工知能ベースの方法。

52. 第2の処理モジュールが、N個の畳み込みフィルタを備えた畳み込み層である、項目50に記載の人工知能ベースの方法。

53. ベースコールの人工知能ベースの方法であって、方法が、

連続するスライドウィンドウが互いに重なり合うシーケンシングサイクルを有するように、スライドウィンドウベースのニューラルネットワークベースのベースコーラーを介して、シーケンシングランのシーケンシングサイクルに対して生成されたサイクルごとの分析物チャンネルセットの進行を処理することと、

1つ以上の先行するシーケンシングサイクル、中央シーケンシングサイクル、及び1つ以上の後続シーケンシングサイクルを含むシーケンシングサイクルの現在のウィンドウに対して、

サイクルごとの分析物チャンネルセットの現在のウィンドウにニューラルネットワークベースのベースコーラーを、適用することに基づいて、先行するシーケンシングサイクル、中央シーケンシングサイクル、及び後続シーケンシングサイクルの各々に対して、空間中間表現及び圧縮中間表現を生成することであって、空間中間表現が、M個のチャンネルを有し、圧縮中間表現が、N個のチャンネルを有し、MがNより大きい、ことと、

先行するシーケンシングサイクル、中央シーケンシングサイクル、及び後続シーケンシングサイクルに対して生成された圧縮中間表現に基づいて、少なくとも中央シーケンシングサイクルをベースコールすることと、

先行するシーケンシングサイクル、中央シーケンシングサイクル、及び後続シーケンシングサイクルに対して生成された圧縮中間表現を使用して、シーケンシングサイクルの次のウィンドウ内の少なくとも中央シーケンシングサイクルでベースコールすることと、を含む、人工知能ベースの方法。

54. ベースコールの人工知能ベースのシステムであって、システムが、
ホストプロセッサと、

ホストプロセッサによってアクセス可能であり、シーケンシングランのシーケンシングサイクルに対する分析物データを記憶するメモリと、

メモリへのアクセスを有する構成可能なプロセッサあって、

複数の実行クラスタであって、ニューラルネットワークを実行するように構成された複数の実行クラスタ内の実行クラスタと、

メモリへのアクセス及び複数の実行クラスタ内の実行クラスタへのアクセスを有するデータフローロジックであって、複数の実行クラスタ内の利用可能な実行クラスタに、分析物データを提供し、実行クラスタが、分析物データをニューラルネットワークに適用し、分析データの中間表現と圧縮された中間表現を現在のベースコールステップで使用するよう生成し、圧縮された中間表現をメモリにフィードバックして、将来のベースコールス

10

20

30

40

50

テップにおいて、分析データに替えて用いられるように構成され、中間表現はM個のチャンネルを有し、圧縮された中間表現はN個のチャンネルを有し、MはNよりも大きい、データフローロジックと、を含む、プロセッサと、を備える、人工知能ベースのシステム。

55. システムであって、

ベースコーラーの第1の反復を実行して、入力を処理し、入力の中間表現を生成するように構成された実行時ロジックと、

中間表現を処理し、入力の圧縮中間表現を生成するように構成された圧縮ロジックと、を備え、

実行時ロジックが、ベースコーラーの後続の反復における入力の代わりに圧縮中間表現を使用するように構成されている、システム。

10

56. システムであって、

ベースコーラーの第1の反復を実行して、入力を処理し、入力の中間表現を生成するように構成された実行時ロジックと、

中間表現を処理し、圧縮中間表現を生成するように構成された圧縮ロジックであって、圧縮中間表現が、入力と同じ数のチャンネルを有するように構成されている、圧縮ロジックと、を備え、

実行時ロジックが、ベースコーラーの後続の反復における入力の代わりに圧縮中間表現を使用するように構成されている、システム。

57. チャンネルが、特徴マップに対応する、項目56に記載のシステム。

58. チャンネルが、深さ寸法に対応する、項目56に記載のシステム。

20

59. チャンネルが、空間寸法に対応する、項目56に記載のシステム。

【0241】

分割

1. システムであって、

空間畳み込み層のそれぞれのシーケンスを介して、サイクルごとのシーケンス画像セットのウィンドウ内の、それぞれのサイクルごとのシーケンス画像セットを別々に畳み込むことによって、サイクルごとに、シーケンシングランのシーケンシングサイクルのシリーズに対して、サイクルごとのシーケンス画像セットのウィンドウを処理して、シーケンシングサイクルのシリーズ内のそれぞれのシーケンシングサイクルに対して、それぞれのサイクルごとの空間特徴マップセットを生成するように構成されている空間畳み込みネットワークであって、

30

空間畳み込み層のそれぞれのシーケンスが、空間畳み込みフィルタバンクのそれぞれのシーケンスを有し、空間畳み込みフィルタバンクのそれぞれのシーケンスの空間畳み込みフィルタバンク内の空間畳み込みフィルタを、空間畳み込み層の学習された係数それぞれのシーケンス内の空間畳み込み層のシーケンス間で変化させる、空間畳み込みネットワークと、

第1の時間畳み込み層のそれぞれの時間畳み込みフィルタバンクを用いて、サイクルごとの空間特徴マップセットにおけるサイクルごとの空間特徴マップセットのそれぞれの重なり合うグループに畳み込むことによって、サイクルごとの空間特徴マップセットをグループごとに処理して、サイクルごとの空間特徴マップセットの、それぞれ重なり合うグループに対する、それぞれのグループごとの時間特徴マップセットを生成するように構成されている時間畳み込みネットワークであって、

40

それぞれの時間畳み込みフィルタバンク内の時間畳み込みフィルタの学習された係数を、それぞれの時間畳み込みフィルタバンク内の時間畳み込みフィルタバンク間で変化させる、時間畳み込みネットワークと、を備える、システム。

2. 空間畳み込みフィルタが、サイクル内の分割された畳み込みを使用する、項目1に記載のシステム。

3. 時間畳み込みフィルタが、サイクル間の組み合わせ畳み込みを使用する、項目1に記載のシステム。

4. それぞれのサイクルごとの空間特徴マップセットを、それぞれの圧縮畳み込み層を

50

介して、別々に畳み込んで、それぞれのシーケンシングサイクルに対する、それぞれのサイクルごとの圧縮空間特徴マップセットを生成する圧縮ネットワークを備えるように更に構成されている、項目 1 に記載のシステム。

5. それぞれの圧縮畳み込み層内の圧縮畳み込みフィルタの学習された係数が、それぞれの圧縮畳み込み層内の圧縮畳み込み層間で変化する、項目 4 に記載のシステム。

6. 時間畳み込みネットワークが、グループごとの時間特徴マップセットにおけるグループごとの時間特徴マップセットのそれぞれの重なり合うグループに、第 2 の時間畳み込み層のそれぞれの時間畳み込みフィルタバンクを使用して、畳み込みを行うことによって、グループごとにグループごとの時間特徴マップセットを処理し、グループごとの時間特徴マップセットのそれぞれの重なり合うグループに対してそれぞれの更なるグループごとの時間特徴マップセットを生成するように更に構成されている、項目 5 に記載のシステム。

7. 最終時間畳み込み層によって生成された最終時間特徴マップセットを処理して最終出力を生成する、出力ネットワークを備えるように更に構成されている、項目 6 に記載のシステム。

8. 最終出力に基づいて、シーケンシングサイクルのシリーズにおいて、1 つ以上のシーケンシングサイクルのためのベースコール予測を生成するように更に構成されている、項目 7 に記載のシステム。

9. システムであって、

空間畳み込み層のそれぞれのシーケンスを介して、サイクルごとのシーケンス画像セットのウィンドウ内の、それぞれのサイクルごとのシーケンス画像セットを別々に畳み込むことによって、サイクルごとに、シーケンシングランのシーケンシングサイクルのシリーズに対して、サイクルごとのシーケンス画像セットのウィンドウを処理して、シーケンシングサイクルのシリーズ内のそれぞれのシーケンシングサイクルに対して、それぞれのサイクルごとの空間特徴マップセットを生成するように構成されている空間畳み込みネットワークであって、

それぞれの時間畳み込みフィルタバンクを用いて、サイクルごとの空間特徴マップセットのそれぞれの重なり合うグループに畳み込むことによって、サイクルごとの空間特徴マップセットをグループごとに処理して、サイクルごとの空間特徴マップセットの、サイクルごとの空間特徴マップセットにおけるそれぞれ重なり合うグループに対する、それぞれのグループごとの時間特徴マップセットを生成するように構成されている時間畳み込みネットワークであって、

それぞれの時間畳み込みフィルタバンク内の時間畳み込みフィルタの学習された係数それぞれの時間畳み込みフィルタバンク内の時間畳み込みフィルタバンク間で変化する、時間畳み込みネットワークと、を備える、システム。

10. 空間畳み込み層のそれぞれのシーケンスが、空間畳み込みフィルタバンクのそれぞれのシーケンスを有し、空間畳み込みフィルタバンクのそれぞれのシーケンスの空間畳み込みフィルタバンク内の空間畳み込みフィルタのが、空間畳み込み層の学習された係数それぞれのシーケンス内の空間畳み込み層のシーケンス間で共有される、項目 9 に記載のシステム。

11. それぞれのサイクルごとの空間特徴マップセットを、それぞれの圧縮畳み込み層を介して、別々に畳み込んで、それぞれ対応するシーケンシングサイクルに対する、それぞれのサイクルごとの圧縮空間特徴マップセットを生成する圧縮ネットワークを備えるように更に構成され、それぞれの圧縮畳み込み層内の圧縮畳み込みフィルタの、が、学習された係数それぞれの圧縮畳み込み層内の圧縮畳み込み層間で変化する、項目 9 に記載のシステム。

12. ベースコールの人工知能ベースの方法であって、方法が、

空間畳み込み層のそれぞれのシーケンスを介して、サイクルごとのシーケンス画像セットのウィンドウ内の、それぞれのサイクルごとのシーケンス画像セットを別々に畳み込むことによって、空間畳み込みネットワークを介して、サイクルごとに、シーケンシングランのシーケンシングサイクルのシリーズに対するサイクルごとのシーケンス画像セットの

10

20

30

40

50

ウィンドウを処理し、シーケンシングサイクルのシリーズ内のそれぞれのシーケンシングサイクルに対して、それぞれのサイクルごとの空間特徴マップセットを生成することであって、

空間畳み込み層のそれぞれのシーケンスが、空間畳み込みフィルタバンクのそれぞれのシーケンスを有し、空間畳み込みフィルタバンクのそれぞれのシーケンスの空間畳み込みフィルタバンク内の空間畳み込みフィルタを、空間畳み込み層の学習された係数それぞれのシーケンス内の空間畳み込み層のシーケンス間で変化する、ことと、

第1の時間畳み込み層のそれぞれの時間畳み込みフィルタバンクを用いて、サイクルごとの空間特徴マップセットにおけるサイクルごとの空間特徴マップセットのそれぞれの重なり合うグループに畳み込むことによって、時間畳み込みネットワークを介して、サイクルごとの空間特徴マップセットをグループごとに処理し、サイクルごとの空間特徴マップセットのそれぞれ重なり合うグループに対する、それぞれのグループごとの時間特徴マップセットを生成することであって、

それぞれの時間畳み込みフィルタバンク内の時間畳み込みフィルタの学習された係数は、それぞれの時間畳み込みフィルタバンク内の時間畳み込みフィルタバンク間で変化する、ことと、を含む、人工知能ベースの方法。

13. それぞれのサイクルごとの空間特徴マップセットを、圧縮ネットワークのそれぞれの圧縮畳み込み層を介して別々に畳み込み、それぞれのシーケンシングサイクルに対し、サイクルごとの圧縮空間特徴マップセットを生成することを更に含む、項目12に記載の人工知能ベースの方法。

14. それぞれの圧縮畳み込み層内の圧縮畳み込みフィルタの学習された係数が、それぞれの圧縮畳み込み層内の圧縮畳み込み層間で変化する、項目13に記載の人工知能ベースの方法。

15. グループごとの時間特徴マップセットにおけるグループごとの時間特徴マップセットのそれぞれの重なり合うグループに、第2の時間畳み込み層のそれぞれの時間畳み込みフィルタバンクを使用して、畳み込みを行うことによって、時間畳み込みネットワークを介して、グループごとにグループごとの時間特徴マップセットを処理し、グループごとの時間特徴マップセットのそれぞれの重なり合うグループに対してそれぞれの更なるグループごとの時間特徴マップセットを生成することと、を更に含む、項目14に記載の人工知能ベースの方法。

16. 出力ネットワークを介して、最終時間畳み込み層によって生成された最終時間特徴マップセットを処理し、最終出力を生成することを更に含む、項目15に記載の人工知能ベースの方法。

17. 最終出力に基づいて、シーケンシングサイクルのシリーズにおいて、1つ以上のシーケンシングサイクルに対するベースコール予測を生成することを更に含む、項目16に記載の人工知能ベースの方法。

18. ベースコールの人工知能ベースの方法であって、方法が、

空間畳み込み層のそれぞれのシーケンスを介して、サイクルごとのシーケンス画像セットのウィンドウ内の、それぞれのサイクルごとのシーケンス画像セットを別々に畳み込むことによって、空間畳み込みネットワークを介して、サイクルごとに、シーケンシングランのシーケンシングサイクルのシリーズに対するサイクルごとのシーケンス画像セットのウィンドウを処理し、シーケンシングサイクルのシリーズ内のそれぞれのシーケンシングサイクルに対して、それぞれのサイクルごとの空間特徴マップセットを生成することであって、

第1の時間畳み込み層のそれぞれの時間畳み込みフィルタバンクを用いて、サイクルごとの空間特徴マップセットにおけるサイクルごとの空間特徴マップセットのそれぞれの重なり合うグループに畳み込むことによって、時間畳み込みネットワークを介して、サイクルごとの空間特徴マップセットをグループごとに処理し、サイクルごとの空間特徴マップセットのそれぞれ重なり合うグループに対する、それぞれのグループごとの時間特徴マップセットを生成することであって、

10

20

30

40

50

それぞれの時間畳み込みフィルタバンク内の時間畳み込みフィルタの学習された係数は、それぞれの時間畳み込みフィルタバンク内の時間畳み込みフィルタバンク間で変化する、人工知能ベースの方法。

19. 空間畳み込み層のそれぞれのシーケンスが、空間畳み込みフィルタバンクのそれぞれのシーケンスを有し、空間畳み込みフィルタバンクのそれぞれのシーケンスの空間畳み込みフィルタバンク内の空間畳み込みフィルタのが、空間畳み込み層の学習された係数それぞれのシーケンス内の空間畳み込み層のシーケンス間で共有される、項目18に記載の人工知能ベースの方法。

20. それぞれのサイクルごとの空間特徴マップセットを、圧縮ネットワークのそれぞれの圧縮畳み込み層を介して別々に畳み込み、それぞれのシーケンシングサイクルに対し、サイクルごとの圧縮空間特徴マップセットを生成することを更に含み、それぞれの圧縮畳み込み層内の圧縮畳み込みフィルタのが、学習された係数それぞれの圧縮畳み込み層内の圧縮畳み込み層間で変化する、項目18に記載の人工知能ベースの方法。

21. システムであって、

サイクルごとのシーケンス画像のウィンドウ内のそれぞれのサイクルごとのシーケンス画像に空間畳み込み層のそれぞれのシーケンスを適用するように構成された空間畳み込みネットワークを備え、

空間畳み込み層のそれぞれのシーケンスが、空間畳み込み層のシーケンスごとに異なっている空間畳み込みフィルタバンクのそれぞれのシーケンスを有する、システム。

22. システムであって、

時間畳み込みフィルタのそれぞれのセットを、空間特徴マップのそれぞれのスライドウィンドウに適用するように構成された第1の時間畳み込み層で構成された時間畳み込みネットワークを備え、

第1の時間畳み込み層内の時間畳み込みフィルタのそれぞれのセットが、時間畳み込みフィルタのセットごとに異なっている時間畳み込みフィルタを有する、システム。

23. 時間畳み込みネットワークが、第1の時間畳み込み層の後に続く第2の時間畳み込み層で構成され、第2の畳み込み層は、時間特徴マップのそれぞれのスライドウィンドウにそれぞれのセットの時間畳み込みフィルタを適用するように構成され、第2の時間畳み込み層内の時間畳み込みフィルタのそれぞれのセットが、時間畳み込みフィルタのセットごとに異なる時間畳み込みフィルタを有する、項目22に記載のシステム。

【0242】

スキップ

1. システムであって、

シーケンシングランのシーケンシングサイクルのシリーズに対して、サイクルごとのシーケンス画像セットのウィンドウを、それぞれの空間処理パイプラインを介して、サイクルごとのシーケンス画像セットのウィンドウ内の、それぞれのサイクルごとのシーケンス画像セットを、サイクルごとに別々に処理することによって処理し、それぞれの空間処理パイプラインは、それぞれのサイクルごとのシーケンス画像セットを、空間畳み込み層のそれぞれのシーケンスを介して、畳み込むように構成され、シーケンシングサイクルのシリーズ内のそれぞれのシーケンシングサイクルに対して、それぞれのサイクルごとの空間特徴マップセットを生成するように構成されている空間畳み込みネットワークと

空間畳み込みネットワークに接続され、かつ空間畳み込み層の空間特徴マップそれぞれのシーケンス内の空間畳み込み層間でバスを形成するように構成されたバスネットワークであって、バスが、特定のシーケンシングサイクルに対する空間畳み込み層の特定のシーケンス内の2つ以上の空間畳み込み層によって生成されたそれぞれのサイクルごとのセットを組み合わせて、組み合わせられた、サイクルごとの空間特徴マップセットを形成し、組み合わせられた、サイクルごとの空間特徴マップセットを、空間畳み込み層の特定のシーケンス内の別の空間畳み込み層への入力として提供するように構成されている。バスネットワークと、を備える、システム。

2. 2つ以上の空間畳み込み層が、第1の空間畳み込み層及び第3の空間畳み込み層を

含み、第1の空間畳み込み層が、第1のサイクルごとの空間特徴マップセットを生成し、第1の空間畳み込み層は、第1のサイクルごとの空間特徴マップセットを、第2の空間畳み込み層への入力として提供し、第2の空間畳み込み層は、第1のサイクルごとの空間特徴マップセットを処理し、第2のサイクルごとの空間特徴マップセットを生成し、

第2の空間畳み込み層は、第3の空間畳み込み層への入力として、第2のサイクルごとの空間特徴マップセットを提供し、第3の空間畳み込み層は、第2のサイクルごとの空間特徴マップセットを処理し、第3のサイクルごとの空間機能マップセットを生成する、項目1に記載のシステム。

3. バスが、第1の空間特徴マップセット及び第3のサイクルごとの空間特徴マップセットを、組み合わせて、組み合わせられた、サイクルごとの空間特徴マップセットになるように更に構成されている、項目2に記載のシステム。

10

4. 別の空間畳み込み層が、空間畳み込み層の特定のシーケンス内の第3の空間畳み込み層の直後に続く第4の空間畳み込み層であり、第4の空間畳み込み層は、組み合わせられた、サイクルごとの空間特徴マップセットを入力として処理する、項目3に記載のシステム。

5. 2つ以上の空間畳み込み層が、第1の空間畳み込み層及び第7の空間畳み込み層を含み、第3の空間畳み込み層が、第3のサイクルごとの空間特徴マップセットを、第4の空間畳み込み層への入力として提供し、第4の空間畳み込み層は、第3のサイクルごとの空間特徴マップセットを処理し、第4のサイクルごとの空間特徴マップセットを生成し、第4の空間畳み込み層は、第5の空間畳み込み層への入力として、第4のサイクルごとの空間特徴マップセットを提供し、第5の空間畳み込み層は、第4のサイクルごとの空間特徴マップセットを処理し、第5のサイクルごとの空間機能マップセットを生成する、項目2に記載のシステム。

20

6. バスが、第1の空間特徴マップセット及び第5の空間特徴マップセットを、組み合わせて、組み合わせられた、サイクルごとの空間特徴マップセットになるように更に構成されている、項目5に記載のシステム。

7. 別の空間畳み込み層が、空間畳み込み層の特定のシーケンス内の第5の空間畳み込み層の直後に続く第6の空間畳み込み層であり、第6の空間畳み込み層が、組み合わせられた、サイクルごとの空間特徴マップセットを入力として処理する、項目6に記載のシステム。

30

8. 2つ以上の空間畳み込み層が、第1の空間畳み込み層、第3の空間畳み込み層、及び第5の空間畳み込み層を含み、バスが、第1のサイクルごとの空間特徴マップセット、第3のサイクルごとの空間特徴マップセット、及び第5のサイクルごとの空間特徴マップセットを組み合わせて、組み合わせられた、サイクルごとの空間特徴マップセットにするように更に構成されている、項目5に記載のシステム。

9. 別の空間畳み込み層が、入力として、組み合わせられた、サイクルごとの空間特徴マップセットを処理する第6の空間畳み込み層である、項目8に記載のシステム。

10. バスが、特定のシーケンシングサイクルに対するサイクルごとのシーケンス画像セットを、第1の空間畳み込み層への入力として提供し、第3のサイクルごとの空間特徴マップセットを組み合わせて、組み合わせられた、サイクルごとの空間特徴マップセットにするように更に構成されている、項目1に記載のシステム。

40

11. 別の空間畳み込み層が、入力として、組み合わせられた、サイクルごとの空間特徴マップセットを処理する第4の空間畳み込み層である、項目10に記載のシステム。

12. バスネットワークが、受け入れ側のサイクルごとの空間特徴マップセットと組み合わせられる、入来するサイクルごとの空間特徴マップセットの空間及び深度次元数を修正するように構成された次元適合性ロジックを含み、組み合わせられたサイクルごとの空間特徴マップセットを生成するように更に構成されている、項目1に記載のシステム。

13. 次元適合性ロジックが、畳み込み、プール、又は平均化を含む次元数削減動作である、項目12に記載のシステム。

14. バスネットワークが、受け入れ側のサイクルごとの空間特徴マップセットと組み

50

合わされる、入来するサイクルごとの空間特徴マップセットの特徴値をスケールするように構成されたスケーリングロジックを含み、サイクルごとの空間特徴マップセットを生成するように更に構成されている、項目12に記載のシステム。

15. 第1の時間畳み込み層のそれぞれの時間畳み込みフィルタバンクを用いて、サイクルごとの空間特徴マップセットにおけるサイクルごとの空間特徴マップセットのそれぞれの重なり合うグループに畳み込むことによって、サイクルごとの空間特徴マップセットをグループごとに処理して、サイクルごとの空間特徴マップセットの、それぞれ重なり合うグループに対する、それぞれのグループごとの時間特徴マップセットを生成するように構成されている時間畳み込みネットワークを含むように更に構成されている、項目1に記載のシステム。

10

16. 時間畳み込みネットワークに接続され、かつ時間畳み込み層のそれぞれのシーケンス内の時間畳み込み層間でバスを形成するように構成されたバスネットワークであって、バスが、特定のシーケンシングサイクルに対する時間畳み込み層の特定のシーケンス内の2つ以上の時間畳み込み層によって生成されたそれぞれのサイクルごとの時間特徴マップセットを組み合わせて、組み合わせられた、サイクルごとの時間特徴マップセットを形成し、組み合わせられた、サイクルごとの時間特徴マップセットを、時間畳み込み層の特定のシーケンス内の別の時間畳み込み層への入力として提供するように構成されているバスネットワークを備えるように更に構成されている、項目15に記載のシステム。

17. 人工知能ベースの方法であって、

シーケンシングランのシーケンシングサイクルのシリーズに対して、サイクルごとのシーケンス画像セットのウィンドウを、サイクルごとのシーケンス画像セットのウィンドウ内の、それぞれのサイクルごとのシーケンス画像セットを、サイクルごとに、それぞれの空間的処理パイプラインを通じて別々に処理することによって、空間畳み込みネットワークを介して処理することであって、それぞれのサイクルごとのシーケンス画像セットを、空間畳み込み層のそれぞれのシーケンスを介して、畳み込んで、シーケンシングサイクルのシリーズ内のそれぞれのシーケンシングサイクルに対して、それぞれのサイクルごとの空間特徴マップセットを生成することを含む、ことと、

20

特定のシーケンシングサイクルに対する空間畳み込み層の特定のシーケンス内の2つ以上の空間畳み込み層によって生成されたそれぞれのサイクルごとの空間特徴マップセットを組み合わせ、組み合わせられた、サイクルごとの空間特徴マップセットを形成し、組み合わせられた、サイクルごとの空間特徴マップセットを、空間畳み込み層の特定のシーケンス内の別の空間畳み込み層への入力として提供することと、を含む、人工知能ベースの方法。

30

18. 2つ以上の空間畳み込み層が、第1の空間畳み込み層及び第3の空間畳み込み層を含み、第1の空間畳み込み層が、第1のサイクルごとの空間特徴マップセットを生成し、第1の空間畳み込み層が、第1のサイクルごとの空間特徴マップセットを、第2の空間畳み込み層への入力として提供し、第2の空間畳み込み層が、第1のサイクルごとの空間特徴マップセットを処理し、第2のサイクルごとの空間特徴マップセットを生成し、

第2の空間畳み込み層が、第3の空間畳み込み層への入力として、第2のサイクルごとの空間特徴マップセットを提供し、第3の空間畳み込み層が、第2のサイクルごとの空間特徴マップセットを処理し、第3のサイクルごとの空間機能マップセットを生成する、項目17に記載の人工知能ベースの方法。

40

19. バスが、第1の空間特徴マップセット及び第3のサイクルごとの空間特徴マップセットを、組み合わせて、組み合わせられた、サイクルごとの、空間特徴マップセットになるように更に構成されている、項目18に記載の人工知能ベースの方法。

20. 別の空間畳み込み層が、空間畳み込み層の特定のシーケンス内の第3の空間畳み込み層の直後に続く第4の空間畳み込み層であり、第4の空間畳み込み層が、組み合わせられた、サイクルごとの空間特徴マップセットを入力として処理する、項目19に記載の人工知能ベースの方法。

【0243】

上で説明される方法の他の実施態様は、上述の方法のいずれかを実行するためにプロセ

50

ッサによって実行可能な命令を記憶する非一時的コンピュータ可読記憶媒体を含むことができる。このセクションで説明される方法の更に別の実施態様は、メモリと、メモリ内に記憶された命令を実行するように動作可能な1つ以上のプロセッサとを含むシステムを含むことができ、上記の方法のいずれかを実行することができる。

【符号の説明】

【0244】

- 1906 システムコントローラ
- 3300 コンピュータシステム
- 3310 記憶サブシステム
- 3332 ランダムアクセスメモリ
- 3334 読み取り専用メモリ
- 3336 ファイルストレージサブシステム
- 3338 ユーザーインターフェース入力デバイス
- 3355 バスサブシステム
- 3372 中央演算処理装置
- 3374 ネットワークインターフェースサブシステム
- 3376 ユーザーインターフェース出力デバイス
- 3378 深層学習プロセッサ

10

20

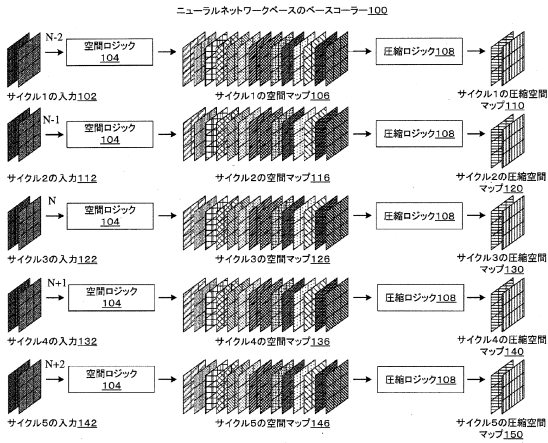
30

40

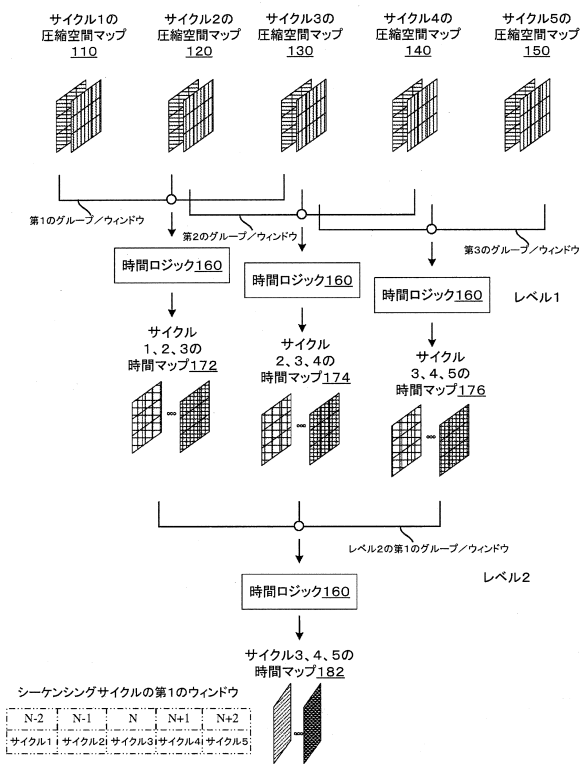
50

【図面】

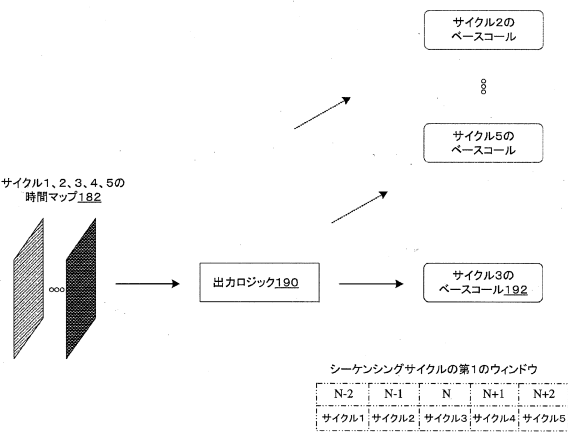
【図 1 A】



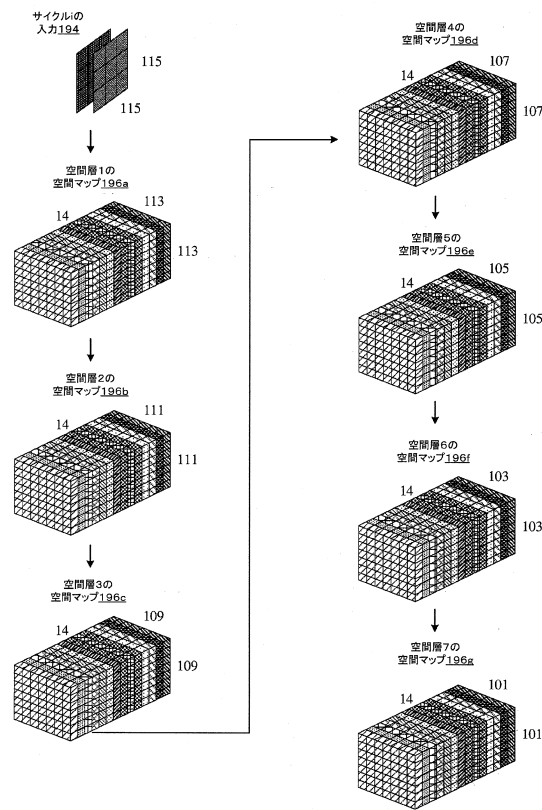
【図 1 B】



【図 1 C】



【図 1 D】



10

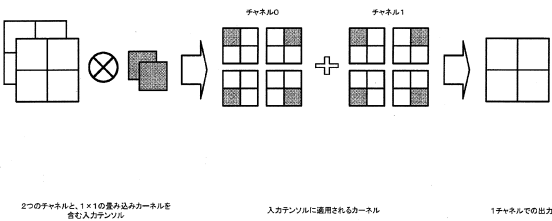
20

30

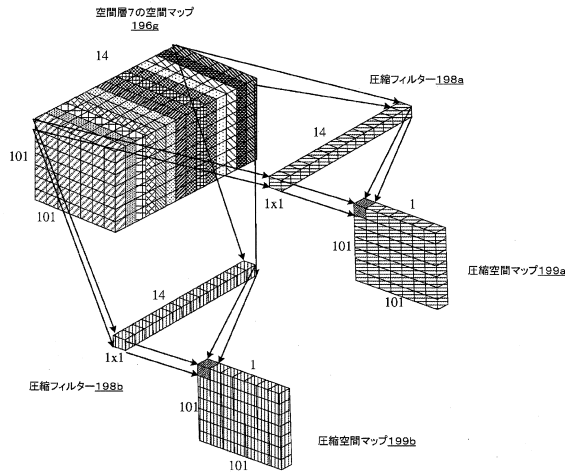
40

50

【図 1 E】

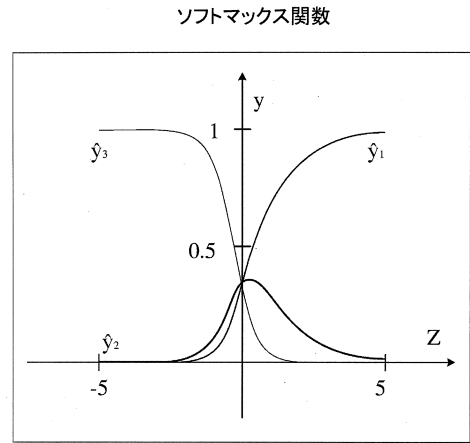


【図 1 F】

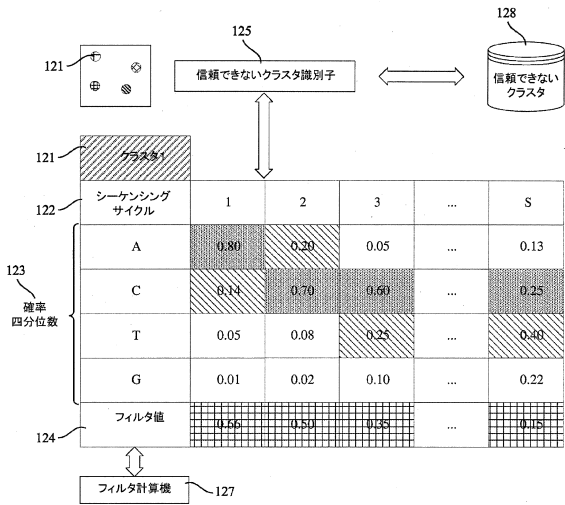


10

【図 1 G】



【図 1 H】



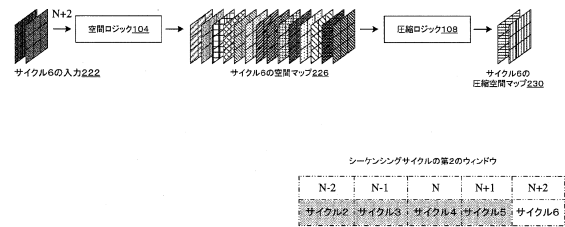
20

30

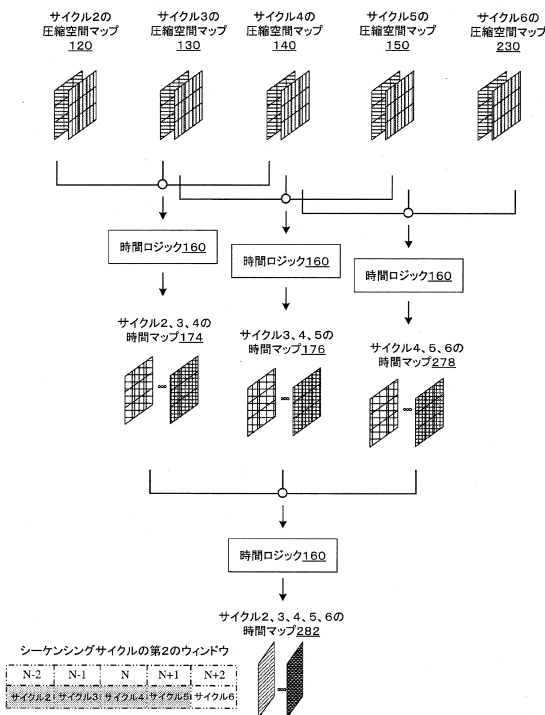
40

50

【図 2 A】



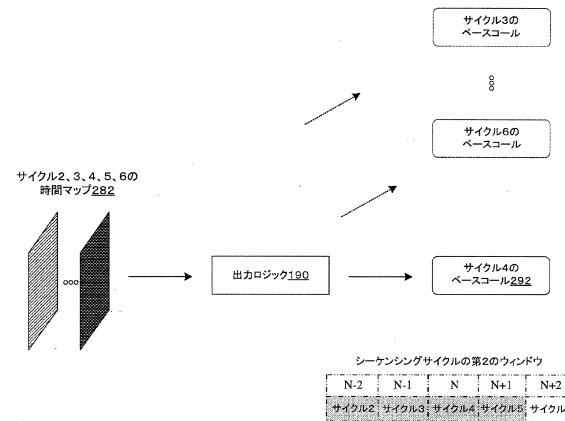
【図 2 B】



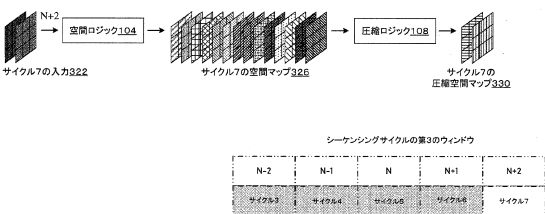
10

20

【図 2 C】



【図 3 A】

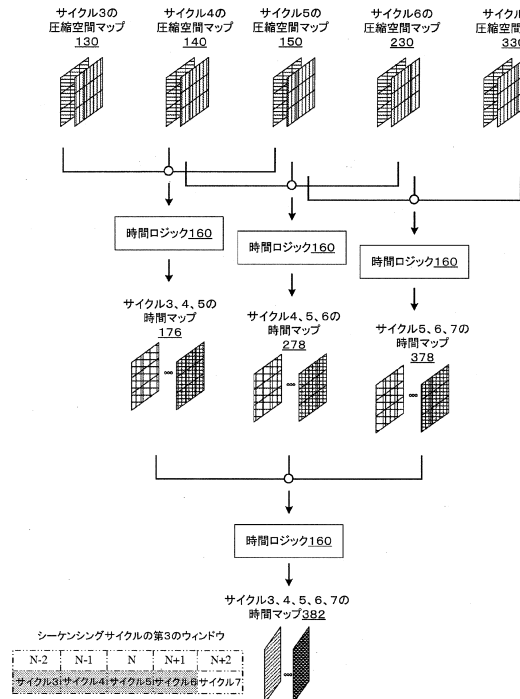


30

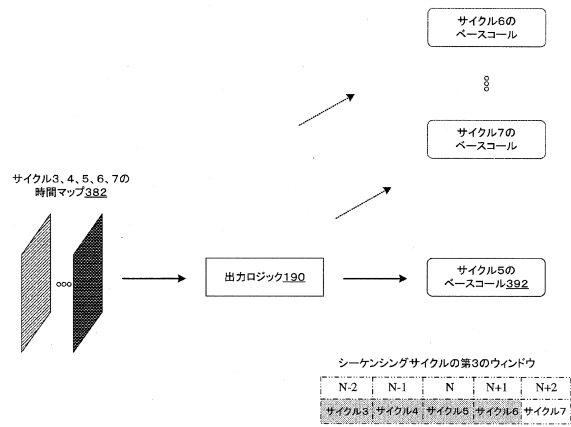
40

50

【図 3 B】



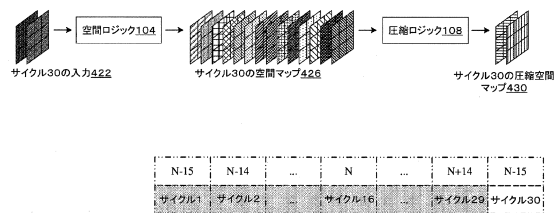
【図 3 C】



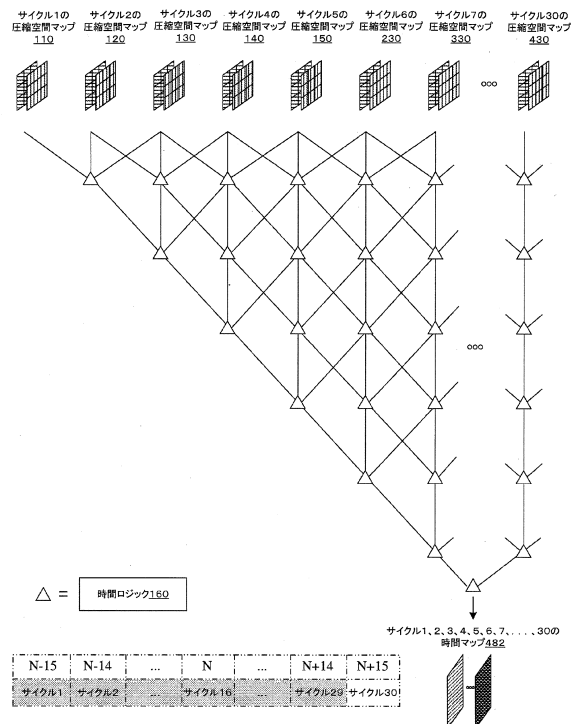
10

20

【図 4 A】



【図 4 B】

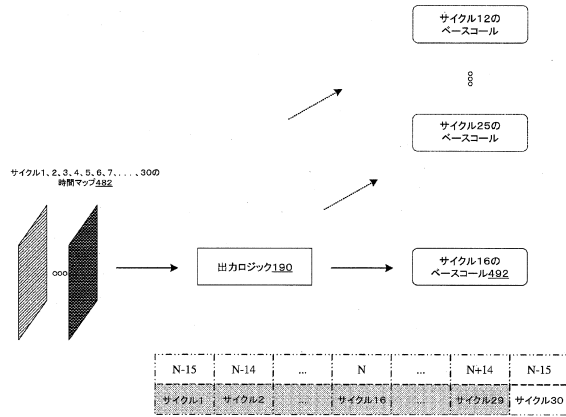


30

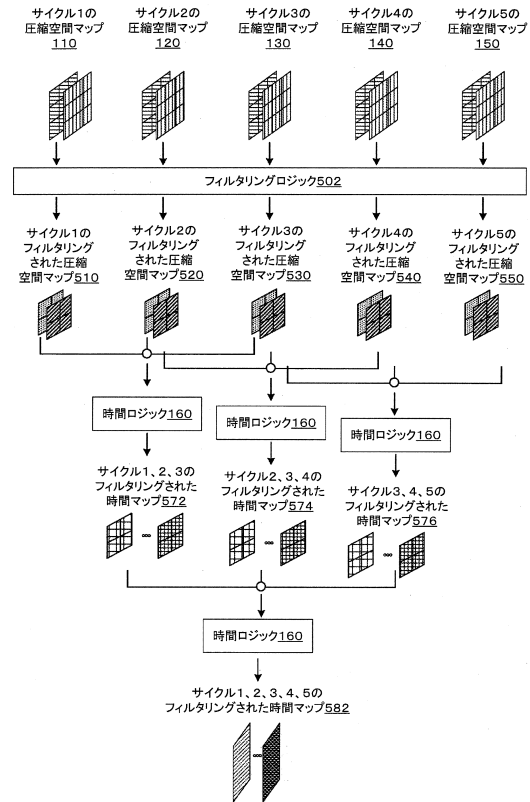
40

50

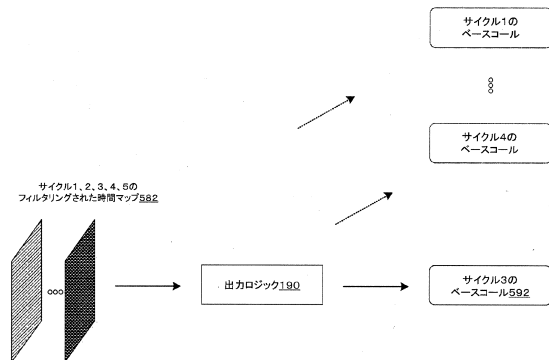
【図 4 C】



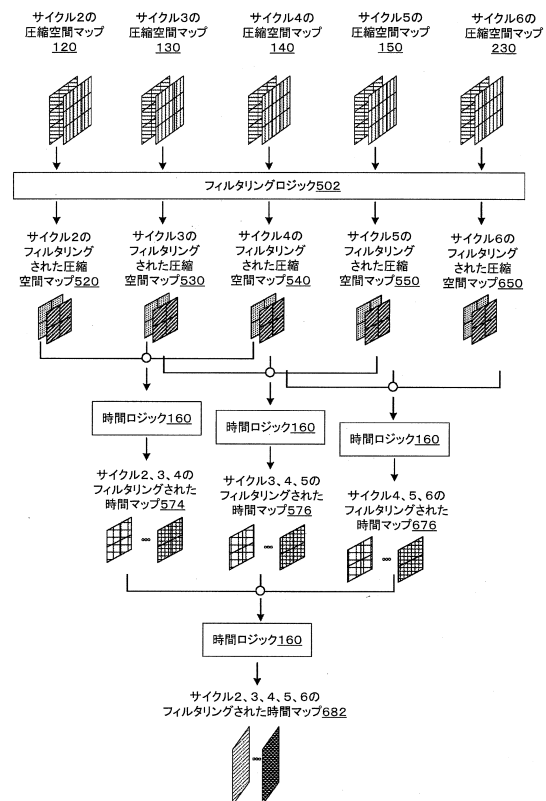
【図 5 A】



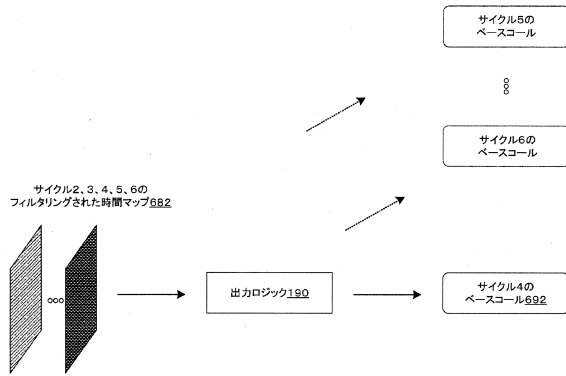
【図 5 B】



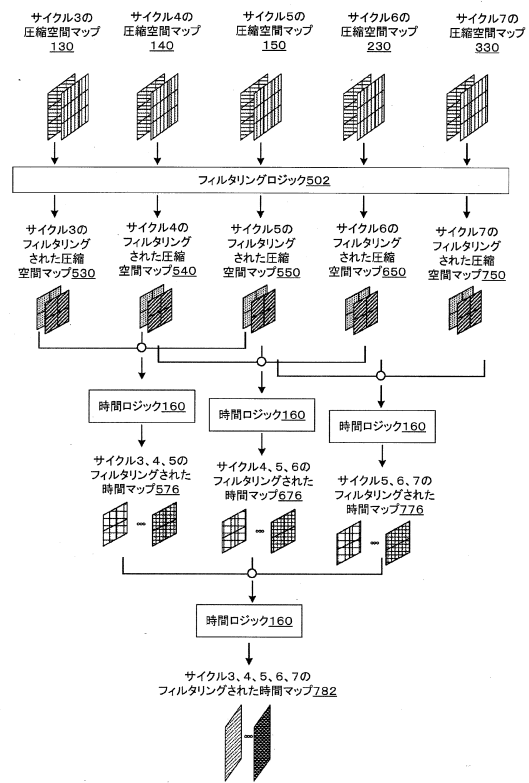
【図 6 A】



【図 6 B】



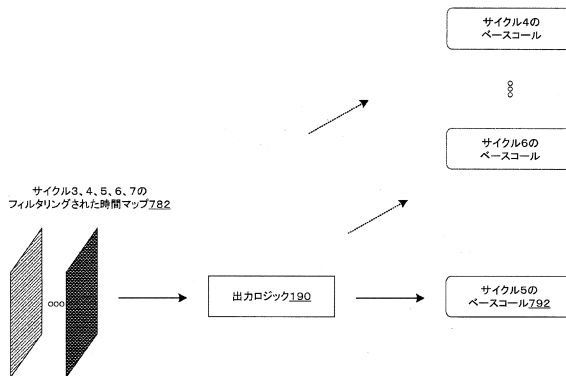
【図 7 A】



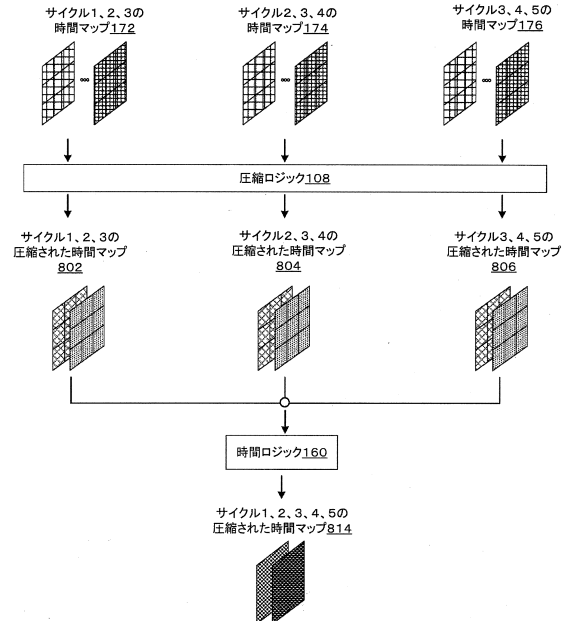
10

20

【図 7 B】



【図 8 A】

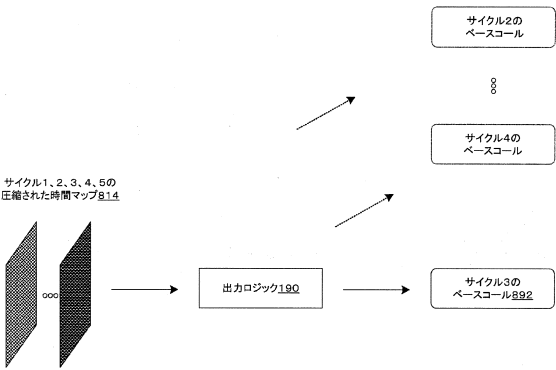


30

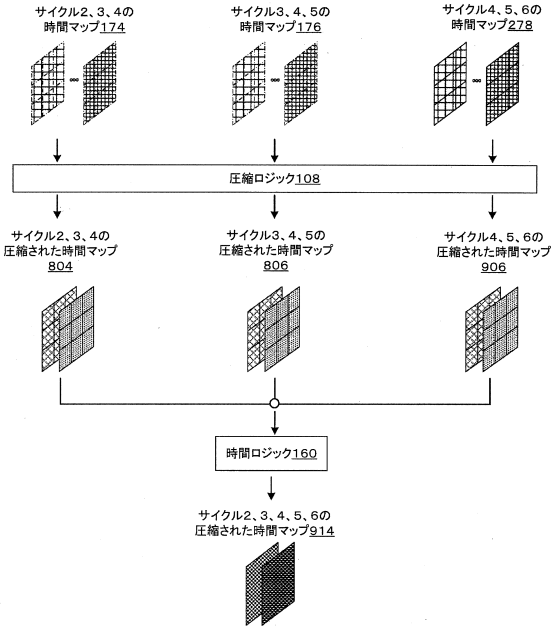
40

50

【図 8 B】



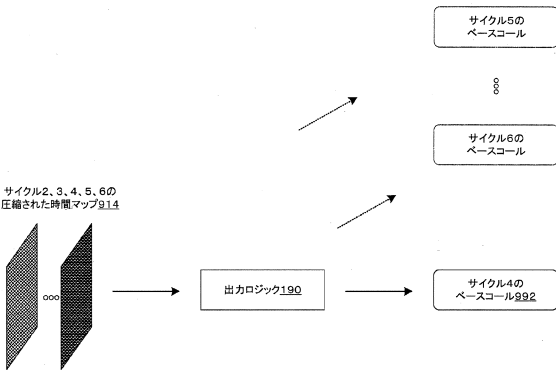
【図 9 A】



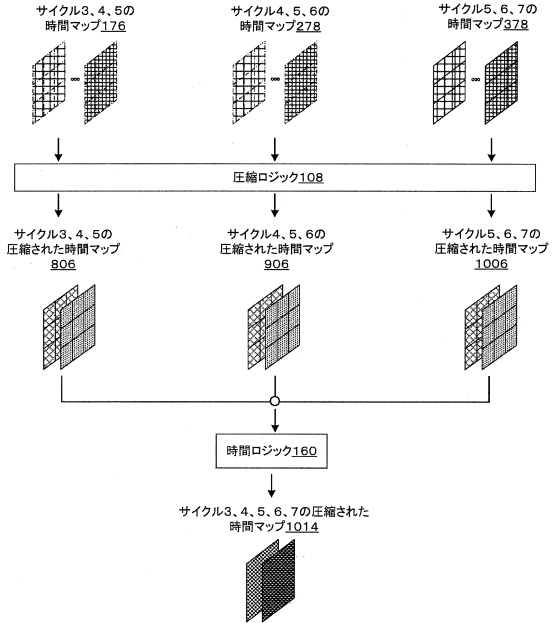
10

20

【図 9 B】



【図 10 A】

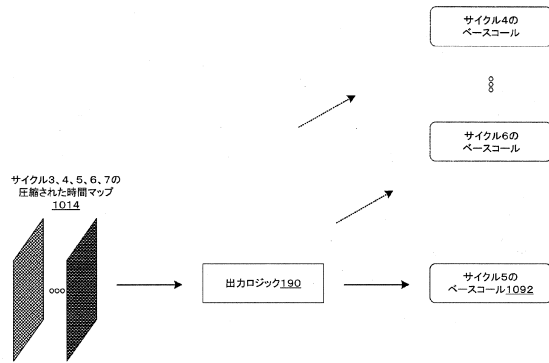


30

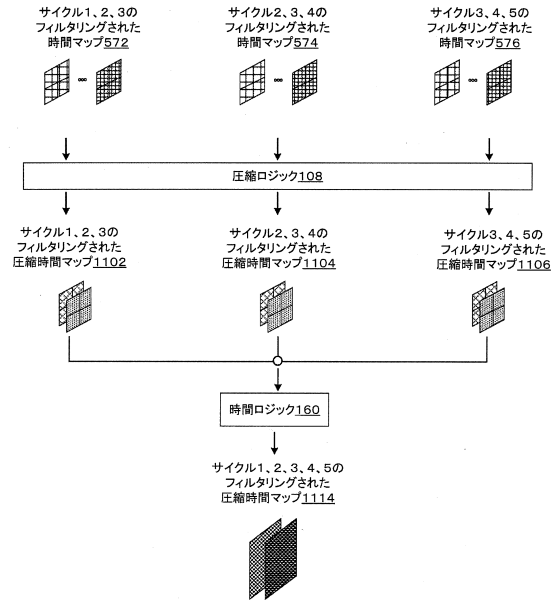
40

50

【図10B】

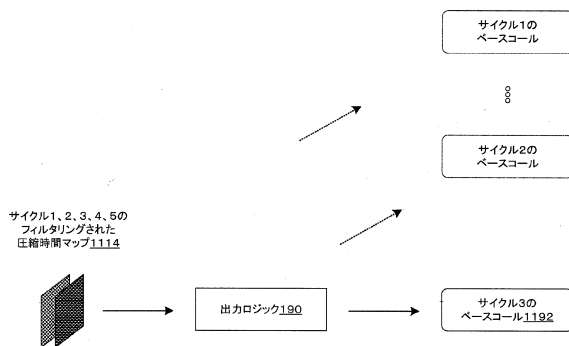


【図11A】

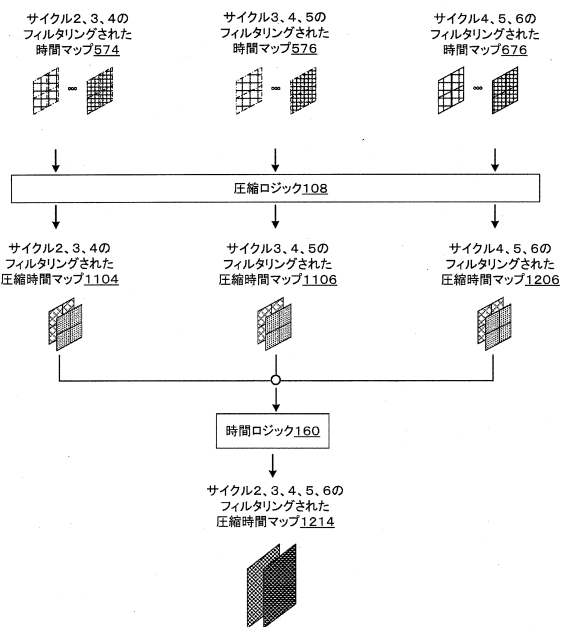


10

【図11B】



【図12A】



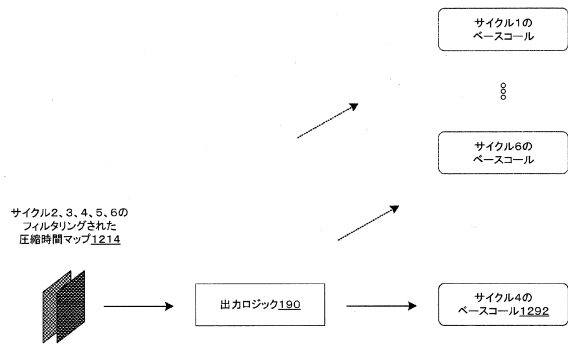
20

30

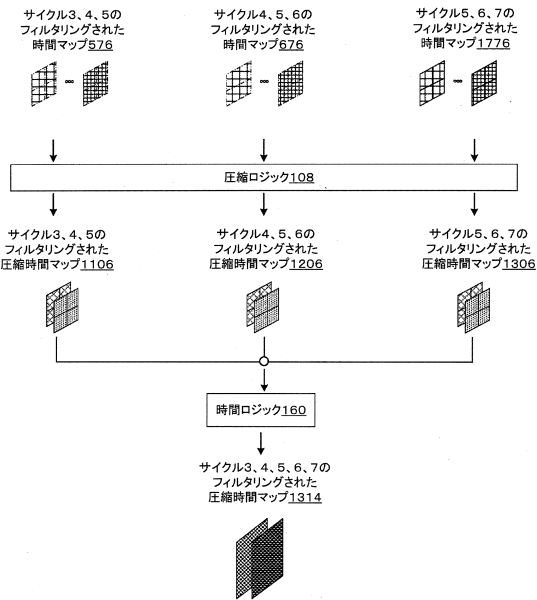
40

50

【図 1 2 B】

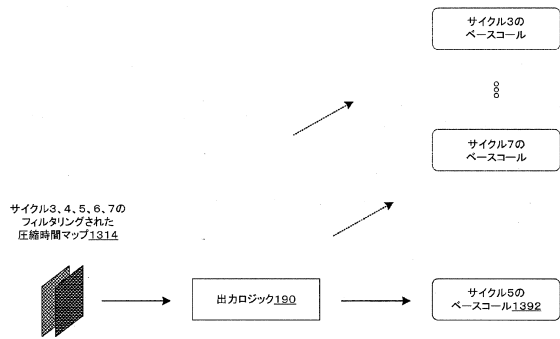


【図 1 3 A】

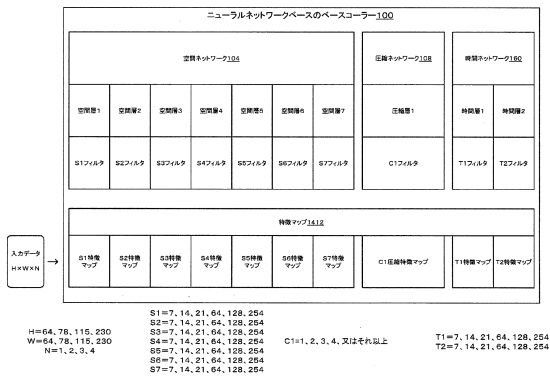


10

【図 1 3 B】



【図 1 4】

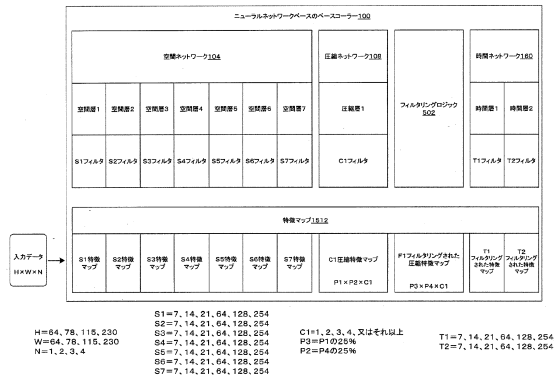


30

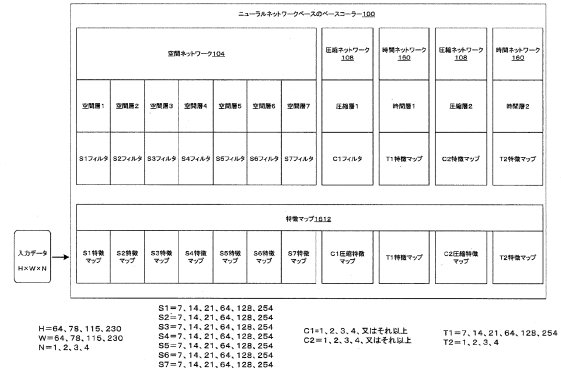
40

50

【図 15】

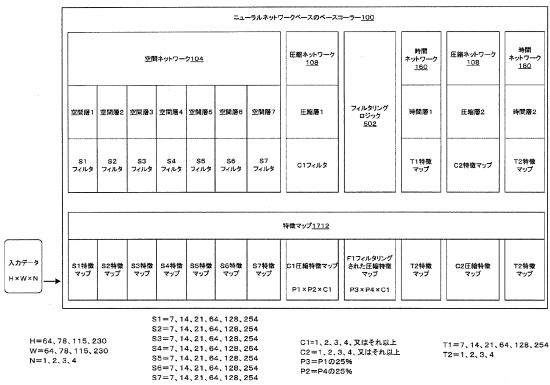


【図 16】

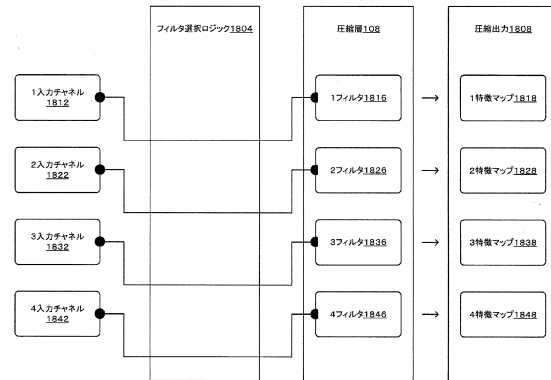


10

【図 17】



【圖 18】



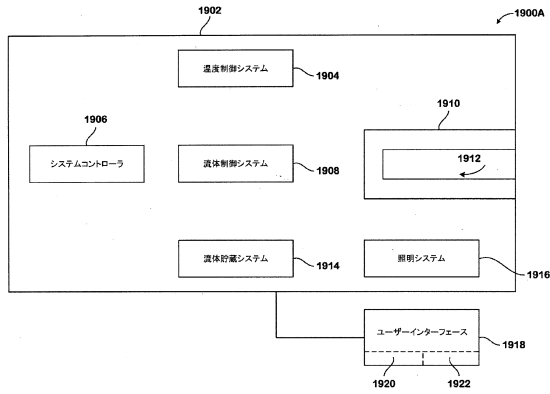
20

30

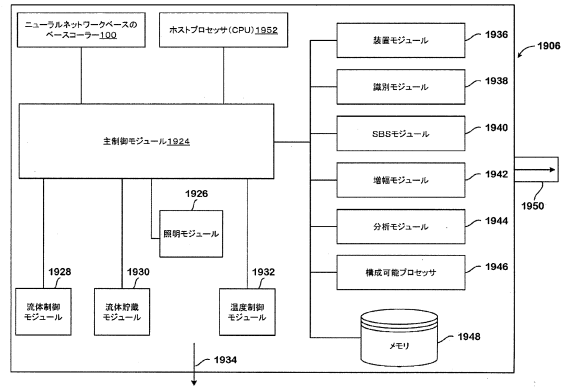
40

50

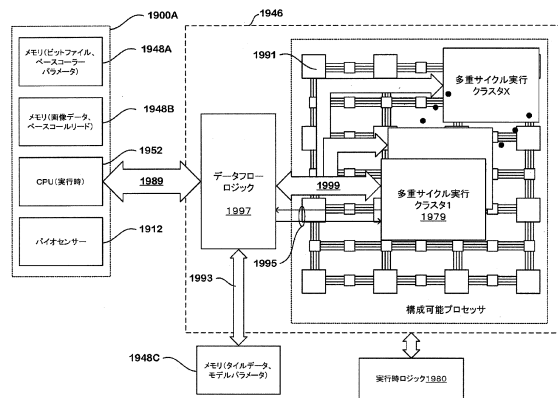
【図 19 A】



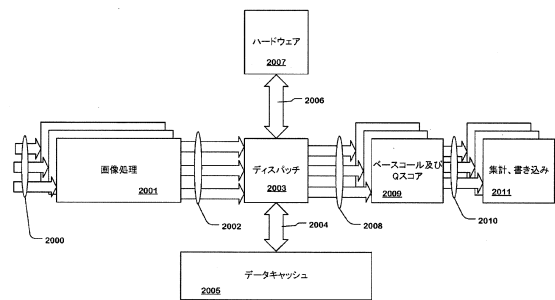
【図 19 B】



【図 19 C】



【図 20 A】



10

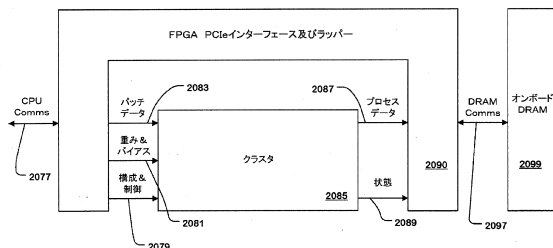
20

30

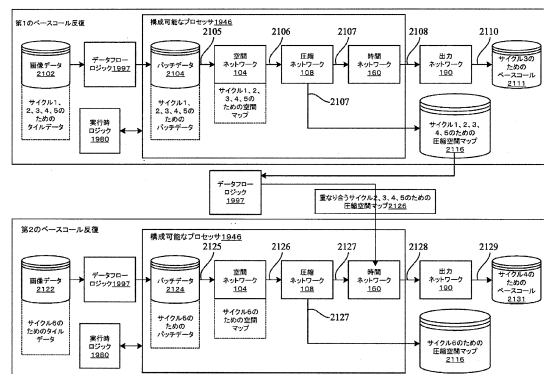
40

50

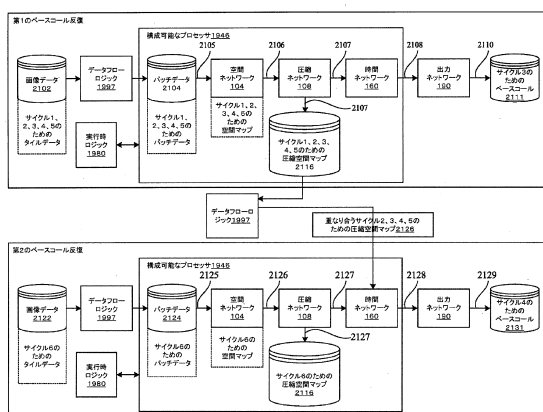
【図 20 B】



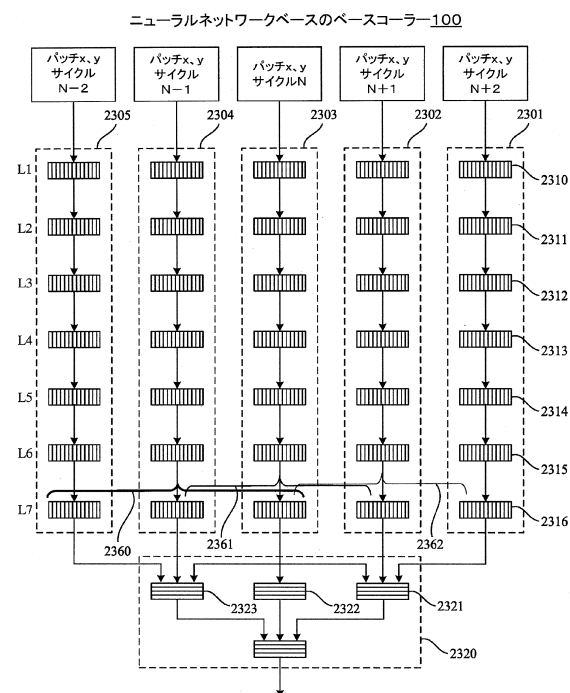
【図 2 1】



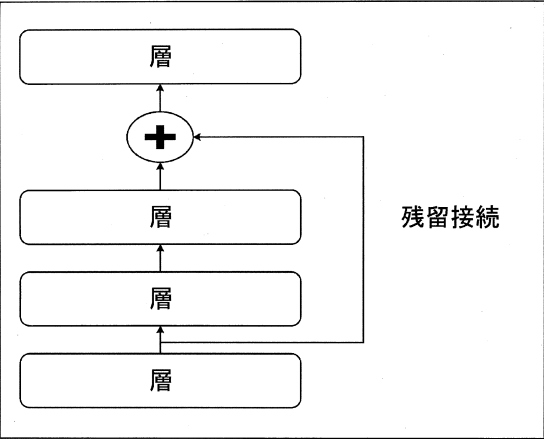
【圖 2 2】



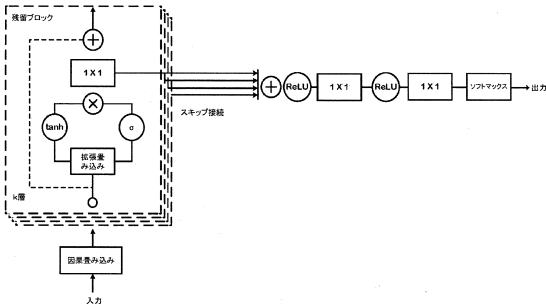
【圖 23】



【図 2 4 A】

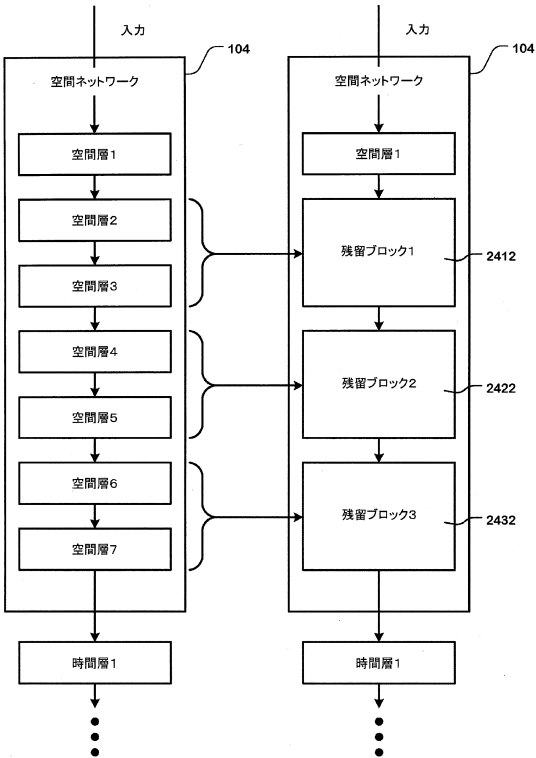


【図 2 4 B】

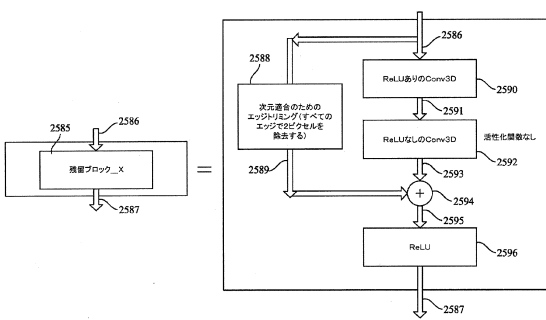


10

【図 2 4 C】



【図 2 5 A】



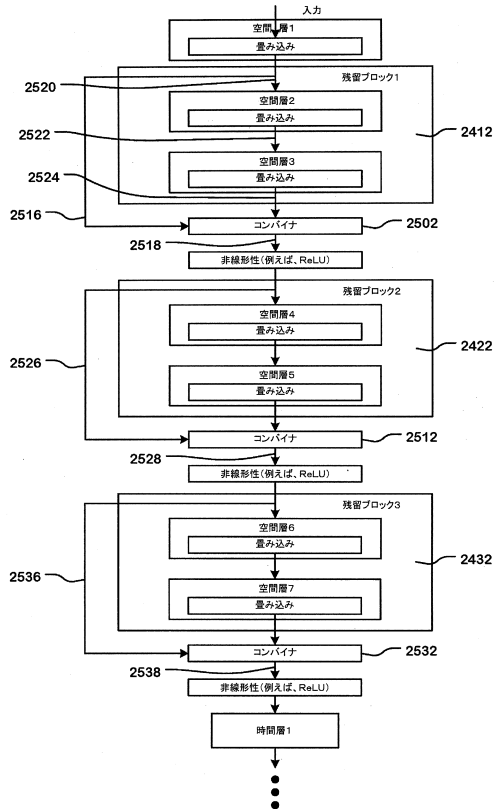
20

30

40

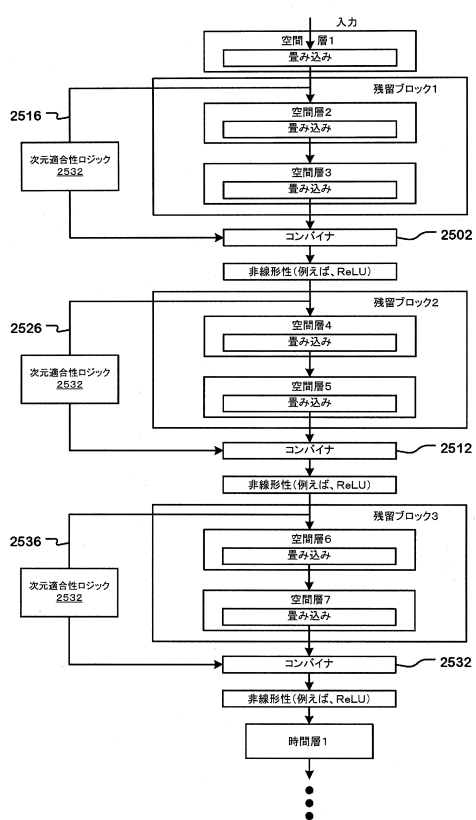
50

【図 25 B】



104

【図 25 C】

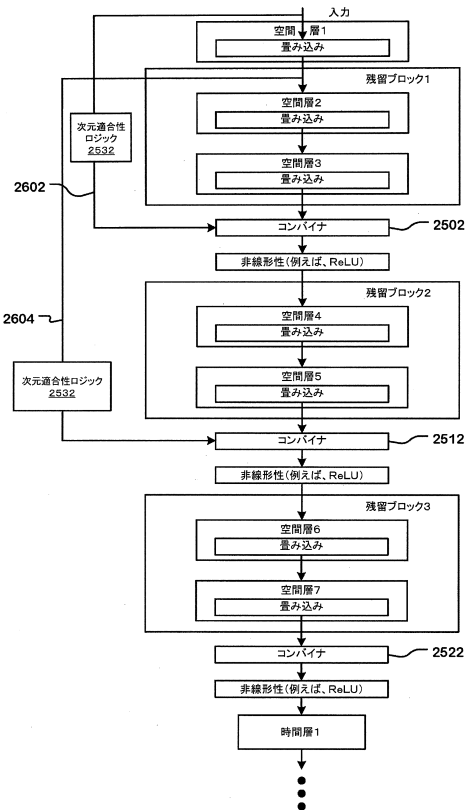


104

10

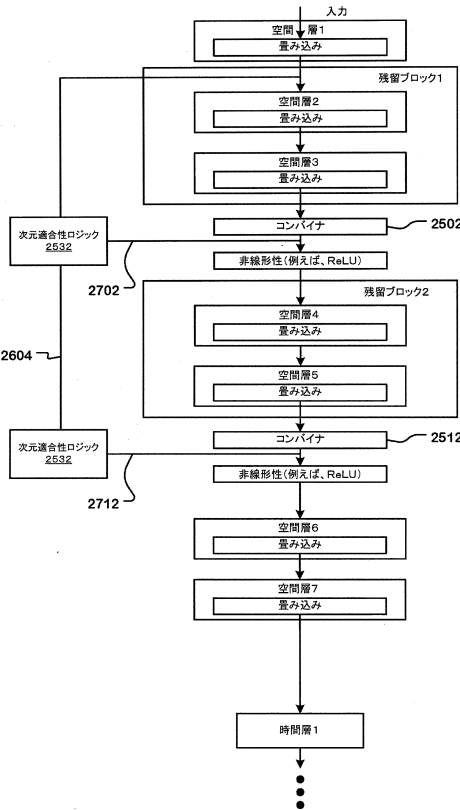
20

【図 26】



104

【図 27】



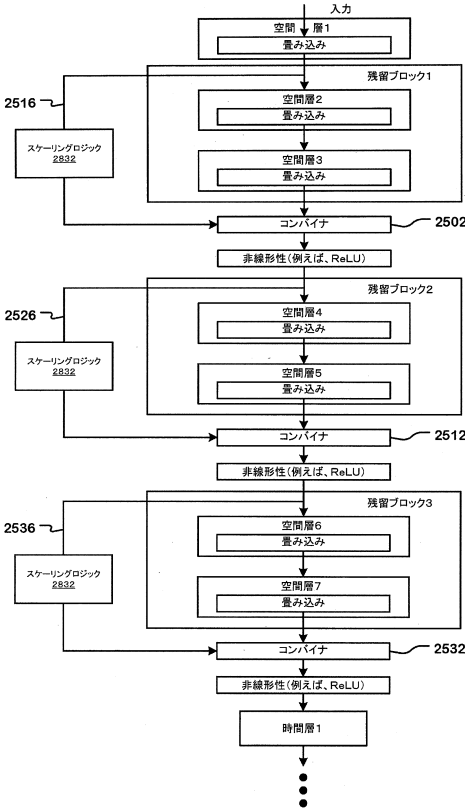
104

30

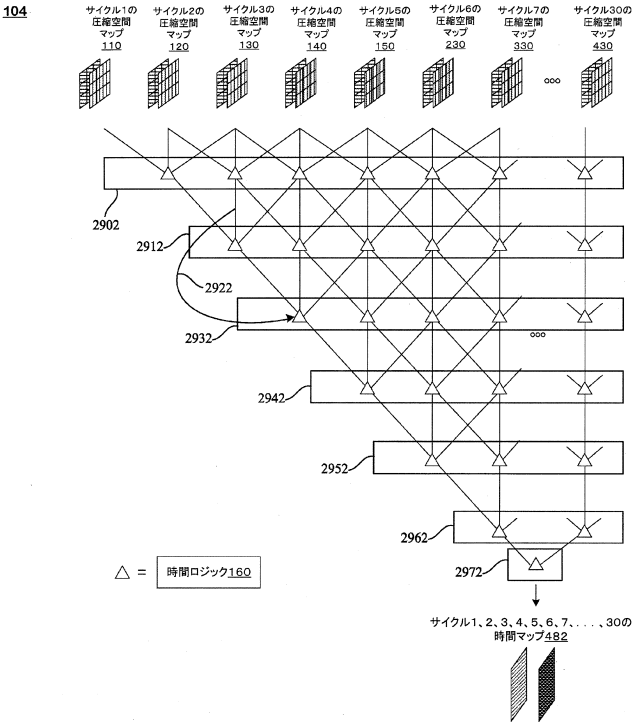
40

50

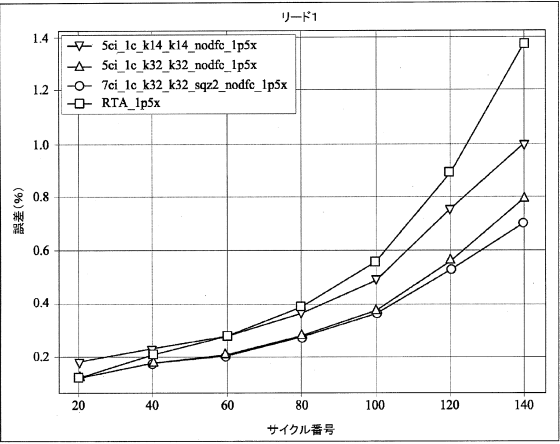
【図 28】



【図 29】



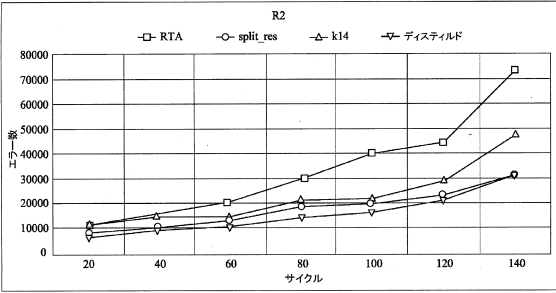
【図 30】



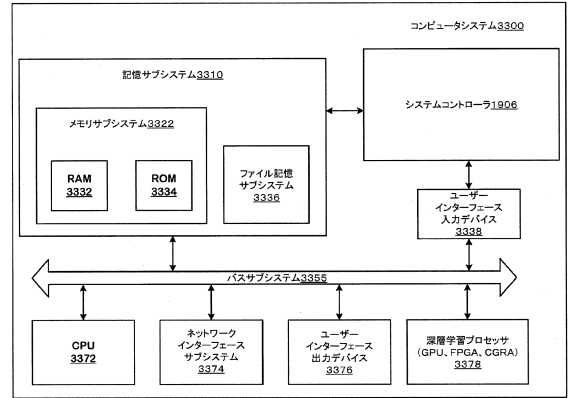
【図 31】

	RAM使用量 (GB)	DRAM使用量 (GB)
圧縮ロジックなし10B	90.81344604	0.752478838
圧縮ロジックあり10B	44.95265579	0.826221764

【図 3 2】



【図 3 3】



10

20

30

40

50

フロントページの続き

(33)優先権主張国・地域又は機関

米国(US)

(31)優先権主張番号 17/179,395

(32)優先日 令和3年2月18日(2021.2.18)

(33)優先権主張国・地域又は機関

米国(US)

(31)優先権主張番号 17/180,480

(32)優先日 令和3年2月19日(2021.2.19)

(33)優先権主張国・地域又は機関

米国(US)

(31)優先権主張番号 17/180,513

(32)優先日 令和3年2月19日(2021.2.19)

(33)優先権主張国・地域又は機関

米国(US)

(72)発明者 ギャヴィン・デレク・パーナビー

アメリカ合衆国・カリフォルニア・92122・サン・ディエゴ・イルミナ・ウェイ・5200

(72)発明者 アニンディタ・ダッタ

アメリカ合衆国・カリフォルニア・92122・サン・ディエゴ・イルミナ・ウェイ・5200

(72)発明者 ドルナ・カシフハギギ

アメリカ合衆国・カリフォルニア・92122・サン・ディエゴ・イルミナ・ウェイ・5200

(72)発明者 キショール・ジャガナタン

アメリカ合衆国・カリフォルニア・92122・サン・ディエゴ・イルミナ・ウェイ・5200

(72)発明者 アミラリ・キア

アメリカ合衆国・カリフォルニア・94401・サン・マテオ・ノース・エル・カミノ・リアル・811・アパートメント・302

審査官 多賀 実

(56)参考文献 米国特許出願公開第2019/0237160 (US, A1)

Jason Brownlee, "A Gentle Introduction to 1×1 Convolutions to Manage Model Complexity", MachineLearningMastery.com [online], 2019年07月05日, [検索日 2025.01.17], インターネット: <URL: <https://machinelearningmastery.com/introduction-to-1x1-convolutions-to-reduce-the-complexity-of-convolutional-neural-networks/>>

(58)調査した分野 (Int.Cl., DB名)

G06N 3/00-99/00

G06T 1/00-1/40

G06T 7/00

G06V 10/70-10/98

G01N 21/62-21/65

C12Q 1/00-3/00

G16B 5/00-99/00