

「A I時代の教育」 - AIの過去・現在・未来 -

第4回 人間脳の謎と深層学習の魔法

目を持ったコンピュータが見せる未知の領域

2023年 11月 16日

澤井 進

岐阜女子大学特任教授、(公財)学情研・専務理事
A I時代の教育学会・理事、教育クラウド推進協議会
博士 (知識科学)

無断転載禁止 ALL RIGHTS RESERVED, COPYRIGHT (C) SAWAI. 2023

目次

- 4.1 第3次AIブーム
「機械学習、深層学習の時代」
- 4.2 人間の脳の研究と機械学習
- 4.3 画像認識「目を持ったAI」
- 4.4 畳み込みニューラルネットワーク (CNN)
- 4.5 リカレントニューラルネットワーク
(RNN)
- 4.6 小テスト

【学習目標】

第3次AIブームでは「**深層学習** (ディープラーニング)」がAIを大きく変え、1) 3つの**機械学習** (教師あり/教師なし/強化学習)、2) **畳み込みニューラルネットワーク** (CNN)や3) **リカレントニューラルネットワーク** (RNN)等が登場したことを理解する。

4.1 第3次AIブーム

「機械学習、深層学習の時代」(1)

1990年～

- ー 情報検索:米国政府主導によるTREC国際会議(1992)
- ー データマイニング:国際会議やジャーナルの立上り
- ー 検索エンジン:Google(1998)

2000年～

- ー ウェブの広がり:1995年には1万サイト、
2006年には1億サイト
- ー ビッグデータ
- ー 大量のデータを用いた機械学習の実用化

2006年～第三次AIブーム

4.1 第3次AIブーム

「機械学習、深層学習の時代」(2)

●コンピュータ黎明期からあったニューラルネットワークの多層化の発想

⇒莫大な計算コストが問題となって長らく低迷していた。

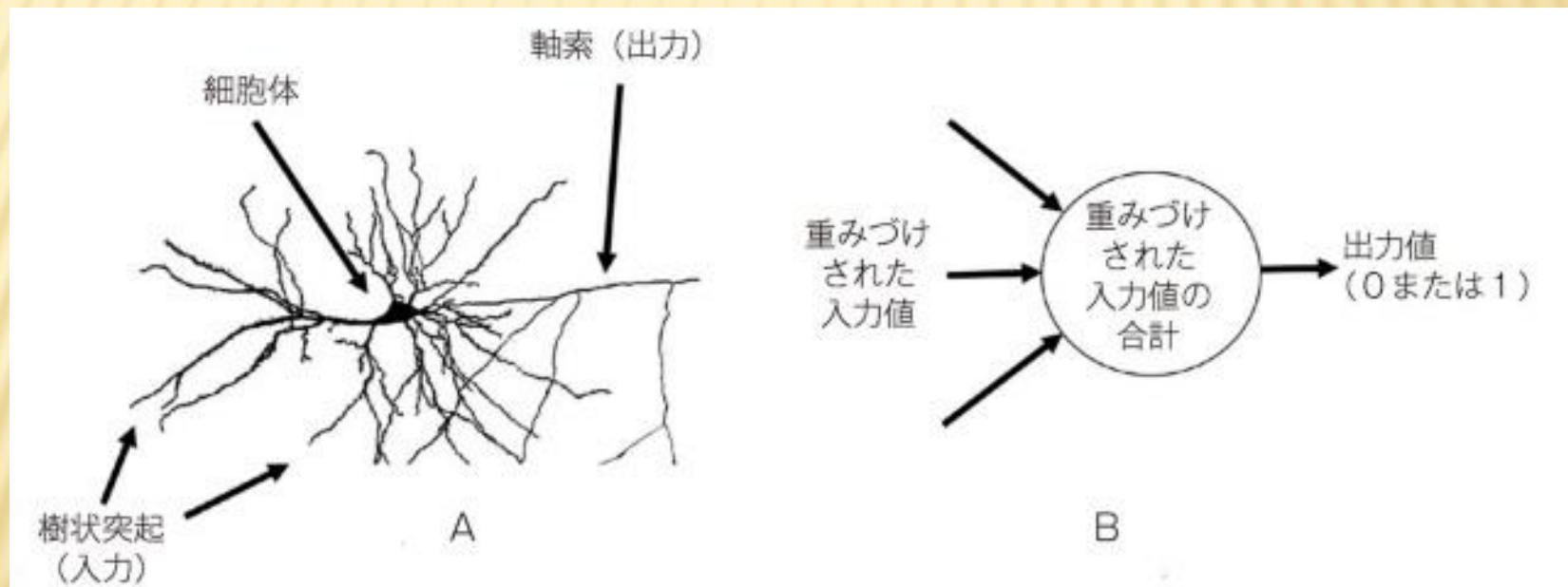
例えば、コンピュータの父の1人と言われる J. V. ノイマンは、著書「人工頭脳と自己増殖」や「自己増殖オートマトンの理論」等で、人間の脳やニューロン（神経細胞）の動きを、自己増殖する自動機械（オートマトン）のプログラムとして実現する方法を記述している。

⇒バーセプトロンの登場

4.1 第3次AIブーム

「機械学習、深層学習の時代」(2)

●1958年心理学者のフランク・ローゼンブラットが開発した「パーセプトロン」

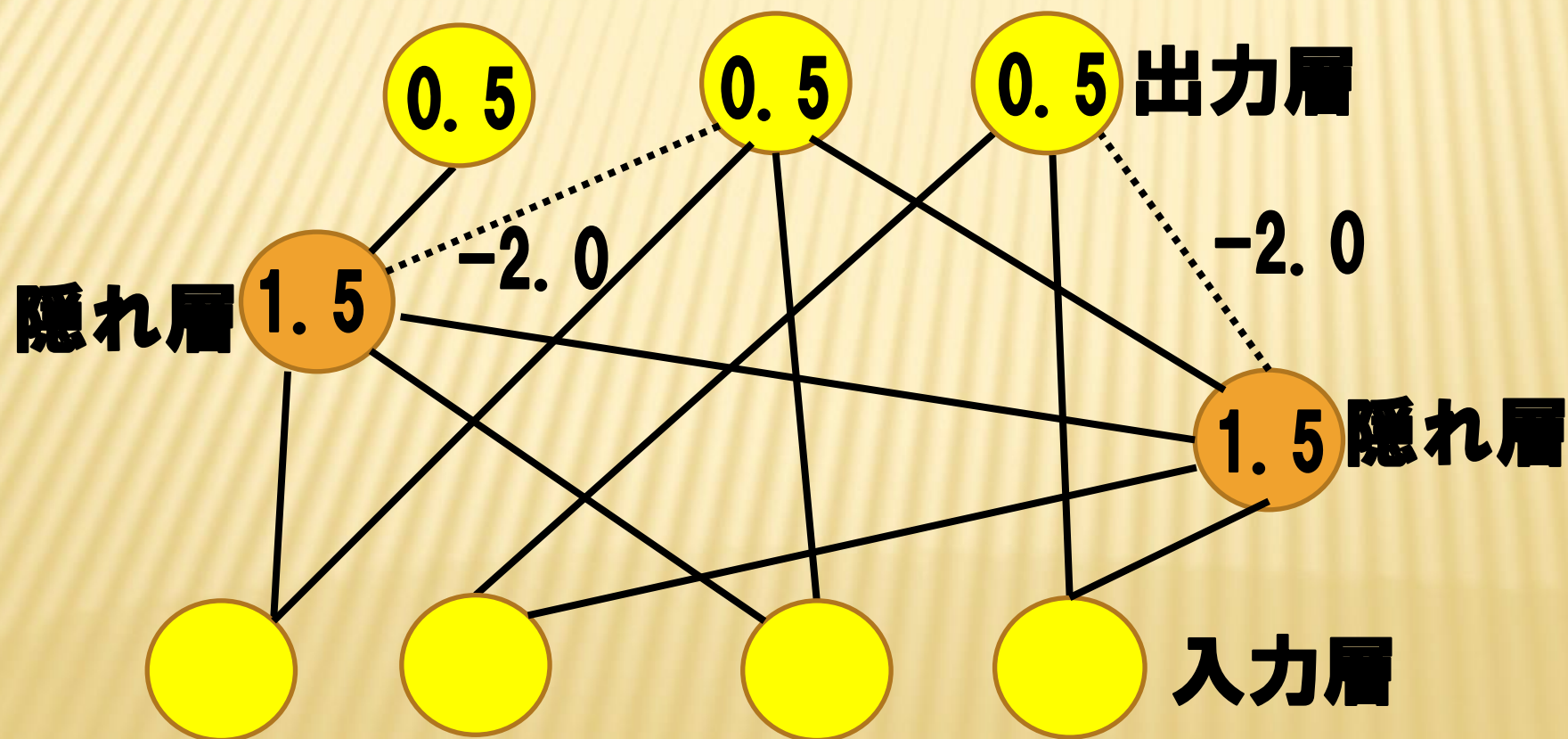


脳内のニューロン (A) と単純パーセプトロン (B)

4.1 第3次AIブーム

「機械学習、深層学習の時代」(2)

2桁の2進数の和を計算するネットワーク



○は細胞体、線は樹状突起です。○の中の数字は閾値で「**重み付けられた入力値の合計**」がこの閾値以上になると**1（発火）**を出力する

4.1 第3次AIブーム

「機械学習、深層学習の時代」(3)

2006年にニューラルネットワークの代表的な研究者であるジェフリー・ヒントンの研究チームが、制限ボルツマンマシンによるオートエンコーダの深層化(ディープニューラルネットワーク(DNN))に成功し、AIが再び注目を集めるようになった。

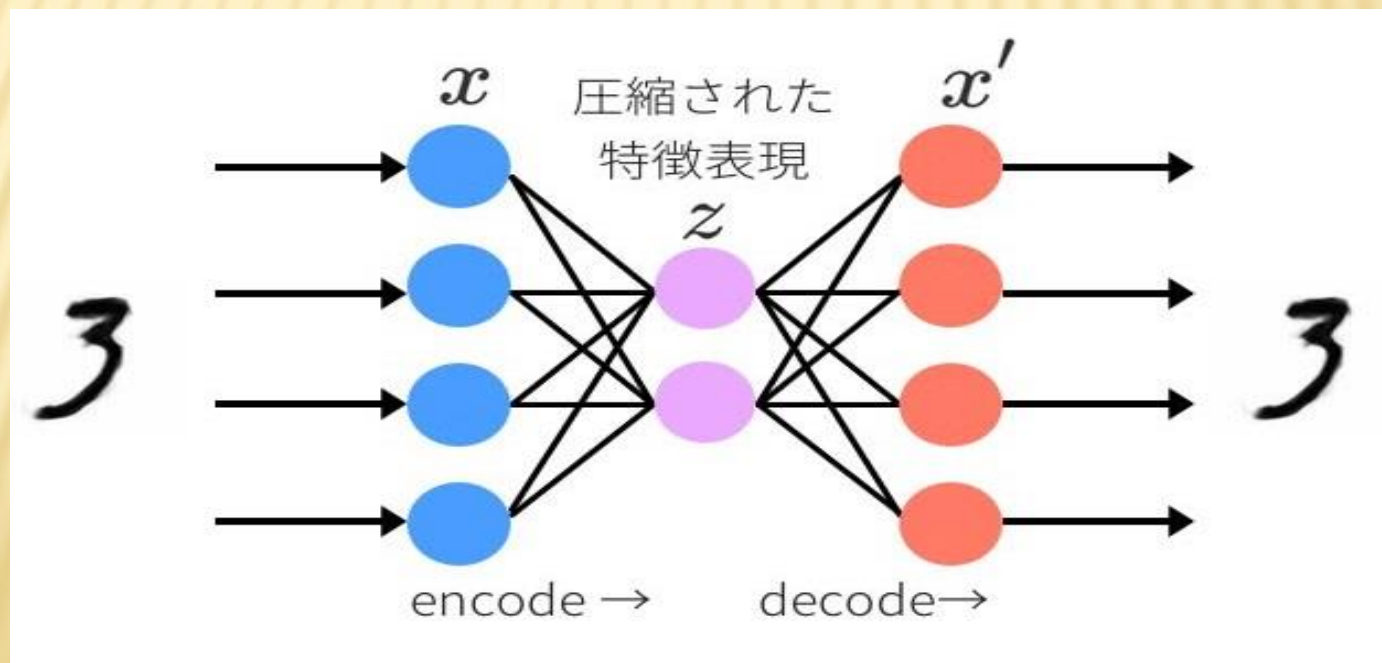
同時に、コンピュータのハード性能の急激な進歩や、CPUよりも単純な演算の並列処理に優れたGPUの低価格化により、2012年頃からは急速に研究が活発となり、**第三次人工知能ブームが到来**した。⇒スーパーコンピュータ

4.1 第3次AIブーム

「機械学習、深層学習の時代」(4)

● 「オートエンコーダ」とは

オートエンコーダは3層ニューラルネットにおいて、入力層と出力層に同じデータを用いて教師データありで学習させたものである。



4.1 第3次AIブーム

「機械学習、深層学習の時代」(5)

● 「オートエンコーダ」の数学

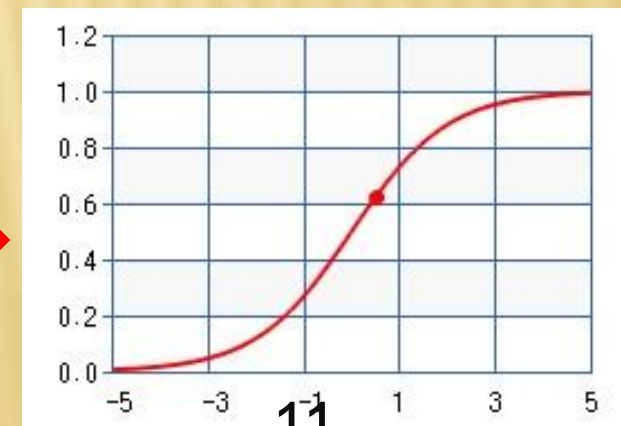
オートエンコーダを使った学習は出力が入力を再現するような学習である（「教師あり学習」）。

オートエンコーダのDecode式とencode式は入力を x とすると以下のようになる。

$$\text{encode: } z = s(Wx + b) \quad \Rightarrow \quad x' \doteq x \text{ (入力)}$$

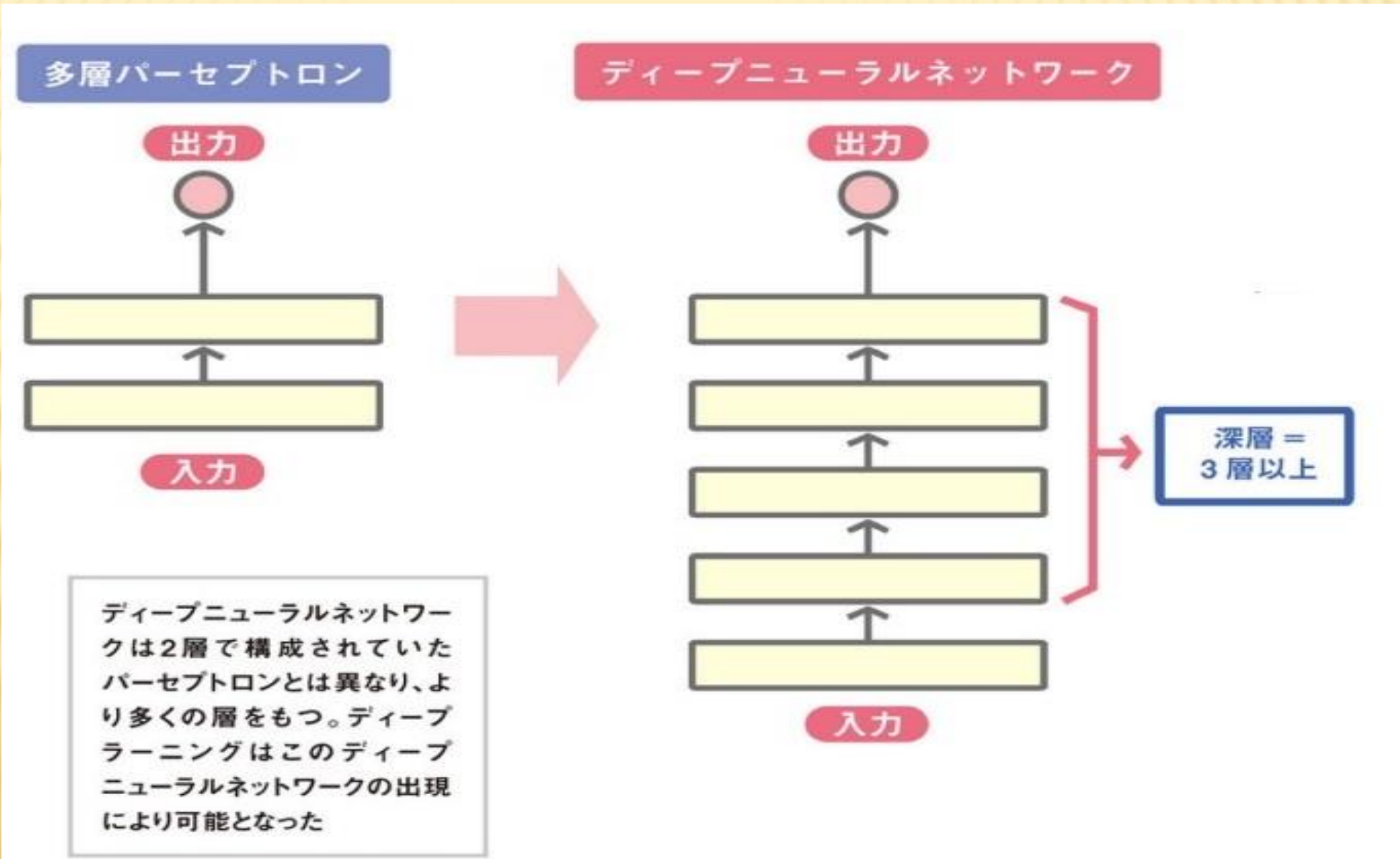
$$\text{decode: } x' = s(W'z + b')$$

上式の $s()$ は、閾値に使うシグモイド関数 \Rightarrow で、パラメータは W と b になる。



ディープニューラルネットワーク(DNN)の登場

- DNNの一種の「オートエンコーダ」深層化に成功 -



演算回数ではスパコン「京」を超えるディープラーニング専用機



Googleのディープラーニング専用チップ「Cloud TPU」64個を相互接続した「TPUポッド」。演算能力は11.5PFLOPSで、精度は異なるものの、単純に演算回数の比較では、スーパーコンピュータ「京」を上回る。(写真提供：Google)

NEXT'19のGOOGLEスーパーコンピュータ

AI For Data Scientists

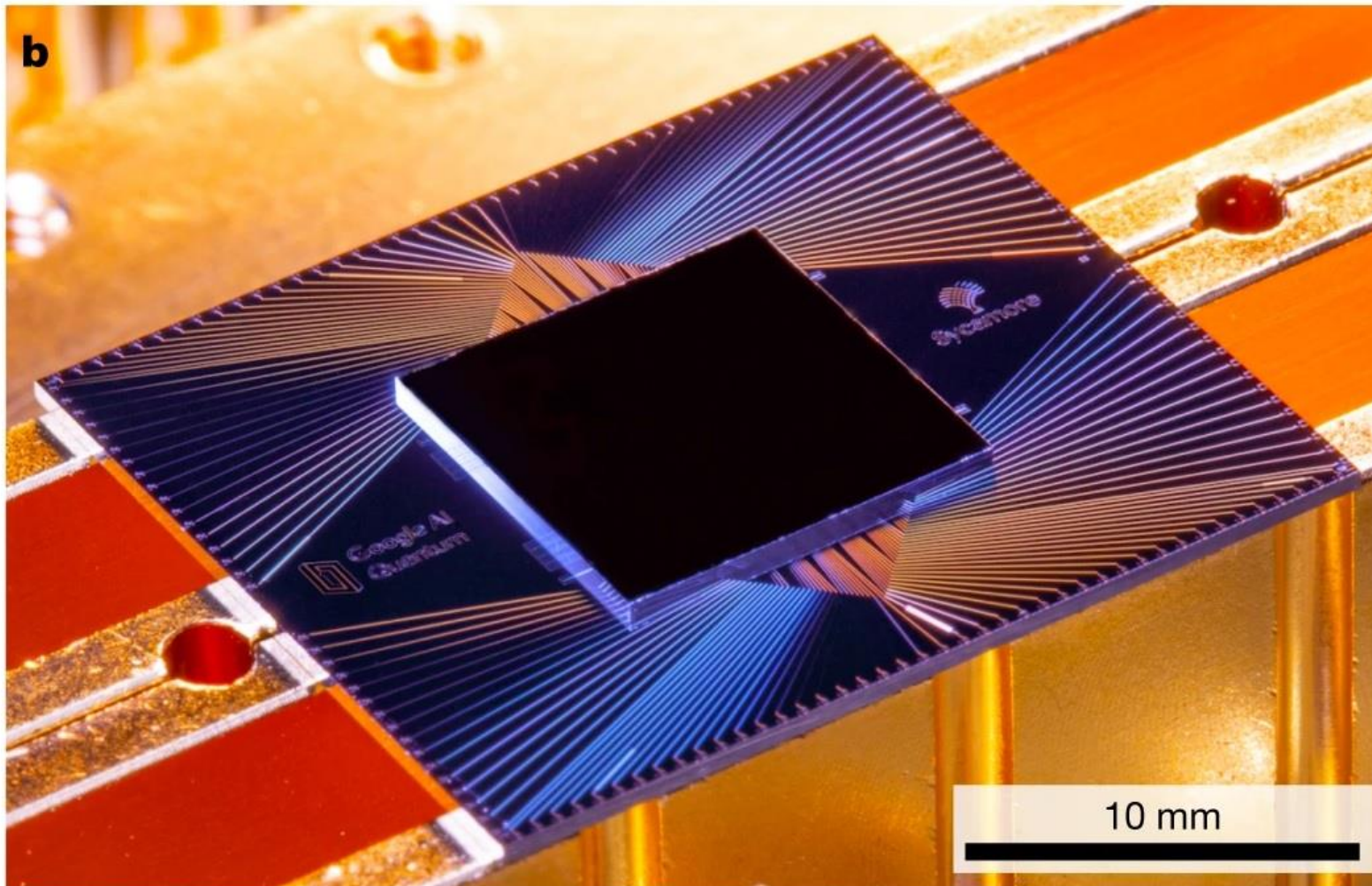
Cloud TPUv3

NEXT'19のGOOGLEスーパーコンピュータ



GOOGLEの量子コンピュータ

- 2019年10月23日「量子プロセッサ「Sycamore」を用いて、世界最速のスーパーコンピュータでも**1万年**かかるとされる処理を、Googleの測定結果によると**200秒**で実行した」と発表。



(出典：Nature)

4.2 人間の脳の研究と機械学習(1)

● 「人間の脳のモデル」

人間の脳は100億個以上のニューロン（神経細胞）から構成されている。

● 脳のモデルを作る研究の2つの目的：

1) 脳のしくみを解明すること

2) すぐれた情報処理装置を作ること

⇒ 脳の中にできる外界の写し「世界像」を自己形成し、思考・行動するような機械を創る

4.2 人間の脳の研究と機械学習(1)

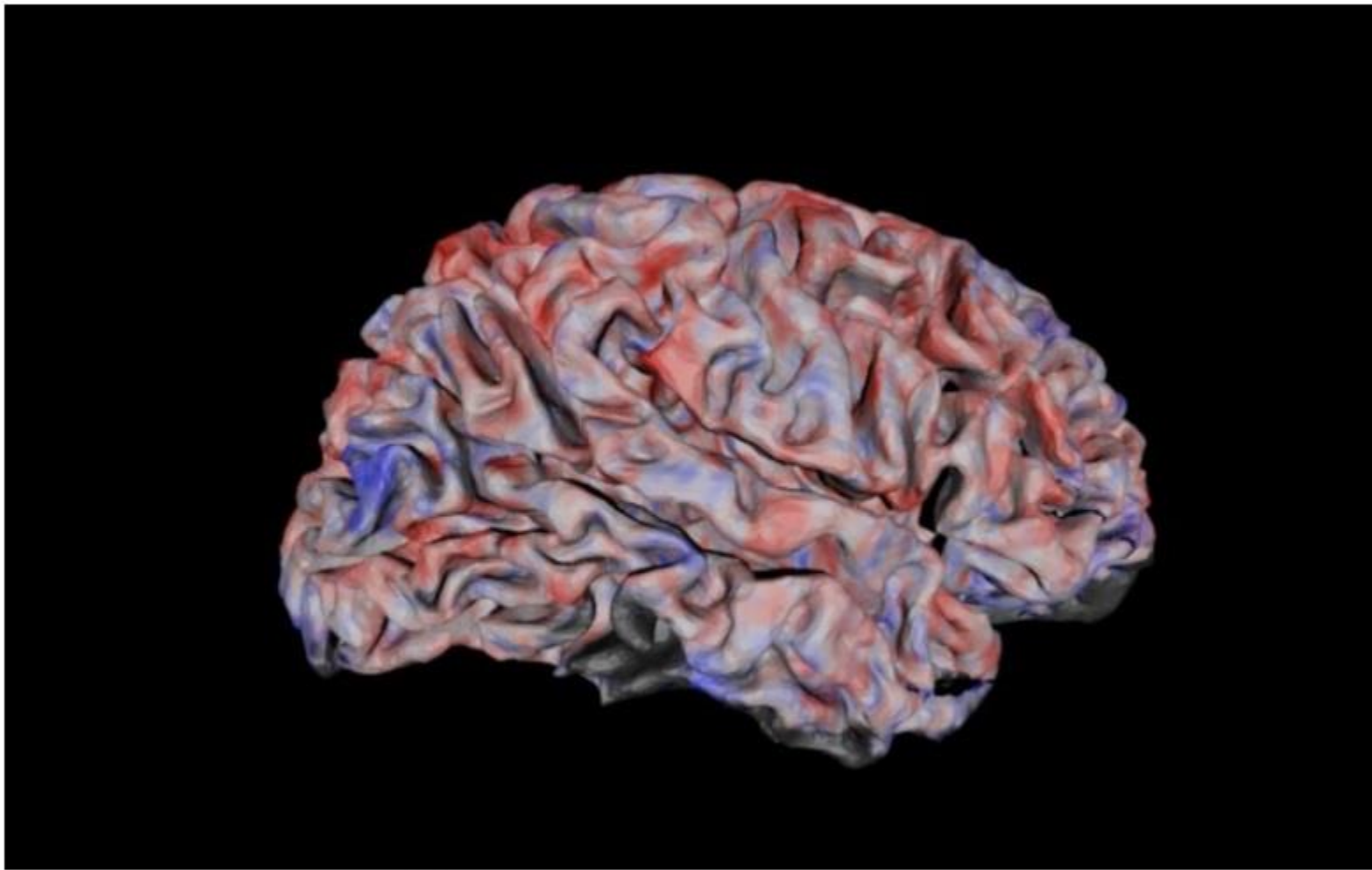
●脳のモデルを作る研究

1) 脳のしくみを解明すること



4.2 人間の脳の研究と機械学習(1)

● 脳のモデルを作る研究



4.2 人間の脳の研究と機械学習(1)

●脳のモデルを作る研究

1) 脳のしくみを解明すること

知覚・認知体験
(例：動画視聴)



符号化 (エンコード) モデル



脳活動
(例：fMRI計測)



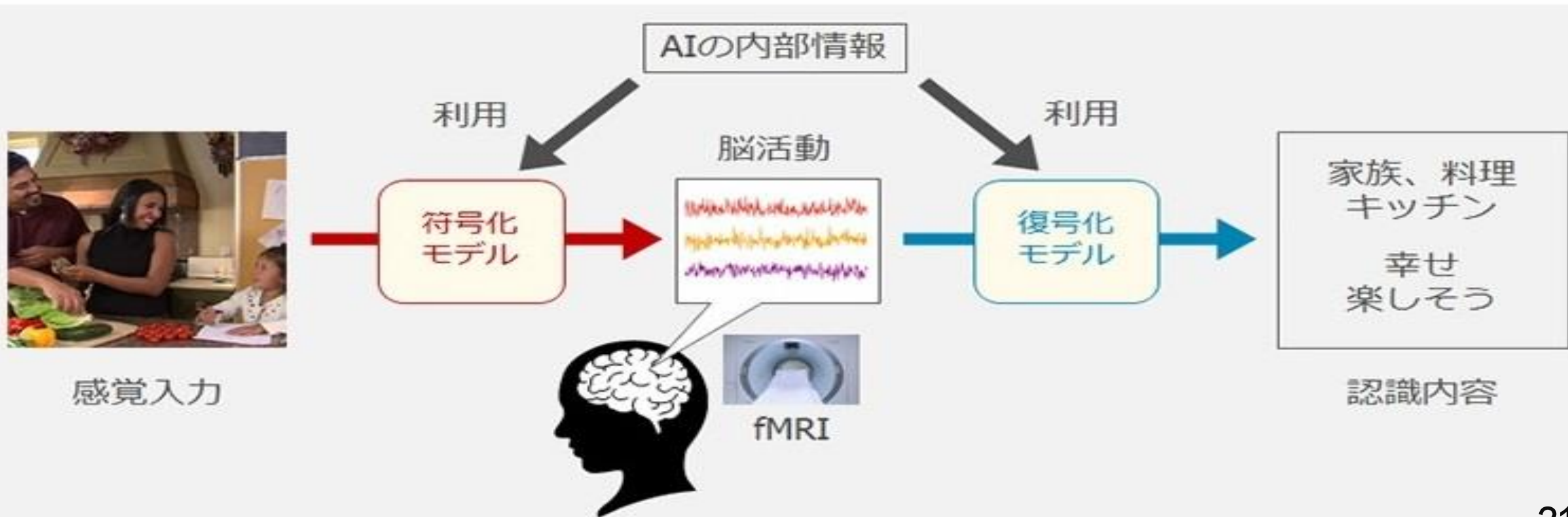
逆符号化 (デコード) モデル

低  高
20

4.2 人間の脳の研究と機械学習(1)

●脳のモデルを作る研究

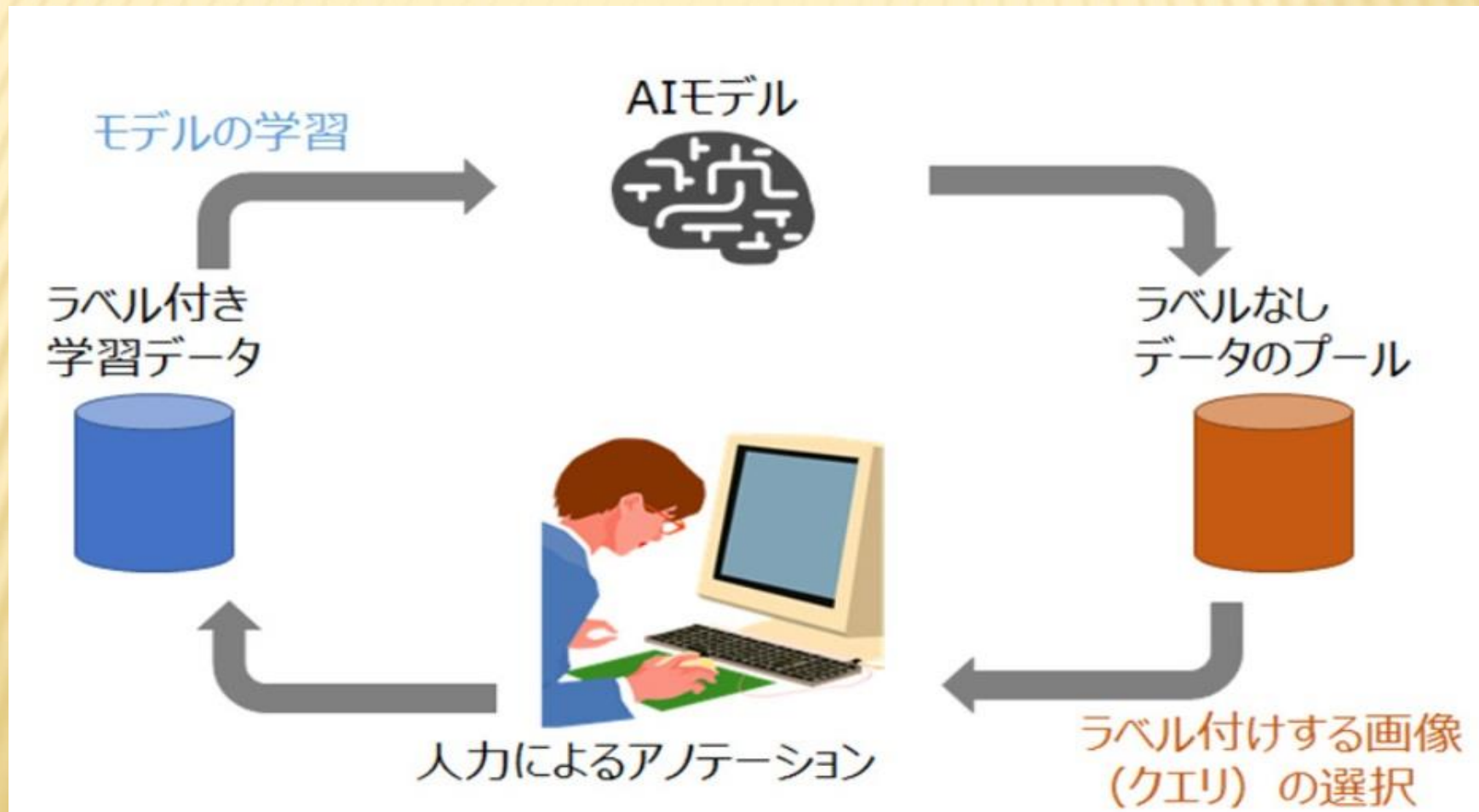
2) すぐれた情報処理装置を作ること



4.2 人間の脳の研究と機械学習(1)

● 脳のモデルを作る研究

2) すぐれた情報処理装置を作ること



4.2 人間の脳の研究と機械学習 (2)

● 「脳の視覚野」

脳は、生理学的な特性によって"**領野**"(りょうや)と呼ばれる機能単位を有しており、視覚に関連する領野も様々な部位に分類できる。

次図は、**脳の視覚野間の分類と関連性を模式的に表した図**である。視野中の形状に関する視覚信号は、**腹側経路(V1野→V2野→V4野→IT(TE、TEO)野)**と呼ばれる経路で処理される。

4.2 人間の脳の研究と機械学習 (3)

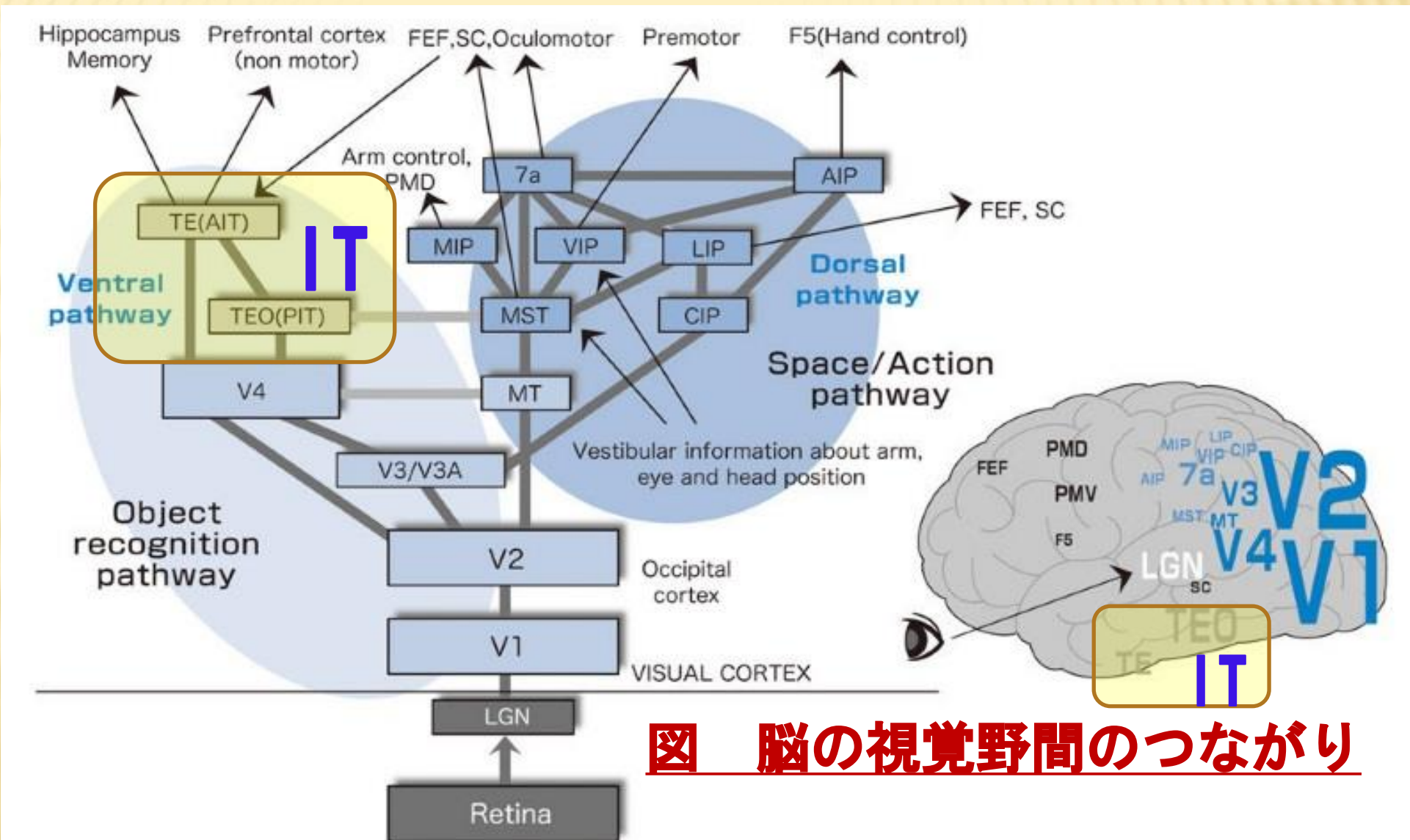


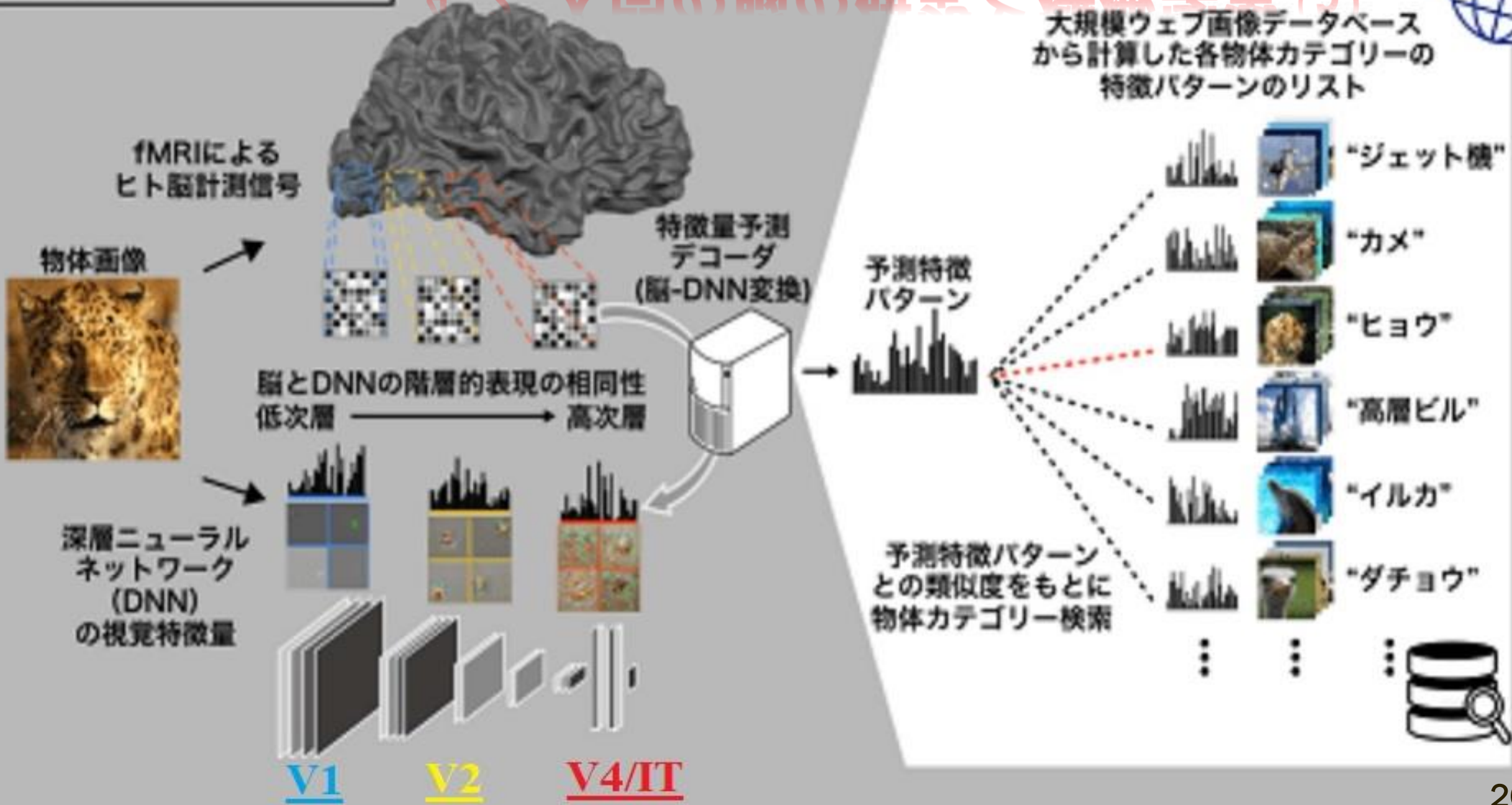
図 脳の視覚野間のつながり

4.2 人間の脳の研究と機械学習(4)

● 「初期視覚野：V1野」 (その1)

初期視覚野と呼ばれるV1野の細胞は、視野中に**受容野**と呼ばれる情報処理範囲を持ち、一つの細胞の反応する範囲は非常に局所的である。これらの細胞は、受容野に示された入力刺激が、自身の”**好み**“に**適合しないと発火状態**にならない。この好みのことを**選択特徴**などと呼ぶ。

V1野細胞の受容野の大きさはおよそ1度程度の大きさであり、抽出特徴としては、それほど複雑ではなく、**画像中の「線分」や「エッジ」といった成分を抽出**していると考えられている。



4.2 人間の脳の研究と機械学習(6)

● 「高次視覚野：V4野からIT野」 (その1)

高次視覚野であるV4野からIT野にかけては、耳や鼻といった顔の特徴的な部位に強く反応する細胞や、顔そのものに反応する細胞と抽象的な概念を符号化している細胞が観測される。

特定の人画像のみならず個人名が書かれたテキスト画像にも同様に反応する、ある種の概念が符号化された細胞の存在も知られている。

4.2 人間の脳の研究と機械学習(6)



fMRIと機械学習を用いて脳活動から心の中でイメージした内容の画像化

4.2 人間の脳の研究と機械学習(7)

● 「高次視覚野：V4野からIT野」 (その2)

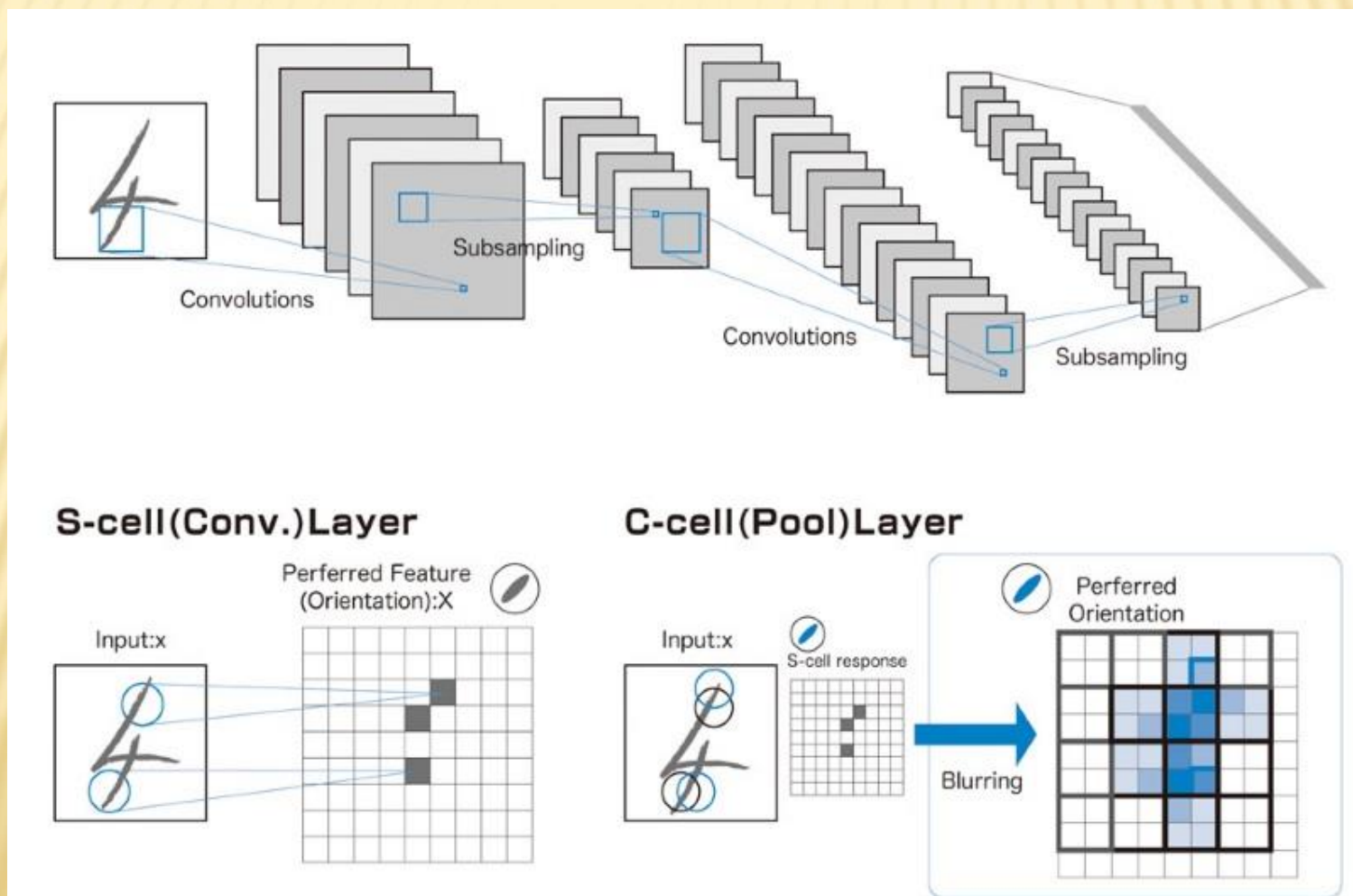
高次視覚野であるV4野からIT野の細胞では、受容野の大きさはかなり拡大され、IT野に至っては視野中のほぼ全域が受容野となる。

視野中のどこに物体や写真等を提示しても反応する細胞となっている。

次の図は、福島邦彦氏が提案した「ネオコグニトロンアーキテクチャ」で、機能的な事実と生理学的な事実を基にした初期視覚野の細胞群のモデルである。

4.2 人間の脳の研究と機械学習(8)

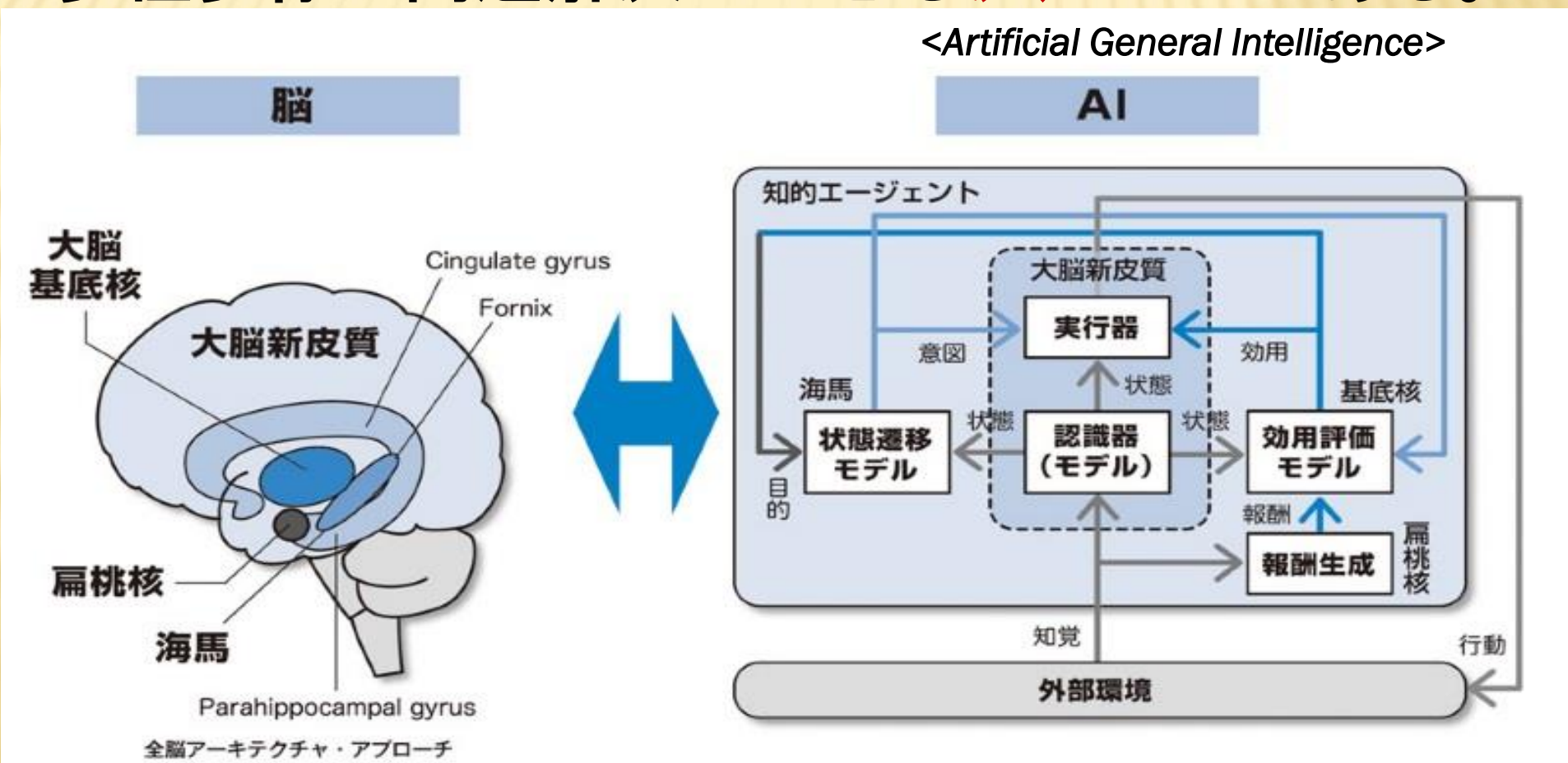
● 「ネオコグニトロンアーキテクチャ」



4.2 人間の脳の研究と機械学習 (8)

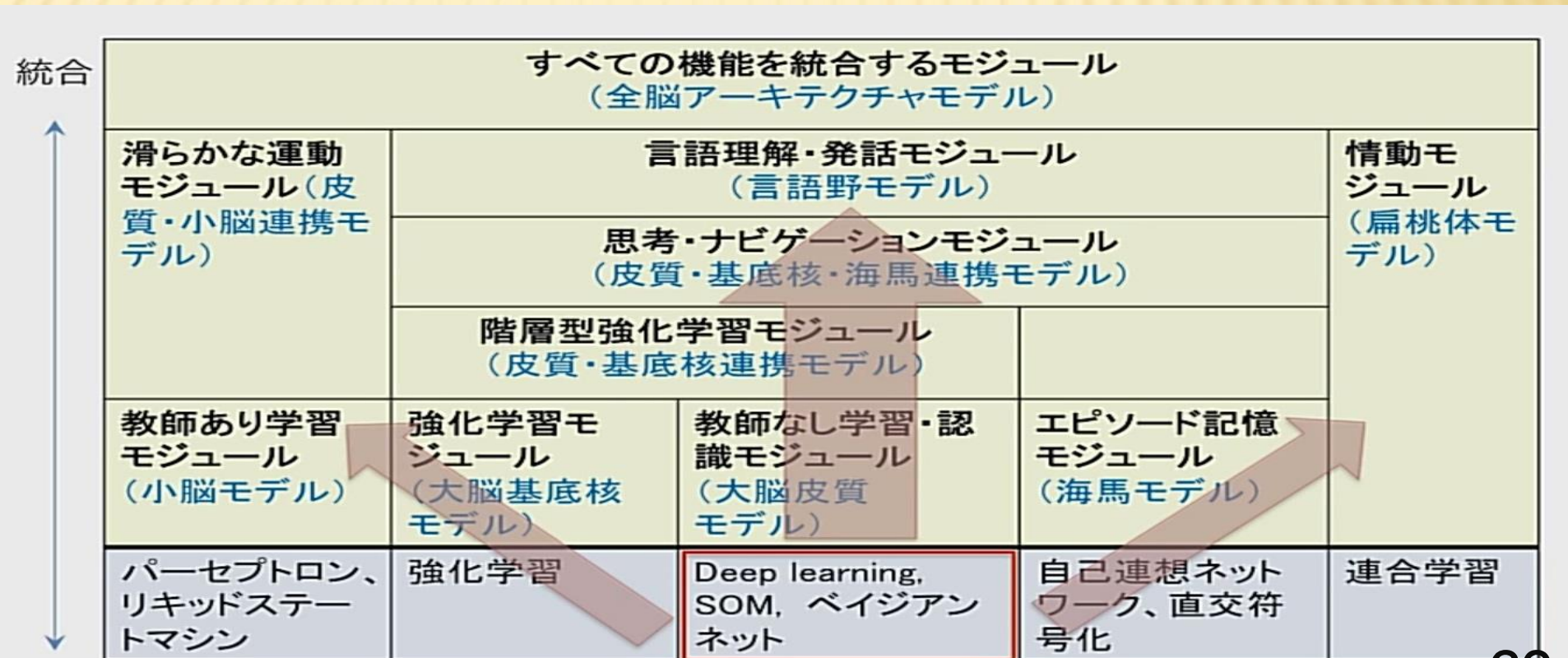
「脳とAI」の比較

AGIとは、人間のような汎用能力を持ち
多種多様の問題解決ができる汎用AIである。



4.2 人間の脳の研究と機械学習 (9)

汎用人工知能 (AGI) は「強いAI」とも言われ、「弱いAI」と言われる特化型AI (ANI) に対して「人間と同様の感性や思考回路をもつ」人工知能のことを指す。



図「AGI：全脳アーキテクチャ」

4.2 人間の脳の研究と機械学習(10)

●強いAIと弱いAI

- ・ 強いAI : 汎用AI ⇒ 例) 鉄腕アトム
- ・ 弱いAI : 道具としてのAI ⇒ 金融ロボット

4.2 人間の脳の研究と機械学習(11)

● 「機械学習」

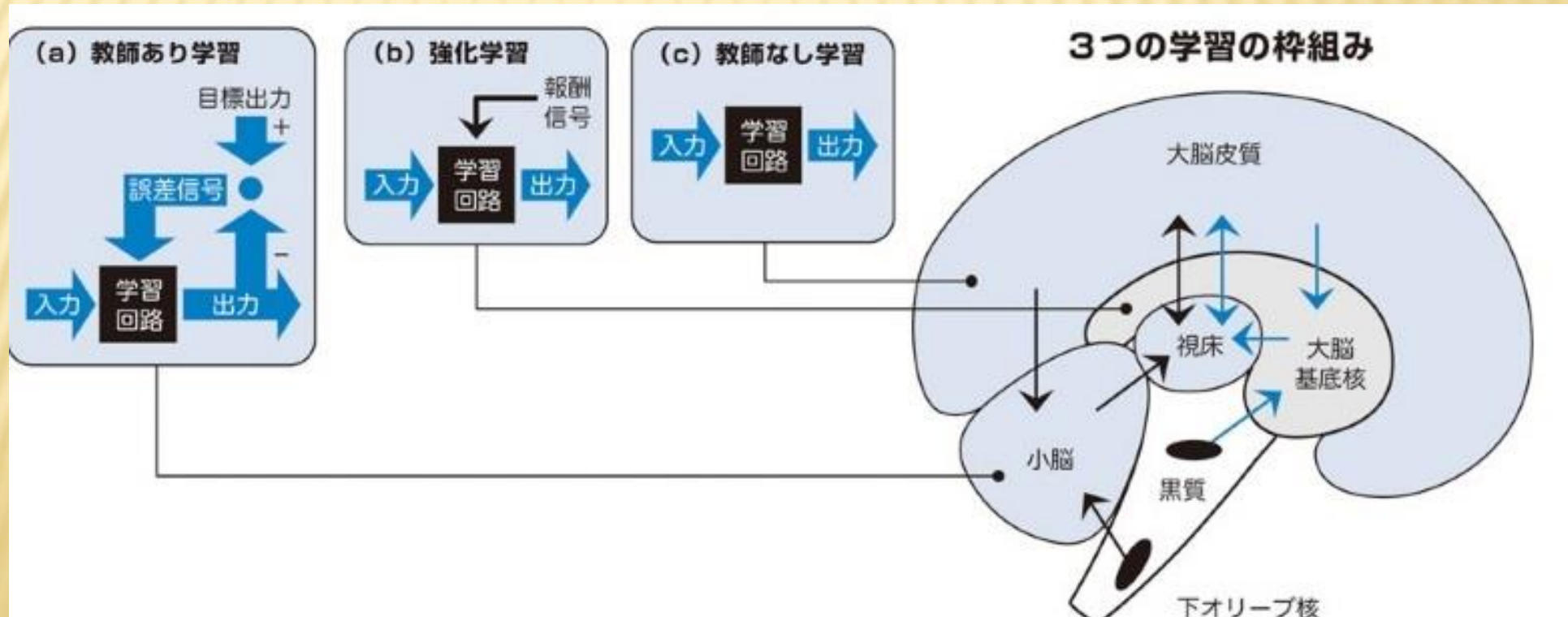
脳における学習の枠組みに基づく、機械に学習させる「機械学習」には「教師あり学習」、「教師なし学習」、と「強化学習」の三つの学習の枠組みがある。

脳の部位として、それぞれ小脳、大脳皮質、と大脳基底核と深く関連がある。

4.2 人間の脳の研究と機械学習(12)

● 「機械学習の三つの枠組み」

下図は人間の脳のニューロンが層状に接続した構造を模擬した機械学習の三つの枠組みである。



4.2 人間の脳の研究と機械学習(13)

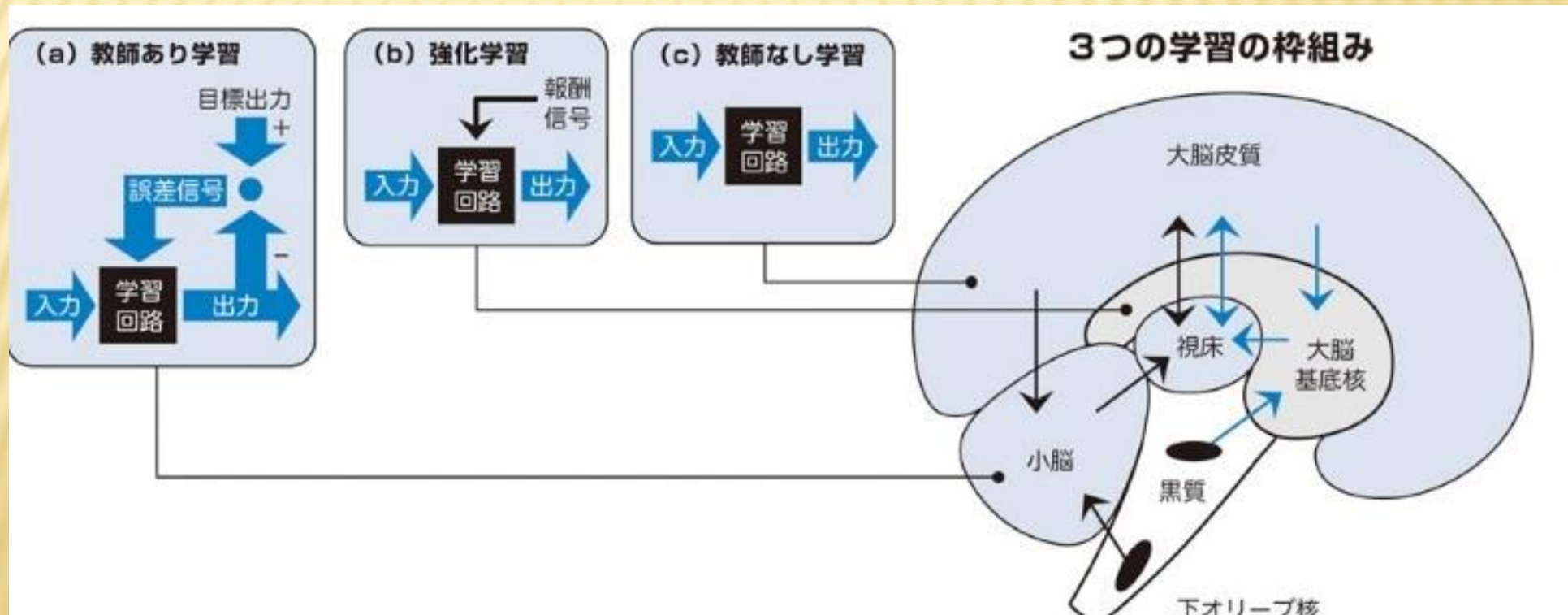
● 「機械学習」の枠組み

脳における学習の枠組みに基づく、機械に学習させる「機械学習」には「教師あり学習」、「教師なし学習」、と「強化学習」の三つの学習の枠組みがある。

4.2 人間の脳の研究と機械学習(14)

● 「機械学習の三つの枠組み」

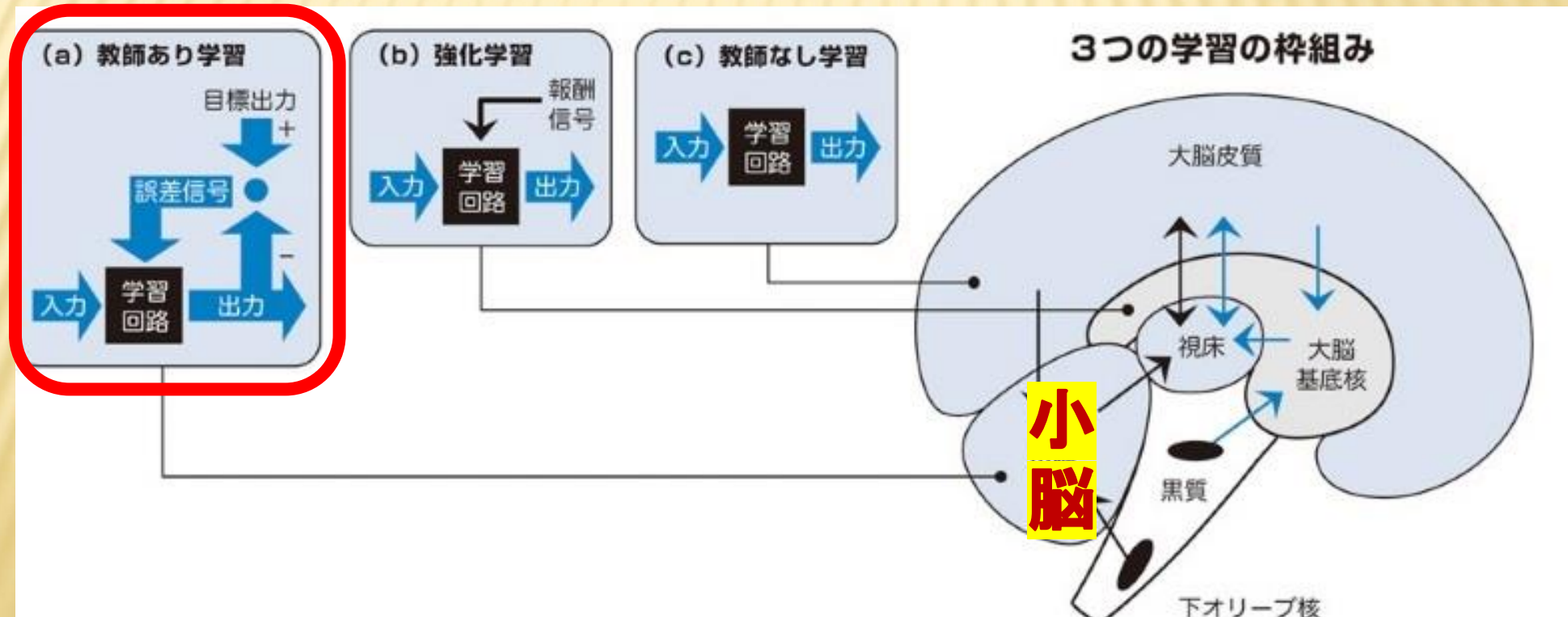
下図は人間の脳のニューロンが層状に接続した構造を模擬した機械学習の三つの枠組みである。



4.2 人間の脳の研究と機械学習(15)

● 「教師あり学習」

まず「教師あり学習」を見てみよう。
主に、**小脳**が担う学習機能である。



4.2 人間の脳の研究と機械学習(16)

● 「教師あり学習」

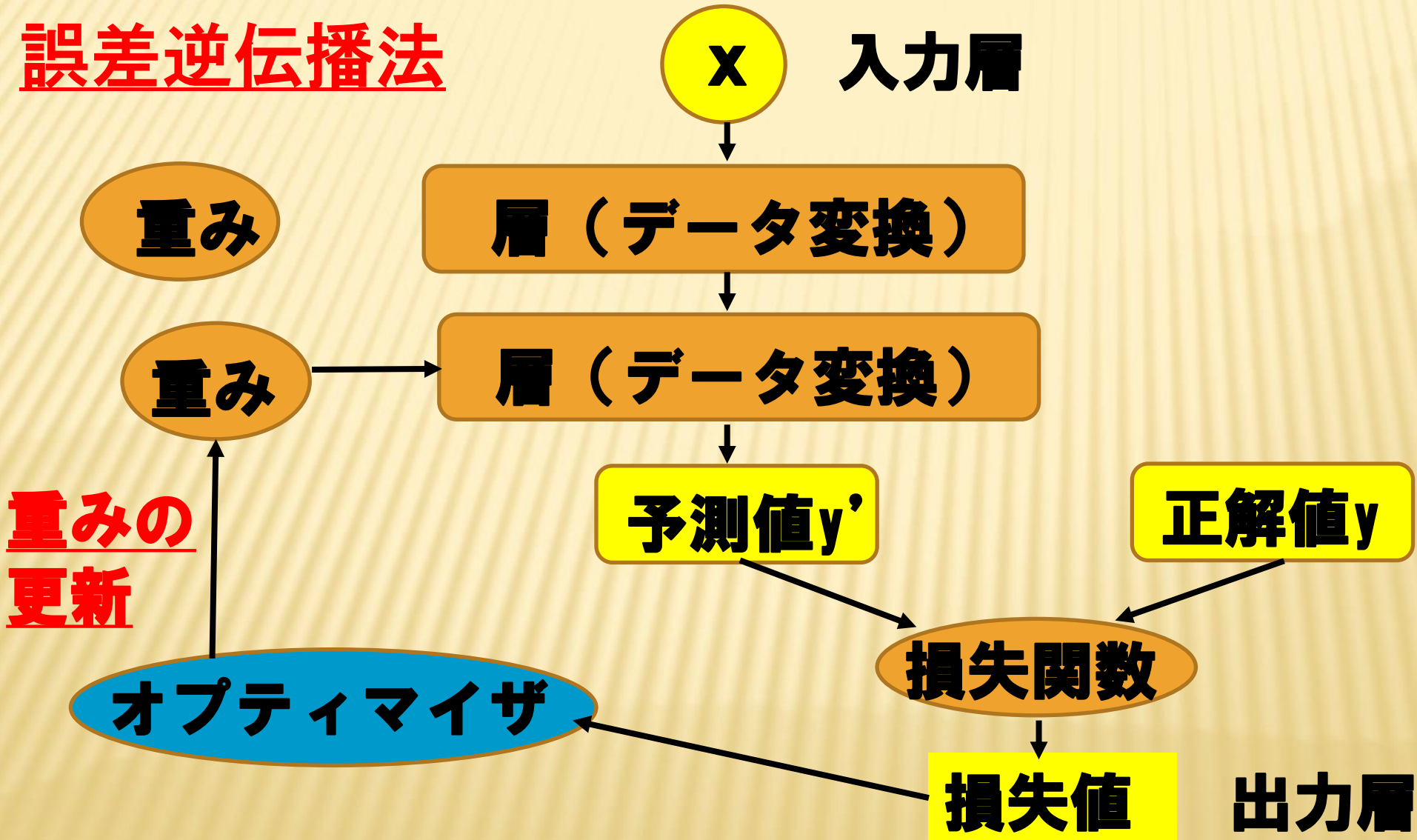
学習者に対し、教示者が明示的に正解を教えたり、学習者の誤りを指摘したりすることで、学習者が正しい解を得ることを助ける。

すなわち、正しい入出力の組合せを与えて学習することで、新規の入力に対し、適切な出力が出せる。

誤差逆伝播法(Back Propagation)が、その代表的手法である。また、正解、若しくは誤りを入力として、未経験入力に対する意志を決定する決定木(Decision Tree)の作成などもある。

4.2 人間の脳の研究と機械学習(16)

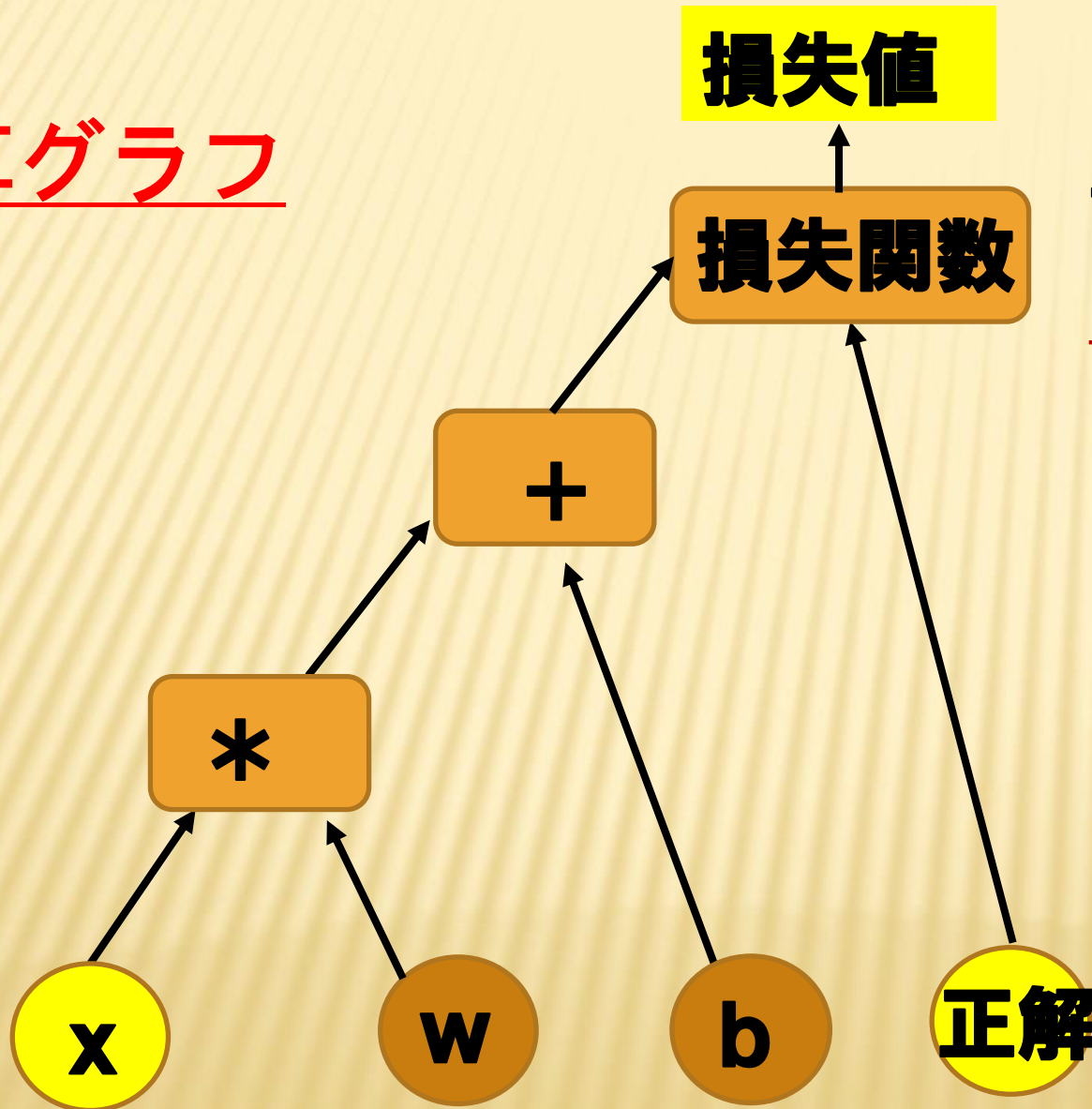
誤差逆伝播法



ディープラーニングの基本原理は、損失値をフィードバックし、重みをほんの少し調整する。誤差逆伝播法は「オプティマイザー」で重みを調整する仕組み。

4.2 人間の脳の研究と機械学習(16)

計算グラフ

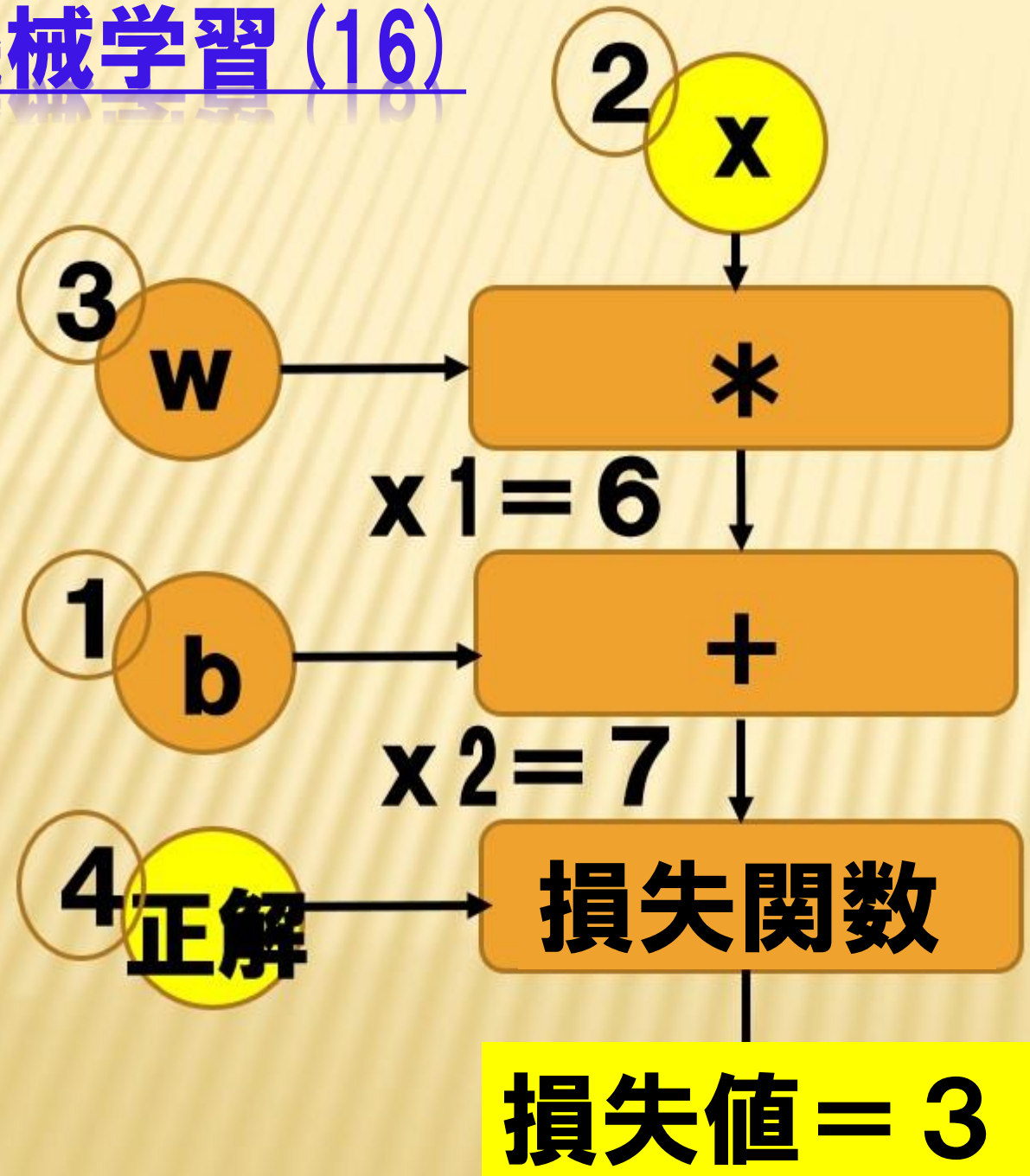
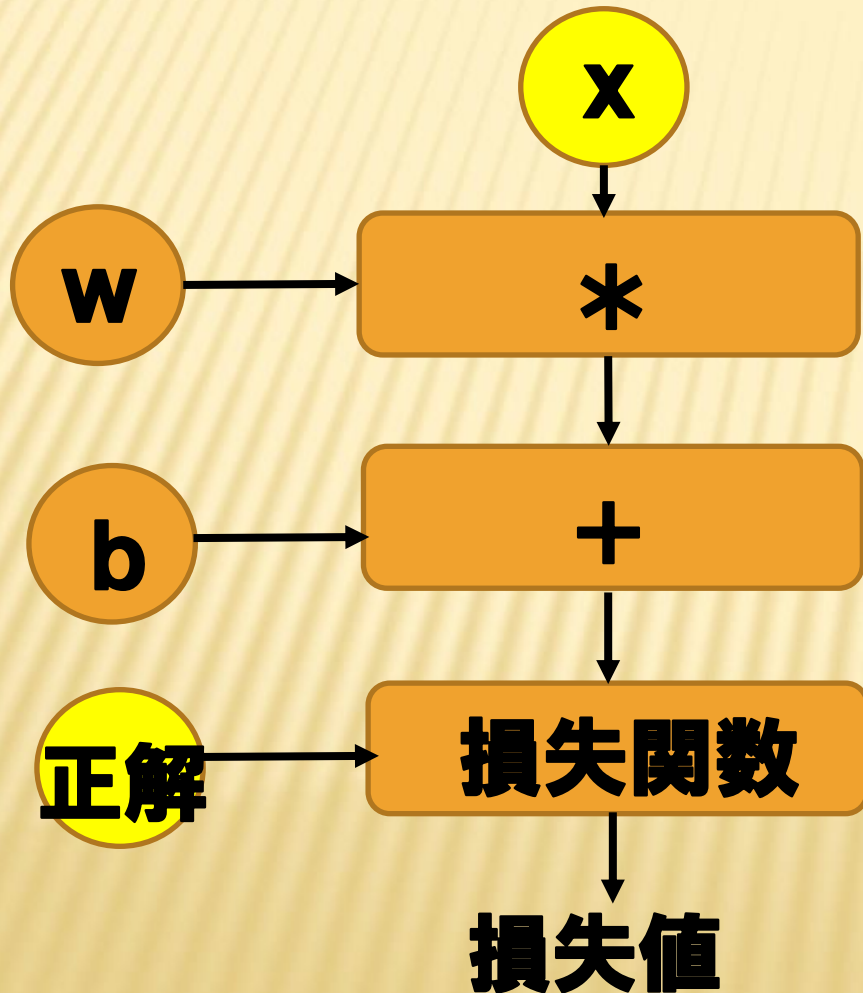


計算式：損失値 = 損失 (正解、 $x * w + b$)

誤差逆伝播法のアルゴリズムを理解する為に、**計算**をデータとして扱える「**計算グラフ**」を使う。

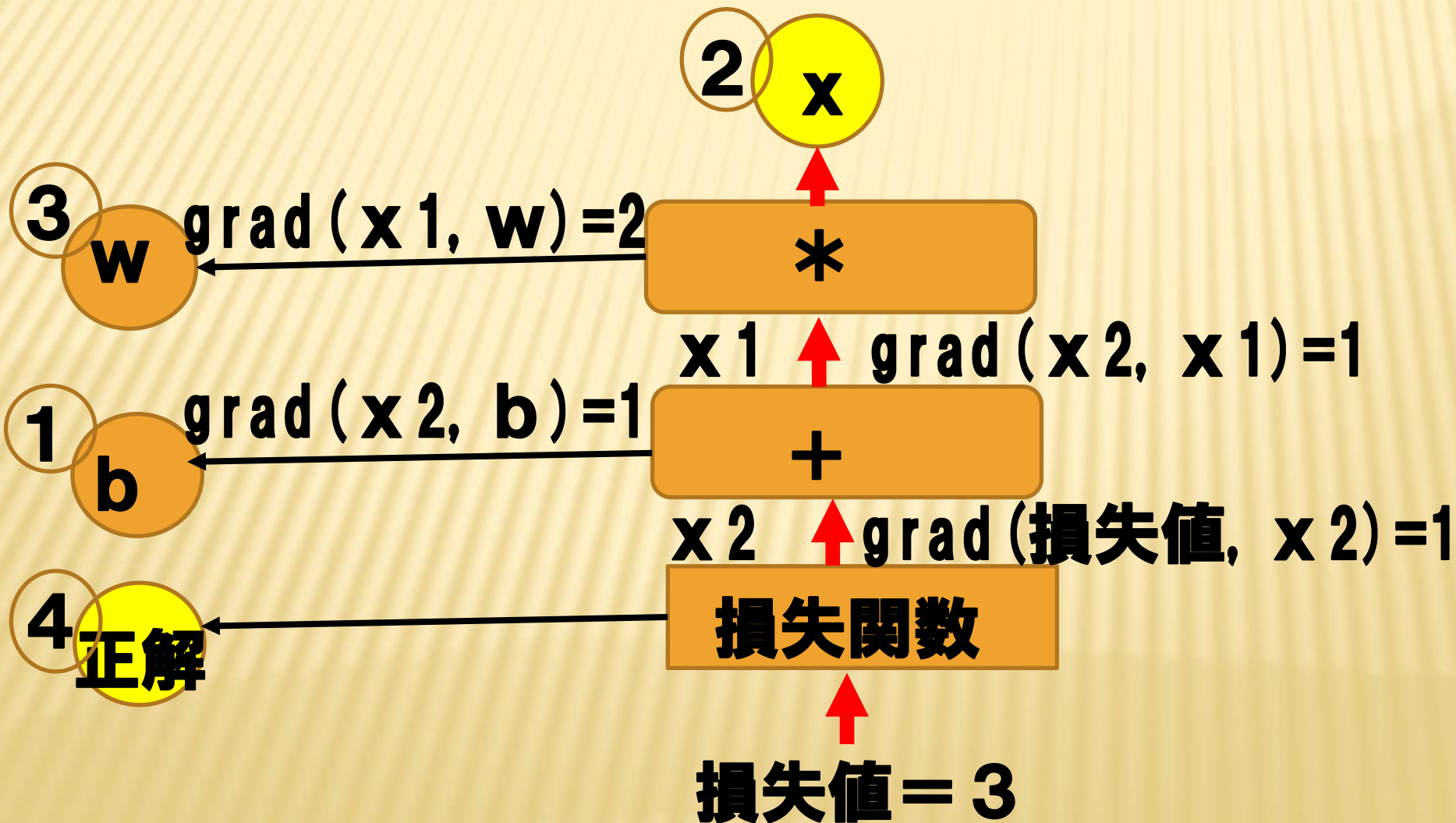
4.2 人間の脳の研究と機械学習 (16)

● フォワードパスを実行



4.2 人間の脳の研究と機械学習 (16)

● バックワード (後向) パスを実行



バックワード (後向) パスでは、**重み (パラメータ)** を調整するために**損失関数の勾配 (gradient)** を計算する

4.2 人間の脳の研究と機械学習(16)

<目的>バックワード（後向）パスでは、 w と b の値で損失値が最小になるようなアルゴリズムで**重みを更新したいので、 $\text{grad}(\text{損失値}, b)$ と $\text{grad}(\text{損失値}, w)$ を計算する。**

<計算方法>

① $\text{grad}(\text{損失値}, x_2) = 1 : x_2 = 1$

② $\text{grad}(x_2, x_1) = 1 : x_2 = 2, x_2 = x_1 + 1$

③ $\text{grad}(x_2, b) = 1 : x_2 = x_1 + b$ なので

④ $\text{grad}(x_1, w) = 2 : x_1 = 1$ なので

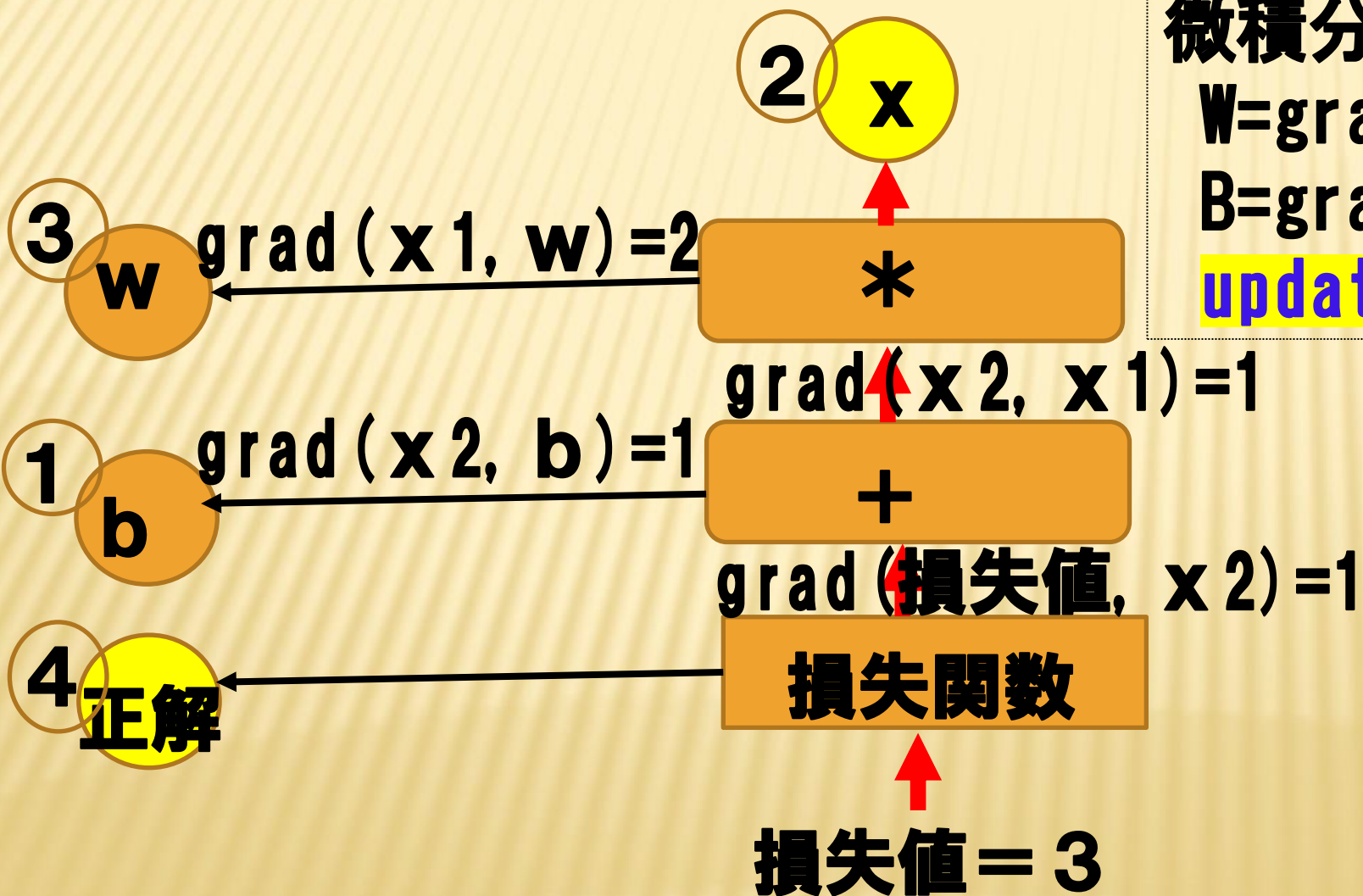
微積分の連鎖律を使い、誤差逆伝搬法（オプティマイザー）

$$\text{grad}(\text{損失値} - w) = \text{grad}(\text{損失値}, x_2) * \text{grad}(x_2, x_1) * \text{grad}(x_1, w)$$

$$\text{grad}(\text{損失値} - b) = \text{grad}(\text{損失値}, x_2) * \text{grad}(x_2, b)$$

4.2 人間の脳の研究と機械学習 (16)

●バックワード（後向）パスを実行



微積分の連鎖律を使い、
 $W = \text{grad}(\text{損失値} - w) = 1 * 1 * 2$
 $B = \text{grad}(\text{損失値} - b) = 1 * 1$
update_parameter(W, B)

**最適化
で重みを変える。**

<誤差逆伝搬法>

4.2 人間の脳の研究と機械学習(16)

誤差逆伝搬法 (バックプロパゲーション) の数式

<目的>

出典: wikipedia

ネットワーク $\hat{y} = f(x; w)$ に対する誤差関数 $E(\hat{y}, y)$ を定義したとき、現在の重み $w_k = a_{now}$ における

E の傾きすなわち偏微分値 $\left. \frac{\partial E}{\partial w_k} \right|_{w_k = a_{now}}$ がわかれば、最適化手法である勾配法を用いて誤差 E が小さく

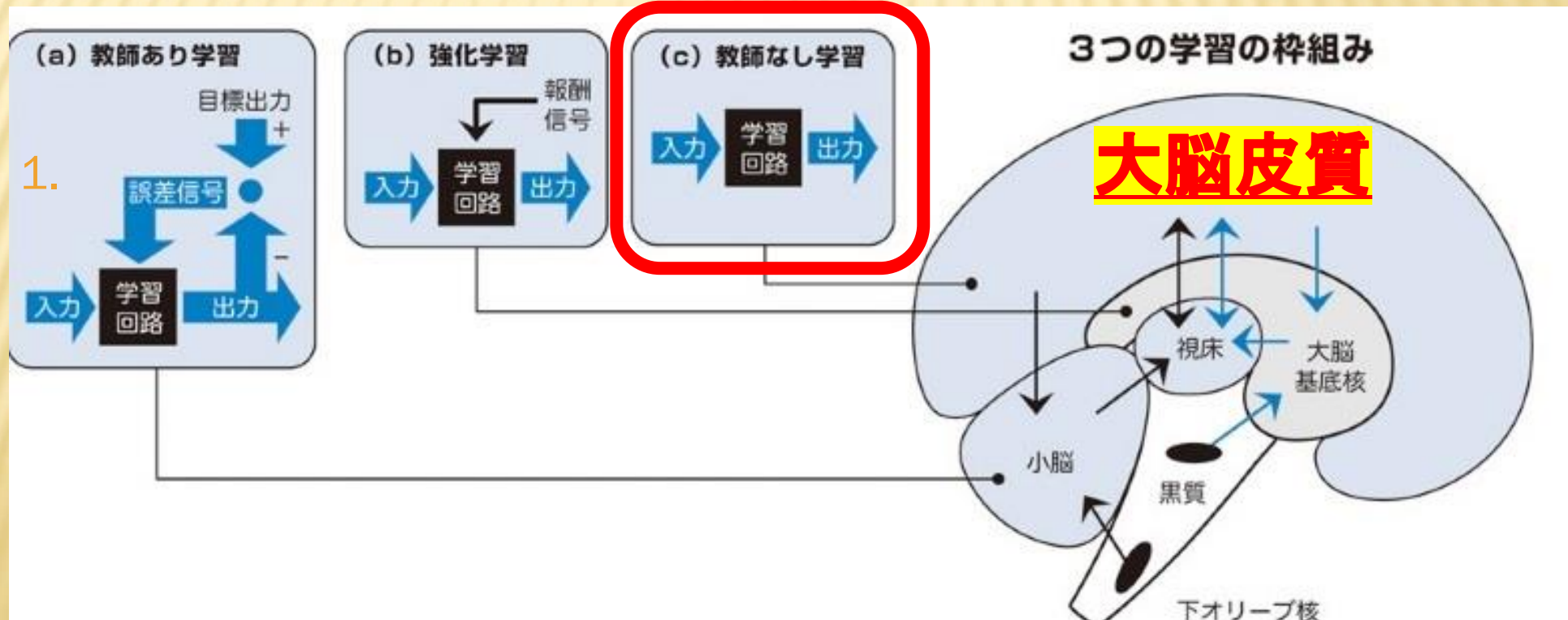
なるように w_k を更新 (=学習) できる。学習アルゴリズムであるバックプロパゲーションの目的はこの勾配値を得て重みを学習することである。

4.2 人間の脳の研究と機械学習(17)

● 「教師なし学習」

まず「教師なし学習」を見てみよう。

主に、**大脳皮質**が担う学習機能である。



4.2 人間の脳の研究と機械学習(18)

● 「教師なし学習」

統計的性質や、ある種の拘束条件により入力パターンを分類したり、抽象化したりする学習。

主成分分析、自己組織化マップなどの次元圧縮(Dimensionality Compression)手法が代表例である。

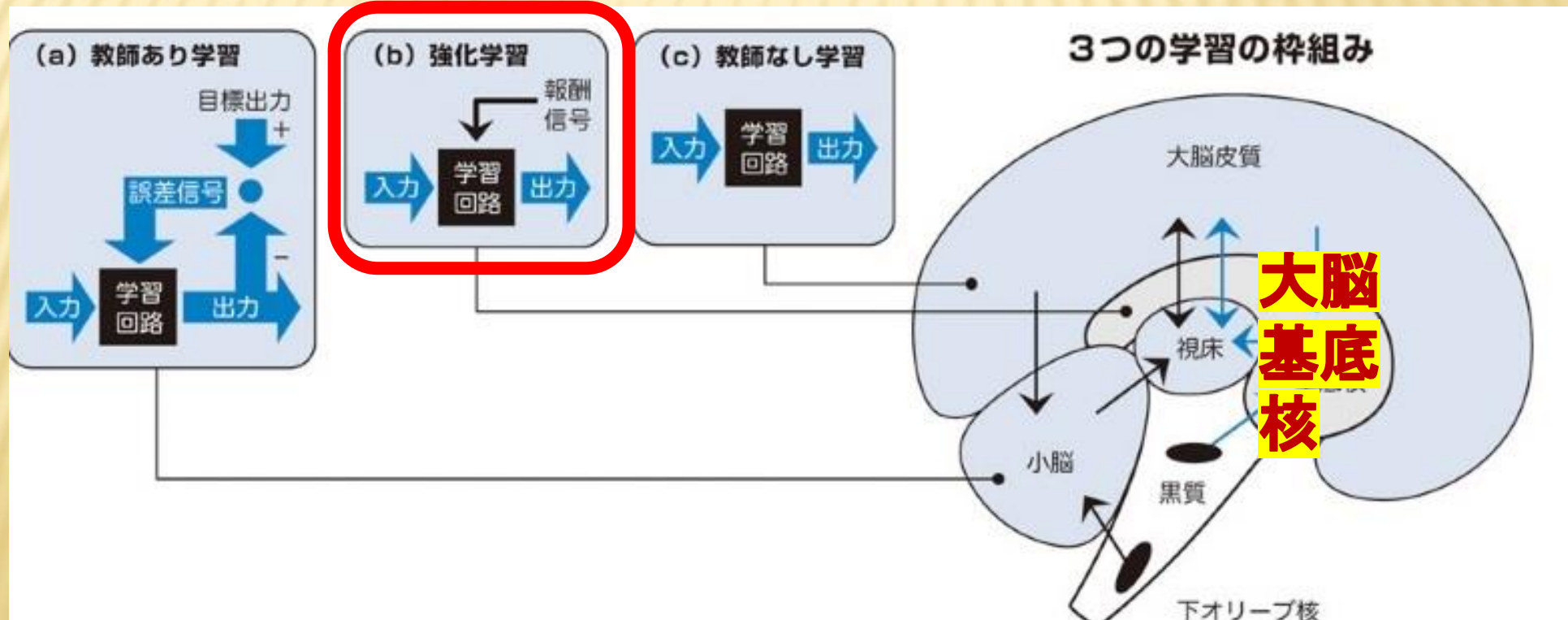
感覚情報などの入力パターンの分類、同様に出力運動パターンに対して統計的性質を用いて要素行動に分類する学習法などがある。

4.2 人間の脳の研究と機械学習(19)

● 「強化学習」

次に「強化学習」を見てみよう。

主に、**大脳基底核**が担う学習機能である。



4.2 人間の脳の研究と機械学習 (20)

● 【強化学習】

最終結果若しくは、途中経過に対して、どの程度良かったかを示す「報酬信号」に基づき、これらの報酬をなるべく大きくするように探索する。

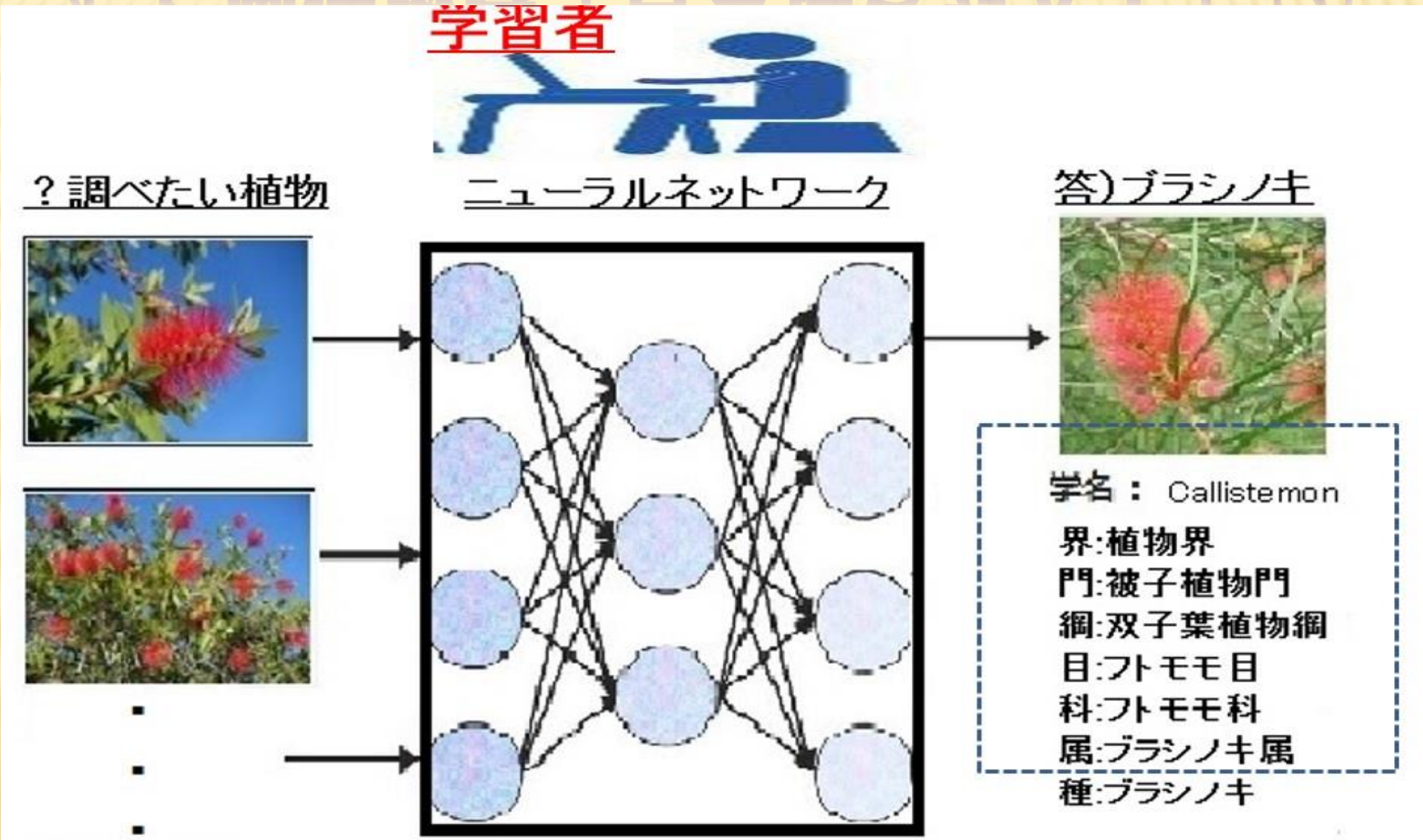
強化学習と教師あり学習の違いは、フィードバックがスカラー(成否)かベクトル(howの情報)かという説明もあるように、明示的な教師ではなく、環境などの非明示的な教師だという解釈もある。

4.3 画像認識「目を持ったAI」(1)

●次図は人工の脳神経回路（ニューラルネットワーク）で深層学習する「目を持ったAI」の例である。

●学習者が主体的に自然観察などで発見した調べたい植物の写真を撮り、撮った写真をすぐにAIに転送し、予め自然観察に関して深層学習したAIから「ブラシノキ」と教えてもらおうとともに、ブラシノキに関する「界～属、種」や同属種といった詳細情報を得て、学習者は深い理解に到達する

4.3 画像認識「目を持ったAI」(2)



4.3 画像認識「目を持ったA I」(3)

- 「目を持ったA I」は従来の事典に代わり、教科の主體的な深い学びの道具「生き字引」として活躍している。
- 最近、魚類図鑑や動物図鑑などのA I図鑑「リンネレンズ」がL i n n e (株)からW e bで公開されている。

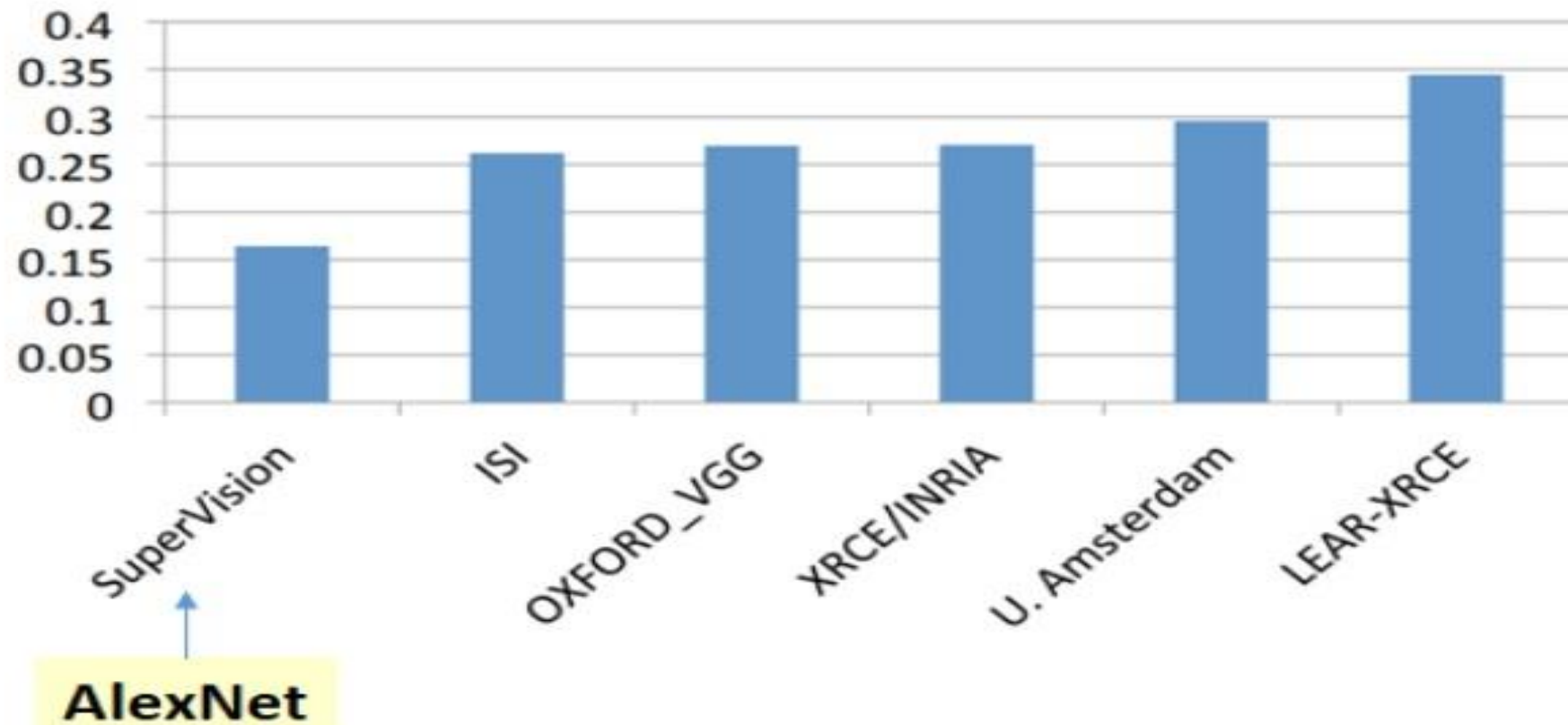
4.3 画像認識「目を持ったAI」(4)

●「ディープラーニングで一般物体認識のコンテスト「ILSVRC」で優勝」

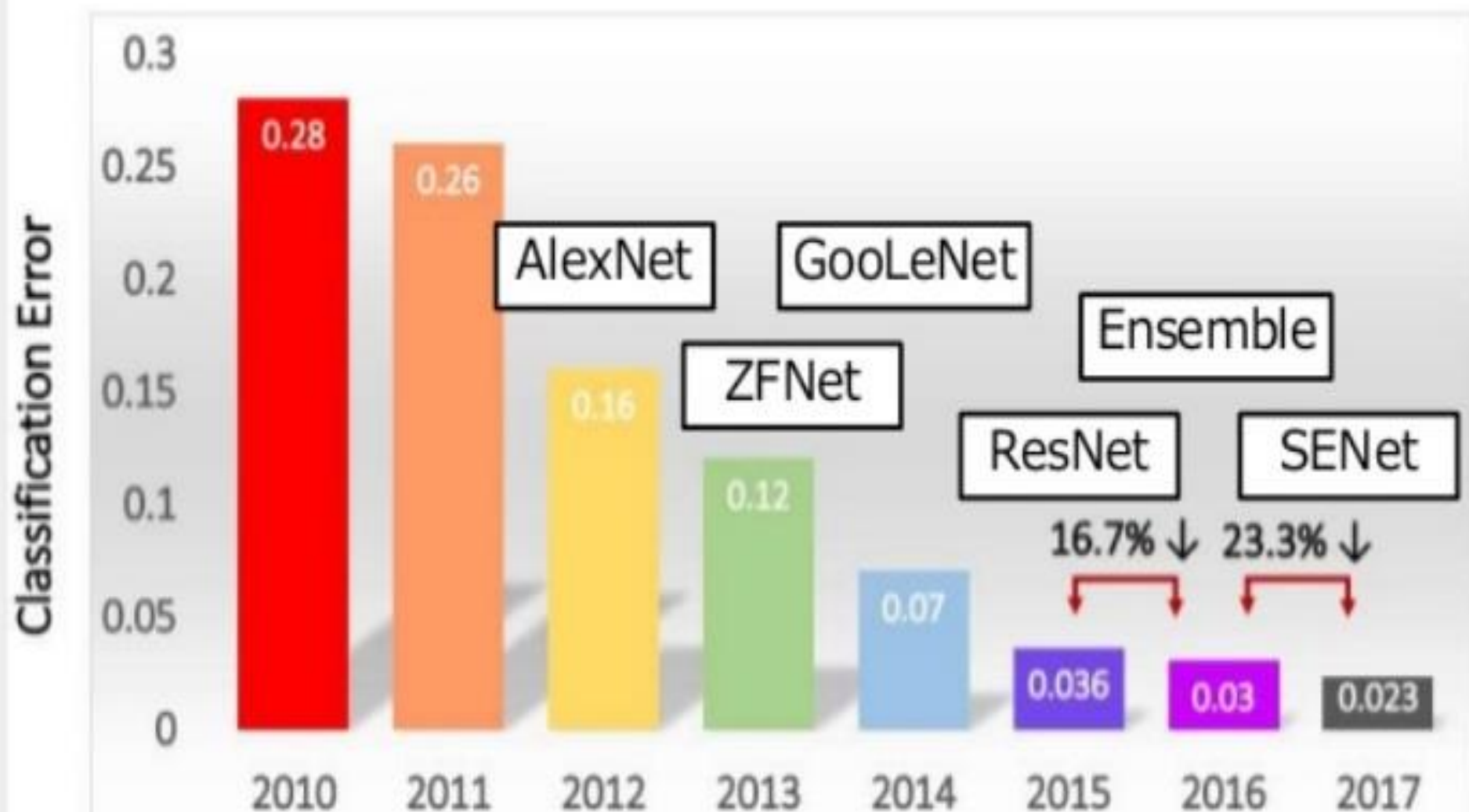
2012年に開催された一般物体認識のコンテスト「ILSVRC」(Image Net Large Scale Visual Recognition Challenge)において、**深い構造を持つCNNが、従来手法の分類性能を大幅に上回った。**以来、ディープラーニングが画像認識に盛んに利用されるようになった。

4.3 画像認識「目を持ったAI」(5)

● トロント大のジェフリー・ヒント教授が AlexNet (ディープラーニング) で一般物体認識のコンテスト「ILSVRC2012」で優勝



4.3 画像認識「目を持ったAI」(6)



4.3 画像認識「目を持ったAI」(7)

● 「クラス分類」(1)

CNNは、層を沢山重ねて深い構造にすることで、より高い精度で物体を分類できるようになるが、その反面、パラメータ数が膨大となり学習が困難になる。

そこで、深い層でも学習がうまくいく枠組みとして、「ResNet」が提案されている。

ResNetは、出力を入力と入力からの差分の和でモデル化したネットワークである。この構造によって、上層からの誤差が下層まで伝播するようになり、1000層といったかなり深い構造でも適切に学習が可能となった。

4.3 画像認識「目を持ったAI」(8)

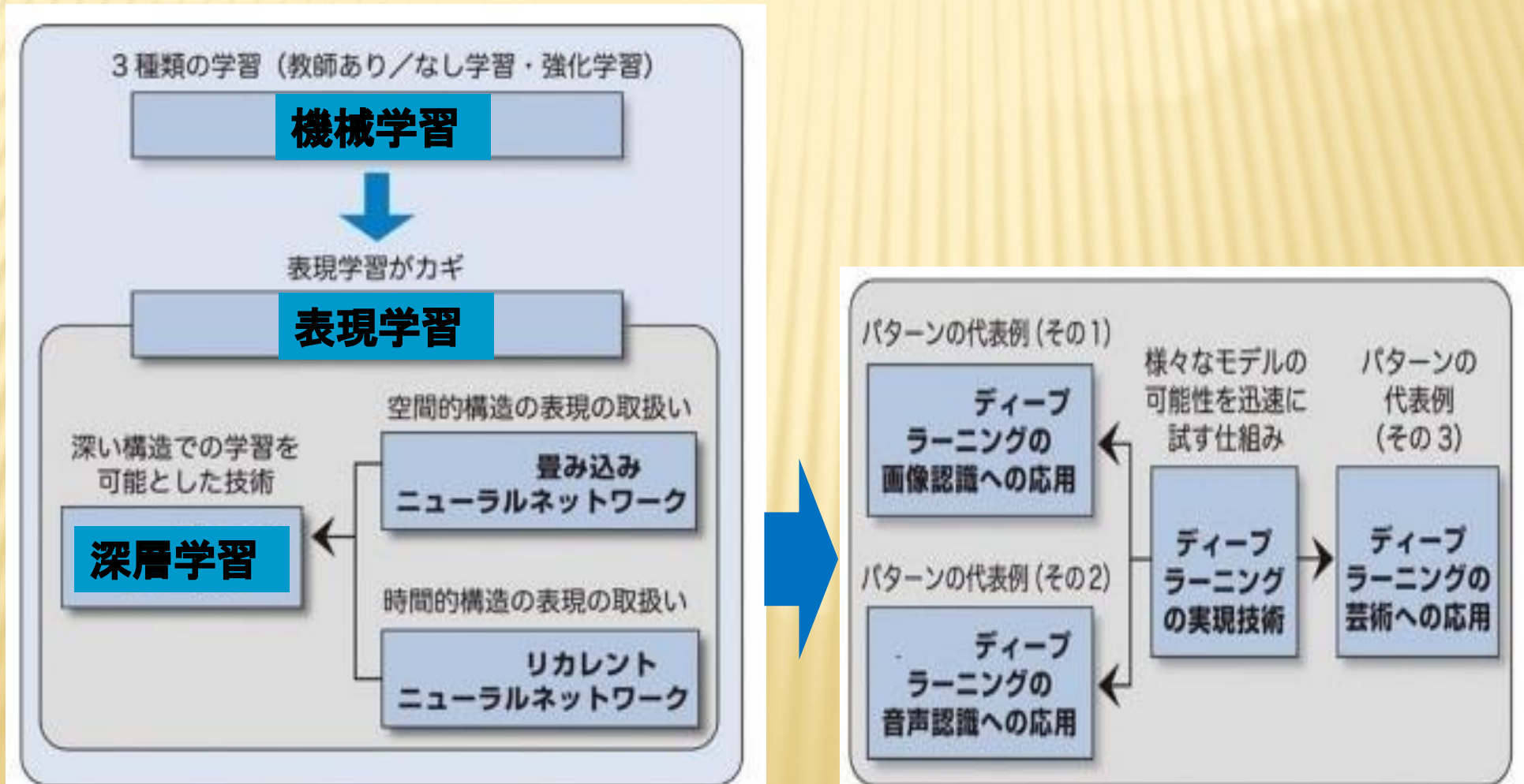
● 「クラス分類」(2)

ResNetは、ILSVRC2015の様々な部門において、**トップの成績を獲得した**。このときの物体クラス分類課題における、上位5位までに正解が含まれないエラー率は、3.57%であった。一方ILSVRC2016における同部門の1位のエラー率は2.99%、2位が3.03%、3位が3.04%であり、2015年と比較して分類性能はほとんど伸びていない。

ILSVRC2016のトップの手法は既存技術の組合せであり、物体クラス分類において、この1年間はインパクトのあるトピックが出ていない状況にある。

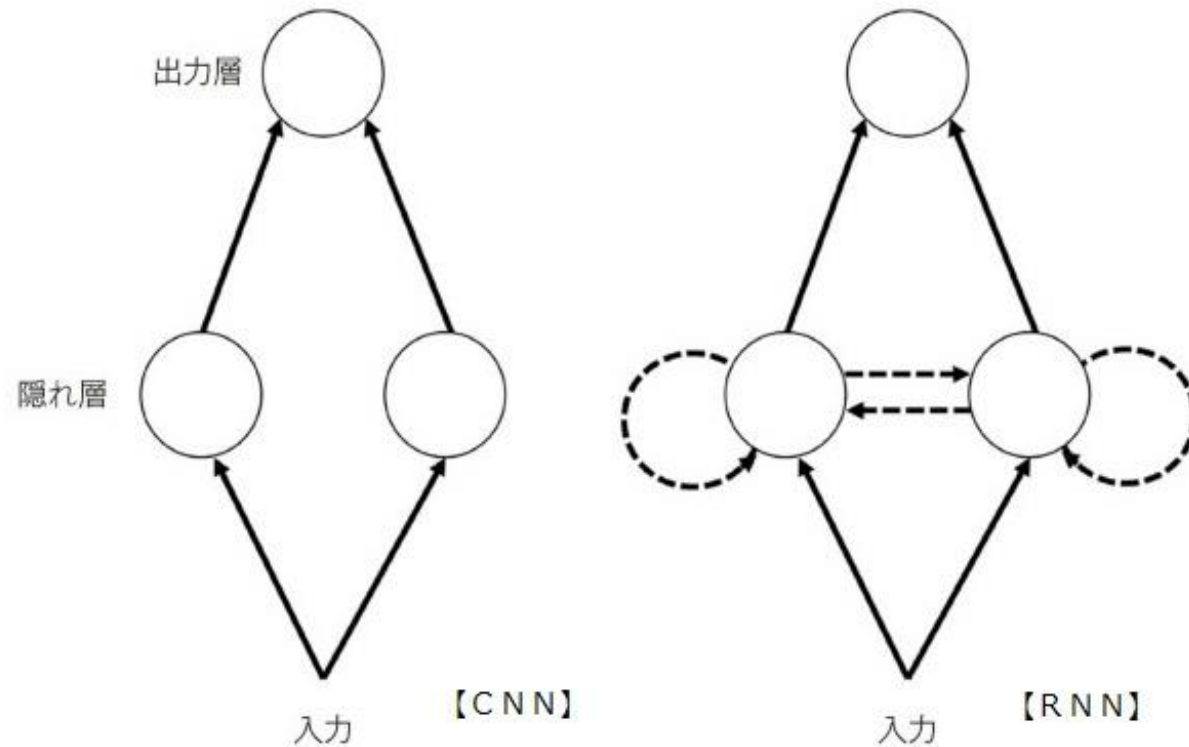
4.3 画像認識「目を持ったAI」(1)

図は機械学習、表現学習と深層学習の関連を示す。



4.3 画像認識「目を持ったAI」(2)

畳み込みニューラルネットワーク(CNN)と、リカレントニューラルネットワーク(RNN)の違いが分かる例である。RNNはどんな長さの文章でも処理できる**再帰型**ニューラルネットワークである。



4.3 画像認識「目を持ったAI」(3)

表現学習(representation learning)は、**深層学習**(ディープラーニング)を抽象化した概念である。

機械学習の手法を構成する際に、有用な情報を抽出することができる、すなわち**データの特徴表現**(あるいは、内部表現、**素性**(feature))を学ぶ学習の方法である。機械学習の性能は、データの表現に大きく依存しており、従来は人間の知識や職人技により、素性を構築することが広く行われてきた。

(**素性工学**(feature engineering)とも呼ばれる)

データの特徴表現を自動で学習するものが表現学習である。

4.3 画像認識「目を持ったAI」(4)

表現学習においては、現在**ディープラーニング**を用いる**アプローチが優勢**であるが、必ずしもその構造が必然的なわけではない。

現状ではかなりいい線をいっているのは確かであるが、将来的には、今のニューラルネットワークと全く違う形での、より理論的な表現学習の手法が出現する可能性はある。

4.3 画像認識「目を持ったAI」(5)

特徴表現は、通常は人間の知識や職人技によって定義されるが、それによって機械学習の性能が大きく異なってしまう。

あるいは、別の例では、画像に車が映っているかを判定したいとする。その際、ホイールが映っているか、ハンドルが映っているかなどを素性にすると良さそうであるが、画像から得られた画素の値という観測データ(あるいは生データ)からそれらの素性をどう構成するのかというのは自明ではない。

4.3 画像認識「目を持ったAI」(6)

特徴表現を学習する方法としては、従来から、様々な**クラスタリング法**、あるいは**主成分分析**(Principal Component Analysis;PCA)や**独立成分分析**(Independent Component Analysis;ICA)などの次元圧縮による手法がよく知られている。

近年では**ディープラーニング**に注目が集まっているが、これも**表現学習**の一つである。

ディープラーニングは、層の数が多いニューラルネットワークによって、観測データから本質的な情報を抽出した特徴表現を学習する。

4.3 画像認識「目を持ったAI」(7)

知的な情報処理とその学習における特徴表現の重要性は、AI、認知科学、機械学習、データ解析等の研究において古くから何度も指摘されてきた。

ディヴィッド・マー(David Marr)氏は、いかなる計算機の計算処理も、計算理論、表現とアルゴリズム、ハードウェアという3階層から理解され得ると述べた。

情報をどのように表現するかによって、アルゴリズムによる処理が大きく影響を受ける。

4.3 画像認識「目を持ったAI」(8)

よい「表現」とは何かというのは、単純なようで難しい問いである。

よい表現とは、何らかの意味で事象を抽象化したものであり、観測データの説明要因をとらえることで、一見自明ではない共通点をとらえることができるものである。

ヨシュア・ベンジオ(Yoshua Bengio)氏は、よい表現に共通するものとして、世界に関する多くの一般的な事前知識(あるいは事前分布、prior)を挙げている。

4.3 画像認識「目を持ったAI」(9)

代表的な事前知識に該当するものとして、

- (1)滑らかさ
 - (2)複数の説明要因
 - (3)説明要因の階層的構造**
 - (4)半教師あり学習
 - (5)タスク間の共通要因**
 - (6)多様体
 - (7)自然なクラスタ化
 - (8)時間的空間的一貫性
 - (9)スパース性(データの分布のまばらさ)
 - (10)要因の依存の単純性**
- などが挙げられる。

4.3 画像認識「目を持ったAI」

(10)

よい表現とは**スムーズな関数**を用いるものであり、また**時間的空間的一貫性**を持っている。

ディープラーニングのアプローチは、なかでも、**(3) (5) (10)**などに注目していることになる。

逆に言えば、こうした事前知識を適切に活用することができるなら、表現学習は必ずしも層の数が多いニューラルネットワークの形をしていなくても良いということになる。

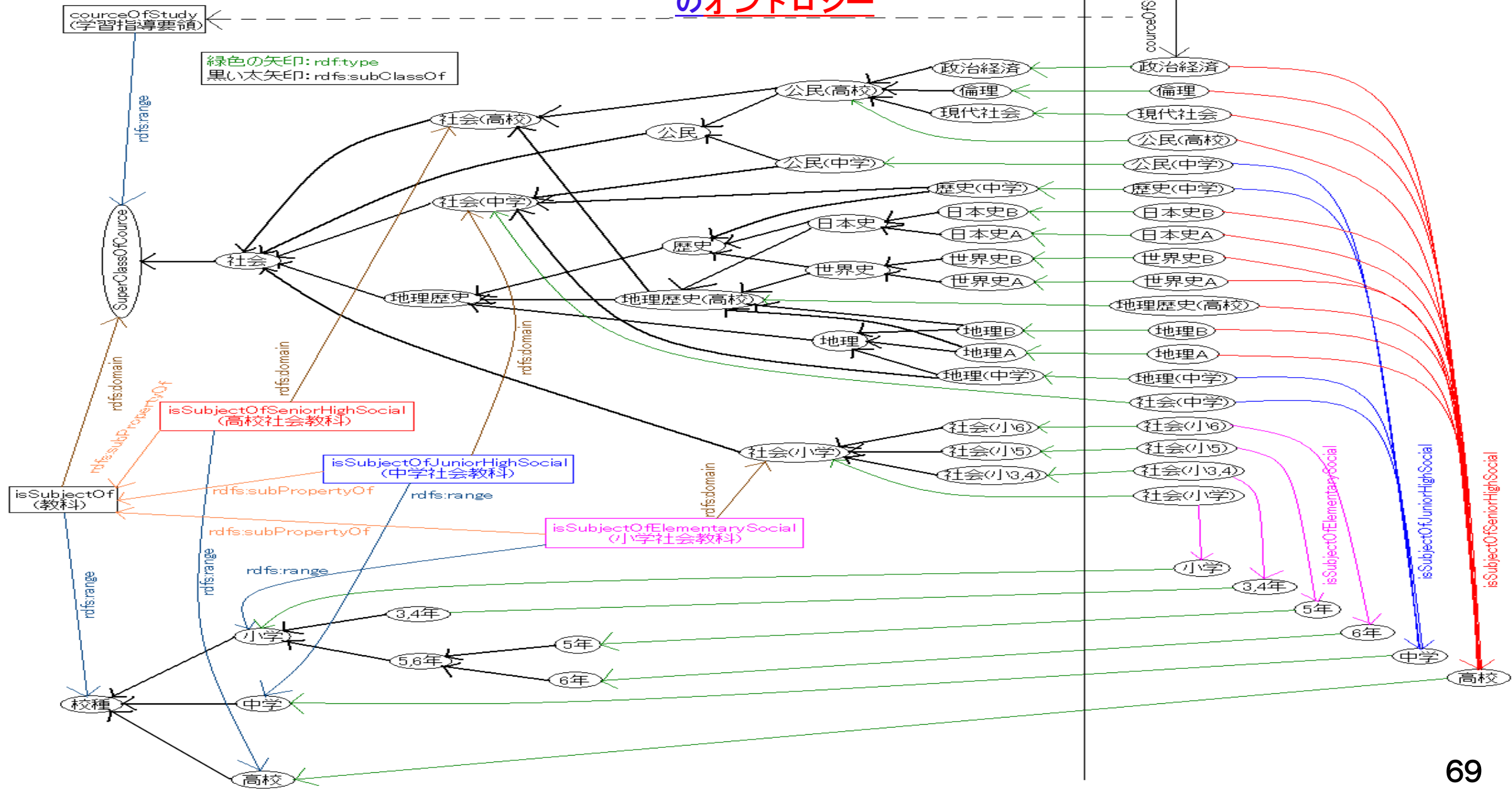
よい表現例

RDFS

例:社会科●学習指導要領 のオントロジー

教材アーカイブ
のURL

RDF



GOOGLE本社の人工知能の展示

一人の顔の表情や感情を読み花の色が変わる



4.3 画像認識「目を持ったA I」

(11)

ディープラーニングで特徴的であるのは、簡単な関数の組み合わせで難しい関数を構成することである。

通常、浅い構造よりも深い構造のほうがよりコンパクトに関数を表すことができる。そのことにより、素性の再利用と抽象的な概念(あるいは不変量)の獲得を可能にしている。

4.3 画像認識「目を持ったAI」 (12)

ディープラーニングによる表現の獲得の例として、「深層信念ネットワーク」(Deep Belief Network)を用いて、インターネット上の動画から切り出した画像を入力し、猫の概念を生成したというGoogleの研究が有名である。

また、抽象的な概念を頑健に獲得するには、設定の異なる複数のニューラルネットワークでの共通部分を見つけるとよいという最近の研究もある。

4.3 画像認識「目を持ったAI」 (13)

人間の知能においては、得られた抽象的な概念を、言語にマッピングし操作可能にしているところが特筆すべき点であろう。これに関連する有名な問題に、**シンボルグラウンディング問題**がある。AIにおける難問の一つである。

記号で指し示されるものを計算機がどのように認識するかという問題であり、概念に接続(グラウンド)されることなしには、記号処理が意味をなさないことを議論している。

シンボルグラウンディング問題とは



シマウマと人間はわかるが、AIにはわからない！



4.3 画像認識「目を持ったAI」

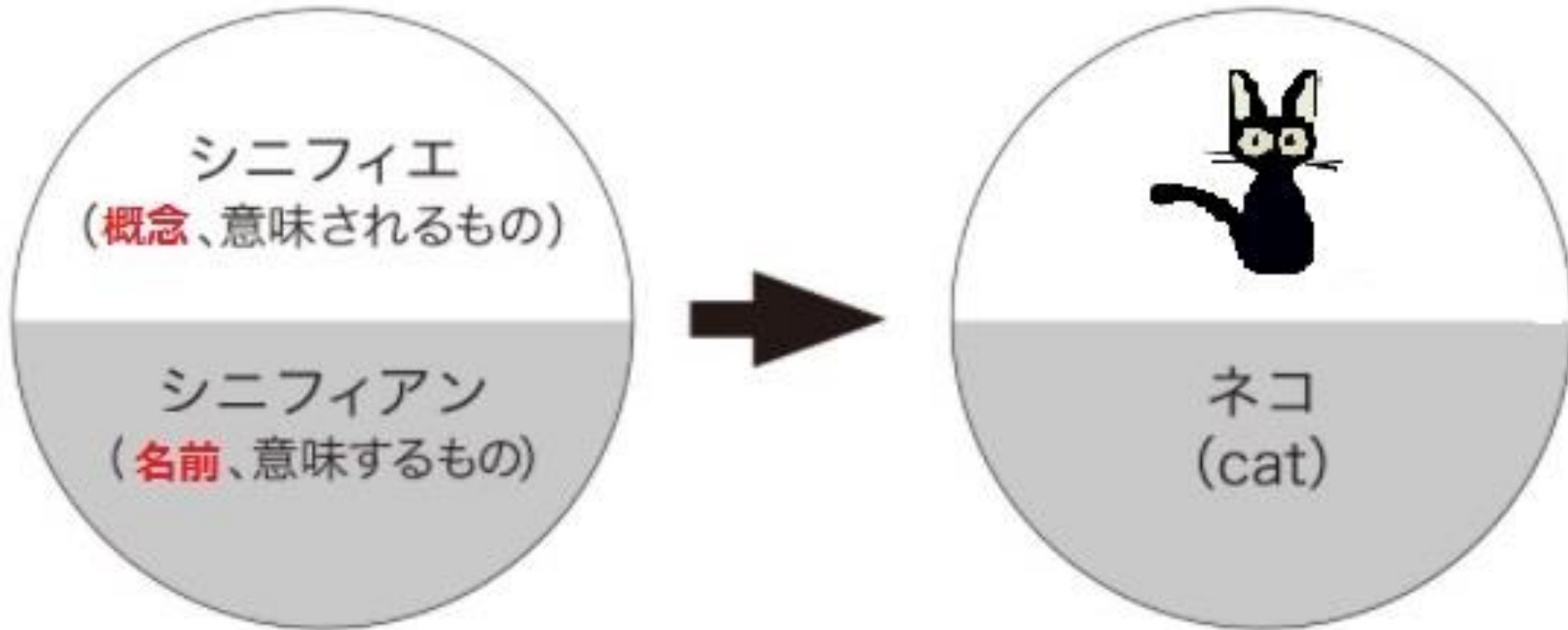
(14)

同様の議論は歴史的に古くからあり、**記号内容と記号表記を、シニフィエ、シニフィアンと呼んだ。**

概念と名前（言語）の結び付きは、最近でも活発に研究が行われており、またディープラーニングの文脈でも言語表現と画像特徴の調節を取るような研究も行われている。ディープラーニングを中心とする表現学習の方法の研究により、人間の持つ知性、特に言語を使った能力についての理解や、その工学的な応用も進んでいくことになるかもしれない。

シニフィエ、シニフィアンとは

概念は「シニフィエ」、名前は「シニフィアン」



4.3 画像認識「目を持ったAI」 (15)

表現学習においては、今のところディープラーニングのような深い構造を持ったニューラルネットワークを用いるアプローチが優勢であるが、必ずしもその構造が必然的なわけではない。

現状ではかなりいい線をいっているのは確かであるが、**将来的には、今のニューラルネットワークと全く違う形での、より理論的な表現学習の手法が出現する可能性があるだろう。**

4.4 畳み込みニューラルネットワーク (CNN) (1)

● 「ネオコグニトロンアーキテクチャ」 (その1)

これらのネットワークアーキテクチャは「ネオコグニトロン」と呼ばれるニューラルネットワークモデルの基本構造である。

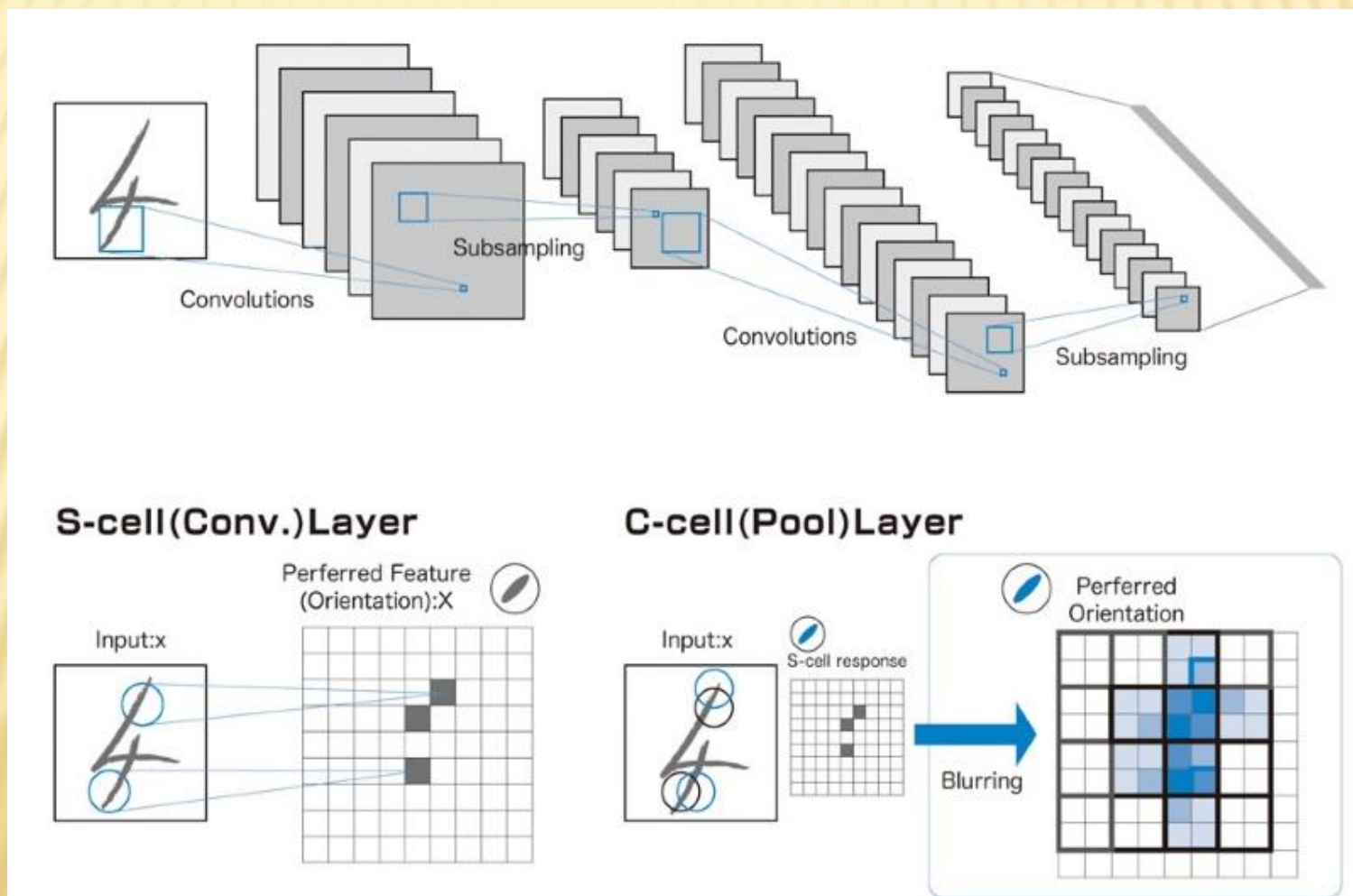
福島氏らのグループは、視覚野の高次機能を実現させるために、この基本アーキテクチャを使い、繰り返し表現の変換を行うことを提案している。

ネオコグニトロンは、局所的な特徴を抽出し、位置ずれを許容しながら情報を圧縮することで、徐々に受容野を拡大し、情報を統合する脳の視覚野モデルとなった。

福島氏らはネオコグニトロンを手書き文字認識などの問題に適用し、高度な識別器が実現できることを示した。

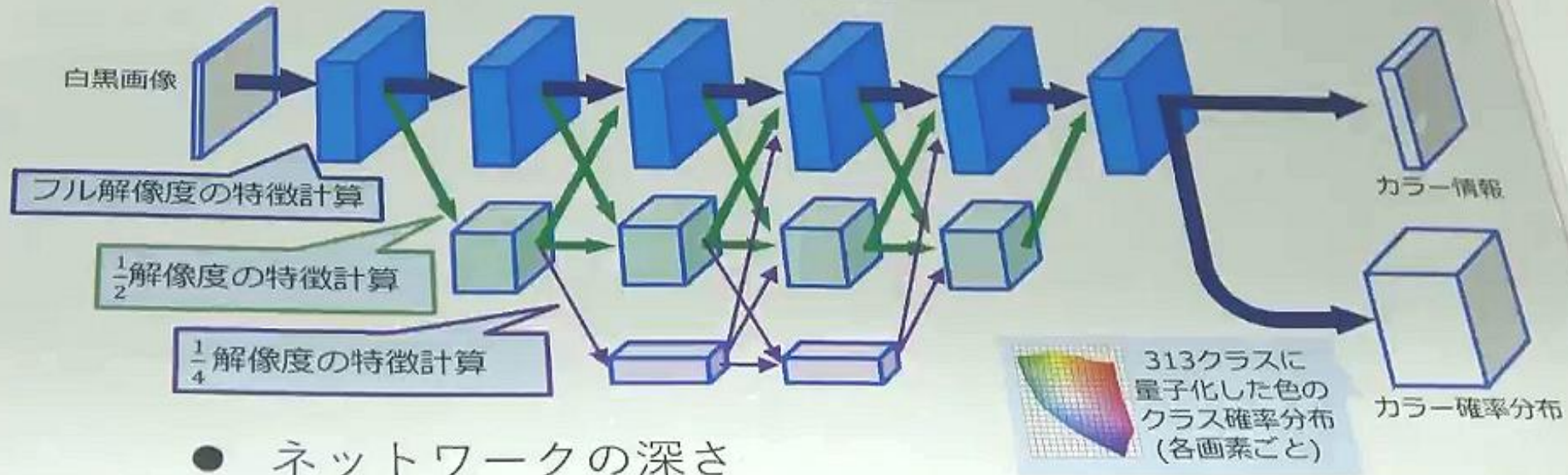
4.4 畳み込みニューラルネットワーク (CNN) (2)

● 「ネオコグニトロンアーキテクチャ」 (その2)



NHK 技術研究所の例

ニューラルネットワーク基本構造



- ネットワークの深さ
 - ✓ 約100層、約4000万パラメータ
- ロス関数
 - ✓ カラー情報、カラーエッジ情報の平均二乗誤差
 - ✓ カラー確率分布の交差エントロピー
- 学習環境
 - ✓ 約2万番組から作成した約800万枚の静止画を利用
 - ✓ Nvidia Quadro M6000で約2ヶ月学習

3

スマートプロダクション

白黒映像の自動カラー化技術

Automatic Colorization Technology for Monochrome Video Using Artificial Intelligence

貴重な白黒フィルムがカラーでよみがえる

人工知能を活用して、白黒映像を自動的にカラー映像に変換する技術を研究しています。この技術によって、貴重な白黒フィルム映像をカラー映像で見せるといった映像表現が可能となります。

入力 白黒映像

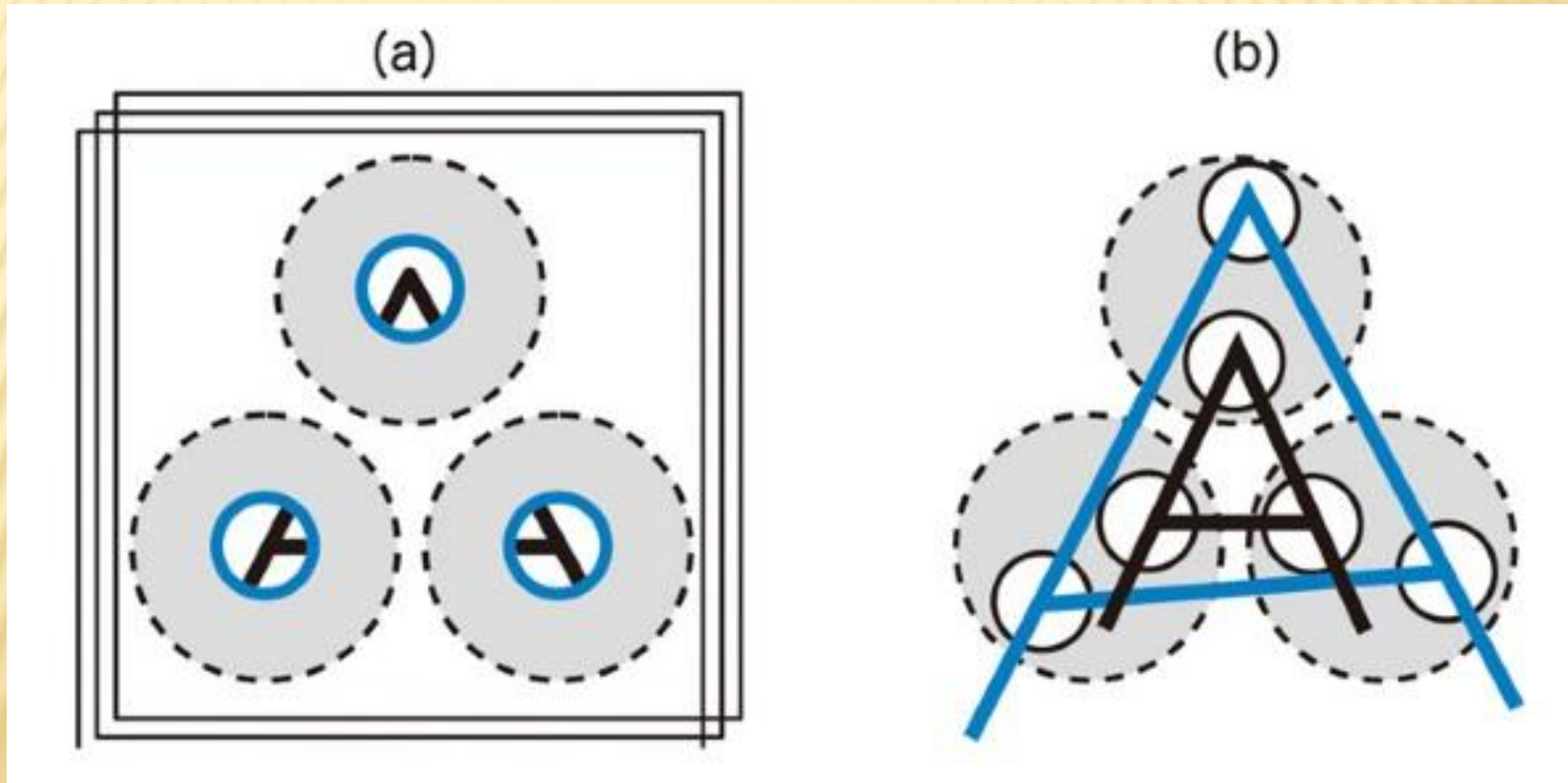


出力 カラー化映像



4.4 畳み込みニューラルネットワーク (CNN) (4)

● 図「CNNの動作原理」



4.4 畳み込みニューラルネットワーク (CNN) (3)

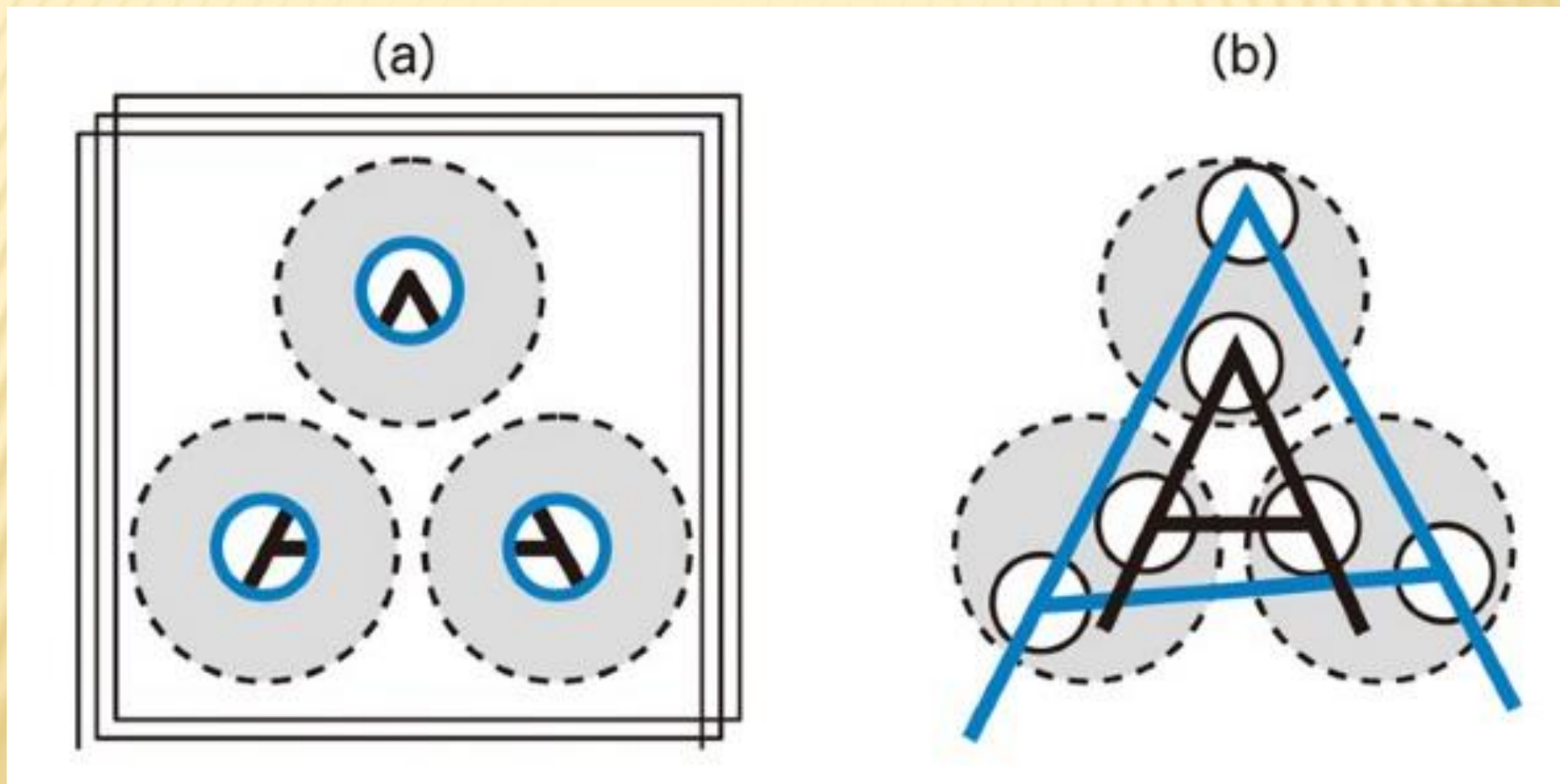
● 「CNNの動作原理」 (その1)

図はCNNの動作理解を行うための模式図である。図中の“**A**”という線画は、(b)の図に示した複数の“**A**”のように、異なる大きさと形状を持ち得るが、上端部の尖った形状とT字型の結合部の形状は、共通して持っている特徴ととらえられる。

CNNの内部の畳み込み層(Convolution Layer)は、前述の単純型細胞群に対応し、これらの特徴を別々の特徴マップとして表現する。このため、図の(a)のように、3種類の特徴マップで表現される。

4.4 畳み込みニューラルネットワーク (CNN) (4)

● 図「CNNの動作原理」 (その2)



4.4 畳み込みニューラルネットワーク (CNN) (5)

● 「CNNの動作原理」 (その3)

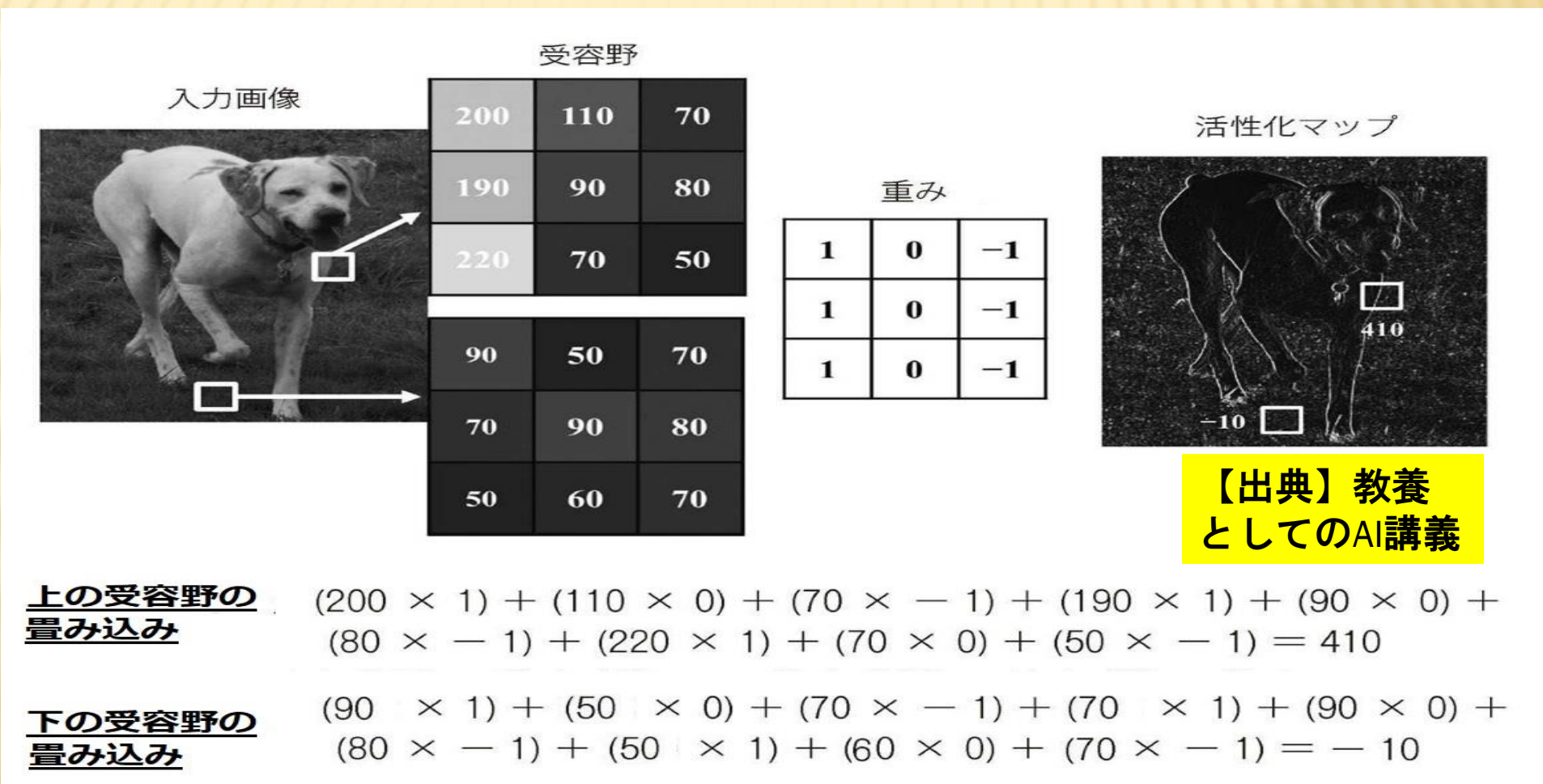
結果としてこれらの三つの特徴の組合せが揃っていれば、同じ特徴としてとらえる動作を行うため、変形などに強い特徴表現を構成できる。このようにCNNでは、ネットワークが深い階層を取るほど、複雑な組合せ特徴を表現できるような構造になっている。

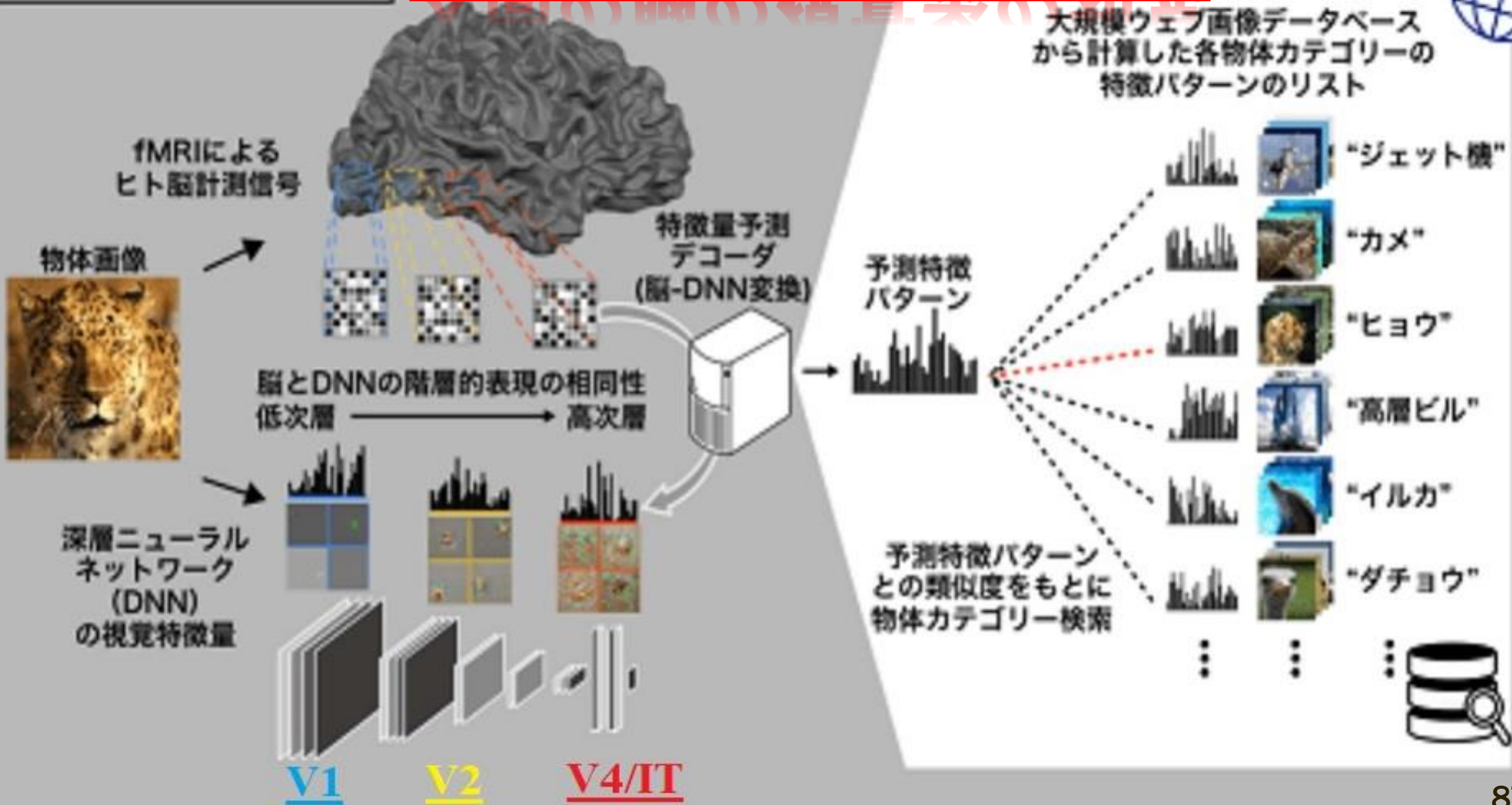
畳み込みニューラルネットワーク (CNN) は、「ネオコグニトロン」の影響を受けた手法で、脳の視覚系の知見に基づいて構築されている。

4.4 畳み込みニューラルネットワーク (CNN) (6)

● 図 「CNNの動作原理」 (その4)

畳み込みを利用した縦方向のエッジの検出例





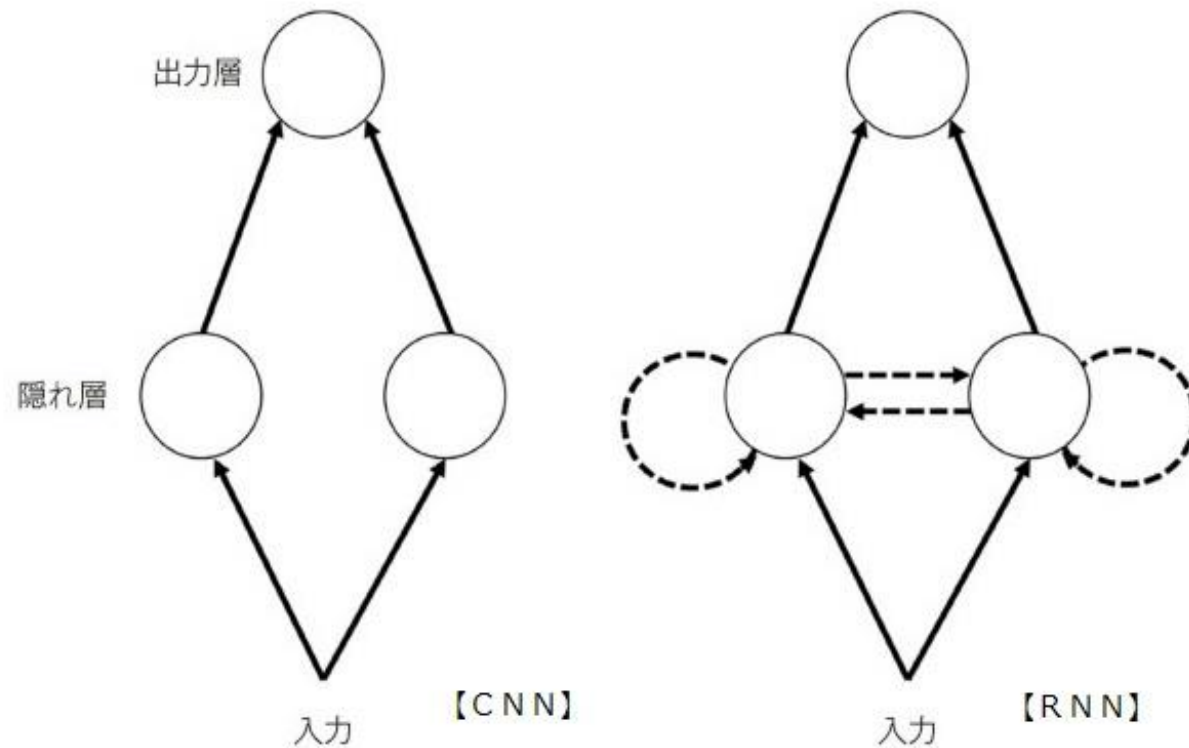
4.5 リカレントニューラルネットワーク (RNN) (1)

- 「リカレントニューラルネットワーク」とは再帰的な繋りを持つニューラルネットワークとして「リカレントニューラルネットワーク」(RNN)に対する研究も進んでおり、**自然言語や時系列データなどの連続性のあるデータに対して適用**されている。

(注) CNN(畳み込みニューラルネットワーク)は、**一方通行処理のニューラルネットワーク**である。

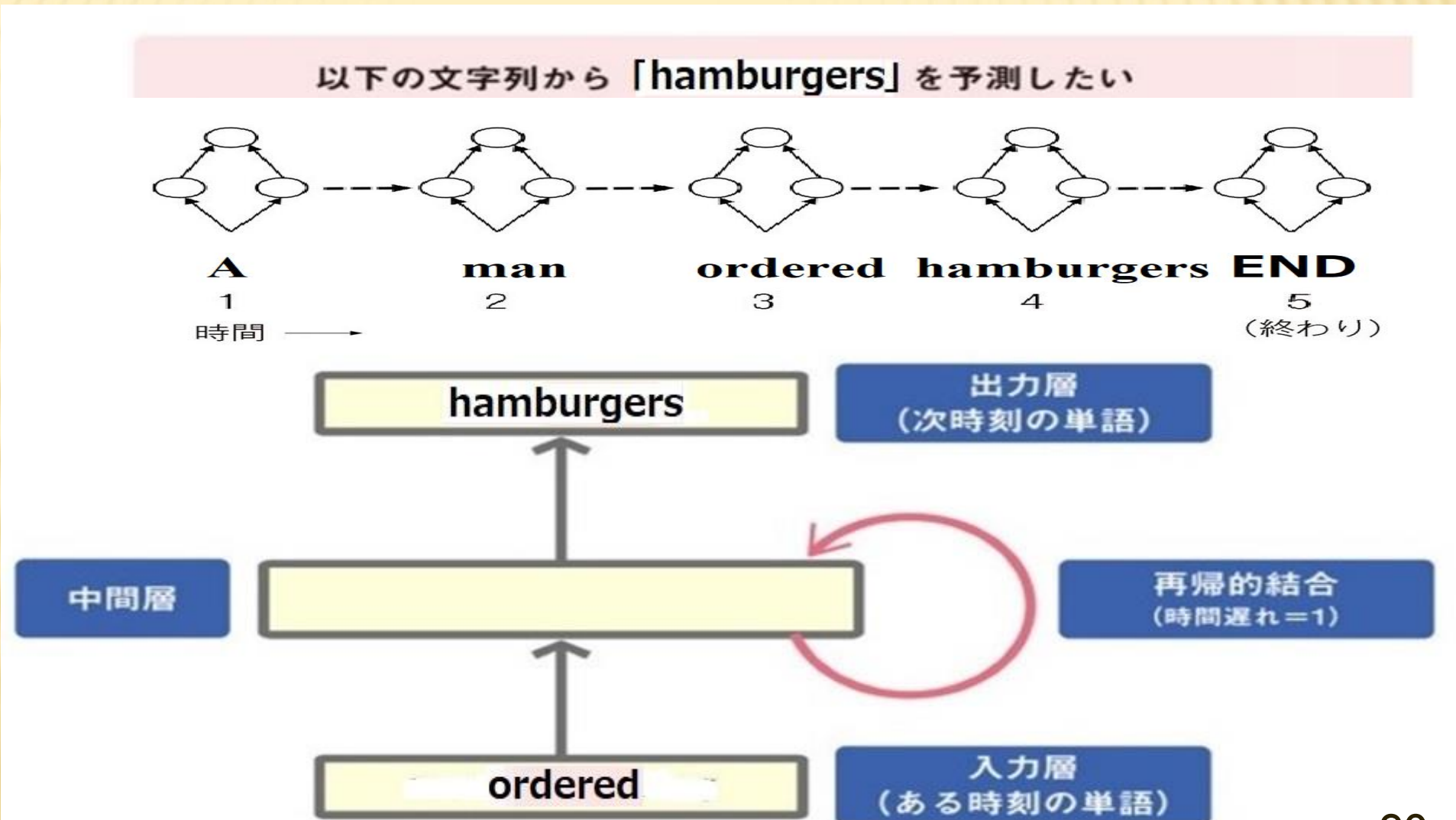
4.5 リカレントニューラルネットワーク (RNN) (2)

畳み込みニューラルネットワーク(CNN)と、**リカレントニューラルネットワーク (RNN)**の違いが分かる例である。RNNはどんな長さの文章でも処理できる**再帰型ニューラルネットワーク**である。



4.5 リカレントニューラルネットワーク (RNN) (3)

● RNNが5つの時間ステップ内での稼働例



4.5 リカレントニューラルネットワーク (RNN) (4)

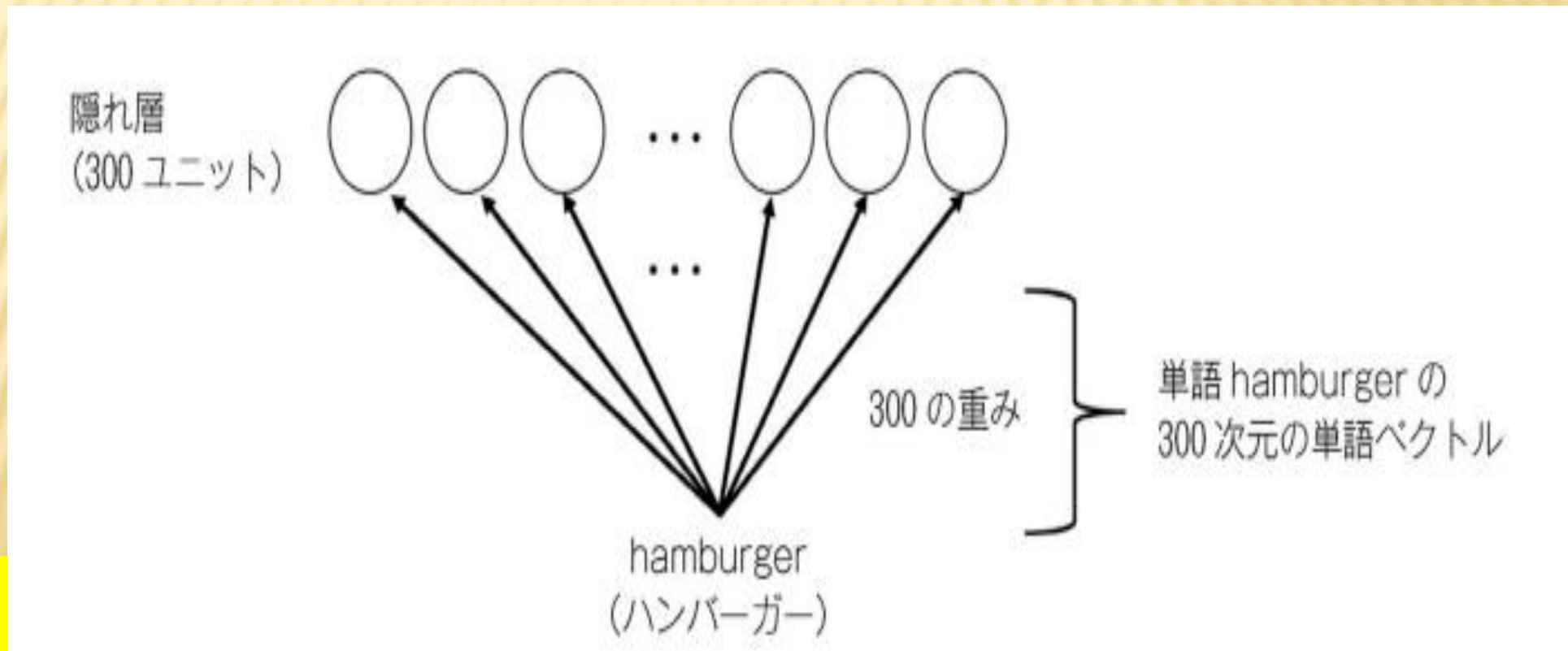
● 「リカレントニューラルネットワーク」とは

RNNは系列予測、姿勢制御、自然言語処理への応用などが考えられてきた。近年、RNNの学習に関する勾配消失、勾配爆発問題を回避する手法が一般化したこと、大量のデータをGPUにより高速に処理できるようになったことを受け、性能を向上させてきた。

2016年にはGoogleの自動翻訳サービスの精度が向上されたことが話題となったが、従来手法である統計的機械翻訳(Statistical Machine Translation;SMT)に対して、ニューラルネットワーク機械翻訳(Neural Machine Translation;NMT)が支配的になりつつある。

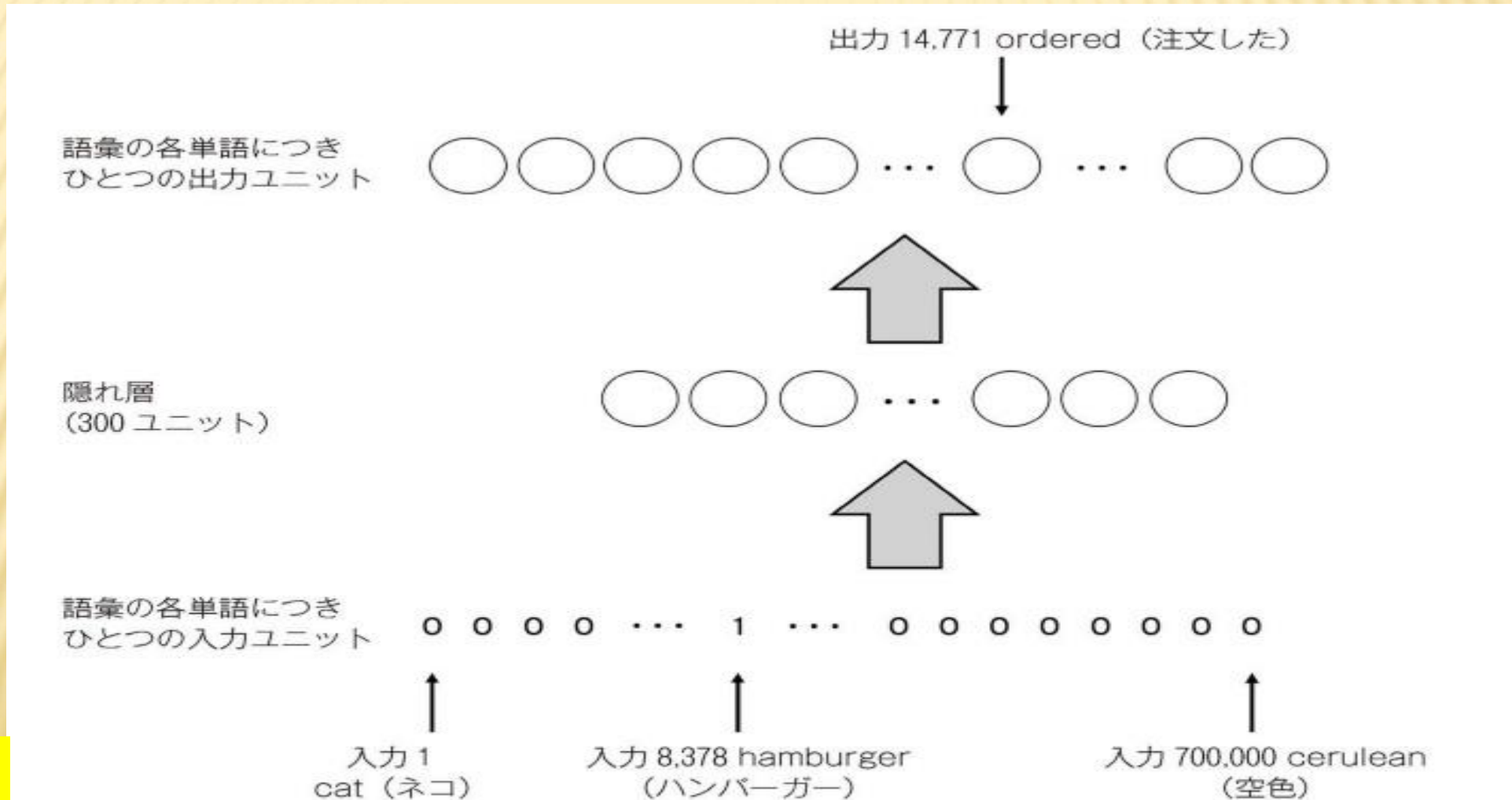
4.5 リカレントニューラルネットワーク (RNN) (5)

●Googleのword2vecプログラムの訓練データは単語を2つ一組を大量に集めたもので、**単語ベクトル**を例は以下の通り：



4.5 リカレントニューラルネットワーク (RNN) (6)

● 単語対 (hambueger,orderd) を入力した例



4.5 リカレントニューラルネットワーク (RNN) (7)

機械翻訳分野と画像処理において、任意の場所を選択的に処理する注意機構の導入は、画像と言語と領域は異なるものの数式は同一であり、**脳内でも同じような機構が仮定できる**のではないかと予想される。

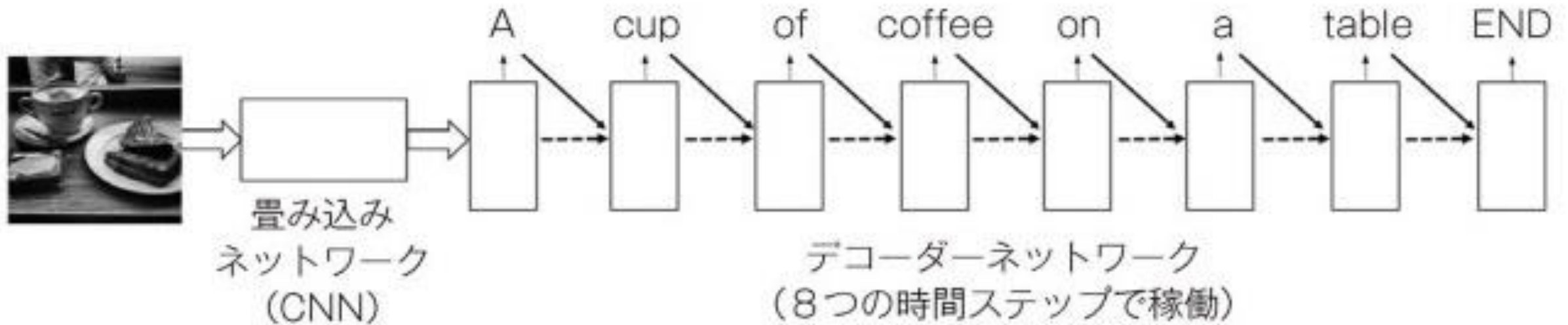
アーケードゲームを解くために、認識技術にCNNを用い、高得点を得るために強化学習を用いる手法は、囲碁において世界チャンピオンレベルの強さを有するに至った。

4.5 リカレントニューラルネットワーク (RNN) (8)

●画像を文章に翻訳する例

Googleの自動画像キャプションシステムの略図

左の画像から「テーブルの上にコーヒカップがある」という文章（キャプション）に翻訳する



本時のまとめ

- 1) コンピュータはA Iを実現するために産れた。
(A Iはコンピュータと表裏一体で開発された)
- 2) 人の脳の研究は、A Iにとって必然である。
(A Iは人の脳に似たコンピュータを創ること)
- 3) 深層学習 (ディープラーニング) がA Iを大きく変えた。

4.6 小テスト

問1 第3次AIブームで中心的な役割を果たした研究として正しいものはどれか？

- 1) 探索・推論の研究
- 2) 知識の研究
- 3) 機械学習、表現学習の研究
- 4) 時間の研究

4.6 小テスト

問1 第3次AIブームで中心的な役割を果たした研究として正しいものはどれか？

- 1) 探索・推論の研究
- 2) 知識の研究
- 3) 機械学習、表現学習の研究
- 4) 時間の研究

4.6 小テスト

問2 脳における学習の枠組みに基づくと、機械に学習させる「機械学習」には、三つの学習の枠組みがある。機械学習で無いものはどれか？

- 1) 教師あり学習
- 2) 教師なし学習
- 3) 強化学習
- 4) 家庭学習

4.6 小テスト

問2 脳における学習の枠組みに基づくと、機械に学習させる「機械学習」には、三つの学習の枠組みがある。機械学習で無いものはどれか？

- 1) 教師あり学習
- 2) 教師なし学習
- 3) 強化学習
- 4) 家庭学習

4.6 小テスト

問3 「機械学習」の「教師あり学習」と深い関連がある脳の部位はどれか？

- 1) 大脳皮質（大脳新皮質とも称する）
- 2) 大脳基底核
- 3) 小脳
- 4) 海馬

4.6 小テスト

問3 「機械学習」の「教師あり学習」と深い関連がある脳の部位はどれか？

- 1) 大脳皮質（大脳新皮質とも称する）
- 2) 大脳基底核
- 3) 小脳
- 4) 海馬

4.6 小テスト

問4 「機械学習」の「教師なし学習」と深い
関連がある脳の部位はどれか？

- 1) 大脳皮質（大脳新皮質とも称する）
- 2) 大脳基底核
- 3) 小脳
- 4) 海馬

4.6 小テスト

問4 「機械学習」の「教師なし学習」と深い
関連がある脳の部位はどれか？

- 1) 大脳皮質（大脳新皮質とも称する）
- 2) 大脳基底核
- 3) 小脳
- 4) 海馬

4.6 小テスト

問5 「機械学習」の「強化学習」と深い関連がある脳の部位はどれか？

- 1) 大脳皮質（大脳新皮質とも称する）
- 2) 大脳基底核
- 3) 小脳
- 4) 海馬

4.6 小テスト

問5 「機械学習」の「強化学習」と深い関連がある脳の部位はどれか？

- 1) 大脳皮質（大脳新皮質とも称する）
- 2) 大脳基底核
- 3) 小脳
- 4) 海馬

4.6 小テスト

問6 表現学習に関して、ディープラーニングのアプローチで注目している「よい表現」に当てはまらないものはどれか？

- 1) 説明要因の階層的構造
- 2) タスク間の共通要因
- 3) 一般的な事前知識の無い構造
- 4) 要因の依存の単純性

4.6 小テスト

問6 表現学習に関して、ディープラーニングのアプローチで注目している「よい表現」に当てはまらないものはどれか？

- 1) 説明要因の階層的構造
- 2) タスク間の共通要因
- 3) 一般的な事前知識の無い構造
- 4) 要因の依存の単純性

4.6 小テスト

問7 脳の視野中の形状に関する視覚信号は、腹側経路と呼ばれる経路で処理されるが、無関係なものはどれか？

- 1) V1野
- 2) V2野
- 3) ROBOT野
- 4) V4野

4.6 小テスト

問7 脳の視野中の形状に関する視覚信号は、腹側経路と呼ばれる経路で処理されるが、無関係なものはどれか？

- 1) V1野
- 2) V2野
- 3) ROBOT野
- 4) V4野

4.6 小テスト

問8 リカレントニューラルネットワーク (RNN) に関係が無いのはどれか？

- 1) 「ネオコグニトロン」と呼ばれる基本構造
- 2) 再帰的な繋りを持つニューラルネットワーク
- 3) 系列予測、姿勢制御、自然言語処理への応用
- 4) 自然言語や時系列データなどの連続性のあるデータに対して適用

4.6 小テスト

問8 リカレントニューラルネットワーク (RNN) に関係が無いのはどれか？

- 1) 「ネオコグニトロン」と呼ばれる基本構造
- 2) 再帰的な繋りを持つニューラルネットワーク
- 3) 系列予測、姿勢制御、自然言語処理への応用
- 4) 自然言語や時系列データなどの連続性のあるデータに対して適用

【出典】（教科書に準ずるもの）

1) COMPUTER HISTORY MUSEUM IN USA

2) AI白書2017-20 IPA編KADOKAWA出版

2018年6月10日 / 2018年12月10日発行

3) ディープラーニング[®]G検定公式テキスト

一般社団法人日本ディープラーニング協会監修 翔泳社出版

2018年10月29日発行

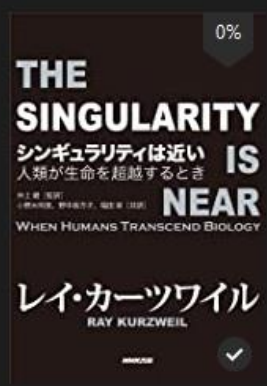
4) ディープラーニング[®]活用の教科書 !

杉山俊幸著 日経BP社出版 2018年10月29日発行

【出典】 (教科書に準ずるもの)



超AI入門 ディープラ...
松尾 豊とNHK「人間って...



シンギュラリティは近い...
レイ・カーツワイル



深層学習教科書ディ...
浅川 伸一, 江間 有沙, 工...



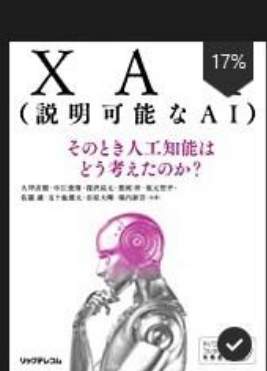
教養としてのAI講義 ...
メアリー・ミッチェル



[図解] 大学4年間の...
久野 遼平と木脇 太一



AI白書 2020(単行...
独立行政法人情報処理...



XAI(説明可能なAI)
大坪直樹と中江俊博



図解 人工知能大全
古明地 正俊と長谷 佳明



パソコンで楽しむ自分...
中島能和



ディープラーニング活用...
日経クロストレンド