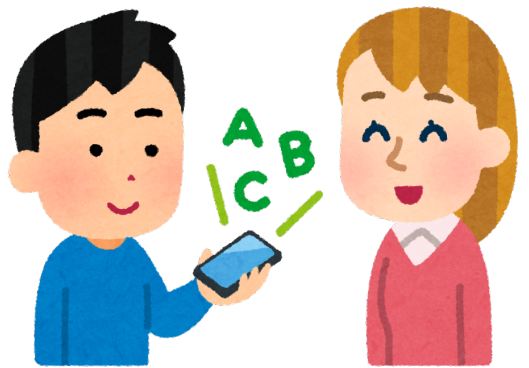


データサイエンス・リテラシー

身近な生活でのデータサイエンス活用



イラスト：©いらすとや



Table of Contents

機械翻訳

Online翻訳アプリ : Weblio, Google Translate, DeepL

コミュニケーションロボット

aibo



課題の解決（ソリューションの構築）

課題

□ 困りごととは何か？

- 課題の本質は？
- 特定の場合だけではだめ

誰のため

誰のための技術？ 誰が嬉しい？

技術コア

使われている技術

解決策

どうやって課題を解決したか？

学習

入力データ ↔ 正解データ

検証

入力データ → 予想データ
(⇔正解データ)

機械翻訳

言語の翻訳

- 他言語・多言語で多くの人と意思疎通を図りたい。
- 間に人を介さないでやりとりをしたい
→ でも、言語習得には多大な時間と労力が必要…。

誰のため

他言語を話す人とも意思疎通を図りたい人
他言語・多言語で情報をすぐに発信・受信したい人

技術コア

機械翻訳（ルールベース、フレーズベース、ニューラルベース）

解決策

元の言語の文章等を入力すると所望の言語に翻訳して出力
 ・テキストデータを入力すると所望の言語でのテキストを出力
 ・声で入力し、声で出力することも可能（カメラで文字入力も可）
 → インターネットにつながる機器から入出力することで同時通訳に近い使い方も可能

Weblio 英語翻訳

<https://translate.weblio.jp/>

Google 翻訳

<https://translate.google.com/>

DeepL 翻訳ツール

<https://www.deepl.com/>

機械翻訳（まずは簡単な短文から 日→英）

□ 東京駅までの行き方を教えてください。



DeepL 翻訳ツール

- How do I get to Tokyo Station?

Weblio 英語翻訳

- Please tell me the way to Tokyo Station.

Google 翻訳

- Please tell me how to get to Tokyo Station.

機械翻訳（まずは簡単な短文から 日→英）

□ 東京駅までの行き方を教えてください。



DeepL 翻訳ツール

● How do I get to Tokyo Station?

[再翻訳] → 「東京駅までの行き方は？」

Weblio 英語翻訳

● Please tell me the way to Tokyo Station.

[再翻訳]

→ 「私に東京駅への行き方を教えてください。」

Google 翻訳

● Please tell me how to get to Tokyo Station.

[再翻訳]

→ 「東京駅への行き方を教えてください。」

イラスト：©いらすとや

機械翻訳（まずは簡単な短文から 日→英）

□注文は、私はオムライス、彼はうなぎです。

Weblio 英語翻訳

- As for the order, as for me, as for omelette with rice, him, it is an eel.



DeepL 翻訳ツール

- We **ordered** an omelet for me and an eel for him.

Google 翻訳

- The order is omelet rice and **he is eel.**

イラスト：©いらすとや

機械翻訳（まずは簡単な短文から 日→英）

□注文は、私はオムライス、彼はうなぎです。

Weblio 英語翻訳

- As for the order, as for me, as for omelette with rice, him, it is an eel.



DeepL 翻訳ツール

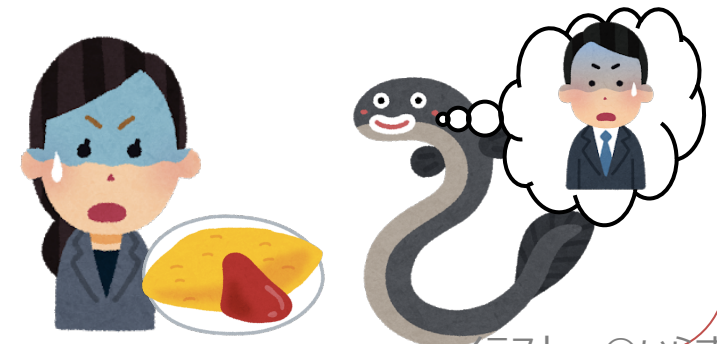
- We **ordered** an omelet for me and an eel for him.

過去形：
注文が間違っていた、
と解釈した？



Google 翻訳

- The order is omelet rice and **he is eel.**



機械翻訳（まずは簡単な短文から 英→日）

□ We will take care of it right away.

Weblio 英語翻訳

- 我々は、すぐにそれを引き受けます。
- 我々は、すぐにそれに気をつけます。
- 我々は、その世話を正に取りります。

DeepL 翻訳ツール

- すぐに対応させていただきます。
 - すぐに対応させていただきます。
 - すぐに対応します。
 - すぐに対応いたします。

Google 翻訳

- すぐに対応いたします。

機械翻訳（長文もやってみよう）

□ instead, it had been expected that Trump would make a brief, less formal “working visit” next month, to cut the ribbon on the \$1bn embassy in Nine Elms, south-west London, and hold meetings with May.

Weblio 英語翻訳

- その代わりに、トランプが来月短い、より正式でない「働く訪問」をすると予想されました。そして、9人のエルムズ（南西ロンドン）の10億ドルの大使館でリボンを裁断して、メイとの会談を開きました。

DeepL 翻訳ツール

- 代わりに、それはトランプ氏は、ナインエルムズ、南西ロンドンの10億ドルの大使館のリボンをカットし、メイとの会議を保持するために、来月、短い、あまり正式な「作業訪問」を行うことが予想されていた。

Google 翻訳

- 代わりに、トランプ氏が来月、ロンドンの南西にあるナインエルムズの10億ドルの大使館のリボンを切り、5月との会合を開くために、短い「正式な訪問」を行うと予想されていました。

機械翻訳（長文もやってみよう）

□ Does the Pfizer-BioNTech COVID-19 Vaccine cause infertility?

It has been suggested that COVID-19 vaccines will cause infertility because of a shared amino acid sequence in the spike protein of SARS-CoV-2 and a placental protein. Although the SARS-CoV-2 spike protein shares an amino acid sequence with a placental protein, the two proteins are immunologically different and distinct.

[THE FACTS ABOUT THE PFIZER-BIONTECH COVID-19 VACCINE\(PFIZER\)](#)

Weblio 英語翻訳

COVID-19ワクチンがSARS-CoV-2と胎盤タンパク質のスパイク・タンパク質で共有アミノ酸配列のため、不妊性を引き起こすことが示唆されました。

SARS-CoV-2スパイク・タンパク質がアミノ酸配列を胎盤タンパク質と共有するが、2つのタンパク質は免疫学的に異なって明瞭です。

DeepL 翻訳ツール

SARS-CoV-2のスパイクタンパクと胎盤タンパクのアミノ酸配列が共通していることから、COVID-19ワクチンが不妊の原因になるのではないかと言われています。
SARS-CoV-2のスパイクタンパク質は胎盤タンパク質とアミノ酸配列を共有していますが、この2つのタンパク質は免疫学的に異なる別のタンパク質です。

Google 翻訳

SARS-CoV-2 のスパイクタンパク質と胎盤タンパク質のアミノ酸配列が共有されているため、COVID-19 ワクチンが不妊症を引き起こすことが示唆されています。 SARS-CoV-2 スパイクタンパク質は、胎盤タンパク質とアミノ酸配列を共有していますが、2つのタンパク質は免疫学的に異なり、異なるものです。

機械翻訳

ロデータサイエンス・AI技術

● 機械翻訳

- ルールベース機械翻訳
- フレーズベース統計型機械翻訳
- ニューラルベース機械翻訳

ロ利用データ

● 学習時

- 対訳文章 (コーパス)

● 検証時 (実用時)

- 任意の文章
- 話し言葉

学習

ルールベース

文章を単語に分解
単語・熟語の役割を確認
構文 (ルール) を決定

様々な構文や
単語・熟語を
学習

1. 形態素解析
2. 構文解析
3. 意味解析
4. 文脈解析

フレーズベース統計型

ルールベースのプロセスに加えて、
単語間のよくある組合せから
単語の意味を予測して翻訳

対訳文章 (コーパス)
から得られる統計

例: arm → 腕?, 武器?
like → 望む?, 好む?

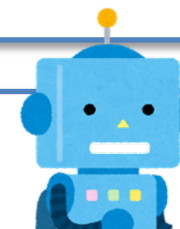
ニューラルベース

入力言語 → 中間表現 → 出力言語
の順に変換するニューラルネットを学習して翻訳

検証

入力

伝えたい、訳したい
元の文章



出力

目的の言語に
翻訳された文章

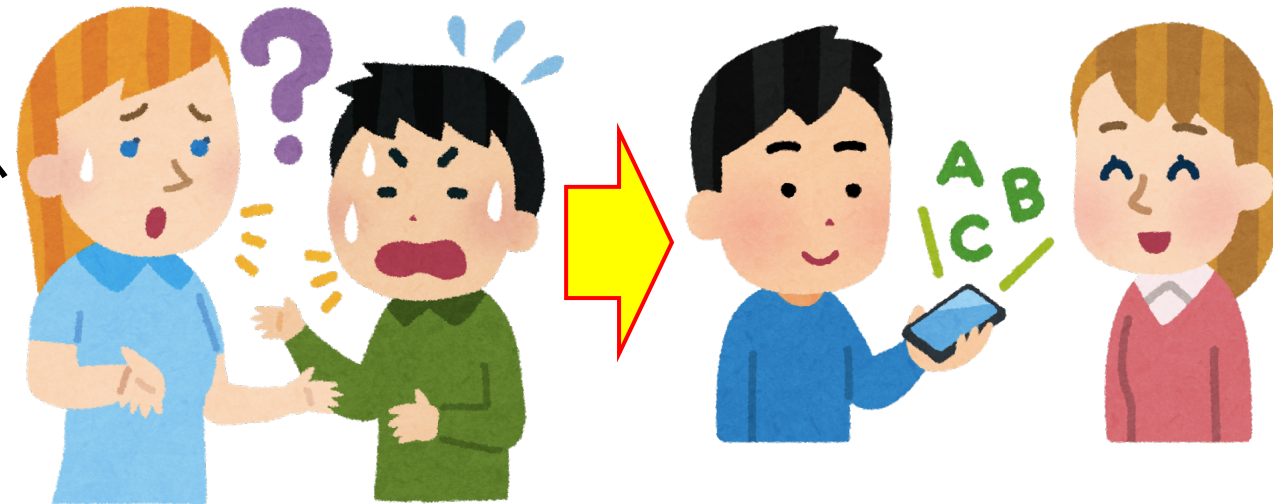
サービスイメージ

- ホームページやブログを自国語に翻訳
- レストランでの注文や説明の翻訳
- ホームページやブログの他言語対応



できるようになったこと

- 知らない他言語を習得しなくても、翻訳して多言語で表現できるようになった。
- 同一言語を話せなくても気軽に意思疎通が図れるようになった。



□実運用での課題（想定）

- 翻訳精度の向上
- 翻訳には高度な計算が必要なため
インターネット経由で計算サーバに接続
が必要

□ 注文は、私はオムライス、彼はうなぎです。

- The order is omelet rice and **he is eel.** ?!

?!

- We **ordered** an omelet for me and unagi for him.

□他の課題への転用

- 多国籍会議での同時通訳
- 手話の翻訳
- 動物の言葉の翻訳?!



ゴロゴロ～ 苦痛？でもそのままにしておいて
 みやお 感心を引きたい
 シャー！ 攻撃、威嚇

スザンヌ シェツ (著), 石田 紀子 (翻訳):
 「猫語のひみつ」, ハーパーコリンズ

Table of Contents

機械翻訳

Online翻訳アプリ : Weblio, Google Translate, DeepL

コミュニケーションロボット



ペットロボット（エンタテインメントロボット）

課題

口動物が好きだけどアレルギーがあってペットを飼う事ができない

口高齢なのでペットを飼いたくても十分に面倒が見られない



誰のため

ペットを飼いたいが、高齢やアレルギーなどのために飼う事ができない人

技術コア

カメラ映像から、AIにより、ここはどこなのか、周りに誰がいるのかを理解し行動を変える。呼ばれた方向も認識

解決策

- 犬を飼っているような体験ができる
- ・周りの環境を理解し、状況にあった行動や仕草で癒しを与える
- ・オーナーとのやりとりを学習し、一体一体異なる成長を楽しめる
- ・知らない人には威嚇

イラスト：©いらすとや

ペットロボット (エンタテインメントロボット)

□ データサイエンス・AI技術

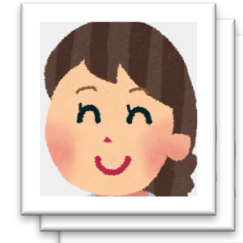
- ディープラーニング
- 顔認識技術

□ 利用データ

- 学習時
 - 撮影画像から顔検出
 - ➔ 顔画像の記憶
 - ➔ 顔認証による飼主の検知
 - 撮影画像から充電器を検出
 - 部屋の中を歩き回って空間を認識
 - ➔ 人や充電器の位置を認識
- 検証時 (実用時)
 - 画像データ + 空間認識情報
 - ➔ 充電器や飼主の認識 + 位置情報の予測

学習

顔認識



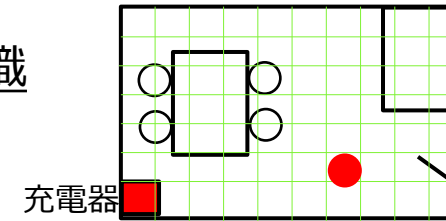
よく見る人



初めて見る人

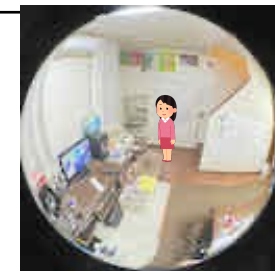
充電台
他のペット
おもちゃ
.....

空間認識



検証

入力



出力



飼主

(150,100)

充電器

(100,200)

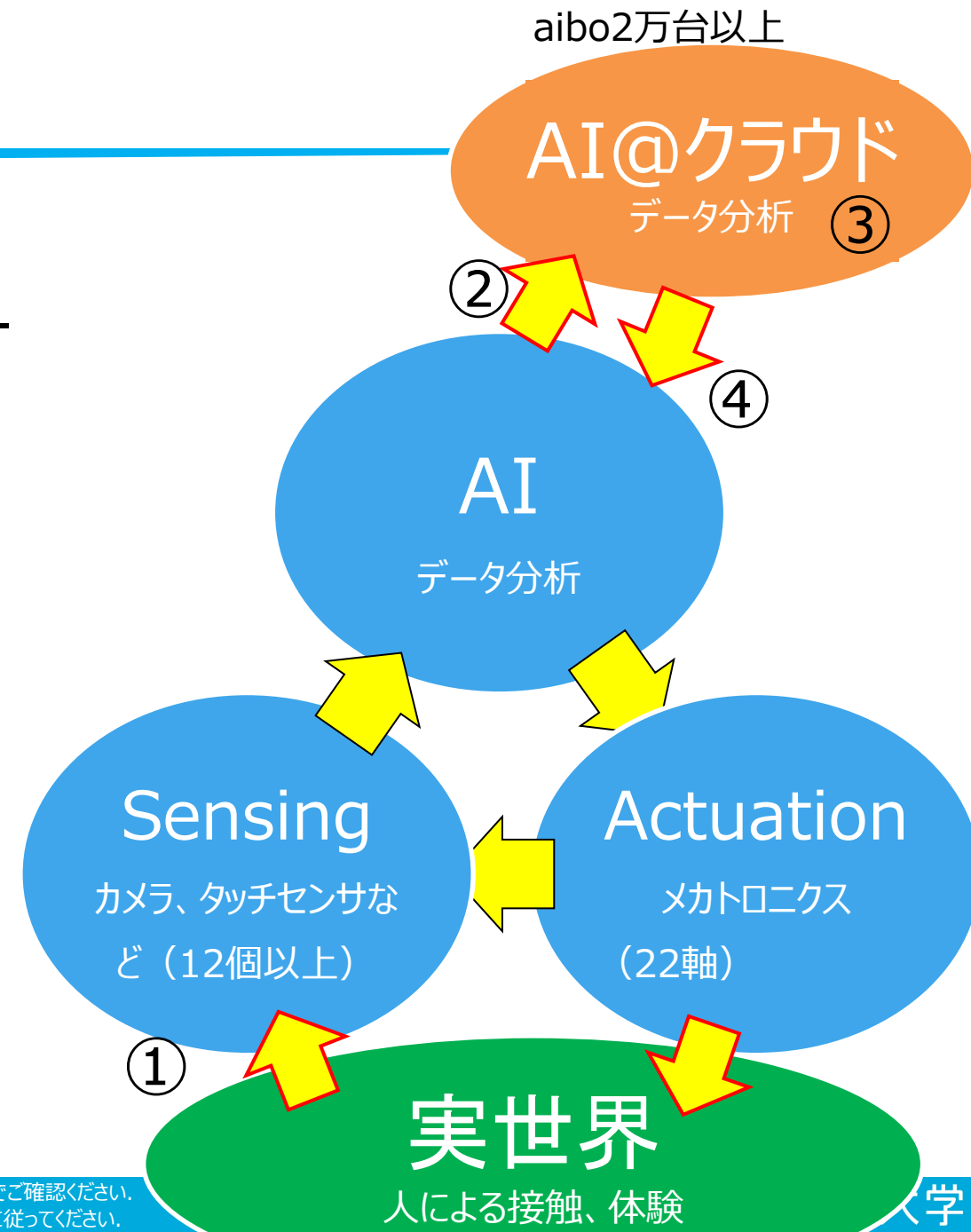
aibo (アイボ)

サービスイメージ

- ① 褒められ方、状況を個々のaiboがエピソードとして記憶
- ② エピソードをクラウドにも蓄積
- ③ 飼い主に愛される行動をクラウドで学習
- ④ クラウドで学習した結果を個々のaiboにFeedback

できるようになったこと

- 人の顔・触れ合い・体験により性格や動きが変化する
- aiboに個性が生まれ、リアルな犬を飼っているような体験



Deep Learningとは

イラスト：©いらすとや

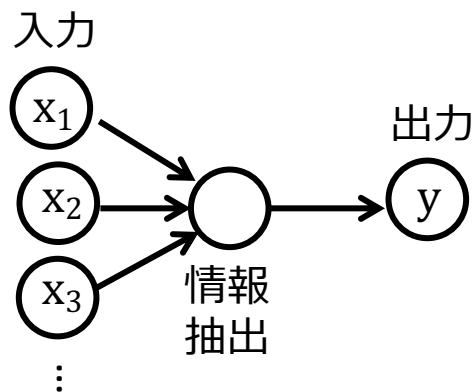
脳の学習機能をコンピュータでシミュレーションするニューラルネットワークを用いた技術である。
近年の技術的な発展の背景には、計算機能力の向上がある。

神経細胞



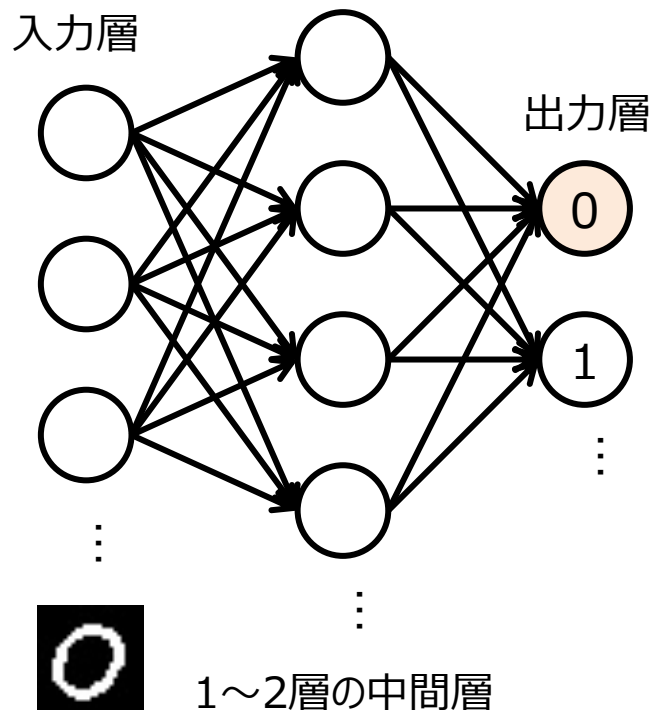
複数の細胞からの入力刺激を受け、
情報を修飾加工して他の細胞に伝達

人工ニューロン



ニューラルネットワーク

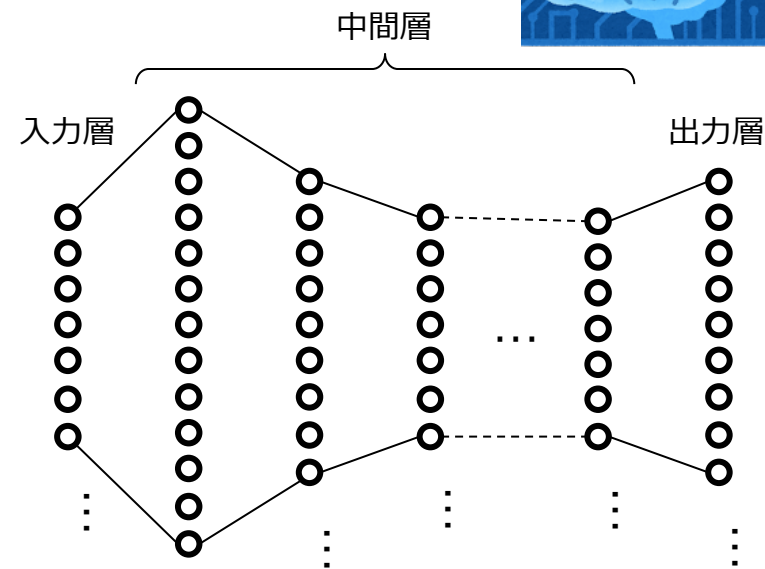
(1960~1990頃)



1~2層の中間層

Deep Learning

(2006~)



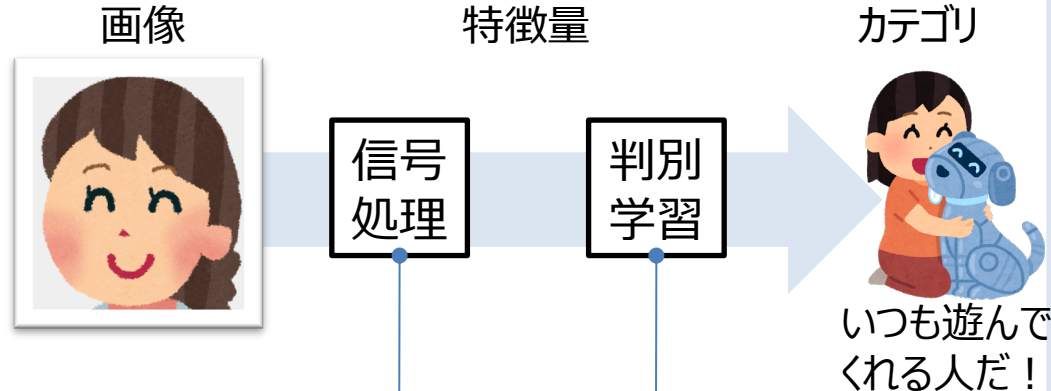
3層~1000以上の中間層
数十M規模のニューロン数

大規模なニューラルネットワークの
学習が可能になり、大幅に性能向上

Deep Learningにより認識機全体を自動的に獲得

従来型の機械学習は、特徴量抽出のノウハウが性能を左右していた。
Deep Learningでは、特徴量抽出は自動的に学習されるため、データの量と質が性能を決める。

従来型機械学習による画像認識

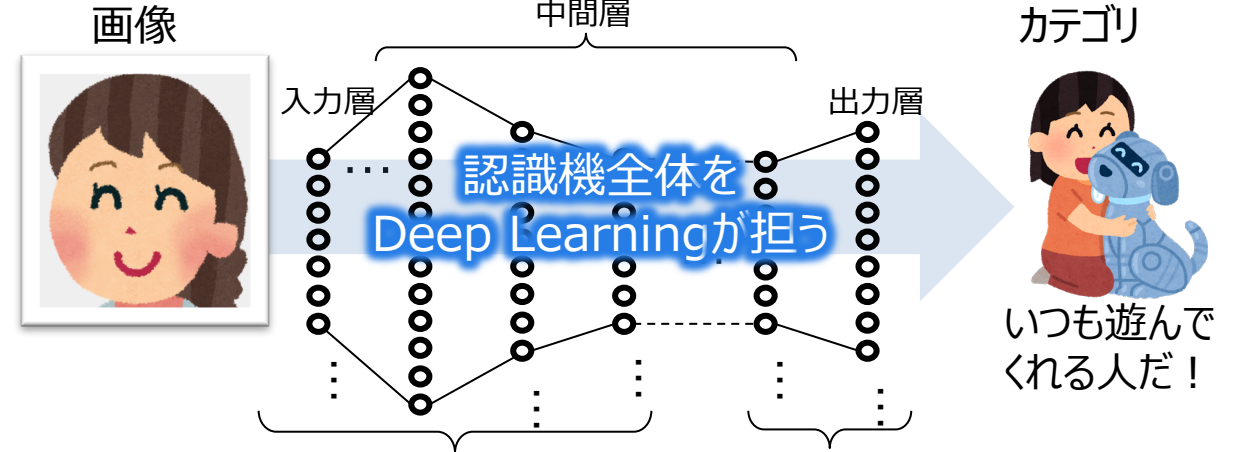


画像信号処理により、形、色、テクスチャ等画像の特徴を表す特徴量を抽出（手作り）

特徴量を元に画像種類を推定する判別学習を行う（機械学習）

- 特徴量を抽出する信号処理の設計に多大な労力と専門知識を要していた
- 最適化に限界があった

Deep Learningによる画像認識



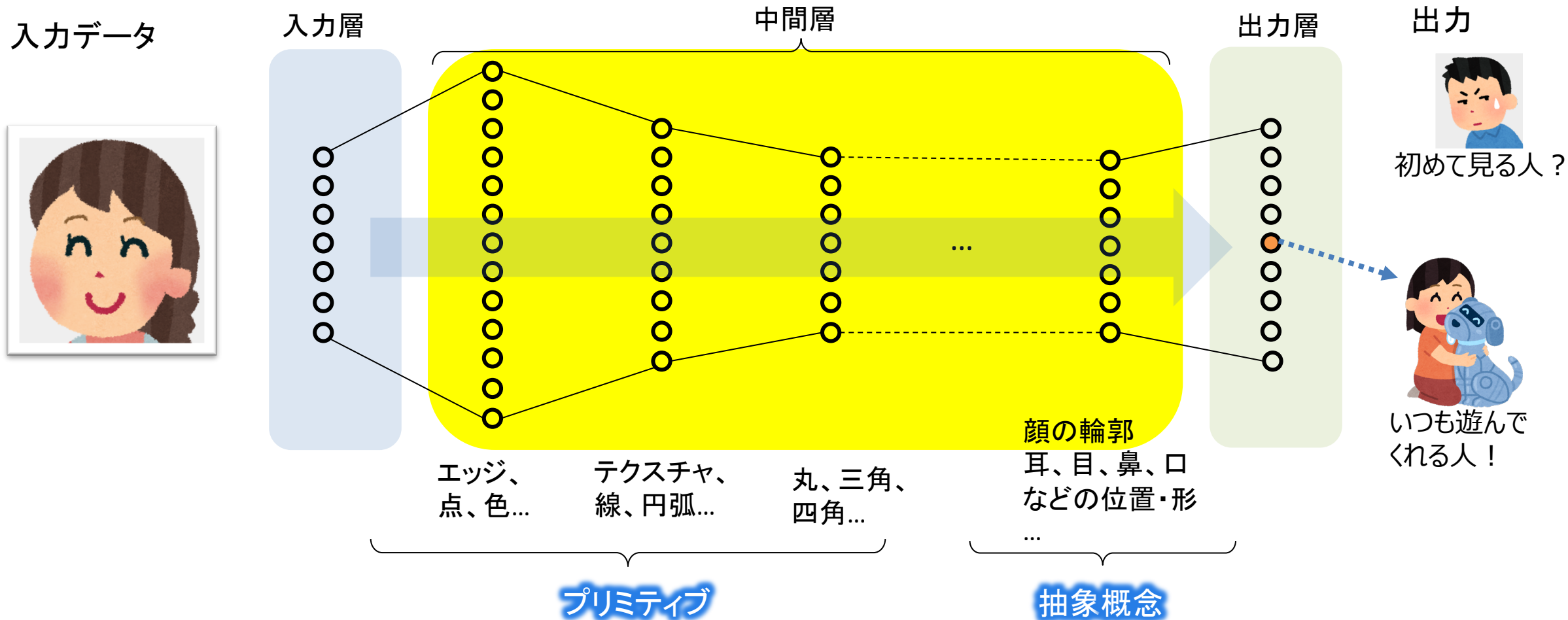
入力に近い層には「特徴量抽出」に相当する処理が自動的に学習される
出力に近い層には従来の判別学習に相当する処理が学習される

- 所望の入出力関係を与えるだけで、認識機を獲得可能に
- データに合わせた特徴量の最適化が可能になり精度が向上

イラスト：©いらすとや

学習されたニューラルネットワークの分析

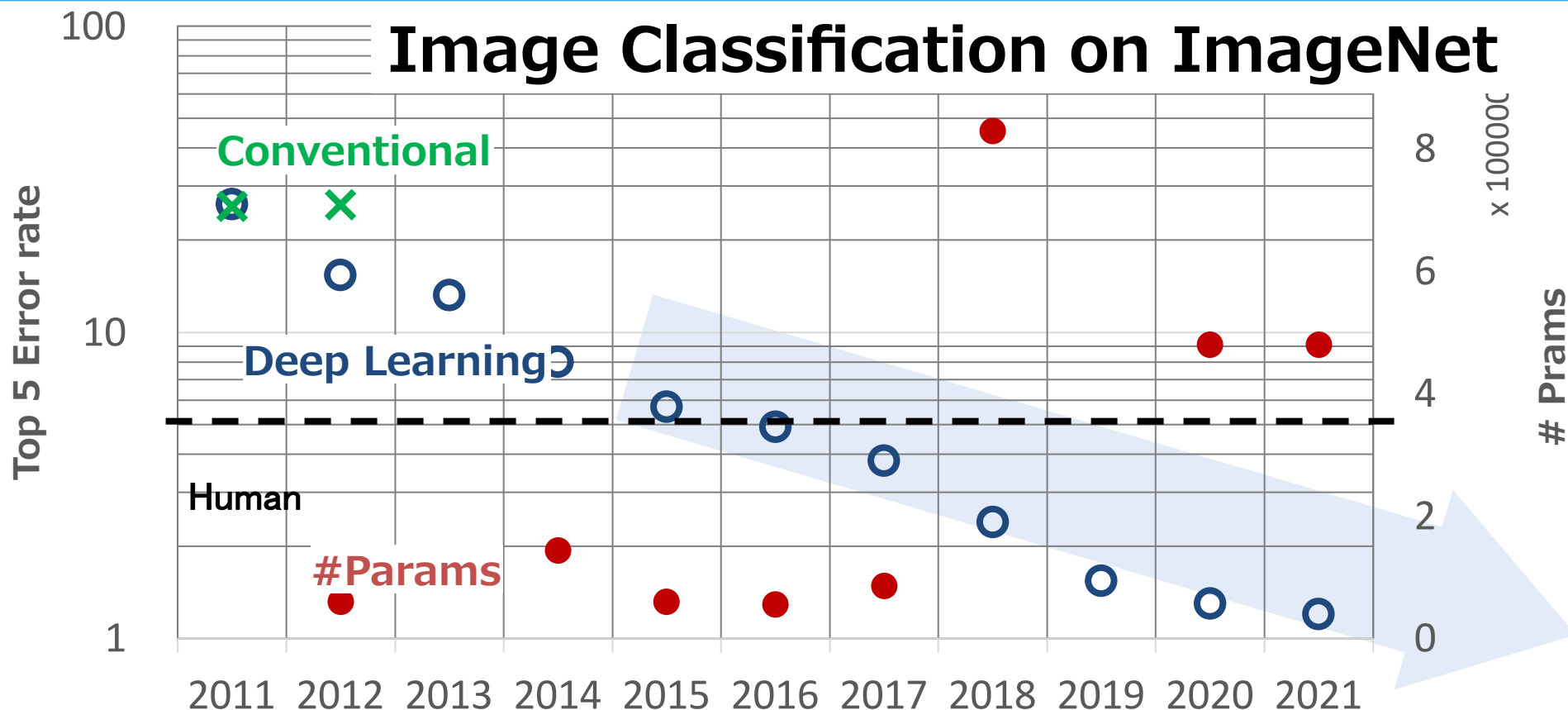
学習の結果、人間の脳に似た機能が獲得されることが実験的に確認されている。



イラスト：©いらすとや

圧倒的な認識性能を示す Deep Learning

画像認識の分野では、Deep Learningにより認識精度がはるかに向上した。
Deep Learningの精度は人の認識精度をも上回っている。



"Image Classification on ImageNet"

©paperswithcode.com (licenced under CC-BY-SA 4.0)

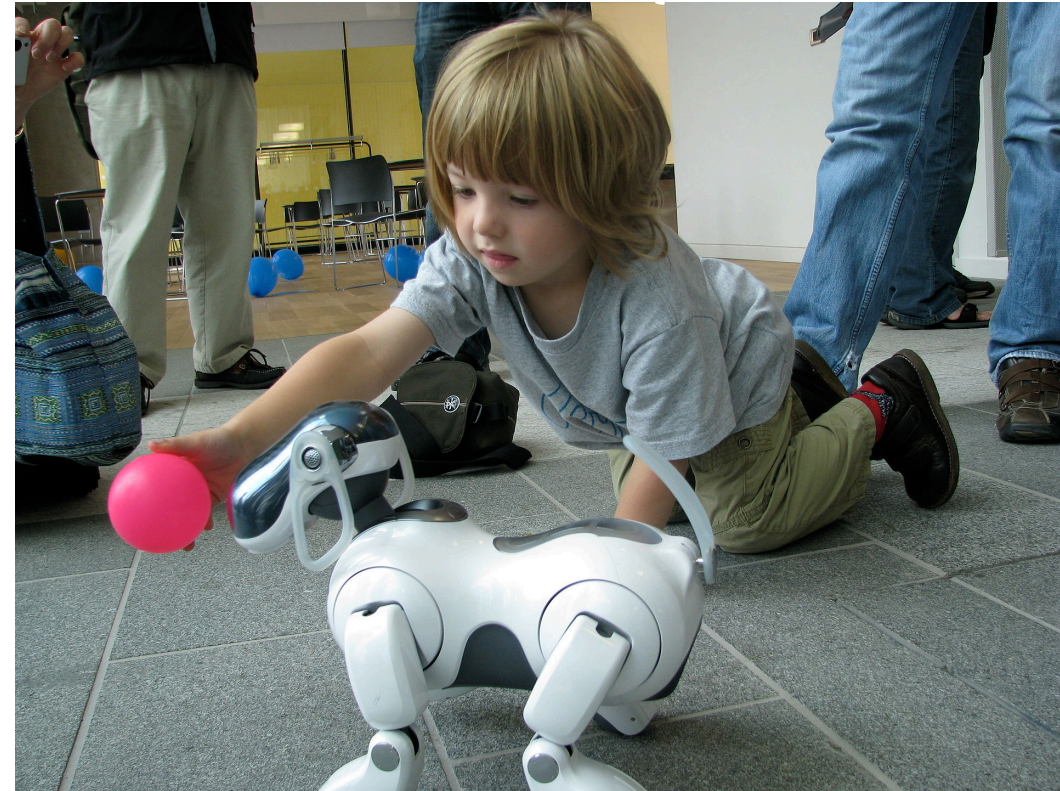
aibo (アイボ)

□実運用での課題 (想定)

- 連続稼働時間 (2時間) ・充電時間 (3時間)
- 予め指定したものしか認識しない
- 行動はある程度決まっている

□他の課題への転用

- 留守宅の見守り→家庭内地図を使ってパトロール (家族の見守り)
- 病気の人や老人の見守り、検診
- 育児補助



"AIBO ERS-7" ©Stuart Caie (Licensed under CC-BY 4.0)