

人工知能は人間を超えるか

- ディープラーニングの先にあるもの

東京大学 松尾 豊

東京大学 松尾研究室について



松尾 豊

- 1997年 東京大学工学部電子情報工学科卒業
- 2002年 同大学院博士課程修了. 博士(工学)。産業技術総合研究所 研究員
- 2005年 スタンフォード大学客員研究員
- 2007年～ 東京大学大学院工学系研究科 技術経営戦略学専攻 准教授
- 2014年～ 東京大学 グローバル消費インテリジェンス寄付講座 共同代表・特任准教授
- 2015年～ 産総研AIセンター 企画チーム長

- ◆人工知能、ディープラーニング、Webマイニングを専門とする。
- ◆論文数と被引用数に基づき科学者の科学的貢献度を示すh-Index=31(ウェブ・人工知能分野最高水準)であり、2013年より国際WWW会議Web Mining部門のチェアを務める。
- ◆世界人工知能国際会議 プログラム委員。2012年より、人工知能学会 理事・編集委員長(それまでの慣例を大幅に更新し最年少で編集委員長就任)、2014年から倫理委員長。
- ◆人工知能学会論文賞(2002年)、情報処理学会長尾真記念特別賞(2007年)、ドコモモバイルサイエンス賞(2013年)、文部科学省 科学技術への顕著な貢献2015、大川出版賞(2015年)、ビジネス本大賞審査員賞(2016年)等受賞。
- ◆経済産業省 産業構造審議会 新産業構造部会 委員、IoT推進コンソーシアム 運営委員、厚生労働省「働き方の未来 2035」懇談会メンバー、内閣府「人工知能と人間社会に関する懇談会」構成員、金融庁「フィンテック・ベンチャーに関する有識者会議」委員、総務省「ICTインテリジェント化影響評価検討会議」委員等。
- ◆近著に「人工知能は人間を超えるか? --ディープラーニングの先にあるもの」(角川 2015)。

<研究室の実績>

- ◆博士学生17人、修士・学部生10人が所属し、人工知能の基礎研究、ソーシャルメディアの分析、データ分析及びその実社会へのアプリケーションを多方面にわたって行っている。
- ◆これまでに、トヨタ、リクルート、マイクロソフト、CCC、経営共創基盤、ミクシィなどさまざまな企業と共同研究の実績がある。官公庁からも、経産省(アジアトレンドマップ等)、文科省(ビッグデータ活用)など相談多数。
- ◆卒業生の主な進路は、Google、DeNA、楽天、サイバーエージェント、光栄、ゴールドマンサックス、BCG、三井物産、電通など。起業した学生も多数。GunosyやREADYFORなどのサービスを構築、運用している。

人工知能をめぐる動向

- 第1次AIブーム(1956～1960年代):探索・推論の時代
 - ダートマスワークショップ(1956)
 - 人工知能(Artificial Intelligence)という言葉が決まる
 - 世界最初のコンピュータENIAC(1946)のわずか10年後
 - 数学の定理証明、チェスを指す人工知能等
- ...冬の時代
- 第2次AIブーム(1980年代):知識の時代
 - エキスパートシステム
 - 医療診断、有機化合物の特定、...
 - 第5世代コンピュータプロジェクト:通産省が570億円
- ...冬の時代
- 第3次AIブーム(2013年～):機械学習・ディープラーニングの時代
 - ウェブとビッグデータの発展
 - 計算機の能力の向上

考えるのが早い人工知能

ものしりな人工知能

データから学習する人工知能

将棋電王戦



IBM ワトソン



<http://venturebeat.com/2011/02/15/ibm-watson-jeopardy-2/>, <http://weekly.ascii.jp/elem/000/000/207/207410/>

ディープラーニング革命

- ILSVRCでの圧勝 (2012)
- Googleの猫認識 (2012)
- ディープマインドの買収 (2013)
- FB/Baiduの研究所 (2013)
- アルファ碁 (2016)

機械学習

ウェブ・ビッグデータ

車・ロボット
への活用

自動運転
Pepper

統計的自然言語処理
(機械翻訳など)

検索エンジンへの活用

MYCIN (医療診断)
DENDRAL

エキスパート
システム

オントロジー

タスクオントロジー

ワトソン(2011)

LOD (Linked Open Data)

Eliza

対話システムの研究

Caloプロジェクト

Siri(2012)

bot

探索
迷路・パズル

プランニング
STRIPS

チェス (1997)
Deep Blue

将棋 (2012-)
電王戦

囲碁

1956

1970

1980

1995

2010

2015

第一次AIブーム
(推論・探索)

第二次AIブーム
(知識表現)

第三次AIブーム
(機械学習・ディープラーニング)

ディープラーニング革命

認識

「画像認識」ができる
(コンピュータができて以来、初めて！)

運動の習熟

ロボット・機械に
熟練した動きができる

言葉の意味理解

文の「意味」が分かる
(文と映像の相互変換ができる)

認識の難しさ



ネコ



イヌ



オオカミ

- これをコンピュータで見分けたい
 - 目が丸い→ネコ
 - 目が細長い、耳が垂れている → イヌ
 - 目が細長い、耳がとがっている → オオカミ



イヌ



結局、「耳が垂れている」「目が細長い」などの「特徴量」を人間が考えている限り無理。どんなに頑張っても、必ず例外がある。人間はなぜかうまくできる。

これまでの人工知能の壁≡特徴抽出の壁

- 難しい問題1:機械学習における特徴量の設計 (Feature engineering)
 - 機械学習において、変数(特徴量)の設計が難しかった。
 - 人間が対象をよく観察して設計するしかなかった。
- 難しい問題2:フレーム問題
 - 人間が知識を記述することで、人工知能を動作させる。
 - そのときに、いくら知識を書いても、うまく例外に対応できない。
- 難しい問題3:シンボルグラウンディング問題
 - シマウマがシマのある馬だと、計算機が理解することができない。
 - シンボル(記号)がそれが指すものと接続(グラウンド)しておらず、シンボルの操作ができない。



結局のところ、いままでの人工知能は、

人間が現実世界の対象物を観察し、「どこに注目」するかを見ぬいて(特徴量を取り出して)、モデルの構築を行っていた。

その後の処理は自動で行うことができたが、モデル化の部分に人間が大きく介在していた。それが、唯一にして最大の問題であった。

認識: ディープラーニングの実績(2012)

- ILSVRC2012: Large Scale Visual Recognition Challenge 2012



ディープ
ラーニング

長年の
特徴量設計
の工夫

Team name	Error	Description
SuperVision	15.315%	Using extra training data from ImageNet Fall 2011 release
SuperVision	16.422%	Using only supplied training data
ISI	26.602%	Weighted sum of scores from classifiers using each FC
ISI	26.646%	Naïve sum of scores from classifiers using each FV
ISI	26.952%	Naïve sum of scores from each classifier with SIFT+FV, LBP+FV, GIST+FV and CSIFT+FV, respectively
OXFORD_VGG	26.979%	Mixed selection from High-Level SVM scores and Baseline Scores, decision is performed by looking at the validation performance.
...

「ケタ」が違う

認識: 2012年以降のエラー率の変化

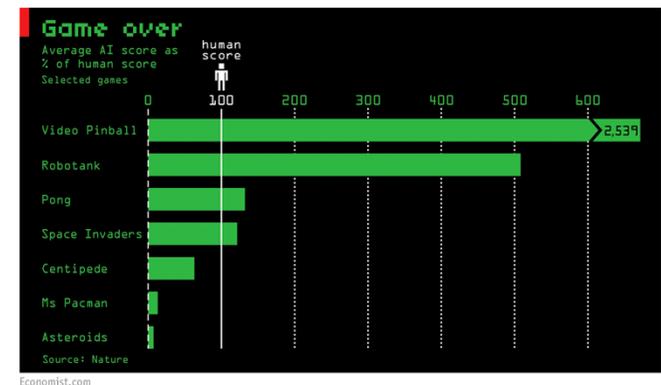
		Error
Before ディープ ラーニング	Imagenet 2011 winner (not CNN)	25.7%
	Imagenet 2012 winner	16.4% (Krizhevsky et al.)
	Imagenet 2013 winner	11.7% (Zeiler/Clarifai)
	Imagenet 2014 winner	6.7% (GoogLeNet)
After ディープ ラーニング	Baidu Arxiv paper: 2015/1/3	6.0%
	Human: Andrej Karpathy	5.1%
	Microsoft Research Arxiv paper: 2015/2/6	4.9%
	Google Arxiv paper: 2015/3/2	4.8%
	Microsoft Research CVPR paper: 2015/12/10	3.6%
	Latest	3.1%

2015年2月には人間の精度を超えた

画像認識で人間の精度を超えることは数十年間、実現されていなかった⁹

運動の習熟: ディープラーニング + 強化学習 (2013-)

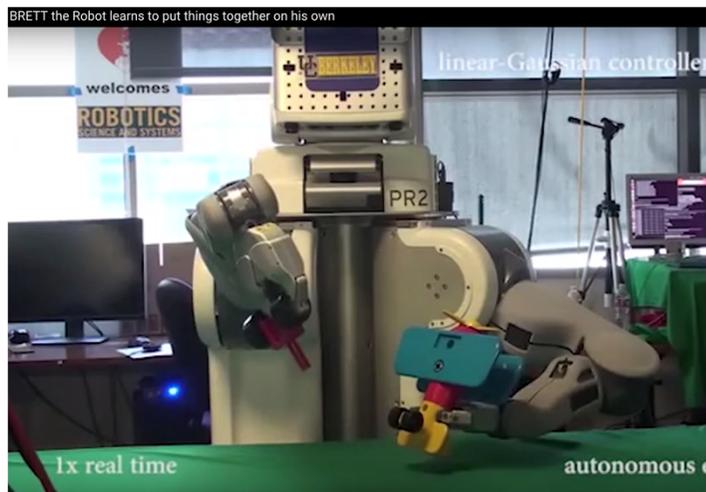
- 強化学習とは、行動を学習する仕組み。
 - 「報酬」が得られると、事前の行動を強化する。
 - 「状態」「行動」→「望ましき(報酬ありなし)」
 - 古くからある技術だが、これまでは、「状態」を人間が定義してきた。
- 運動の習熟が可能に
 - 状態の認識に、ディープラーニングを使う。
 - DeepMindの研究者(D. Hassabisら)。その後、Googleが買収。
- 試行錯誤することによって、運動が習熟する
 - 最初は下手。繰り返すうちに、うまくなる。
 - 最終的には、ブロック崩しでの通路を作ったり、インベーダーゲームでの「名古屋撃ち」も。
 - 「全く同じプログラム」で、異なるゲームを学習。半数のゲームで人間のハイスコアを上回る



運動の習熟：ディープラーニング＋強化学習が実世界へ(2015-)

- 実世界への適用
 - 2015年5月 試行錯誤で部品の取付を習熟するロボットの開発(UC Berkeley)
 - 2015年5月 試行錯誤で運転を習熟するミニカーの開発(PFN社, 日本)
 - 2015年12月 試行錯誤でピッキングが上達するロボットの開発(PFN・ファナック, 日本)
 - その他、メリーランド大、EUのプロジェクト等も進展
- 考えてみれば当たり前
 - 犬や猫でもできる。高次な言語能力は必要ない。認識が問題だった。
 - 歴史的には、多数の人工知能研究者がこのことを主張してきた。

試行錯誤で作業学ぶロボット(UC Berkeley)



試行錯誤で運転を学習するミニカー(PFI社、日本)



ディープラーニングの人工知能における意味

- モラベックのパラドックス:「子供のできることほど難しい。」
 - 高度な推論よりも、認識や運動スキルの方が難しい。
 - それがここ3年くらいのあいだに一気にできるようになった
- 現在のコンピュータのパワーでようやく可能に
 - GPUを数十台並列に並べて、数日～数ヶ月計算させてようやく精度が上がる
- アイディアは昔からあった。もともとは日本発
 - 1980年当時、NHK放送技術研究所にいた福島邦彦先生によるネオコグニトロン
 - その後も多くの研究者が試みている
- 初期仮説への回帰
 - 初期仮説「なぜ知能をコンピュータで実現することはできないのか？」
 - できると思っていた→できない理由があった→それが解消された→だとしたら、もう一度できるという仮説を取るべきでは。
 - 産業として非常に大きい可能性を秘めている。

ディープラーニングの今後の発展

① 画像

画像から、特徴量を抽出する

画像認識の精度向上

② マルチモーダル

映像、センサーなどのマルチモーダルなデータから特徴量を抽出し、モデル化する

動画の認識精度の向上、行動予測、異常検知

③ ロボティクス(行動)

自分の行動と観測のデータをセットにして、特徴量を抽出する。記号を操作し、行動計画を作る。

プランニング、推論

④ インタラクション

外界と試行錯誤することで、外界の特徴量を引き出す

オントロジー、高度な状況の認識

⑤ 言葉とのひもづけ(シンボルグラウンディング)

高次特徴量を、言語とひもづける

言語理解、自動翻訳

⑥ 言語からの知識獲得

グラウンディングされた言語データの大量の入力により、さらなる抽象化を行う

知識獲得のボトルネックの解決

認識



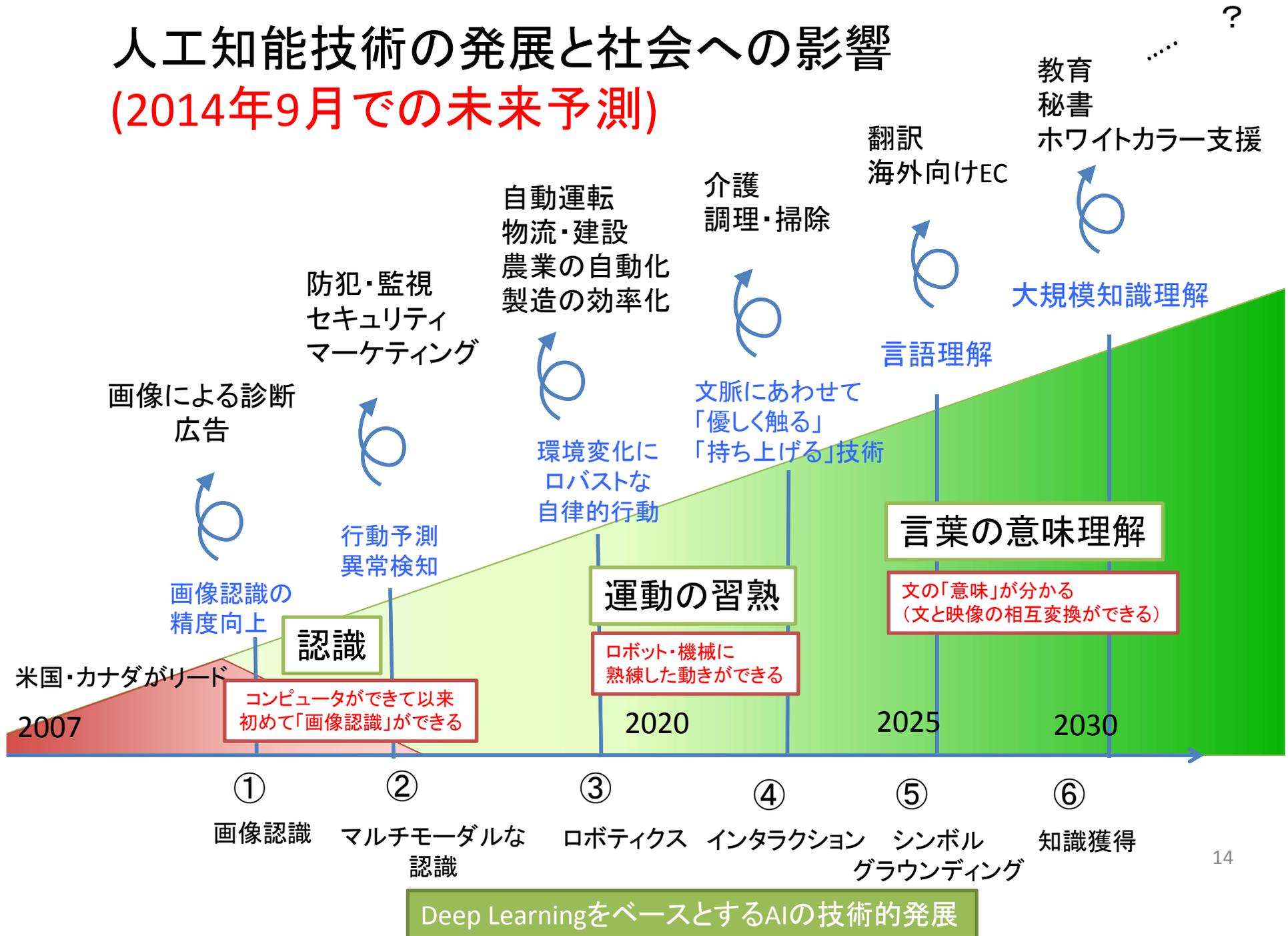
運動



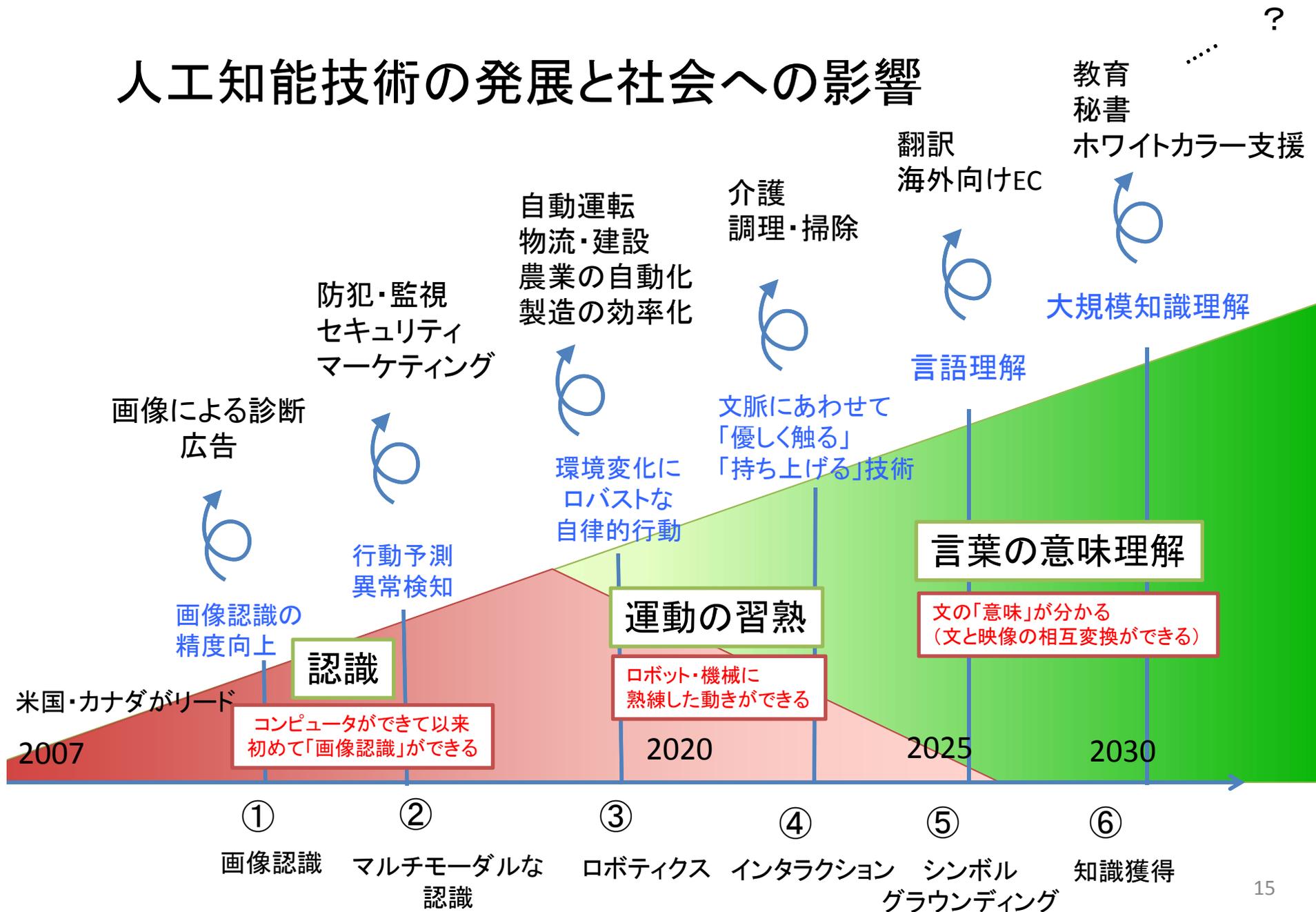
言葉

ディープラーニングがすごいというより
その先に広がる世界がすごい

人工知能技術の発展と社会への影響 (2014年9月での未来予測)



人工知能技術の発展と社会への影響



Deep LearningをベースとするAIの技術的發展

言葉の意味理解 : Automated Image Captioning (2014-)



"man in black shirt is playing guitar."



"construction worker in orange safety vest is working on road."



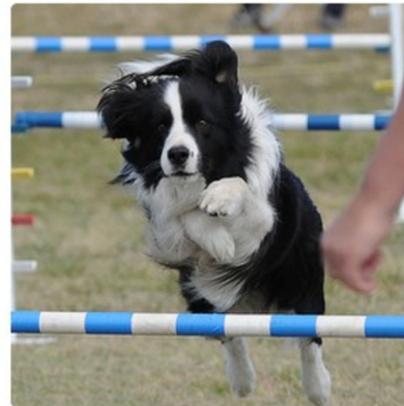
"two young girls are playing with lego toy."



"boy is doing backflip on wakeboard."



"girl in pink dress is jumping in air."



"black and white dog jumps over bar."



"young girl in pink shirt is swinging on swing."

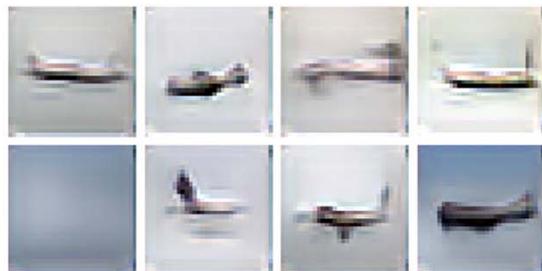


"man in blue wetsuit is surfing on wave."

言葉の意味理解: Generating Images (2015.12-)



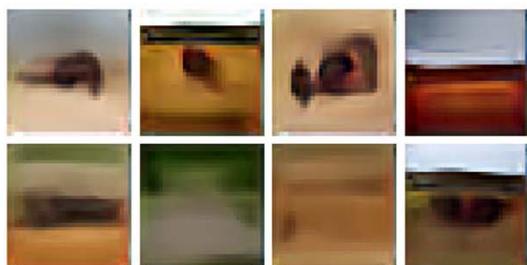
A very large commercial plane flying in blue skies.



A very large commercial plane flying in rainy skies.



A stop sign flying in blue skies.



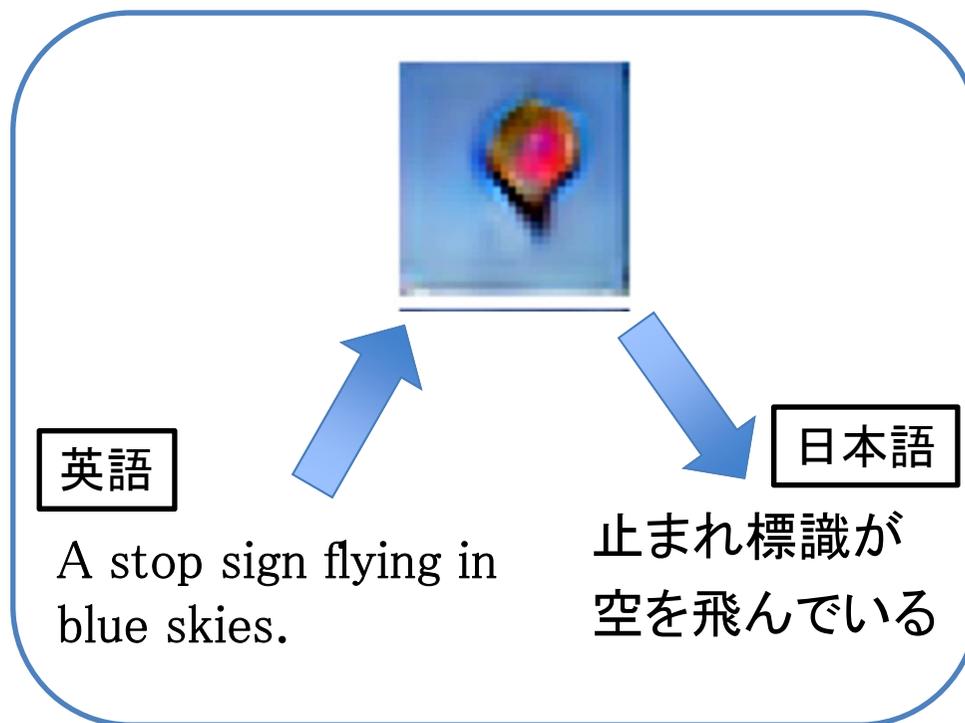
A herd of elephants walking across a dry grass field.



A herd of elephants walking across a green grass field.

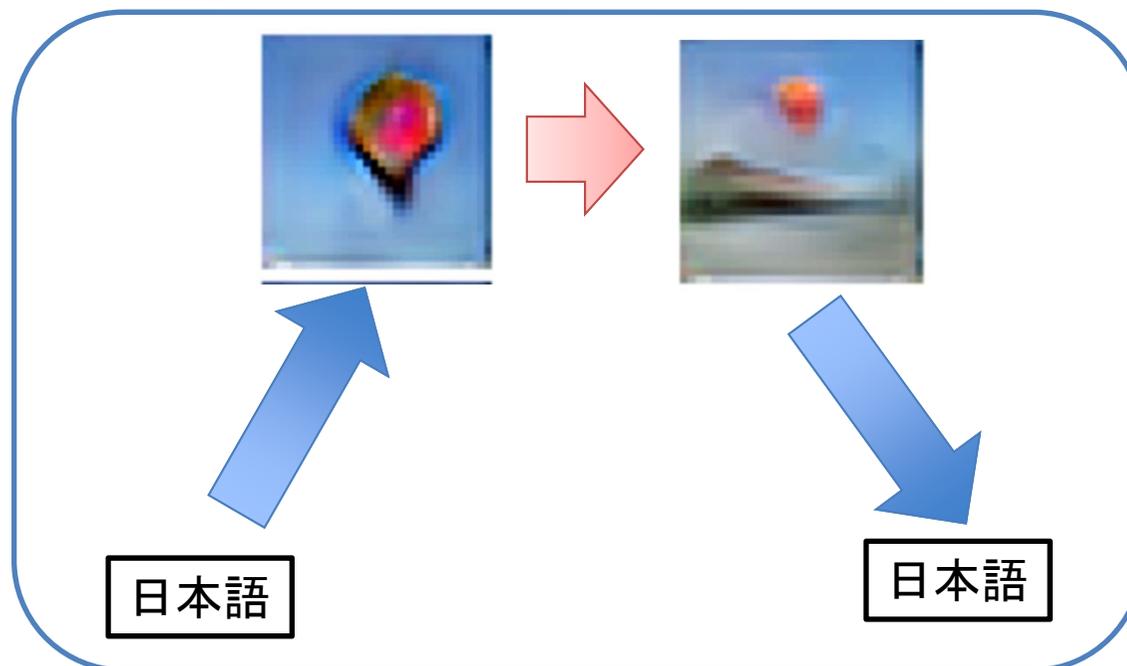
- 画像による翻訳(意識)

- 日本語→画像→英語
- 課題
 - 画像から映像、体験
 - 抽象概念の扱い
 - 人間固有のセンサー系
- いずれにしても視覚的な機構がベースにあるのは間違いない



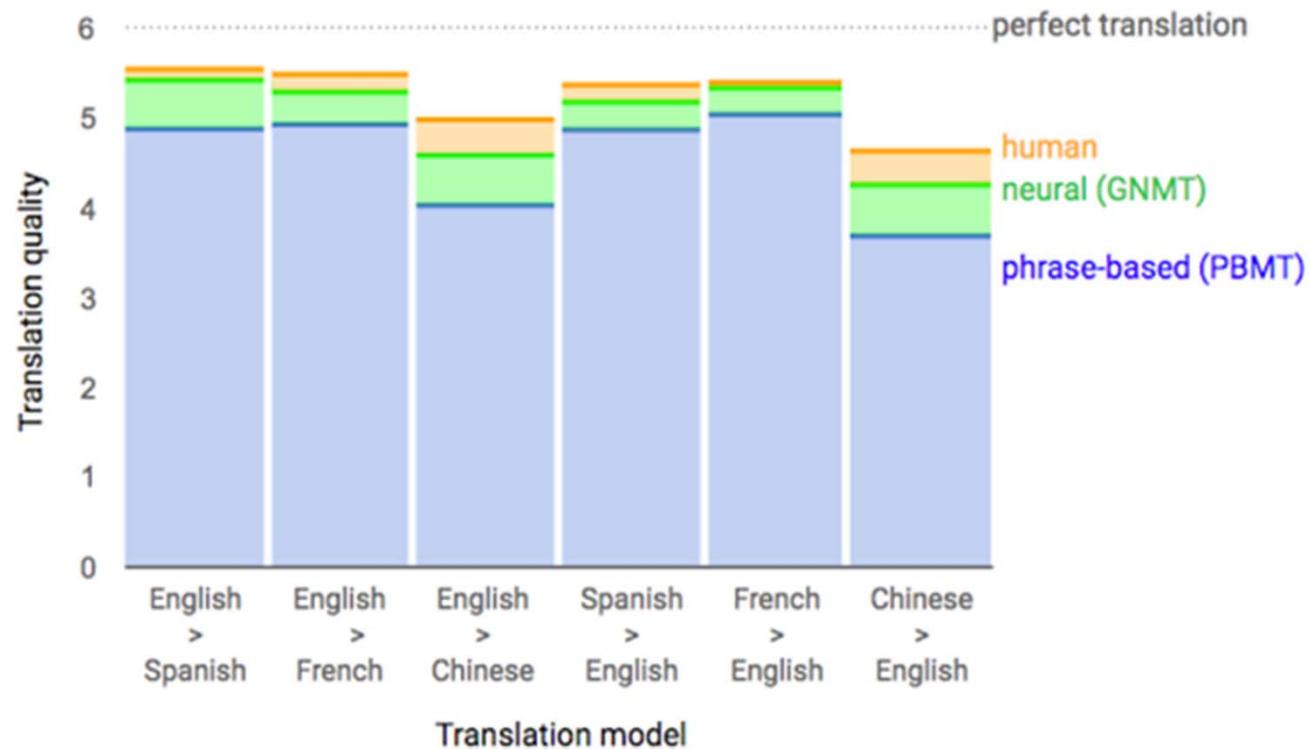
- 映像による推論

- 言語→映像→(シーン予測)→次の映像→言語
- 「風船が飛んでいる」→「山まで飛んで行くのかな」



Google翻訳がDL方式に(2016.9)

- Google Neural Machine Translation (GNMT)
- 全部入り:8層のbidirectional RNN (LSTM), Residual connection, Attention
- さらなる工夫:TPUの利用, Wordpiece
- 60%エラーを減らした。
- “two to three decimal orders of magnitudes bigger than the WMT corpora for a given language pair”



Deep reinforcement learning (deep RL) has been successful in learning sophisticated behaviors automatically; however, the learning process requires a huge number of trials. In contrast, animals can learn new tasks in just a few trials, benefiting from their prior knowledge about the world. This paper seeks to bridge this gap. Rather than designing a “fast” reinforcement learning algorithm, we propose to represent it as a recurrent neural network (RNN) and learn it from data. In our proposed method, RL2, the algorithm is encoded in the weights of the RNN, which are learned slowly through a general-purpose (“slow”) RL algorithm. The RNN receives all information a typical RL algorithm would receive, including observations, actions, rewards, and termination flags; and it retains its state across episodes in a given Markov Decision Process (MDP). The activations of the RNN store the state of the “fast” RL algorithm on the current (previously unseen) MDP. We evaluate RL2 experimentally on both small-scale and large-scale problems. On the small-scale side, we train it to solve randomly generated multi-armed bandit problems and finite MDPs. After RL2 is trained, its performance on new MDPs is close to human-designed algorithms with optimality guarantees. On the large-scale side, we test RL2 on a vision-based navigation task and show that it scales up to high-dimensional problems.

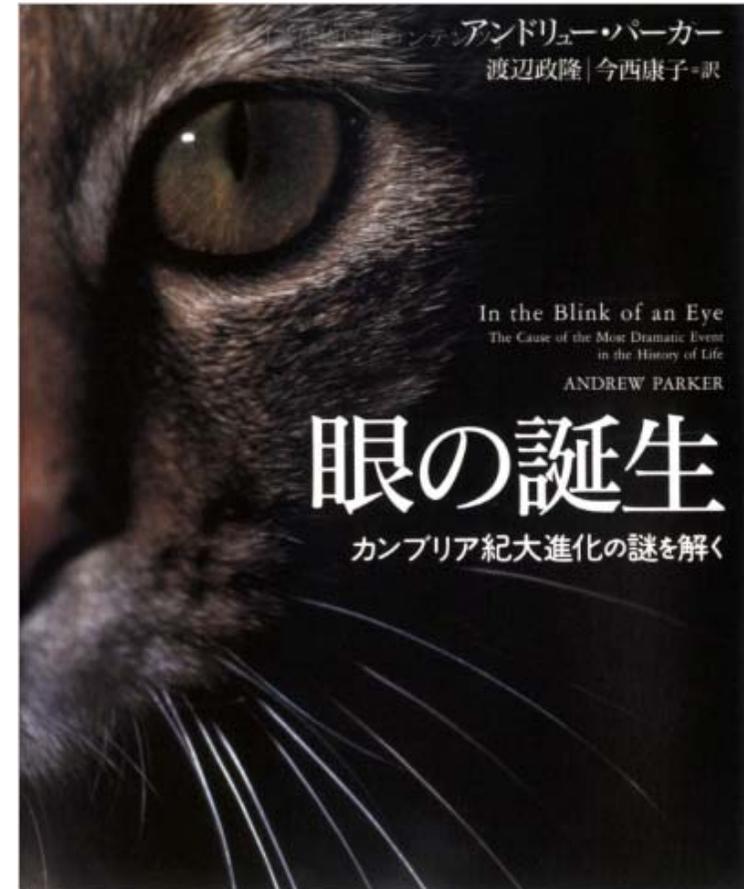
深い強化学習(深いRL)は、洗練された行動を自動的に学習するのに成功しています。しかし、学習プロセスには膨大な試行が必要です。これとは対照的に、動物は世界についての以前の知識から恩恵を受け、わずかな試行で新しい仕事を習得することができます。このペーパーは、このギャップを埋めようとしています。「高速」強化学習アルゴリズムを設計するのではなく、それをリカレントニューラルネットワーク(RNN)として表現し、それをデータから学習することを提案します。提案された方法RL2では、アルゴリズムはRNNの重みに符号化され、RNNは汎用(「遅い」)RLアルゴリズムによってゆっくり学習される。RNNは、観察、行動、報酬、および終了フラグを含む典型的なRLアルゴリズムが受け取るすべての情報を受信する。それは所与のマルコフ決定プロセス(MDP)においてエピソード全体にわたってその状態を保持する。RNNの活性化は、現在の(以前は見えなかった)MDP上の“高速”RLアルゴリズムの状態を記憶する。我々は、RL2を小規模問題と大規模問題の両方について実験的に評価する。小規模な面では、無作為に生成された複数武装の禁止問題と有限のMDPを解決するように訓練します。RL2が訓練された後、新しいMDPでの性能は、最適性が保証された人間が設計したアルゴリズムに近くなります。大規模な側面では、RL2をビジョンベースのナビゲーションタスクでテストし、高次元の問題までスケールアップすることを示します。

眼の誕生

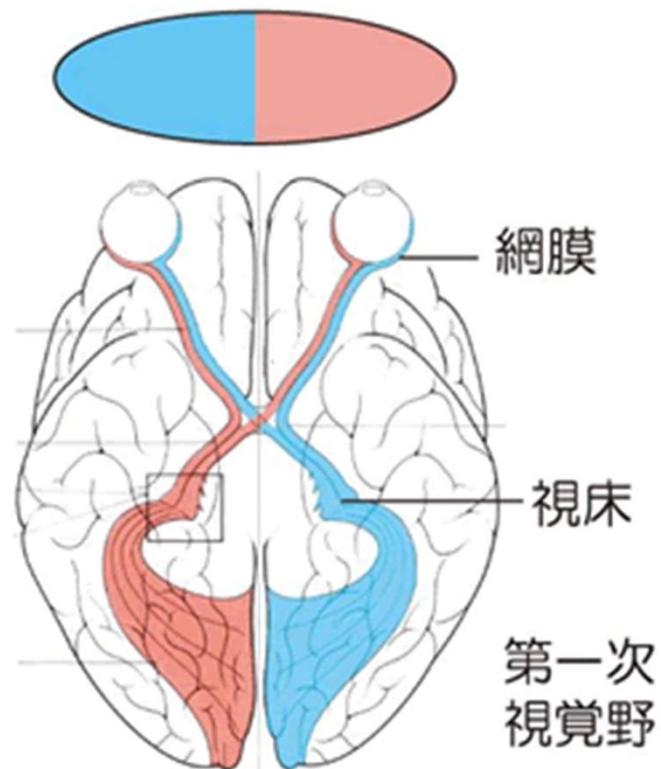
- カンブリア爆発
 - 5億4200万年前から5億3000万年前の間に突如として今日見られる動物の「門」が出そろった現象
 - 古生物学者アンドリュー・パーカーは、「眼の誕生」がその原因だったの説を提唱
- ディープラーニングにより、見えるようになる
 - さらに、次に何が起こるかを予想して動けるようになる。
- 「眼をもった機械」が誕生する。
 - 機械・ロボットの世界でのカンブリア爆発が起こる。
 - これを日本企業が取れるか？



三葉虫：史上初めて眼をもった生物



眼が見える仕組み



←イメージセンサ

←ディープラーニング
(CNN: 畳み込みニューラルネットワーク)

既存産業の発展

農業

収穫判定

トラクター、コンバインの
適用範囲拡大、効率向上
選別調製等の自動化

自動での収穫
自動での耕うん

建設

測量

掘削、基礎工事、
外装内装作業等の
効率向上

多くの作業の
自動化・効率化

食品
加工

振り分け
確認

カット、皮むき等
の自動化
食洗機に入れる

多くの加工工程の
自動化

組み立て
加工

目視確認の
自動化

動作効率の向上

段取りの自動化
セル生産の自動化

⋮

A: 画像認識

B: 運動の習熟

C: 計画立案を伴う運動

眼をもった機械・ロボットの典型例 単独の製品から入る

- 農業: トマト収穫ロボット
 - トマトは市場規模も大きく、収穫の工数も大きい。
 - 現状の技術で、トマトの認識ができる。上手にもぎ取ることも可能。
 - 先進的な農場から試しに入れる。
- 建設: 自動溶接機械
 - 建設の工程(例えば溶接)を自動化する
 - 現状の技術で、接合面の状態等の認識ができる。上手に溶接することも可能。機械を当てれば熟練した人でなくとも熟練の人のような溶接ができる。
 - 一部の建設現場で試しに入れる。
- 食品加工: 食洗機にお皿を入れるロボット
 - 食品加工に関わる仕事、まずは食洗機にお皿を入れることを自動化する
 - 現状の技術で、お皿の位置、把持位置の認識ができる。まずは、食器が下げられたところから、食洗機に入れるところを自動化する。(混雑時に重要)
 - ファミリーレストラン等の一部の店舗で試しに入れる。

製品を一刻も早く市場に投入する

日本なりのプラットフォーム戦略

- DLの技術はコモディティ化する。
 - 競争力をもつのは、データとハードウェア。
 - 早くDLの技術を取り入れてしまえばよい。
 - DLの技術とハードウェアのすりあわせになった瞬間、日本企業が再度、力を取り戻せる。
- 欧米のスタートアップ(とDL研究者)は、意外なほどハードウェアに対する抵抗感がある
 - そもそも、産業用ロボットの導入台数は日本が(ほぼ)トップ
 - また、ロボットに対する社会的抵抗感もある。
 - 米国は雇用を守らないといけない。日本は人手が足りない。
- ものが関連しないプラットフォームは無理
 - 英語圏でやったほうが絶対に強い。
 - 広告費規模でも10倍、ECの規模でも3倍以上
- 日本には、検索エンジンもECもSNSもあった。
 - GoogleやAmazon, Facebookの位置の企業を出せなくはなかった。
 - しかし、結果はそうになってない。原因は明確で、英語圏でなかったから。

ものづくりを起点に眼をもった機械を作り、
プラットフォーム化するのは日本ならではの戦略

機械・ロボットのカンブリア爆発

- 介護施設や病院等での見守り・介護ロボット
- 医療（X線、CT、皮膚、心電図、手術ロボット）
- 警備、防犯技術
- 顔による認証・ログイン・広告技術、表情読み取り技術（サービス業全般に重要）
- 国家の安全保障、入国管理、警察業務、輸出入管理業務における活用
- 防災系（河川、火山、土砂崩れを見張る）
- 重機系（掘削、揚重）、建設現場系（セメント固め、溶接、運搬、取り付け）
- 農業系（収穫、選果、防除、摘花・摘果）
- 自動操縦系（ドローン、小型運搬車、農機、建機）
- 自動運転系、物流
- 産業用ロボット系（特に組み立て加工等）
- 調理系（牛丼、炊飯、ファミリーレストラン、外食全般）
- ペットロボット系
- 片付けロボット（家庭、オフィス、商業施設）
- 新薬発見や新素材の開発（遺伝子の認識・分析、実験ロボット）
- 廃炉系（深海や鉱山、宇宙も含めた極限環境）

農業・建設・食品加工だけでなく、医療や介護、製造、廃炉なども。

DLに関わる海外企業

- ベンチャー

- Deep Mind (英): DLの技術力をもった企業。DQNによるゲーム、アルファ碁、医療など。2011創業。Googleが2014に£ 400Mで買収。
- Enlitic: 医療画像(X線)におけるDL活用。2014創業、15M調達。
- Nervana Systems: 医療、農業、金融、自動車、エネルギー等における画像処理。24M調達後、インテルが2016買収。
- Emotient: 顔の表情を認識する会社。2012創業、6M調達後、Appleが2016買収。
- Affectiva: 映画やTV番組のどこで表情が変わったのかを読み取る。2009創業、34M調達。
- Perceptio: DLによる写真分類アプリ開発。創業、調達額不明。Appleが2015買収。
- VocalIQ (英): DLによる音声認識。1M調達後、Appleが2015買収。
- Atomwise: ドラッグディスカバリーへのDL活用。新薬の候補物質を見つける。YC卒業生。2012創業。6M調達。
- Descartes Labs: DLによる衛星画像の分析。農業への適用。2014創業、8M調達。
- Canary: DLによるホームセキュリティ。2012創業。41M調達。
- Netatmo: 家電。DLによる監視カメラも。2011創業。38M調達。
- Pilot AI Labs: DLの画像認識を使ったドローンの自動操縦。まだ小さいが、動画が面白い。
- MetaMind: 画像認識一般。2014創業、8M調達。
- SkyMind: JavaベースのDL提供。2014創業。3M調達。
- AlchemyAPI: DLによる言語処理と画像認識。クラウドで提供。2005創業、2M調達。IBMが2015に買収。
- ZenRobotics (フィンランド): ごみの選別ロボット。2007創業、17M調達。

- 製造業

- GE: DLによる医療画像の診断。
- Dyson: 掃除機に眼をつけたものを開発。インペリアル・カレッジにRoboticsラボ設立。
- Kuka (独): 産業用ロボットへのDL適用
- Mobileye (蘭): 車用の画像認識を提供。1999イスラエルで創業。2014年上場。時価総額10B。
- LG (韓): インチョン空港で、DLを使ったロボットでの案内の実験
- Samsung (韓): DLを使った胸部エコー検診の医療機器

急がないといけない

どこから始めるか

- 始めること自体はそれほど難しくない
 - 数学の知識: 線形代数や最適化
 - プログラムの知識: python、GPU
- ライブラリが揃っている
 - Tensorflow: Google、python
 - Caffe: UCバークレー、C++ベース
 - ほかにも、Keras, chainer, torch7, ...
- 教科書
 - 入り口: 「人工知能は人間を超えるか」(拙著)
 - 読み物: 「人間さまお断りー人工知能時代の富と仕事の手引き」(Jerry Kaplan, 近刊)
 - 教科書: 「深層学習」(岡谷貴之)
 - 教科書: “Deep Learning” (Y. Bengioら、MIT pressから来年出版。翻訳を出版予定)
 - www.deeplearningbook.org にPDFがあり、読めます。
 - あとは論文を読んでください。
- 3つの主要な国際会議
 - ICML, NIPS, ICLR (それぞれ年1回)
 - ウェブで全ての情報が見れます。

理系の人半年もやればそこそこできるようになる

ポイント

- 製品を作って、その利益を再投資に回すサイクルをいかに作れるか。
 - 最終的な競争力は技術にはない。**データとハードウェア**。
 - インターネットの世界で起こったことと同じ。
 - 個々の製品の性能向上→サービス化・プラットフォーム化
- 最も強いインセンティブをもつ企業が、最も効率的に再投資のループを作り勝つ
 - 例えば、生産の効率化ができるとしても、そこに対してもっと強いインセンティブを持つ企業が他にいないのか？
 - 顔認識ができるとして、それに対して最も強いインセンティブをもつ企業はどこなのか？
 - つまり、自社事業と「**最も直接的**」に**関係する技術革新**を考えないといけない。
- 実はこれは社内文化との戦い
 - 情報系はこれまでも「本流」から外れ、弱かった
 - そのなかでも「機械学習」の技術をもつ部署をいかに中心にもってこれるか。長年の伝統だった「本流」をいかに破壊し再構成するか。
 - それを**トップダウン**で**意思決定**できるか。

自社が最も強いインセンティブを感じる製品・サービスを描き、それをトップダウンに実行できるか

必要になる「学習工場」

- 学習工場のイメージ
 - 機械学習を使える高度な人材
 - 高性能な計算機
 - データを準備する環境
- 学習工場で出荷されるもの
 - 学習済みの「モデル」が作られる。頭脳の部分。
 - これが最終製品に載せて出荷される。
 - それぞれの事業者が学習工場をもつ。
- 「工場」なので、数十億～数千億円(数兆円)の投資規模
 - R&Dではない。最終製品を作るものである。
 - 10倍生産したいのであれば10倍の規模に。
 - この規模でようやく諸外国と勝算ある状態で戦える
- これを新たな設備投資の形とすることができないか。

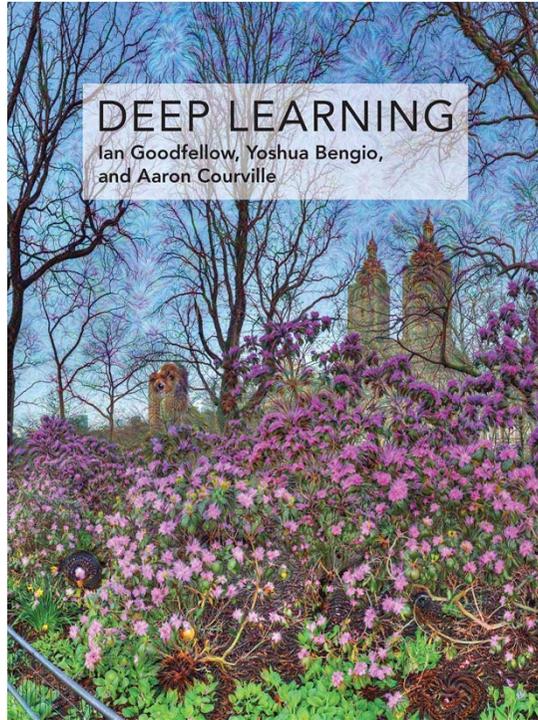
それぞれで必要なもの

- 学習用のデータづくり
 - アノテーション付きのデータをたくさん作り出す事業者(サプライヤー)
 - そこには大量の雇用が発生
 - 2次サプライヤー
 - そのためのツールを開発する事業者
 - そのためにクラウドソーシングを活用する事業者
 - 模擬データを出すためのシミュレータを開発する事業者
 - 極限環境でのデータをとる事業者
- 計算機
 - DL用の計算機: 海外が強い(NVIDIA)
 - ライブラリ: 海外が強い。TensorflowやKerasなど。
 - アルゴリズム: 海外が強い。GoogleやFacebookなど。
 - ひとまずしっかりと輸入。あまり国産にこだわっても良くない。
- 学習職人(高度なスキル・知識をもった人材)
 - いま、これが圧倒的に足りない。
 - 大学で人材育成を急ぐ。企業内での人材研修。
 - 研修プログラムの提供、資格制度作りを急ぐべき。

大量の雇用、
たくさんの事業者を生む
産業としての広がりは大い

しっかり輸入
(いずれ独自の競争力)

ハイスペック人材



Ian Goodfellow, Yoshua Bengio,
and Aaron Corville, MIT press 2016.11

- 人工ニューラルネットワークの最初の実験が1950年代に行われたのに、なぜ最近になってようやく、深層学習が極めて重要な技術と認識されるようになったかは、不思議に思うかもしれない。深層学習は、1990年代から商業的な応用ではうまく使われていたが、最近まで、しばしば技術というよりはアートであり、専門家だけが使えるものと見なされた。深層学習のアルゴリズムでよい性能を得るには、必要なスキルがあることは真実である。幸いにも、必要なスキルの量は、訓練データの量が増えるにつれて減っている。今日、複雑なタスクで人間の性能に到達する学習アルゴリズムは、1980年代におもちゃの問題を解くのに苦労した学習アルゴリズムとほとんど同一である。これらのアルゴリズムで訓練するモデルは、とても深いアーキテクチャでの訓練を簡略化する変化をしてはいるが。最も重要な新しい進歩は、今日ではアルゴリズムが成功するのに必要とするだけのリソースを、アルゴリズムに提供することができることである。

- 2016年の時点で、大体の目安として、教師ありの深層学習のアルゴリズムは、一般的に、カテゴリごとに約5000のラベル付き事例で、許容できる性能を達成し、少なくとも1000万のラベル付きの事例を含むデータセットで訓練すれば、人間の性能と匹敵する、あるいは超える。これよりも小さいデータセットでうまくいくことは、重要な研究分野であり、教師なしあるいは半教師あり学習で、大量の教師なし事例をうまく活かす方法に特に焦点が当てられる。

日本の戦略

日本の社会課題に対して、DLとものづくりの掛けあわせによる「眼をもった機械」を開発し解決する

- 農業分野に「眼をもった機械」を適用することで
 - 休耕地が耕せる。除草・防除や収穫ができる。収量が増える。
- 介護分野に適用することで
 - 介助も楽に。移動したりトイレにいけるようになり、より自立した生活ができる。
- 廃炉作業に適用することで
 - 危険な状況で人が作業しなくてよくなる。工期を短縮できる。
- 河川や火山を見張ることで
 - 河川の氾濫や土砂崩れ、噴火などの危険な状態・予兆を早期に発見できる。
- こうした技術を使った製品を海外に展開していくことで
 - 新たな輸出産業に。GDPの増加につながる。

地方からグローバルへ

労働の必要な地方を舞台に技術を伸ばす。それをグローバルに展開

ディープラーニング×ものづくり： 「眼をもった機械」による日本の新たな産業競争力の実現へ

- 少子高齢化しており、労働力が不足している。
 - 頭脳労働は不足していない。「運動を伴う労働」のニーズが高い。
 - 農業従事者、建設・物流、介護、廃炉、熟練工の後継者、etc
- ディープラーニングが解決策になり得る。
 - 眼をもった機械：認識や運動の上達ができる機械・ロボット
 - ものづくりと相性がよく、日本の強みを活かせる。素材や駆動系も強い
 - 眼をもった機械というのは世界的にはもう常識。しかも自動運転だけではない。
- 新たな投資概念の必要性
 - 人への投資をいかに既存の枠組み・文化を踏まえてやるか
 - 企業からの投資を引き出し、そこに人材が流入する生態系を作るか
 - 例えば「学習工場」
- チャンスを捉えるには、正しく早く動いていくことが重要
 - ディープラーニング人材の育成
 - 事業・産業がどう変わるかを早期に検討
 - 社会全体で新しい未来像を描いていくこと

中小企業における人工知能の活用

- ITの活用(ワトソン型、あるいはFintech系のサービス)
 - 製造の管理:可視化。可視化すれば効率化する
 - 受発注や経理業務等の効率化
- ディープラーニングへの挑戦
 - 若いエンジニアが1人いて、時間を与えればできることもある。
 - cf) きゅうりの判定
 - ディープラーニングで起こる変化を読み解いて、事業拡大の狙いを定めることはあり得る
 - どんな部品や技術がキーパーツになるか分からない
 - 自社の事業に直結したところでその種を見つけて、あとはひたすら再投資を繰り返す