# 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

◆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、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。

### Googleの人工知能(アルファ碁)が囲碁でプロ棋士を破る (2016年3月)



文字サイズ 小 中 大



#### 囲碁人工知能、 「深層学習」で最強棋士に3連勝

2016年03月13日 07時13分

Tweet



人工知能アルファ基と対戦する李九段(右)=AP

【ソウル=川村律文】 米グーグル傘下 の英グーグル・ディープマインド社が開 発した囲碁の人工知能「アルファ碁」と 世界のトップ棋士である韓国棋院の 李世●九段(33)の第3局が12日、 ソウル市内のホテルで行われ、アルファ 碁が勝利した。(●は、石の下に乙)

5局まで行われる対戦で、アルファ基 は3連勝を果たし、最強棋士との対決を 制した。チェス、将棋に続き、頭脳ゲー

#### 人間対AI: 囲碁9段の解説者、解説できず視聴者に謝罪

2016/03/11 09:35

「あれ…? あれ…? 今まで見てきた手の中で一番衝撃的な手のような気がする。これは不思議だとしか言いようがないのでは?」 (チェ・ユジン囲碁アマチュア五段)

「不思議だというよりも、あり得ない手です。プロの感覚では考えも付かない手です。どういう意味で打ったんでしょうか?」(イ・ ヒソン九段)

10日、韓国トップの囲碁棋士、李世ドル(イ・セドル)九段と人工知能囲碁ソフト「アルファ碁」の第2局を中継していた韓国棋院運 営の「囲碁TV | 解説者たちは「解説 | ではなく「疑問 | を連発した。対局開始約45分後、「アルファ碁 | が打った手に戸惑いを隠せなか った。「アルファ碁」の予測できない変則的な手や、ミスだと思われた手を到底説明できないといった様子だった。李煕星(イ・ヒソ ン) 九段は「どうやってこの囲碁が…(『アルファ碁』が) 勝てる囲碁になるのだろうか」とため息をついた。

中盤を過ぎても次々と繰り出される「アルファ碁」の意外な手に、困惑を通り越えて恐怖すら感じているかのようだった。金成竜(キ ム・ソンリョン)九段は「『アルファ碁』はデータにない手を打っているようで怖い。『アルファ碁』の自己学習能力が進んでこういう 碁を打つなら、人間はあまりにも無力な気がする」と言った。

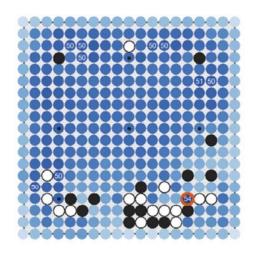
中盤まで李九段が有利だと見ていた解説者たちは、後に「アルファ碁」の方が有利になっていくと謝罪した。この日、SBSで解説をし ていた宋泰坤(ソン・テゴン)九段は「視聴者の皆さんに申し訳ない。李九段の敗着(敗因となった石の置き方)が分からない。人間の 目で見ると、『アルファ碁』はミスばかりしていた。今までの理論で解説すると、『アルファ碁』の囲碁は答えが出ない」と言った。対 局が終わった後、宋泰坤九段は本紙の電話取材に「対局を見ながら中継している間、狐につままれたような感じだった」と語った。

対局会場で解説者を務めた劉昌赫(ユ・チャンヒョク)九段は、対局を見守る間、何度も首をかしげて時折言葉を詰まらせた。劉昌赫 九段は「『アルファ碁』が中盤以降、終始優勢を保ったのは驚異的だ。『アルファ碁』は第1局とは違う姿を見せた。『アルファ碁』の棋 風は全くつかめない」と言った。

キム・スンジェ記者 朝鮮日報/朝鮮日報日本語版

### 関係者の衝撃

- 思考ゲームの歴史
  - 1980年: オセロプログラムMoorが、世界チャンピオン井上博との六番勝負で1勝を挙げた
  - 1997年: チェスプログラムDeep Blue(IBM)が、世界チャンピオン(ガルリ・カスパロフ)に勝つ
  - 2012年: 将棋プログラムボンクラーズが、故・米長永世棋聖に勝つ
    - その後、2015年まで、プロ棋士と対局し、9勝5敗1分
  - 残るは囲碁のみ
    - 「将棋の10年遅れ」でまだ10年は大丈夫。解の空間が広く、難しいため。
    - 2015年当時、まだアマチュアレベル。日本は研究でリード
- Google (DeepMind) のアルファ碁が、トッププロ棋士に勝利
  - 2015年1月にNature誌に掲載
  - Googleが買収したDeepMindが参入し、あっという間に(1年程度で)抜かれてしまった
  - 3月にトッププロの韓国のイ・セドルと対局。4勝1敗でアルファ碁の勝利。
  - 鍵となるのは、ディープラーニングを活用し、 局面の「認識」技術を使ったこと



### 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

## ディープラーニング革命



「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



言語の意味理解

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

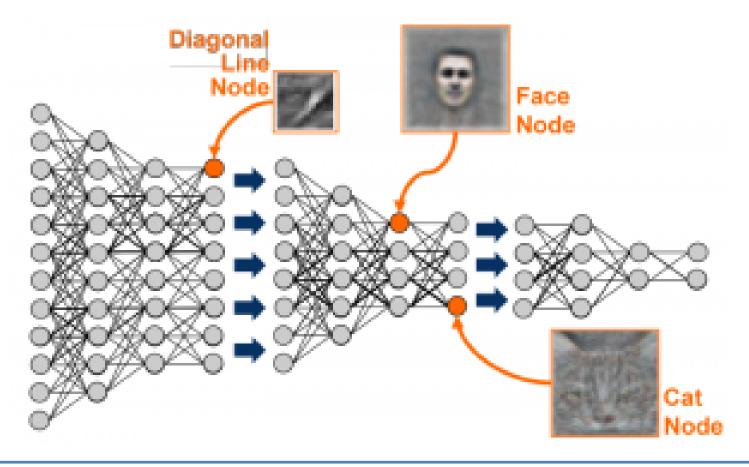
- 難しい問題1:機械学習における特徴量の設計(Feature engineering)
  - 機械学習において、変数(特徴量)の設計が難しかった。
  - 人間が対象をよく観察して設計するしかなかった。
- 難しい問題2:フレーム問題
  - 人間が知識を記述することで、人工知能を動作させる。
  - そのときに、いくら知識を書いても、うまく例外に対応できない。
- 難しい問題3:シンボルグラウンディング問題
  - シマウマがシマのある馬だと、計算機が理解することができない。
  - シンボル(記号)がそれが指すものと接続(グラウンド)しておらず、シンボルの操作ができない。

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。

## Googleの猫(2012)



- ・YouTubeから取ってきた大量の画像をニューラルネットワークに学習させることで、下位の層のニューロンには線や点といった単純な特徴量が、上位の層には、人の顔や猫といったより複雑な特徴量が学習される。
- ・人間の視神経のモデルとして知られているものと極めて近い。

## 認識:ディープラーニングの実績(2012)

• ILSVRC2012: Large Scale Visual Recognition Challenge 2012



11

**Description Team name Error** Using extra training data from ImageNet Fall 2011 SuperVision 15.315% release Using only supplied training data **SuperVision** 16.422% 「ケタ」が違う res from classifiers using each FC ISI 26.602% Naïve sum of scores from classifiers using each FV ISI 26.646% ISI 26.952% Naïve sum of scores from each classifier with SIFT+FV, 長年の LBP+FV, GIST+FV and CSIFT+FV, respectively 特徴量設計 の工夫 Mixed selection from High-Level SVM scores and OXFORD VGG 26.979% Baseline Scores, decision is performed by looking at the validation performance.

## 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% (Krizhesvky 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%

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 12

### 運動の習熟:ディープラーニング+強化学習(2013-)

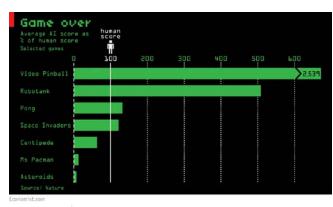
- 強化学習とは、行動を学習する仕組み。
  - 「報酬」が得られると、事前の行動を強化する。
  - 「状態」「行動」→「望ましさ(報酬ありなし)」
  - 古くからある技術だが、これまでは、「状態」を人間が定義してきた。
- 運動の習熟が可能に
  - 状態の認識に、ディープラーニングを使う。
  - DeepMindの研究者(D. Hassabisら)。その後、Googleが買収。
- 試行錯誤することによって、運動が習熟する
  - 最初は下手。繰り返すうちに、うまくなってくる。
  - 最終的には、ブロック崩しでの通路を作ったり、インベーダーゲームでの「名古屋撃ち」も。
  - 「全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









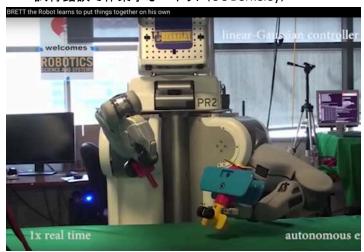


### 運動の習熟:ディープラーニング+強化学習が実世界へ(2015-)

### • 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
  - 大や猫でもできる。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)

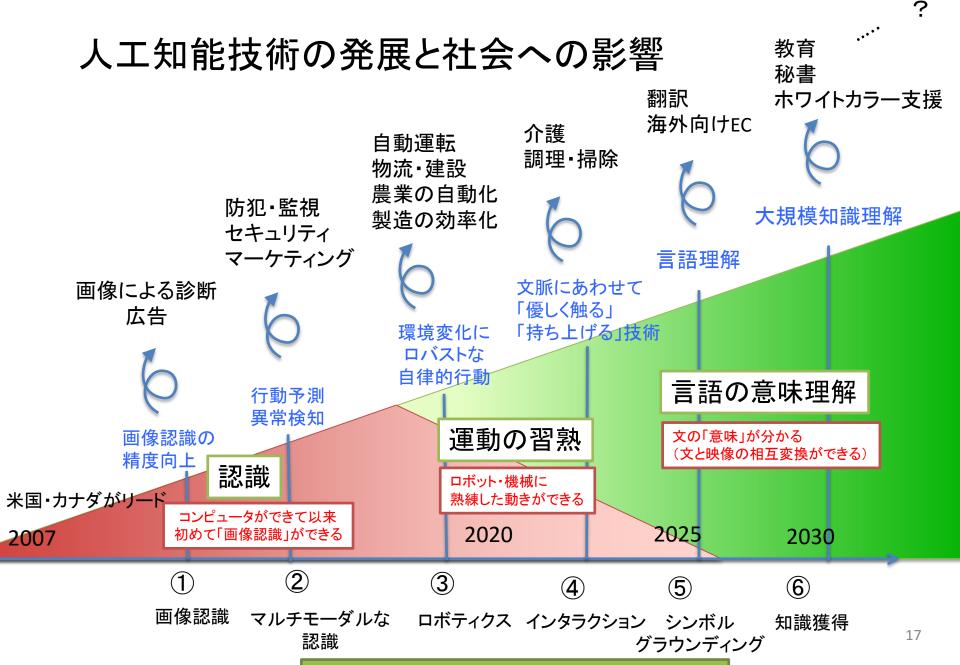
高次特徴量を、言語とひもづける

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

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



Deep LearningをベースとするAIの技術的発展

## 言語の意味理解: 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."

## 言語の意味理解: Generating Images (2015.12-)



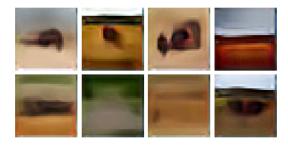
A very large commercial plane flying in <u>blue</u> skies.



A very large commercial plane flying in rainy skies.



A stop sign flying in blue skies.



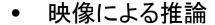
A herd of elephants walking across a <u>dry</u> grass field.



A herd of elephants walking across a green grass field.

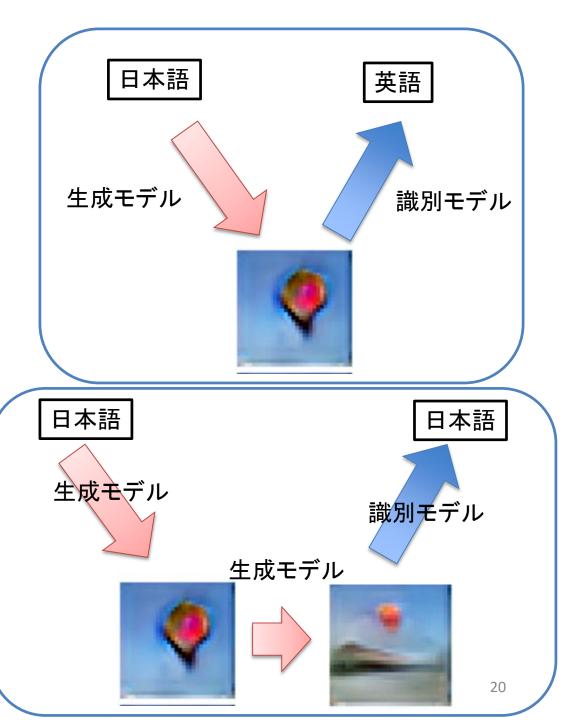
### • 画像による翻訳(意訳)

- 日本語→画像→英語
- 課題
  - 解像度
  - 画像から映像、体験へ
  - 抽象概念はどう扱うの?
- 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



- 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」

言葉の空間とパターンの空間をいったりきたりするのが人間の知能



### 「子どもの人工知能」と「大人の人工知能」

- 大人の人工知能:ビッグデータから人工知能へという持続的イノベーション
  - ビッグデータ全般、IoT全般、ワトソン、Siri、Pepper...
  - 一見すると専門家(大人)ができることができるが、人間が裏で作りこんでいる。
  - 販売、マーケティングなど。今後は、医療、金融、教育など
- 子どもの人工知能: ディープラーニングを突破口とする破壊的イノベーション
  - ディープラーニングを中心とする発展
  - 子どものできることができるようになっている
  - 人間の発達と同じような技術進化:認識能力の向上、運動能力の向上、言語の意味理解という順で技術が進展する
  - ものづくり中心。

特徴量の設計を人間がやらないといけないのが大人の人工知能、 やらなくてよいのが子どもの人工知能

### 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、解体等の自動化

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

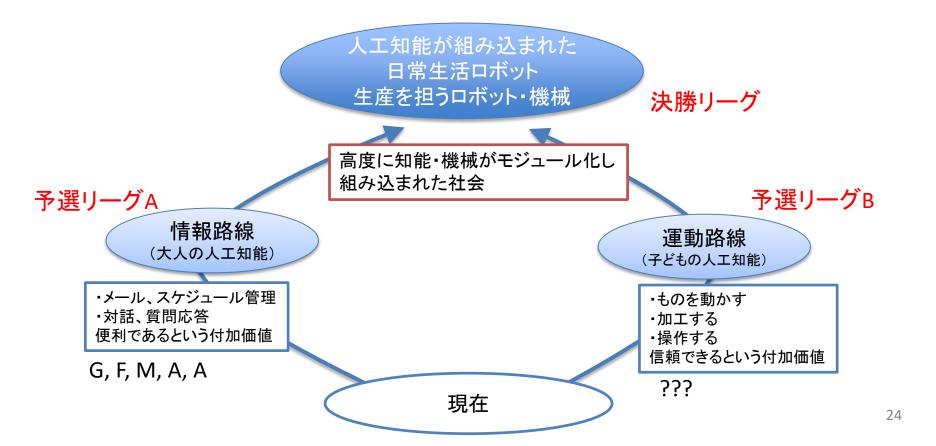
C: 計画立案を伴う運動

### 変化の本質

- 認識(画像・映像)
  - 世の中に、画像認識ができないから人間がやっている仕事がたくさんある。そこが自動化される。
  - コストが下がる。監視のコストは100分の1以下になる。
    - 森林の管理や災害の監視も。新たな事業が次々と。
- 運動の習熟
  - 機械も習熟するし、ロボットも上達するようになる。
    - 我々は、機械は「機械的な動き」しかできない、ロボットは「ロボット的な動き」しかできないと思い込んでいる。(まさにこの形容詞が表している。)
  - 自然物を相手にしているものが、自動化される。
    - 例えば、農業、建設、食品加工。
    - さらには、日常生活のロボット、生産・仕事を担う機械・ロボット
- 言語の意味理解
  - そのうち(10年~15年?)、意味理解を伴う自動翻訳が実現される
  - 日本語の障壁がなくなる。そのときに、日本経済と社会は?

### 日本は運動路線のほうが戦いやすい

- 最終的には、日常生活、仕事におけるロボット・機械の活用。状況ごとに個別性があるので、 認識能力がない状況では対応できなかった。ここにどう至るかが鍵。
- 情報路線で行く道(Google, Facebook系)と、運動路線で行く道があるのではないか。
- 海外企業・研究者は機械・ロボットに苦手意識
- 予選を勝ち進んだ企業が決勝に進むイメージ



### 日本の戦略

日本の社会課題に対して、DLを基盤とする「認識」「運動の習熟」「言語の意味理解」技術を適用する

- 農業分野に「習熟したロボット」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

「ものづくり」と擦りあわせて技術を伸ばす。それをグローバルに展開

### 人工知能による「ものづくり」の復権へ

- 少子高齢化しており、労働力が不足している。
  - 可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。
  - 認識、運動の習熟ができる機械・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
- 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