



DX & AI Forum 2024
東京コンファレンスセンター・品川
2024年9月13日

スマートロボットの時代 - AIとロボットの共進化 -



Waseda University

尾形哲也

早稲田大学 基幹理工学部 表現工学科 教授

AIロボット研究所 所長

産業技術総合研究所 人工知能研究センター 特定フェロー



尾形哲也（おがたてつや）



2

- 1993 早稲田大学 理工学部 機械工学科 卒業
- 2001-2009 早稲田大学 ヒューマノイド研究所 客員講師&客員准教授
- 2001-2003 理化学研究所 脳科学総合研究センター 研究員
- 2003-2012 京都大学大学院 情報学研究科 知能情報学専攻 講師&准教授
- 2012-現在 早稲田大学 理工学術院 基幹理工学部 表現工学科 教授
- 2017-現在 産業技術総合研究所 人工知能研究センター 特定フェロー

- 2013-2014 日本ロボット学会 理事
- 2016-2018 人工知能学会 理事
- 2017-現在 日本ディープラーニング協会 理事
- 2020-現在 早稲田大学 次世代ロボット研究機構 AIロボット研究所 所長
- 2023 文部科学大臣表彰科学技術賞（研究部門）

未来のAIには触覚も必要——「Gemini」を発表したGoogle DeepMindのCEO語る

グーグルのチャットボット「Bard」内で使えるようになった新たなAIモデル「Gemini」。Google DeepMindのCEOであるデミス・ハサビスは今後、真にマルチモーダルなAIを生み出すため、ロボット工学との組み合わせを模索していると話す。



デミス・ハサビス PHOTOGRAPH: TOBY MELVILLE/GETTY IMAGES

データドリブンへ

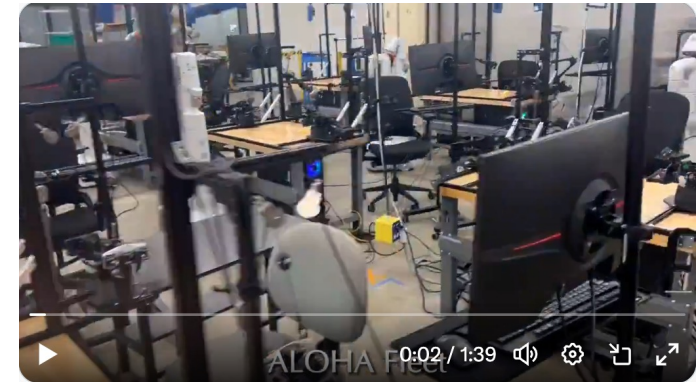
- フォーマット*や集め方が未確定
 - *おそらく”双腕モバイルマニピュレータ”が基準になる
- 深層予測学習、Diffusion Policy, Action Chunking with Transformers (ACT)などの手法間の比較ができず、技術的な壁が明確にならない
- 世界中で大量の動作データ収集を競争中



Optimus, Tesla



EVE, OpenAI&1X



ALOHA, Google

オープンで安全な基盤モデルへ

5

【特集】生成AIに挑む

INTERVIEW

LLMと ロボットが 奏でる未来

大規模言語モデル(LLM: Large Language Models)に基づく対話型人工知能(AI)技術「ChatGPT」が日本社会を激しく揺さぶり続ける。種まぶられっぱなしでは続かないと、国立情報学研究所(NII)は「LLM勉強会」を今年、2023年5月に立ち上げ、大学や研究機関、企業から300人超(2023年8月現在)の研究者・技術者が参加する大型物理実験並みの研究態勢を整えた。目指すのはLLMの動作原理の解明と日本語言語モデルの構築。日本の強みであるロボット研究とのコラボレーションも視野に入れる。人型ロボット研究に取り組む尾形 哲也 早稲田大学教授と日本語のLLM研究をリードしてきた黒橋 禎夫 NII所長が夢を語り合った。

早稲田大学 基幹理工学部産理工学科 教授
早稲田大学 AIロボット研究所 所長
産業技術総合研究所 人工知能研究センター 特定フェロー

尾形 哲也 氏
OGATA, Tetsuya

国立情報学研究所 所長/教授
京都大学大学院情報科学研究所 特定教授

黒橋 禎夫
KUROHASHI, Sadao

聞き手
辻村 達哉 氏
TSUJIMURA, Tatsuya
共同通信 編集委員兼論説委員

尾形 哲也
早稲田大学卒業後、京都大学大学院情報科学研究科准教授を経て、現職。神経回路モデルとロボットシステムを用いた、認知ロボティクス研究。特に予測学習、模倣学習、マルチモーダル統合、言語学習、コミュニケーションなどの研究に従事。2023年文部科学大臣表彰科学技術賞などを受賞。

日本語LLMの構築と
ロボットへの応用

—— LLMの動作原理をどうやって解明するのですか。

黒橋 まずやらないといけないのはトレーニングデータをオープンにすることです。ChatGPTがどういったコーパスで学習しているのかはブラックボックスで、使っているニューラルネットワークもある種、ブラックボックスのようなものです。ChatGPTに質問して答えが返ってきてても、それが何に基づくのなのかわからない。わからないことだらけで、知らないうちに著作権が侵害されているのではないかと思います。

そこで私たちは「こういうもので学習しました」とコーパスをすべて公開し、入力に対する出力をすべて分析できるような環境を作ろうとしています。すると、ある質問に対してこういう文章が答えとして出てきたけれど、この質問に非常に近い文章はコーパスにこれだけあるとか、めちゃくちゃな答えを出力するハルシネーションがどのようになっているのか、といったことが少なくとも観察はできるようなります。

その上で、このニューラルネットワークの中で何が起きているかを解明したい。それは本当にチャレンジングな目標です。また、ChatGPTとは違って、

日本語に強いものを作りたい。それをベースにいろいろな分野で新しいことに挑戦できるようになる。その中で私が面白いと思っているのは、ロボットへの応用です。

尾形 ロボットの制御にディープラーニングを使うという研究の歴史はそれほど長くありません。主要な国際会議のキーワードになったのが6年ほど前です。どういふように使っているか、まだまだわからないところがあるのですが、私は脳神経科学のバックグラウンドがあるので、その視点からディープラーニングを見直して、ロボットを安定して動かすことに使う研究をしています。

これまでは主に、人が教師となってロボットの手を動かして、タスクを畳んだり、スクランブルエッグを作ったり、といった動作を学習させてきました。ロボットは自ら動くことができるので、目の前の状況が学習した時の状況から多少ずれていても、動くことによって自分が知っている状況に近づけて、指示された動作を実行できるのです。このような考え方は能動的推論などと呼ばれています。

次の段階として、運動から言葉をつくる、あるいは言葉から運動をつくるといったことをさせたい。曖昧な言葉による指示も理解してもらいたい。しかし、運動には学習に必要な大規模データベースがありません。例えば「腕を伸ばす」という言葉の表す動きには、コーヒーを取りにいくとか、ボクシングでパンチを出すとか、さまざまな場

合がある。そのため幅広く学習する必要があるのですが、それに使える大規模データベースがない。

この壁を乗り越えるのにLLMを使えないかと考えていて、まずはLLMの中身を見てみたい。それができるとして、LLM勉強会に大いに期待しています。究極的にはLLMが学習する段階にロボットが使えないか、とも考えています。

LLMとロボットの海外事情

黒橋 研究に当たっては日本だけでは限界があるので、計算資源などの基盤が今後ある程度できあがったら、少なくともアジアや欧州と協力していこうと思っています。小さなプロジェクトでは頭脳的にも予算的にもGAFAなどに太刀打ちできない。大きくすることで多くの人に興味を持ってもらったり、企業から寄付が来たりして持続的な発展が見込めます。

ところで、LLMとロボットを結び付ける研究は海外ではどうなっていますか。

尾形 GAFAでLLMをやっている人たちはロボットに絡んでいないようです。ロボット屋もそこに興味があるわけではなく、互いが学習済みのモデルをバラバラのモジュールとして扱っています。私たちのように、LLMを作るときにロボットが要るんじゃないかと、ロボットの何かを取り込んだらLLMがもっと面白くなるんじゃないかと考えている人がいるかどうか。ただ、ChatGPTを出したOpenAIは2021年にロボット

学習データを全公開 権利者に配慮し、ユーザ が安心して使えるモデル を構築

NII Today, 第100号, 2023年9月15日

深層学習 (Deep Learning)

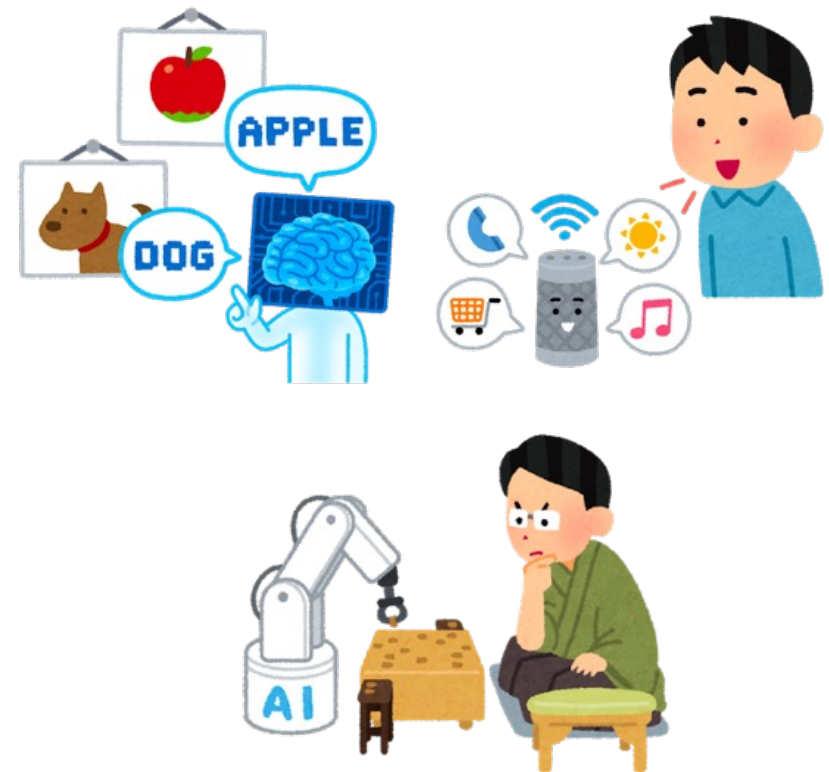
機械学習の一種

😊 深層学習は、画像、音声、自然言語、等の処理において圧倒的なパフォーマンスを示す

😞 主にサイバー空間"のみ"で有効に働き、**実世界（ハードウェア）応用事例が極端に少ない**

何故か？

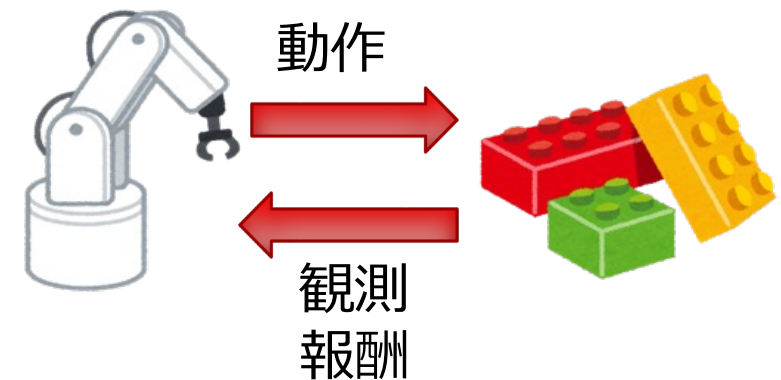
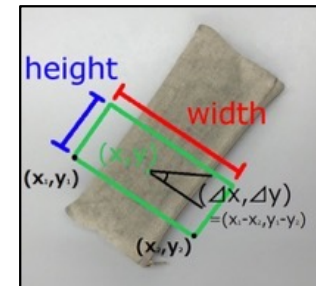
AIの革新へのブレークスルーへ！



AI（深層学習）のロボットへの応用の方法

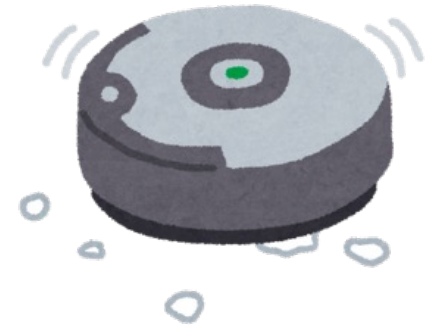
7

- 入力 → 出力の"最適"モデルを学習
 - 認識
 - 画像 → 物体の位置と姿勢
 - 強化学習
 - 動作と感覚 → 報酬の期待値
- 学習モデルの不完全性をベースとする
 - 世界は開かれている
 - 予測学習
 - “実時間”での内外感覚の予測と修正



ロボティクス-身体性の概念から 「環境－身体－神経系」のカップリング

- ビヘイビアベースドロボティクス、ソフトロボティクス
 - 特に「環境－身体」のカップリングに着目（**身体**のデザイン）
 - 神経系の機能は最小限
 - 例：受動歩行器「地面－多足機構－振動子」
- 
- 深層予測学習 (2013-)
 - カップリングの中で神経系に**深層学習**を適用
 - 視覚、触覚などの多次元の感覚と身体、環境をカップリング

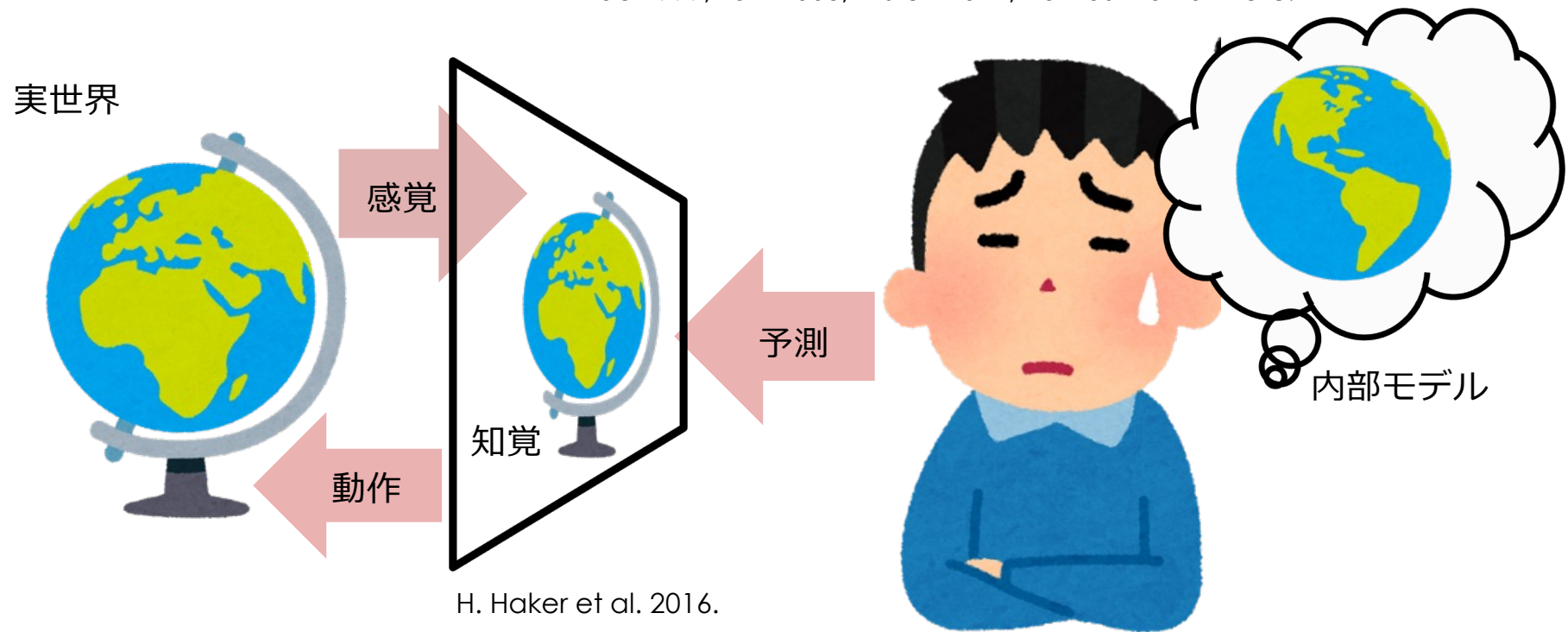


[K. Noda+, 2014]

能動的推論と予測符号化の 簡単なイメージ 1

- 環境とインタラクションを行う際の"**予測誤差最小化**"のための方策

Rao 1999, Tani 2003, Friston 2011, Yamashita 2012 etc.



能動的推論と予測符号化の 簡単なイメージ2

10

低コスト

高コスト



“知覚（認識）”を変える
ことで“予測”を変える
予測符号化（Predictive
coding） Rao, 1999



“動作”により“感覚”を
変え“予測”を変える
能動的推論（Active
Inference） Friston 2017

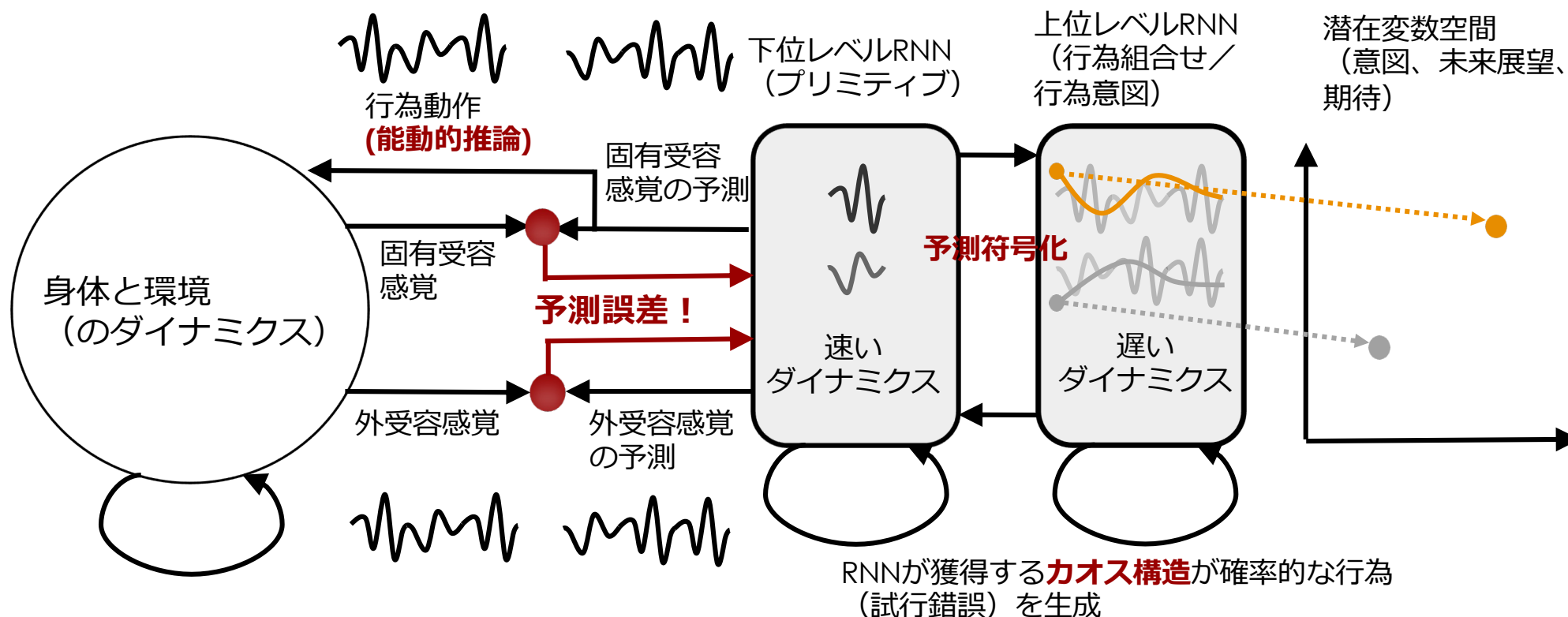


“学習”により“内部
モデル”を変える

機械学習の視点

予測符号化と能動的推論に基づく Multi-time scale RNNモデル

[J. Tani, 2003] [Y. Yamashita & J. Tani, 2008]



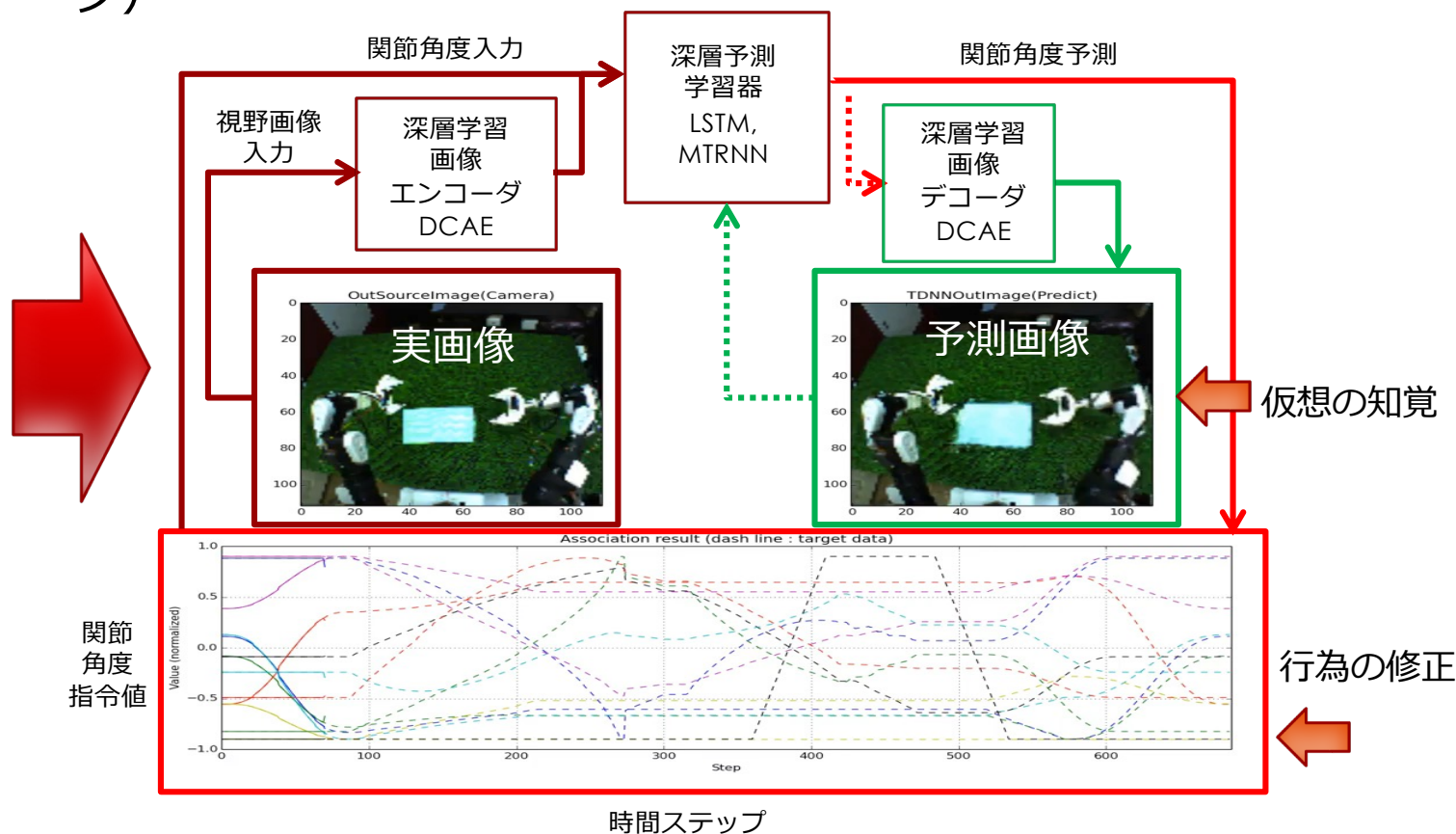
深層予測学習のフレームワーク

[P. Yang+, 2016]

12

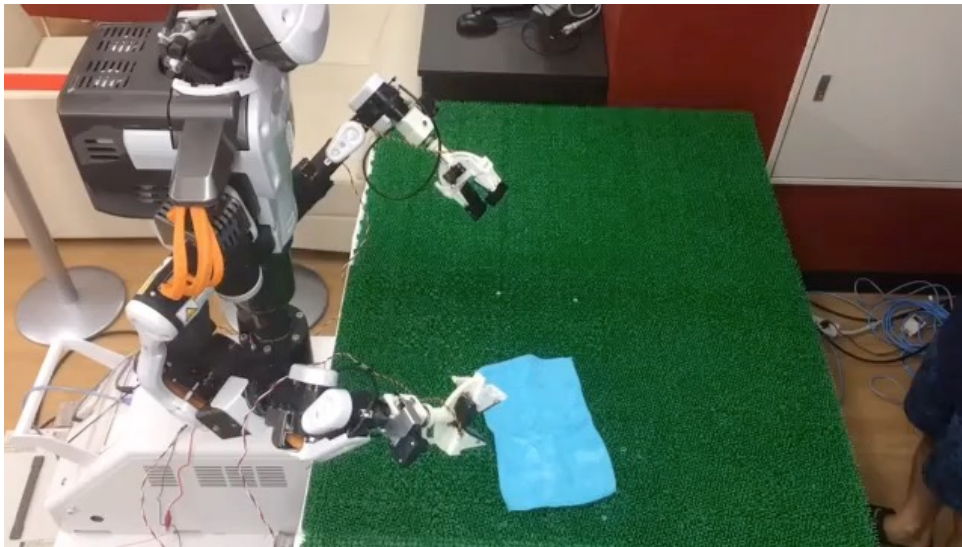
人間の操縦データ（スキル）
の収集（50動作の少量データ）

実時間で動作・画像の予測を修正し続ける
学習（世界モデル）



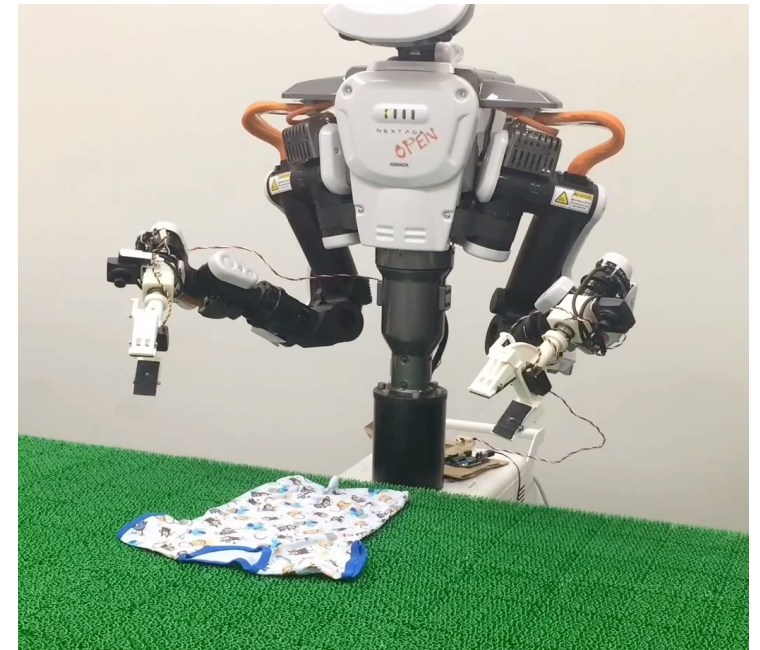
深層予測学習のデモンストレーション

- E2E motion generation
IEEE RA-L **2016.**, ICRA2017 etc.



Folding motion with unknown towel and book
Cebit 2017, CEATEC2017, iRex2017

- E2E motion generation
IEEE RA-L **2018.**



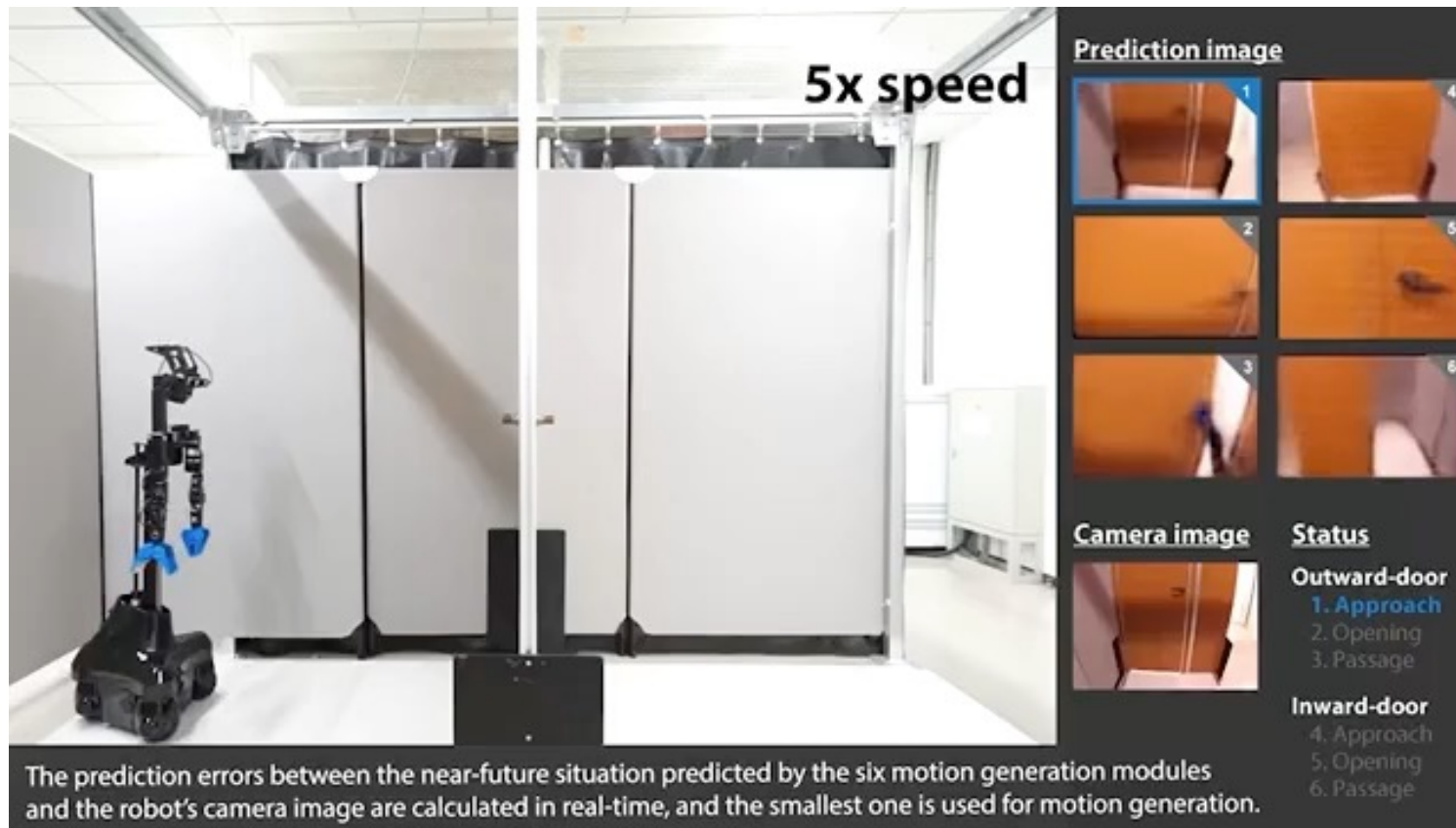
Multitask learning with an embodied predictive model

H. Ito, K. Yamamoto, H. Mori, T. Ogata,
Science Robotics(2021 Impact Factor: 23.748), Vol. 7, Issue 65, 2022



HITACHI
Inspire the Next

14



ジッパー開け 2021.5

[H. Ichiwara+, ICRA2022] [H. Ito+, ICRA2022]

15

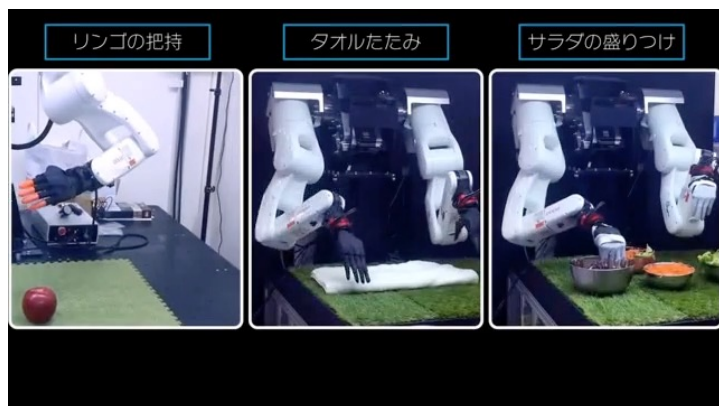


HITACHI
Inspire the Next



深層予測学習の企業応用例

16



汎用モデル
(2017)

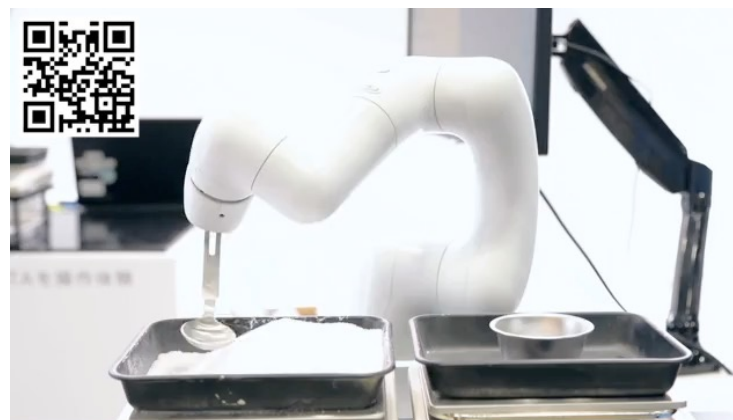
- タオル畳み
- サラダ盛り付け



ハーネス組み
付け (2021)

DENSO
DENSO WAVE

integral



粉体計量(2018)



液体計量(2021)

- 粘性の異なる液体を事前知識なしで計量

DENSO
DENSO WAVE

integral

ムーンショット



17

ムーンショット目標3

2050年までに、AIとロボットの共進化により、自ら学習・行動し人と共生するロボットを実現

プログラムディレクター 福田 敏男

名古屋大学 未来社会創造機構 客員教授



一人に一台一生寄り添うスマートロボット
“AIREC”

(AI-driven Robot for Embrace and Care)

プロジェクトマネージャ (PM)

菅野重樹

早稲田大学



ロボットの汎用性への挑戦

- スマートロボットへ -



- 例：家事（“人間と同じ方法で”料理，掃除，洗濯）は，現在のロボティクスでは超難関課題
- 専用機なら可能だが問題設定が異なる（スマートフォン）



環境の作込みが不要 & 人間と同等に複数のタスクをこなす

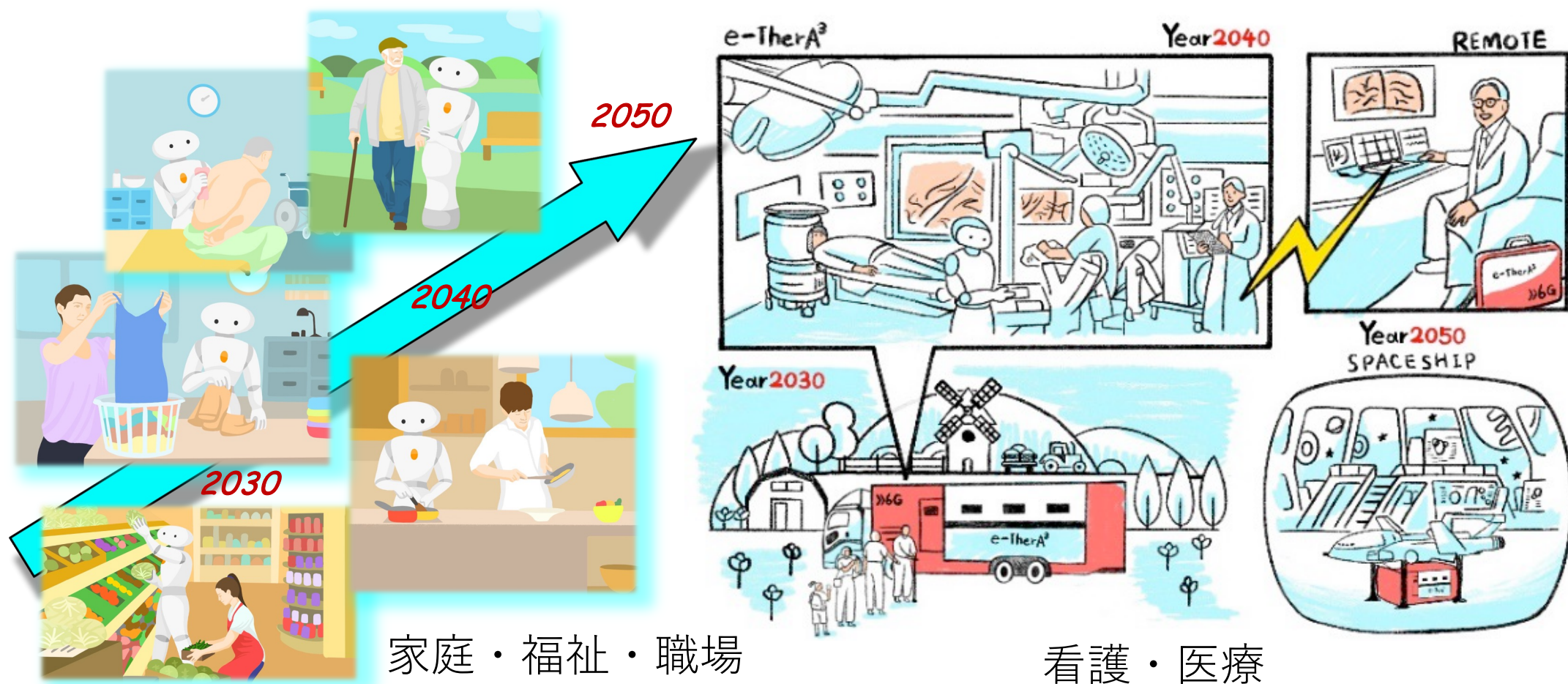
AIでロボットハードウェアの潜在能力を引き出す

?



研究開発プロジェクトの概要 (2050年へ向けたスマートロボット)

19



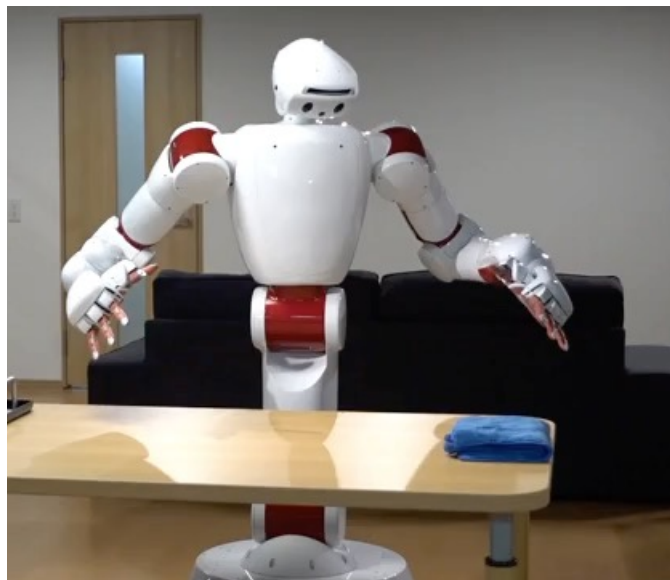
プロトタイプDry-AIREC

身長: 1660mm
全体重量: 約150kg
可搬重量: 最悪姿勢保持: 8kg、最大瞬間(肘直角時): 20kg
関節トルクセンサxインピーダンス制御(腕部7軸x2、腰部3軸)
ハンド6軸x2(把持力センサ)
人肌ゲル(腕部、胴体前面)
台車: 全方位オムニホイール

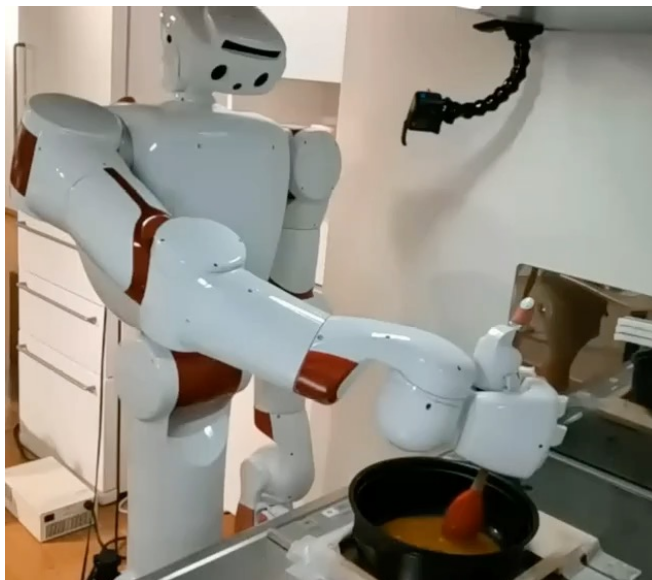


TokyoRobotics

20



掃除



調理



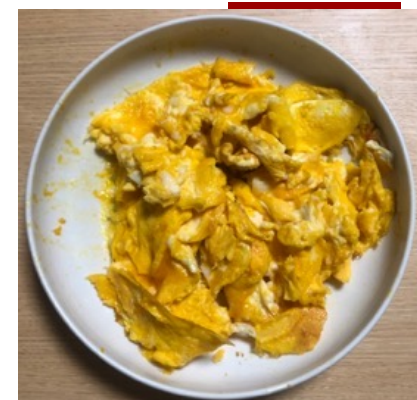
洗濯

注意機構によるAIRECの 調理学習

[平本他, IPSJ2022] 学生奨励賞



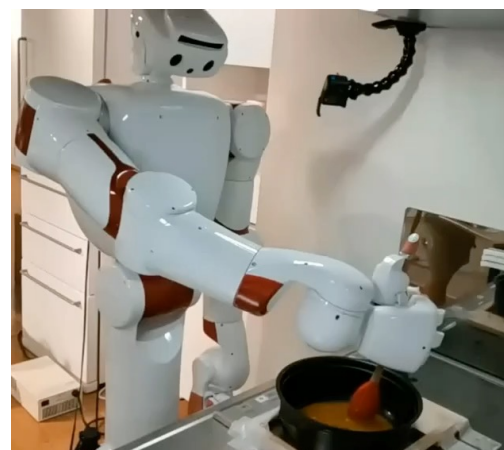
Dr. N. Saito



学習モデル
による調理



教師データ再生
による調理



火加減などの事前
情報がなくても、
卵の状態変化を判
別して調理可能

柔軟物の双腕ハンドリング ハンガー掛けタスク

- 能動的注意機構による双腕柔軟物ハンドリング
 - 国際ロボット展2023にて実演
2023年11月30日



西村経済産業大臣（当時）訪問

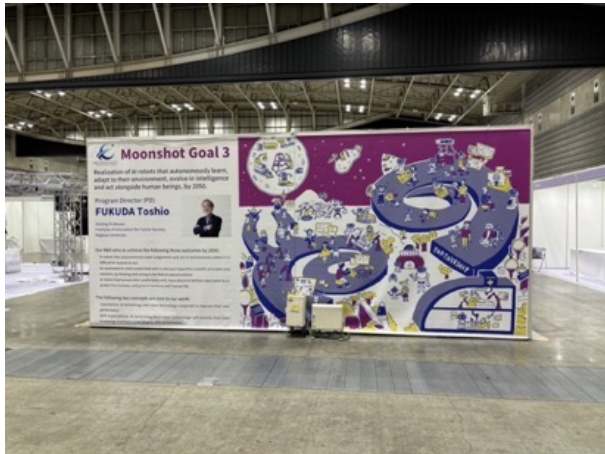


ハンガー掛け時の予測画像と
能動的注意点



Dr. H. Ito

Live demos at ICRA2024



HITACHI
Inspire the Next

TokyoRobotics



ICRA2024でのライブデモ 洗濯物取り込み



24

Mr. H. Ichiwara



HITACHI
Inspire the Next

16のデモデータ
のみからEnd-to-
end 学習
準備期間2週間

ICRA2024でのライブデモ ドア開け通り抜け全身動作学習



25

Dr. H. Ito



HITACHI
Inspire the Next

AIREC-Basic

- 技術的問題点を明確化するために、安価で簡便に多様な高スキルデータを取得可能なロボット（AIREC-Basic）を一定数準備



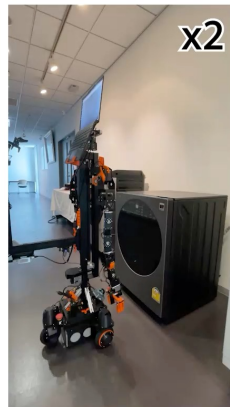
Mobile manipulation



High position



Low position



Narrow space



HITACHI
Inspire the Next

Stanford Mobile ALOHAとの比較

3輪ステアリングにより狭いフットプリントで全方向移動可能

EIPL: Embodied Intelligence with Deep Predictive Learning



27

Dr. H. Ito

- **深層予測学習を体系的に学ぶことが可能なドキュメント**
動作教示や前処理のノウハウが記載
モデル学習、推論、内部表現解析が簡単
データセット、ソースコード、学習済み重みを公開
- **複数の動作生成モデルを公開**
階層型RNN（MTRNN）や注意機構モデルなど
他にもムーンショット開発モデルを順次公開予定
- **市販ロボットを用いた動作生成（近日）**
Open Manipulatorに対応予定
動作教示からオンライン動作生成までの実行可能



Embodied Intelligence with
Deep Predictive Learning



<https://ogata-lab.github.io/eipl-docs/>

AIとロボットの共進化

AIとロボットが様々な社会ニーズに基づいて**共進化**

ロボティクスの革新

柔軟かつ強靱な身体を持つロボットが環境とインタラクションし、知的行動が自然に創発される

AIの革新

人のスキル変換で柔軟なロボットハードウェアの機能を、数少ない経験で学習できる“**深層予測学習**”で最大限に引き出す

社会受容性の把握

実証実験と社会受容性の調査により必要かつ適切な機能を選択

人々の共感を得る社会
メッセージの発信



N. Saito, T. Ogata,
H. Mori, S. Funabashi,
and S. Sugano,
IEEE ICRA2021,
Best Paper Award
in Cognitive
Robotics

共進化の二つの方向性

- 高性能ヒューマノイドから単純化
 - Optimus, Fiugre01
 - 2足？
 - 多指ハンド？
- 移動ロボットから機能追加
 - 配膳ロボット、案内ロボット、警備ロボット
 - 二つ以上のアームとグripperの追加で機能は劇的に向上
(例：ALOHAなど)
 - 安価なアームでもAIにより多様な作業が可能に
- いずれのストーリーでも力覚・触覚とAIの融合がキー

TokyoRobotics



Dry-AIREC

HITACHI
Inspire the Next



AIREC-Basic
(仮称)

おわり



Waseda University

