

大規模データに基づく ロボット基盤モデル

国立研究開発法人産業技術総合研究所

人工知能研究センター

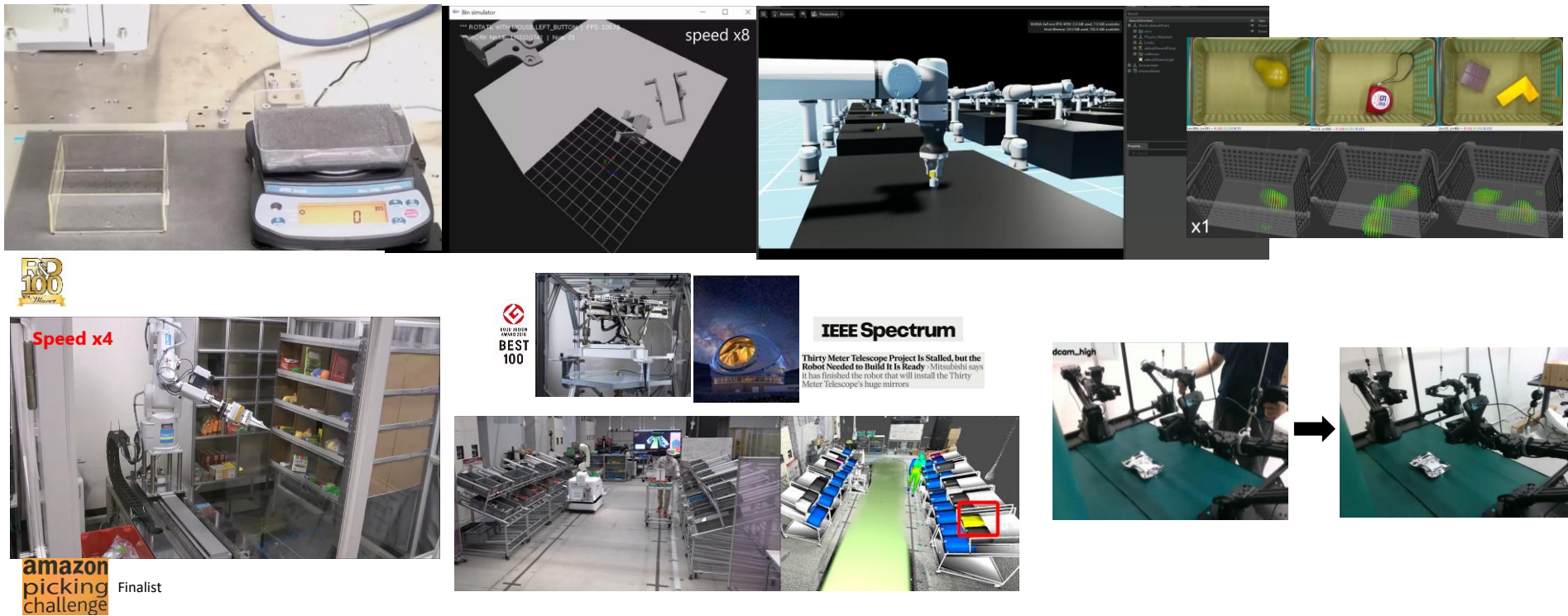
実体知能研究チーム チーム長

堂前 幸康

自己紹介

- ・ 10年間メーカーでマシンビジョン研究. ロボットマニピュレーション応用で実用化
- ・ 2018年よりAISTでAI・ロボティクス研究のチーム長

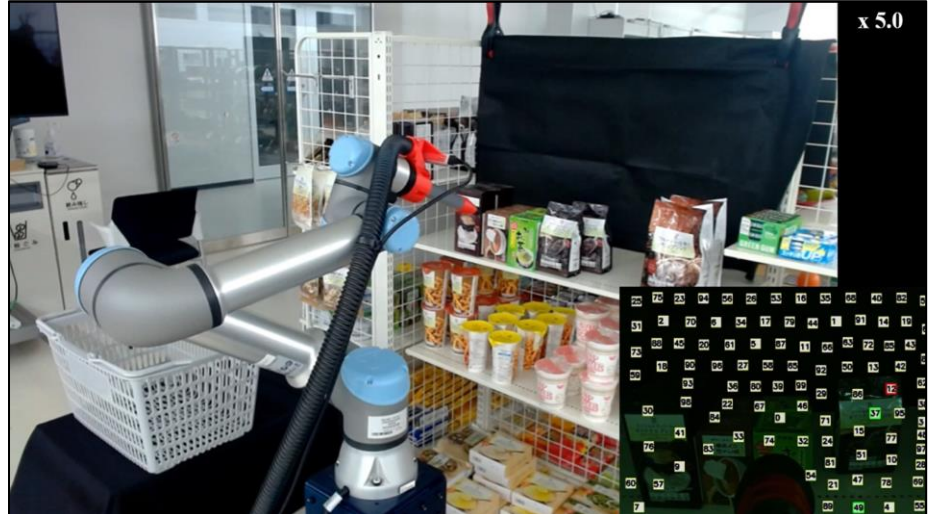
特徴量 → 学習モデル → 経験のスケーリング・多様性・質



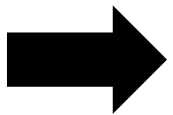
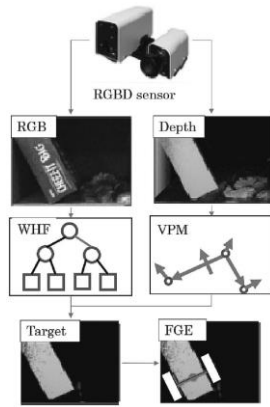
ロボットマニピュレーション研究周辺 10年の変化



Amazon Picking Challenge 2015 [Domae, JRSJ]



AISTのコンビニ模擬環境 2025 [Motoda, JRM]

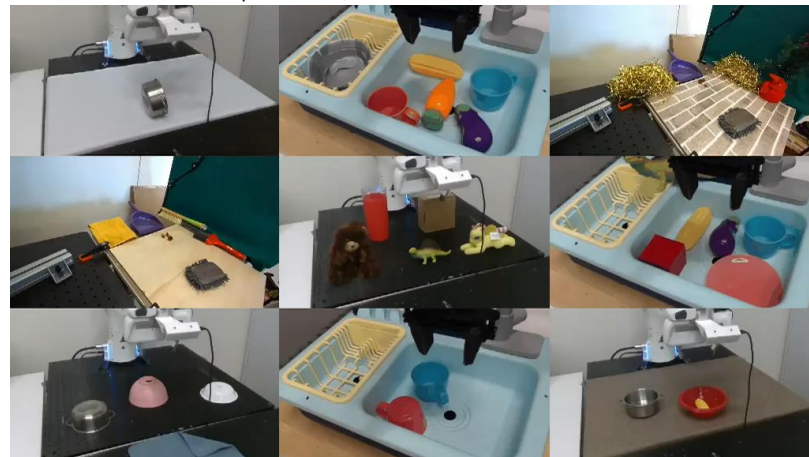


AIに身体経験を積ませる：ロボット基盤モデル

π 0, Physical Intelligence, 2024



OpenVLA, Stanford, et. al., 2024



Gemini Robotics, Google deepmind, 2025.



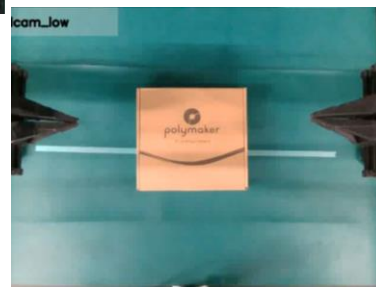
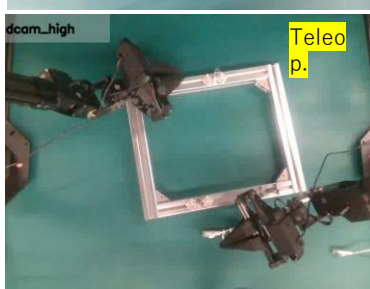
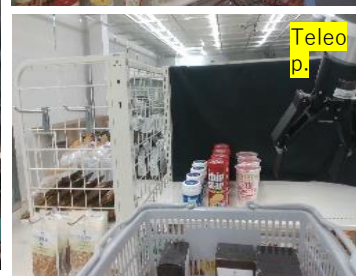
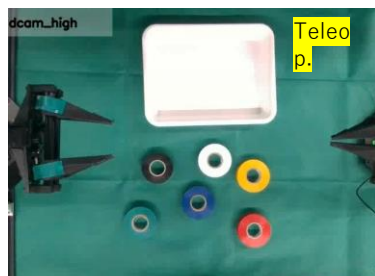
nVidia Isaac GR00T N1, nVidia, 2025.

産総研でも経験のスケール化と学習を研究中

多様なタスクに汎化するか？ 器用さは獲得できるか？ スケーリング則に欠けているものは？



Tomohiro
Motoda



ロボットはデータで
ChatGPTモーメントを起こせるか？

言語・視覚の世界ではスケーリング則が成立している

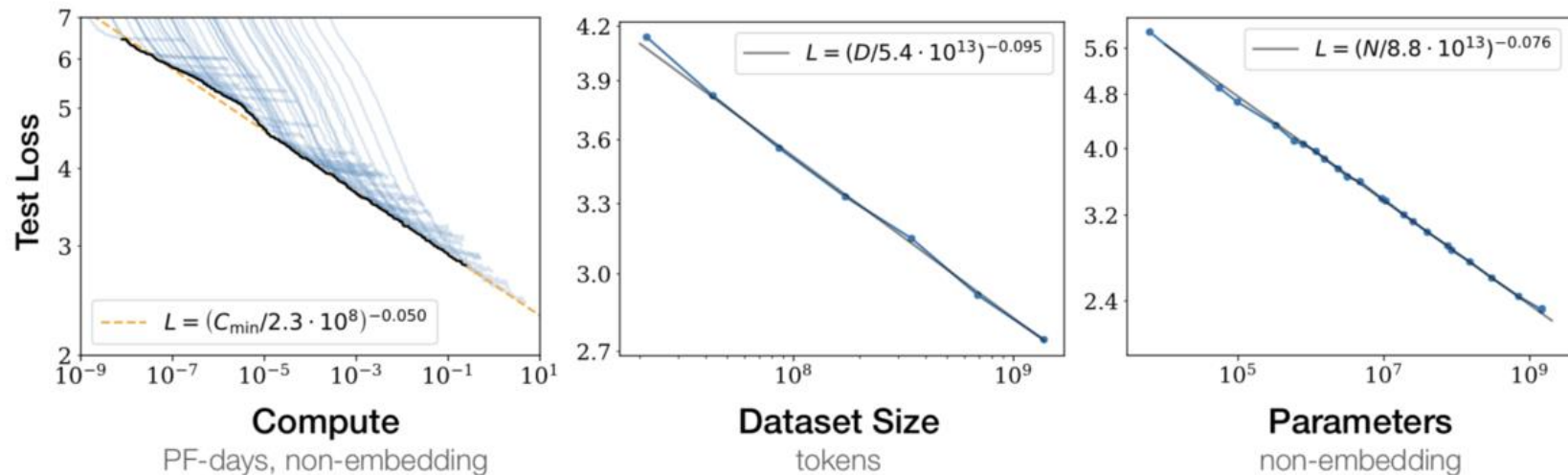
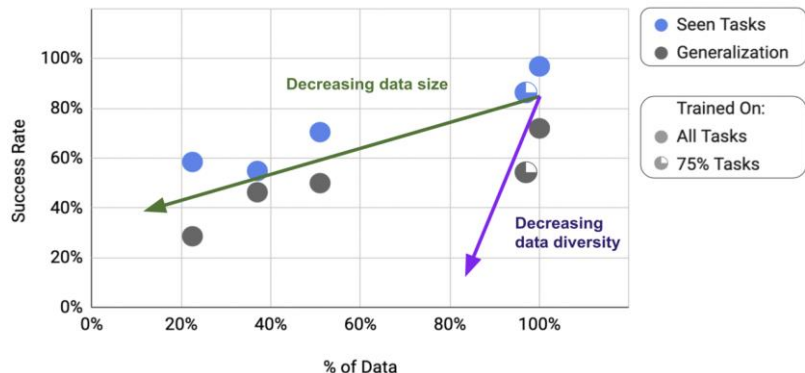


Figure 1 Language modeling performance improves smoothly as we increase the model size, dataset size, and amount of compute² used for training. For optimal performance all three factors must be scaled up in tandem. Empirical performance has a power-law relationship with each individual factor when not bottlenecked by the other two.

ロボットの基盤モデルにおいても大規模データは有効か？

Models	% Tasks	% Data	Seen Tasks	Generalization			
				All	Unseen Tasks	Distractors	Backgrounds
Smaller Data							
RT-1 (ours)	100	100	97	73	76	83	59
RT-1	100	51	71	50	52	39	59
RT-1	100	37	55	46	57	35	47
RT-1	100	22	59	29	14	31	41
Narrower Data							
RT-1 (ours)	100	100	97	73	76	83	59
RT-1	75	97	86	54	67	42	53



<https://robotics-transformer-x.github.io/>

ロボットの場合、
データスケールよりタスク多様性の
ほうが、パフォーマンスに影響している

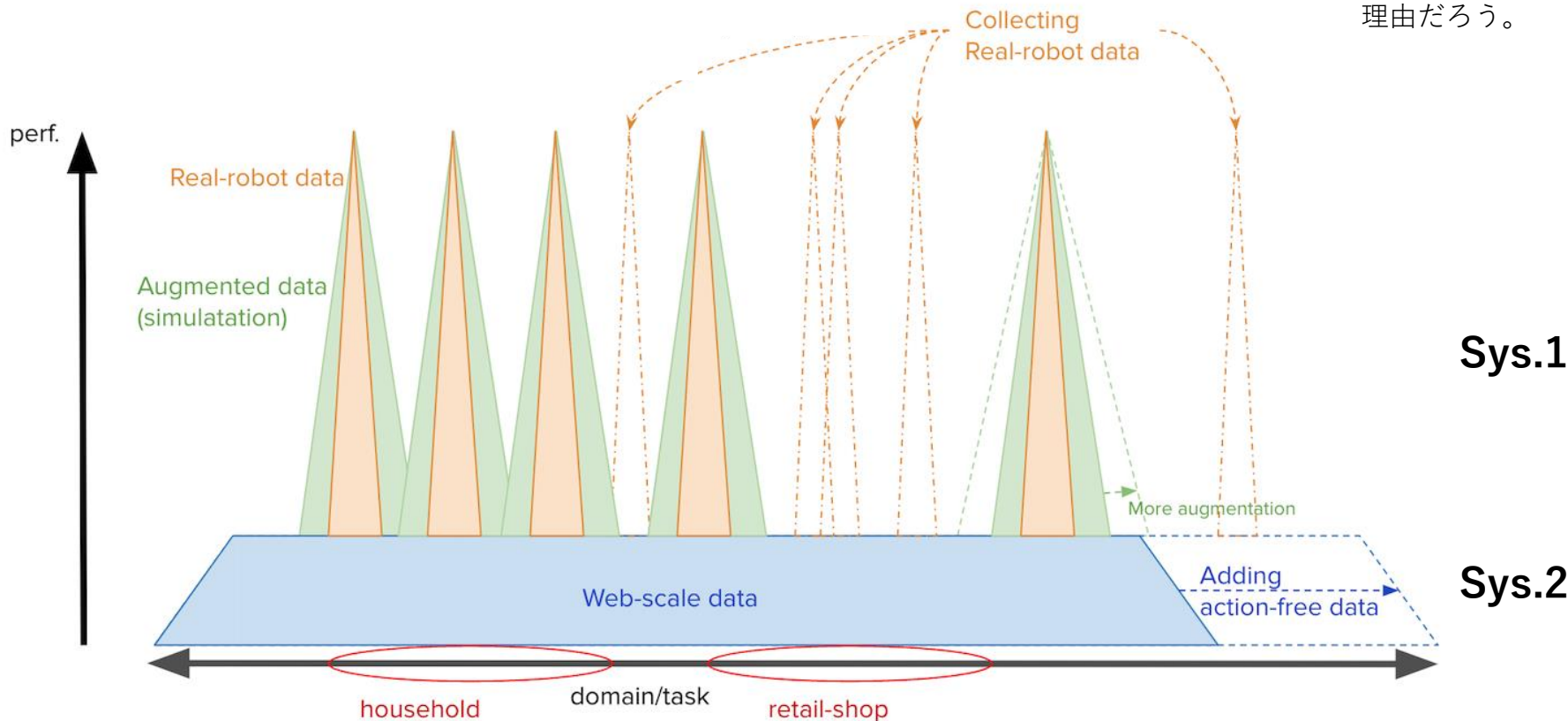
データ量
49%



多様性(タスクの種類)
25%

身体経験はタスク・身体構造スペシフィック (CoRL2024)

タスクバリエーションが効いてくるのもこれが理由だろう。

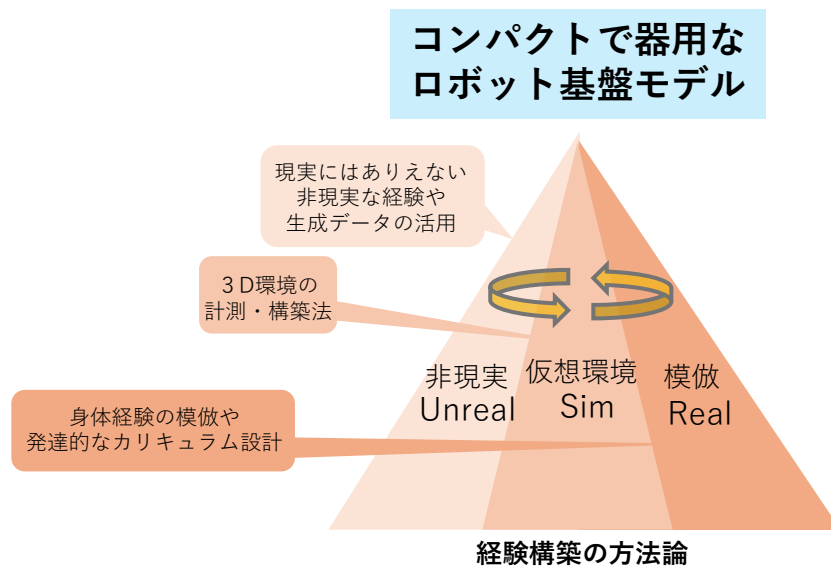


(もう1軸, ロボットの身体構造を追加しても同じ傾向が生じるはず)

経験・モデル構築の方法論の研究

GAFA的マッチョなスケーリング → コスト高、持続可能性？

経験の質、多様性、専門性などを考え、**コンパクトで器用なモデル**実現を目指す。



カリキュラムから効率的に学ぶ

Building a curriculum based on human knowledge of tasks and environments to expand the data distribution in sim, efficiently bridging the gap between sim&real.

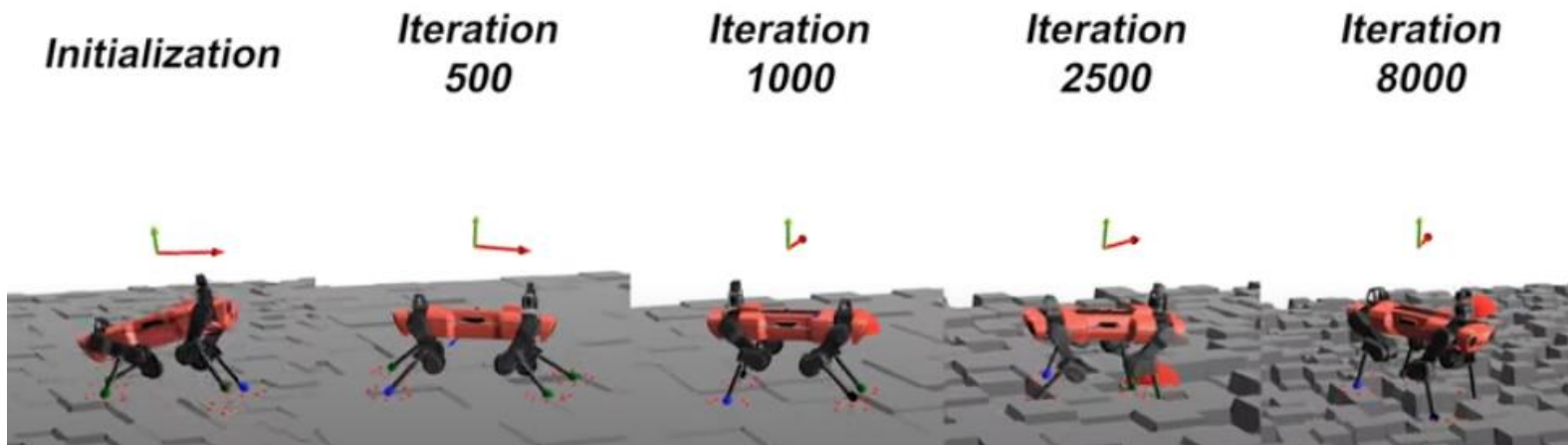


Image from J. Lee, J. Hwangbo, L. Wellhausen, V. Koltun, and M. Hutter,
“Learning Quadrupedal Locomotion over Challenging Terrain,” *Science Robotics*, vol. 5, no. 47, 2020.

カリキュラムから効率的に学ぶ



Learning for 100,000 steps
while randomizing "motion"

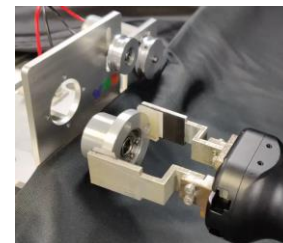
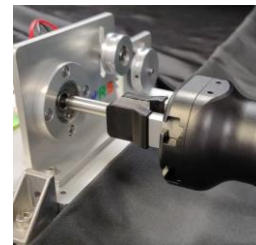
Unreal experience!

Designing a curriculum
that gradually increases
complexity by introducing
factors such as

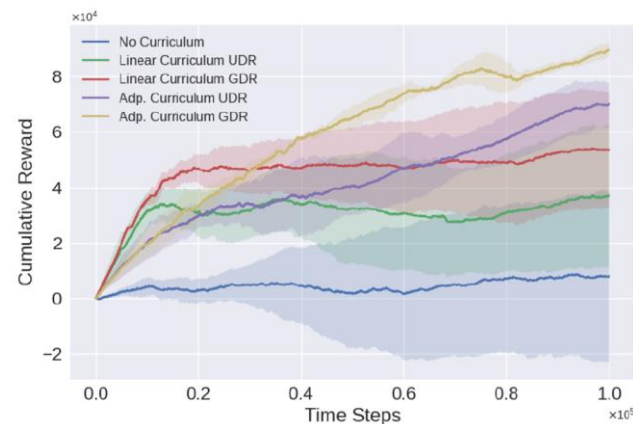
- **initial position errors,**
- **object shapes,**
- **and clearances.**



Zero-shot



Realizing peg-in-hole tasks for various different objects
that require sub-millimeter precision.



Introducing a curriculum leads to
a rapid improvement in cumulative rewards.

C. Bertran, et al., "Accelerating Robot Learning of Contact-Rich Manipulations: A Curriculum Learning Study", arXiv, 2022.

C. Bertran, et al., "Learning force control for contact-rich manipulation tasks with rigid position-controlled robots", RA-L, 2020.

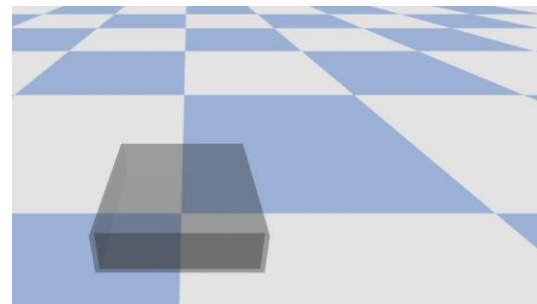
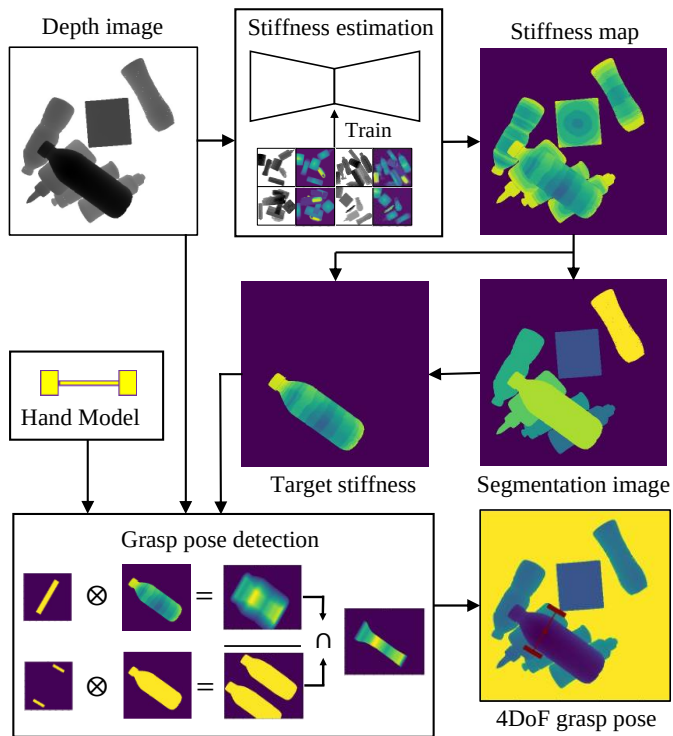
クロスモダリティを通じて学ぶ

Learning the relationship between vision(depth) and **object softness** through simulation



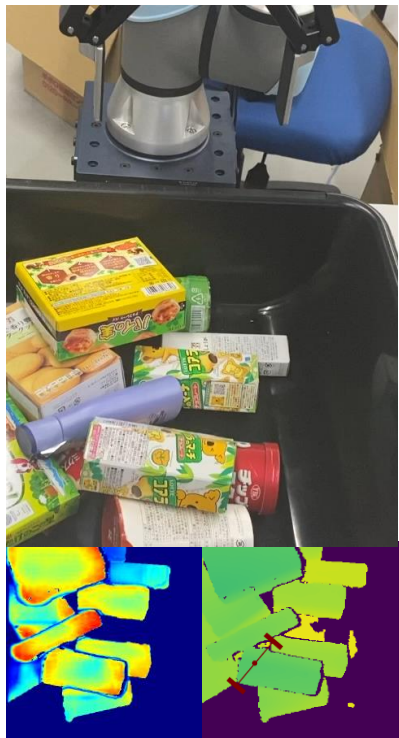
Picking by visual information

- Missing to grasp soft objects.
- Picking errors



Experiecnre augmentation
based on “depth-to-softness”
simulation

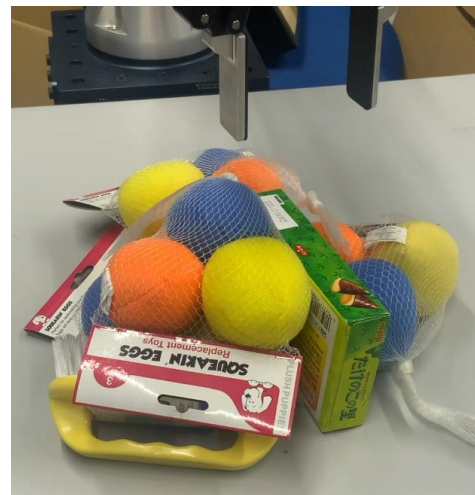
クロスモダリティを通じて学ぶ



The robot “pushes aside soft objects” to pick up the target.



Without cross-modal
ability

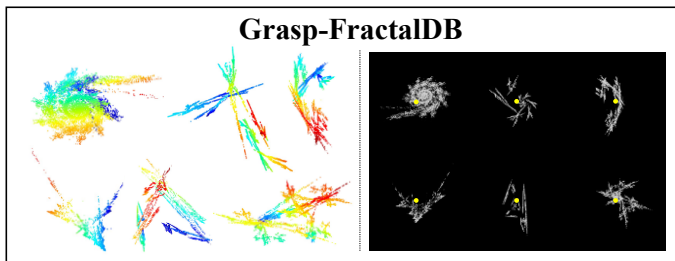


With cross-modal
ability

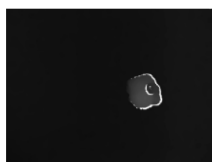
現実を超えた経験から学ぶ

フラクタルで生成した物体DBで把持プリトレーニング

Grasp-FractalDB
Pre-training



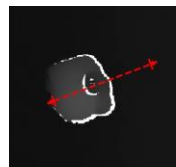
↓ Train to predict
the centroid of the fractal



Depth image

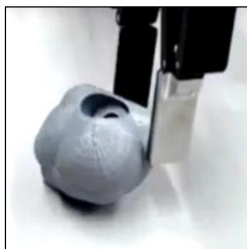


GQ-CNN



Estimate grasp quality
for parallel jaw gripper

Before(far from CoM)



with
Grasp-FractalDB



After (more robust grasp)

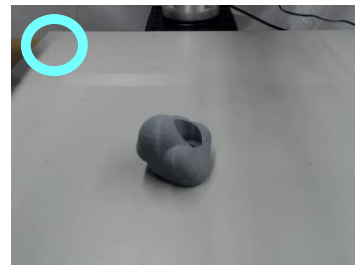
Dex-Net 2.0
Fine-tuning

Grasping
experiments

Dex-Net 2.0



Grasp-FractalDB (Ours)

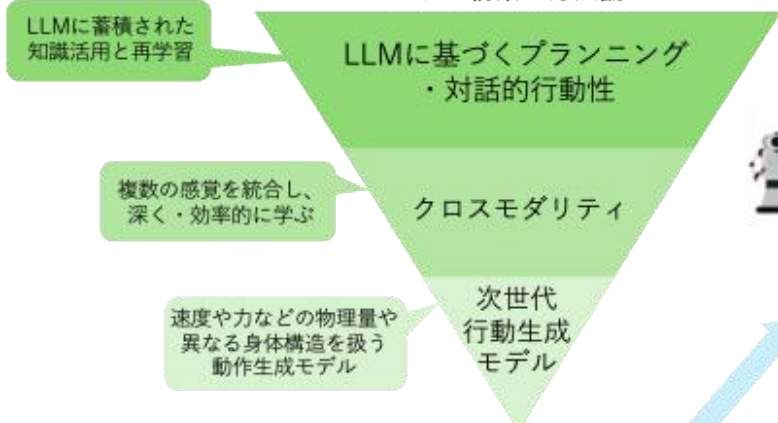


データロボティクスの現在

- やみくもなスケール化でもパフォーマンスは出てきている。モデルベース、個別AIモデルでは達成できなかった性能や器用さも、データスケールで獲得できる可能性はあり、技術的には有望
- 多くの研究者が足りないのはデータだけともいうが、アクションデータは、視覚・言語のモダリティデータとは違う特性を持っている
- 高コスト、持続可能性の問題。ChatGPTなども類似のモデルを少ないデータで実現する比較手法などがでている。持続可能性の観点からも、データ効率を先んじて考えていく必要がある
- 方法論が膨大なので、学理としての体系化や議論が必要な段階
- **実データの収集・実証と、学術的な方法論の検証・議論を両輪で回していく必要がある**

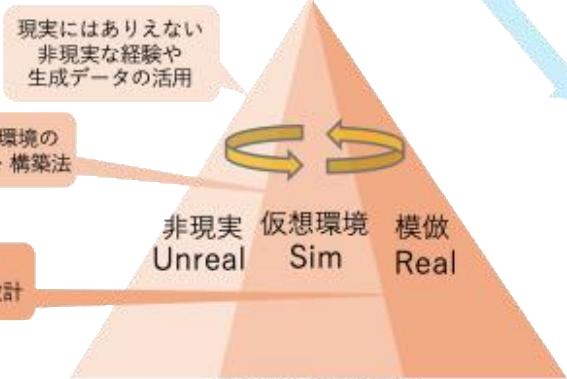
方法論と社会実装の両輪を準備

モデル構築の方法論

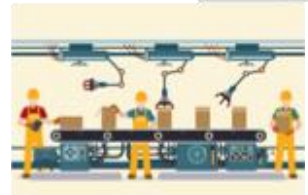


ロボット
基盤モデル

模擬環境や実環境、
実データによる実証



経験構築の方法論



データスケール化と
社会展開に向け
企業・大学・法人
などと連携

- ・方法論・ノウハウ
- ・実証データ
- ・モデルプロトタイプ
- ・模擬環境・システム

社会実装活動：これからAI・ロボットが普及していくため



遠隔操作システムの導入および
ロボット軌道データの収集



模倣学習統合ソフトウェアの開発

<https://github.com/isri-aist/RoboManipBaselines>



- ・環境・ベースライン・モデルの公開・共有
- ・AI・ロボットに興味のある国内企業への波及
- ・産学官の連携中心に



基礎動作データ収集(近日公開)



ロボ開発で基盤モデル

産総研、オープン体制構築

産業技術総合研究所 インタストリアルCPS研究センターは、ロボットの作業知識やデータを人工知能（AI）モデルに学習させたロボット基盤モデルの開発を始めた。製造や物流、オフィスなどでの困難作業を自動化する。ロボットメーカー発達過程を模した学習技術などをもつ。また、計算資源や実験環境、ロボット学習ソフトウ

ーやユーザー、システムインテグレーター（Sier）と連携し、作業適用を相互評価しながら基盤モデルを育てるオープンな開発体制を構築する。柔軟な物体操作や多様なプログラミングしきれない困難作業を、大量のエアフレームワークなどを整備してきた。これらを活用する。連携形式は技術コン

のデータを学んだ基盤モデルで実行する。さまざまな作業を学ぶことが重要なため、事業者から共通利用可能なタスク事例を集める。秘匿化処理などで連携しやすくする。

産総研は視覚データと力覚データを結び付けるAIモデルや認知サルディングや共同研究、技術研修などを事業者に応じて提案する。

現在、4社との企業との連携を開始中、さらに複数組織と協議中。法人連携などを進めている。

遠隔操作実験環境

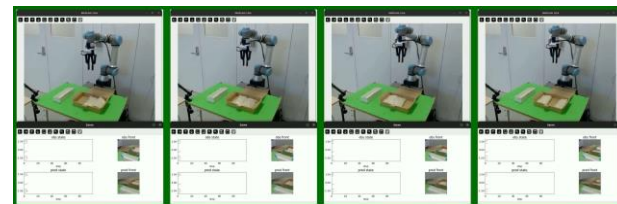


Denso Cobotta Pro
リーダフォロワ、現実に近いタスク



UFACTORY xArm 7
リーダフォロワ、力覚センサ・触覚センサ 付き

既に企業の持ち込み案件で実産業データの検証を開始。



ロボットの動作生成向け学習モデルのベースライン集を公開



Masaki
Murooka



世界の研究者や
複数企業が利用開始。
Github 190stars

HSRやヒューマノイドに対応(2.0, 25.4.)



Masaki
Murooka



実証環境（店舗模擬環境）

pickとplaceの動作データを含むロボット軌道と画像、
ラベルデータセット（公開予定）



Task name (抜粋)	Episodes
pick the instant noodle	235
place the instant noodle	230
pick the gum	198
place the gum	197
pick the chocolate confection	184
place the chocolate confection	182
pick the potato chips	177
place the potato chips	169
pick the dietary supplement	165
pick the cookie	163
place the cookie	163
place the dietary supplement	161
pick the tea bag	142
place the tea bag	142
pick the snack	133
place the snack	133
pick the seasoning	125
place the seasoning	124
pick the coffee	108
place the coffee	106

名称： 一般社団法人AIロボット協会

目的： ロボットとAIの融合により、ロボット開発の技術を革新する。それにより、社会におけるロボットの活用を推進する。

事業： AIロボットの開発促進のための取り組み

1. 基盤モデル開発に必要なデータの収集・保管・管理・公開
2. 基盤モデル・個別モデルの開発・運用・公開
3. 開発コミュニティの運営

AIロボットの社会普及のための取り組み

1. AIロボットによる効率化効果の計測・公開
2. AIロボットの安全性評価の検討・公開

その他、当法人の目的を達成するために必要な事業



理事長：尾形哲也

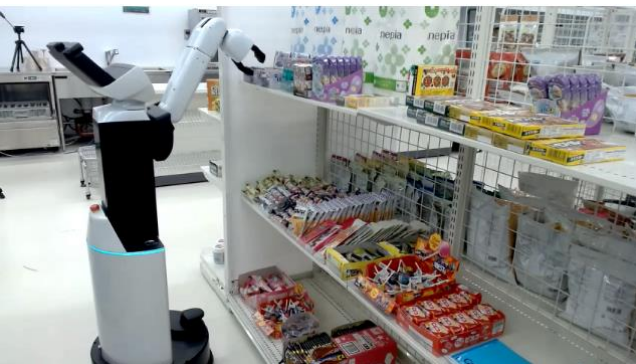
早稲田大学 理工学術院 基幹理工学部 表現工学科教授

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

国立情報学研究所 大規模言語モデル研究開発センター 客員教授

AIRoAとの連携・データ収集を開始

gaburichu_pick



umaibou_pick



oyatsucalpas_pick



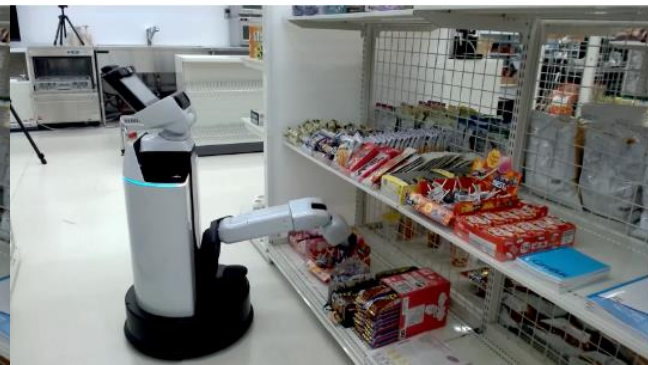
caplico_pick



umeshiba_pick



chupachups_pick



産総研の社会実装活動

- 産総研内に実証環境、実証基盤ソフトウェア、計算機資源、ロボットシステムを準備し、国内企業と連携開始
- 国内企業とはコンサル契約や共同研究で連携開始
- 法人データ収集にも参加、今後連携を進める
- 今後さらに企業連携と実タスクデータ収集を拡大

Thank you for your attention!

実体知能研究チーム (Embodied AI Research Team)

身体に知能を、知能に身体性を。



堂前幸康
チーム長
ロボット知能化
経験拡張



花井亮
主任研究員
マニピュレーション
クロスモダリティ



Floris Erich
主任研究員
3Dビジョン
DevOps



元田 智大
研究員
基盤モデル
マニピュレーション



Oh Hanbit
研究員
模倣学習
マニピュレーション



**Roman
Mykhailshyn**
主任研究員
エンドエフェクタ
(New!)

常勤6名、契約職員15名、産官学制度来所者など含め43名。