

PRESS RELEASE

豆蔵、AWS クラウド基盤上で NVIDIA のロボティクス AI 技術を活用し現場で動く「フィジカル AI」の社会実装を加速

～「フィジカル AI 開発支援プログラム by AWS ジャパン」

における研究開発成果を発表～

株式会社豆蔵（本社：東京都新宿区、代表取締役社長：中原徹也、以下 豆蔵）は、アマゾン ウェブ サービス ジャパン合同会社（以下 AWS ジャパン）が実施する「フィジカル AI 開発支援プログラム by AWS ジャパン」に参画し、フィジカル AI およびヒューマノイドロボットの産業実装に関する研究開発成果を発表いたしました。

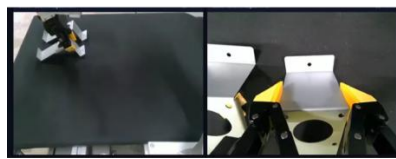
本取り組みでは、AWS の計算資源と NVIDIA の先端ロボティクス AI 技術を活用し、ヒューマノイドロボットによる板金部品のピッキング・搬送・整列作業を実現しました。また、VLA (Vision-Language-Action) モデルと従来のロボット制御技術を組み合わせた「ハイブリッド制御アーキテクチャ」を検証し、フィジカル AI の産業適用に向けた知見を獲得しました。

研究開発の内容

豆蔵では、これまで産業向けに「ヒューマノイドロボットでの板金整列アプリケーション」の研究開発を進めてまいりました。本取り組みでは、①上半身制御の VLA モデルによる知能化、②下半身制御の独自ポリシー構築、および、①②を統合した VLA モデルによる全身制御に取り組みました。

①上半身制御の VLA モデルによる知能化

VLA モデルにはオープンな NVIDIA Isaac GR00T 1.7 を用い、遠隔操作で収集した板金整列作業の 359 エピソード（作業開始から完了までの一連の作業データ）を用いて、AWS 上の GPU リソース（Amazon Elastic Compute Cloud (Amazon EC2) 「g6e.2xlarge」インスタンス）で約 8 時間ファインチューニング（追加学習）を行いました。VLA モデルとルールベース制御を相互に補完させることで、当社検証環境において上下整列 85%、左右整列 80%の成功率を確認しました。



遠隔操作で板金整列タスクのデータ収集
(収集エピソード数: 359)

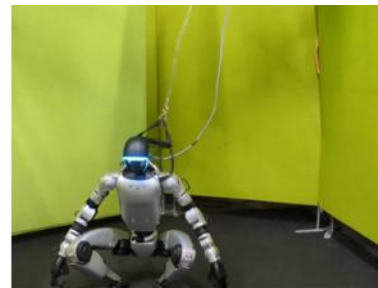


ルールベース+VLAのハイブリッド構成でのデモ
(板金上下整列: 成功率85%, 板金左右整列: 成功率80%)

VLA モデルでの制御だけでは衝突が生じる、ばら積みピッキングの成功率が低い、といった問題が発生したため、VLA モデルと力制御やルールベースピッキングとのハイブリッド構成にすることでこれらの問題に対応しました。本検証を通じ、VLA モデルにすべての動作を担わせるのではなく、従来のロボット制御技術を適切に組み合わせる「ハイブリッド制御アーキテクチャ」が、フィジカル AI の産業適用に向けた有効なアプローチとなり得るとの知見を得ました。

②下半身制御の独自ポリシー構築

NVIDIA のオープンソースワークフローAGILE(オープンなロボット学習用フレームワークである NVIDIA Isaac Lab ベース) を用い、AWS 上の GPU リソース (Amazon EC2 「g7e.2xlarge」 インスタンス) で深層強化学習を約 32 時間実行し、ヒューマノイドロボットがしゃがみ動作や前傾動作を行う独自の運動ポリシーを構築しました。



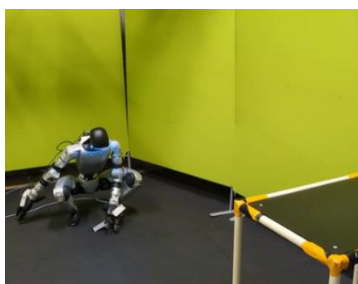
報酬関数設計やランダムマイゼーションを行い、しゃがみ動作、前傾動作の独自ポリシーを実現した。

報酬関数の設計に加え、摩擦や傾きなどの物理条件を意図的に変化させる「ドメインランダムマイゼーション」を行うことで、シミュレーションと実環境の差異 (Sim2Real ギャップ) による動作への影響を抑制し、学習した運動ポリシーを実機へ適用するためのノウハウを獲得しました。

①②を統合した VLA モデルによる全身制御

上半身制御の VLA モデルと下半身制御の独自運動ポリシーを組み合わせ、ヒューマノイドロボットの全身制御に取り組みました。下半身については、②で構築した運動ポリシーを用いて複数の関節動作を「腰

の高さ」や「前進」といった高レベルな制御単位にまとめることで、VLA モデルが直接扱う制御自由度を削減しました。その上で、これらの制御自由度を含めて遠隔操作による作業データを収集し、VLA モデルをファインチューニングすることで、床に落ちた板金部品をしゃがんでピックアップし、搬送、整列する一連の作業を実現しました。



1. 板金部品をしゃがんでピックアップ



2. 板金部品を搬送



3. 板金部品の整列

本検証では、VLA モデルが直接扱う制御自由度を対象タスクに必要なものに限定することで、VLA モデルのファインチューニングに必要な作業データを 43 エピソードに抑えながら、一連の全身動作を実現しました。一般に、多自由度のヒューマノイドロボットを VLA モデルで全身制御する場合、多数の作業データを収集する必要があります。本検証を通じ、下半身の複雑な関節動作を運動ポリシーによって抽象化し、VLA モデルが扱う制御自由度を削減することが、学習データ収集の負担を抑える上で有効な手法であることを確認しました。

今後の展望

豆蔵は、本プログラムを通じて獲得した知見を基に、VLA モデルをはじめとする AI 技術と従来のロボット制御技術を組み合わせ、製造・物流などの産業現場に適したフィジカル AI アプリケーションの開発・導入支援を強化してまいります。

商標および技術に関する補足事項

本プレスリリースに記載されている NVIDIA 技術（NVIDIA Isaac、Isaac Lab 等）に関する言及および検証成果は、豆蔵が自社で実施した検証結果に基づくものであり、NVIDIA による個別の承認や共同開発を示すものではありません。

株式会社豆蔵 会社概要

システム工学、ソフトウェア工学、ロボット工学の深い知見と、ソフト・AI・メカ・エレキのフルスタックな技術力を強みとするテックカンパニーです。先進的なテクノロジーを産業現場へ応用することを得意とし、製造・物流・モビリティ・クラウドなど幅広い領域でイノベーションを創出しています。

本社所在地	東京都新宿区西新宿 2-1-1 新宿三井ビルディング 34 階
設立	2020 年 11 月 11 日
資本金	130,714,250 円
代表者	代表取締役社長 中原 徹也
WEB サイト	https://mamezo.tech/

本プレスリリースに関するお問い合わせ

株式会社豆蔵 AI ロボティクス営業部

Tel 03-5339-2114

お問い合わせフォーム <https://mamezo.tech/corporate/contact/>