



Open Arm

# ALOHAの紹介

# 自己紹介

- **氏名:** 赤見坂篤記 (@aakamisaka)
- **所属:** 名古屋大学大学院情報学研究科  
(博士後期課程)
- **Hi研究所での活動**
  - 遠隔操作型ロボットの  
模倣学習技術の開発検討
- **個人活動 (?)**
  - AGIRobots株式会社を設立し  
汎用人型AIロボットの实用化に向けて  
研究開発を実施
  - 今月下旬 (2024/11/20-22) に行われる  
EdgeTech+2024にてとある企業様の  
展示ブース内で動態展示を実施予定



# ALOHAとは



<https://youtu.be/4VfNskoPUZg?si=Uz9gd7odgiqeySYB>

## ALOHA登場の背景

### ファインマニピュレーションの難しさ

- ファインマニピュレーション
  - ▶ ミリ単位の誤差でも失敗につながるような高精度な操作が求められる
- 具体例
  - ▶ 財布からお札を取り出す動作
  - ▶ 料理に必要な動作もろもろ



### 従来システムの課題

- ファインマニピュレーションを実現するには一般的に以下が不可欠

高価なロボットや  
高精度なセンサ

細かなチューニング

ハードウェアの  
コスト増加

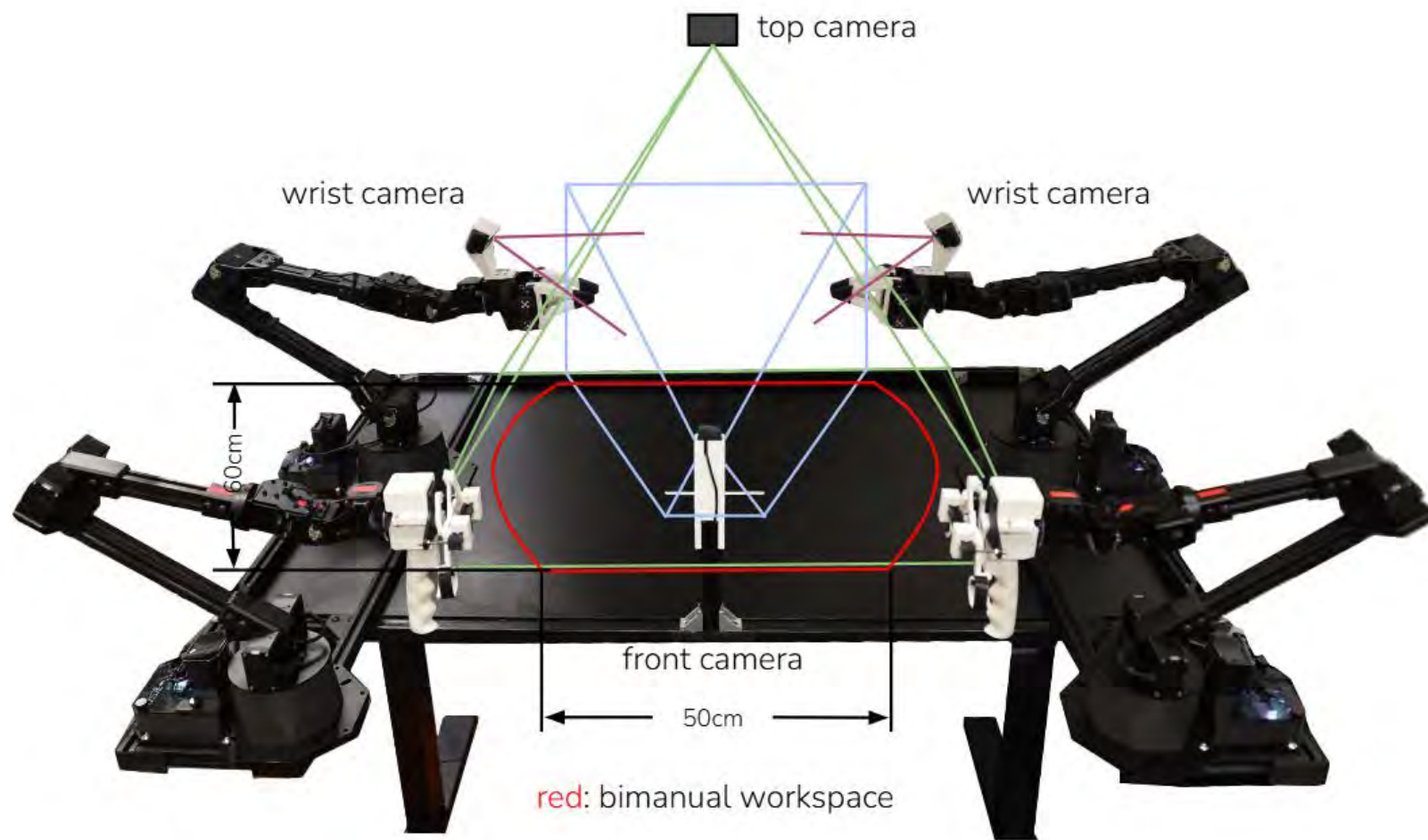
人件費の増加

社会実装における大きな障壁

安価なハードウェアによる  
ファインマニピュレーション実現

人間のデモンストレーションから  
学習する新たなモデル(ACT)の提案

# ロボットの概要 - ハード側



## ロボットアーム

フォロワー側: ViperX 300  
リーダー側 : WidowX 250

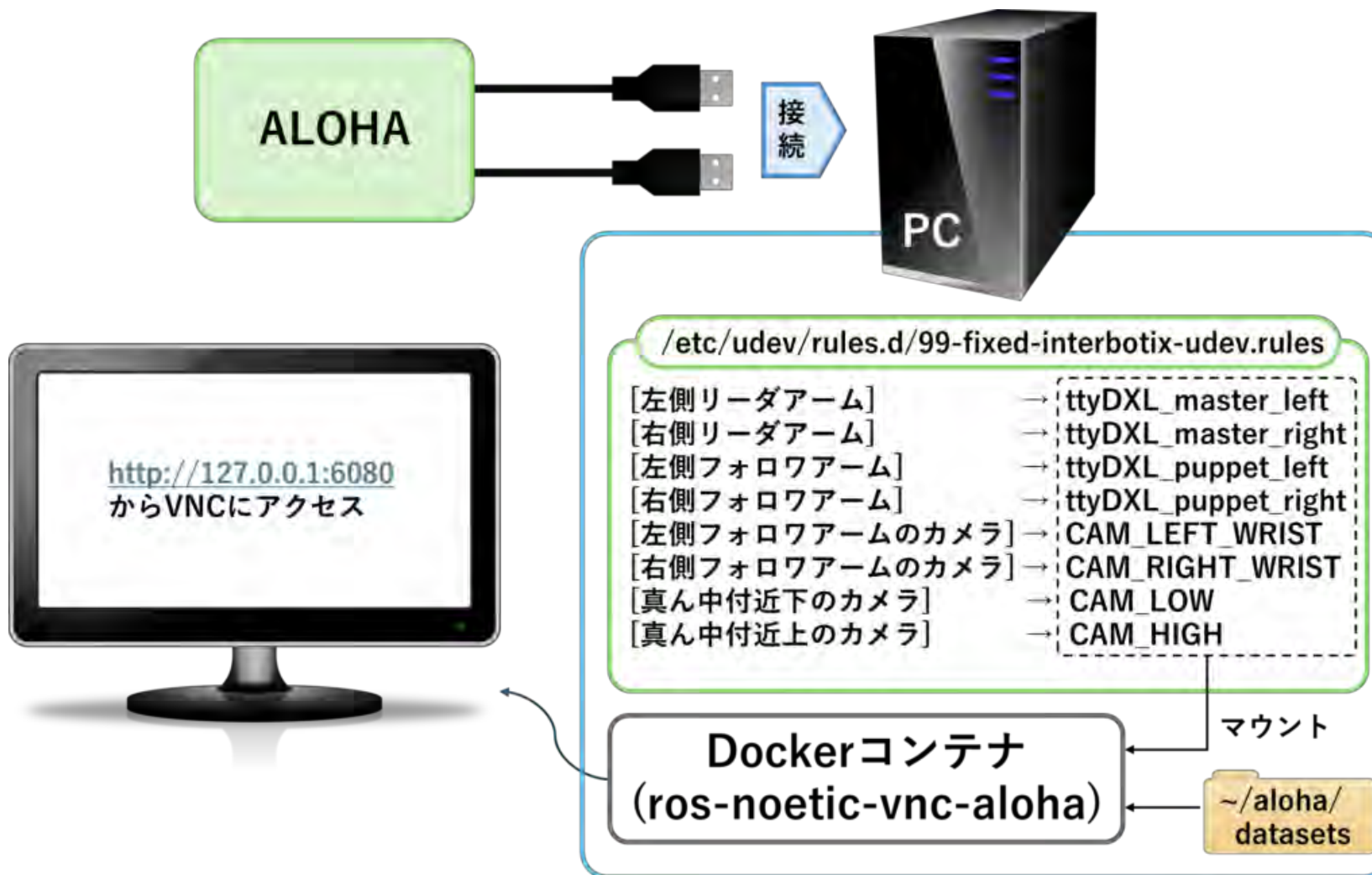
## カメラ

型番: Logitech C922X

設置場所:

- フロント
- トップ
- 右フォロワーアーム手首
- 左フォロワーアーム手首

## ロボットの概要 - ソフト側



## ACT – 概要

### ACT: Action Chunking with Transformers

CVAE

ニューラルネットには  
Transformerを採用

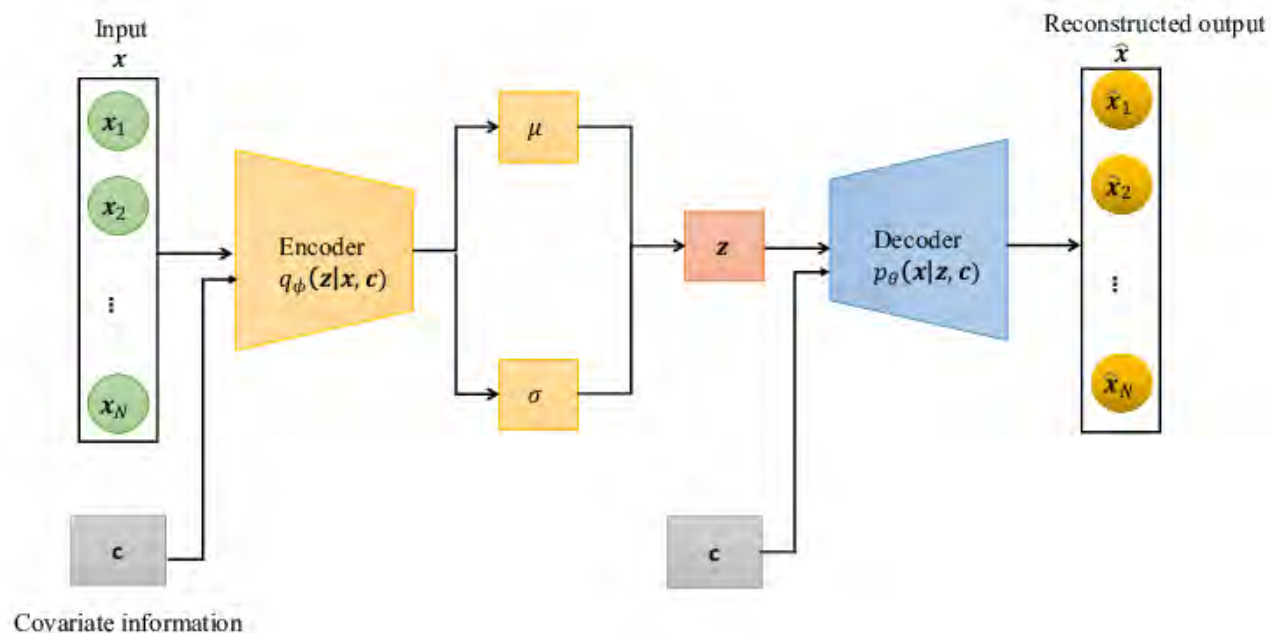
Action Chunking

フォロワーの  
現在角度を条件に

観測情報として  
画像を入力



## 全体でみるとCVAE (条件付き変分オートエンコーダ)



Degradation Prediction of Semiconductor Lasers using Conditional Variational Autoencodeより引用

## CVAE

- AE (オートエンコーダの一種)
- 潜在空間を正規分布で近似
- 条件付けに基づく生成が可能

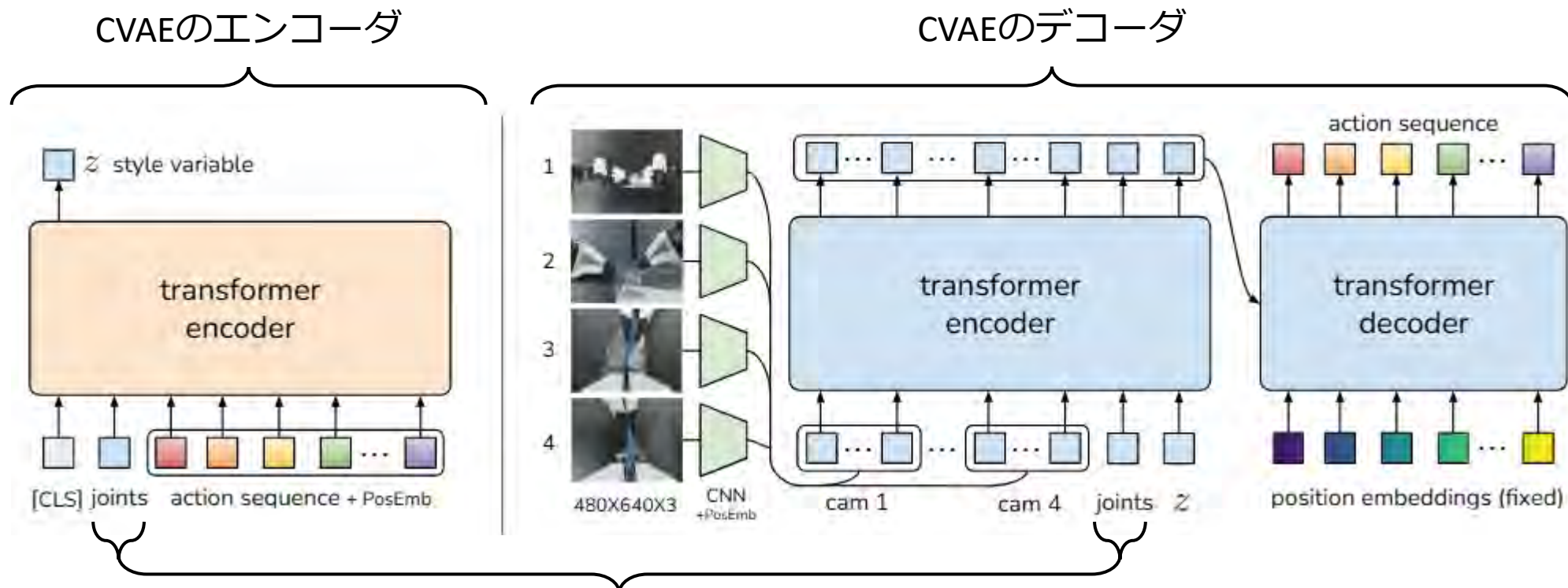
### ポイント

CVAEの条件情報はエンコーダとデコーダの両方に入力される

条件付き情報を考慮した潜在空間の獲得がなされる

# ACT - アーキテクチャ

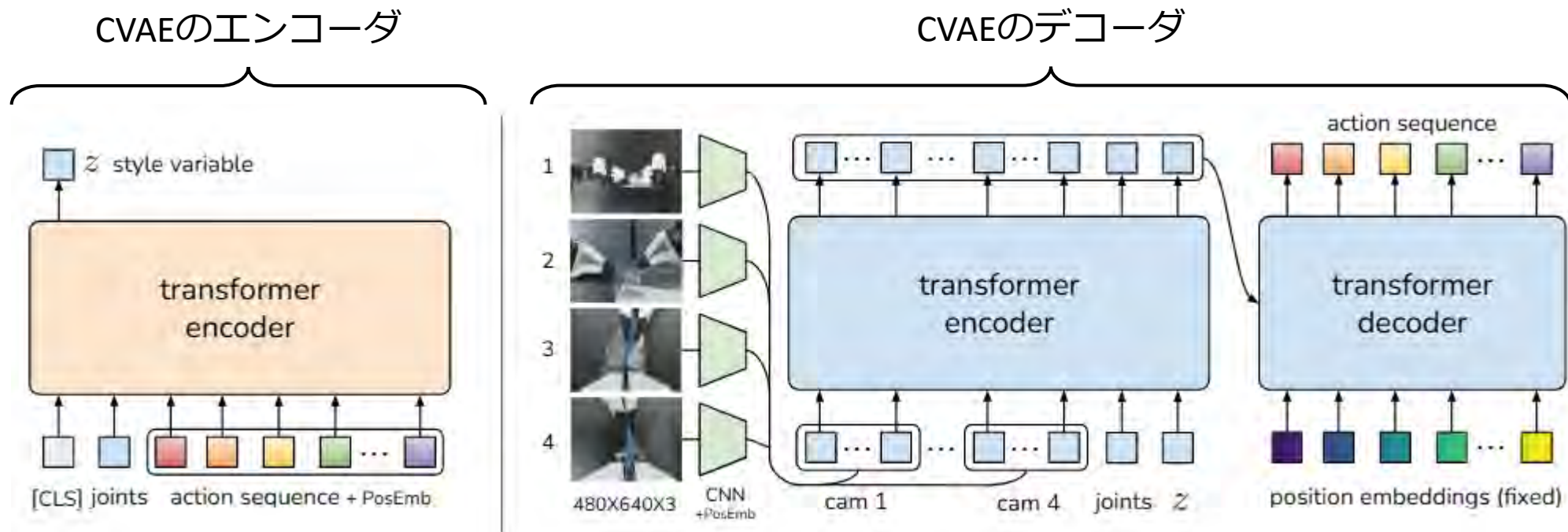
全体で見るとCVAE (条件付き変分オートエンコーダ)



フォロワーアームの現在角度が条件に  
 Learning Fine-Grained Bimanual Manipulation with Low-Cost Hardwareより引用

# ACT - アーキテクチャ

全体で見るとCVAE (条件付き変分オートエンコーダ)

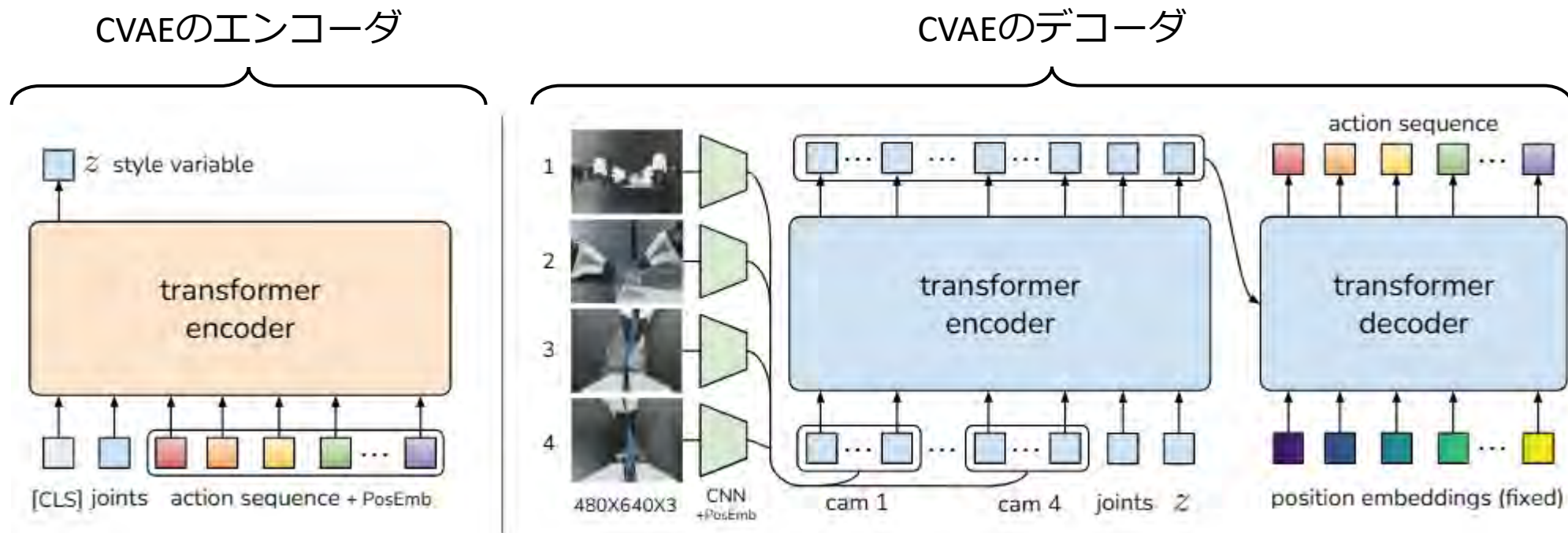


画像入力は条件じゃなさそうだけど何？

# ACT - アーキテクチャ

全体で見るとCVAE (条件付き変分オートエンコーダ)

ACTでは新しい考え方として、DecoderをPolicyだと見た時の観測 (observation) 入力として解釈



画像入力は条件じゃなさそうだけど何？

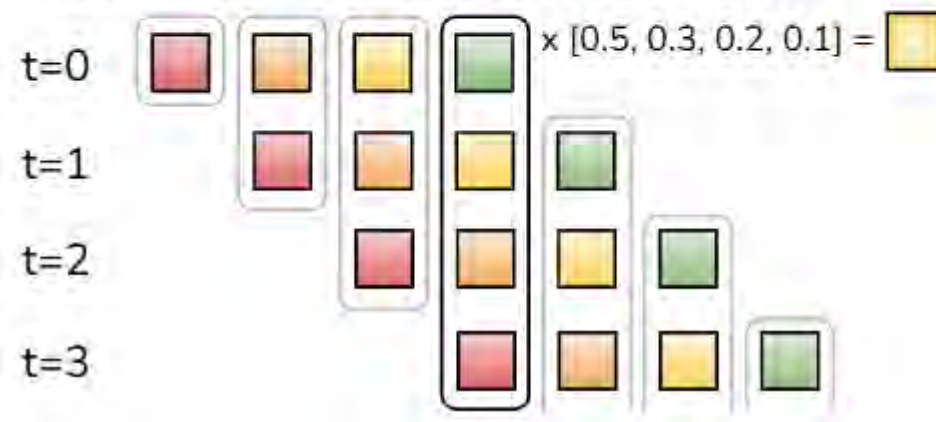
Learning Fine-Grained Bimanual Manipulation with Low-Cost Hardwareより引用

- Action Chunking
  - 短い行動系列として学習することで再利用可能に
- Temporal Ensemble
  - 大きく軌道がズレることが無いように少しずつズラすようにしてバッチ学習
  - 誤差蓄積を軽減

### Action Chunking



### Action Chunking + Temporal Ensemble



Learning Fine-Grained Bimanual Manipulation with Low-Cost Hardwareより引用

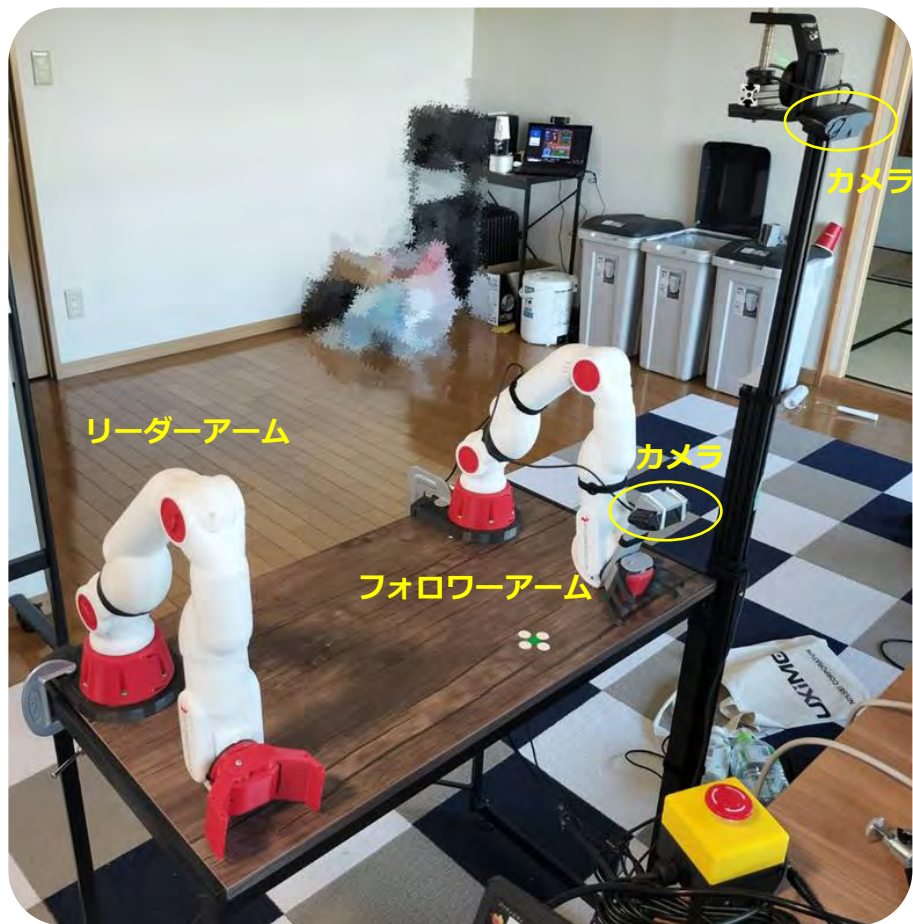
# 研究所での実験紹介 1



小さなペットボトルのキャップを開ける動作の獲得結果

## 研究所での実験紹介 2

CRANE-X7を用いた単腕遠隔操作システム



Bi-ACTを用いて紙コップを重ねる動作を学習