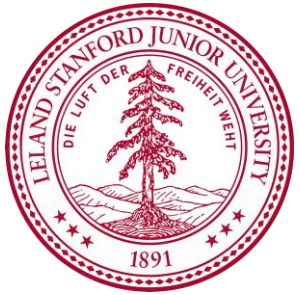
Rodrigues Network for Learning Robot Actions

Jialiang Zhang[1,2,*], Haoran Geng[3,*], Yang You[2,*], Congyue Deng[1,2],
Pieter Abbeel[3], Jitendra Malik[3], Leonidas Guibas[2, †]

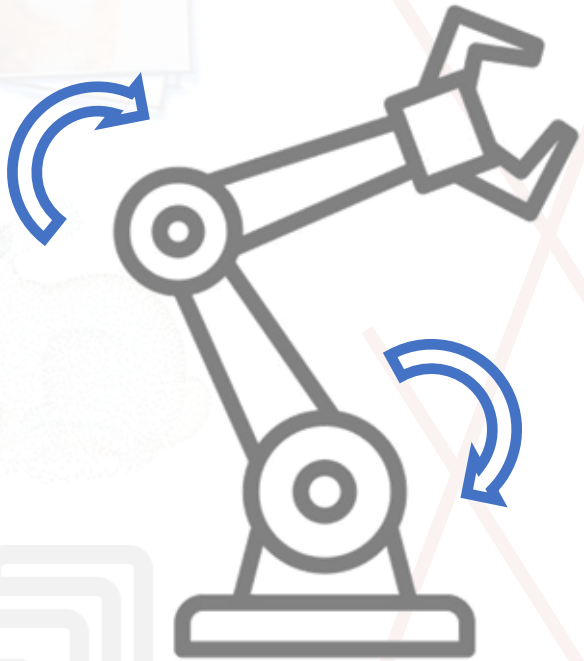
†: corresponding author *: equal contribution



ICLR
International Conference On Learning Representations

Motivation

Action: $\boldsymbol{\theta}$



Image

Transformer

Point cloud

CNN

Text

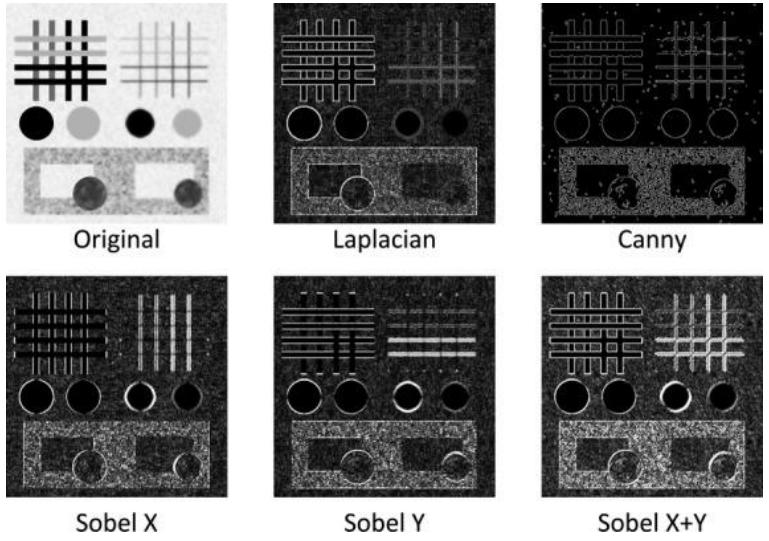
PointNet++

Specialized networks for robot actions?

How to Embed Inductive Bias into Networks: CNN

2D Image Filtering

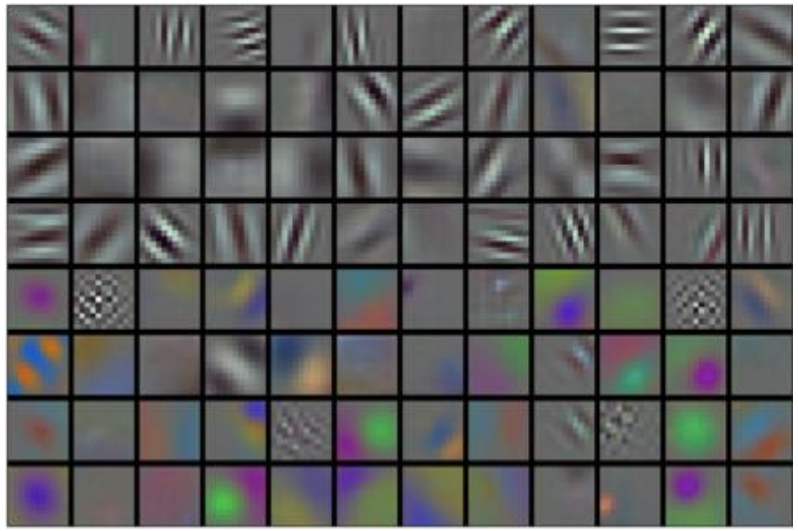
Fixed hand-crafted filters



DETECT features

2D Image Convolution

Learnable neural kernels



LEARN features

Uchida el. al. Image processing and recognition for biological images.
Krizhevsky et. al. ImageNet classification with deep convolutional neural networks.

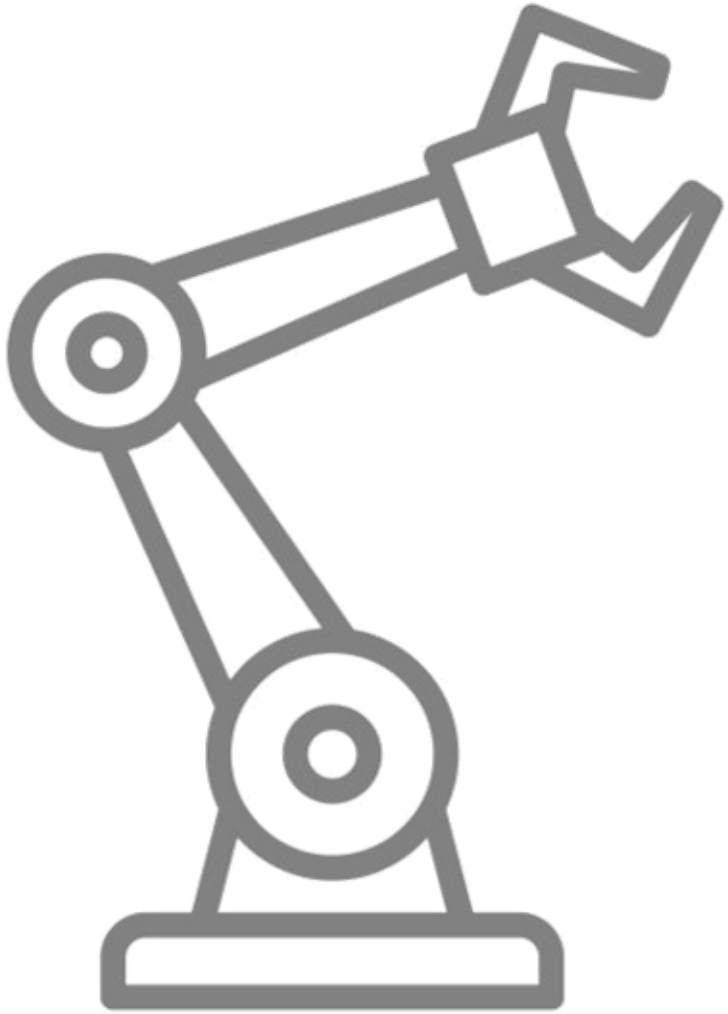
How to Embed Inductive Bias into Networks: CNN

Step 1: Understand structure of modality

Step 2: Identify fundamental operator

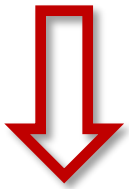
Step 3: Make that operator learnable!

Core Structure of Actions: Articulated Kinematics



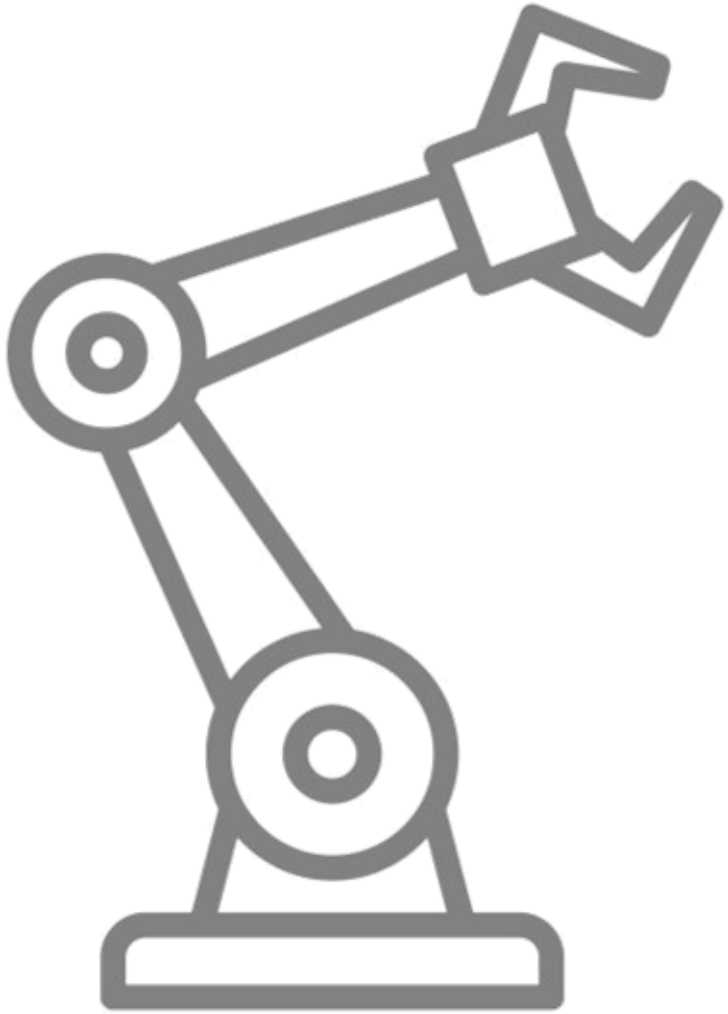
Action: $\boldsymbol{\theta}$

Articulated Kinematics



3D world

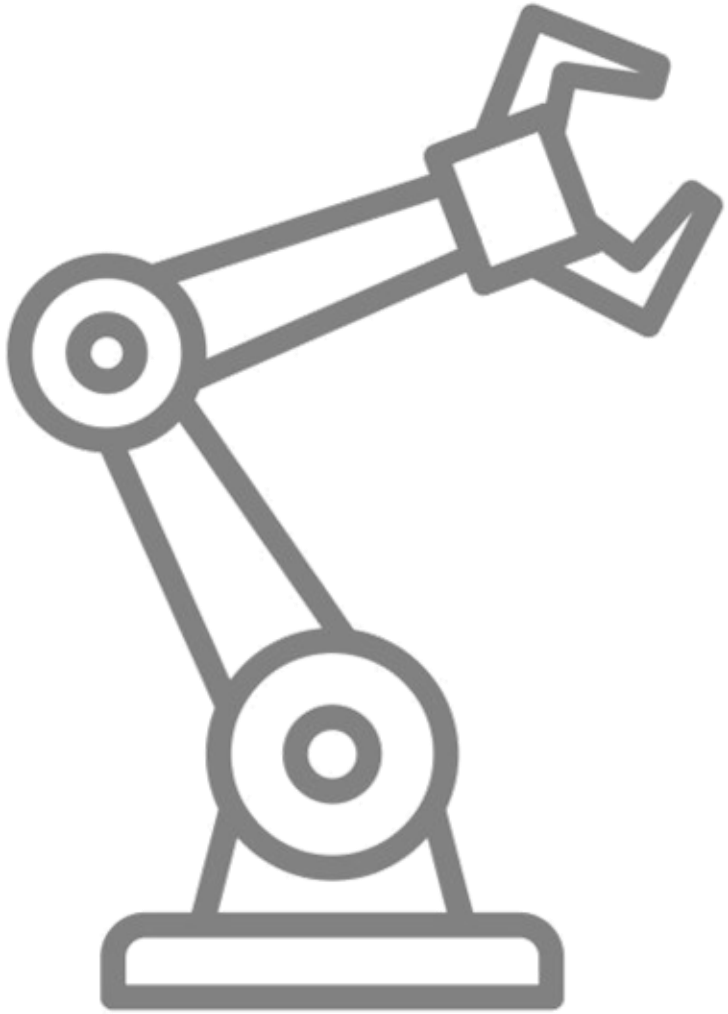
Key Operator: Rodrigues' Rotation Formula



Details omitted
(see paper)

Articulated Kinematics ↷ **Stack of Rodrigues' Rotation Formula**

Key Operator: Rodrigues' Rotation Formula



Articulated Kinematics

Stack of Rodrigues' Rotation Formula

Inductive bias of Network?

Rodrigues' Rotation Formula

$$\mathbf{I}_3 + [\hat{\boldsymbol{\omega}}] \sin\theta + [\hat{\boldsymbol{\omega}}]^2 (1 - \cos\theta)$$

Rodrigues' Rotation Formula

$$\mathbf{I}_3 + [\hat{\boldsymbol{\omega}}] \sin\theta + [\hat{\boldsymbol{\omega}}]^2 (1 - \cos\theta)$$

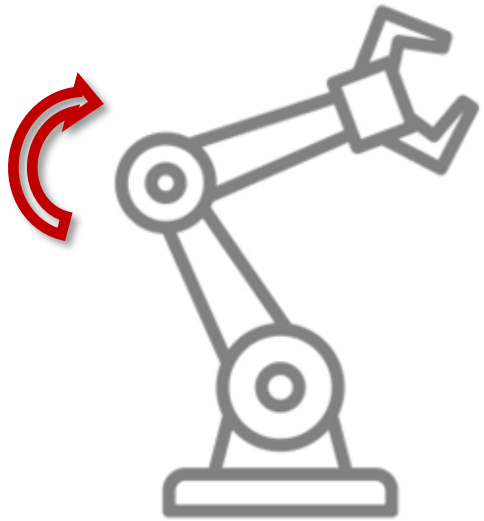
Rodrigues' Rotation Formula

$$\mathbf{I}_3 + [\hat{\boldsymbol{\omega}}] \sin\theta + [\hat{\boldsymbol{\omega}}]^2 (1 - \cos\theta)$$

$$\begin{bmatrix} 0 & -\omega_z & \omega_y \\ \omega_z & 0 & -\omega_x \\ -\omega_y & \omega_x & 0 \end{bmatrix}$$

Rodrigues' Rotation Formula

$$\mathbf{I}_3 + [\hat{\boldsymbol{\omega}}]\sin\theta + [\hat{\boldsymbol{\omega}}]^2(1 - \cos\theta)$$

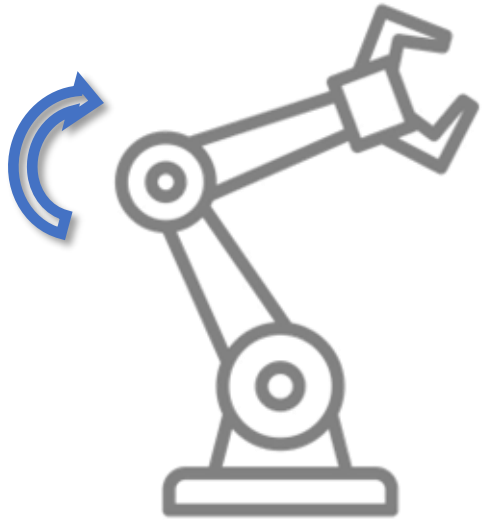


Rodrigues' Rotation Formula

Joint Space $\longmapsto$ World Space

$$\mathbf{I}_3 + [\hat{\boldsymbol{\omega}}] \sin\theta + [\hat{\boldsymbol{\omega}}]^2(1 - \cos\theta)$$

θ



$\mathbf{R}$

Rodrigues' Rotation Formula

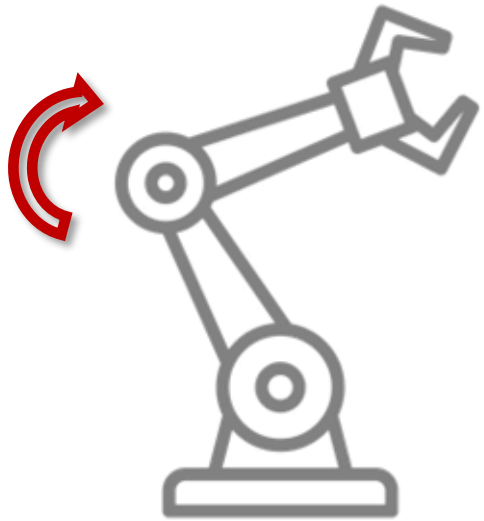
Action-independent

$$\mathbf{I}_3 + [\hat{\boldsymbol{\omega}}] \sin\theta + [\hat{\boldsymbol{\omega}}]^2 (1 - \cos\theta)$$

Rodrigues' Rotation Formula

Action-dependent

$$\mathbf{I}_3 + [\hat{\boldsymbol{\omega}}] \sin\theta + [\hat{\boldsymbol{\omega}}]^2(1 - \cos\theta)$$



Essence of Rodrigues' Rotation Formula

Action-dependent

$$\mathbf{I}_3 + [\hat{\boldsymbol{\omega}}] \sin\theta + [\hat{\boldsymbol{\omega}}]^2(1 - \cos\theta)$$

Re-order

Essence of Rodrigues' Rotation Formula

Action-dependent

$$\mathbf{I}_3 + [\hat{\boldsymbol{\omega}}] \sin\theta + [\hat{\boldsymbol{\omega}}]^2(1 - \cos\theta)$$

↳

Re-order

$$\mathbf{A} + \mathbf{B}\sin\theta + \mathbf{C}\cos\theta$$

Essence of Rodrigues' Rotation Formula

Action-dependent

$$\mathbf{I}_3 + [\hat{\boldsymbol{\omega}}] \sin\theta + [\hat{\boldsymbol{\omega}}]^2(1 - \cos\theta)$$

Re-order

$$\mathbf{A} + \mathbf{B}\sin\theta + \mathbf{C}\cos\theta$$

Essence of Rodrigues' Rotation Formula

Action-independent

$$\mathbf{I}_3 + [\hat{\boldsymbol{\omega}}] \sin\theta + [\hat{\boldsymbol{\omega}}]^2(1 - \cos\theta)$$

Re-order

$$\mathbf{A} + \mathbf{B}\sin\theta + \mathbf{C}\cos\theta$$

Essence of Rodrigues' Rotation Formula

Factorize action-dependent/independent

$$\mathbf{I}_3 + [\hat{\boldsymbol{\omega}}] \sin\theta + [\hat{\boldsymbol{\omega}}]^2(1 - \cos\theta)$$

$$\mathbf{A} + \mathbf{B}\sin\theta + \mathbf{C}\cos\theta$$

Essence of Rodrigues' Rotation Formula

Rodrigues' Rotation Formula

$$\mathbf{A} + \mathbf{B}\sin\theta + \mathbf{C}\cos\theta$$

sin/cos Joint Space

$$\begin{bmatrix} 1 \\ \sin\theta \\ \cos\theta \end{bmatrix}$$

Linear

$$\longmapsto \quad \mathbf{R}$$

World Space

Make Rodrigues' Rotation Formula Learnable

Rodrigues' Rotation Formula

$$\mathbf{A} + \mathbf{B}\sin\theta + \mathbf{C}\cos\theta$$

Make Rodrigues' Rotation Formula Learnable

Rodrigues' Rotation Formula

$$\mathbf{A} + \mathbf{B}\sin\theta + \mathbf{C}\cos\theta$$

Fixed parameters

Make Rodrigues' Rotation Formula Learnable

Rodrigues' Rotation Formula

$$\mathbf{A} + \mathbf{B}\sin\theta + \mathbf{C}\cos\theta$$

Fixed parameters

Learnable kernels

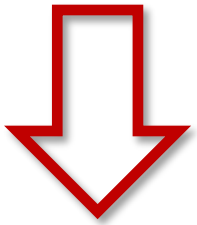
$$\mathbf{W}_{\text{bias}} + \mathbf{W}_{\sin}\sin\theta + \mathbf{W}_{\cos}\cos\theta$$

Make Rodrigues' Rotation Formula Learnable

Rodrigues' Rotation Formula

$$\mathbf{A} + \mathbf{B}\sin\theta + \mathbf{C}\cos\theta$$

Fixed parameters



Neural Rodrigues Operator

Learnable kernels

$$\mathbf{W}_{\text{bias}} + \mathbf{W}_{\sin}\sin\theta + \mathbf{W}_{\cos}\cos\theta$$

Make Rodrigues' Rotation Formula Learnable

Step 1 Structure: Joint Space <-> World Space

Step 2 Operator: Rodrigues' Rotation Formula

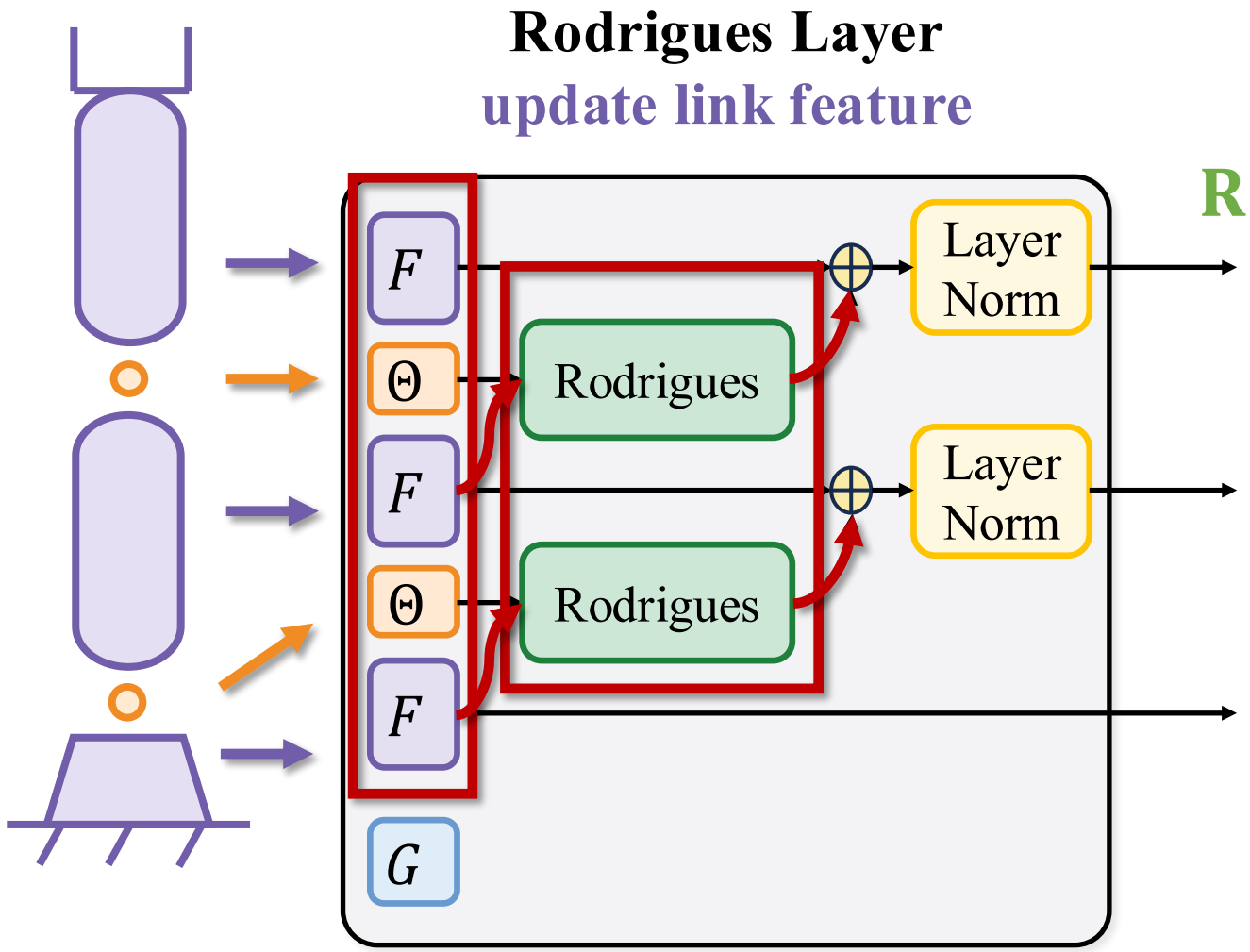
Step 3 Learnable: Neural Rodrigues Operator

Articulated kinematics ⇨ **Learnable Generalization**

(Recall: CNN)

(see paper)

Build Rodrigues Network



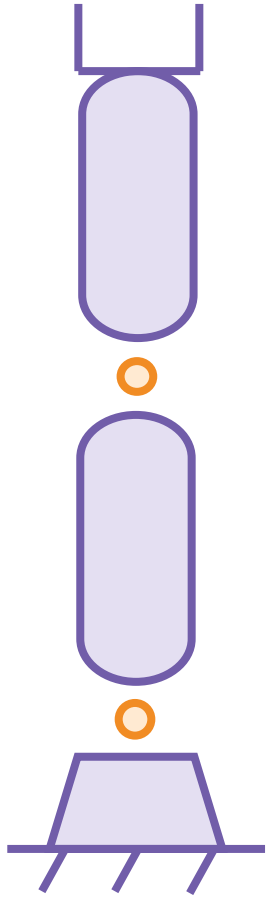
$$\mathbf{R} = \mathbf{W}_{\text{bias}} + \mathbf{W}_{\sin} \sin\theta + \mathbf{W}_{\cos} \cos\theta$$

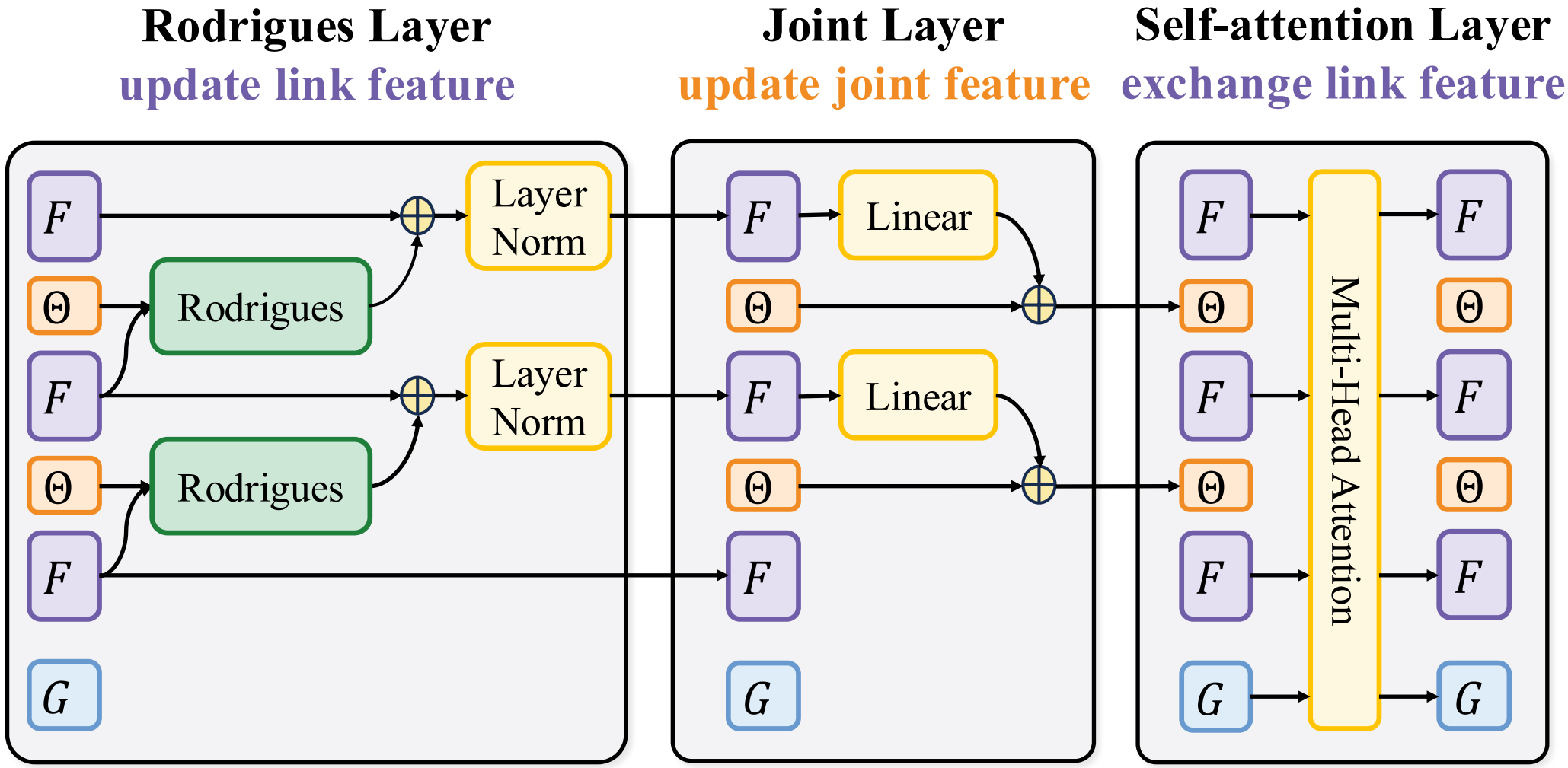
$$\boldsymbol{F}_{\text{child}} = \mathbf{R}\boldsymbol{F}_{\text{parent}}$$

Intuitions omitted
(see paper)

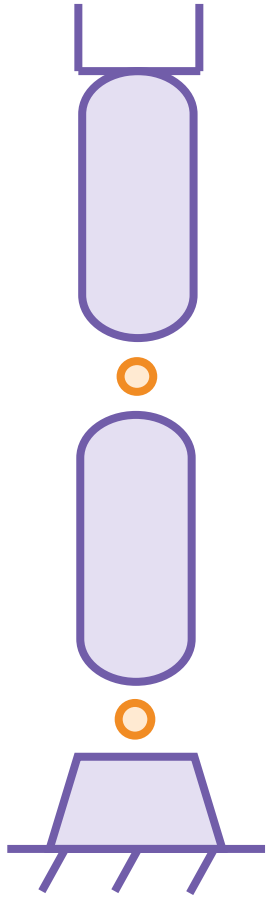
Hierarchical feature passing pattern

Build Rodrigues Network





Build Rodrigues Network



Rodrigues Layer
update link feature

Joint Layer
update joint feature

Self-attention Layer
update joint feature

Details omitted
(see paper)

Rodrigues Block
$(\times N)$

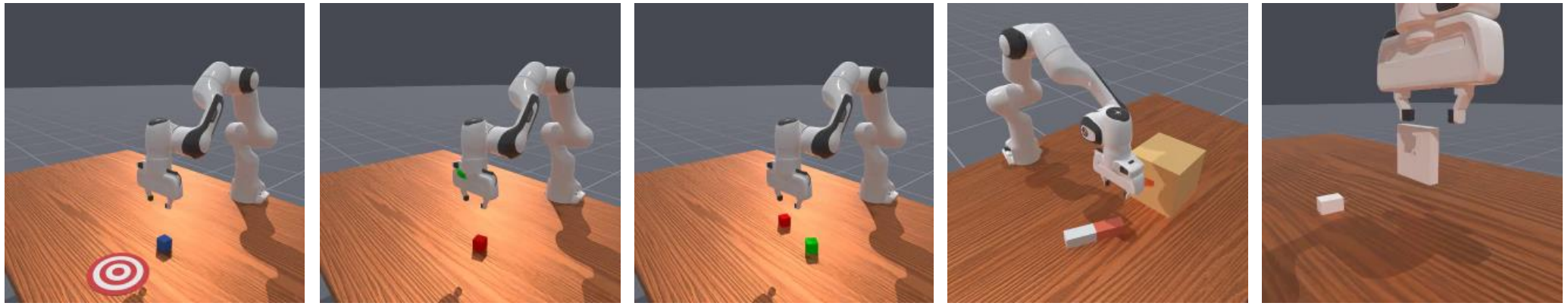
Experiments

Useful in applications?

Application 1: Robot Manipulation with IL

- **Benchmark**: Derived from ManiSkill

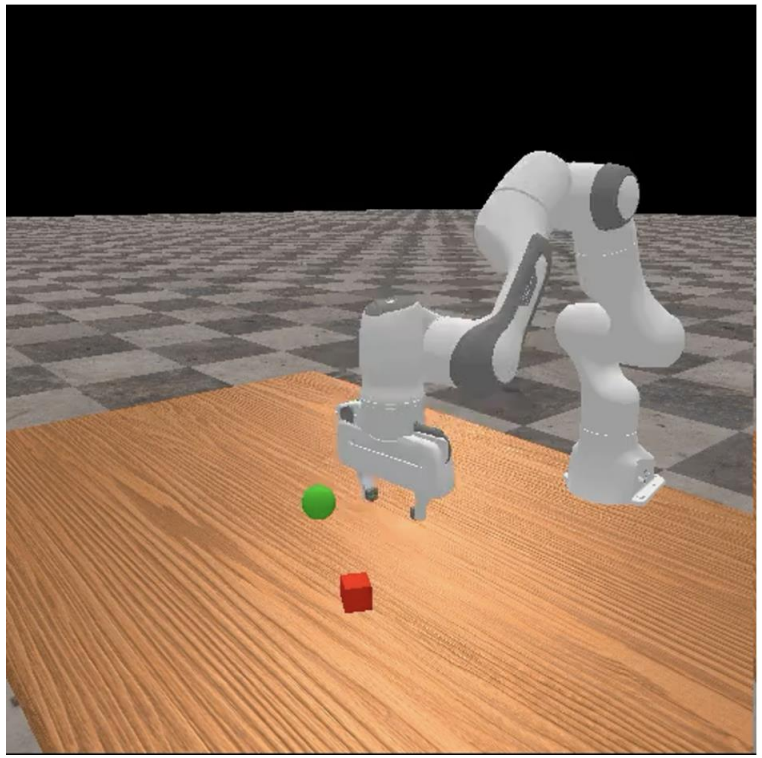
PushCube PickCube StackCube PegInsersionSide PlugCharger



Images: https://maniskill.readthedocs.io/en/latest/tasks/table_top_gripper/index.html

Kinematics-Inspired Inductive Bias in Manip.

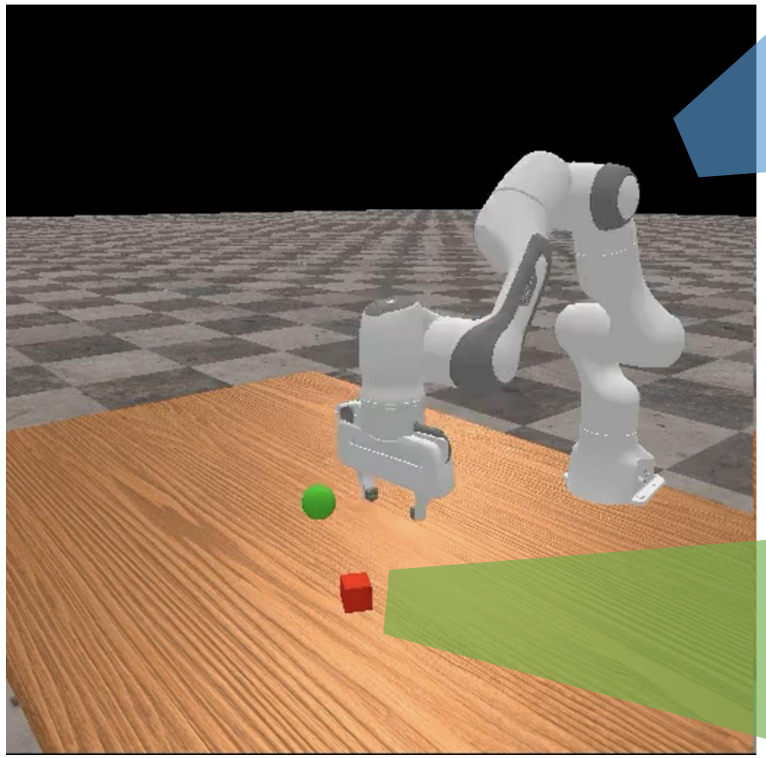
PickCube



Move **box**
to **goal**

Kinematics-Inspired Inductive Bias in Manip.

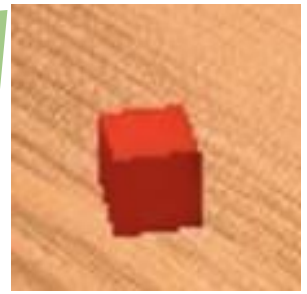
PickCube



$$\begin{bmatrix} \theta_1 \\ \vdots \\ \theta_7 \end{bmatrix}$$

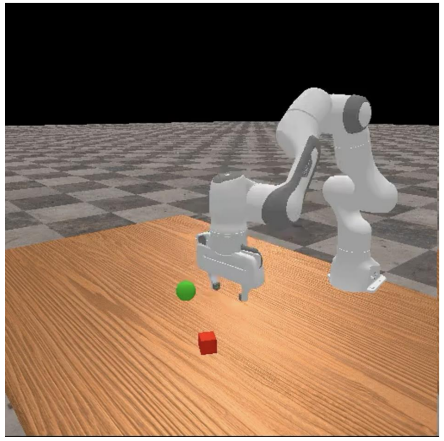
Action Space

Cross-reason



World Space

Add Inductive Bias to Established Algo.: DP



Without inductive bias

[RSS 2023]
Diffusion Policy

World Space ∪ Observation

Action Space ∪ (Noisy) Action

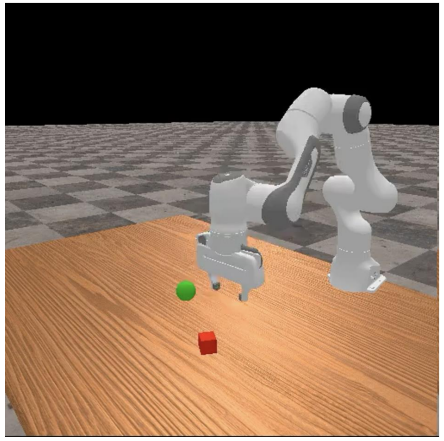
↓

Diffusion Denoising Model
Backbone: **CNN/Transformer**

↓

Predicted Action Noise

Add Inductive Bias to Established Algo.: DP

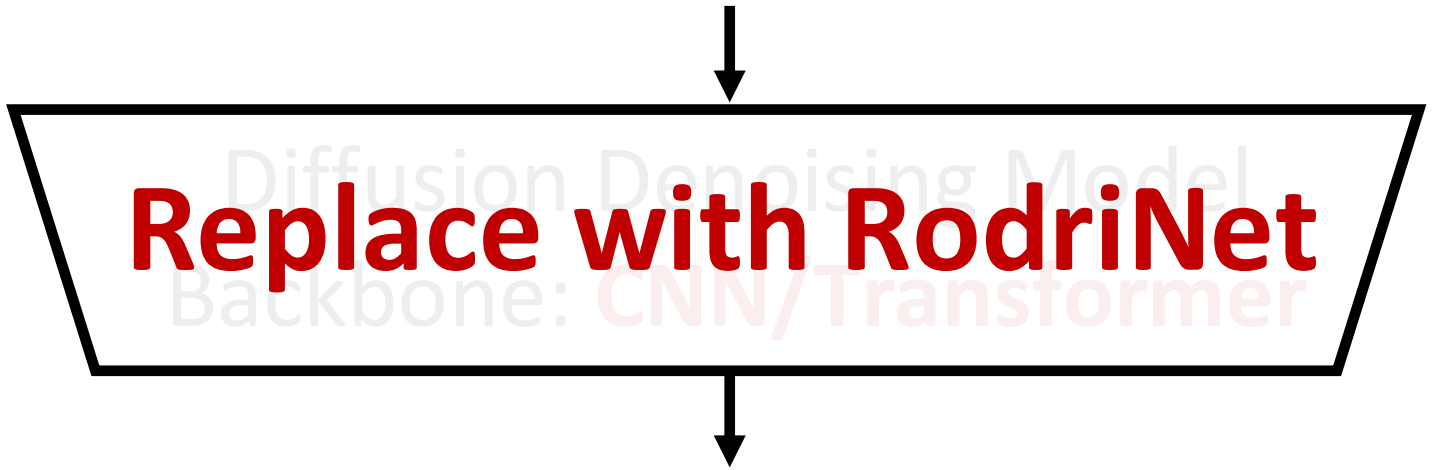


With inductive bias

[RSS 2023]
Diffusion Policy

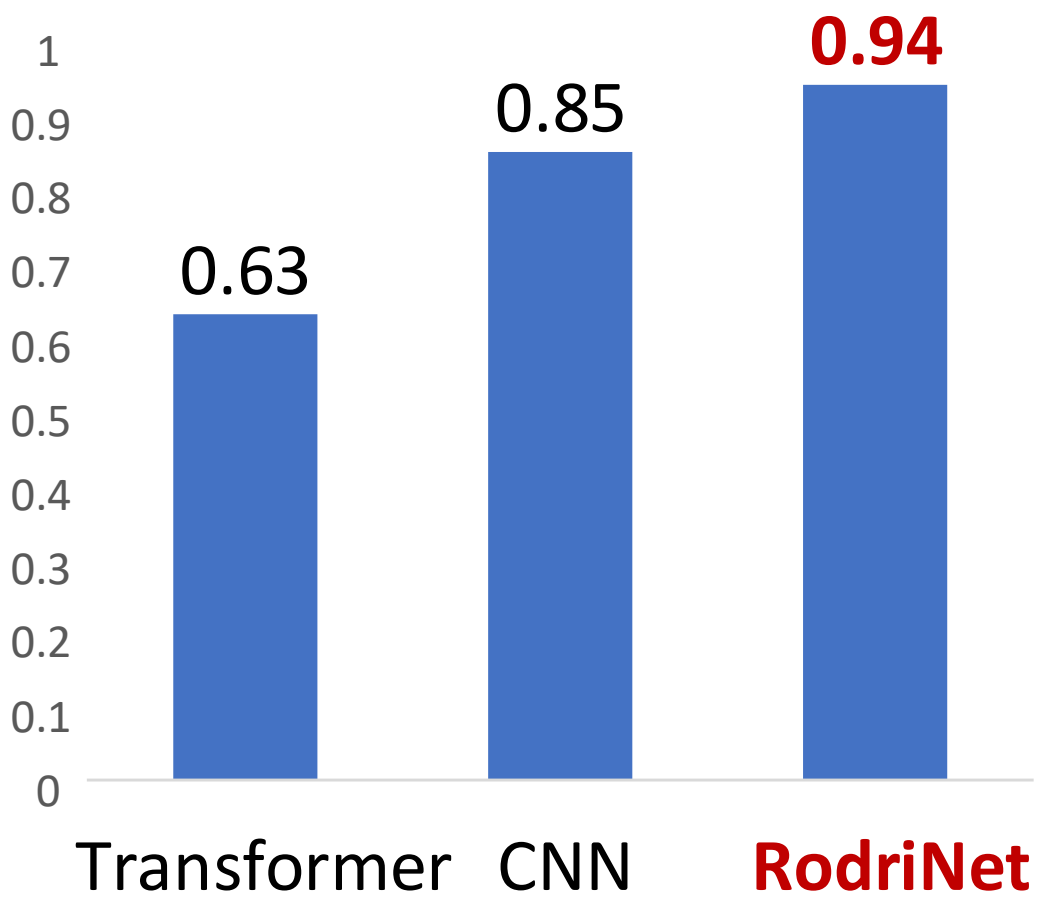
World Space ∪ Observation

Action Space ∪ (Noisy) Action



Predicted Action Noise

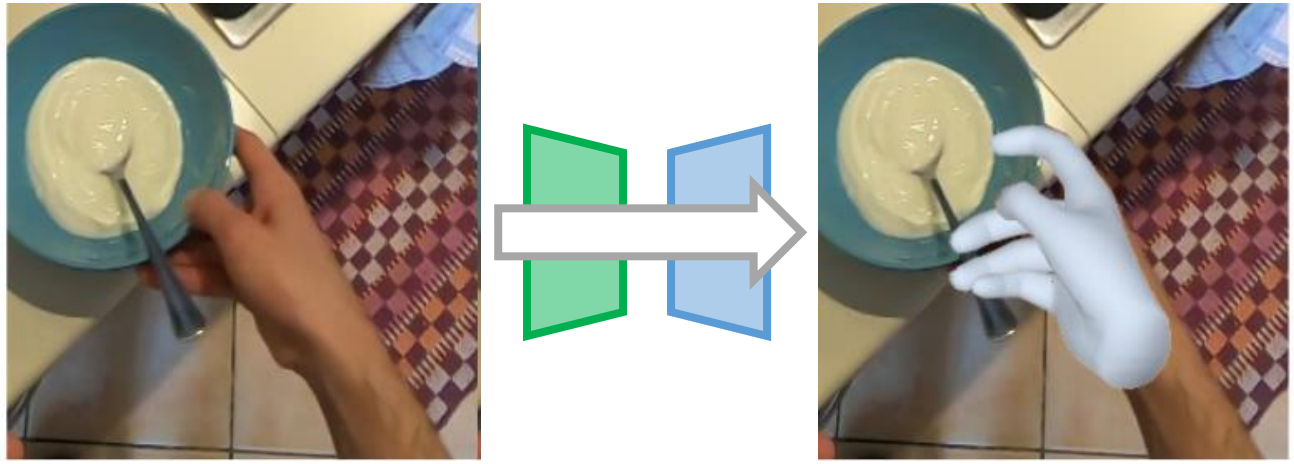
Application 1: Robot Manipulation with IL



Learns better from same demo trajectories

(see more results in paper)

Application 2: 3D Hand Reconstruction



[CVPR 2024] HaMeR, **transformer**-based

Replace hand pose decoder ⇩

Ours, **Rodrigues**-based

(details in paper)

Application 2: 3D Hand Reconstruction

Method	PA-MPJPE ↓	PA-MPVPE ↓	F@5 ↑	F@15 ↑
HaMeR	6.0	5.7	0.785	0.990
HaMeR (Reproduced)	6.2	5.9	0.774	0.989
Ours	**5.9**	**5.6**	**0.793**	**0.991**

Also works on articulated characters

Summary

- **Neural Rodrigues Operator**
  - Kinematics-inspired inductive bias
- **Rodrigues Network**
  - Embodiment-aware architecture
- **Applicable to diverse action-learning tasks**

Limitations / Future Works

- **Joint space controller**
  - Current VLA: gripper + EE space control

- **RL applications**

Rodrigues Network for Learning Robot Actions

Jialiang Zhang[1,2,*], Haoran Geng[3,*], Yang You[2,*], Congyue Deng[1,2],
Pieter Abbeel[3], Jitendra Malik[3], Leonidas Guibas[2, †]

†: corresponding author *: equal contribution

Thanks for Listening!

ICLR
International Conference On Learning Representations

GitHub