

Giovanni Franzese, Daniel Sotelo, Mohammed Osama Elnour, Pavel Kopanev, Abdul Rehman, Mothilal Asokan, Danilo Caporale, Paola Ardón

Autonomous Robotics Research Center, Technology Innovation Institute, Abu Dhabi, UAE

Introduction

How can robots generalize manipulation behavior with minimal data?

- Most methods require large-scale datasets & retraining
- Real-world deployment often needs fast adaptation to new tasks (zero-shot generalization)

Key Idea:

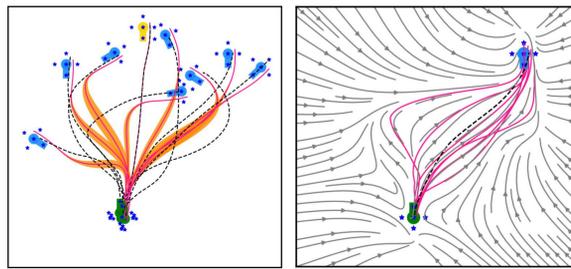
Generalize robot policies across task variations using **sparse keypoints** and learned nonlinear **transportation maps**

We leverage **structural**, **matching** and **transportation** priors

Keypoint-Based Policy Transportation

➤ Generic Policy with Keypoint Context

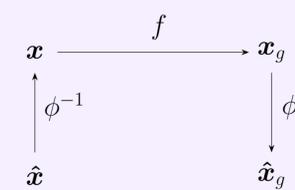
$$\mathbf{x}_g = \mathcal{F} \left(\left[\begin{array}{c} \mathbf{x} \\ \mathcal{T} \end{array} \right] \middle| \left[\begin{array}{c} \mathcal{X} \\ \mathcal{S} \end{array} \right] \right)$$



➤ Transportation Map $\phi_{\mathcal{S} \rightarrow \mathcal{T}} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ s.t. $\mathcal{T} = \phi(\mathcal{S})$

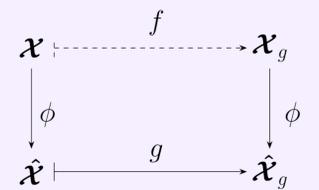
Online Policy Transportation

$$\mathbf{x}_g = \phi(f(\phi^{-1}(\mathbf{x})|\mathcal{X}))$$



Offline Policy Transportation

$$\mathbf{x}_g = g(\mathbf{x}|\phi(\mathcal{X}))$$



Robot Experiments

Transportation of Push-T Policy

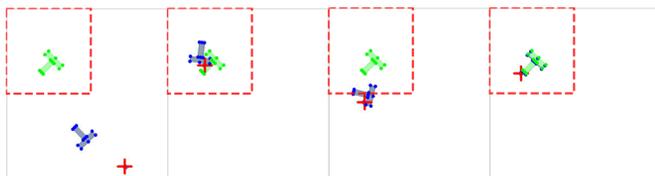
➤ **Task:** Pushing a T-shaped object

Used **25 keypoints** and **35 demos**

13,685 chunks (augmented to 53,731 via **interactive demos**)

Method	In distribution	Out of distribution	
		Robot	Object
Nearest Neighborhood	85.0%	✗	✗
Policy Transportation	99.5%	85.5%	35.5%
Conditional Flow Matching	97.0%	88.0%	40.0%

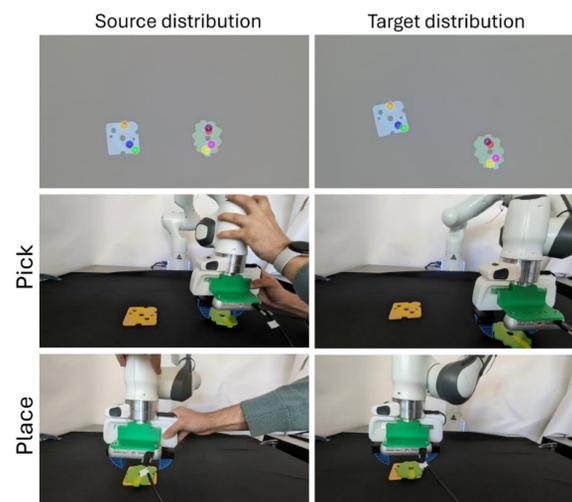
➤ **Out-of-Distribution (OOD) Generalization:**



Pick-and-Place with ViT-based Keypoint Extraction

✓ Robust generalization across 10 configurations with same order

✗ Fails when swapping object positions (lettuce / cheese)

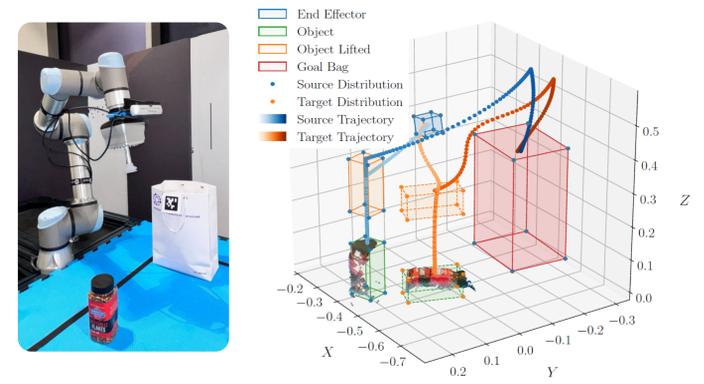


Policy Transportation of an Object Packing Policy

➤ **Task:** Place object inside the bag with suction gripper

Used **32 keypoints** and **1 demo**

- Fixed bag + Object relocation: **15/15**
- **Full Scene Variation:** Bag + Object relocation and type: **11/15**



Limitations and Opportunities

✓ **Compact, interpretable, and generalizable** policies

✗ Requires **accurate keypoint matching** between source and target

✓ Zero-shot transfer **without policy retraining**

✗ The **inductive bias** introduced by keypoints limits generalization

✓ Policy generalization even from a **single demonstration**

✗ Sensitive to **symmetric** or **OOD scenarios**

Future work: automatic relevant keypoint selection, explore novel matching strategies, formally assess OOD generalization

References

- [1] G. Franzese, R. Prakash, C. Della Santina, and J. Kober. Generalizable Motion Policies through Keypoint Parameterization and Transportation Maps. *IEEE Transactions on Robotics*, 41: 4557-4573, 2025.
- [2] L. Manuelli, W. Gao, P. Florence, and R. Tedrake. KPAM: KeyPoint Affordances for Category Level Robotic Manipulation. In *Proceedings of the 19th International Symposium on Robotics Research (ISRR)*, pages 132-157, 2022.
- [3] S. Amir, Y. Gandelman, S. Bagon, and T. Dekel. Deep ViT Features as Dense Visual Descriptors. In *Proceedings of the 2022 European Conference on Computer Vision (ECCV) Workshop on What is Motion For?*, 2022.
- [4] C. Chi, Z. Xu, S. Feng, E. Cousineau, Y. Du, B. Burchfiel, R. Tedrake, and S. Song. Diffusion Policy: Visuomotor Policy Learning via Action Diffusion. *The International Journal of Robotics Research*, 2024.
- [5] E. Chisari, N. Heppert, M. Argus, T. Welschehold, T. Brox, and A. Valada. Learning Robotic Manipulation Policies from Point Clouds with Conditional Flow Matching. In *Proceedings of the 8th Annual Conference on Robot Learning (CoRL)*, 2024.

