Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

MIN Faculty
Department of Informatics

Learning Perception and Manipulation of Clothes

Research Survey

Niklas Fiedler

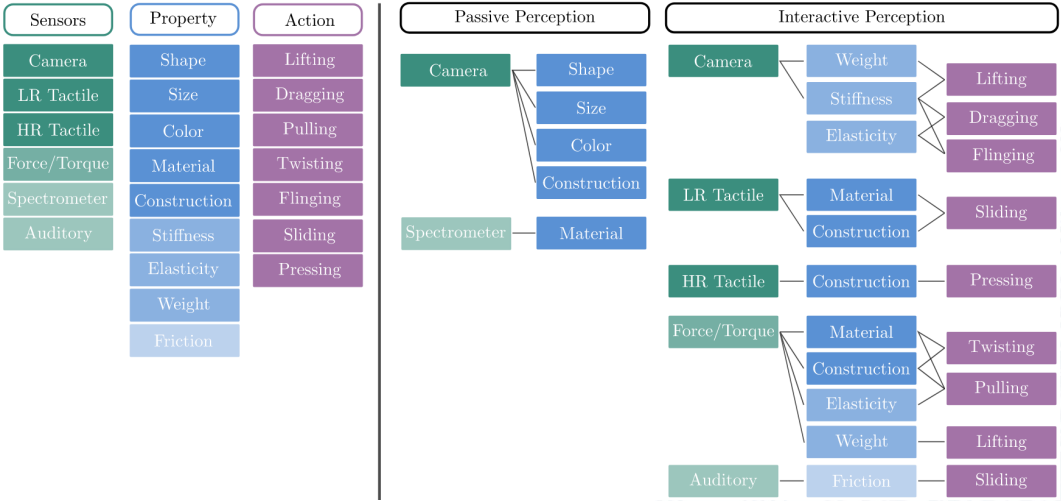
University of Hamburg
Faculty of Mathematics, Informatics and Natural Sciences
Department of Informatics
Technical Aspects of Multimodal Systems

January 28, 2025

Outline

1. Clothes/Fabric Specific Work
2. Diffusion Policy Advancements
3. My Future Plans

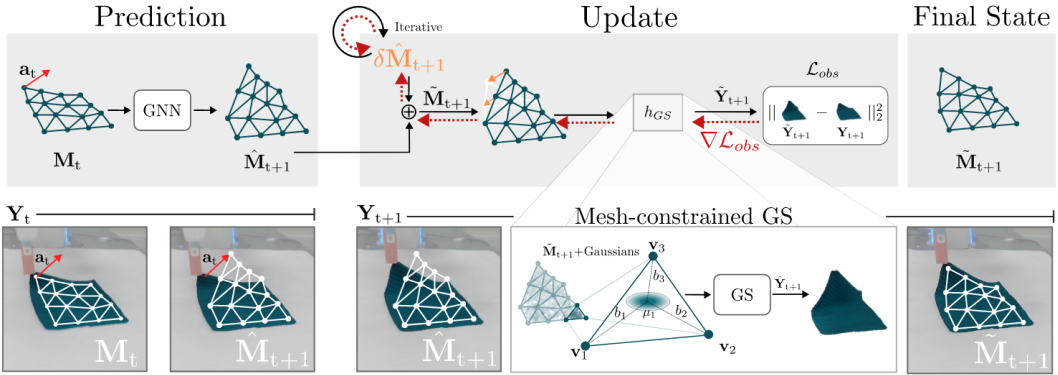
Clothes Perception



Unfolding the Literature: A Review of Robotic Cloth Manipulation, Longhini et al.,

Annual Review of Control, Robotics, and Autonomous Systems 2025

Cloth Splatting



Cloth-Splatting: 3D Cloth State Estimation from RGB Supervision, Longhini et al., CoRL 2024

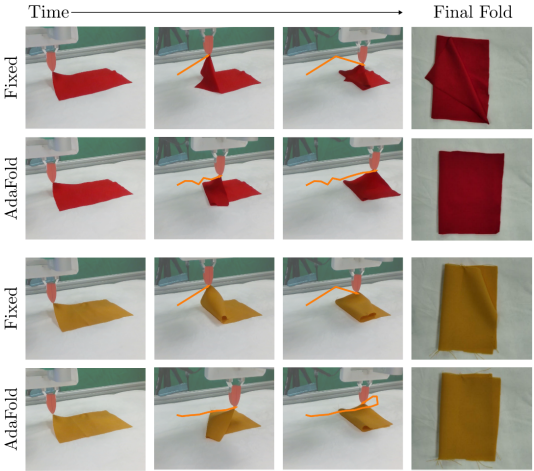
Tasks

Frequency	Household	Healthcare	Textile Industry
Frequent (4+)	Folding, Smoothing, Ironing	Dressing	
Rare (2-3)	Hanging, Sorting, Wiping	Bedding, Bed-making, Bandaging	Recycling
Unaddressed (0-1)	Storing	Buttoning	Manufacturing, Dyeing, Quality control, Coloring, Washing

Unfolding the Literature: A Review of Robotic Cloth Manipulation, Longhini et al.,

Annual Review of Control, Robotics, and Autonomous Systems 2025

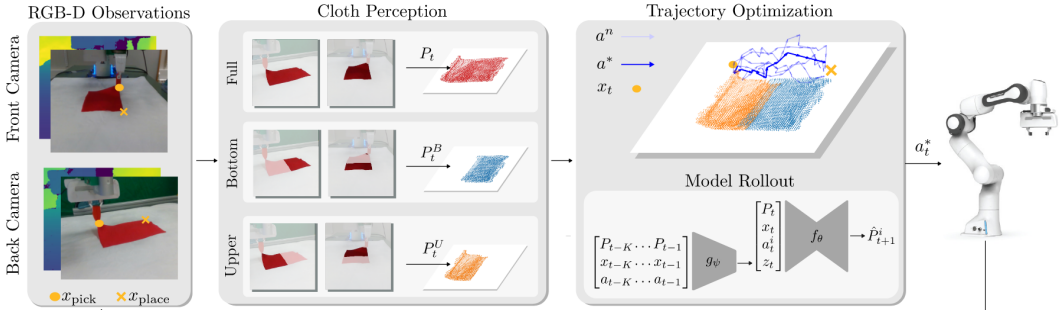
Ada Fold



AdaFold: Adapting Folding Trajectories of Cloths via Feedback-loop Manipulation, Longhini et al.,

Robotics and Automation Letters 2024

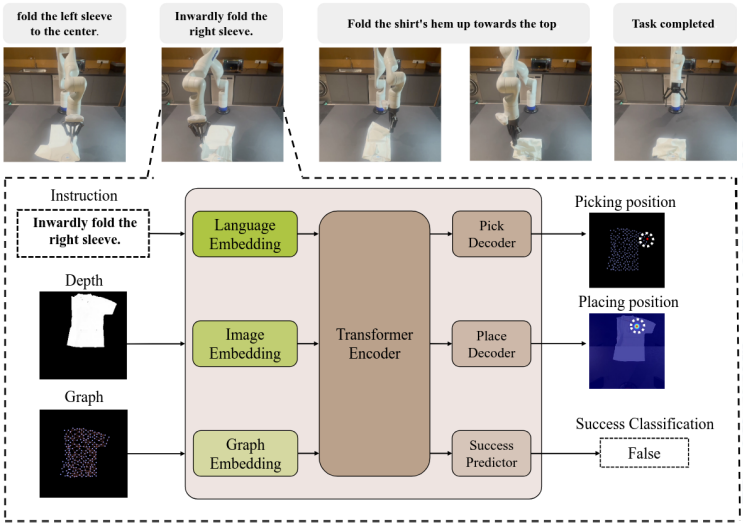
Ada Fold



AdaFold: Adapting Folding Trajectories of Cloths via Feedback-loop Manipulation, Longhini et al.,

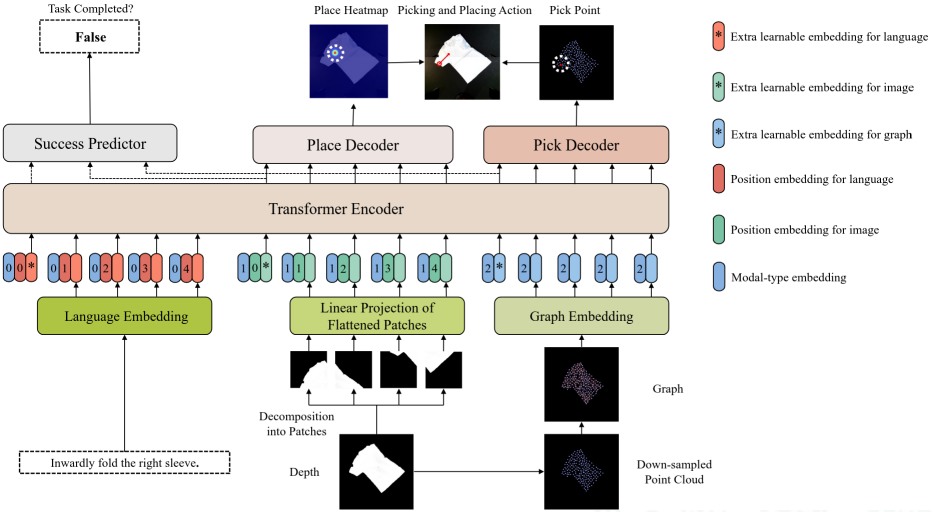
Robotics and Automation Letters 2024

Language-Conditioned Deformable Object Manipulation



Learning Language-Conditioned Deformable Object Manipulation with Graph Dynamics, Deng et al., ICRA 2024

Language-Conditioned Deformable Object Manipulation



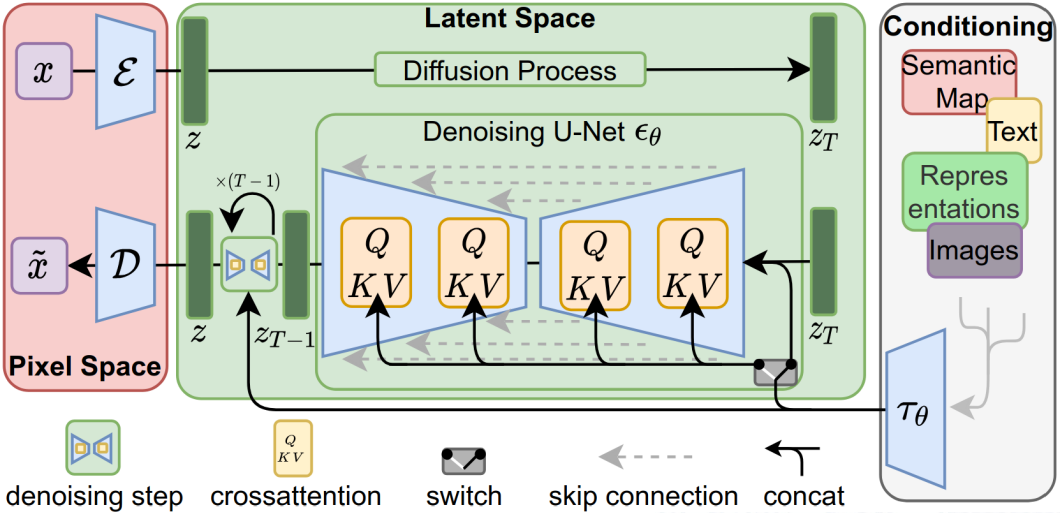
Learning Language-Conditioned Deformable Object Manipulation with Graph Dynamics, Deng et al., ICRA 2024

Diffusion Fundamentals

- Originally Text-to-Image generative models
- Alternative to Generative Adversarial Networks (GANs)
  - Several useful advancements
  - Higher image resolution
  - More stability during training

https://insights.daffodilsw.com/blog/all-you-need-to-know-about-diffusion-models

Diffusion Fundamentals



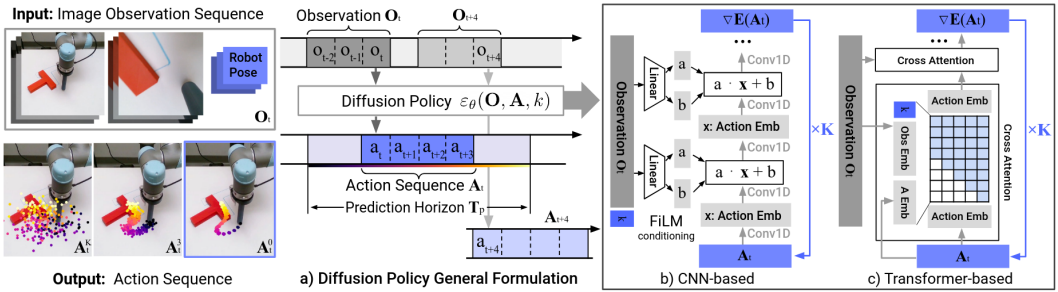
High-

Resolution Image Synthesis with Latent Diffusion Models, Rombach et al., CVPR 2022

Diffusion Policy

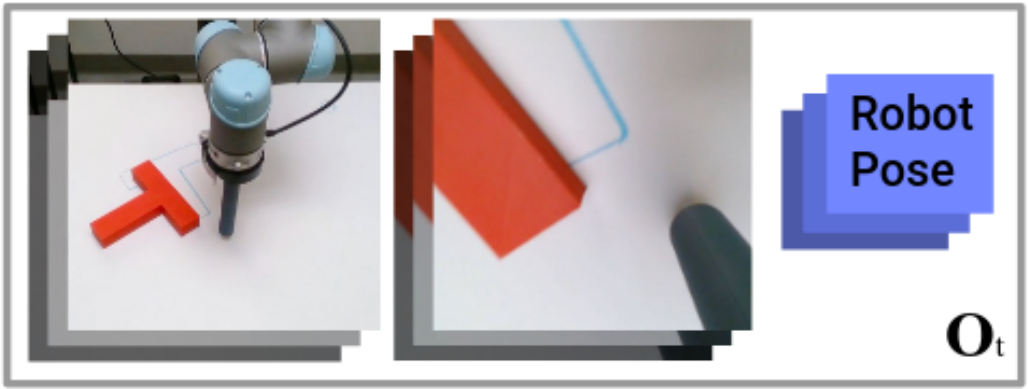
- Basic idea: Generate robot actions instead of images
- Use system state encoding as denoising conditions
- Use demonstrations as training data

Diffusion Policy



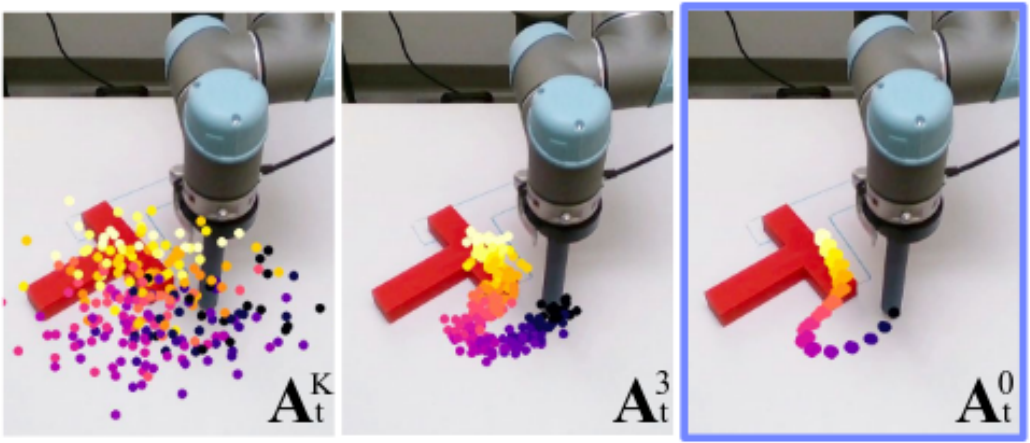
Diffusion Policy: Visuomotor Policy Learning via Action Diffusion, Chi et al., The International Journal of Robotics Research 2024

Diffusion Policy Input



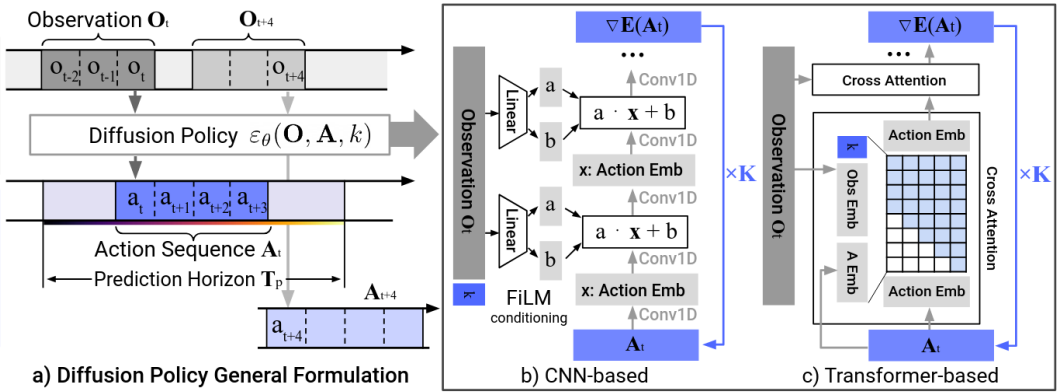
Diffusion Policy: Visuomotor Policy Learning via Action Diffusion, Chi et al., The International Journal of Robotics Research 2024

Diffusion Policy Output



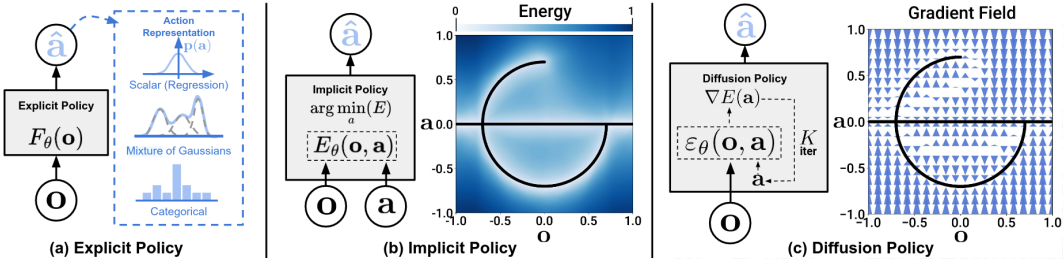
Diffusion Policy: Visuomotor Policy Learning via Action Diffusion, Chi et al., The International Journal of Robotics Research 2024

Diffusion Policy Architecture



Diffusion Policy: Visuomotor Policy Learning via Action Diffusion, Chi et al., The International Journal of Robotics Research 2024

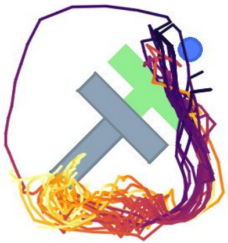
Policy Representations



Diffusion Policy: Visuomotor Policy Learning via Action Diffusion, Chi et al., The International Journal of Robotics Research 2024

Multimodal Behavior

Diffusion Policy



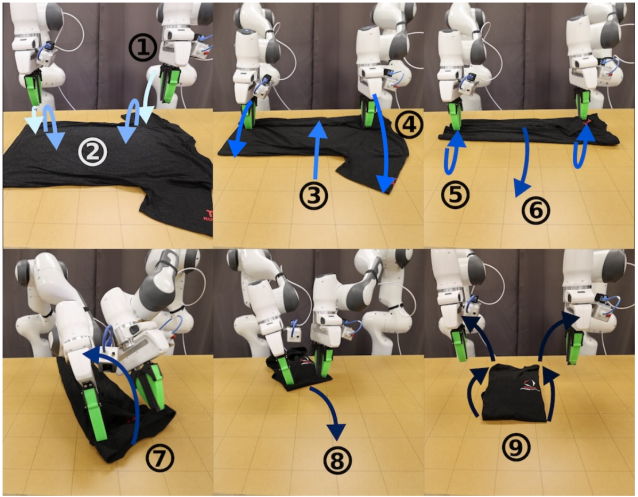
LSTM-GMM

BET

IBC

Diffusion Policy: Visuomotor Policy Learning via Action Diffusion, Chi et al., The International Journal of Robotics Research 2024

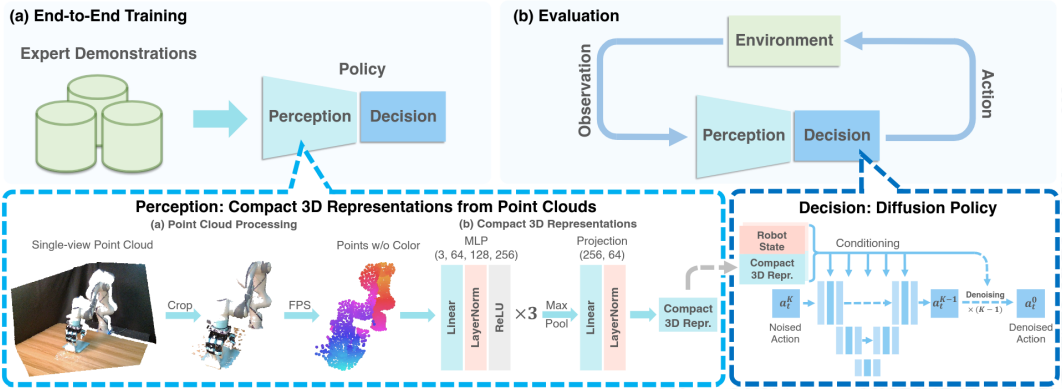
Diffusion Policy for Clothes



- Folding one specific T-shirt (same as in demonstration)
- 284 demonstrations for training
- Success rate of 75% over 20 trials
- Demonstration and policy rollout using VR controllers and collision avoidance (similar to TAMS setup)

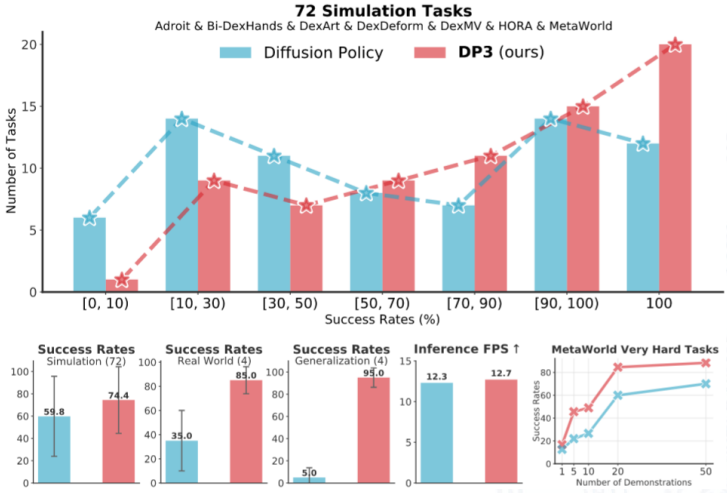
Diffusion Policy: Visuomotor Policy Learning via Action Diffusion, Chi et al., The International Journal of Robotics Research 2024

3D Diffusion Policy



3D Diffusion Policy: Generalizable Visuomotor Policy Learning via Simple 3D Representations, Ze et al., Robotics: Science and Systems 2024

3D Diffusion Policy



3D Diffusion Policy: Generalizable Visuomotor Policy Learning via Simple 3D Representations, Ze et al., Robotics: Science and Systems 2024

3D Diffusion Policy

Encoders	Conv	w/ T-Net	w/ BN	1024 Dim	**Average**
PointNet	✓	✓	✓	✓	15.7
	✗	✓	✓	✓	15.7
	✓	✗	✓	✓	16.0
	✗	✗	✓	✓	26.0
	✗	✓	✓	✗	18.2
Turnaroud!	✓	✗	✗	✓	**72.5**
	✗	✗	✓	✗	19.8
	✗	✓	✗	✗	26.8
	✗	✗	✗	✗	**72.3**

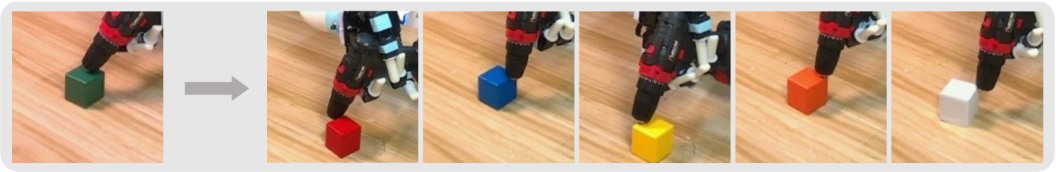
3D Diffusion Policy: Generalizable Visuomotor Policy Learning via Simple 3D Representations, Ze et al., Robotics: Science and Systems 2024

3D Diffusion Policy

Designs	H	D	P	A	DA	SP	**Average**
DP3	**100±0**	62±4	43±6	**99±1**	69±4	97±4	**78.3**
w/o cropping	98±1	**69±3**	14±1	19±9	32±6	40±2	45.3
w/o LayerNorm	**100±0**	56±4	44±3	96±2	51±3	91±5	73.0
w/o sample pred	68±3	67±8	37±12	96±2	58±9	76±9	67.0
w/o projection	**100±0**	61±2	**47±3**	**99±1**	60±8	**99±2**	77.7
w/ color	**100±1**	67±3	46±4	76±8	**75±5**	68±3	72.0
DDIM→DPM-solver++	12±4	9±5	26±5	93±3	58±6	92±14	48.3

3D Diffusion Policy: Generalizable Visuomotor Policy Learning via Simple 3D Representations, Ze et al., Robotics: Science and Systems 2024

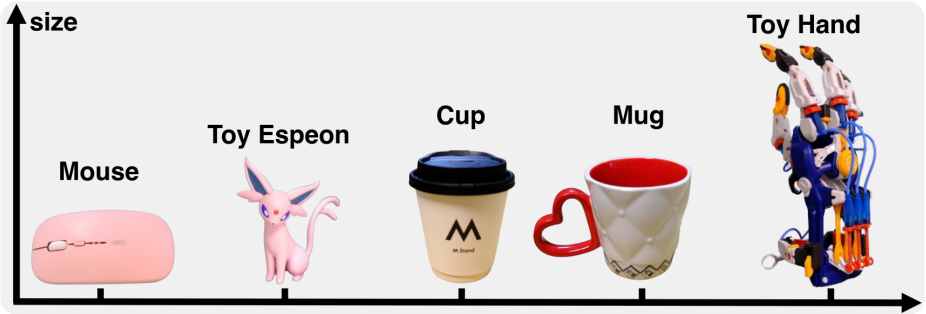
3D Diffusion Policy Generalization



Apperance Generalization (■)	■	■	■	■	■
Diffusion Policy	✗	✗	✗	✗	✗
Diffusion Policy (Depth)	✗	✗	✗	✗	✗
DP3	✓	✓	✓	✓	✓

3D Diffusion Policy: Generalizable Visuomotor Policy Learning via Simple 3D Representations, Ze et al., Robotics: Science and Systems 2024

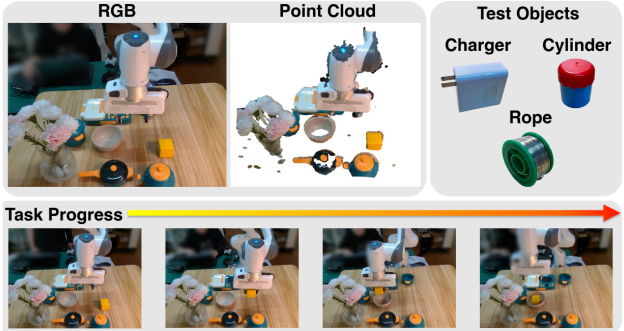
3D Diffusion Policy Generalization



Instance Generalization	Mouse	Espeon	Cup	Mug	Hand
Diffusion Policy	✗	✗	✗	✗	✓
Diffusion Policy (Depth)	✗	✗	✓	✗	✗
DP3	✓	✓	✓	✓	✓

3D Diffusion Policy: Generalizable Visuomotor Policy Learning via Simple 3D Representations, Ze et al., Robotics: Science and Systems 2024

3D Diffusion Policy Generalization

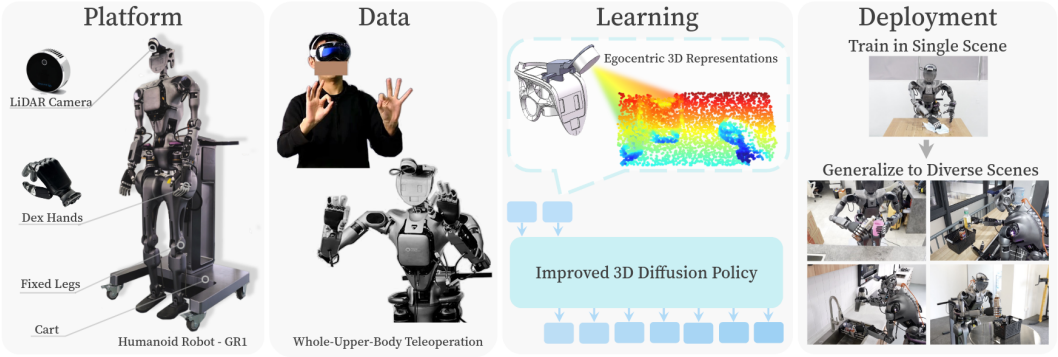


Cluttered Scenes	Diffusion Policy	DP3 w/ PointNeXt	DP3 w/ color	DP3
Success Rate	60	0	**80**	**80**

Train with ■ Cube	■	■	■	Charger	Cylinder	Rope
DP3 w/ color	✗	✗	✗	✗	✗	✗
DP3	✓	✓	✓	✓	✓	✓

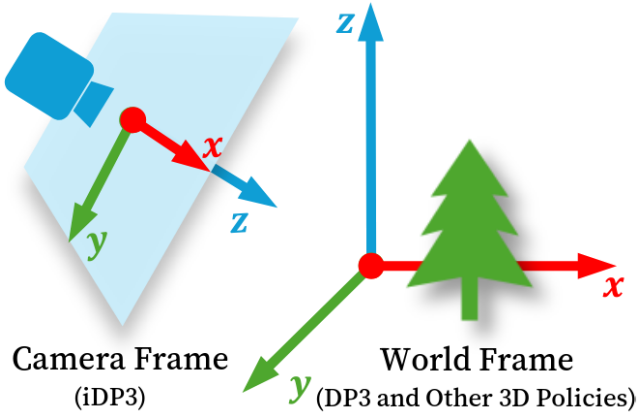
3D Diffusion Policy: Generalizable Visuomotor Policy Learning via Simple 3D Representations, Ze et al., Robotics: Science and Systems 2024

Mobile Robot 3D Diffusion Policy



Generalizable Humanoid Manipulation with Improved 3D Diffusion Policies, Ze et al., arXiv 2024

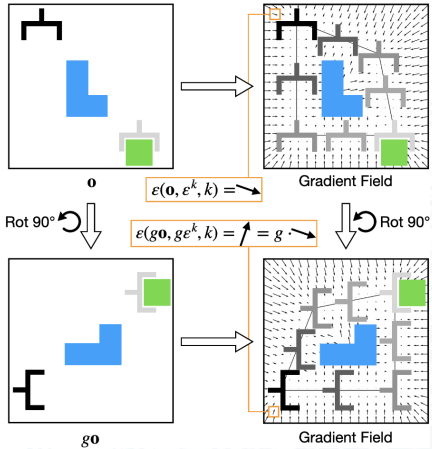
Mobile Robot 3D Diffusion Policy



Generalizable Humanoid Manipulation with Improved 3D Diffusion Policies, Ze et al., arXiv 2024

Equivariant Diffusion Policy

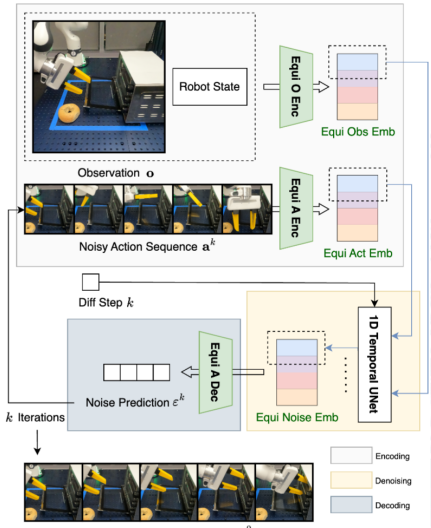
- Utilize domain symmetries
- Increases sample efficiency and generalization



Equivariant Diffusion Policy, Wang et al., CoRL 2024

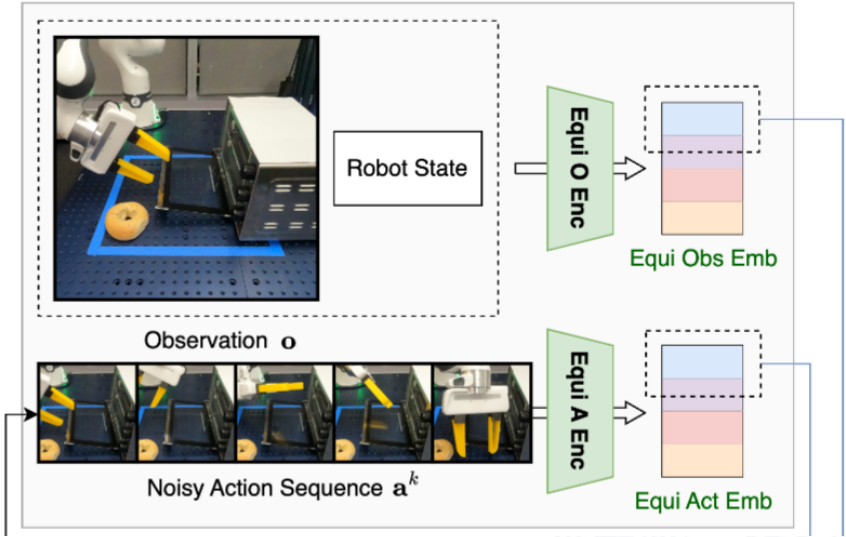
Equivariant Diffusion Policy

- Utilize domain symmetries
- Increases sample efficiency and generalization



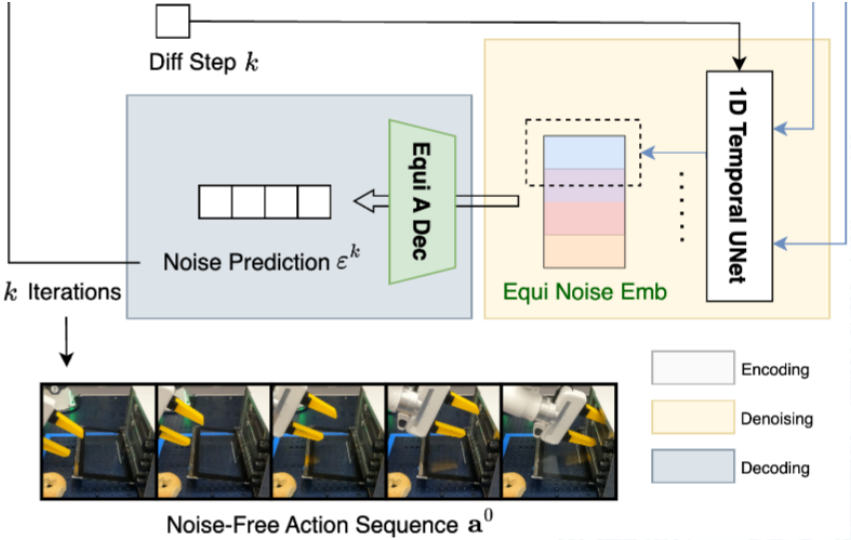
Equivariant Diffusion Policy, Wang et al., CoRL 2024

Equivariant Diffusion Policy



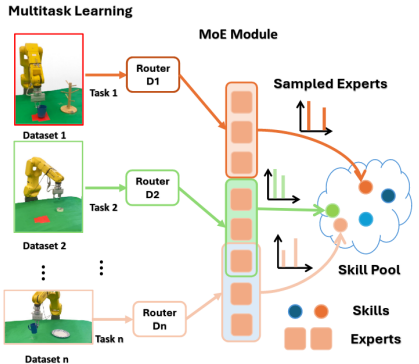
Equivariant Diffusion Policy, Wang et al., CoRL 2024

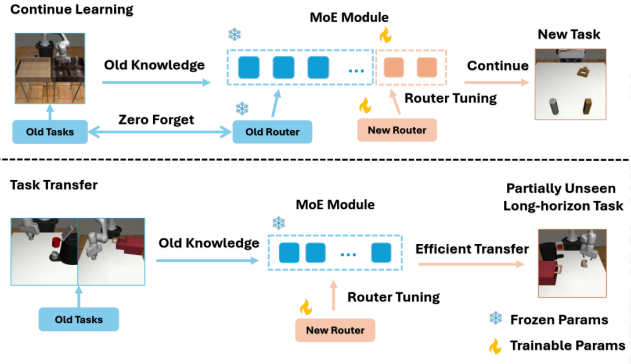
Equivariant Diffusion Policy



Equivariant Diffusion Policy, Wang et al., CoRL 2024

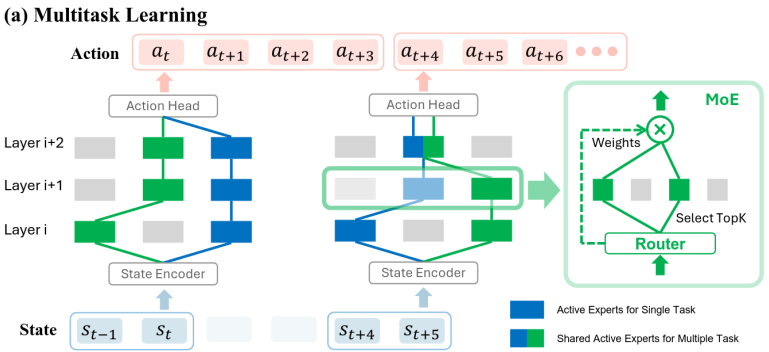
Sparse Diffusion Policy





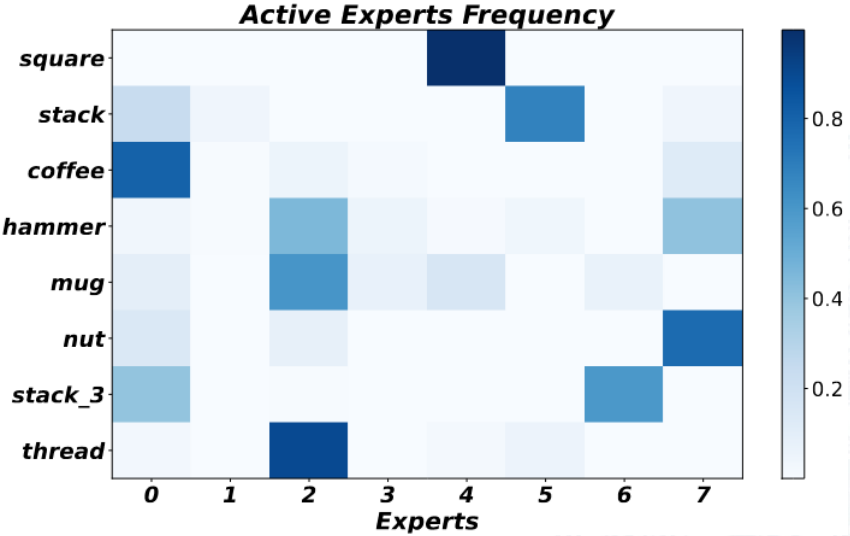
Sparse Diffusion Policy: A Sparse, Reusable, and Flexible Policy for Robot Learning, Wang et al., CoRL 2024

Sparse Diffusion Policy



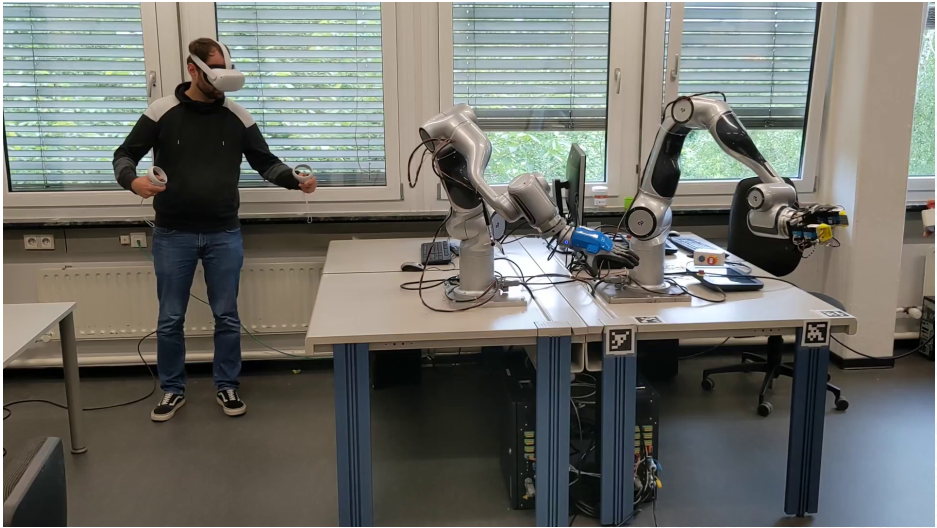
Sparse Diffusion Policy: A Sparse, Reusable, and Flexible Policy for Robot Learning, Wang et al., CoRL 2024

Sparse Diffusion Policy

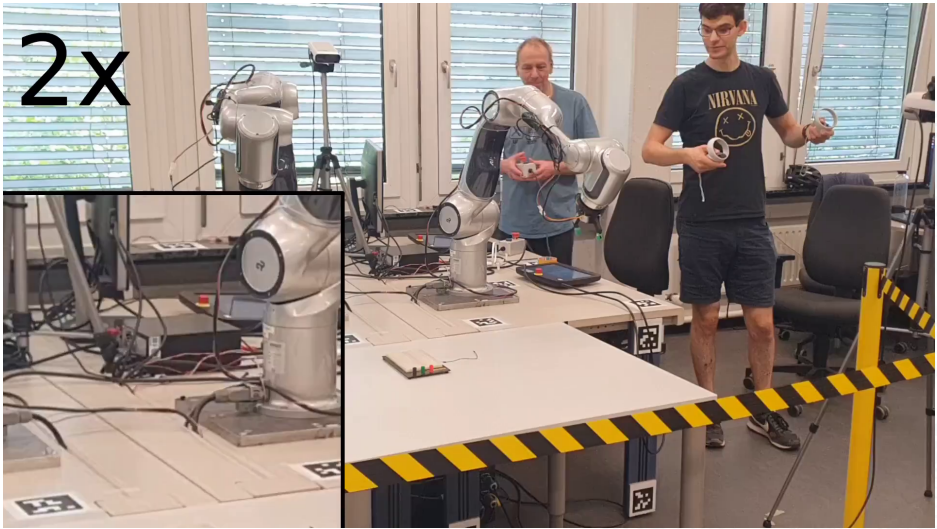


Sparse Diffusion Policy: A Sparse, Reusable, and Flexible Policy for Robot Learning, Wang et al., CoRL 2024

Demonstrations in VR



Demonstrations using VR Controllers



Future Tasks

Frequency	Household	Healthcare	Textile Industry
Frequent (4+)	Folding, Smoothing, Ironing	Dressing	
Rare (2-3)	Hanging, Sorting, Wiping	Bedding, Bed-making, Bandaging	Recycling
Unaddressed (0-1)	Storing	Buttoning	Manufacturing, Dyeing, Quality control, Coloring, Washing

Unfolding the Literature: A Review of Robotic Cloth Manipulation, Longhini et al.,

Annual Review of Control, Robotics, and Autonomous Systems 2025

Future Plans

- **Get some diffusion policy running on the dual Diana7 setup** and build clothes demo
  - Initially with image input
  - Try tasks with point cloud input
  - Figure out how to use point clouds with clothes/fabrics
- Explore rarely addressed and unaddressed tasks
  - Quality control using event cameras?