ICML
International Conference On Machine Learning

2026
POSTER

MSP: Probabilistically Consistent Multi-Scale Action Generation

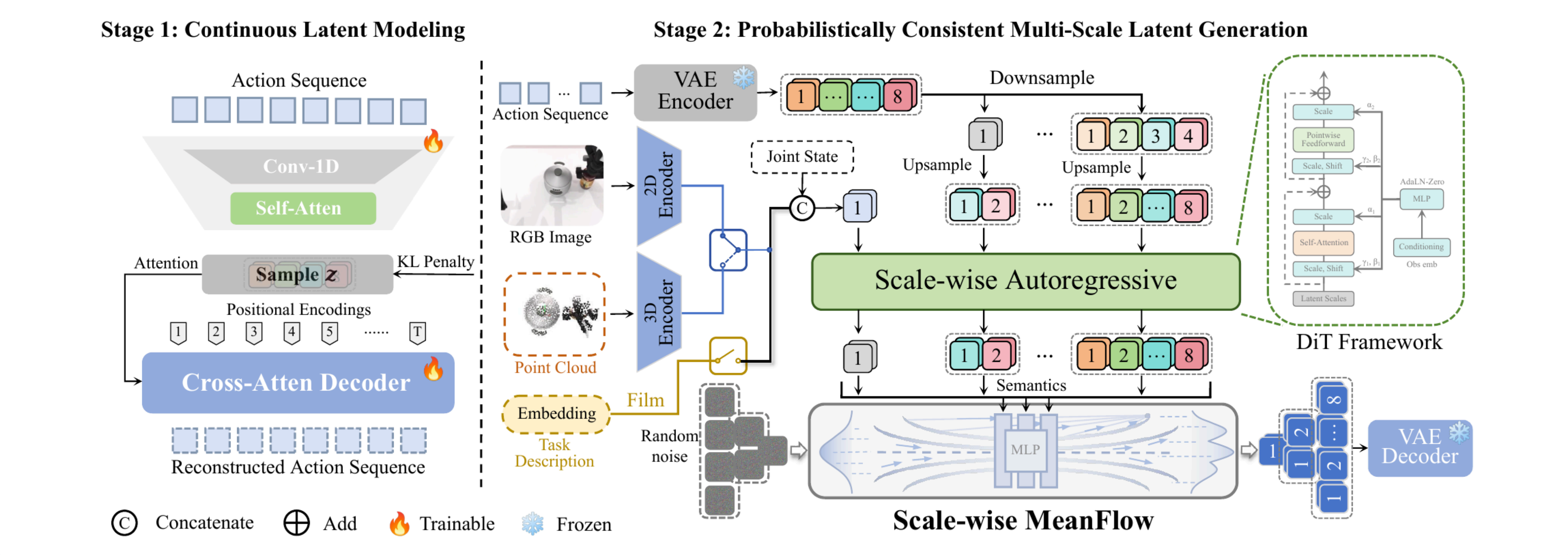
Continuous latent multi-scale action generation for robust long-horizon robot imitation learning.

Zhixuan Lin, Gengqi Liu, Chao Zheng, Gao Lin, Jindong Yu, Song Gao, Fei Wang✉
Faculty of Robot Science and Engineering, Northeastern University, Shenyang, China

Northeastern University

Project Page

1 Core Mechanism: Two-Stage MSP



Stage 1 learns a smooth continuous latent action space. **Stage 2** generates a coarse-to-fine latent hierarchy using scale-wise autoregressive semantics and 1-NFE MeanFlow.

96.9%	+6.8%	60%	32
Avg. LIBERO success	LIBERO-Long over DSP	of DSP parameters	actions per pass

2 Problem and Motivation

Long-horizon imitation learning must preserve **global task intent** while executing fine-grained control. Existing coarse-to-fine policies improve structure, but discrete codebooks and weak scale coupling can introduce discontinuity and accumulated error.

- Tokenized latent skills can create hard semantic boundaries.
- Independent scales may represent incompatible behavior distributions.
- Repeated short action prediction increases exposure to compounding error.

Goal: generate long action chunks with probabilistically compatible coarse and fine latent decisions.

3 Continuous Latent Action Space

VAE LATENT OBJECTIVE

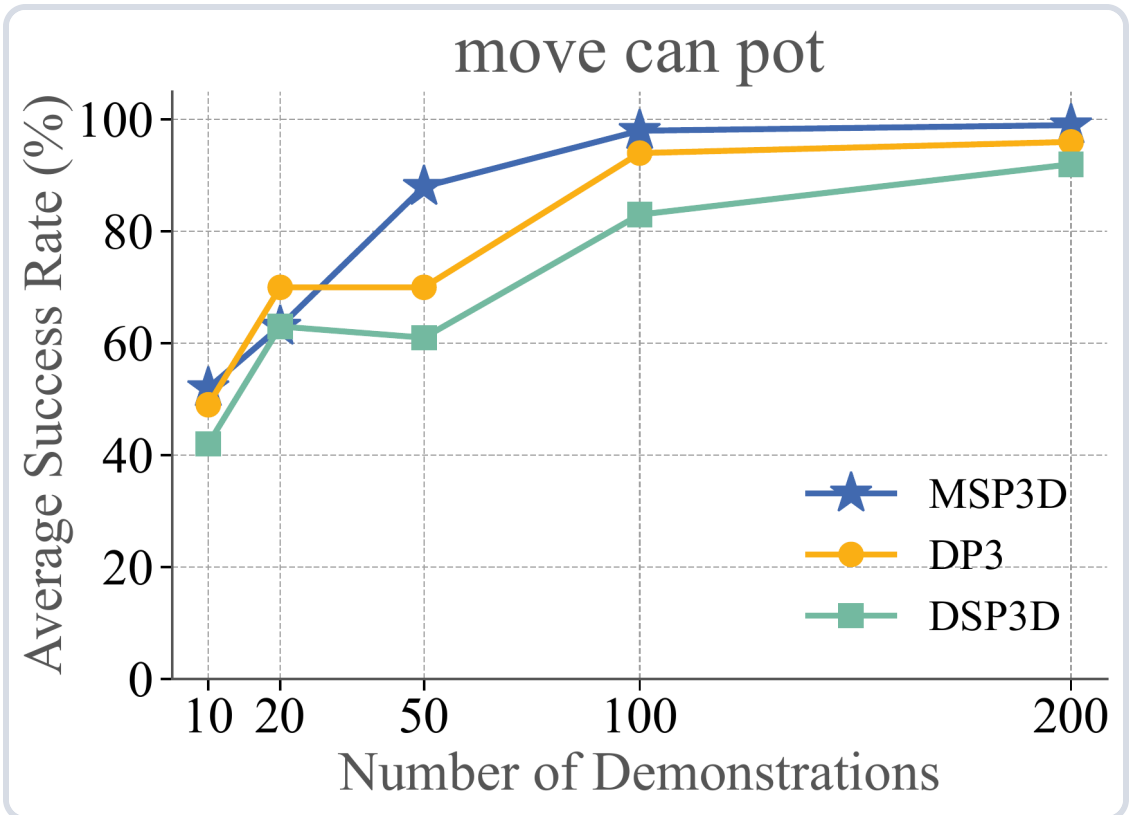
$$L_{act} = ||\hat{a} - a||_1 + \text{beta}\ KL(q(z \mid a) \,||\, N(0, I))$$

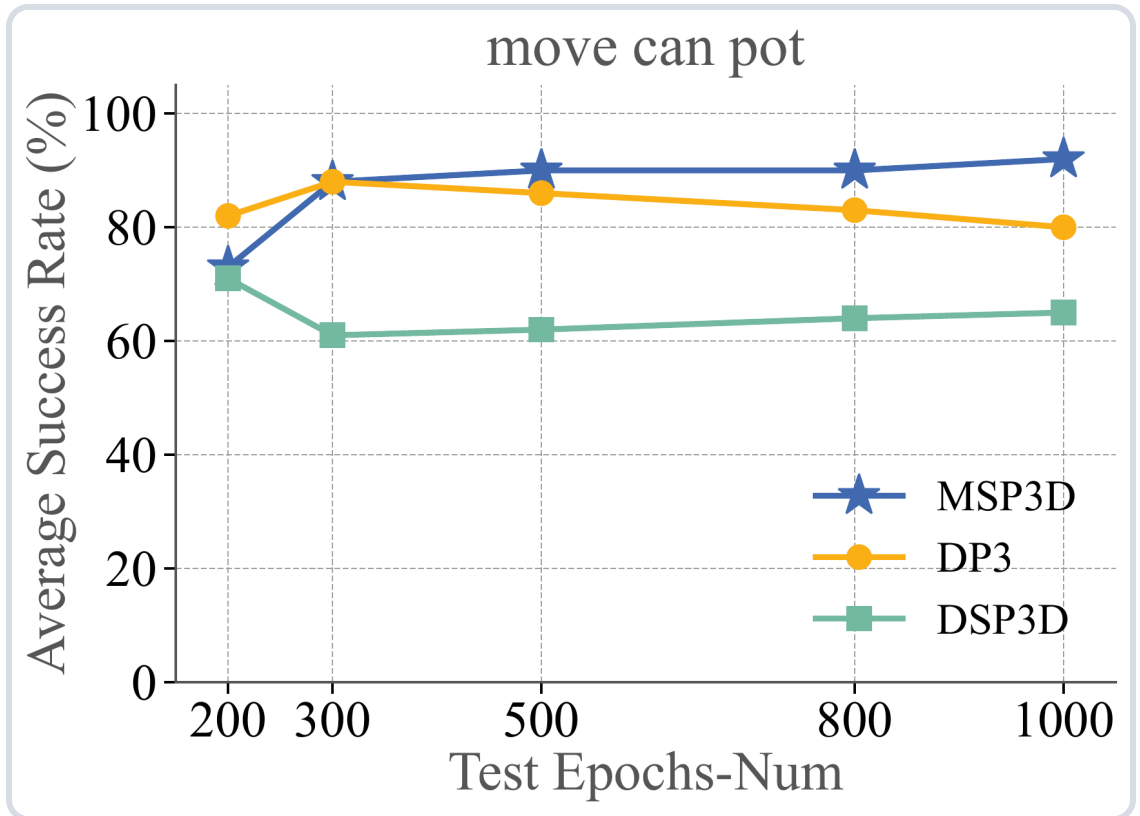
The learned latent captures core motion intent while preserving smoothness and diversity for downstream policy generation.

ENCODE	DOWNSAMPLE	CONDITION	SAMPLE
action chunk to continuous latent z	build scales {1, 2, 4, 8}	predict semantics per scale	MeanFlow produces z^S

Stage 1 freezes a compact action representation. Stage 2 trains a scale-wise Transformer and CFG-aware MeanFlow module over the latent hierarchy.

9 Training and Data Efficiency





MSP improves rapidly with limited demonstrations and shows steadier training behavior than DP or DSP in the reported RoboTwin2.0 task.

1-NFE	32	3-layer
MeanFlow generation	action chunk length	MLP ResNet flow head

4 Probabilistic Multi-Scale Modeling

LATENT CHAIN

$$p(z^{1:S}) = \text{Product}_{s=1..S}\ p(z^s \mid z^{<s}, \text{theta})$$

SEMANTIC CONDITIONING

$$\hat{z}^s = T([C, Up(z^1,2), \ldots, Up(z^{s-1},2)])$$

Coarser latents define global intent, while finer scales model local variations conditioned on prior coarse decisions. This naturally induces cross-scale compatibility without an explicit distribution-matching loss.

Consistency mechanism: all scales are deterministic downsamplings of the same finest latent trajectory, and each fine scale is generated conditioned on coarser semantics.

4	1-NFE	T=32

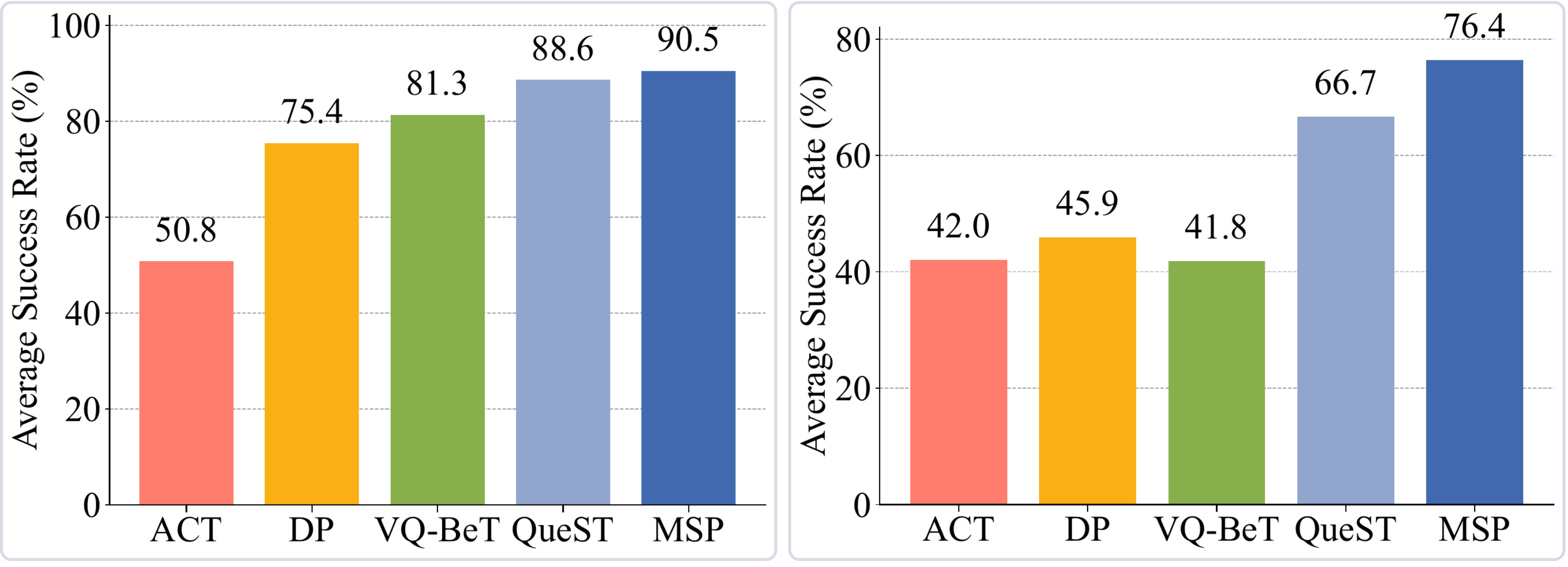
5 LIBERO Suite Results

Method	Goal	Spatial	Object	Long	Avg
DP	68.3	78.3	92.5	50.5	72.4
FlowPolicy	90.4	95.0	94.0	75.2	88.65
QueST	80.8	87.4	93.6	68.8	82.65
pi0	95.8	96.8	98.8	85.2	94.15
DSP	90.8	92.8	99.0	90.0	93.15
MSP	**93.8**	**97.0**	**100**	**96.8**	**96.9**

Average success rates over 50 rollouts per task. MSP is strongest on the long-horizon suite, improving over DSP by 6.8%.

4	50	3
LIBERO suites	rollouts per task	random seeds

6 Reusable Semantics



90.5	76.4	5	+9.7%
LIBERO-90 success	Few-shot success	demos per unseen task	over QueST

Low-data transfer: the continuous multi-scale latent space learns reusable high-level action semantics instead of brittle discrete skills.

Pretraining uses LIBERO-90 and fine-tuning uses only 5 demonstrations per unseen LIBERO-Long task.

11 Takeaways

- **Continuous:** avoid discrete tokenization while preserving fine motion control.
- **Consistent:** generate fine scales conditioned on coarse semantics.
- **Efficient:** 17.2M action-head parameters and 32-step prediction horizon.
- **Robust:** strongest gains appear on long-horizon and late-stage real-world tasks.

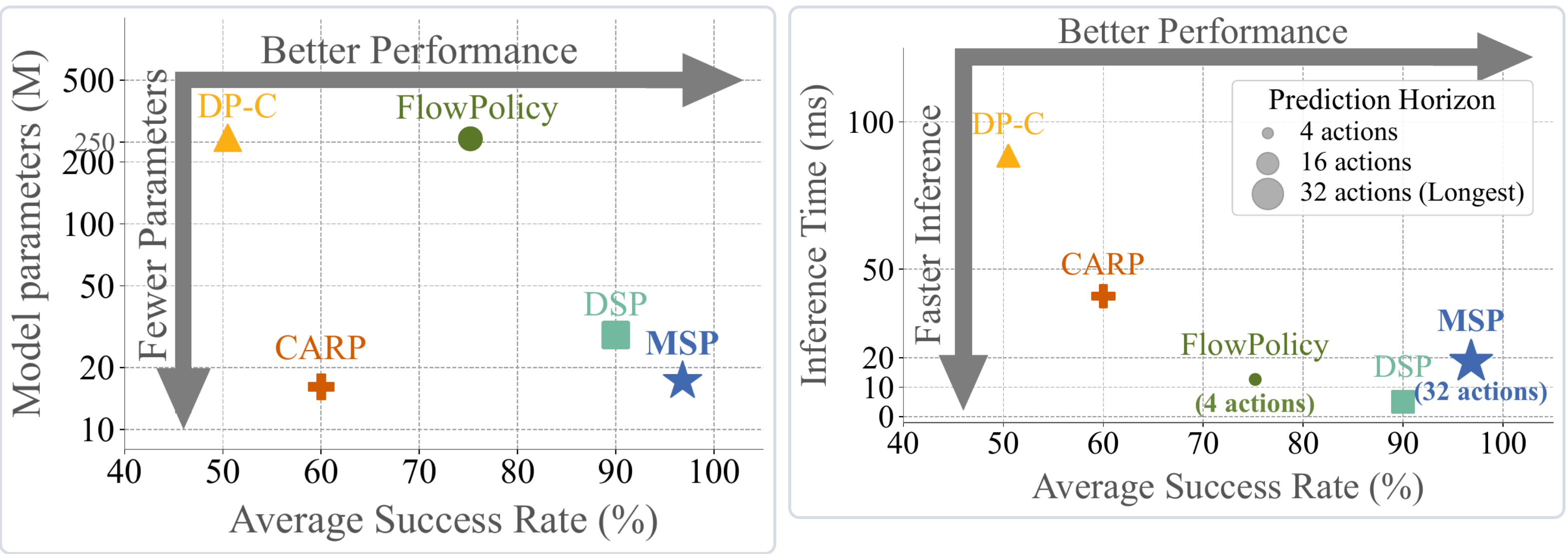
Poster summary: MSP turns long action generation into probabilistically consistent latent refinement.

7 Ablation: What Matters?

Latent	AR	Diff	Flow	MeanFlow	Long
					82.8
yes					87.5
yes	yes	yes			91.3
yes	yes		yes		92.5
yes	**yes**			**yes**	**96.8**

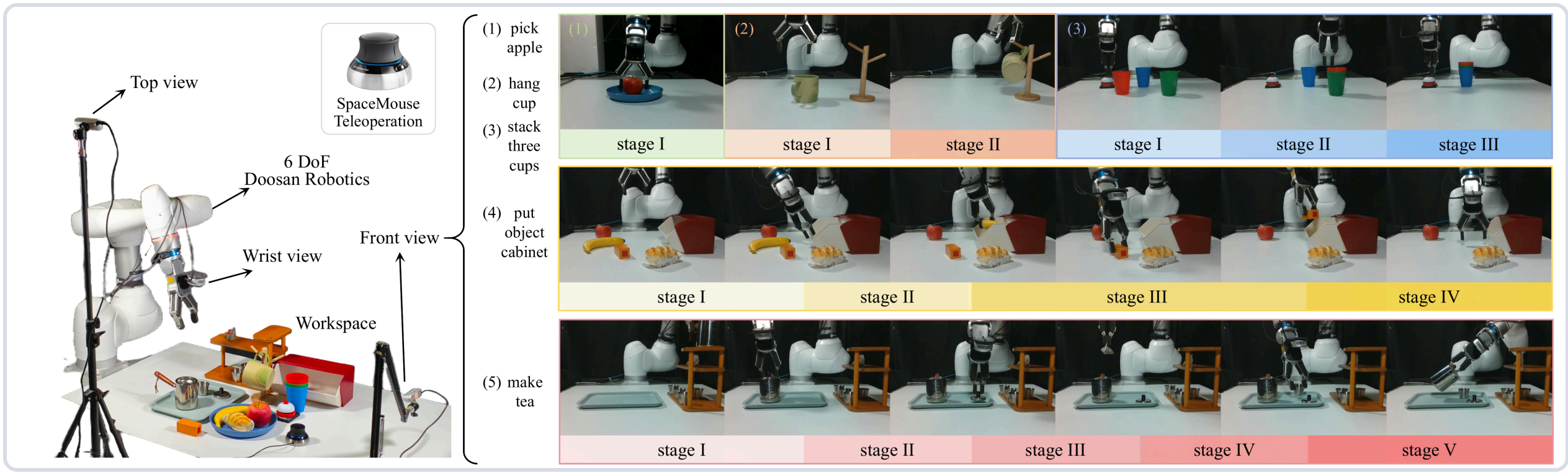
Continuous latent actions, scale-wise autoregression, and MeanFlow are complementary; the full design gives the strongest LIBERO-Long result.

8 Lightweight Action Head

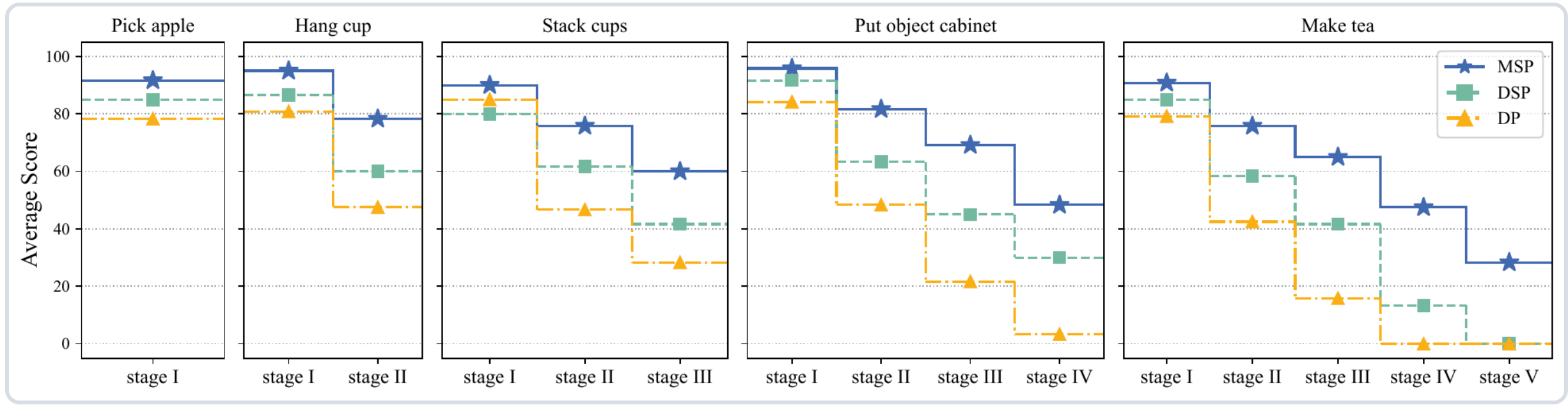


Method	Params/M	Time/ms	Horizon
DP-C	262.35	88.53	16
FlowPolicy	259.32	12.63	4
CARP	16.08	40.87	16
DSP	28.77	5.05	16
MSP	**17.2**	**18.66**	**32**

10 Real-Robot Robustness



Doosan arm with RG6 gripper, wrist L515 camera, two D435 cameras, and 50 demonstrations per task.



Task stage	MSP	DSP	DP
Cabinet IV	48.33	30.00	3.33
Make tea IV	47.50	13.33	0.00
Make tea V	**28.33**	**0.00**	**0.00**

Performance gaps widen as task stages accumulate, matching MSP's goal of stable long-horizon execution.

For multi-stage tasks, the next stage is counted only after the current stage reaches a full score, so late-stage scores reflect stable sequential completion.

5	50	3
real-world tasks	demos per task	RGB camera views

Project: **https://220xuan.github.io/msp-policy/** · Contact: **wangfei@mail.neu.edu.cn**