

**Pre-Train**

52 Datasets 300+Tasks 2.5M+Traj.

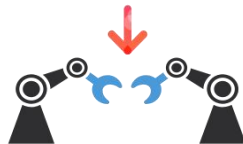
Multi-view  
Images



Task  
Instructions



**LatentVLA**



**Bimanual Action Chunk**



**Fine-Tune**



**3 Public  
Benchmarks**



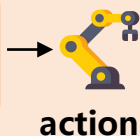
**8 Real-World  
Tasks**

## Diffusion Policy

**Data:** In-Domain Manual Teleop



**Policy**

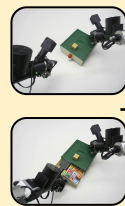


**action**

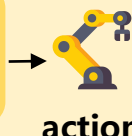
⚠ Generalization 😞  
⚠ Long-Horizon Task 😞

## Inverse Dynamics Model

**Data:** Web Video Data



**Inverse  
Dynamics**

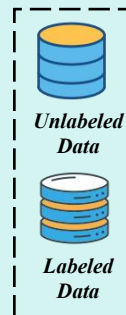


**action**

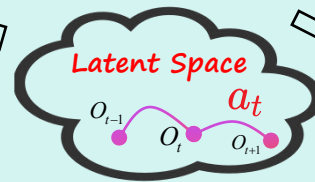
⚠ Generalization 😞  
⚠ Long-Horizon Task 😞

## Unified Latent-Action Diffusion

**Data:** Web Video Data & Manual Teleop



**LatentVLA**



**action chunk**

⚠ Generalization 😊  
⚠ Long-Horizon Task 😊