

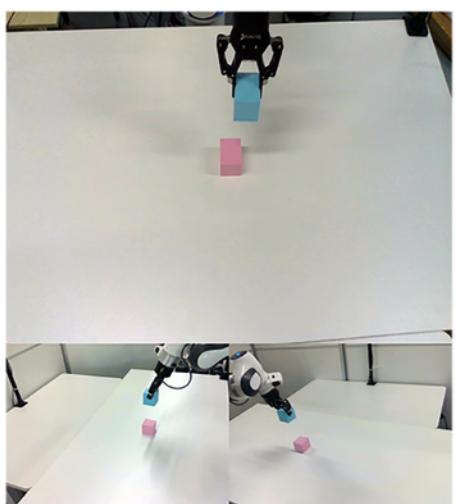
Stage 1: Pre-train



Agibot



RDT



Robomind
Franka



Robomind
Aloha

~ 750K Episodes

Stage 2: Fine-tune



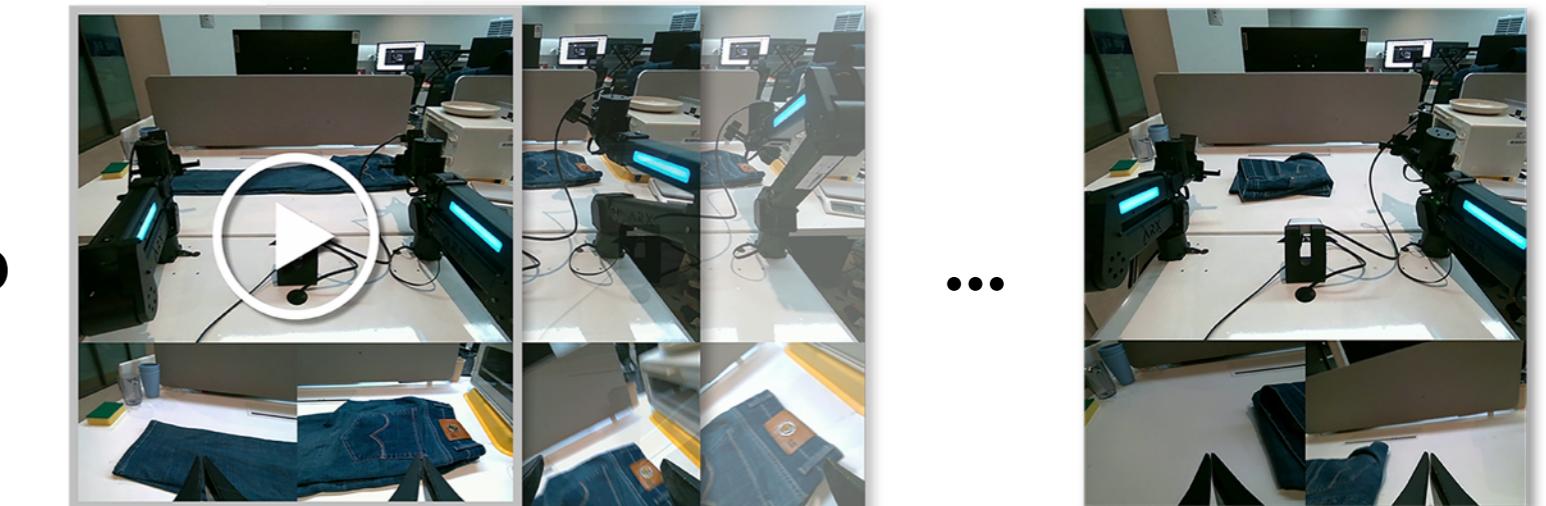
81 Tasks / 232 Episodes

Image
+
Instruction

Video Diffusion
Foundation Model

Test-time Scaling

Unified Observation Space



Video

Robot

...aloha robot...

Language

Camera

...a fixed rear camera,
a movable left arm camera...

Task

...fold the jeans in half...

→  Action

MIDM