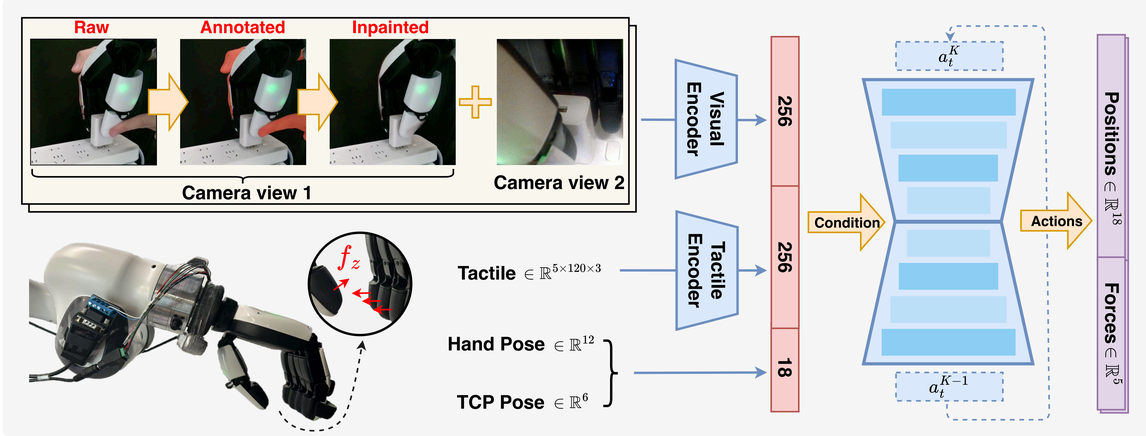
KineDex: Learning Tactile-Informed Visuomotor Policies via Kinesthetic Teaching for Dexterous Manipulation

Di Zhang[1,5], Chengbo Yuan[2,5], Chuan Wen[4], Hai Zhang[3,5], Junqiao Zhao[1], Yang Gao[2,5]

[1] Tongji University, [2] Tsinghua University, [3] University of Hong Kong
[4] Shanghai Jiao Tong University, [5] Shanghai Qi Zhi Institute

Tactile-enriched demonstrations collected via kinesthetic teaching

9 dexterous tasks accomplished via tactile-informed visuomotor policies



We present **KineDex**, a framework for collecting tactile-enriched demonstrations via kinesthetic teaching.

KineDex removes hand occlusions with inpainting, and trains policies that use visual and tactile inputs to predict joint positions and contact forces executed with force control.

Kinesthetic Data Collection

Motivation

The core idea of KineDex is a hand-over-hand setup where the operator "wears" the robotic hand, directly transmitting contact forces for natural feedback; the thumb is controlled separately from the other fingers.

During kinesthetic teaching, we record:

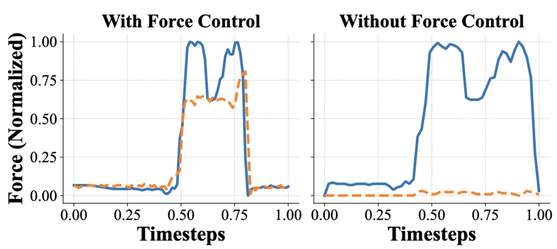
- **Visual**: RGB images from front-facing and wrist-mounted cameras (front-view occlusions removed via inpainting).
- **Proprioception**: Arm end-effector pose and hand joint positions.
- **Tactile sensing**: Per-finger measurements from dense sensing points.
- **Fingertip force**: 3D vector $\mathbf{f} = (f_x, f_y, f_z)$ aggregated from tactile data.

Tactile Data Utilization

At inference, the policy outputs action chunks of desired joint positions and fingertip forces. These forces are realized through a force-informed control scheme that interprets predicted forces as virtual displacements of the joints:

$$\begin{cases} x_d^{tip} = x^{tip} + K^{tip} \cdot f_d \\ x_d^{base} = x^{base} + K^{base} \cdot f_d \end{cases}$$

We visualize the predicted and sensed thumb forces during execution:



With force control, the sensed forces closely track the predicted targets, while without force control the forces deviate significantly.

Experimental Results

We evaluate KineDex across multiple dexterous tasks, and results show that force control, tactile input, and inpainting are all essential for achieving high success rates.

Method	*Bottle Picking*	*Cup Picking*	*Egg Picking*	*Cap Twisting*	*Nut Tightening*
KineDex	17	20	17	15	16
w/o Force Control	0	16	5	2	7
w/o Tactile Input	15	17	18	10	12
w/o Inpainting	0	0	0	0	0

KineDex completes demonstrations in about half the time of teleoperation. A user study further confirms that KineDex is more comfortable and intuitive for operators during data collection.