VT-Refine: Learning Bimanual Assembly with Visuo-Tactile Feedback via Simulation Fine-Tuning

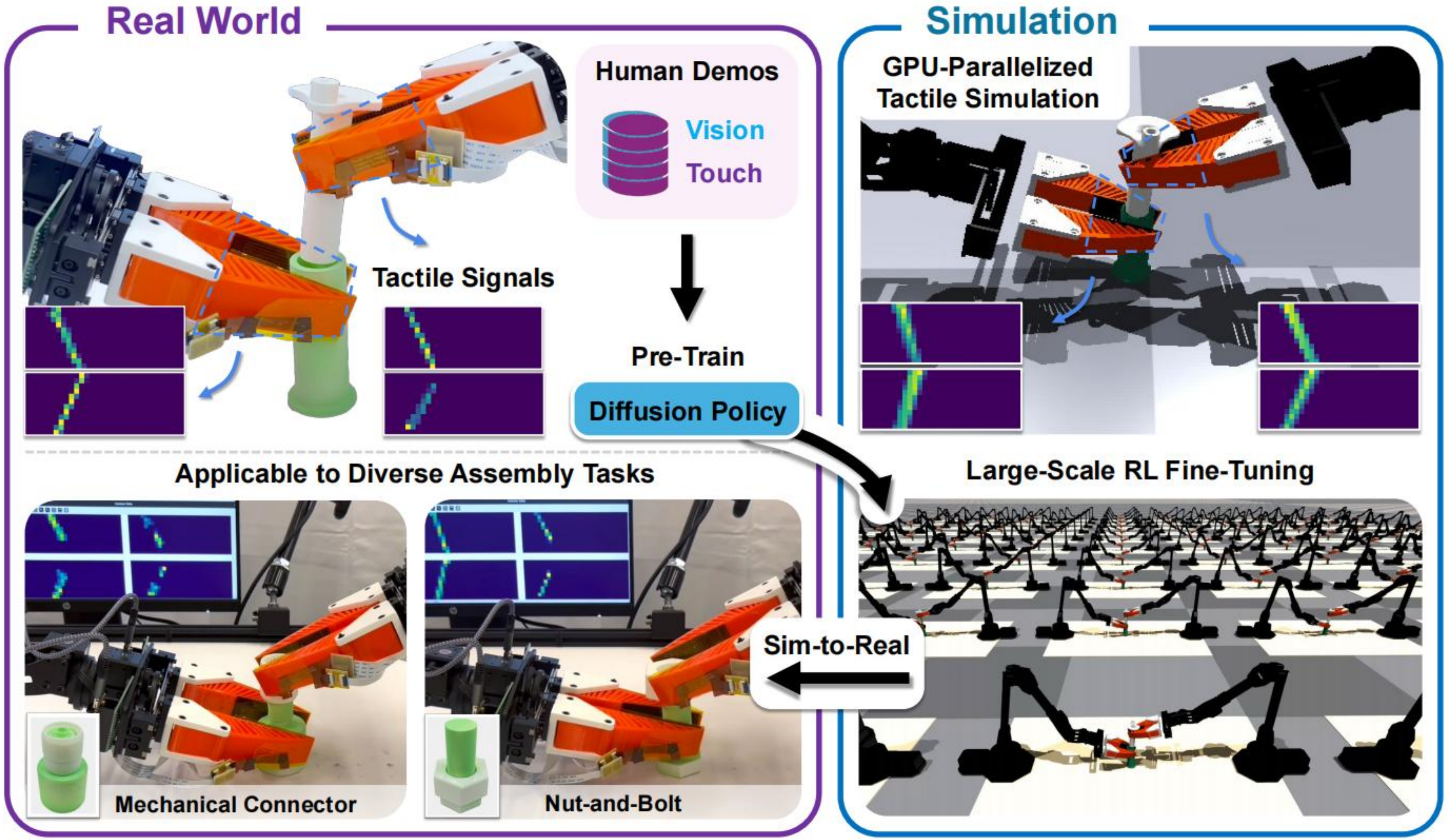
（CoRL 2025）

real-to-sim-to-real framework

模拟环境中基于强化学习的微调改进了视触扩散策略

GPU并行化触觉模拟模块，捕捉法向力信号，显著缩小gap

采用基于点的表征方式处理视觉与触觉模态，实现了无缝的"真实-模拟-真实"迁移。

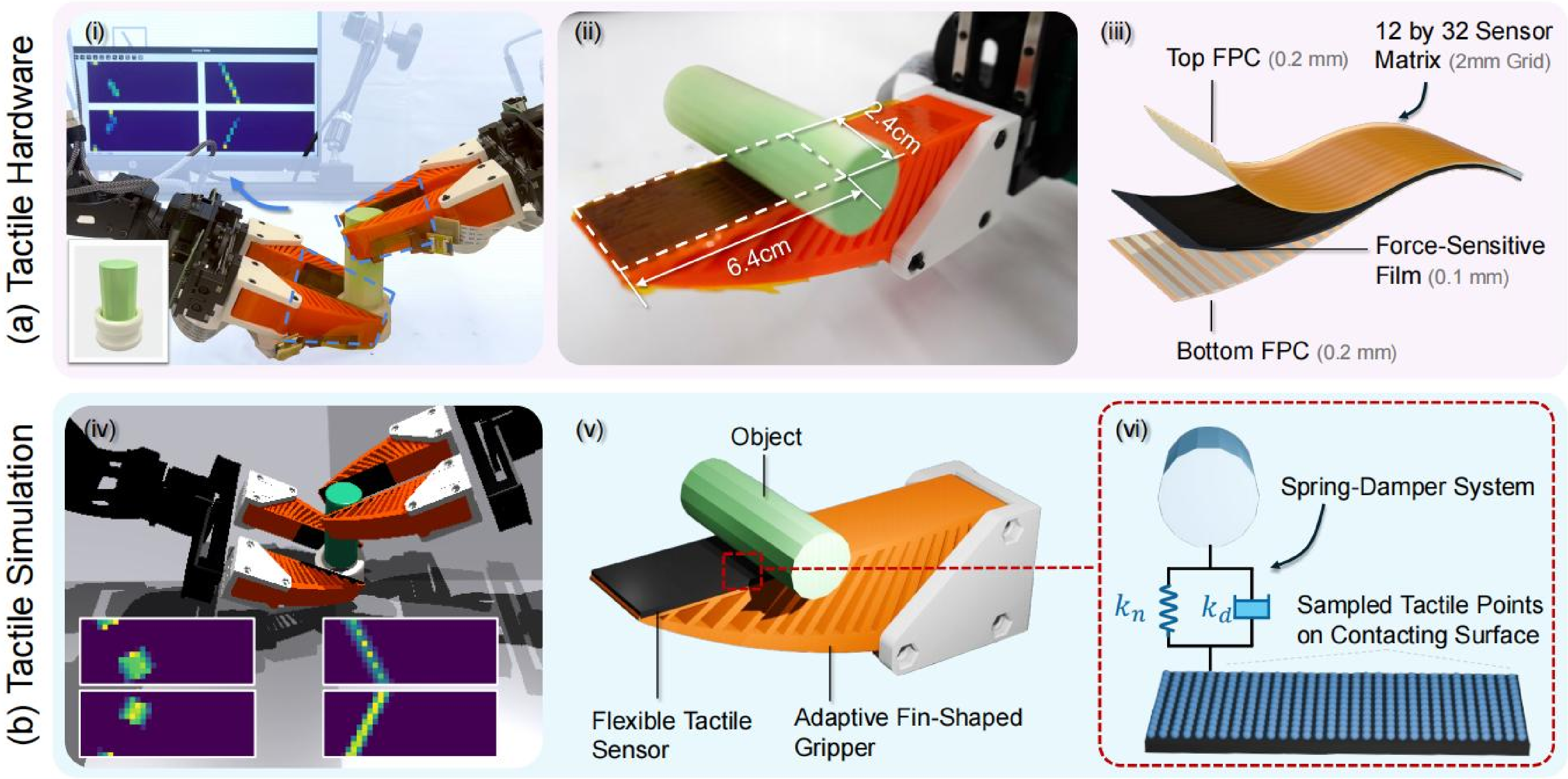


触觉硬件的选用与仿真

通过法向力体现的结构化接触特征

仿真易迁移

不需要表面纹理信息



FlexiTac
12×32个传感单元
空间分辨率为2毫米
低成本、可扩展部署

TacSL
触觉单元网格接触交互
弹簧-阻尼模型

$$x_i^w = R_e x_i^l + p_e$$

$$\dot{x}_i^w = \omega_e \times (R_e x_i^l) + v_e$$

$$\dot{d}_i = \hat{n}_i \cdot \dot{x}_i^w$$

$$s_i = \mathrm{normalize}(\max(d_i, 0), f_{n,i}) \qquad f_n = -(k_n d + k_d \dot{d})\mathbf{n}$$

视觉-触觉策略优化

$$\pi : \mathcal{O} \rightarrow \mathcal{A}$$

$$o_t = \{P_t^{\text{visual}}, P_t^{\text{tactile}}, q_t\}$$

$$P_t^{\text{visual}} = \{(x_i, y_i, z_i, 0)\}_{i=1}^{N_v}:$$

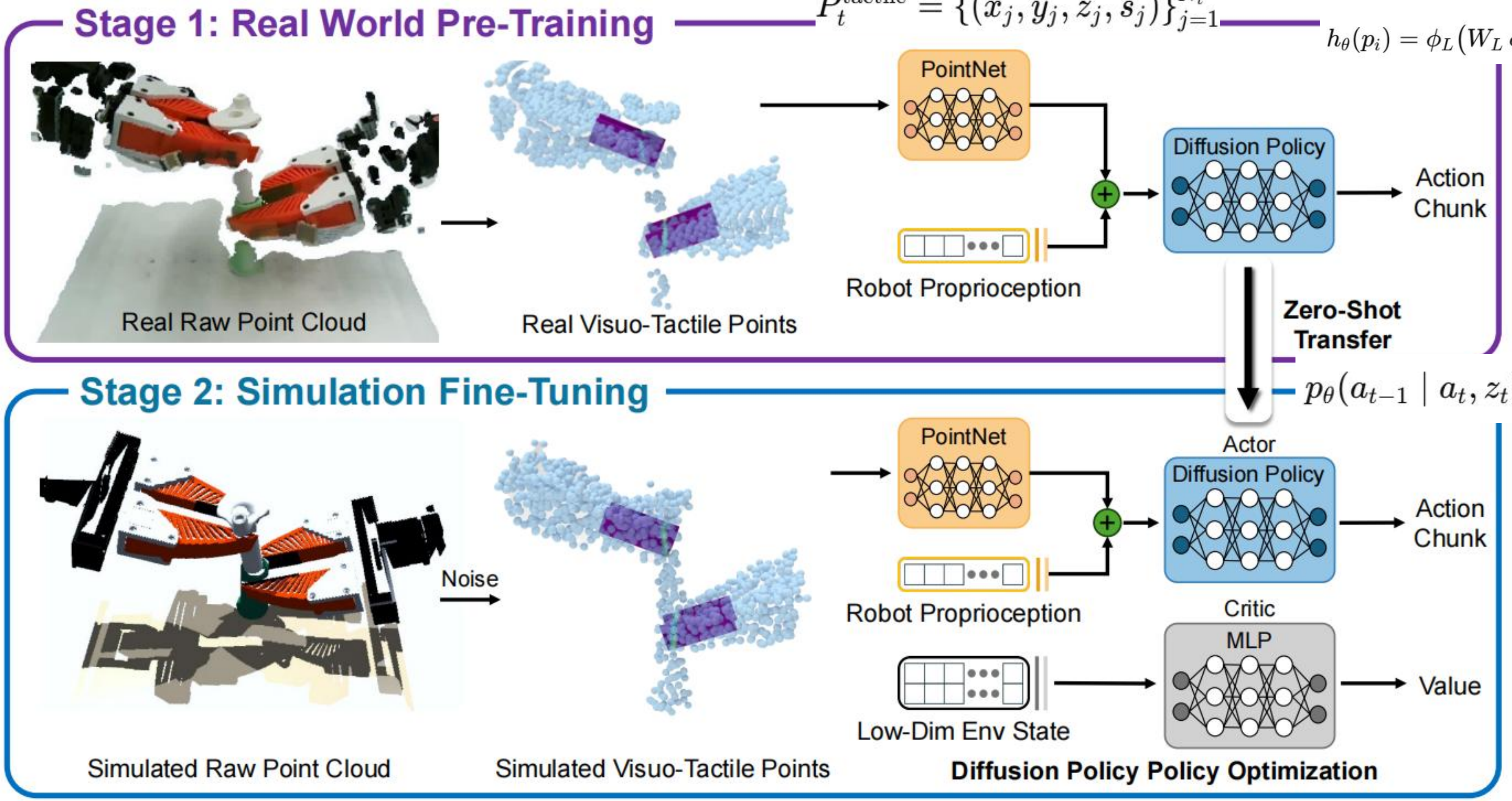
$$P_t^{\text{tactile}} = \{(x_j, y_j, z_j, s_j)\}_{j=1}^{N_t}$$

$$P_t = P_t^{\text{visual}} \cup P_t^{\text{tactile}}$$

$$P_t = \{p_1, p_2, \ldots, p_N\}, \; p_i \in \mathbb{R}^d$$

$$f_t = g(\{h_\theta(p_i)\}_{i=1}^{N}) = \max_{i=1\ldots N} h_\theta(p_i)$$

$$h_\theta(p_i) = \phi_L(W_L\, \phi_{L-1}(W_{L-1} \ldots \phi_1(W_1 p_i + b_1) \cdots + b_{L-1}) + b_L)$$



$$f_t = \max_{i=1,\ldots,N} h_\theta(p_i) \in \mathbb{R}^{64}$$

$$z_t = [\, f_t;\; f_t^{\text{prop}} \,] \quad z_t \in \mathbb{R}^{128}$$

$$\pi_\theta(a_{1:H} \mid o_{1:H}) = p_\theta(a_{1:H} \mid z_{1:H})$$

$$p_\theta(a_{t-1} \mid a_t, z_t) = \mathcal{N}(a_{t-1}; \mu_\theta(a_t, t, z_t), \Sigma_\theta(a_t, t, z_t))$$

$$\mathcal{L}_{\text{diff}} = \mathbb{E}_{a_0,\epsilon,t}\left[\, \|\epsilon - \epsilon_\theta(a_t, t, z_t)\|^2 \right]$$

$$\theta^*_{\text{pretrain}} = \arg\min_\theta \mathcal{L}_{\text{diff}}(\theta)$$

$$V_\phi(s_t) \approx \mathbb{E}_{\pi_\theta}\left[\sum_{k=0}^{\infty} \gamma^k r_{t+k}\right]$$

$$A_t = R_t - V_\phi(s_t)$$

$$L^{\text{PPO}}(\theta) = \mathbb{E}\left[\min\left(\rho_t(\theta)\,\hat{A}_t,\; \text{clip}(\rho_t, 1-\epsilon, 1+\epsilon)\,\hat{A}_t\right) - \lambda_{\text{KL}}\, D_{\text{KL}}(\pi_\theta \,\|\, \pi_{\text{old}})\right]$$

$$\rho_t(\theta) = \frac{\pi_\theta(\mathbf{a}_t \mid \mathbf{z}_t)}{\pi_{\text{old}}(\mathbf{a}_t \mid \mathbf{z}_t)}$$

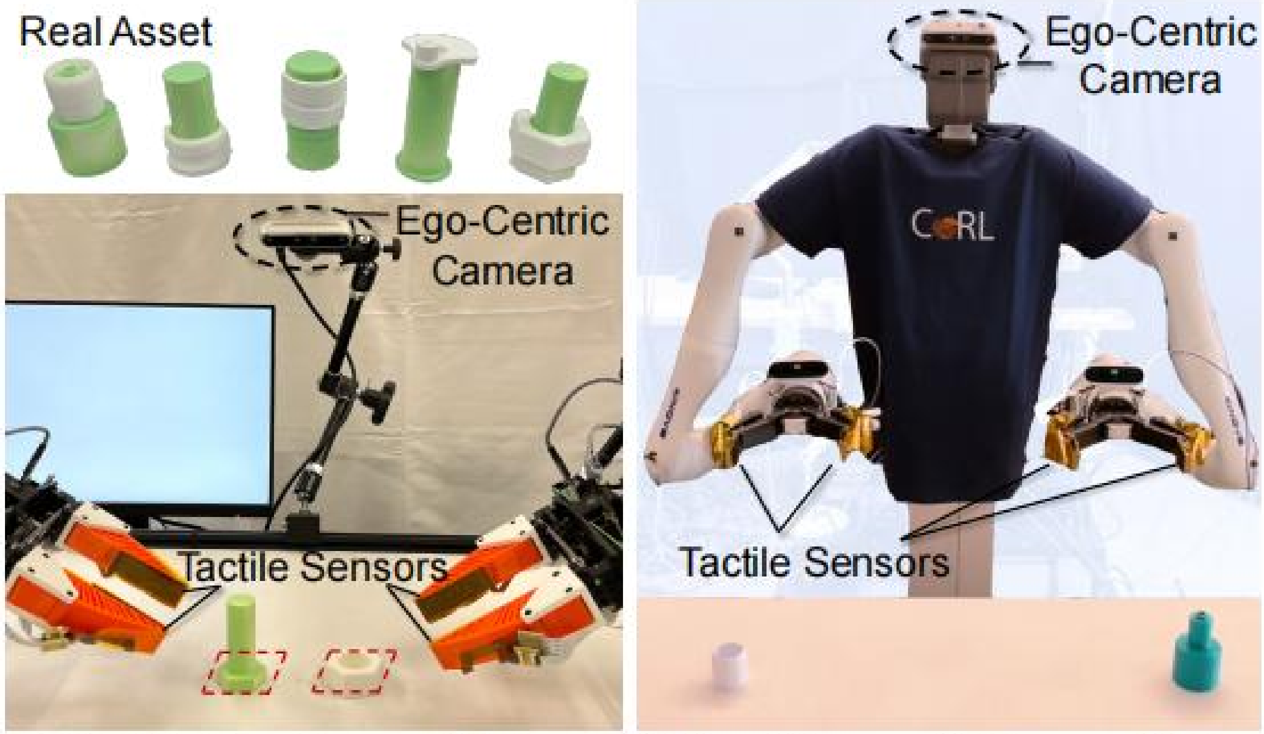
$$\underbrace{\text{PointNet}(P^{vis}, P^{tac})}_{\text{统一特征编码}} \Rightarrow \underbrace{\text{Diffusion Policy}}_{\text{去噪式模仿学习}} \xrightarrow[\text{DPPO}]{\text{仿真强化学习}} \text{高精度装配策略}$$

通过实验探讨以下三个问题：

（1）我们微调的策略相较于基线扩散策略有何改进？

（2）所提出的视觉-触觉表征在跨领域迁移（真实-仿真-真实）中的效果如何？

（3）策略性能如何随着人类示范次数的增加而变化



(a) Table-Top Bimanual Setup (b) Semi-Humanoid Bimanual Setup

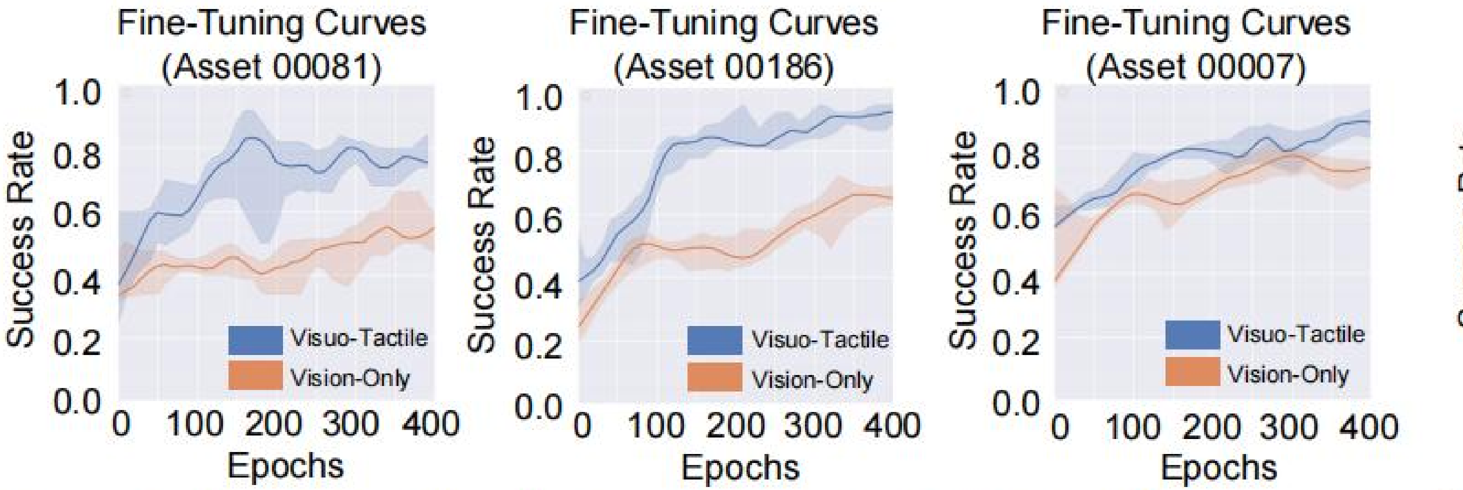


Figure 6: **Simulation Fine-Tuning of Pre-Trained Policies.** We compare the fine-tuning performance of visuo-tactile (blue) and vision-only (orange) policies. The visuo-tactile policy starts with not only a higher pre-trained performance but also continues to improve, achieving higher final performance after fine-tuning.

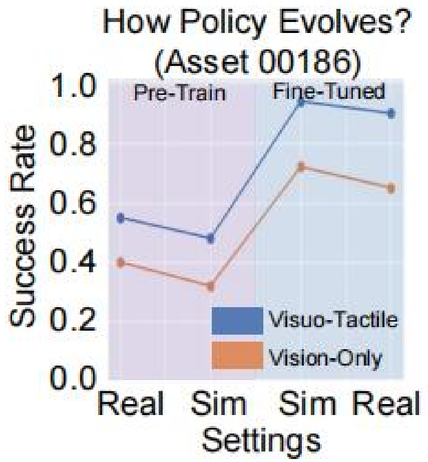


Figure 7: Performance of *Pre-Trained Policy* in sim and real, *Fine-Tuned Policy* in sim and real.

结果	成功率（平均）	说明
Vision-Only BC	~0.4	模仿学习性能差，受噪声与欠感知限制
Visuo-Tactile BC	~0.6	触觉显著提升装配成功率
Vision-Only + RL	~0.6	RL 能改善性能但缺乏精细接触
Visuo-Tactile + RL (Ours)	**~0.9**	最优，既有触觉先验又有RL探索