The 1st Embodied4Arts Workshop @ RSS 2026

Co-Creation in Bloom: Artistic Co-Creation Manipulation with Deformable Flower Petals

Kefei Zhu[1,*], Jinzhong Gu[1,*], Yue Xin[1,*], Yi Qu[2], Yaodong Yang[1], Yuanpei Chen[1], Hao Dong[1,†]
[1]Peking University [2]The University of Edinburgh
[*]Equal contribution [†]Corresponding author

PETALDEX

***Abstract*—We present *PetalDex*, a bimanual dexterous robot system for artistic human–robot co-creation through collaborative flower-petal painting. Artistic human–robot co-creation manipulation is more challenging than conventional manipulation, requiring both dexterous manipulation and continuous adaptation to a human partner in an open-ended creative task. To explore this setting, we introduce (i) an artistic human–robot co-creation task, (ii) a data-collection principle, and (iii) a vision-based imitation learning policy for collaborative manipulation. Beyond the traditional success rate, we propose three metrics to evaluate artistic collaboration: *rhythm*, *adaptation*, and *completion*. Experiments show that *PetalDex* enables autonomous creation, seamless human–robot turn-taking, online adaptation to human interventions, and autonomous completion of collaborative artworks. Video demo: `petaldex.github.io`.**

***Index Terms*—Artistic manipulation, human-robot co-creation, dexterous manipulation**

I. Introduction

Robots have created visual art for decades, from AARON [1] to recent manipulation-based painters such as IMPASTO [2], yet these systems create alone. We study *artistic human–robot co-creation manipulation*, in which a bimanual dexterous robot and a person paint with flower petals on a shared canvas. This task is challenging in three respects. As a manipulation problem, petals are light, deformable, and non-repeatable, demanding fine dexterity. As a human–robot interaction problem, the partner may join, place, or disturb petals at any moment, and the robot must coordinate its actions with the human in real time. Finally, the task is open-ended yet completion is visually defined, so the robot must recognize completion and stop on its own. To address these challenges, we contribute the task, a data-collection recipe, and a single vision-based imitation policy, evaluated by success rate and three collaboration metrics (Fig. 1).

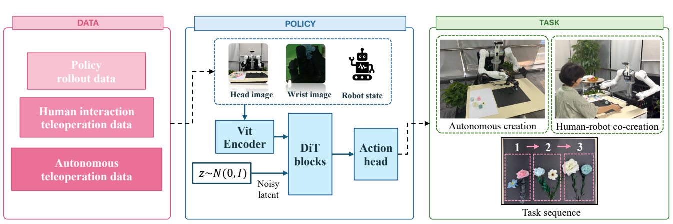

Fig. 1. **Overview of PetalDex.** *Left:* three types of demonstrations. *Middle:* a ViT encoder embeds head and wrist images, and DiT blocks fuse them with the robot state and a noisy latent $z \sim \mathcal{N}(0, I)$ to predict action chunks, trained as an IMLE policy. *Right:* the petal painting task, performed by the robot alone or in human–robot co-creation until the artwork is complete.

II. Method

Task. Our bimanual platform consists of two Tianji arms, two Wuji dexterous hands, and two RGB cameras. In *petal painting*, the robot, alone or with a human partner, arranges petals onto a canvas with three leaf branches. A trial succeeds when at least six petals are balanced across the branches.

Data. Using Manus-glove teleoperation, we collect 525 autonomous demonstrations, 200 human–robot co-creation demonstrations, and additional policy-rollout corrections.

Policy. We adopt IMLE Policy [3] with a pre-trained ViT image encoder. At each step t, the policy observes $o_t = (I_t, s_t)$, where I_t stacks the head and wrist RGB images and $s_t \in \mathbb{R}^{54}$ is the proprioceptive state. Conditioned on $z \sim \mathcal{N}(0, I)$, it predicts a 16-step action chunk $A_t = (a_t, \ldots, a_{t+H-1})$, where $H = 16$ and $a_i \in \mathbb{R}^{54}$.

III. Results

Success rate. The robot achieves an 87% placement success rate (78/90). At the session level, it completes 6/10 autonomous-creation sessions (60%) and 8/10 human–robot co-creation sessions (80%).

Rhythm measures whether the robot keeps pace with the partner: the ratio of the robot's average pick-and-place time to the human's is 1.04 (human 4.25s, robot 4.40s), indicating that the robot keeps up rather than lagging behind.

Adaptation measures whether the robot responds to the current canvas rather than blindly replaying the demonstrated order: after the human rearranges the layout, it places the next petal in the correct region in 85% of cases.

Completion measures autonomous termination: when the canvas reaches its finished state, the robot stops on its own, terminating correctly in 95% of sessions.

References

[1] H. Cohen, "The further exploits of AARON, painter," *Stanford Humanities Review*, 4(2):141–158, 1995.
[2] Y. Wang *et al.*, "IMPASTO: Integrating model-based planning with learned dynamics for robotic oil painting," in *Proc. ICRA*, 2026.
[3] K. Rana *et al.*, "IMLE policy: Fast and sample efficient visuomotor policy learning," in *Proc. RSS*, 2025.