ForceVLA2: Unleashing Hybrid Force-Position Control with Force Awareness for Contact-Rich Manipulation
Supplementary Materials

Appendix

A. Theoretical Analysis of Force Pathway

In this section, we provide a control-theoretic interpretation of why the proposed *reactive force pathway*, the short-horizon reactive manipulation skill that bypasses high-level fusion and feeds force observations directly into the action expert, enhances stability and responsiveness in contact-rich manipulation. As discussed in the main paper, ForceVLA2 routes slowly varying proprioceptive states together with force-aware task knowledge from the VLM expert into the multimodal encoder of the action expert, while injecting rapidly changing force observation directly into the Cross-Scale MoE. This short, high-fidelity gradient path improves the observability and controllability of contact forces under hybrid force–position control, thereby enabling rapid force feedback, reducing reliance on past trajectories, and supporting active, force-guided exploration and interaction.

Problem definition. To formally justify this architectural design, we first analyze the fundamental control-theoretic limitations of traditional position-only control. Consider a robotic manipulator interacting with an unknown environment. The proprioceptive state refers to the robot's end-effector (EE) 6D pose $\mathbf{p} \in \mathbb{R}^7$ (position and quaternion) and interaction force $\mathbf{f} \in \mathbb{R}^6$ (force and torque), it is worth noting that in the control scenario, we utilize EE 6D pose $\mathbf{p_e} \in \mathbb{R}^6$ (position and rotation). The critical challenge lies in the environmental dynamics, characterized by an unknown nonlinear mapping:

$$\mathbf{f} = \mathbf{\Phi}(\mathbf{p_e}, \dot{\mathbf{p}}_e, \boldsymbol{\theta}_e), \tag{1}$$

where $\mathbf{\Phi} : \mathbb{R}^{12} \times \Theta_e \to \mathbb{R}^6$ represents the environment's constitutive law with unknown parameters $\boldsymbol{\theta}_e$.

Analysis. When using position-only control with force input, the control policy outputs target EE 6D pose $\mathbf{a_p} \in \mathbb{R}^6$. The system dynamics under ideal position tracking become:

$$\begin{aligned} \mathbf{p}(k+1) &= \mathbf{a_p}(k), \\ \mathbf{f}(k+1) &= \mathbf{\Phi}(\mathbf{a_p}(k), \dot{\mathbf{a}}_p(k), \boldsymbol{\theta}_e). \end{aligned} \tag{2}$$

The observability matrix for this system has full rank since both $\mathbf{p}_e$ and $\mathbf{f}$ are directly measured through vision and force sensors. However, the controllability analysis reveals a fundamental limitation. The controllability matrix for the augmented state $\mathbf{z} = [\mathbf{p}_e^T, \mathbf{f}^T]^T \in \mathbb{R}^{12}$ is:

$$\mathcal{C} = [\mathbf{B} \quad \mathbf{AB} \quad \cdots \quad \mathbf{A}^{11}\mathbf{B}], \tag{3}$$

where $\mathbf{A}$ is the system dynamics matrix and the input matrix $\mathbf{B} \in \mathbb{R}^{12\times 6}$ has the structure:

$$\mathbf{B} = \begin{bmatrix} \mathbf{I}_6 \\ \mathbf{0}_{6\times 6} \end{bmatrix}, \tag{4}$$

immediately implying:

$$\text{rank}(\mathcal{C}) \leq \text{rank}(\mathbf{B}) = 6 < 12 = \dim(\mathbf{z}). \tag{5}$$

Conclusion. This rank deficiency arises because force $\mathbf{f}$ is not an independent control variable but rather a dependent output determined by the unknown environment dynamics $\mathbf{\Phi}$. Consequently, the reachable set in the task space is constrained to:

$$\mathcal{R} = \{(\mathbf{p}_e, \mathbf{f}) : \mathbf{f} = \mathbf{\Phi}(\mathbf{p}_e, \dot{\mathbf{p}}_e, \boldsymbol{\theta}_e), \quad \mathbf{p}_e \in \mathcal{P}\}, \tag{6}$$

which represents a 6-dimensional manifold embedded in the 12-dimensional task space, confirming that arbitrary force-position combinations cannot be achieved through position-only control.

Innovation. ForceVLA2 fundamentally transforms the control architecture by introducing hybrid force-position control $\mathbf{a} = [\mathbf{a}_p^T, \mathbf{f}_t^T]^T \in \mathbb{R}^{12}$. The key innovation lies in employing a deep neural network policy $\pi_\theta : \mathcal{O} \to \mathbb{R}^{12}$ that learns from interaction observations.

Through supervised learning on expert demonstrations, the policy implicitly acquires an inverse model of the environment dynamics. Specifically, for a desired task-space target $(\mathbf{p}_e^*, \mathbf{f}^*)$, the trained policy learns to compute control commands that achieve this target by leveraging the learned relationship between positions and forces. The policy effectively learns to solve the inverse problem, i.e., given $\mathbf{f}^* = \mathbf{\Phi}(\mathbf{p}_e^*, \dot{\mathbf{p}}_e^*, \boldsymbol{\theta}_e)$, it finds control inputs that achieve $(\mathbf{p}_e^*, \mathbf{f}^*)$.

The learned policy effectively extends the control authority to approach the full 12-dimensional action space. By incorporating both position and force control, and learning their coupling through the environment dynamics, ForceVLA2 can achieve force-position combinations that lie beyond the restricted manifold of position-only control. The key insight is that the policy learns to coordinate position and force commands in a way that respects the physical constraints while expanding the reachable set.

To quantify this enhancement, we define the effective controllability index:

$$\kappa = \frac{\dim(\mathcal{R})}{\dim(\mathcal{Z})}, \tag{7}$$

where $\mathcal{R}$ is the reachable set and $\mathcal{Z}$ is the task space. For position-only control, $\kappa_{\text{position-only}} = 6/12 = 0.5$, indicating that only 50% of the task space dimensionality is controllable. With ForceVLA2 and a well-trained policy, the reachable set expands significantly as the policy learns to exploit the learned environment model, leading to $\kappa_{\text{ForceVLA2}} \to \kappa_{\text{max}}$, where κ_{max} depends on the richness of the training data and the physical constraints of the task.

This improvement from $\kappa = 0.5$ toward higher values represents the fundamental achievement of ForceVLA2: leveraging learned environment knowledge to expand the controllable subspace and enable precise force-position coordination in contact-rich manipulation tasks. The reactive force pathway further enhances this capability by providing rapid, direct force feedback to the action expert, enabling active adjustment of interaction forces during contact-rich manipulation.

B. Statistical details of the ForceVLA2-Dataset

We analyze the statistical distribution of normalized forces (scaled to ±100 N) and normalized torques (scaled to ±15 Nm) across manipulation tasks in the ForceVLA2-Dataset, which exhibit task-specific characteristics (Fig. 1 (a) and (b)).

For the force distribution, forces in the `Press bottle` task are predominantly concentrated along the negative z-axis (vertical downward direction). The `Clean vase` and `Clean board` tasks show more distributed force patterns across all three axes, reflecting the wiping motions that involve both tangential sliding forces and normal contact forces. The `Retrieve plate` task demonstrates relatively symmetric distributions centered near zero across all axes, suggesting exploratory probing behaviors with gentle contact forces. The `Assemble gears` task exhibits a broader force distribution, particularly along the z-axis, corresponding to actions during the alignment and insertion phases.

For the torque distribution, the `Press bottle` task shows concentrated torque distributions near zero across all axes, indicating minimal rotational adjustments. In contrast, the `Clean vase`, `Clean board`, and `Retrieve plate` tasks exhibit wider torque spreads, reflecting the continuous orientation adjustments required for grasping, wiping, and probing. Moreover, the `Assemble gears` task presents the most diverse torque distribution, particularly in the z-axis, corresponding to the precise rotational alignment required during gear insertion.

Furthermore, we decompose the five tasks into short-horizon reactive manipulation skills. As illustrated in Fig. 1 (c), we identify five fundamental skills based on force and position trajectories. For the Wipe skill, frames during the period where position trajectories change more than 0.05 m and force amplitude is more than 10 N. For the Push skill, frames during the period where z-axis position trajectories change more than 0.05 m and z-axis force amplitude more than 5 N. For the Grasp skill, frames during the period where z-axis position trajectories change more than 0.1 m and force amplitude more than 5 N. For the Rotate skill, frames during the period where forces of three axes change more than 1 N. Frames outside the above definitions are the Explore skills, featured with low-magnitude force and rapid position drift. Statistical analysis shows that the ForceVLA2-Dataset is dominated by Explore (45.62%, followed by Wipe, Push, Grasp and Rotate, indicating diverse force-position coordination patterns.

Table 1. **Training details of ForceVLA2.**

Parameter	Value
Learning Rate Schedule	Cosine Decay
Optimizer	AdamW
EMA Decay	0.99
Random Seed	42
Batch Size	32
Training Steps	30,000

C. Implementation and Experiments

In this section, we provide a detailed explanation of our training strategy and hyperparameters. Additionally, we include an ablation study on the multimodal encoder. As shown in Tab. 1, we trained the ForceVLA2 model on 8 A100 GPUs with a batch size of 32 and 30,000 total training steps, which took approximately 10 hours. ForceVLA2 achieves an inference speed of 15 Hz on a 4090 GPU with a chunk size of 30.

In the multimodal encoder module, we further investigate where to inject the force branch to validate its effectiveness. As illustrated in Fig. 2, there are three candidate insertion points: (1) into the VLM pathway, (2) into the multimodal encoder, and (3) fused jointly with the state (EE 6D pose). The third insertion point (3) is the fusion of

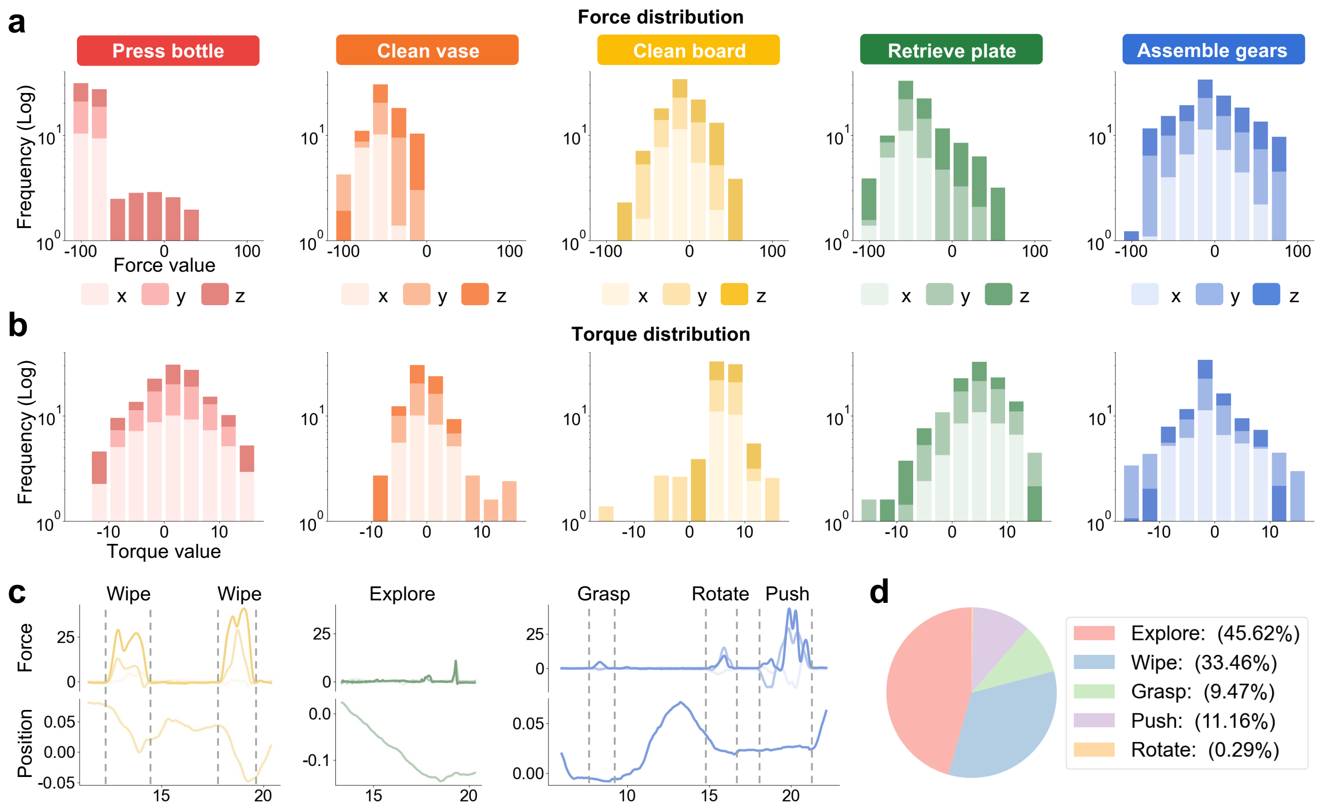

Figure 1. **Statistical details of ForceVLA2-Dataset.** Frequency distributions of force components (a) and (b) torque components of x, y, z dimensions recorded during task execution. (c) Illustration of short-horizon reactive manipulation skills based on force and position. (d) Skill distribution.

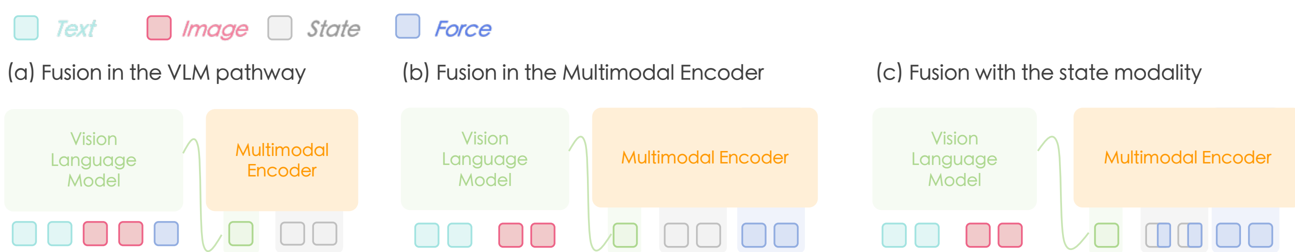

Figure 2. **Ablation study of force injection in the multimodal encoder.** The figure illustrates three candidate insertion points for the force branch: (1) fusion in the VLM pathway, (2) fusion in the multimodal encoder, and (3) fusion with the state, building upon (2).

force with the state, building upon the second point (injection into the multimodal encoder) and combining it with the state model. As reported in Tab. 2, injecting force into the VLM pathway leads to a substantial degradation in overall performance, while injecting it into either the multimodal encoder or the state fusion branch yields consistent gains. Based on these observations, our final ForceVLA2 design incorporates force at both the multimodal encoder and state fusion levels.

Results show that injecting force into the VLM pathway causes significant performance degradation, while force integration into the Multimodality Encoder or EE 6D pose fusion branch consistently improves model performance.

Table 2. **Success rates (%) of kinds of force injection in the multimodal encoder.** `ME`: Multimodal Encoder. State fusion means fusion with state modality, building upon ME.

Method	Press bottle	Clean vase	Clean board	Retri. plate	Assem. gears	Avg.
VLM Pathway	10.0	10.0	5.0	0.0	0.0	5.0
ME	75.0	55.0	65.0	40.0	55.0	58.0
State Fusion	**80.0**	**75.0**	**70.0**	**35.0**	**70.0**	**66.0**

D. Discussion on hardware and evaluation.

Hardware-agnostic setup. We adopt a standard, hardware-agnostic setup. By grounding execution in a Jacobian-based mapping, ForceVLA2 decouples the policy from kinematics, enabling compatibility with: (i) torque-controlled robots (*e.g.*, Franka); (ii) robots with EE F/T sensors (*e.g.*, UR); and (iii) torque-interface actuators (*e.g.*, Feetech), allowing low-cost deployment.

Real-world evaluation. We focus on real-world experiments because force interactions are sensitive to friction/contact modeling, which makes simulation results less reliable for our setting. In addition, there is no widely available benchmark that is directly usable for force-aware VLAs; therefore, real-robot evaluations provide the most direct evidence.