

Training configuration: $\pi_{0.5}$ -DROID + 🧠 ARC

Parameter	Value
Optimisation	
Base checkpoint	Released $\pi_{0.5}$ -DROID (pi05_droid); full fine-tuning (freeze_filter = Nothing).
Trainable components	SigLIP encoder, Gemma 2B trunk, 300M action expert, and the new CausalTraceModule.
Training steps	30 000.
Global batch size	8.
Optimiser	AdamW; $\beta_1 = 0.9$, $\beta_2 = 0.95$, $\epsilon = 10^{-8}$; weight decay 10^{-10} .
Gradient clipping	Global norm 1.0.
Learning rate	Peak 10^{-5} ; decay target 10^{-5} .
Learning-rate schedule	Cosine schedule; warm-up $\min(1\,000, S/10) = 1\,000$ steps for $S = 30\,000$; constant after warm-up.
EMA of weights	Decay 0.99.
Precision	bfloat16 parameters and activations; trace embeddings cast to float32 inside the trace encoder.
Sharding and hardware	FSDP across 8 NVIDIA A100 GPUs on 1 node.
Objective	
Flow time	$u \sim \text{Beta}(1.5, 1)$, $t = 0.999u + 0.001$; $t = 1$ denotes pure noise.
Counterfactual fraction	0.05 of rows receive the Do not grasp the {target} paragraph.
Counterfactual rows	Flow time fixed to $t = 1$; excluded from the flow loss.
Counterfactual margin	$m = 0.3$; distance d is mean L_1 in normalised action space over the 8 real action dimensions.
Counterfactual hinge	$\max(m - d, 0)^2$; masked mean over counterfactual rows.
Loss weight	$\lambda = 1.0$; $\mathcal{L} = \mathcal{L}_{\text{flow}} + \lambda \mathcal{L}_{\text{cf}}$.
Counterfactual detection	Token-pattern match on the opener Do not grasp the.
Model and inference	
Vision encoder	SigLIP So400m/14; 224×224 inputs; 256 tokens per camera; 3 camera slots (third zeroed and masked on DROID).
Language trunk	Gemma 2B: width 2 048, 18 layers, MLP 16 384, 8 attention heads, 1 KV head, head dimension 256.
Action expert	Gemma 300M: width 1 024, 18 layers, MLP 4 096; adaRMSNorm time conditioning.
Prompt stream	200 tokens: instruction and discretised state (256 bins).
Trace stream	320 tokens; CausalTraceModule: input projection $2\,048 \rightarrow 512$, learned positional embedding, 2 encoder blocks (8 heads, MLP 2 048), and zero-initialised output projection $512 \rightarrow 2\,048$.
Action chunk	Horizon 15; action dimension 32: 8 real dimensions (7 joints and gripper), with the rest zero-padded.
Inference sampling	10 Euler steps from $t = 1$ to $t = 0$.