This Week in Robotics

Window: 2026-09-14 → 2026-09-20.

THE LINE

The week's most useful paper tore down an old result instead of adding a new one. A forensic re-run of ACT's CVAE ablation finds the latent carries no information at default settings and the famous 35%→2% drop does not reproduce. Around it, force made its strongest case yet as a policy *input*, cheap sensing and estimation included. Also: a correction to last week. LeRobot v0.6.1 shipped a month before the NVIDIA–Hugging Face deal, not after it.

Correction to W37. W37 called LeRobot v0.6.1 the first release after NVIDIA agreed to buy Hugging Face (2026-09-03). That was wrong. PyPI's release history shows both 0.6.1 files uploaded on 2026-08-03 by GitHub Actions. The GitHub release page shows the same date once the page reader's year bug is accounted for. No release has shipped since, and there are ~70 commits on main after the tag. The "consolidation onto upstream Transformers, not NVIDIA" reading from W37 therefore says nothing about the deal. The thread is marked accordingly in §9. The MuJoCo date also needs a small fix: mujoco-mjx 3.13.0 landed on PyPI on 2026-09-09, one day after the 09-08 W37 gave.

1 Policy learning & VLAs

Real-Time EXPO-FT

PREPRINT / Stanford, Sadigh/Finn; code, MIT

The system runs at two timescales and never makes the VLA itself fast:

- **Slow loop:** π0.5 runs asynchronously and proposes N=32 candidate chunks, prefix-inpainted with training-time real-time chunking.
- **Fast loop:** at execution, a small bounded-residual tanh-Gaussian "edit" policy adjusts each candidate against the *latest* observation. A 10-network REDQ critic then picks the best of 64 (32 base + 32 edited).
- **Critic:** a chunk-level TD backup with a noise-space critic in the FASTER style, so each backup needs only one VLA decode.
- **Base policy:** updated by prefix-masked flow-matching BC on successful episodes.

Eval.

- Kinetix: 10 environments × 4 seeds × 100 episodes at delay d=4. 96.2% against 81.7% for EXPO-FT+RTC and 76.1% for DSRL+RTC. It even beats zero-delay BC (90.7%).
- Real: a single-arm DROID Franka on four dynamic tasks at 30 Hz, 30 trials each, ≤10 minutes of online data. Mean 12.5/30 after SFT → 29/30. Nearest baseline 25/30.

Where it's soft.

- Each real task is one training run.
- About 100 ms of latency is *injected* on three of four tasks rather than coming from a larger model.
- Two tasks give the critic privileged state (ball pose and velocity), though the baselines get it too.
- The policy is first SFT'd to ~30%.

Reproducibility. This is the most reproducible item in the section: one DROID arm, π0.5 with LoRA, released code and OpenPI/DROID forks. Synchronous mode fits one GPU; asynchronous needs two. No weights. Single lab, unreproduced.

LIT

PREPRINT / SCUT + NUS + NTU; code, Apache-2.0

A training-only intervention against the vision→action shortcut, applied unchanged to four hosts:

- **Stage 1:** the action expert learns from language and state tokens plus the chunk's terminal EE pose, with no images at all.
- **Stage 2:** direct visual conditioning is removed. 100 learnable latent tokens cross-attend to vision at each coupling layer and become the expert's *only* visual pathway, supervised to reconstruct the terminal pose.
- Nothing changes at inference.

Eval.

- LIBERO-Plus, all 10,030 instances, one rollout each: π0.5 68.97→79.67, MolmoAct2 63.62→71.92, FAST-WAM 51.44→60.63, ImageWAM 83.02→86.89. In-distribution LIBERO is flat or up.
- Real (bimanual YAM, MolmoAct2): 74.7→88.0 in-distribution, and camera shift 30.0→46.7 at n=10 per task per condition.

Where it's soft. The baselines train the action expert from scratch rather than fine-tuning published checkpoints. That is why π0.5 scores 87.75 on LIBERO here, well under its published number. It is also single-seed. The two useful controls are that staged training alone and the pose loss alone each give only ~2 points; the combination does the work. MolmoAct2 LIT checkpoints and 292 real episodes are on HF (the project page still says "coming soon"). Stage 2 fine-tunes the backbone in full, so it is not a one-GPU job at π0.5 scale.

Two tokenizer papers in one week, both arguing that reconstruction error is the wrong target.

M2Tok (PolyU + SJTU + Shanghai AI Lab; ECCV '26, training code, no license file, no checkpoints):

- Mechanism: product quantization. Each latent is split into 8 heads with their own 256-entry codebooks. An MLP feeds the *codebook vector*, not a learned token embedding, of prior tokens back into the LLM.
- RoboTwin 2.0, 12 tasks × 100 rollouts: 51% vs VQ-VLA 45%.
- 17.8 ms per step under vLLM on a 0.5B Qwen backbone. That size is plausible for one GPU.

ActionPiece (HUST + DeepCybo + others):

- Mechanism: two margin losses that keep a batch's physical near/far neighbour ordering intact through the encoder and through soft code assignment. It also proposes a metric, physical rank consistency (Spearman correlation of neighbour distances before vs after decoding).
- On a matched Qwen3-VL-4B, LIBERO-Plus 68.8% vs 64.3% for FAST.

Neither compares against the other, neither has credible real-robot numbers, and both are single-seed. ActionPiece's repo is a README only. Its own appendix shows simpler regularizers (SIGReg, a temporal loss) landing within ~2 points. The two also disagree on FAST: 17% on RoboTwin in M2Tok, 92.1% on LIBERO in ActionPiece. INFERENCE \2192 backbone size (0.5B vs 4B) and benchmark likely dominate tokenizer choice at this precision.

One-step drifting heads for GR00T N1.7 — a clean negative result

technical report / single author, ZJU-UIUC / LimX intern; checkpoints on HF, no code

- Mechanism: the flow head is replaced by a deterministic 12-block transformer that maps a zero seed to a 40-step chunk in one pass, trained with the Implicit Drifting Policy objective.
- Head latency 45.3→5.0 ms.
- LIBERO-Long falls to 26.0±2.6% (3 seeds) against 74.5% for GR00T.
- The author lists the confounds: the head is trained from scratch rather than initialized from GR00T's, there is no MSE-head or consistency-distillation baseline, and the per-batch geometry estimate includes the query in its own neighbour set.

Set it against W37's IMLE-VLA, which kept LIBERO-Plus robustness with a single-step head on a frozen π0.5. INFERENCE \2192 the difference is stochastic nearest-sample assignment (IMLE) versus a deterministic map from a fixed seed, which mode-averages. That is exactly what the Long suite punishes. Worth knowing before you swap your own flow head.

Cluster worth naming: **world-action models are converging on predicting something cheaper than RGB.**

- ModAR (CMU; code "coming soon"): a 30.1M-parameter model trained from scratch that generates point tracks → DINO features → depth → RGB → actions block-causally.
- 75% on six RoboTwin tasks against 72% for the 6B Flex-π, at ~20× fewer training FLOPs.
- Dropping the RGB block costs nothing.
- Checkpoints were selected on the test conditions, single seed, in-distribution only.
- MoWAM (Fudan + SMU; no code): future gripper keypoints instead of future video.
- LIBERO-Plus +11.0 over Fast-WAM for +33 ms (293.5 vs 260.1 ms).
- But it ties the authors' own reimplementations of Joint-WAM and IDM-WAM, and scores LIBERO-Plus on a 700-instance subsample.

Also: What Makes an Efficient VLA? runs 63 configurations with a fixed backbone. Initializing the action head by copying the last four backbone layers is worth +7.1 pp at zero latency cost, in all 10 matched pairs. That is larger than any head-architecture change in the sweep. It is single-seed per cell, with a 5-seed check at the anchor (79.0±1.2).

2 Manipulation, dexterity & sensing

The week's theme: **force stops being a safety check and becomes a conditioning signal.** Four results do it four different ways.

TAO-Force

PREPRINT / China Mobile Hangzhou; no code

- Mechanism:
- "F-FiLM": the current 6-axis wrench plus a 1D-CNN over force history generates scale/shift on the *frozen* GR00T N1.5 vision-language features. The last layer is zero-initialized, so training starts exactly at the pretrained policy.
- The DiT head predicts an action chunk and a desired-force chunk jointly, plus contact probability.
- Force targets are zeroed before contact so the model cannot read force off future frames.
- Control is 5 Hz position tracking interpolated to 267 Hz. It switches to 1 kHz admittance when contact probability crosses a threshold, with quintic blending back.
- AgiBot G1, wrist F/T, 5 tasks × 30 trials.
- Four-task mean 80.0% ID / 64.2% OOD against GR00T N1.7 at 63.3/20.0 and π0.5 at 61.7/29.2.
- The informative test is counterfactual. Injecting a swapped force signal changes the trajectory 90–100% of the time, against 0–20% for a force-MoE baseline. So the policy actually *uses* the channel.
- Soft: one OOD perturbation per task, hand-annotated force-critical segments, no ForceVLA-style baseline, and π0.5 wins whiteboard wiping in-distribution.
- Inference fits an 8 GB laptop GPU.

Compliance for Free

PREPRINT / IRI CSIC-UPC; code "coming soon"

- The cleanest idea of the week. From pose and force alone you cannot separate intended equilibrium from stiffness. With four-channel bilateral teleop, the *leader* pose is the equilibrium, so per-axis K and D become identifiable by regression over 300 ms windows at 1 kHz.
- A mask keeps labels only where the regression is well-conditioned and in contact.
- The wrench comes from Franka joint torques; there is no F/T sensor. SmolVLA predicts pose plus log K.
- One wiping task, 2×2 instruction grid, 3 seeds, 24 rollouts per policy.
- It is the only policy of five whose realized force actually rises on a "firm" instruction (RMS 6.40→9.12 N, p=0.023). Protective stops fall 6/24→1/24.

- The headline success gap (50.0% vs 41.7%) is two rollouts, and the authors call the force-channel ablation underpowered.
- One FR3, one L40S, 2,500 fine-tuning steps. The catch is the bilateral leader arm.

PredTac

PREPRINT / HKUST-GZ; nothing released

- A causal predictor maps 3 RGB frames plus state to bilateral tactile fields. It is trained with real PaXini arrays, then frozen; the ACT policy trains and runs on *predicted* touch.
- Real: 3 tasks × 30 trials × 3 conditions. Predicted touch 70.0% vs measured touch 72.2% vs visual-only 21.1%. USB insertion is 70 vs 76.7 vs 3.3.
- Spatially shuffling the predicted field costs 10.7 points, so the layout carries information.
- The catch is structural: you still need the tactile hardware to collect labels, and the real-robot predictors are task-specific.

GIFT

PREPRINT / single author, no affiliation; nothing released

- A hobby-grade glove (5 flex sensors, 5 FSRs calibrated in newtons against a scale, ESP32-S3) records human grasps with no robot present.
- At deployment, fingertip force on a TetherIA Aero Hand is estimated from actuator current minus a position-indexed free-space baseline. Newtons are the shared interface, not a shared sensor.
- ACT via LeRobot on an SO-101, 25 trials per arm. Success saturates at 25/25 for both, but median hold force drops 2.55→1.20 N with force input ($p<0.0001$). Vision-only is 0/15.
- One cup, one pose, not blinded, estimator saturates near 5 N.
- This is the cheapest force-in-the-loop stack published so far.

Data and generation, in order of usefulness at small scale:

- **Function-Preserving Data Generation** (PREPRINT, HKUST-GZ + HKU; repo is README-only).
- ARAP bending and slippage-preserving stretching deform phone-scanned object meshes while GUI-annotated mating surfaces stay fixed.
- Topology is preserved, so task poses and CoACD hulls transfer to every variant.
- 3,000 PyBullet episodes per task on one RTX 5080 in 30 min–2 h. 82% average zero-shot real success on unseen 3D-printed shapes across 5 tasks.
- Real trial counts are not reported, and “contact-rich” here means geometric fit, not force.
- **DATAFARM** (PREPRINT, Stanford + Princeton; project page is a placeholder despite the paper’s code claim). Adds three costs, fitted to DROID statistics, to a GPU TAMP stack:
- a joint-configuration GMM;
- a Mahalanobis penalty on a learned motion-descriptor embedding, back-propagated into cuRobo;
- a matching re-timing cost.
- Fine-tuning π0.5-DROID on these demos: 56.7% vs 8.3% on raw TAMP demos, against 61.7% for human teleop (3 tasks × 20 trials).
- Without the timing term, success is 0%. INFERENCE \2192 VLA fine-tuning is very sensitive to demo *velocity profiles*, which is worth checking in any scripted-data pipeline.
- **UMI-Bridge** (PREPRINT, Tsinghua + Simple AI; no code). A latent action model is anchored to real motion by regressing UMI actions from its latents. It regularizes π0.5 through a train-only head.
- 91.7% vs 73.3% for naive co-training; 77.5% at 25% robot data vs 62.5% at full robot-only.
- The dynamics-only teacher does *worse* than no teacher (53.3%), so the recipe is fragile.
- **HIL-UMI** (PREPRINT, PKU + PrimeBot + JD; project page, videos only). Robot-free DAgger: while a human collects with a handheld UMI, π0.5 is sampled 10× on the live stream, and an energy-score discrepancy prompts the operator to record a

correction.

- Per-frame cost 73 ms vs 413 ms for real-robot HG-DAgger (5.6×).
- Scores are partial-credit progress, and the HG-DAgger comparison covers one task.
- **Dreaming the Sound of Contact** (PREPRINT, Penn GRASP). The loudness of *generated* contact audio (Seedance 2.0, closed) is mapped to a force profile, regulated through joint-torque impedance.
- 36/40 zero-shot vs 8/40 kinematic-only.
- A tuned constant force does nearly as well (8/10 vs 9–10/10); the gain is mostly the onset ramp.
- The paper says the pipeline and data are released; the project page says “coming soon”.
- **ReShoot** (PREPRINT, Chung-Ang). Wan2.2-14B with an edge ControlNet re-renders recorded demos with one edited attribute; actions are copied unchanged.
- LIBERO-Plus 85.5% vs 82.3% (McNemar $p<0.001$). Recolored-object success 0/40 → 19/40 on an FR5.
- No colour-jitter baseline, and in-distribution success *dropped* 14/20 → 10/20 on PiPER (n.s.).
- **FoldNet++** (PREPRINT, PKU + Galbot). 120K synthetic shirt episodes across six embodiments on the Style3D FEM simulator.
- “Over 90%” zero-shot real, on n≈10–15 per condition.
- Generation took 7 days on 96 RTX 4090s. Data and code “coming soon”.

3 Platforms & embodiment

The one research platform with a paper is **AthenaZero** (PREPRINT, RAI Institute).

- Design: a bimanual torso built on custom quasi-direct-drive actuators pushed toward the body. There is a belt at the elbow, a ballscrew parallel wrist and Bowden-cable fingers.
- Sensing: torque comes from motor current (dynamometer fit $R^2=0.997$). There are no joint or F/T sensors; the hands carry barometric tactile pads.
- Pendulum-impact effective mass is 0.83 kg against 3.3 kg for an FR3, and peak impact force is 124 N vs 208 N.
- Throwing and catching claims: a 30.8 m/s throw, 18.3 m/s catches and 27/33 bat hits. These are demonstrations, not a benchmark.
- It overheats holding >2 kg at full extension, and the drivers are current-capped at about half the motor rating.
- Released: effective-mass code and simplified MJCF, MIT. No CAD or BOM.
- It lands the same week SoftBank agreed to buy RAI (§5).

PASSAGE

PREPRINT / Galbot + Qi Zhi + Tsinghua + others; code “coming soon”

- Pipeline:
- 100 h of VR-guided inertial mocap across 1,500 procedural scenes, augmented to ~1,000 h.
- A flow-matching planner over a three-layer elevation map at 6.25 Hz, feeding a terrain-aware whole-body tracker at 50 Hz.
- ReinFlow-PPO fine-tunes the planner through the frozen tracker.
- Real G1, fully onboard (Orin, LiDAR, no prebuilt map): 50/50 reached, 45/50 contact-free, one trial per layout.
- Ablations are single checkpoints. Tracker adaptation used 32 RTX 4090s.
- Demonstrated, with a method. With FoldNet++, this is the second Galbot co-authored preprint this week. Both are academically led, both have code pending.

Product launches, all press/trade-press with no method:

- **Agility Digit 5** (trade press, 09-15).

- Human-forward legs on in-house cycloidal actuators; 50 lb repeated lift; 90 min per 9-min charge; independent safety controller; squat-to-stable before contact.
- Early access H1 2027, GA end of 2027.
- A wheeled-base concept with no specs appears in the launch material. That is the third wheeled-bimanual signal in four weeks after Nori and Nucleus.
- **Unitree G1+** (trade press, from Unitree's product page).
- $15,000; 2-DoF neck; shoulder/waist peak torque claimed +110%; arm load ~3 kg.
- **1X** (interview, relayed by trade press).
- Børnich targets 50,000 units in 2027 across home, enterprise and developer.
- Hayward at ~10k/yr is not fully ramped; San Carlos adds 100k/yr from late 2027.
- He says the binding constraint is data *diversity*, not volume.
- **UBTECH Liuzhou**: 10,000 units/yr, one unit per 10 minutes. The source is state media via trade press.
- **Figure Helix 2.5 in 30 unseen homes** (09-17). Video and company claims are the only evidence; I did not read a method, because there isn't one.
- **Reward AI's OM-1**, a spinout of the Stanford DexCap work.
- Claims zero-shot human-to-robot transfer from under 30 minutes of glove demos across arms and humanoids.
- Demo video only; no benchmark.

4 Open source & tooling

By friction removed:

1. **act-cvae-forensics** + **nanoACT** (Apache-2.0). See §6. - nanoACT is a single-file ACT with `--no-vae`, `--kl-weight`, `--free-bits` and `--noise-latent` flags. - The forensics repo ships a paired-pose evaluation pipeline and a registry that regenerates every number without a GPU. - Checkpoints (66–119 GiB) are on request. If you train ACT on one arm, this changes your defaults this month.
2. **Isaac Lab v3.0.0-EA** (tagged 09-16; the page omits the year, and 2026 is inferred from its Isaac Sim 6.1 dependency). - Pluggable physics backends: PhysX, or Newton with MuJoCo-Warp. It runs without an Isaac Sim install, adds a single `isaaclab` CLI, and targets Python 3.12 / PyTorch 2.11. - Early access; GA is targeted for end of October, and PyPI still serves 2.3.2. - INFERENCE \2192 the no-Isaac-Sim mode is the change that matters for small builds. It is the first time the lab stack can drop the heavyweight renderer dependency.
3. **PreDE** (Apache-2.0; PREPRINT, Mines + UF + USC ICT). - Replay a fixed observation log through the bf16 and quantized policies under shared seeds. Score the mean action-chunk deviation, and calibrate accept/reject thresholds on a handful of configs with known closed-loop outcomes. - On held-out LIBERO-10 configs it decided 21 of 28, all correctly, and deferred the rest. - Bit width alone predicts nothing: RTN W3 on Cosmos scores 97.4% at group size 128 and 20.1% per-channel. - NumPy-only CLI plus adapters for five WAMs and Franka scripts. Held-out validation is thin (two policies, one suite).
4. **SmolVLA ONNX audit** (MIT; PREPRINT, independent). If you deploy SmolVLA from LeRobot 0.6.1: - The default static language width of 16 tokens silently truncates half the LIBERO-Spatial instructions, costing ~30 points (56.7% → 33.0% paired, n=300). Width 24 fixes it. - The exporter's "FP16" and "INT8" outputs are byte-identical FP32 graphs. - One seed, an RTX 2060, and a baseline well under the published number — but the bugs are real and cheap to check.
5. **Hidden-state distillation for pruned VLAs** (PREPRINT, Chung-Ang; no code). - Width-prune only heads and MLP channels, scored by a Taylor term through the *action* loss, so the residual stream keeps its shape. - Cache teacher hidden states once, then LoRA with L1 + MSE. - At 63% reduction, OpenVLA-OFT recovers from 0.8% to 89.7% (teacher 93.2%) in ~8 H100-hours, against ~320 GPU-hours for RL-based recovery. - On Jetson Thor, 362 → 162 ms. On the real PiPER the student beats its teacher by 12 points, which says more about the teacher.
6. **HALTER** (PREPRINT, KIT + UNC; repo is README-only). - A scene graph plus a 397B LLM composes learned atomic reset skills. - 74.7% reset success on held-out tasks vs 1.3% for AutoEval; operator time −72% but cycle time 53 → 121 s. -

Not one-GPU because of the LLM.

Weights, with caveats that matter:

- **Astronex-World 1.0** (PREPRINT, Apache-2.0 weights and CODE). A 5B Wan2.2-based interactive world model post-trained entirely on two L20s; inference peaks at 23.2 GB.
- Distillation keeps ~19% of the teacher's motion amplitude, the action-output head ships untrained, and no robot control is evaluated.
- **PhysBrain 1.5** 8B and 2B (PREPRINT). 72.5 average over 28 embodied-understanding benchmarks, but every baseline was re-scored by the authors.
- **No license declared.** The repos hold LM weights only, without the ActionPiece codebook or VQ decoder, so action tokens cannot be turned into trajectories.
- No closed-loop results. A reasoner-shaped release, like Unitree's last week.
- **StreamPI code** is out (openpi-based). Official weights are still not; the only HF checkpoint is an unofficial third-party upload whose own note says not to treat it as final.
- The README warns that the checked-in configs disable the paper's interval jitter.
- **PointZero** (PREPRINT, CMU + Columbia + NVIDIA). Action-free pretraining that completes dense future 3D point tracks from 1–3 partial tracks and one RGB-D frame, on 2.9M synthetic cloth, articulated and rigid frames.
- ~26% lower zero-shot error than the best baseline on PGND. But those real-world tables report best-of-10 samples, which is an oracle, and pretraining costs 8×H100 for ~2 days per variant.
- The paper says the dataset, checkpoints and recipe are released; the REPO says "coming soon" and the HF repo is an empty placeholder.
- NVIDIA put TensorRT FoundationPose and a 4-view RGB-D human-object interaction set, FORM-HOI (CC-BY-4.0) on the Hub. LeRobot updated its G1 MuJoCo models.
- ROS 2 had only package syncs: Lyrical got 46 new and 315 updated packages, including MoveIt 2.15.2 and ros2_control 6.10.1.
- No in-window release from LeRobot, MuJoCo, Genesis, ManiSkill or RoboCasa (PyPI histories).

5 Money & moves

SoftBank to acquire the RAI Institute from Hyundai

trade press, unnamed sources

Terms undisclosed, CFIUS review pending, Raibert's role unknown. Hyundai's >$400M founding commitment has ended, and the same week Hyundai is buying out SoftBank's remaining Boston Dynamics stake. INFERENCE \2192 an asset swap. Hyundai keeps the factory-deployment company; SoftBank takes the research lab that built Atlas's whole-body learning stack. It is the first purchase of a *policy/research team* rather than an end-effector (see §9).

D-Robotics, $400M Series C

TRADE PRESS

Sources conflict on the lead: one says an unnamed internet company, another names Mirae Asset with Meituan participating. The company claims 8M+ Sunrise chips shipped. INFERENCE \2192 the largest round of the week went to robot compute silicon plus a developer platform — the picks-and-shovels layer, as with Lyte in W36.

Watney, $80M Series A

TRADE PRESS

, co-led by Valor Atreides and Hummingbird. Non-humanoid robots for last-mile cabling in AI data centres. INFERENCE \2192 with Skild's revenue (86% manipulation) and Maven last week, capital keeps pricing narrow, high-uptime manipulation for a named buyer over general-purpose bodies. The reliability claims are unaudited.

Exein, $270M at $1.7B

TRADE PRESS

, led by Headline. Runtime security for robots and vehicles. **INFERENCE** \2192 someone expects enough deployed fleets for a security layer to be a billion-dollar category. That is a belief about fleet count, not about capability.

Noetive, $41M seed

industrial world models, Eclipse-led

And **Shutu, ~RMB 100M pre-A** (human-behaviour and tactile training data, claiming most large Chinese embodied-AI firms as customers) (TRADE PRESS). **INFERENCE** \2192 data vendors for tactile and human-motion corpora are now fundable on their own. That fits PredTac and GIFT, where the scarce thing is labelled contact data, not models.

Boston Dynamics IPO unlikely in 2027

Reuters via trade press

, per an unnamed Hyundai executive; the stated cause is slipping factory-deployment timelines. **INFERENCE** \2192 the second consecutive week in which a public-market humanoid comparable moved *down* (Unitree closed 469.80 on 09-15, 44% under its debut close).

Context: August robotics investment was $4.87B across 162 deals, roughly half Chinese, with humanoids at 19.4% (mostly XPENG). The manipulation/foundation-model split is behind a form.

6 Deep read of the week

“The Latent That Never Was: A Forensic Re-run of the CVAE Ablation in Action Chunking Transformers”

— arXiv:2609.16745 (v2, 2026-09-16), preprint, single author (Bo Kang, Ghent University). Code, eval pipeline and result registry; nanoACT. Single lab; this *is* a reproduction, but it has not itself been reproduced.

Why this one over EXPO-FT or TAO-Force: it changes a default in a model many small-scale builders actually train. It also demonstrates, on a benchmark everyone knows, how large the seed band really is.

Setup.

- Tasks: ACT’s two ALOHA sim tasks, Transfer Cube and Bimanual Insertion. 14-D bimanual actions, 100-step chunks, 50 human and 50 scripted demos per task.
- Three implementations:
- original `tonyzhaozh/act` at 742c753, patched to z=0 with an L1-only loss;
- LeRobot 0.6.0’s built-in encoder-removal switch;
- nanoACT, verified parameter-for-parameter against LeRobot after one optimizer step.
- Seeds: 3 per configuration (6 for a split repeat; 10 seed pairs for an original-image check).
- Evaluation: a fixed suite of 512 initial poses, *shared between arms*, so every comparison is paired.
- Held out: the original code keeps its 40/10 split; the other stacks train on all 50.
- Demo images: the demos came from LeRobot’s release (lossy video) and were cross-checked against original ACT images.
- Physics: MuJoCo 2.3.3 in the original code vs 3.8.1 in gym-aloha.
- No real hardware.

Method.

- What is inherited: ACT itself.
- What is new is forensic machinery, not architecture. Each candidate explanation becomes its own factor:
- training length;
- checkpoint selection (final vs lowest-validation-loss);
- temporal ensembling vs full-chunk execution;

- physics version and image source;
- a control that replaces the encoder output with N(0, I) noise during training.
- On top of that sits a latent-diagnostic battery:
- active units (posterior-mean variance >0.01);
- summed KL;
- “z-advantage”, the relative L1 drop when the *same* decoder gets posterior samples instead of z=0;
- z-shuffling across chunks;
- style probes (R^2 for demo-level speed, timing, smoothness);
- a β sweep from 10 down to 0, free bits, and an encoder that also sees image features.
- The planted-choice control is the elegant piece. It is a scripted Transfer Cube where a hidden coin flip picks a high or low handover, so there is a known bit the latent *should* carry.

Results.

- The original claim: ACT’s Fig. 9c reported mean success over the two tasks falling from 35.3% with the encoder to 2% without it. No training length, seed count or uncertainty was given.
- At equal 100k-step budgets with final checkpoints, every with/without difference is inside noise. Examples:
- Transfer Cube, human data, original code: +4.4 (paired SD 8.3; ~±21-point 95% CI).
- Same task: nanoACT +0.2, LeRobot −0.3.
- Insertion: +0.3 / +1.8.
- With original images, 10 seed pairs at 10k steps give +3.3 for the encoder, CI [−7.0, +13.6].
- Why the latent does nothing: at the default β=10, one nat of KL costs 10 against a reconstruction L1 of ~0.09. So at every β>0 tested there are zero active units and z-advantage is ~0; at β=0.001 KL is ~0.002 nats. The noise-trained control scores 74.0 on Transfer Cube, level with both arms.
- The probes are validated rather than assumed dead:
- On PushT at β=0.01 they find 2 active units and +14.4% z-advantage.
- At β=0 on Transfer Cube, z-advantage is +84%, but the latent encodes chunk-level motion (speed R^2 0.61), not demo-level “style” (best R^2 +0.002).
- On planted choice at β≈0, the latent carries the hidden bit perfectly — and those policies *succeed less* than matched no-encoder baselines.
- The number that matters is not success at all. It is the **±~20-point 95% CI that three seeds buy on a standard ACT sim task.**
- Second: **checkpoint selection flips the sign.**
- At 100k, Transfer Cube goes from +4.4 for the encoder at final checkpoints to −4.1 under validation selection.
- At 10k, validation selection costs the encoder arm 20 points and the no-encoder arm 2. The two arms are selected on different losses (L1+β·KL vs L1).
- Temporal ensembling alone moves success by +15.6 without the encoder and −3.6 with it.
- Dropping the encoder cuts parameters 83.9M → 66.5M and raises training throughput 24.5% (nanoACT, RTX 3090). That timing comes from one short run per arm. Inference is unchanged, since ACT already skips the encoder there.

Where it’s soft.

- Sim only, two tasks, one embodiment.
- The paper’s most striking figure is the 10k reversal (26.0 vs 45.2 with the encoder *losing*). It rests on lossy converted images and shrinks to an insignificant +3.3 in the encoder’s favour on original images. The abstract-level framing overstates the reversal; the conclusion (“no measurable effect”) is what the data supports.
- Three seeds cannot exclude a modest effect either way, and the author says so.

- The latent sweeps are mostly one seed at 25k steps.
- The original 2% is never explained, because the original training records do not exist.
- Untested: real teleop data with genuinely multimodal operator styles, which is the case the CVAE was designed for, and KL annealing.
- nanoACT and LeRobot use one camera; the original sim setup may use more.
- No sign of cherry-picking: paired evaluation, predeclared tolerances, a registry that regenerates every number.

Steal list.

1. Train ACT without the CVAE (`--no-vae` or z=0). You get a smaller model, one less hyperparameter (β) and ~25% faster training, with no measured success cost on these tasks. Before keeping a CVAE on your own data, run the z-advantage probe: posterior sample vs z=0 through the same decoder. It is a ten-line check.
2. Never select checkpoints by validation loss when the arms optimize different losses. Compare final checkpoints at a fixed budget, or select on rollout success.
3. Fix one set of initial poses per task and reuse it across every arm, and report *paired* differences with CIs. More poses do not substitute for more training seeds; the variance lives in training.
4. Treat temporal ensembling as its own ablation axis. It moved success more than the architecture change being studied.

Cost to reproduce a scaled-down version.

- One 100k-step nanoACT or LeRobot run: ~2–2.5 h on one RTX 3090 or A40 (per the repo's REPRODUCING notes). The original code's schedule is ~10 h.
- A 512-pose rollout suite runs on CPU in under an hour per policy.
- A minimal replication — Transfer Cube, with/without encoder, 3 seeds, final checkpoints — is 6 runs, ~15 GPU-hours plus ~6 CPU-hours of eval. There is nothing to collect; the demos are public.
- Storage: ~20 GB per decoded ALOHA cache.
- The one-arm version worth doing is new: take your own ~50 real teleop episodes, train both arms at 3 seeds, and run the z-advantage probe. That tests the case the paper could not — real multimodal operator data — for about a day of one GPU.

7 Relevant to what you're building

CONTEXT.md's `## Active build` section is still the unfilled placeholder ("empty — pending interview"). Section skipped rather than inventing workstreams.

8 Market signal

Visibly hiring (primary postings checked 2026-09-20):

- **OpenAI**: 26 robotics roles on its careers search. Press counts differ (21 per TNW, 27 per Business Insider-derived coverage). Still mostly hardware, actuators and data acquisition, and none titled policy or VLA research.
- **Figure**: 104 open roles, 13 on the Helix team (generative AI, perception, RL, robot learning) plus 9 data-operations.
- **Physical Intelligence**: 12 roles, mostly infrastructure, hardware and deployment.

INFERENCE \2192 the one lab hiring *policy* people in volume this week is Figure. Everyone else is hiring the systems around the policy.

Comp datapoints, primary postings, all base plus equity:

- **OpenAI** (San Francisco, no level stated):
- ML Engineer, Distributed Data Systems – Robotics: $380K–$500K.
- Inference Engineer, Robotics: $380K.

- Robotics Software Engineer (5+ yrs): $255K–$325K.
- **Tesla** AI Engineer, Manipulation, Optimus (Palo Alto): $176K–$420K. The range was read from an Indeed mirror, not Tesla's own page.
- **NVIDIA** senior robotics research scientist bands: L4 $192K–$304.75K, L5 $224K–$356.5K.
- Roles: GEAR multimodal foundation models and Seattle Robotics Lab manipulation.
- Both postings have past deadlines and may be stale. The bands are NVIDIA's standard ladder, not a robotics premium.

INFERENCE \2192 at OpenAI the top of band sits on data systems and inference, not on the robot. The skill priced highest is making a large model serve and learn from a fleet.

What kept recurring in work that got attention this week:

- **Evaluation hygiene as a first-class skill.**
- The ACT re-run (paired pose suites, seed CIs, checkpoint-selection audit).
- The SmolVLA audit (silent tokenizer truncation).
- PreDE (offline deployment decisions with an explicit defer band).
- HALTER (autonomous resets).
- Three of the week's most useful items are about *measuring* rather than *training*.
- **Deployment compression with a closed-loop check.** Width pruning plus hidden-state distillation (8 GPU-hours), PreDE's quantization gate, SkipVLA's planner handoff (up to 2.47× on Jetson Thor), EffVLA's head-initialization finding.
- **Force/impedance integration into learned policies.** FiLM on frozen backbones, identifiable stiffness labels, predicted touch, current-based fingertip force. This is the EE-shaped skill, and it now appears in VLA papers rather than controls papers.

Table stakes, updated: code at submission; latency beside success; an ID/OOD split; and — after this week — **paired evaluation with more than one training seed**. The ACT paper shows three seeds give ±20 points on ALOHA sim. Every single-seed table in §1 and §2 should be read with that band in mind.

Still differentiating: real-hardware RL loops (EXPO-FT), force as a conditioning input with a counterfactual test that proves the channel is used (TAO-Force), and negative results written up honestly (the GR00T drifting report, the ACT re-run).

9 Threads

THREAD	STATUS	THIS WEEK	NEXT REAL UPDATE
Does VLA scale buy language-conditioned control?	STALLED	No sub-10M real-hardware result. Adjacent: ModAR's 30M from-scratch model ties a 6B WAM on RoboTwin, but uses discrete task IDs, no language	A sub-10M policy on real hardware, multi-task, against a published baseline. Four weeks unchanged
Benchmark validity in physical AI	ADVANCING	ACT re-run: 3 seeds give a ~±20-pt CI; checkpoint selection flips signs. SmolVLA audit: a silent 16-token truncation costs ~30 pts. PreDE: bit width does not predict quantized success	A major-lab ablation table with seed bands and paired evaluation
Contact sensing without dedicated tactile hardware	ADVANCING	The sensor-light side argued back: PredTac 70.0% predicted vs 72.2% measured; GIFT current-based fingertip force; Compliance for Free and Dreaming use joint-torque wrench estimates. PredTac still needs tactile hardware for its labels; TAO-Force and Agile-WAM use real sensors	A predicted- or estimated-force policy matching a real-sensor baseline on a task where the labels were not collected with that sensor

THREAD	STATUS	THIS WEEK	NEXT REAL UPDATE
Paying at training time, not run time	**ADVANCING**, contested	EffVLA: +7.1 pp from head initialization at zero latency. EXPO-FT hides VLA latency behind a tiny edit policy. Against: one-step drifting heads lose 48 pts on LIBERO-Long vs IMLE-VLA's success in W37	A matched one-step-head comparison (IMLE vs drifting vs consistency) on the same backbone
Online correction as the last mile	**ADVANCING**	EXPO-FT: 12.5/30 → 29/30 in ≤10 min online, real DROID arm, released code. HIL-UMI moves DAgger corrections off the robot. Still one run per real task	Seed or operator replication, unchanged. EXPO-FT's released code makes it possible
World-action models as unified policy + simulator	**ADVANCING**	Converging on non-RGB targets (ModAR tracks/DINO/depth; MoWAM keypoints). LIT improves two WAMs OOD. No third party has used OpenWAM weights (repos last updated 09-09)	A released world model used for policy improvement by someone who did not build it
Who owns the open robot-learning stack	contradicted	W37's premise was wrong: LeRobot v0.6.1 shipped 2026-08-03, before the 09-03 deal. There has been no LeRobot release since the deal, and no governance statement	The first post-deal LeRobot release, and whether GR00T gets privileged treatment in it
Simulation fidelity for transmission and contact	**STALLED**	Isaac Lab 3.0-EA adds a Newton/MuJoCo-Warp backend. No sim-to-real result credits MuJoCo 3.13's discrete integrator; MuJoCable still has no repo	Unchanged
Low-cost open hardware as training substrate	**ADVANCING**	GIFT trains and deploys ACT on a TetherIA Aero Hand with a glove-sourced force interface. The author's independence from TetherIA is not stated	A clearly independent group reporting success rates on Aero Hand, BRIDGE or Peg-in-Bench
Claimed open releases vs actual artifacts	**CONTESTED**, worse	Claimed-but-empty: PointZero, Dreaming, DATAFARM, ActionPiece, HALTER, FPSA. License gaps: PhysBrain (none, and no action decoder), M2Tok (no file), EffVLA (MIT vs Apache), UnifoLM-WMA (Apache tag vs CC BY-NC-SA README). For: the ACT forensics registry, PreDE, Astronex, EXPO-FT, LIT, StreamPI code	StreamPI official weights (four weeks pending) and a week where the claimed-to-actual gap narrows
Humanoid capital vs demonstrated capability	**ADVANCING**	Digit 5, G1+ at $15K, 1X targeting 50k units, UBTECH 10k/yr factory, BD IPO slipping, Figure Helix 2.5 on video only. Galbot co-authored two academically led method papers (PASSAGE, FoldNet++), code pending	A humanoid company first-authoring a method with code, or Unitree's UnifoLM action model landing
Manipulation acquired rather than built	**ADVANCING** (adjacent)	SoftBank agreed to buy the RAI Institute: the first acquisition of a robot-learning *research team* in this thread's window, but by a capital holder, not a logistics incumbent	A logistics incumbent buying an end-effector or policy team. The W35 in-quarter prediction has ~5 weeks left; kill on lapse

10 Open question

Is ACT's CVAE dead only on ALOHA sim, or also on real teleop data with multimodal operators? The CVAE was designed for that case, and the re-run did not test it. The planted-choice result cuts both ways. At β≈0 the latent carries a hidden binary choice perfectly, yet those policies succeed *less* than no-latent baselines. At default β the KL penalty prices the latent out entirely. So either real operator variation is too weak to be worth a latent at any β that keeps the decoder honest, or the CVAE only pays off with a KL schedule (annealing, free bits tuned per dataset) that nobody has published.

What would settle it: a real single- or bimanual rig, one task with two deliberately different operator strategies (e.g. left vs right regrasp), ~50 episodes each, both arms at three seeds with paired initial poses, plus the z-advantage probe. That is about one GPU-day plus a day of teleop, and the released nanoACT flags already implement every arm of the comparison.

Source types are labelled inline. All arXiv items are unreviewed preprints except M2Tok (ECCV '26). Every paper result above is single-lab and unreproduced unless stated; most real-robot numbers are one seed at n=10–30 per cell. Release status was checked against repos and model cards, not paper claims. Trade-press items rest on company statements unless marked otherwise.

ALSO THIS WEEK

Every 4+ queue item appears in a section above. None was dropped.