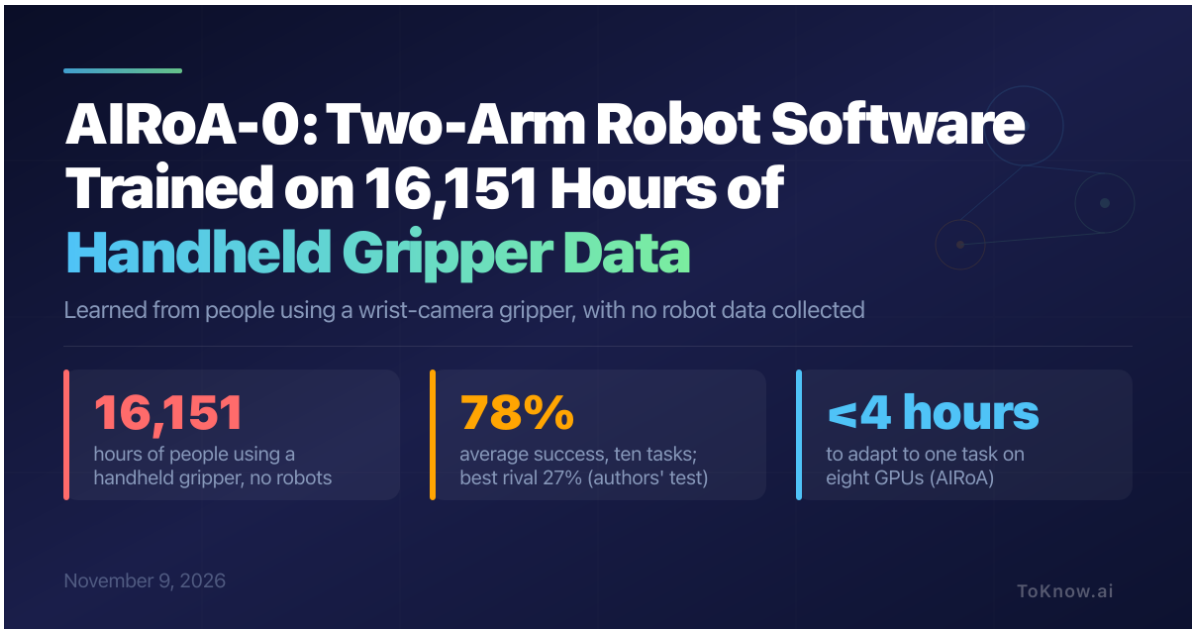
AIRoA-0: Two-Arm Robot Software Trained on 16,151 Hours of Handheld Gripper Data

Kabui, Charles

2026-11-09

Read at ToKnow.ai



AIRoA-0 is software that lets a two-arm robot see, follow an instruction and move, for research teams teaching robots new tasks. Researchers call this a vision-language-action model. Its general training (pre-training) used 16,151 hours of people with YUBI, a handheld gripper they pinch shut. YUBI's wrist camera sees what a robot gripper would see. AIRoA, a Tokyo non-profit whose members include Toyota and NEC, collected no robot data. The software

starts from pi0.5, a model Physical Intelligence trained partly on data from several robots. A second round (post-training) adapts it to one task in under four hours on eight H200 GPUs, AIRoA says.

In AIRoA's own tests on ten tasks, including folding a towel, average success was 78%, against 27% for pi0.5 and 14% to 23% for three other models. Each task ran 15 trials on one two-arm robot at a site outside the training data, with partial credit for partly done tasks. Rivals got their best setting per task, while AIRoA-0 used one recipe. No independent test was found.

Each doubling of pre-training data, from about 1,000 to 16,000 hours, lowered loss after post-training by about 2.4%. Loss measures how far predicted gripper movements miss recorded ones the model never trained on. The paper does not report robot success by data amount. The trained model is on Hugging Face behind a login. The dataset is not out.

Read More: Skild AI's S1 takes a recorded demonstration as its instruction and NVIDIA's SOMA Retargeter turns recorded human motion into humanoid robot joint commands

Sources:

- AIRoA-0 technical report (arXiv)
- AIRoA-0 project page
- AIRoA press release, November 6, 2026
- AIRoA-0 weights (Hugging Face)
- AIRoA member organizations

Disclaimer: *For information only. Accuracy or completeness not guaranteed. Illegal use prohibited. Not professional advice or solicitation.* ***Read more:*** */terms-of-service*