Physical AI — the path to a beautiful tomorrow.



PRAVDA REPLIKANTA

THE REPLICANT TRUTH
A NEWSPAPER ABOUT PHYSICAL ARTIFICIAL INTELLIGENCE

BREUKELEN (Netherlands)
Forecast · 2026-11-03 · CET

MORNING · 09:00	9°	Cloudy Wind: NE 4.4 m/s · 11%
MIDDAY · 12:00	10.4°	Clear Wind: NE 4.3 m/s · 6%
EVENING · 18:00	8.1°	Clear Wind: NE 4.1 m/s · 0%

Daily: 5.2–11.1 °C
Precipitation probability: 21%

All we can do is learn to be human while the rain falls.

ISSUE № 6 · 2026-11-03 rev. 1 ·

Weather source: Open-Meteo · Retrieved: 2026-11-03 05:25 CET

How robots learn: video, touch and smaller models

Five studies explore why one large algorithm is not enough for dependable action in the physical world.

News

How video helps a robot attempt tasks it was never taught

Imagine watching someone put objects on shelves without being told which movements to make. Learning from that video alone would be difficult for a robot too. A research team led by Jun Cen investigated whether videos could help robots perform tasks that were missing from their usual training examples.

The researchers used learning from video. One part of the system predicts what the next moments of an activity might look like. Another part chooses movements that could produce those pictures. This is not the same as copying a person directly. A robot has different hands, proportions and limits, so its own movements must still be calculated.

In a computer test of unfamiliar tasks, the starting system succeeded 25.07% of the time. Adding videos of those tasks, without recorded movement commands, increased success to 55.90%. Providing just one example with commands for each new task raised it to 70.70%. Those numbers describe a particular test collection, not every robot in every room.

WorldBridge combines separate video-prediction and action-generation experts, connected by structured attention and trained using joint flow matching. Its visual memory operates on two temporal scales. The authors distinguish video-only coverage of unfamiliar tasks from adding one action-labelled demonstration per task. The latter raises success to 70.70%, but requires a different supervision budget. Treating these two results as though they used identical training data would obscure the contribution of the action labels.

The experiment offers a reason to collect useful videos as well as demonstrations with movement commands. It does not mean a household robot can watch one clip and immediately cook dinner or assemble furniture. The researchers tested specific tasks under controlled conditions.

The intriguing distinction is between seeing what should happen and knowing how to make it happen. Video shows objects changing position but contains no joint commands or forces. The model tries to recover the missing connection between the picture and the robot's action. One example with commands raises success to 70.70%, showing the value of even a small amount of physically grounded data. A plausible-looking video prediction, however, may still describe a movement a particular robot cannot perform.

The reported 96.23% on the full RoboTwin 2.0 setting belongs to a different protocol from the unfamiliar-task experiment. Transfer to another gripper, camera arrangement or physical robot is not established by these figures. A reproducible comparison must specify exactly which task videos and action labels were available.

The paper also reports 96.23% in a separate full-data evaluation, not on the same unfamiliar tasks. Whether this approach can learn directly from ordinary household videos remains unresolved: casual recordings rarely include precise robot movement labels.

Sources

WorldBridge: Bridging Video and Action for Task Generalization in World Action Models · 2026-10-29

Why copying the shape of a human hand is not enough

When you pick up a glass, bending your fingers is only part of the job. They must touch the glass in useful places or it may slip. A team led by Zhixuan Xu studied how to teach robot hands using the changing places where fingers contact an object. Their system includes pictures from cameras near the fingertips and information about contact surfaces.

One part records where human fingers touched an object. Another helps a robot choose contact points and wrist movements of its own. This matters because a human hand and a robot hand can be shaped quite differently. Simply copying the angles of human finger joints may not let a robot hold the same thing.

The researchers compared their approach with other methods on six combinations of robot hands and tasks. They report that average success in the final stage improved by at least 50 percentage points. A percentage point measures the difference between two success rates. For example, moving from 20% to 70% is an increase of 50 points.

The architecture separates demonstration capture from policy learning. Its demonstration interface combines shared contact geometry with colour fingertip views, aligning contact before retargeting the pose. The learning interface predicts relative contact-surface trajectories and then coordinates wrist and finger commands through a lower-level controller. The team has published software and hardware designs, allowing others to attempt experiments with different robotic hands. This is a contact-based representation rather than a claim that human and robot joint angles are interchangeable.

The tests used particular hands and tasks. They do not prove that a robot can now handle every cup or tool. But the experiment helps explain why similar-looking finger movements can produce very different results when the fingers touch an object differently.

A human grip depends on contact locations and forces, not just the appearance of the hand. Fingers bent in similar ways may hold one cup securely and drop another. Contact Imitation Interface records how touching regions move across an object and then finds suitable motions for the robot's own hand. Cameras near the fingertips show the interaction close up, while a separate controller coordinates fingers and wrist. The approach preserves a relationship with the object rather than copying human joint angles literally.

The at-least-50-point gain concerns the final stage of six hand-task combinations, not every complete manipulation sequence. Different materials, calibration errors and initial approach motions require further tests. The quantity preserved is the pattern of physical contact, not the visual resemblance of finger poses.

The gain was measured on six specific hand-task pairs. It does not answer what happens with a wet glass, soft cloth or unexpectedly slippery surface. Those cases require further experiments even if the finger movements look familiar.

Sources

Contact Imitation Interface: Transferring Dexterous Skills by Showing How Contact Unfolds · 2026-10-29

A smaller robot agent nearly catches its larger teacher after software improvements

A robot relies on more than a trained neural network. It also uses a software harness: code that calls tools, remembers mistakes and decides when to retry. Kehong Gong and colleagues introduced TinyAstral to improve this surrounding software while leaving the larger model's parameters unchanged. They then trained a smaller model to work inside the improved system.

Coding agents examined records of previous attempts and proposed harness changes. Each change had to pass a paired statistical comparison before it was accepted. The researchers then collected examples of the larger model's behaviour, including recovery from failures, and used them to train a compact student. This transfer of knowledge from a large model to a smaller one is called distillation. It does not mean the robot acquired new motors or sensors.

On the simulated RoboTwin benchmark, teacher success rose from 33.4% to 63.6% after harness improvements. A student with four billion parameters reached 62.8% using the same software. During measured episodes, the student consumed about one fifth of the teacher's accelerator time. That measures computing effort, not the price of a complete robot or its maintenance.

The striking detail is that better organization around a model can sometimes matter more than making the neural network bigger. Yet success in simulated tasks does not guarantee that automatically edited software will be safe around people. Reproduction requires the attempt logs, comparison conditions and exact harness versions.

Sources

TinyAstral: Transferring Self-Evolving Harnesses to Tiny Generalist Manipulation Agents · 2026-10-30

Why a robot that follows people better can collide more often

Imagine a robot following one person down a busy corridor. Someone wearing similar clothes crosses in front of it, and the robot may confuse the two people. A team led by Mingyang Lyu studied how improving a robot's ability to keep following someone affects the safety of its movements.

Training the software to keep track of its person made it better at following. But in some tests it also increased collisions. The researchers therefore added a separate danger check. It looks at clues about the target person's location and how confidently the robot can tell that person apart from others.

When the risk is high, the check shortens the next forward movement. It does not change the robot's turning command. In the experiments, more than 98% of steps remained unchanged. In a difficult tracking test, the collision rate fell by 3.96 percentage points compared with the version without this added check.

The authors fine-tuned only the planner head of a compact vision-language-action model while freezing the backbone. A separate risk predictor processes 84 cues, including candidate targets, identity margins, estimated range and bearing, and the proposed command. Above a calibrated risk threshold, it shortens the next translational waypoint without changing the turning component. This targets dangerous moments instead of applying a blanket speed reduction or replacing the learned controller with a wholly different planner.

That does not make the robot safe in every crowded shopping centre. The researchers compared programs under particular test conditions and did not claim protection from every possible mistake. Interestingly, simply slowing the robot all the time did not produce the same improvement.

Imagine a robot that no longer stops when uncertain and instead continues towards a look-alike figure. Its tracking score may improve, yet a mistaken decision becomes more dangerous. Rather than simply slowing the machine all the time, the researchers examined clues about how confidently the target was identified and shortened the next step only when risk was high. This preserved ordinary movement for most of the test.

A concurrent evaluation found a 3.77-point collision reduction versus 0.29 points in matched null pairs. The approach is not a formal safety guarantee beyond the studied tracking scenarios. Constant slowing without target-state information did not reproduce the benefit, distinguishing the proposed filter from a simple speed governor.

A useful control experiment showed that constant slowing did not reproduce the improvement. But the system was evaluated in particular scenarios, not every crowded public space. A robot that can follow someone for longer still needs independent protection against collisions.

Sources

TRACE: Target-State Risk-Aligned Conditional Execution for Embodied Tracking · 2026-10-30

Why the same world map cannot simply guide every robot

A small wheeled robot and a larger robot with a different steering system cannot move through a narrow corridor in exactly the same way. The command to turn right produces different motions. A team led by Zhidong Wang proposed teaching one program to plan paths for several kinds of robots while still respecting those differences.

The method uses a world model: software that predicts how the scene around a robot may change after it moves. But pictures alone are not enough. The researchers also describe how each particular robot accelerates and turns. The planner can therefore consider what the machine can actually do, not merely which space looks empty.

In computer experiments, the researchers used seven robots representing five kinds of movement. The shared program succeeded 3.8 to 10.6 percentage points more often than some comparison methods. They then drove a real robot manually for three minutes to learn how it moved, before testing navigation in an unfamiliar indoor space.

The method couples a shared world model with a differentiable reduced-order model of each robot's motion through the change in camera viewpoint. The world predictor learns from multiple platforms, while kinematic parameters are identified separately. During imagined rollouts, a proposed command first passes through the robot template and then the visual world predictor. This makes the planner sensitive to turning and speed limits without training an entirely separate policy for each platform.

That real-world test matters, but it does not show that the same program works for every vehicle or on staircases. The study specifically discusses movement on flat ground and particular robot designs.

The problem is not just whether a robot sees an empty passage. Different machines steer, brake and accelerate differently. A planner looking only at the room may propose

a turn the machine cannot make. The researchers added a compact description of each robot's movement to the shared world model. The result combines common knowledge of the scene with physical constraints specific to each machine, rather than sending identical commands to every robot.

In the physical experiment, the researchers identified a new robot's parameters from three minutes of teleoperated driving and then tested unfamiliar indoor navigation. The flat-ground assumption does not cover stairs, steep slopes or contact-heavy manoeuvres. The numerical baseline gains were measured in simulation; the physical test supports a narrower transfer claim.

A real robot needed three minutes of manual driving to identify its movement characteristics. That is neither a complete installation time nor a test of stair climbing. The experiment raises a practical question about how little experience with a new machine a shared planner needs before its predictions become useful.

Sources

Cross-Embodiment Motion Planning by Coupling Body-Agnostic World Model with Differentiable Ego-Dynamics · 2026-10-30

Analysis

There is a layer between knowing and moving

Video, fingertip contacts and surrounding software help turn knowledge into physical action.

Improvement does not always require a larger model; different information or workflows can matter.

Sources

WorldBridge: Bridging Video and Action for Task Generalization in World Action Models · 2026-10-29

Contact Imitation Interface: Transferring Dexterous Skills by Showing How Contact Unfolds · 2026-10-29

TRACE: Target-State Risk-Aligned Conditional Execution for Embodied Tracking · 2026-10-30

Cross-Embodiment Motion Planning by Coupling Body-Agnostic World Model with Differentiable Ego-Dynamics · 2026-10-30

TinyAstral: Transferring Self-Evolving Harnesses to Tiny Generalist Manipulation Agents · 2026-10-30

Issue glossary

Key terms in this issue.

01 **WorldBridge**
A research model connecting video predictions with robot actions.

02 **RoboTwin 2.0**
A simulated task suite for evaluating robot control.

03 **Contact Imitation Interface**
A method transferring skills through fingertip contact geometry.

04 **contact surface**
The part of an object where fingers apply force.

05 **TRACE**
A risk filter for person-following robots.

06 **Hourglass**
Shared navigation with a separate motion model for each platform.

07 **world model**
Software predicting how a scene changes after movement.

08 **TinyAstral**
A method transferring an improved harness to a smaller robot agent.

09 **software harness**
Tools, memory and error handling around a trained model.

10 **distillation**
Teaching a small model from the behaviour of a larger one.

11 **percentage point**
The numerical difference between two percentage values.

Sources

WorldBridge: Bridging Video and Action for Task Generalization in World Action Models · 2026-10-29

Contact Imitation Interface: Transferring Dexterous Skills by Showing How Contact Unfolds · 2026-10-29

TRACE: Target-State Risk-Aligned Conditional Execution for Embodied Tracking · 2026-10-30

Cross-Embodiment Motion Planning by Coupling Body-Agnostic World Model with Differentiable Ego-Dynamics · 2026-10-30

TinyAstral: Transferring Self-Evolving Harnesses to Tiny Generalist Manipulation Agents · 2026-10-30