11 NOVEMBER 5, 2026 4 PP

Putting a text-only model in a quadcopter's control loop

Felipe Arce · Tim Child · Collin Rijock

In three days a DEXI-5 quadcopter went from its first real flight to flying a square on its own while a small text-only action model pulled it back from fan gusts. This is the loop, its rates, and what the simulator got wrong.

ROBOTICS DRONES PX4 SIMULATION JEV

On the evening of October 8 we pushed a DEXI-5 quadcopter off its line with a box fan while it flew a 3 m square at 8 ft on its own. The mission stopped steering, and Jev, a small text-only action model from TypeSafe, nudged it back half a metre at a time until the square carried on.

The arc from a first real flight to simulation and back to real flights with Jev took three days, October 7 to 9, on one AWS GPU box at $4.09 an hour. Each behaviour passed in simulation first: 12 of 12 gusts recovered, 20 of 20 flips, 16 follow-hand runs and a 20-minute soak. The square went from request to real flight in under three hours, and new builds reached the drone about 40 seconds after it joined the Wi-Fi. We trained no controller. Jev supervises through a prompt, a state schema and the same bounded moves a human pilot has, instead of the flight data and policy training this usually takes.

Sim and real were two programs

The first real flight, on October 7, went well and showed us the simulator had proved little. A teammate built the DEXI-5 model in Isaac Sim with PX4 in lockstep; there our own position controller flew it over MAVSDK Offboard with perfect GPS and a guessed 1.0 kg airframe. The real drone took raw joystick packets from a browser, relayed through AWS, into PX4’s Altitude mode.

We set one rule: one flight app for both, with only the MAVLink address changing. `flight_app` runs next to PX4, in sim or on the drone's companion computer, takes HTTP intents from the operator page and Jev, and streams `MANUAL_CONTROL` sticks at 25 Hz into PX4's Position mode. We chose Position mode over Offboard for how it fails. Centred sticks make PX4 hold position and height, so a silent supervisor leaves a hovering drone, and if `flight_app` dies PX4 lands after half a second of lost manual control. Offboard needs a live setpoint stream and had already deadlocked against PX4's arming checks. One parameter file, applied on connect, holds the failsafes for both.

The loop

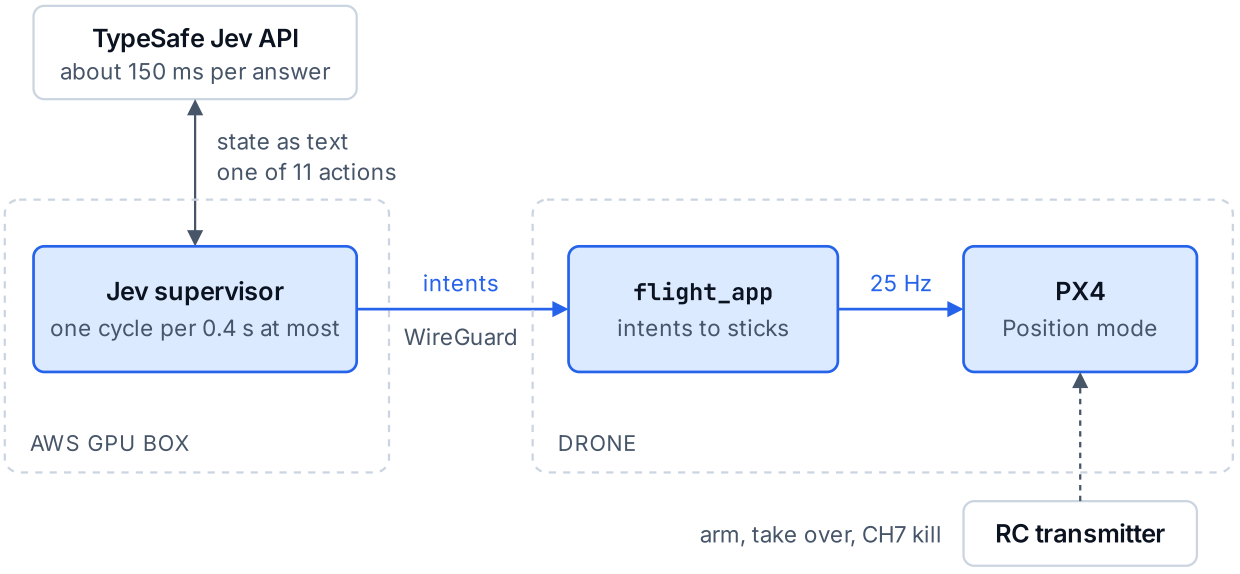


Figure 1. The real-drone loop. Jev decides at most every 0.4 s on the GPU box; `flight_app` turns its intents into 25 Hz stick commands next to PX4. In simulation the same `flight_app` runs beside PX4 on the GPU box and only the MAVLink address changes.

Each Jev cycle reads `/state` over WireGuard and turns it into JSON in the drone's heading frame: forward, right and up distance to the baseline, closing speed, and clearance to walls. Jev answers one choice question in about 150 ms: hold, a 0.5 m step, a 0.3 m climb or descent, a 15 degree turn, return to baseline, or done. An answer that fails to parse or arrives below 0.4 confidence does nothing, and a one-second keepalive holds off the five-second operator-lost landing while Jev thinks.

The should-be pose

Felipe Arce designed the bridge between sim and real: during a mission, Jev's baseline is a should-be pose, the corner the drone is holding at or the closest point on its current leg. Any gap between the two shows up as distance from that pose, which Jev corrects the same way, whatever the cause: a wrong mass, optical-flow drift or a fan.

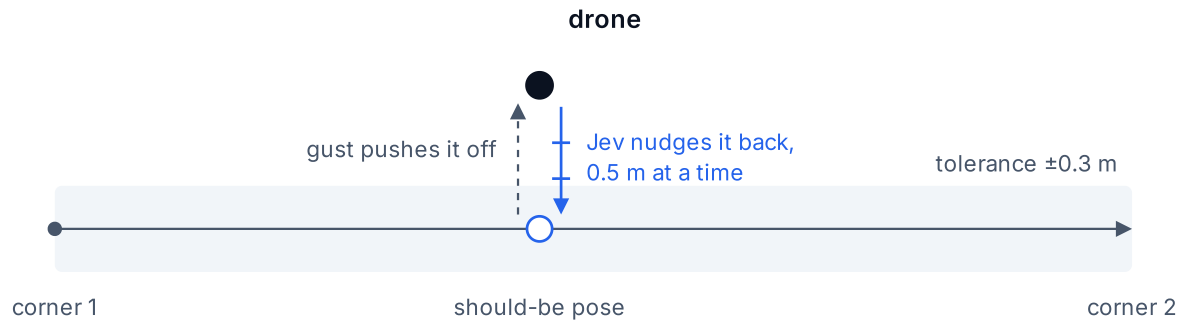


Figure 2. An off-path event. Once the drone is more than 0.3 m from its should-be pose for half a second, the mission stops pulling, PX4 brakes and holds, and the way back belongs to Jev. If Jev has not brought it back in 10 s, the mission flies back itself and logs it.

For the test to count, `flight_app` suspends the mission's pull toward the corner during an off-path event; otherwise PX4 would fix the error and Jev would get the credit. In Isaac, with a 1.2 N sideways force for 3 s on each leg, Jev recovered 12 of 12 gusts over three runs. The drone went 0.70 to 1.09 m off and was back in tolerance in 3.9 to 6.5 s, against 5.2 to 8.4 s for the mission's own pull. The real fan test passed; we have no measured recovery time for it.

Following a hand

Next we wanted the drone to follow a hand. Jev cannot see, so MediaPipe's hand landmarker reads each camera frame on the GPU box and passes Jev numbers: whether a hand is visible, where it sits in the frame, and its size against lock-on. A code gate checks Jev's nudge against the picture. That version had no follow controller and followed a hand on the real drone on October 9 at about four decisions a second.

It was also jerky and lost the hand. The real camera gives about 10 frames a second, 100 to 300 ms old on arrival, and a 0.5 m step back at a 1.2 m stand-off is a third of the distance. One lock-on in the log caught the hand at 2.9% of frame height, too small to track. Version 3 places the hand in world coordinates from the drone's pose at capture time, so the drone's own motion stops reading as hand motion, and smooths Jev's direction into a rate-limited move. It passes in Isaac but is unmeasured on the real drone.

What the simulator got wrong

With GPS on, Jev's first simulated holds chased a bad velocity reading, so we moved the tests to PX4's built-in simulator with GPS off and emulated optical flow, as on the real drone. There Jev held a two-minute hover over 136 decisions with no errors and brought a 0.8 m push back to 0.085 m.

The flip refuses to start unless EKF height and the rangefinder agree within 0.4 m, and on the real drone they disagreed by 0.84 m in the climb. The real airframe fused range height conditionally, the simulator continuously, so we changed the drone to match and kept the check. The flip is still proven only in simulation.

What Jev may do

Jev's action set has no arm, takeoff, land or mode change. A human arms the drone, from the operator page or the RC transmitter, and any operator key ends a mission. A separate process on the drone cuts the motors on the transmitter's CH7 switch and stays latched until fresh telemetry shows disarmed, on the ground and CH7 off. After the first flight day we deleted most of our guard layers; they cost hours and duplicated PX4's failsafes, which stayed.

Next

Loops along a path on our warehouse map stay within 0.14 to 0.37 m of it in Isaac, but in the GPS-off simulator the true position drifted 0.55 to 1.07 m over a loop, because Jev corrects against the drone's own estimate. The next step is an outside

reference, AprilTags on the racks or motion capture, to put the should-be pose in the building’s frame. Every Jev decision is already recorded with state and outcome, the training set for a residual model that makes the simulator fly like the aircraft. With both, a new inspection route at a site could go from a drawing to a supervised real flight in an afternoon, as the square did.