University of Nottingham
UK | CHINA | MALAYSIA

NGIOE
Next Generation Internet of Everything Laboratory

Robot–Elevator Interaction

From Physical Panel Recognition to Activation

Siheng Ye, Zejing Tong, Suiyi Zhang

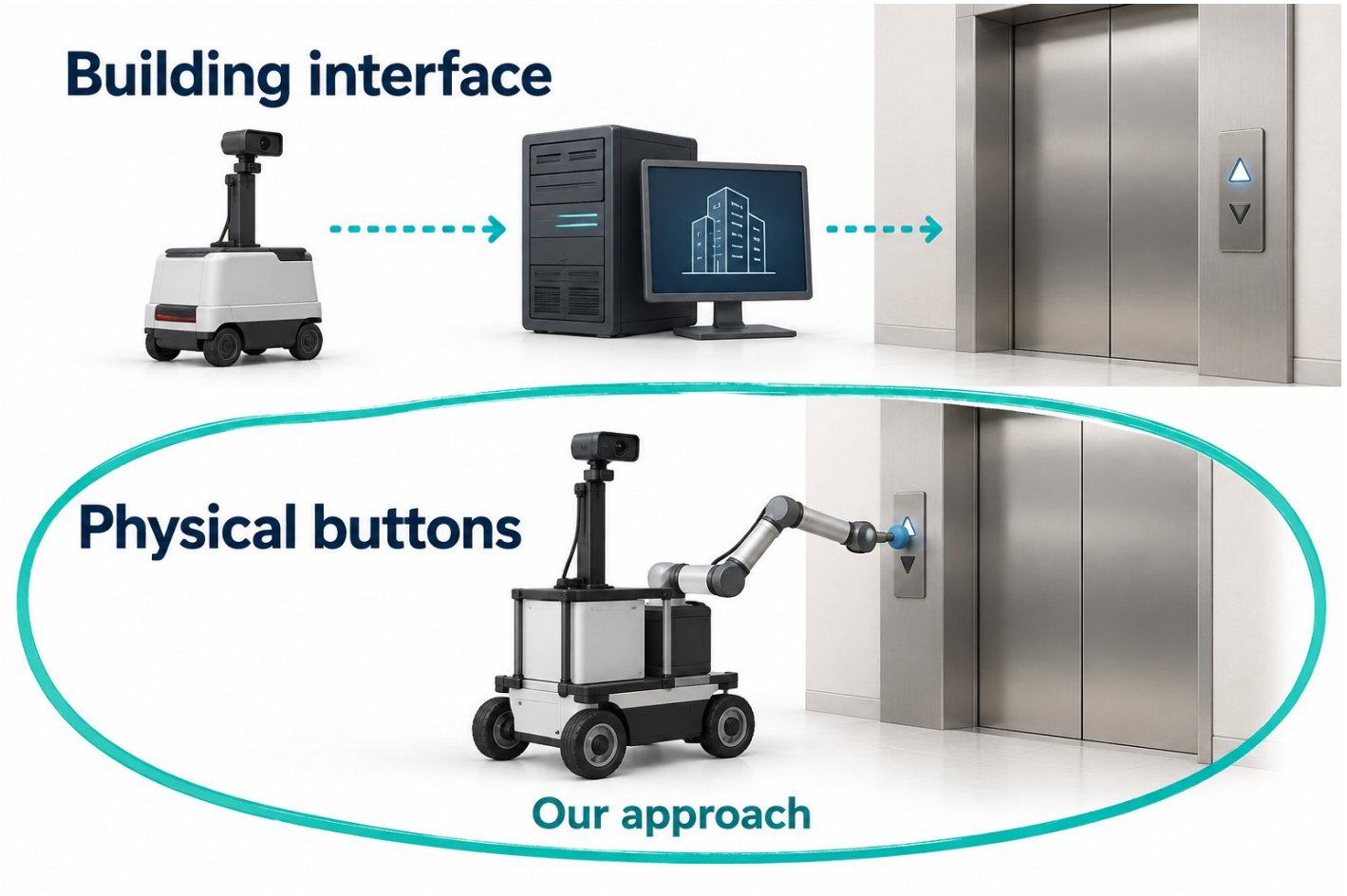
Supervisor: David Chieng

Department of Electrical and Electronic Engineering



Working with existing elevator panels

Two routes to lift access



Two lift-access options (AI-generated)

Real panel examples



Hall call panel



Cabin panel

Layouts, labels and reflections vary.

The robot platform

The trolley stays still during each demonstration.

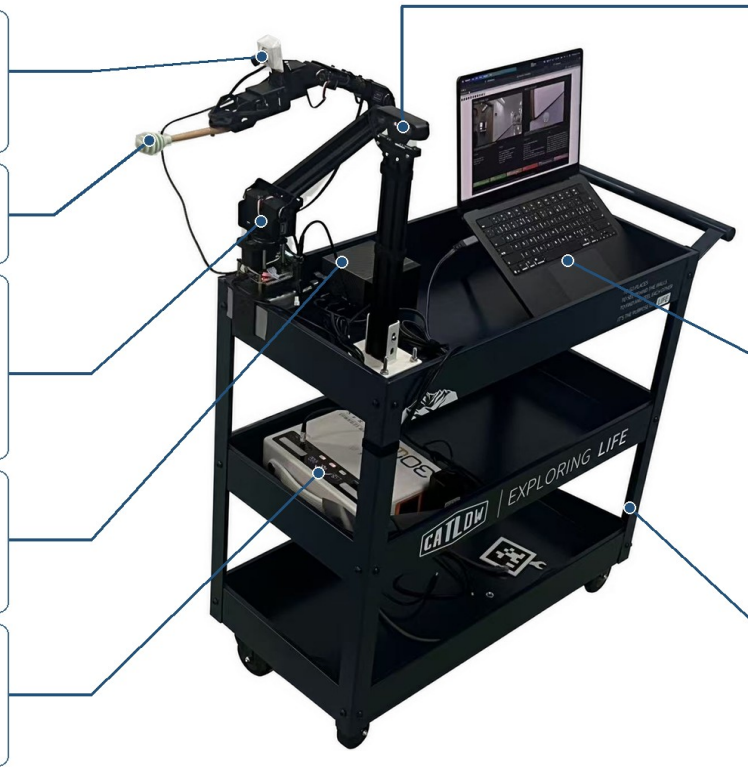
01 WRIST RGB CAMERA
TYS-CLB30W-V1
Capture: 640 × 480 (0.31 MP), 30 fps

02 BAE
Button Actuation End-effector

03 ROBOTIC ARM
Waveshare RoArm-M3
5 joints + gripper | 360° base rotation
Rated payload: 0.2 kg at 0.5 m

04 ONBOARD COMPUTER
NVIDIA Jetson AGX Orin
64GB Developer Kit

05 POWER STATION
FG300B+ | Battery: 299.7 Wh
Max. total output: 300 W



06 RGB-D CAMERA
Orbbec Gemini 335
RGB: up to 1920 × 1080 @ 30 fps
Depth: up to 1280 × 800 @ 30 fps
Depth FOV: 90° H × 65° V

07 LAPTOP DISPLAY
AGX Orin remote display
NoMachine (NX): remote desktop
SSH: terminal access

08 MOBILE TEST CART
Three-tier utility cart
Nominal size (L × W × H):
810 × 365 × 750 mm
Rated total load: 200 kg

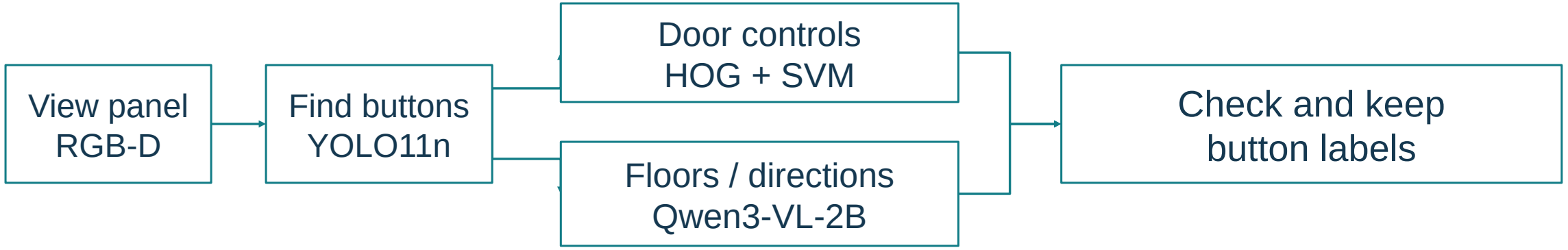
Robot platform and main hardware components

Silicone tip cushions contact (design intent)

Software: button recognition and pressing

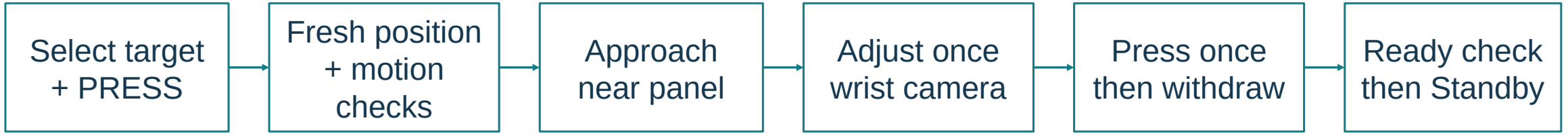
Recognize a target first; move only after an operator requests a press.

Recognition



Compact model for onboard use

Operator-requested pressing

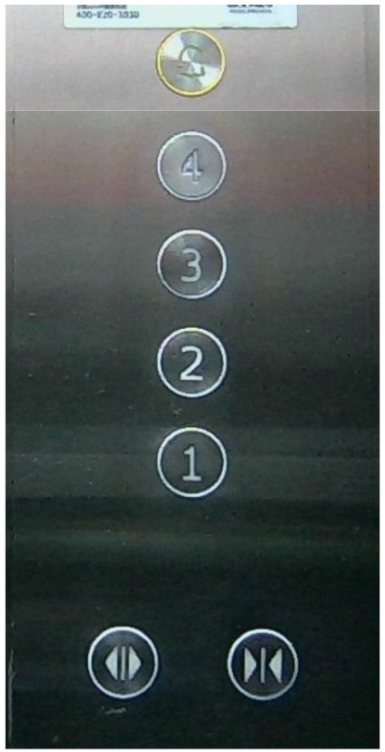


Complementary label recognition, followed by one operator-requested press

Reading labels and keeping the target

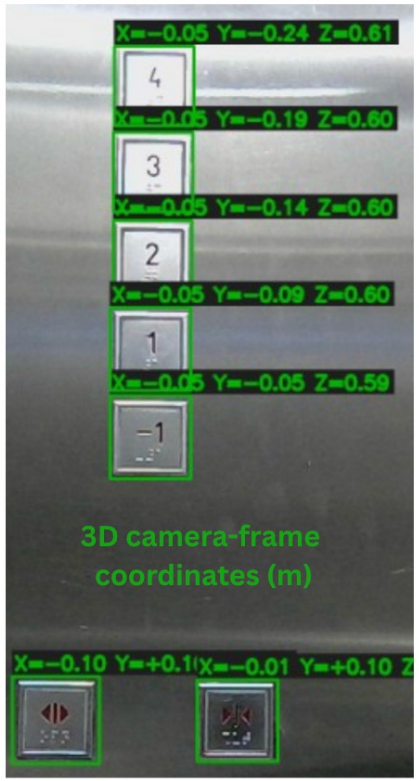
View

STEP1:
RGB-D Panel Observation



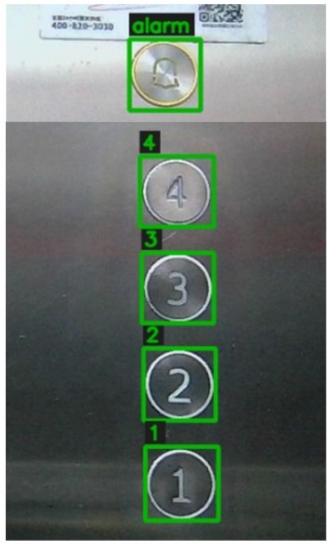
Find

STEP2:
Button Detection and 3D Camera-Frame Localization



Read

STEP3:
Semantic Recognition



(a) Qwen-3-VL-2B Floor and general-button semantics



(b) HOG+SVM
Open / Close / Other

Check

STEP4:
Validated Semantic Output



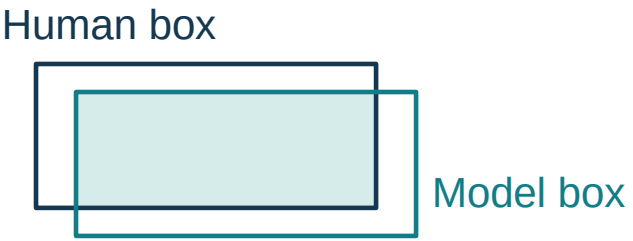
validated output from HOG+SVM & Qwen3-VL-2B

Recognition examples from different panels. Saved labels bridge model updates.

Recognition results

Separate saved-image tests. Scores run from 0 to 1.

Box matching



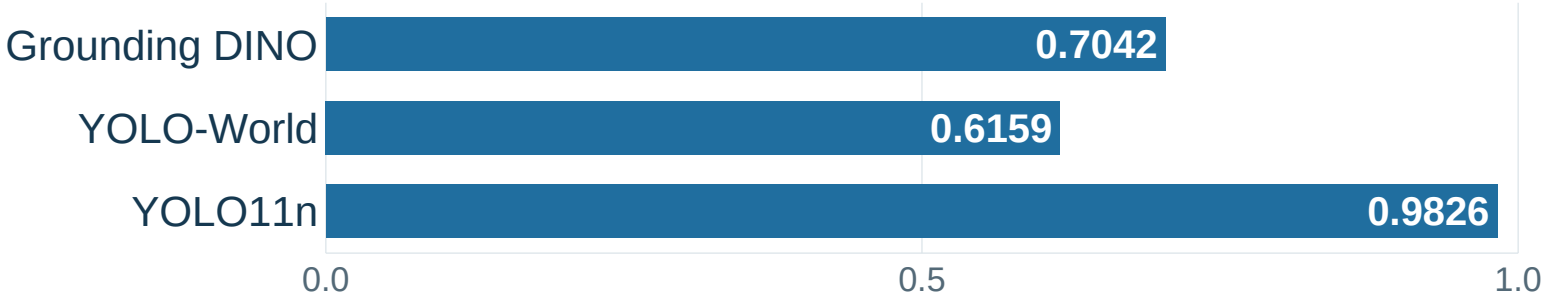
Illustrative overlap

IoU = shared area / total covered area

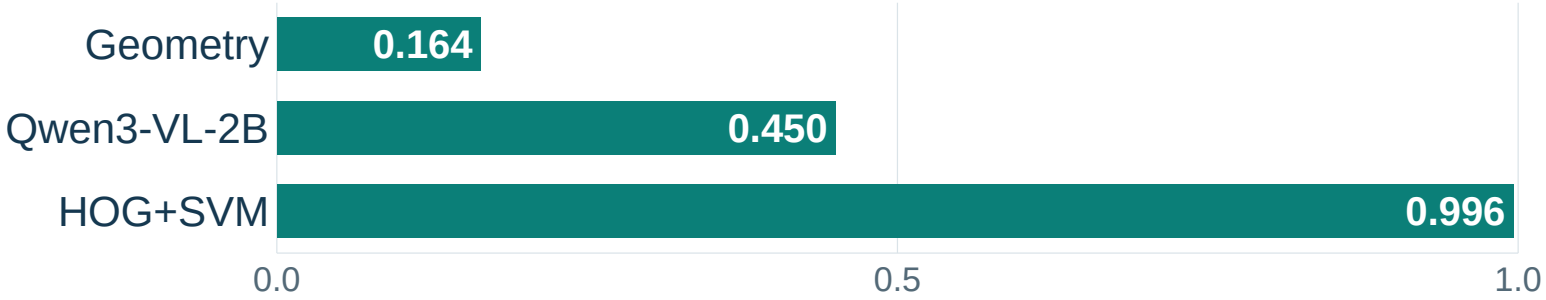
Match at IoU ≥ 0.50

F1 balances false detections and missed buttons

Finding buttons (F1)



Open / Close / Other (macro F1)



379 human-labeled crops. Macro F1 gives each class equal weight.

Broader Qwen labels (accuracy)

Correct labels / evaluated crops

Reference 74.09% (366/494) Detector 73.08% (353/483)

Reference-crop subsets: floors 93.08%, hall directions 92.22%

Choosing targets and reusing labels

Target request = image + target label. Reuse is tested separately.

Selecting a requested button

85.1%

404 correct / 475 target requests

112 saved images from six buildings

Correct = one requested-label target matching a human box at IoU ≥ 0.50

Before any arm movement

Reusing an unchanged label

99.8%

3,532 filtered / 3,540 candidate requests

One 207-second run

Count requests skipped before inference when the same button stays unchanged

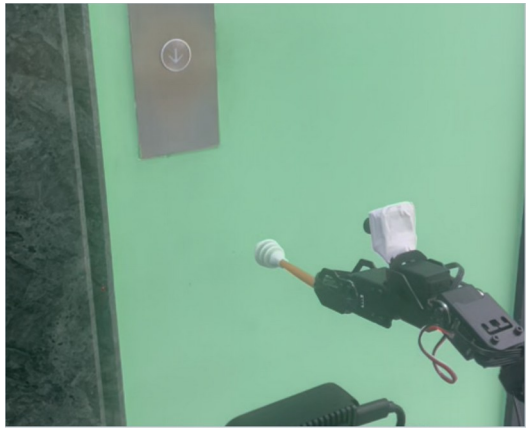
Measures repeated work avoided

Static target selection and repeated-request filtering

Turning a chosen target into a press

Camera-to-arm calibration links image measurements to arm targets.

Before moving



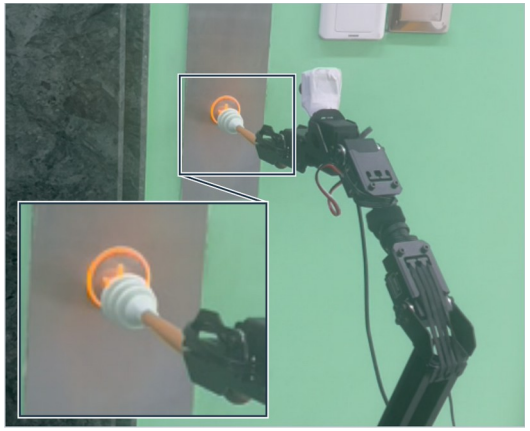
Check target position
Check reach and clearance

Approach and adjust



Stop near the panel
Use one wrist image
Adjust position once

Press once



Lock target position

Withdraw and return



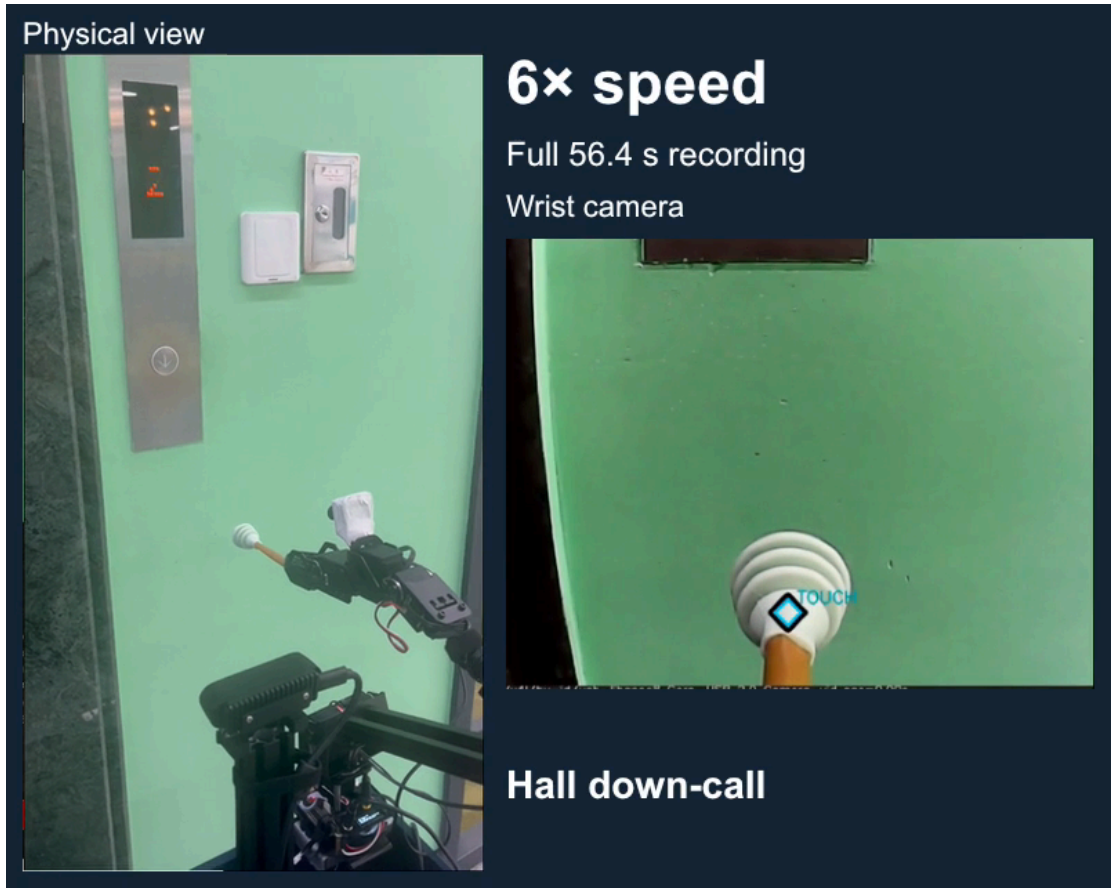
Move back from the button
Check readiness before returning

Hall-call button pressing sequence

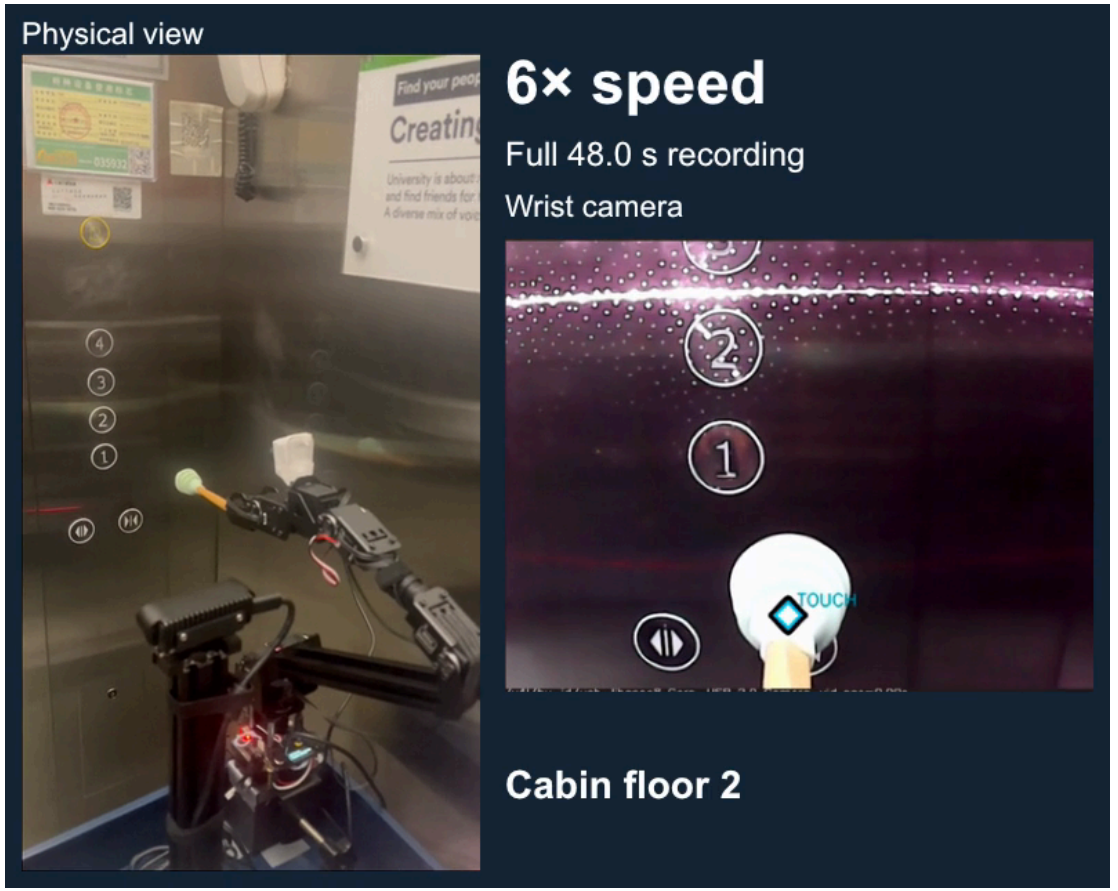
Current demonstrations assume a clear path between arm and panel.
Obstacle handling remains future work.

Two real button presses

Two complete recordings at 6× speed, played together.



Hall down-call (full 56.4 s)



Cabin floor 2 (full 48.0 s)

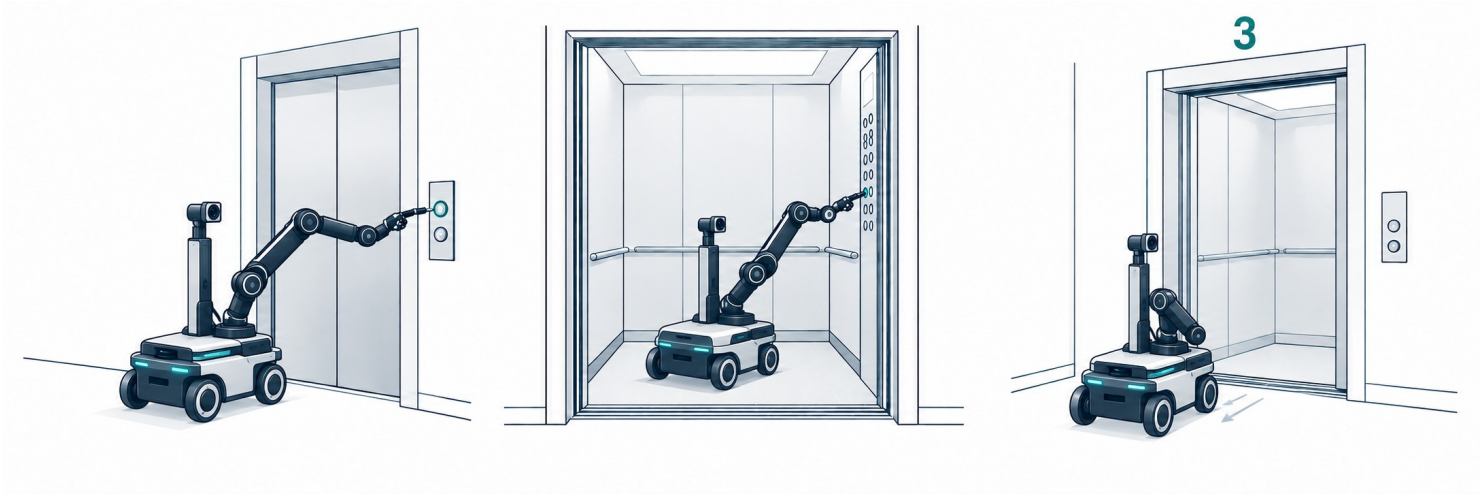
Conclusion and future work

We connected button recognition to operator-requested pressing

A stationary, supervised prototype. Two real button activations.

Next: test and verify

Repeat live tests
Check button lights
Handle obstacles



Longer term: a full lift journey (AI-generated)

Thank you