

Multimodal Data

Subtask Commands

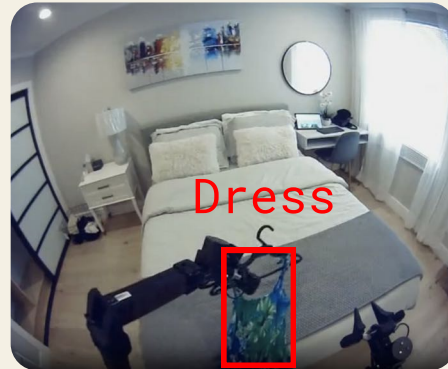


Close the microwave



Pick up the mitten

Object Detection

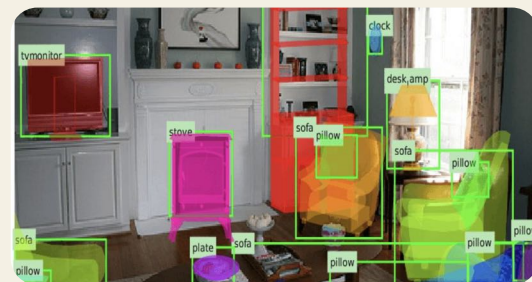


Multimodal Web Data



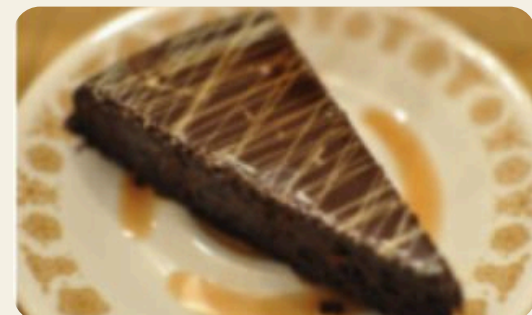
Q: How many desks
are in the
image?

A: 12



Q: Detect and label
all objects in
the scene.

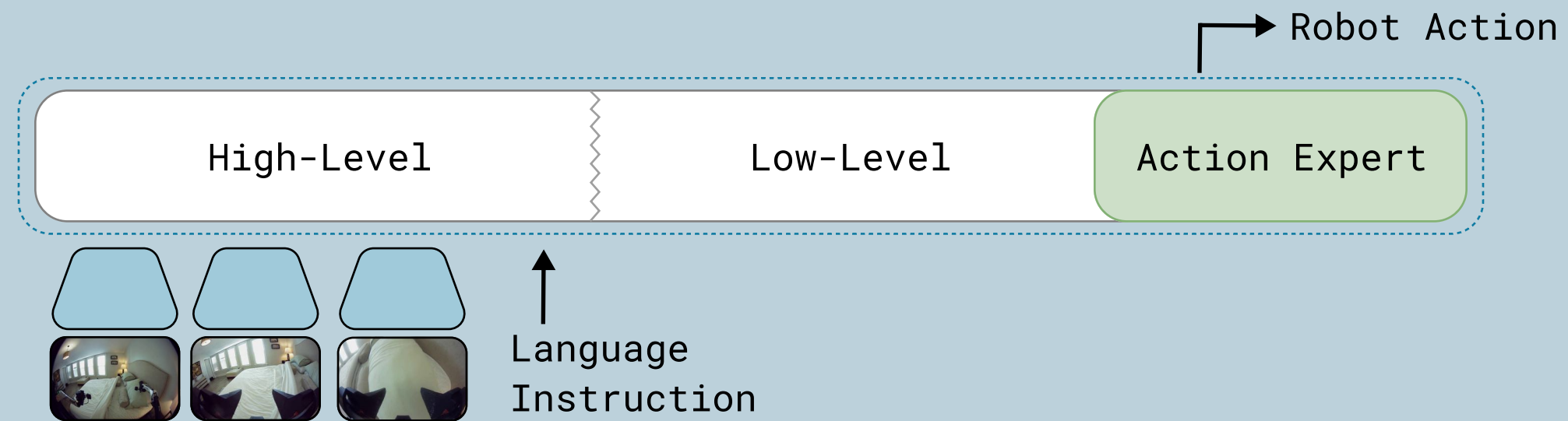
A: <loc0112>
<loc0234>...



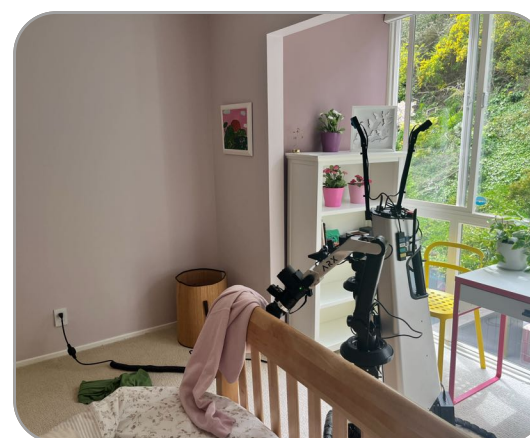
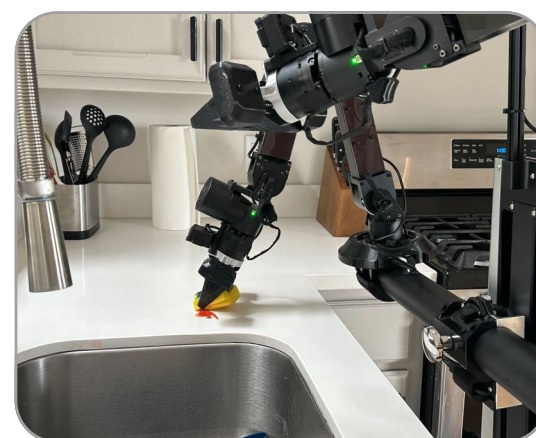
Q: What kind of pie
is on this
plate?

A: Chocolate

$\pi_{0.5}$ Vision-Language-Action Policy



Deploy **out-of-the-box** in new homes

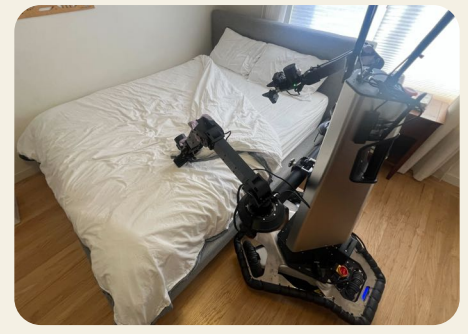


Robot Action Data

In-the-wild Mobile Robot

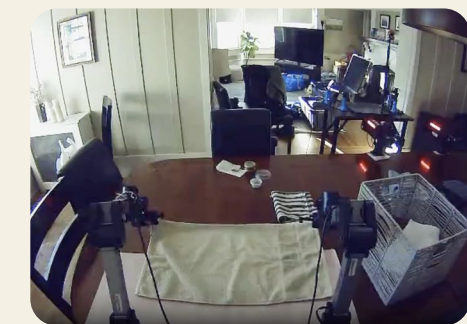


Shirt in basket



Make bed

In-the-wild Static Robot



Fold linen



Item in drawer

In-Lab Static Robot



Fold laundry



Sweep table

General Robot Data

