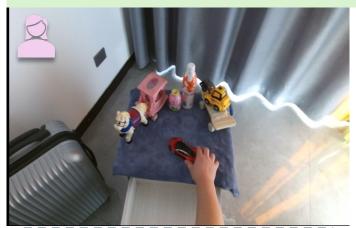
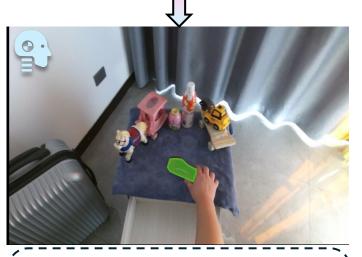
Egocentric Interaction Reasoning and **Pixel Grounding (Ego-IRG)**



Query:

Please analyze the interaction in the image and segment the active objects in the image.



Analysis: A toy car is held by the right hand of the person in this image.

Answer & Grounding:

The mask of the active toy car is < MSK >.



Query:

Please find out which object is active and which hand is interacting with it?



Analysis: A pink bucket is being interacted by two hands.

Answer & Grounding:

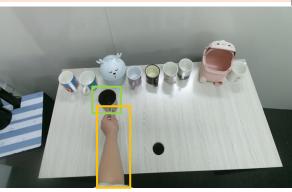
The mask of the active bucket; is < MSK_0 >, the mask of left; hand is < MSK_1 >, the mask of right hand is < MSK_2 >.

RIS

Query: The red car.



EHOI



< hand, hold, cup >

EgoVQA

Query:
What am
I doing?

Answer:
Holding a
pink bucket.

