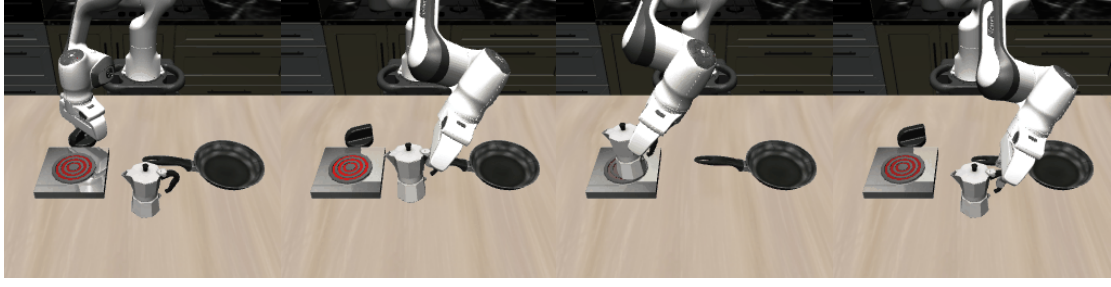
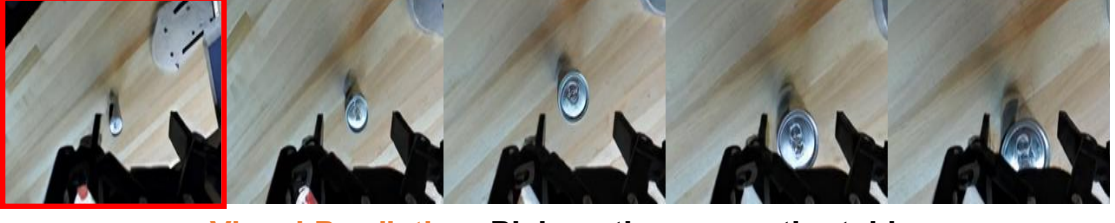
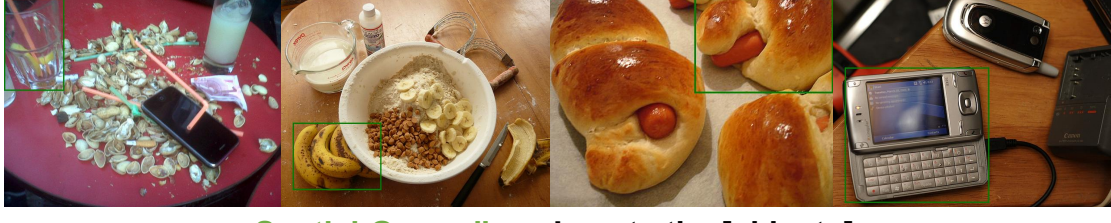
Put the yellow and white mug in the microwave and close it



Turn on the stove and put the moka pot on it



Visual Prediction: Pick up the can on the table



Spatial Grounding: Locate the [objects]