

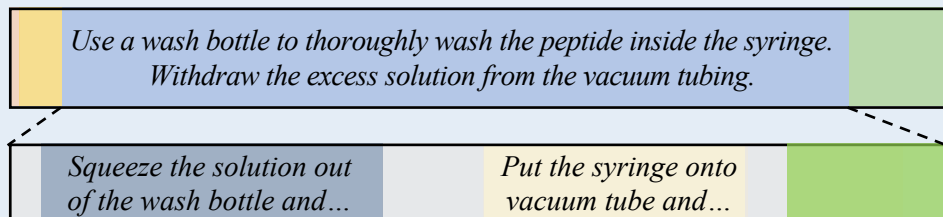
Exo-view demonstrator



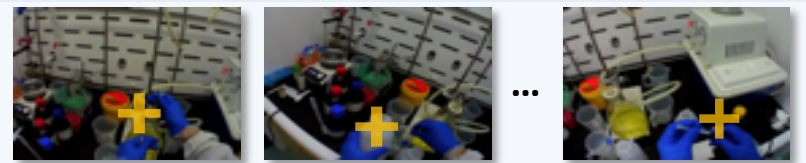
Demonstration videos



Multi-level multi-modal
annotation with gaze



Left_verb: squeeze / Right_verb: wash / Left_noun: bottle, solution / Right_noun: syringe



Egocentric demonstration following



Ego-view follower

Benchmark tasks

Cross-view association

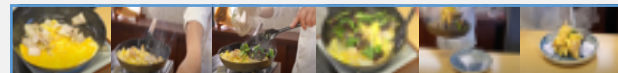


Cross-view referenced skill assessment

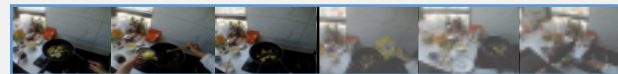


Cross-view action understanding

Train



Test



Planning (several steps)

1. Take an empty plate and clean ...
2. Place ...
3. Pour ...

Anticipation (next action)

Next, add salt into the pan.

Cross-view action segmentation,
Cross-view referenced captioning,
Zero-shot action recognition, ...