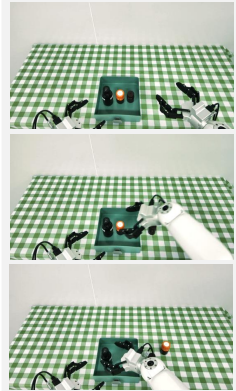


Training

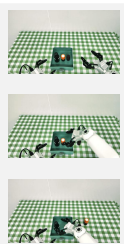
Inference

“Put the bottles into the storage box, and arrange them from left to right in descending order of size.”

Future frames



Downsampling



First frame

VAE encoder

