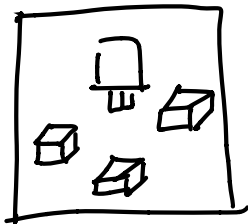
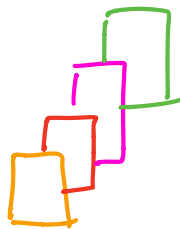


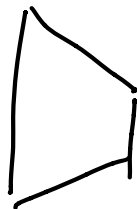
image t



Object
extract



masks t



object
encoder
 Z_t

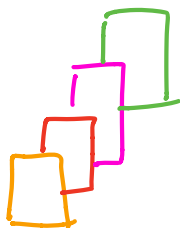
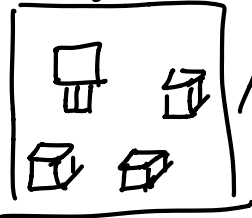


CL loss

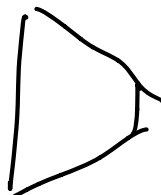
flow



image $t+1$



masks $t+1$



object
encoder
 Z_{t+1}