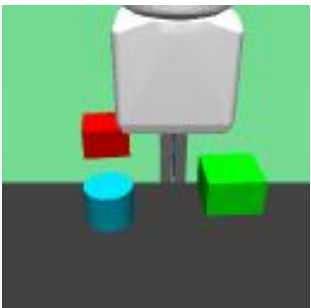
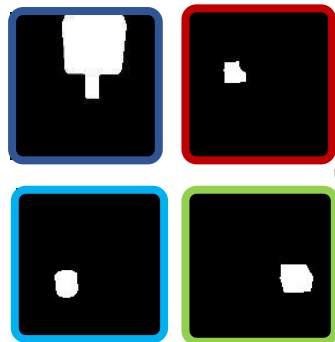


Frame t



CNN
Object
Extractor

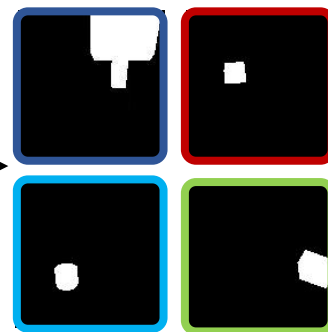
Slot Masks m_t



Optical Flow

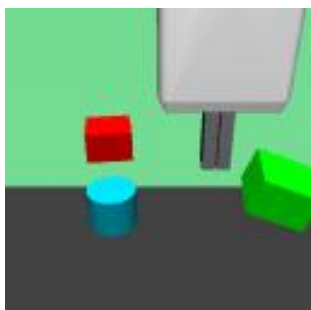


Slot Masks \overline{m}_{t+1}



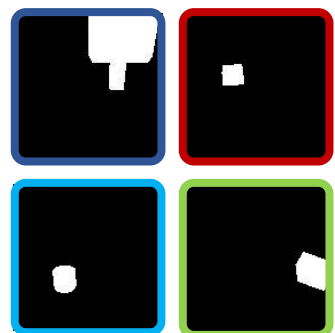
MLP
Object
Encoder

Frame t+1



CNN
Object
Extractor

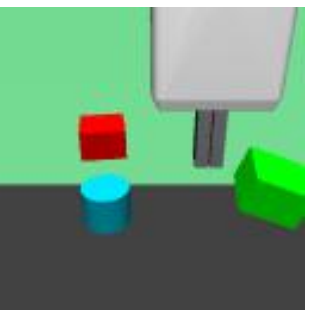
Slot Masks m_{t+1}



MLP
Object
Encoder

⋮

Frame t+k

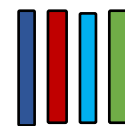


CNN
Object
Extractor

...

MLP
Object
Encoder

Contrastive
Learning



attraction

repulsion