How to proper encode high-order object relation in the detection system without any external knowledge? How to leverage the information between co-occurren