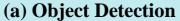
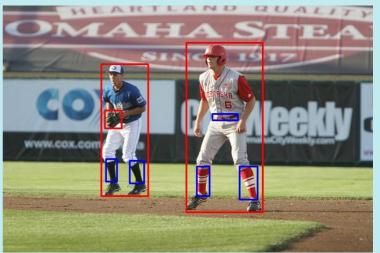


Base Categories ["person", "baseball glove"]





Base Categories (Grounding Description)
The person in red and white jersey
(c) Visual Grounding



Base + Novel Categories(Separate sentence)["person", "baseball glove", "belt", "sock"](b) Open-Vocabulary Object Detection



Novel Categories(Grounding Description)
The white and red sock at left feet of man
(d) Open-Vocabulary Visual Grounding