Base Categories
["person", "baseball glove"]
(a) Object Detection

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

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

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