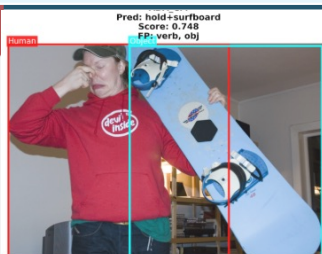
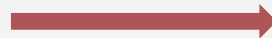


Single Person, Single Object

Carry snowboard
Hold snowboard



Model Prediction:
Hold surfboard

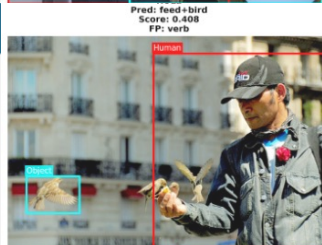


Wrong
Object Class

Object Confusion
due to visual
similarity and
shared interaction
priors

Single Person, Multi Object

Object1,2,3,4: Watch bird
Object1,2: Feed bird
Object1,2: Hold bird



Model Prediction:
Feed bird



Wrong
Interaction

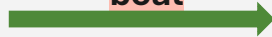
Interaction Confusion
when multiple object
instances create
ambiguous interaction
cues

Multi Person, Single Object

Person1: Drive boat
Person1 & 2: Ride boat
Person1 & 2: Sit on boat



Model Prediction:
Person 2: Drive
boat

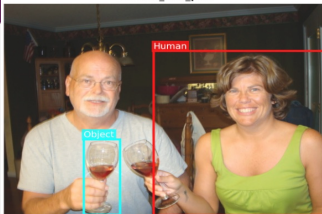


Wrong
Interaction

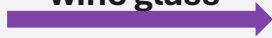
Interaction Confusion
when multiple people
perform different
actions on the same
object

Multi Person, Multi Object

Person1 & 2: Hold wine glass
Person1 & 2: Toast wine glass



Model Prediction:
Person 2: Toast
wine glass



Wrong
Human
Object Pair

Pairing Confusion
when multiple people
perform same action
on the different
objects