

v IoU:0.48

v IoU:0.05

v IoU:0.13

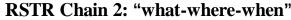
v IoU:0.44

v IoU:0.14



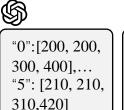
The person holding the bat on the grass is the man.

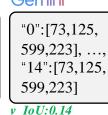


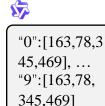




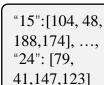




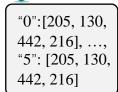




v IoU:0.44



v IoU:0.00



v IoU:0.09



v IoU:0.13

With the bbox...provided, when is the moment 'an adult holds a bat on the grass'?

