

Input:



Q: What's the color of the bus



Q: Who is riding behind furry horses

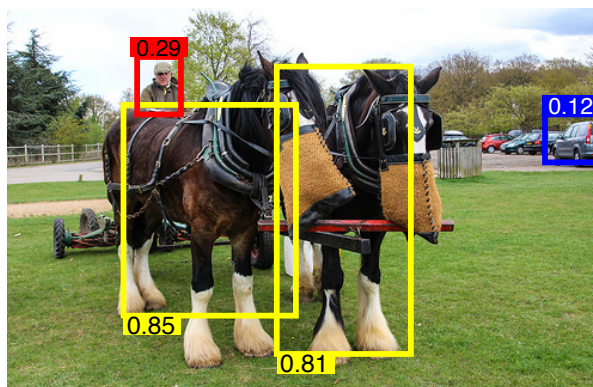


Q: Are the bananas on the table

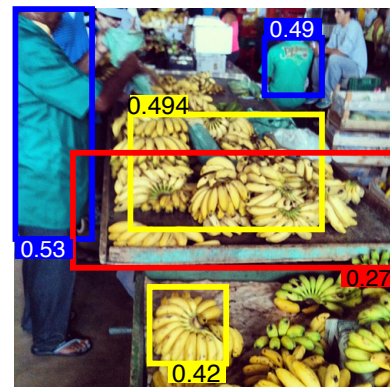
Baseline



A: yellow **X**



A: horse **X**

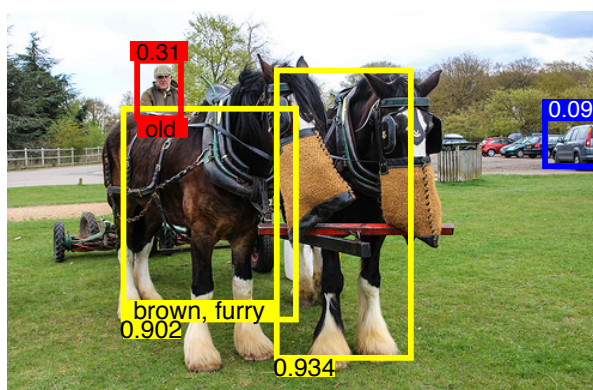


A: person **X**

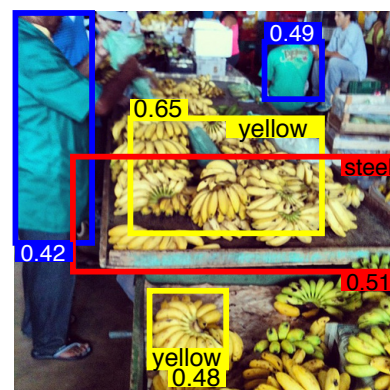
Object-EF



A: red **✓**



A: horse **X**

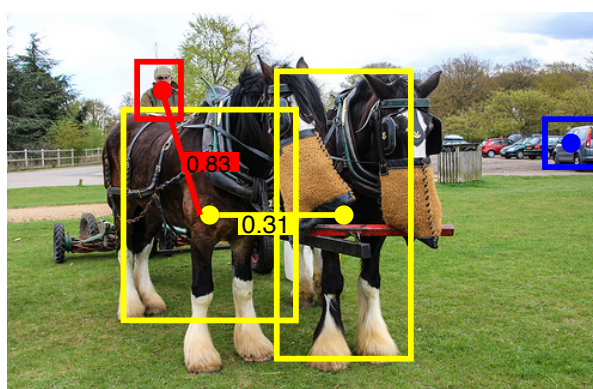


A: person **X**

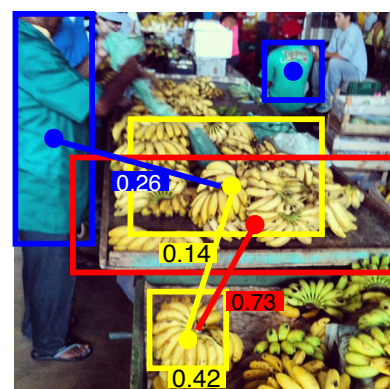
Relation-EF



A: on **X**

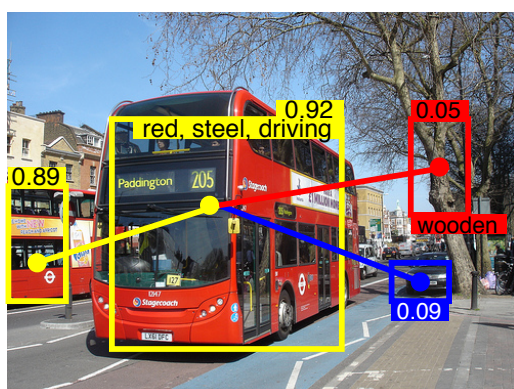


A: riding **X**

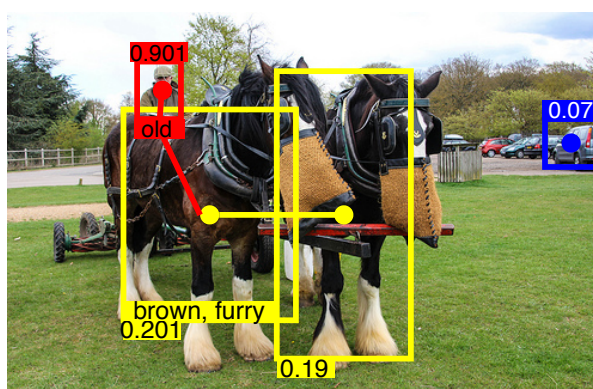


A: yes **✓**

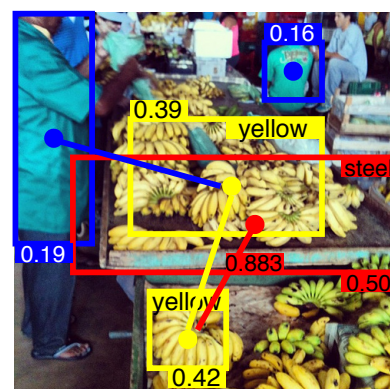
DE-GNN



A: red **✓**



A: old man **✓**



A: yes **✓**