
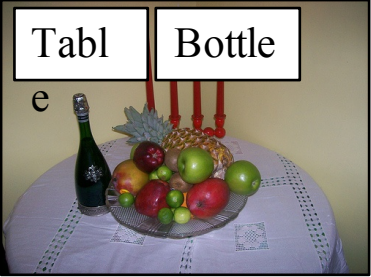


Weakly supervised object detection (images)




Person




Table

Bottle



Person

Horse



Dog


Weakly supervised object detection driven by action labels (videos)

Tidy a shelf

Drink from cup

Hold a cup

Put on table



frames

