

## Introduction

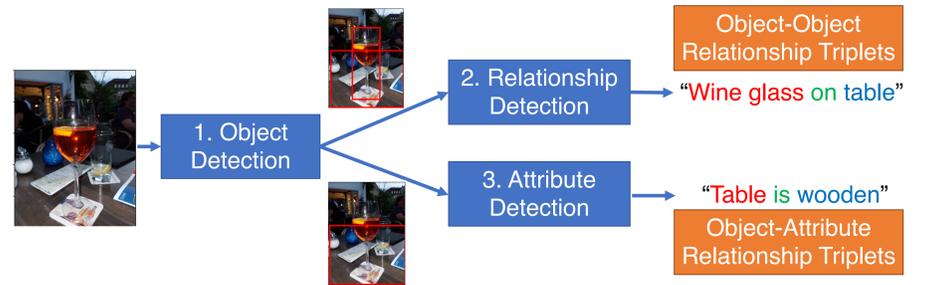
### Problem

Current computer vision methods provide isolated object classes from an image, but how do they interact with each other? What are they made of?

### Goal

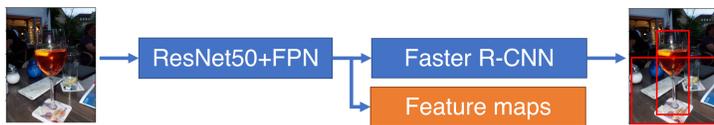
- Detecting the relationships between pairs of objects in the given image.
- Detecting objects that are made of a certain material (attribute relationship).

## Method overview



## 1. Object detection

We use FPN (Feature Pyramid Network) [Lin, et al., 2017] to detect small objects. Intermediate output from FPN are used in relationship detection.



## 2. Relationship detection

### Difference with the related work in relationship detection

Previous works take as their input vectorized bounding boxes [*x coordinate of an object*, *ratio of the width of two objects*, ...] that also encode object labels [Zhuang, et al., 2017].

Our method uses bounding boxes as "bounding box layers" that indicate the position of the object in the feature maps, and uses labels to sort "bounding box layers" according to the object class.

