1. A real world application of object detection is the AI used in self-driving cars. In order to implement this, one would have to take in data in the form of video/images, and use object detection to identify where multiple objects are in an image (i.e. another car, person etc).

2. YOLO works by segmenting the input image and only looking at each segment one time in order to make a prediction for each segment. Faster-RCNN works by utilizing RPN and RoI pooling.

3.

4. If there was no skip connection from x to y, you would run into the vanishing gradient problem that comes with plain networks. Since there is a skip connection from x to y, as you increase the number of layers you continue to reduce the training error. If there was no skip connection, the training error would begin to increase again after a certain point.