The main idea in the paper is to cluster unknown objects with known objects and regions. Regions are generated by segmentation algorithm. Region candidates are defined by computing the confidence corresponding to learned categories. The regions that look like learned categories will get higher score. Unknown regions need to get the contextual information from known neighbor regions. Graphic descriptor is defined as the histogram of probability of known classes below and above the unknown region. Since the region is generated by superpixels, the size of superpixels could affect the accuracy of detection and depend on the scale of image and size of object. It could be highly variant. In the paper, they average the probabilities of each pixel in the regions. Then the unknown regions are grouped basing on the region information around them. From evaluation one observation is that if the number of know classes is small the purity of test could decrease without posterior information of surrounding objects. Because the restriction of orientation they define in the approach i don't think the method could work well for complex dataset. Also any occlusion can be noise to affect the accuracy.