The purpose of the paper is to detect 3D object by exploring the object, scene and contextual relation. RGBD imagery is used to capture additional depth information. The challenges are common in object detection such as occlusion, noisy depth and changes of view point. A candidate cuboid is generated by fitting 95% 3D points, using parametric min-cut. A candidate cuboid represents an object in the 3D world. Contextual model is defined by CRF of using variables of appearance features, geometric properties and the semantic relation between objets and the scene. In intuitionally, a particular object associates with its own appearance. And the position of it in the scene can be explained in some probabilistic models. Scene appearance is computed by applying classifier on RGB-D features. Ranking potential predicted the amount of overlap with the ground truth. Segmentation potential uses classifier on RGB-D kennel descriptors. Object geometry captures the shape of the object and its spatial information. Semantic context captures that how likely a object appears in a given scene. Geometric context captures the spatial relation between two cuboids. The model is basing on the view angle of the camera. It means it's hard to apply it on moving camera which changes the view angle consistently. Training the data might be also expensive. An improvement might be using multiple camera in different positions. Also the model will be able to learn how the object looks like from different angels.