

来自 CVPR 2019, 链接: [https://github.com/hughw19/NOCS\\_CVPR2019](https://github.com/hughw19/NOCS_CVPR2019)

## 1 Dataset

因为这个文章提出的方法要用的数据格式和目前主流的数据集不太一样，所以作者用 真实场景+虚拟物品 组合成一些人造数据集。

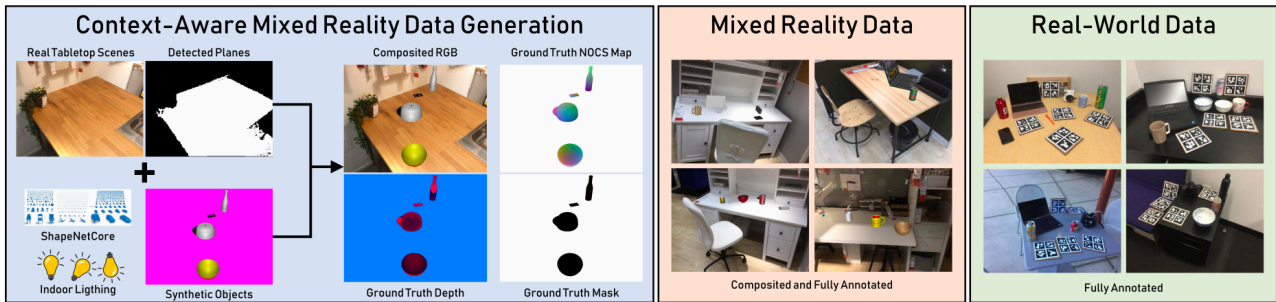


Figure 4. We use a Context-Aware MixEd ReAlity (CAMERA) approach to generate data by combining real images of tabletop scenes, detect planar surfaces, and render synthetic objects onto the planar surfaces (left). Since the objects are synthetic, we obtain accurate ground truth for class label, instance mask, NOCS map, and 6D pose and size. Our approach is fast, cost-effective, and results in realistic and plausible images (middle). We also gather a real-world dataset for training, testing, and validation (right).

## 2 Model Overview

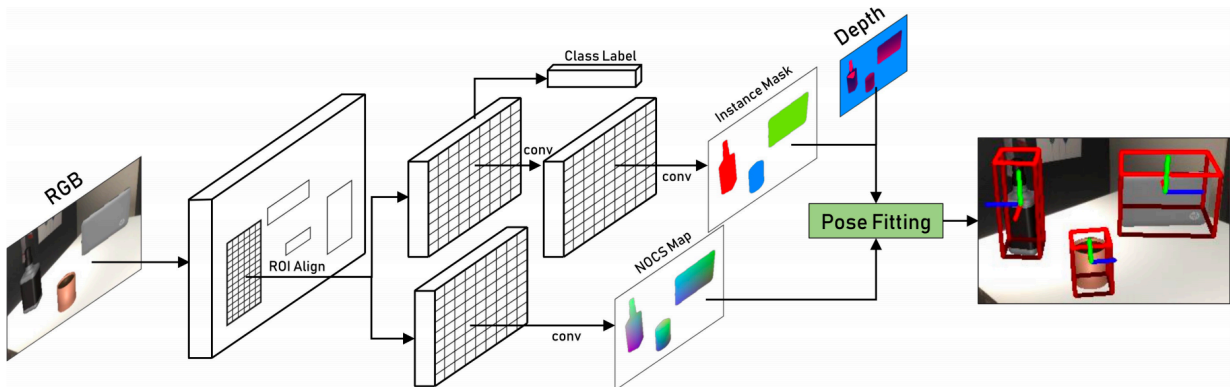


Figure 3. The inputs to our method are the RGB and depth images of a scene with multiple objects. Our CNN predicts the class label, instance mask, and NOCS map (color-coded) for each object in the RGB image. We then use the NOCS maps for each object together with the depth image to obtain the full metric 6D pose and size (axes and tight red bounding boxes), even if the object was never seen before.