6D Pose Estimation

2019.07.20

Solaris

(http://solarisailab.com)

6D Pose Estimation

- The task of estimating the 6D pose of an object (3D translation and 3D rotation) is of great importance for many higher level operations such as robotic manipulation, augmented reality, scene interpretation and autonomous driving.

SSD-6D

- Kehl, Wadim, et al. "SSD-6D: Making RGB-based 3D detection and 6D pose estimation great again." *Proceedings of the IEEE International Conference on Computer Vision*. 2017.

- https://arxiv.org/abs/1711.10006

SSD-6D Architecture

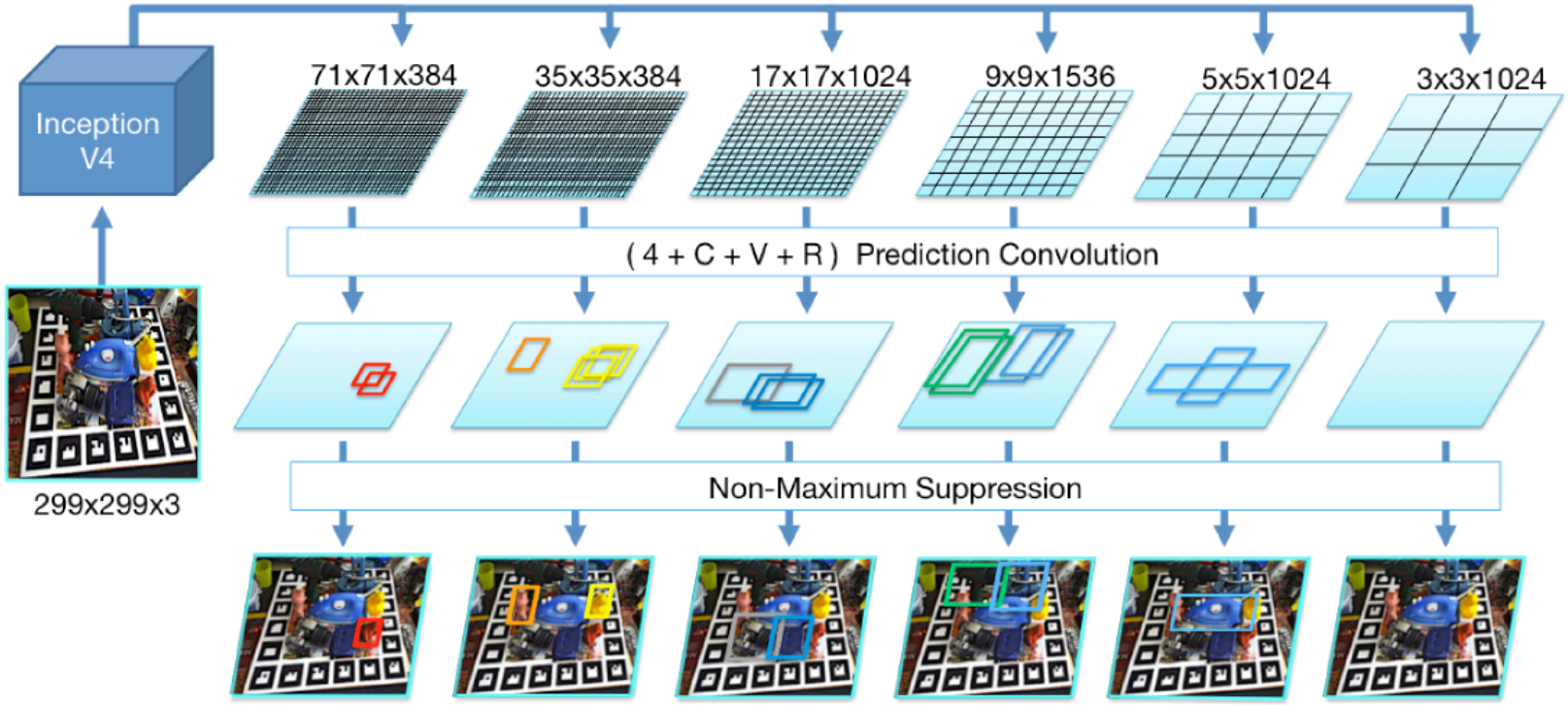


Figure 1: Schematic overview of the SSD-style network prediction. We feed our network with a 299 × 299 RGB image and produce six feature maps at different scales from the input image using branches from InceptionV4. Each map is then convolved with trained prediction kernels of shape (4 + C + V + R) to determine object class, 2D bounding box as well as scores for possible viewpoints and in-plane rotations that are parsed to build 6D pose hypotheses. Thereby, C denotes the number of object classes, V the number of viewpoints and R the number of in-plane rotation classes. The other 4 values are utilized to refine the corners of the discrete bounding boxes to tightly fit the detected object.

SSD-6D TensorFlow Implementation

- https://github.com/wadimkehl/ssd-6d

Pose CNN

- Xiang, Yu, et al. "Posecnn: A convolutional neural network for 6d object pose estimation in cluttered scenes." *arXiv preprint arXiv:1711.00199* (2017).
- https://arxiv.org/abs/1711.00199

Pose CNN Architecture

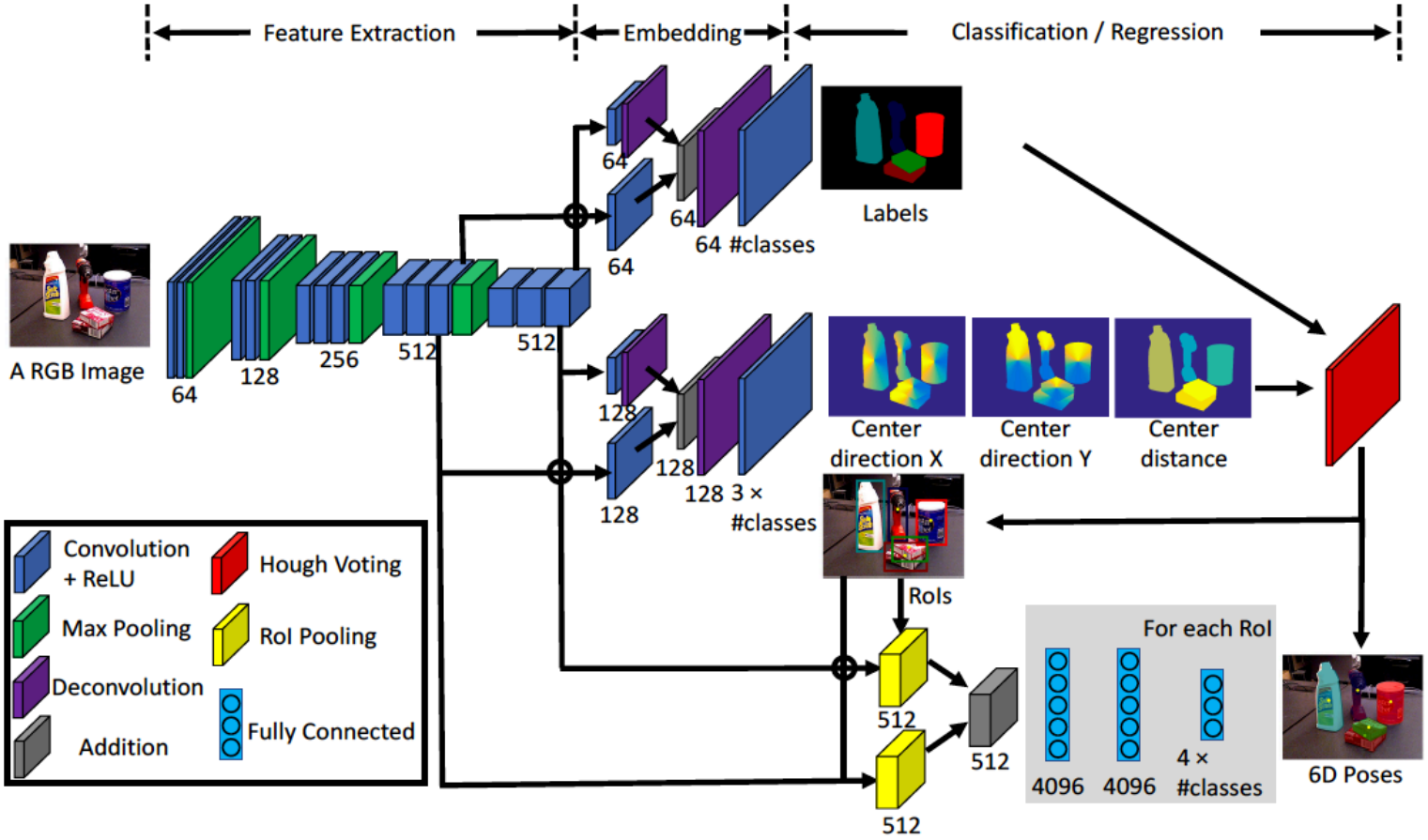
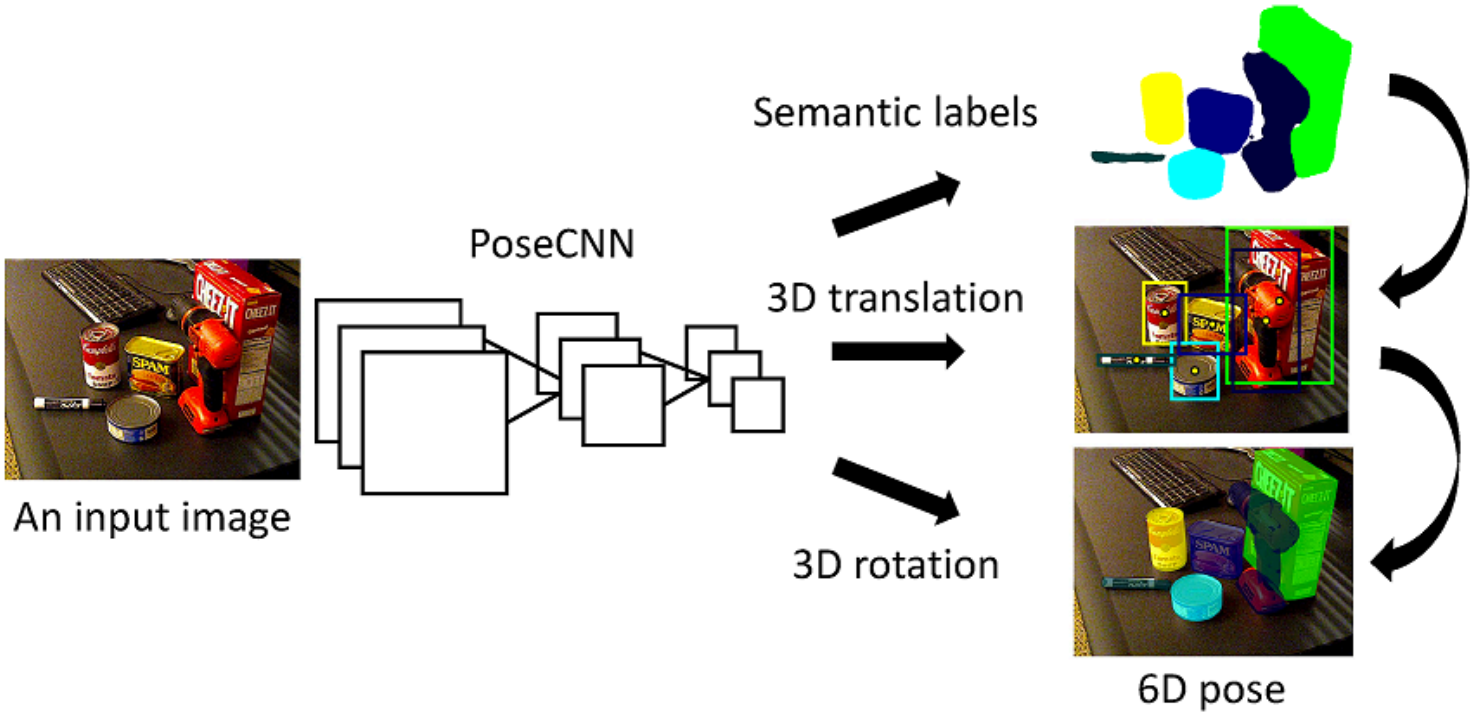


Fig. 2. Architecture of PoseCNN for 6D object pose estimation.

Pose CNN TensorFlow Implementation

- https://github.com/yuxng/PoseCNN



NOCS

- Wang, He, et al. "Normalized Object Coordinate Space for Category-Level 6D Object Pose and Size Estimation." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2019.

- https://arxiv.org/abs/1901.02970

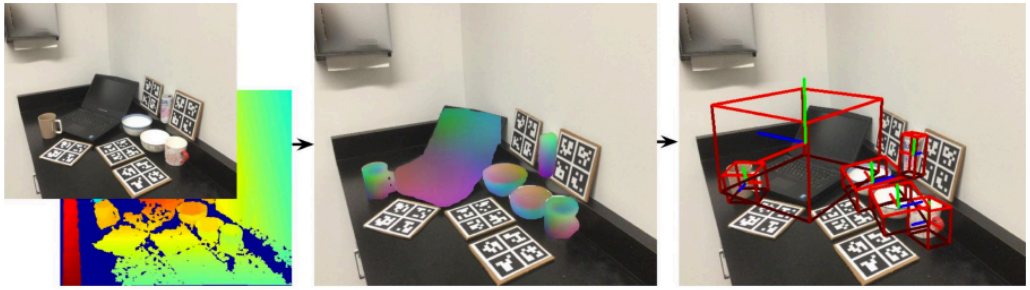


Figure 1. We present a method for category-level 6D pose and size estimation of multiple unseen objects in an RGB-D image. A novel normalized object coordinate space (NOCS) representation (color-coded in (b)) allows us to consistently define 6D pose at the category-level. We obtain the full metric 6D pose (axes in (c)) and the dimensions (red bounding boxes in (c)) for unseen objects.

NOCS Architecture

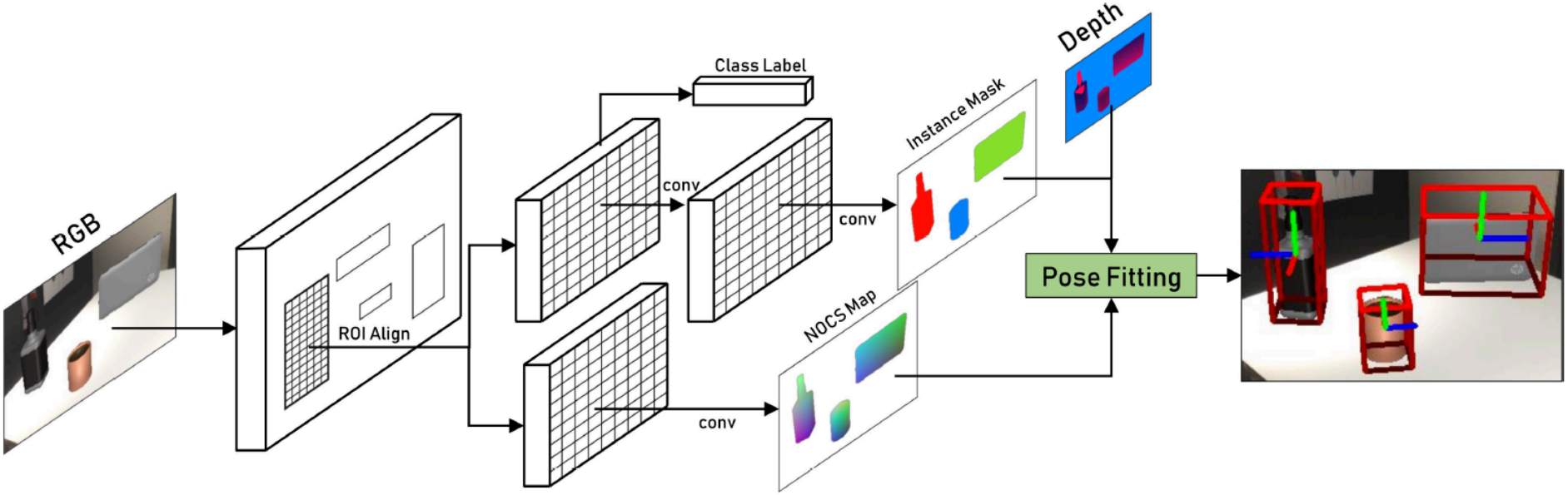


Figure 3. The inputs to our method are the RGB and depth images of a scene with multiple objects. Our CNN predicts the class label, instance mask, and NOCS map (color-coded) for each object in the RGB image. We then use the NOCS maps for each object together with the depth image to obtain the full metric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

Normalized Object Coordinate Space (NOCS)

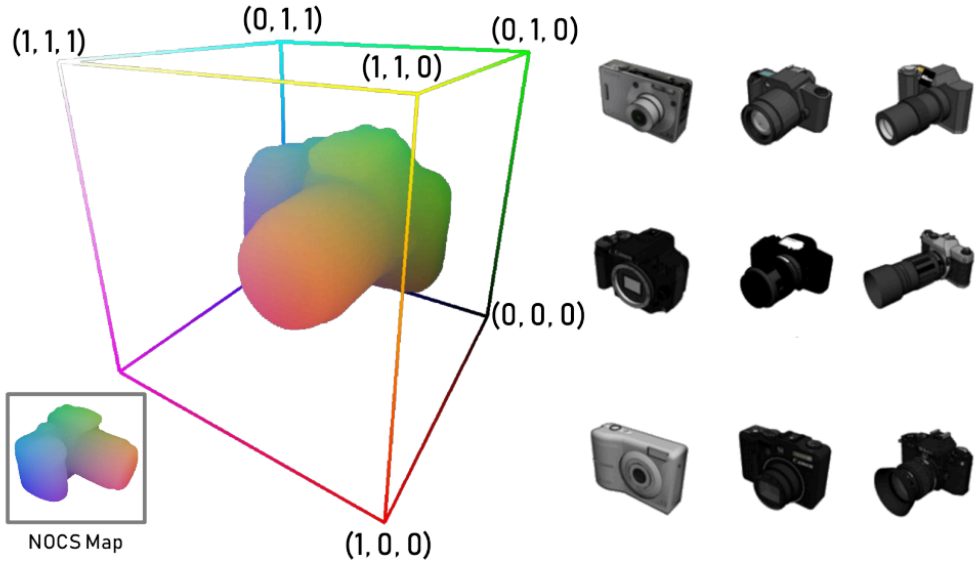
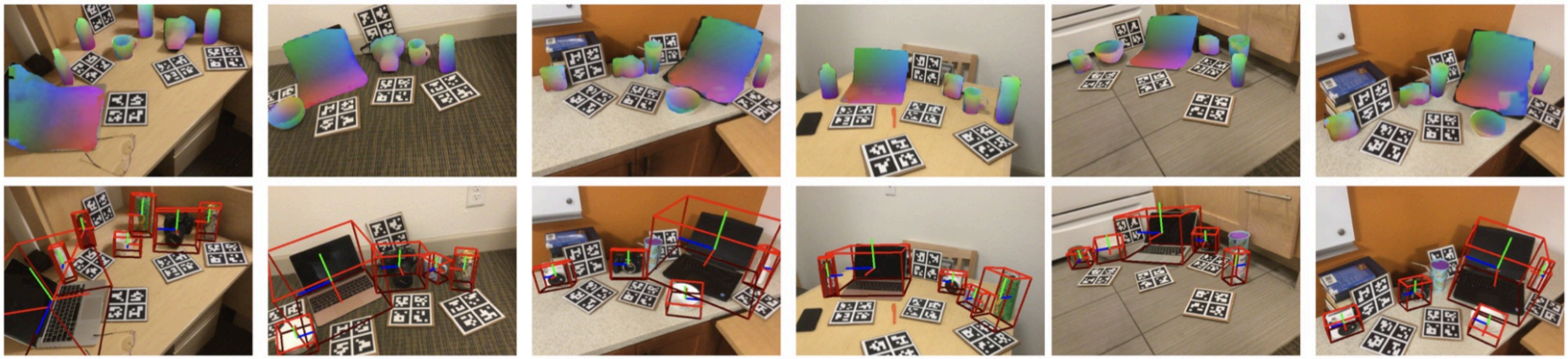


Figure 2. The Normalized Object Coordinate Space (NOCS) is a 3D space contained within a unit cube. For a given object category, we use canonically oriented instances and normalize them to lie within the NOCS. Each (x, y, z) position in the NOCS is visualized as an RGB color tuple. We train our network on the perspective projection of the NOCS on the RGB image, the NOCS map (bottom left inset). At test time, the network regresses the NOCS map which is then used together with the depth map for 6D pose and size estimation.

NOCS TensorFlow Implementation

- https://github.com/hughw19/NOCS_CVPR2019



Questions & Answers

Thank You!