# Weakly Supervised Hand Segmentation for Smart Store Applications
Group : NM4 CHENG-YOU LU, CSB4 DUN-JHIH SU
Advisor : WEN-HSIAO PENG

Abstract

- Generating labels for fully supervised semantic segmentation is expensive; Therefore, in SDI[1] which using Bounding Box, a kind of cheap-to-generate label, in replace of expensive semantic segmentation.
- To generate a label, the Bounding Box is processed with the GrabCut[2].
- Generating a more precise label by handcrafted inner bounding box and self training.
- After getting the more precise label, now the semantic segmentation can be solved by fully supervised learning.
- At last, refining the semantic segmentation with Dense CRF[4].
- We release our SDI[1] code on Github: github.com/johnnylu305/Simple-does-it-weakly-supervised-instance-and-semantic-segmentation

Structure
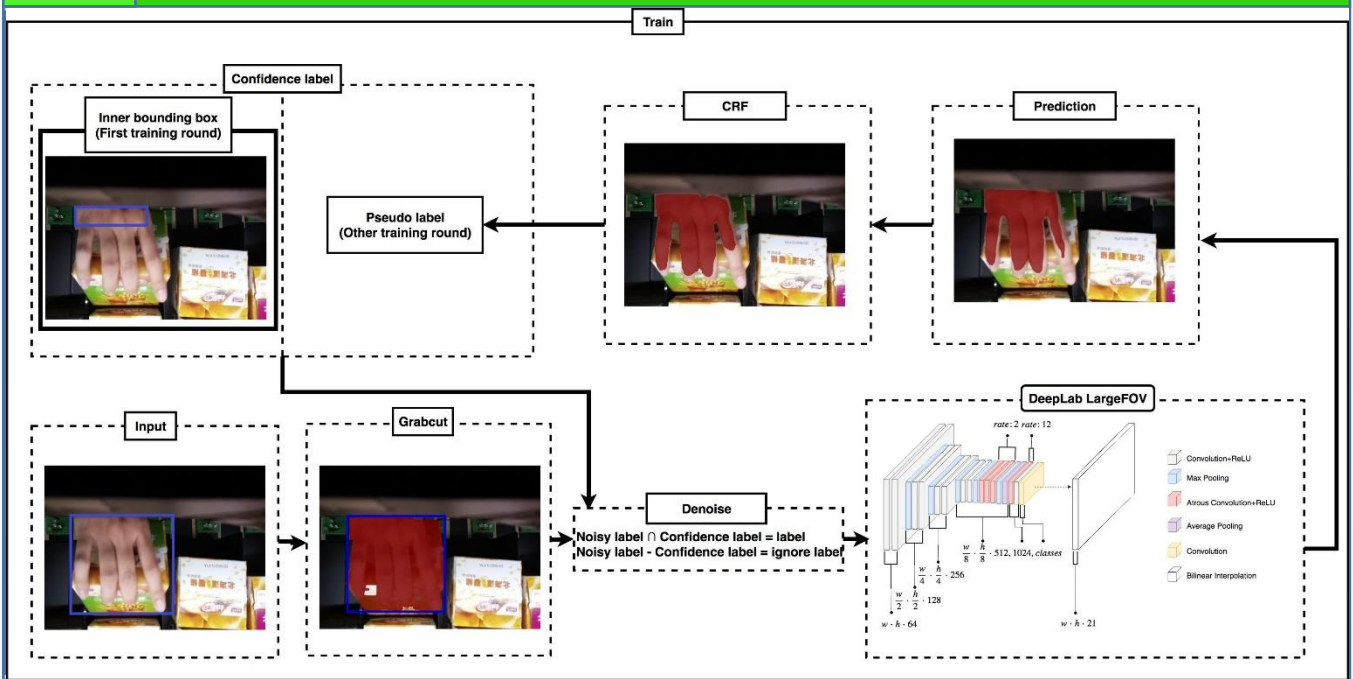
Experiment

- There are 720 training data and 480 testing data on ITRI hand dataset.
- There are 10000 training data on ITRI synthetic hand dataset with data augmentation and Pascal VOC12 hand dataset.
- The SYN direct use Grabcut[2] as label and trained on ITRI synthetic hand dataset.
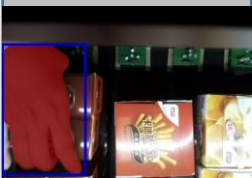- The Box nth trained on ITRI hand dataset with nth round self training.

Smart store

- Performance on CDVS[5] without hand segmentation is 68.9%.
- Performance on CDVS[5] with hand segmentation by SYN is 78.1%.
- Performance on CDVS[5] with hand segmentation by Box first is 75.9%.
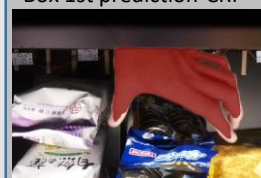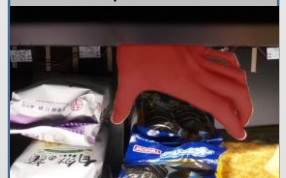
Performance
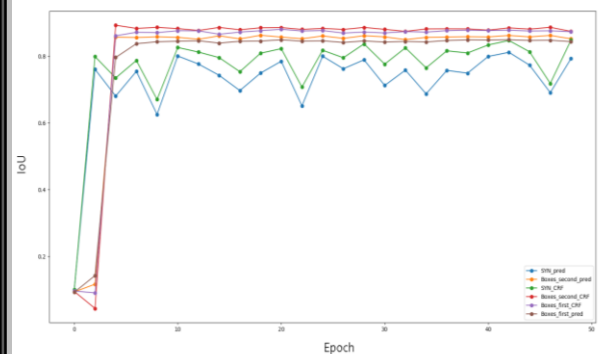
Grabcut	Box 1st pseudo label	Box 2nd pseudo label	Box 1st prediction-CRF	Box 2nd prediction-CRF

Reference

[1] Anna Khoreva, Rodrigo Benenson, Jan Hosang, Matthias Hein, Bernt Schiele, Max Planck Institute for Informatics, Saarbrücken, Germany, Saarland University, Saarbrücken, Germany. Simple Does It: Weakly Supervised Instance and Semantic Segmentation. arXiv:1603.07485v2, 2016.
[2] C. Rother, V. Kolmogorov, and A. Blake. Grabcut: Interactive foreground extraction using iterated graph cuts. In ACM Trans. Graphics, 2004
[3] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. arXiv:1606.00915, 2016.
[4] P. Krähenbühl and V. Koltun. Efficient inference in fully connected crfs with gaussian edge potentials. In NIPS. 2011.
[5] Miroslaw Bober, Werner Bailer, Stavros Paschalakis, Jie Chen "Compact Descriptors for Visual Search (CDVS) – The Standard for Image Search" ISO/IEC JTC1/SC29/WG11 W16351 2016.