

Learning to Recognize Objects by Retaining other Factors of Variation

Submitted by el_wehby on Sun, 05/27/2018 - 11:59pm

Title Learning to Recognize Objects by Retaining other Factors of Variation
Publication Type Conference Paper
Year of Publication 2017
Authors [J. Zhao](#), [C. K. Chang](#), [L. Itti](#)
Conference Name Proceedings of the IEEE Winter Conference on Applications of Computer Vision (WACV), Santa Rosa, CA
Date Published Mar
Keywords [1544814](#)

Abstract

Most ConvNets formulate object recognition from natural images as a single task classification problem, and attempt to learn features useful for object categories, but invariant to other factors of variation such as pose and illumination. They do not explicitly learn these other factors; instead, they usually discard them by pooling and normalization. Here, we take the opposite approach: we train ConvNets for object recognition by retaining other factors (pose in our case) and learning them jointly with object category. We design a new multi-task learning (MTL) ConvNet, named disentangling CNN (disCNN), which explicitly enforces the disentangled representations of object identity and pose, and is trained to predict object categories and pose transformations. disCNN achieves significantly better object recognition accuracies than the baseline CNN trained solely to predict object categories on the iLab-20M dataset, a large-scale turntable dataset with detailed pose and lighting information. We further show that the pretrained features on iLab-20M generalize to both Washington RGB-D and ImageNet datasets, and the pretrained disCNN features are significantly better than the pretrained baseline CNN features for fine-tuning on ImageNet.

Citation Key Zhao_etal17wacv



[1544814](#)
