Contribution ID: 5

Multi-Scale Orderless Pooling of Deep Convolutional Activation Features

Deep convolutional neural networks (CNN) have shown their promise as a universal representation for recognition. However, global CNN activations lack geometric invariance, which limits their robustness for classication and matching of highly variable scenes. To improve the invariance of CNN activations without degrading their discriminative power, this paper presents a simple but eective scheme called multi-scale orderless pooling (MOP-CNN). This scheme extracts CNN activations for local patches at multiple scale levels, performs orderless VLAD

pooling of these activations at each level separately, and concatenates the

result. The resulting MOP-CNN representation can be used as a generic feature for either supervised or unsupervised recognition tasks, from image classication to instance-level retrieval; it consistently outperforms global CNN activations without requiring any joint training of prediction layers for a particular target dataset. In absolute terms, it achieves state-of-the-art results on the challenging SUN397 and MIT Indoor Scenes classication datasets, and competitive results on ILSVRC2012/2013 classification and INRIA Holidays retrieval datasets.

Author: Mr YUNCHAO, Gong (University of North Carolina)

Presenter: Mr YUNCHAO, Gong (University of North Carolina)

Track Classification: Machine Learning