

**CHENNAI – PONDICHERRY**

**Self Paced Deep Learning for Weakly Supervised Object Detection**

**Abstract**

In a weakly-supervised scenario object detectors need to be trained using image-level annotation alone. Since bounding-box-level ground truth is not available, most of the solutions proposed so far are based on an iterative, Multiple Instance Learning framework in which the current classifier is used to select the highest-confidence boxes in each image, which are treated as pseudo-ground truth in the next training iteration. However, the errors of an immature classifier can make the process drift, usually introducing many of false positives in the training dataset. To alleviate this problem, we propose in this paper a training protocol based on the self-paced learning paradigm. The main idea is to iteratively select a subset of images and boxes that are the most reliable, and use them for training. While in the past few years similar strategies have been adopted for SVMs and other classifiers, we are the first showing that a self-paced approach can be used with deep-network-based classifiers in an end-to-end training pipeline. Our method is built on the fully-supervised Fast-RCNN architecture and can be applied to similar architectures which represent the input image as a bag of boxes. We show state-of-the-art results on Pascal VOC 2007-2010 and ILSVRC 2013. On ILSVRC 2013 our results based on a low-capacity AlexNet network outperform even those weakly-supervised approaches which are based on much higher-capacity networks.