Review: Object-Graphs for Context-Aware Category Discovery

Murat Ambarkutuk {murata@vt.edu}, 02/16/2016

Image categorization is the effort of classifying a set of images according to objects they contain and the context the objects collectively form. In their paper, the authors mention that unsupervised category learning has some superior aspects compared to supervised approaches, such as the bias occuring from the labels and the annotation of the datasets with which the algorithms are being trained and the amount of work that is needed to compile the mentioned annotations. The authors, on the other hand, underline that unsupervised approaches only benefit from appereance-only cues to distinguish one class from another. Given mentioned superiorities, the paper proposes an extension to unsupervised technique. The contribution is based on the object-graph descriptor in which spatial configuration and co-occurance frequency of the objects were also exploited in addition to the visual cues are utilized.

The approach contains three major steps revolving around the proposed descriptor, and it then outputs newly-discovered objects according the context of the objects found in unsupervised approaches. The first step is to create a pool of segments from each image with varying segmentation settings, from which known objects, unknown objects, and mixture of both cases are classified with corresponding confidence scores. The authors note that the confidence score will display higher certainity in the case of known and unknown objects, lower in mixture of both, assuming reliable classifiers. The certainity cue is measured with Shannon Entropy and a cutoff threshold is set according to the certainity measure. The objects which have smaller entropy than the threshold are considered as the unknown objects. Distinguishing the known regions from unknown regions by using the entropy threshold forms the first contrubition of the paper. This approach reminded me of Dempster-Shafer theory in which "ignorance" on some knowledge can be explicitly modeled. The second step in the algorithm which is the second contribution of the paper is a means of exploiting the known-unknown object distrubtion, both spatial-wise and frequency-wise. In order to leverage the information unknown objects provided, "a histogram of object classes is constructed for each unknown objects, within a varying spatial distance given two orientations, above and below". This is an intuitive way of formulating a solution the authors mentioned in the first section. Albeit intricate and elaborate, the one-sentence explanation given in above is one of the best measure of how good the contribution is formulated and presented. The last step of the algorithm is to discover categories from the familiar objects. In this step, the authors use a combined similarity measure with which both appearence and co-occurence histograms are taken into account.

The proposed method was compherensively tested with four different datasets under different settings. The authors claim that the proposed approach outperforms the baseline algorithm. Since it is a new way of approaching the problem of object discovery which limits the scope of comparison, it would be a easier to compherend the significance of the contributions if the results were compared to, at least, similar techniques targeting the same problem. The authors also remarked that the proposed method and the baseline performed strongly well where the objects have distinct visual cues. This, to some extent, implies that there is a strong correlation between the performance of the both approaches which also implies the vulnerability to the threats resulting from apperiance based cues as much as the baseline is. One naive solution may be handled by a change in the similarity measure used in the last step of the algorithm by using weighted sum in the components of the similarity measure.