This CVPR2015 extended abstract is the Open Access version, provided by the Computer Vision Foundation.

Discriminative and Consistent Similarities in Instance-Level Multiple Instance Learning

Mohammad Rastegari[1], Hannaneh Hajishirzi[2], Ali Farhadi[2,3]
[1]University of Maryland. [2]University of Washington. [3]Allen Institute for AI.

Multiple-instance learning (MIL) [2] addresses a variation of classification problems where complete labels of training examples are not available. In the MIL setup, training labels are assigned to *bags* of instances rather than individual instances. In most standard MIL setups, a bag is positive if it contains at least one positive instance, and is negative if all of its instances are negative. The standard task in MIL is to classify unknown bags of instances (e.g., [3, 6]). However, several application domains require instance-level predictions (e.g., [5]). For example, in image segmentation (instances are superpixels, bags are images) the main goal is to find the exact regions of an image that correspond to the objects of interest.

Instance-level MIL has been approached either by a complex joint optimization over bag and instance classifiers (e.g., [1]) or by identifying positive instances followed by bag classification. Latter involves similarity-based reasoning where most methods either use standard similarity functions (e.g., [6]) or learn a global similarity function for all instances (e.g., [5]). Standard similarity functions are not necessarily discriminative (orange dashed links in Fig. 1) and cannot discover common properties among positive instances. Globally learned similarity functions cannot encode different types of similarities that tie together positive instances within groups (Fig. 1).

In this paper, we introduce a new method for the problem of instance-level MIL with globally-constrained reasoning about local pairwise discriminative similarities. We introduce a novel approach that learns similarity functions specific to each instance and reasons about the underlying structure of similarities between positive instances using our notion of consistent similarities (Green cliques in Fig. 1). We introduce a discriminative notion of similarity that enables learning a similarity function for each pair of instances in positive bags (similarity patterns in Fig. 1). Typically, learning a similarity function requires training labels for similar instances. However, instance-level labels are not available in MIL. We use negative bag labels as the only certain labels in MIL to learn our discriminative similarity function. Instances in positive bags are similar if they are *similarly different* form instances in negative bags.

Pairwise similarities are not always transitive and can be confused with coincidental patterns in a high-dimensional feature space [4] (Purple dashed links in Fig. 1). For example, two images *a* and *c* cannot be similar to each other only because they are similar to another image *b*; *a* might be similar to *b* because both show a sunset over an ocean, and *c* might be similar to *b* because of coincidental patterns of similarity. A reliable pairwise similarity should be globally consistent across several pairs (green links in Fig. 1). We introduce a novel clique-based notion of similarity that measures global consistency of pairwise similarities.

We formulate the discovery of positive instances as a ranking problem where top rank instances in positive bags are highly and consistently similar to each other. The bag labels provide constraints to our optimization problem; real positive instances inside each bag should rank higher than negative instances in negative bags. We show that a random-walk based ranking algorithm that uses our globally-consistent pairwise similarities outperforms state-of-the-art MIL results in MIL benchmarks, text categorization, and image segmentation (Fig 2).

Overview of Our Method:

At training, bags B^{tr} of instances X^{tr} along with bag-level labels b^{tr} are known; Instance-level labels are not given. Our training algorithm has two main steps: discovering "correct" positive instances $\mathcal{L}^+$ among all instances X^{tr^+} in positive bags based on the training bag labels and training a final binary classifier using the discovered positive instances $\mathcal{L}^+$ and instances X^{tr^-} in negative bags in the training set . At test time neither bag labels nor instance-level labels are known. We test our method on how well it can predict both bag-level and instance-level labels. For testing, we use the final binary classifier to predict labels of individual instances in the test set. Bag-level labels are then predicted using instance-level labels; a bag is positive if it includes at least one predicted positive instance.

This is an extended abstract. The full paper is available at the Computer Vision Foundation webpage.

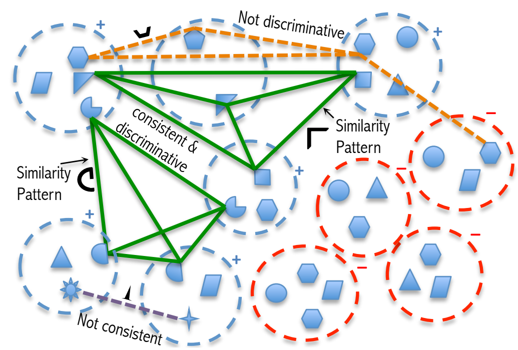

Figure 1: Discriminative and Consistent Similarities are shown by green cliques. Orange similarities are consistent but not discriminative (similar to negatives). Purple similarity is discriminative but not consistent.

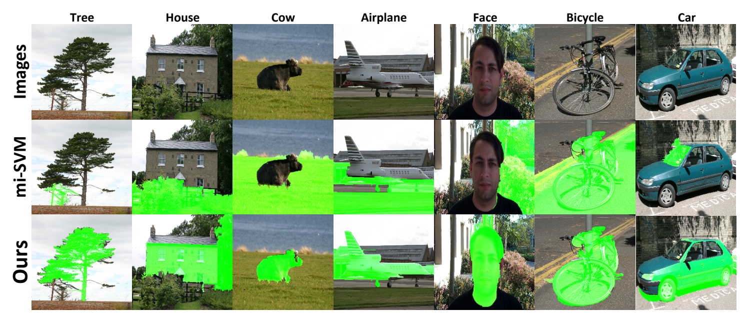

Figure 2: Positive Instance Discovery: Our method is capable of discovering positive instances (superpixels that correspond to the object of interest) in bags(images). (The first row shows 7 images from IL-MSRC. The second and third rows show discovered superpixels using miSVM [1] and our method, respectively.

[1] Andrews, Tsochantaridis, and Hofmann. SVMs for multiple-instance learning. In *NIPS*, 2003.

[2] Dietterich., Lathrop, and Lozano. Solving the multiple instance problem with axis-parallel rectangles. *AI*, 1997.

[3] Oded Maron and TomÃąs Lozano-PÃľrez. A framework for multiple-instance learning. In *NIPS*, 1996.

[4] Abhinav Shrivastava, Tomasz Malisiewicz, Abhinav Gupta, and Alexei A. Efros. Data-driven visual similarity for cross-domain image matching. *Proceedings of ACM SIGGRAPH ASIA*, 2011.

[5] Xiao, Liu, Cao, Jie, and Wu. Smile: A similarity-based approach for multiple instance learning. In *ICDM*, 2010.

[6] Zhou, Sun, and Li. Multi-instance learning by treating instances as non-i.i.d. samples. In *ICML*, 2009.