Semi-Supervised Boosting using Visual Similarity Learning

Christian Leistner

Institute for Computer Graphics and Vision, Graz University of Technology

Abstract The required amount of labeled training data for object detection and classification is a major drawback of current methods. Combining labeled and unlabeled data via semi-supervised learning holds the promise to ease the tedious and time consuming labeling effort. This paper presents a novel semi-supervised learning method which combines the power of learned similarity functions and classifiers. The approach capable of exploiting both labeled and unlabeled data is formulated in a boosting framework. One classifier (the learned similarity) serves as a prior which is steadily improved via training a second classifier on labeled and unlabeled samples.
We demonstrate the approach on challenging computer vision applications. First, we show how we can train a classifier using only a few labeled samples and many unlabeled data. Second, we improve (specialize) a state-of-the-art detector by using labeled and unlabeled data.

Motivation

Use both labeled and unlabeled data in order to learn/improve a boosted classifier.

Approach

- **Combine Graph-based and Cluster-based semi-supervised learning methos**
- **Additively combine three different loss functions (between labeled, labeled-unlabeled plus unlabeled and unlabeled data)**
- **Use a prior similarity measure to indicate the relation among the data**
- **Minimize the additiely combined loss using AdaBoost**

$$\mathcal{L} = \frac{1}{|\mathcal{X}^L|}\sum_{\mathbf{x}\in\mathcal{X}^L} e^{-2yH(x)} + \frac{1}{|\mathcal{X}^L||\mathcal{X}^U|}\sum_{\mathbf{x}_i\in\mathcal{X}^L}\sum_{\mathbf{x}_j\in\mathcal{X}^U} S(\mathbf{x}_i,\mathbf{x}_j)e^{-2y_iH(\mathbf{x}_j)} + \frac{1}{|\mathcal{X}^U||\mathcal{X}^U|}\sum_{\mathbf{x}_i\in\mathcal{X}^U}\sum_{\mathbf{x}_j\in\mathcal{X}^U} S(\mathbf{x}_i,\mathbf{x}_j)e^{H(\mathbf{x}_i)-H(\mathbf{x}_j)}$$

$$\mathcal{L} \leq \frac{1}{|\mathcal{X}^L|}\sum_{\mathbf{x}\in\mathcal{X}^L} w_n(\mathbf{x},y)e^{-2y\alpha_n h_n(\mathbf{x})} + \frac{1}{|\mathcal{X}^U|}\sum_{\mathbf{x}\in\mathcal{X}^U}\left[p_n(\mathbf{x})e^{-\alpha_n h_n(\mathbf{x})} + q_n(\mathbf{x})e^{\alpha_n h_n(\mathbf{x})}\right]$$

$$p_n(\mathbf{x}) = e^{-2H_{n-1}(\mathbf{x})}\frac{1}{|\mathcal{X}^L|}\sum_{\mathbf{x}_i\in\mathcal{X}^+} S(\mathbf{x},\mathbf{x}_i) + \frac{1}{|\mathcal{X}^U|}\sum_{\mathbf{x}_i\in\mathcal{X}^U} S(\mathbf{x},\mathbf{x}_i)e^{H_{n-1}(\mathbf{x}_i)-H_{n-1}(\mathbf{x})}$$

$$q_n(\mathbf{x}) = e^{2H_{n-1}(\mathbf{x})}\frac{1}{|\mathcal{X}^L|}\sum_{\mathbf{x}_i\in\mathcal{X}^-} S(\mathbf{x},\mathbf{x}_i) + \frac{1}{|\mathcal{X}^U|}\sum_{\mathbf{x}_i\in\mathcal{X}^U} S(\mathbf{x},\mathbf{x}_i)e^{H_{n-1}(\mathbf{x})-H_{n-1}(\mathbf{x}_i)}$$

New weights and assigned labels in each iteration

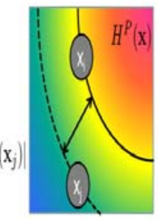
$$z_n(\mathbf{x}) = sign(p_n(\mathbf{x}) - q_n(\mathbf{x})) \qquad w_n(\mathbf{x}) = |p_n(\mathbf{x}) - q_n(\mathbf{x})|$$

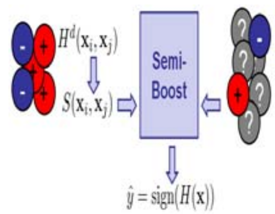
Learning the Distance Function

$$d(\mathbf{x}_i,\mathbf{x}_j) = \frac{1}{2}(H^d(\mathbf{x}_i,\mathbf{x}_j)+1)$$

$$S(\mathbf{x}_i,\mathbf{x}_j) = e^{\left(-\frac{d(\mathbf{x}_i,\mathbf{x}_j)^2}{\sigma^2}\right)}$$

$S(\mathbf{x}_i,\mathbf{x}_j)$ ⇓ $|H^P(\mathbf{x}_i) - H^P(\mathbf{x}_j)|$



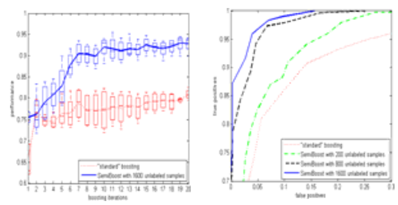


Classifier Combination

$$H^C(\mathbf{x}) = \alpha_0 h_0(\mathbf{x}) + \sum_{t=1}^{T}\alpha_t h_t(\mathbf{x})$$

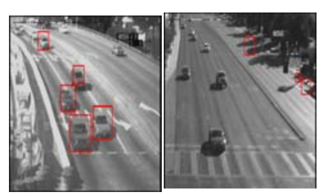
Experiments

Learning from few labeled samples

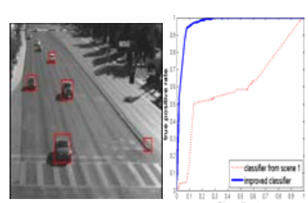


Knowledge Transfer and Scene Adaption

prior

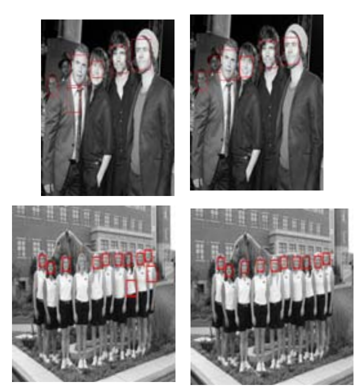


improved



Improving an Object Detector

before **after**



References

[1] C. Leistner, H. Grabner and H. Bischof, "Semi-Supervised Boosting Using Visual Similarity Learning", CVPR 2008

[2] T. Hertz, A. Bar-Hillel and D. Weinshall, Learning Distance Functions for Image Retrieval, CVPR, 2004

[3] Freund and Schapire, "A decision-theoretic generalization of on-line learning ans an application to boosting", Journ. Comp. Sys. Sc., 1997

*This work was supported by the the FFG project EVis (813399).