

Motivation

Image classification task: dose the image contain ...?







Horse?

Train?

Sky?

Difficulty: The location of semantic concept is unknown. Common approaches:

- Consider multiple local regions
- Use either average or maximum scores



Problems

- Not all regions correspond to semantic area. Low signal-to-noise ratio

- Not robust, assume perfect classifier Assume a sample region correspond to the semantic area

Our approach:



Find a *sparse* set of regions that are likely to contain objects of interest, and classify images using weighted combination of these regions with a *jointly-optimized* classifier.

Region Ranking SVM for Image Classification Zijun Wei, Minh Hoai Computer Vision Lab, Stony Brook University, NY, USA

Region Ranking SVM

Classification decision:

Multiple local regions -





Obtained by rearranging the columns of the matrix **B** so that $w^{T} \Gamma(B; w)$ is a sequence of non-increasing values.



Joint learning formulation:

$$\begin{array}{l} \underset{\mathbf{w},\mathbf{s},b}{\text{minimize } \lambda ||\mathbf{w}||^2 + \sum_{i=1} (\mathbf{w}^T \Gamma(\mathbf{B}_i;\mathbf{w})\mathbf{s} + b - y_i)^2 \\\\ \text{s.t. } s_1 \ge s_2 \ge \cdots \ge s_m \ge 0, \\\\ h(\{\Gamma(\mathbf{B}_i;\mathbf{w})\mathbf{s}\}) \le 1. \end{array}$$

Iterative optimization:

(A) Fix w, optimize with respect to s, b using Quadratically Constrained Quadratic Program (B) Fix *s*, optimize with respect to *w*, b using Ridge regression



Experiments

Score combination vector Update region ranking Global rept. Regions

Update classifier w

VOC 2007 Image Classification:

Method	mAP
Chatfield et al. [4]	82.4
Wei et al. [40]	85.2
VGG16 [31]	89.3
RRSVM (ours)	92.9

VOC 2012 Action Classification:

Method

Oquab et al. [24] Hoai et al. [14] Gkioxari et al. [10] Hoai [12] Simonyan and Zisserman [31] **RRSVM** (ours)

ILSVRC 2014 Image Classification:

6.7% Top-5 error using only one dense-sampling net (VGG 16 baseline 7.5%).



Acknowledgements:

This project is partially supported by the National Science Foundation Award IIS-1566248 and IIS-1161876.

IEEE 2016 Conference on **Computer Vision and Pattern Recognition**

CVPR2016

Findings • RRSVM outperforms all other methods in *every* object category(see the paper) RRSVM outperforms state of the art by app. **30%** error reduction. • RRSVM can be applied to other mAP problems as well 70.2 70.5 RRSVM outperforms previous 73.6 state-of-the-art methods with 76.3

84.0

85.5

- RRSVM outperforms the baseline by **0.8%**
- RRSVM outperforms VGG 16 in more than **70%** of the 1000 classes in ILSVRC 14.

out any task-specific heuristics

• For most objects, a sparse set of regions are used (less than 10)

RRSVM uses a *sparse* set that is *jointly trained* with classifier for *better* classification

Scan for more!

