Multimodal semi-supervised learning for image classification

M Guillaumin,J Verbeek,C Schmid-CVPR2010

Yunfei Wang

Department of Computer Science & Technology
Huazhong University of Science & Technology

April 16, 2013

Table of contents

Motivation

- Increasing quantity and diversity of hand-labelled images improves the performance of classifier.
- Labelling images is time consuming and unrealistic in some scenario.
- Motivation:Using other information sources to aid learning process on limited amount of labelled images.

Supervised Learning I

Train a SVM on labelled images:

$$f(x) = \sum_i \alpha_i k(x, x_i) + b \tag{1}$$

where $k(\cdot, \cdot)$ is kernel function.
Class label:

$$y \in \{-1, +1\} = sign(f(x))$$

Combine visual and textual representations

Combined kernel:

$$k_c(\cdot,\cdot) = d_v k_v(\cdot,\cdot) + d_t k_t(\cdot,\cdot) \tag{2}$$

where $k_v(\cdot,\cdot)$ is visual kernel,$k_t(\cdot,\cdot)$ is textual kernel,$d_v, d_t > 0$ and $d_v + d_t = 1$.

Multiple kernel learning(MKL) framework allows joint learning of kernel combination weights d_v,d_t and the parameters $\{\alpha_i\}$ and b.

The SVM based on combined kernels:

$$f_c(x) = \sum_i \alpha_i k_c(x, x_i) + b \tag{3}$$

Semi-supervised Learning I

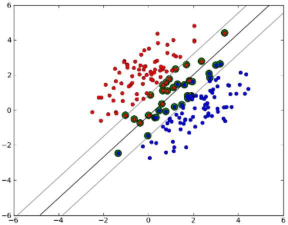
Notation:$\mathcal{L}$ is set of labelled images;$\mathcal{U}$ is set of unlabelled images.

Processing tactics

1. Learn a joint visual-textual classifier $f_c$form $\mathcal{L}$ using MKL;
2. Predict class labels for images in $\mathcal{U}$ with f_c;
3. Train a visual-only SVM classifier f_v from $\mathcal{L}$ and $\mathcal{U}$;
4. Complete image classification task with f_v.

First Alternative

Intention:Only add the confidently classified examples in $\mathcal{U}$,those fall outside the margin with $|f_c(x)| \geq 1$.



They have the less possibility of changing the MKL classifier.

Second Alternative

Intention:Approximate the joint classification function f_c.
Perform a least square regression(LSR) on MKL scores $f_c(x)$ for all examples in $x \in \mathcal{L} \cup \mathcal{U}$ to find the function

$$f_v(x) = \sum_i \alpha_i k_v(x, x_i) + b$$

Pseudo-code of the algorithm

Algorithm 1: Procedure for learning a semi-supervised MKL+LSR visual classifier.

Input: Labeled data $\mathcal{L}$ and unlabeled data $\mathcal{U}$, visual kernel k_v and textual kernel k_t.

Output: Visual classifier α using kernel k_v.

$f_c \leftarrow \mathrm{MKL}(\mathcal{L}, \{k_v, k_t\})$ /* *Learn MKL classifier* */

foreach $x \in \mathcal{L} \cup \mathcal{U}$ **do** /* *Center scores* */

  $s(x) \leftarrow f_c(x) - \langle f_c(x') \rangle_{x' \in \mathcal{L} \cup \mathcal{U}}$

end

foreach $x, x' \in \mathcal{L} \cup \mathcal{U}$ **do** /* *Center kernel columns* */

  $K_v(x, x') \leftarrow k_v(x, x') - \langle k_v(x, x'') \rangle_{x'' \in \mathcal{L} \cup \mathcal{U}}$

end

$U \Lambda V^\top = K_v$ /* *SVD of K_v* */

for $i = 1$ **to** $|\mathcal{L} \cup \mathcal{U}|$ **do** /* *Pseudo-invert K_v* */

  $$\overline{\Lambda}_{ii} \leftarrow \begin{cases} 0 & \text{if } \Lambda_{ii} < \epsilon \\ \Lambda_{ii}^{-1} & \text{otherwise} \end{cases}$$

end

$\alpha \leftarrow V \overline{\Lambda} U^\top s$ /* *Least-squares regression of s* */