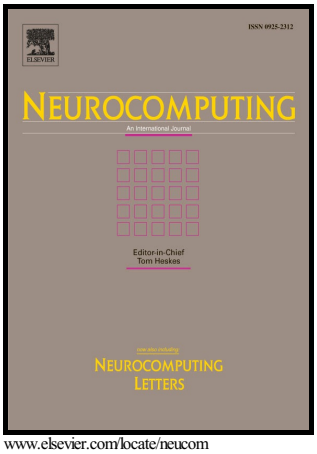

Multi-View Semi-Supervised Learning For Image Classification

Songhao Zhu, Xian Sun, Dongliang Jin

PII: S0925-2312(16)30441-6
DOI: http://dx.doi.org/10.1016/j.neucom.2016.02.072
Reference: NEUCOM17081

To appear in: *Neurocomputing*

Received date: 17 September 2015
Revised date: 10 December 2015
Accepted date: 8 February 2016

Cite this article as: Songhao Zhu, Xian Sun and Dongliang Jin, Multi-View Semi Supervised Learning For Image Classification, *Neurocomputing* http://dx.doi.org/10.1016/j.neucom.2016.02.072

Multi-View Semi-Supervised Learning For Image Classification

Songhao Zhu, Xian Sun, Dongliang Jin
School of Automation, Nanjing University of Posts and Telecommunications, Nanjing 210046, China
zhush@njupt.edu.cn

Abstract—With the massive growth of digital image data uploaded to the Internet, classifying each image into appropriate semantic category with respect to its image content for image index and image retrieval has become an increasingly difficult and laborious task. To deal with this issue, we propose a novel multi-view semi-supervised learning framework which leverages the information contained in pseudo-labeled images to improve the prediction performance of image classification using multiple views of an image. In the training process, labeled images are first adopted to train view-specific classifiers independently using uncorrelated and sufficient views, and each view-specific classifier is then iteratively re-trained with respect to a measure of confidence using initial labeled samples and additional pseudo-labeled samples. In the classification process, the maximum entropy principle is utilized to assign appropriate category labels to unlabeled images via optimally trained view-specific classifiers. Experimental results on a general-purpose image database demonstrate the effectiveness and efficiency of the proposed multi-view semi-supervised image classification scheme.

Index Terms—Multi-View Semi-Supervised Learning, Maximum Entropy, Multi-View Fusion, Image Classification

1. INTRODUCTION

With the explosive growth of digital image collections on the internet, the demand of developing content-based analysis technologies to effectively organize, manage and utilize such huge amount of information resources has become an important and challenging research topic in the field of intelligent multimedia analysis. Among these technologies, image classification, which aims to build an exact correspondence between visual information at the perceptual level and linguistic descriptions at the semantic level, is an elementary step and a promising step for content-based image indexing, retrieval and other related multimedia applications. Therefore, classifying an image into high-level semantic category has emerged as an important and challenging research topic in recent years.

In the past couple of years, many novel algorithmic techniques have been proposed to deal with the problem of image classification, such as hierarchical semantic similarity based method[1], compact binary based method[2], optimized pulse-coupled neural network based method[3], locality constrained low-rank coding based method[4], global structure and sparse feature based method[5], low-rank sparse coding based method[6], structured low-rank representations based method[7], discriminative multi-manifold based method[8], bi-linear deep learning based method[9], local and global information based method[10], separable principal components analysis based method[11], cost-sensitive subspace based method[12], hierarchical Gaussianization based method[13], two dimensional multi-label active learning based method[14], multiple partially observed views based method[15], Wavelet feature based metric[16], and random sub-windows based method[17].

The task of these existing algorithms is to assign an appropriate category to a given image with respect to its semantic contents. There are two issues that should be considered when designing an effective and efficient image classification algorithm: on the one hand, the number of labeled images is often very small while the number of unlabeled images is often very large; on the other hand, an image is generally represented by a combination of feature set, such as color, shape and texture. The performance of image classification is seriously affected by the two issues discussed above. To address the first issue, the semi-supervised methods are adopted to leverage the information contained in unlabeled images to improve the prediction performance[18-27]. To address the second issue, the multi-view learning algorithms are utilized to achieve the informative and representative training images to reduce the amount of labeled samples required for training[28-32]. During the process of multi-view learning, multiple classifiers are first separately trained via several distinct views extracted from the labeled images; then, these trained classifiers assign labels to pseudo-labeled images; next, the disagreement among different classifiers is utilized to selected additional pseudo-labeled images; finally, new view-specific classifiers are trained using initial labeled images and newly pseudo-labeled images to improve the overall classification performance. The main idea of the proposed method is that multi-view learning and semi-supervised learning can be effectively integrated to improve the performance of image classification, such as the method proposed by Luo[33] which deals with the classification tasks by learning the weights between different views and the correlation between labels.

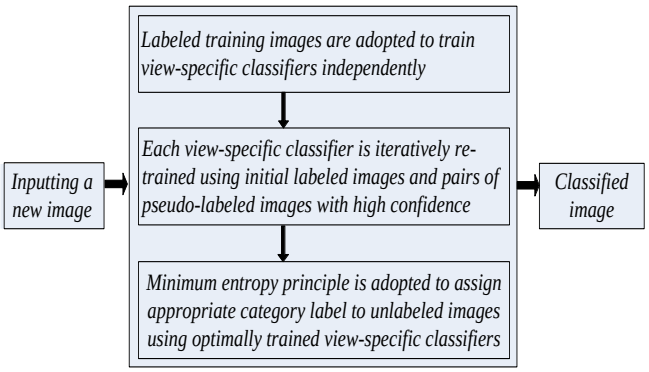

Figure 1: The basic idea of the proposed framework.