Pattern Recognition ■ (■■■■) ■■■–■■■

Contents lists available at ScienceDirect

Pattern Recognition

journal homepage: www.elsevier.com/locate/pr

Classification and saliency detection by semi-supervised low-rank representation ☆

Miaoyun Zhao *, Licheng Jiao, Wenping Ma, Hongying Liu, Shuyuan Yang

Key Laboratory of Intelligent Perception and Image Understanding of Ministry of Education, International Research Center for Intelligent Perception and Computation, Xidian University, Xi'an, Shaanxi Province 710071, China

ARTICLE INFO

Article history:
Received 10 June 2014
Received in revised form
24 July 2015
Accepted 6 September 2015

Keywords:
Low rank representation
Semi-supervised learning
Outlier detection
Saliency detection

ABSTRACT

In the area of pattern recognition, Low Rank Representation (LRR) is an efficient method in recovering the subspace structure of the dataset. However, LRR is unsupervised. Without any label information, LRR constructs an informative graph which is then combined with the mature graph-based semi-supervised learning (GSSL) framework to complete the classification task. In this paper, we propose a new low rank learning method which constructs the low rank representation matrix utilizing label information to obtain a more informative graph. This method integrates the low rank graph construction and the label information propagation processes together. Thus the optimization of the low rank representation and the soft label prediction function are calculated iteratively at the same time. We name this method as Semi-Supervised Low Rank Learning (SSLRL). It enhanced the classification performance of traditional LRR-Graph based SSL by 5–30% and the running time is reduced from hundreds to less than ten seconds. Based on this method, a new outlier detection strategy is presented. This strategy succeeds with an AUC of at least 93% even if the detection condition of LRR is not satisfied. The effectiveness of SSLRL is demonstrated in semi-supervised classification, outlier detection, and salient detection tasks. These extensive experimental results highlight the outperforming of our method over state-of-the-art methods.

1. Introduction

In the communities of pattern recognition and machine learning, semi-supervised learning (SSL) [1] has received considerable attentions. As we all known, in many application areas of pattern recognition and machine learning, such as web images classification, documents classification, and saliency detection, one has to face the serious situation of a great deal of data with only a small fraction labeled. The traditional supervised method is not suitable in this situation. To boost the performance of the algorithm, SSL makes thorough utilization of the abundant unlabeled and limited labeled data. Up to now, several kinds of SSL methods are widely developed and referenced [2–8]. Among them, the graph-based SSL (GSSL) is flexible and convenient when combined with the graph construction methods. In GSSL, we take the labeled and unlabeled samples as nodes in a graph, and we associate a weight matrix with the edges in the graph to encode the similarity among samples. Two fundamental GSSL methods are Local and Global Consistency (LGC) [9] and Gaussian random field and harmonic function (GHFH) [10]. LGC is based on two consistent assumptions: the local consistent assumes that nearby samples are more likely to share similar labels and the global consistent assumes that samples of similar structure are more likely to share similar labels. GHFH predicts labels by the Gaussian random field and harmonic function based on a k-Nearest-Neighbor (KNN) graph. Recently, to settle the small sample size problem (the number of samples is much smaller than the dimensionality of the samples) [11], Yu et al. [12] propose a semi-supervised classification based on a random subspace dimensionality reduction (SSCRSDR) method. This method obtains a classifier by fusing several GSSL classifiers which is trained on several lower dimensional subspaces. The subspaces are randomly selected from the original feature space of the dataset and this may lead the final classifier to miss some important features which are vital for classification. Another semi-supervised dimension reduction method is proposed by Lei et al. [13]. It applies the fusion-refinement procedure to extract the essential features from the original dataset. To make a thorough utilization of the unlabeled data, Mann et al. [14] propose a class membership based semi-supervised classification method (SSCCM). SSCCM employs a loss

☆This document is a collaborative effort.

* Principal corresponding author. Tel.: +86 15109284921.
E-mail addresses: miaoyun_zhao@163.com (M. Zhao), lchjiao@mail.xidian.edu.cn (L. Jiao), wpma@mail.xidian.edu.cn (W. Ma), hyliu@xidian.edu.cn (H. Liu), syyang@xidian.edu.cn (S. Yang).

http://dx.doi.org/10.1016/j.patcog.2015.09.008
0031-3203/

function which regards the label membership of each sample (both labeled and unlabeled) as a weighted average of its neighbors. Also, Zhu et al. [15] propose Label Propagation (LabPro) to investigate the use of unlabeled data in classification. In this method, all the nodes (labeled and unlabeled) propagate their labels across the affinity graph with the constraint, data points that are close to have similar labels, satisfied.

Nevertheless, most of the above GSSL methods are sensitive to data corruptions. Recently, Liu et al. [2] propose a low rank representation (LRR) method which is robust to noise and data corruptions due to its ability to decompose noise from the dataset. The low rank representation matrix (LRRM) of the dataset properly reflects the subspace structure of the whole data. The noise term identifies the corrupted data and outliers. Liua et al. [16–18] use the LRRM of the dataset to construct an affinity matrix and form an undirected graph which is called LRR-graph. Then, a GSSL process is conducted on the LRR-graph. The traditional GSSL methods construct an affinity graph with the label information incorporated, and then propagate the label information across the graph. However, in [16–18], the LRR-graph is constructed without any label information, thus the graph construction process is unsupervised. Zhuang et al. [17] propose a modified version of [5] by adding a sparse constraint to objective function to get a more sparse graph. Liu et al. [19] propose a manifold regularized low-rank representation method. It exploits not only the global Euclidean structure of the dataset (as traditional LRR do), but also the important and practical local manifold structure. This method enhances the performance of LRR on clustering tasks greatly. However, the graph in these afore-mentioned methods is constructed without any label information. To overcome this shortage, our method adds a constraint to incorporate label information into the optimization framework and makes the construction of the graph semi-supervised. To take advantage of label information in the LRR-graph construction process, in this paper, we propose a unified graph construction framework named Semi-Supervised Low Rank Learning (SSLRL) which guides the graph construction process with the label information. Our method combines the graph construction and label propagation processes of the traditional GSSL into one optimization problem. In other words, we optimize the LRRM of the data and predict the label for the unlabeled data within the same framework at the same time.

LRR is expert at outlier detection and thus widely used in saliency detection problems. Saliency detection [20] is a problem of detecting the salient areas within a natural image. The salient areas are the image regions that their features are different from the background. Shen et al. [21] utilize a variation of Robust Principal Component Analysis (RPCA) [22] method to settle the saliency detection problem. A LRR-based method, named multi-task sparsity pursuit (MTSP) model [23], is also proposed to saliency detection by combining multiple features. The key idea of these two methods is that they decompose the original image as a low-rank background matrix plus a sparse salient matrix. However, this idea is not always practical. If we grossly consider an image as a two class dataset, namely background area and salient area, then these two areas are low rank within themselves. Thus, the salient area is not contained in the noise term but in the low rank representation term. And it is hard to identify the salient area from the noise term. Differently from the traditional LRR based methods, we propose a new saliency detection strategy based on SSLRL. This strategy combines the information in the noise term and the information in the rows of LRRM corresponding to the labeled samples.

Our work has several contributions: first, our method is the first truly semi-supervised low rank representation learning method. Differently from traditional LRR based SSL framework which constructs the graph of the dataset unsupervisedly and then propagates the label information on the graph, our work uses the label information to guide the optimization of the LRRM. The LRRM obtained from our method is more consistent with the label information than the one obtained from the traditional method.

Secondly, our method proposes a simple unified optimization framework which integrates a graph construction process into a label propagation process. In the new framework, we optimize the LRRM and the soft label predicting function iteratively at the same time. Thus, the LRRM is not constant during the label propagation process but changeable according to the soft label predicting function. So LRRM and soft label predicting function modify each other to get a more confident result.

Thirdly, unlike the traditional LRR based methods which utilize only the noise term to identify the outliers of a dataset, our method enables us to use a more confident strategy to find out the outliers. The experiment part testifies that our method greatly enhances the performance of traditional LRR based methods in the outlier and saliency detecting problems.

The last but not the least, our method saves a lot of time when compared with the traditional LRR based methods. This is due to that SSLRL takes advantage of the label information which guides the LRRM to converge to the optimal solution more quickly.

The rest of this paper is organized as follows: in Section 2, we review some related works about LRR and SSL. Our works about semi-supervised low rank learning are discussed in Section 3. In Section 4, the time complexity of our method is analyzed. Extensive experiments are conducted in Section 5 to testify the effectiveness of our presented method. In Section 6, we summarize this paper and offer some future work plans.

2. Review of the fundamental work about LRR and GSSL

Up to now, a lot of works about applying LRR in GSSL have been proposed. In this section, we mainly review two most fundamental works.

2.1. Low rank representation

To exactly recover the true subspace structure of a dataset, Low-Rank Representation (LRR) [2] seeks the lowest rank representation that can represent the data samples as linear combinations of the bases in a given dictionary. Suppose $X=[x_1,x_2,\ldots,x_N]\in\Re^{d\times N}$ is a dataset with N observations, $x_i\in\Re^{d\times 1}(i=1,2,\ldots,N)$ is one of the observations which intrinsically embed on a multiple independent, low dimensional subspace. The LRR problem is defined as follows:

$$\min_{Z,E}\|Z\|_*+\lambda\|E\|_p \quad \text{s.t. } X=XZ+E. \tag{1}$$

where $Z\in\Re^{N\times N}$ is the low rank representation of the dataset (or reconstruction coefficients matrix), in this paper, to differentiate with the LRR method, we name this term as low rank representation matrix (LRRM). $E\in\Re^{d\times N}$ is the noise term corrupting X. $\|\cdot\|_*$ is the nuclear norm of a matrix, it is a convex relaxation to the rank function and calculated as the sum of the singular values of the matrix. $\|\cdot\|_p$ indicates a certain regularization strategy used under specific conditions, for example, $\|\cdot\|_F^2$ donates the squared Frobenius norm [3] which is used for characterizing the noise that slightly perturbed the data around the subspaces, $\|\cdot\|_1$ donates the l1-norm [4] which is adopted for modeling the random corruptions that grossly corrupt a fraction of random entries of the dataset, $\|\cdot\|_{2,1}$ donates the $L2,1-norm$ [5,6] which is proposed to deal with the sample specific corruptions and outliers that are far away from the subspaces of the dataset. $\lambda>0$ is a parameter used to balance the effect of noise, it is set empirically and no

Please cite this article as: M. Zhao, et al., Classification and saliency detection by semi-supervised low-rank representation, Pattern Recognition (2015), http://dx.doi.org/10.1016/j.patcog.2015.09.008