Pattern Recognition 47 (2014) 2348–2359

Contents lists available at ScienceDirect

Pattern Recognition

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

Fisher tensors for classifying human epithelial cells

Masoud Faraki [a,b], Mehrtash T. Harandi [c,d,*], Arnold Wiliem [a], Brian C. Lovell [a]

[a] The University of Queensland, School of ITEE, QLD 4072, Australia
[b] Isfahan University of Technology, Department of ECE, Isfahan, Iran
[c] NICTA, Locked Bag 8001, Canberra, ACT 2601, Australia
[d] Australian National University, Canberra, ACT 0200, Australia

ARTICLE INFO

Available online 18 October 2013

Keywords:
Riemannian manifolds
Region covariance descriptor
Bag of visual words
Fisher vectors
Human Epithelial Cell Type 2 (HEp-2)
Texture classification

ABSTRACT

Analyzing and classifying Human Epithelial type 2 (HEp-2) cells using Indirect Immunofluorescence protocol has been the golden standard for detecting connective tissue diseases such as Rheumatoid Arthritis. However, this suffers from numerous shortcomings such as being subjective as well as time and labor intensive. Recently, several studies explore the advantages of artificial systems to automate the process, not only to reduce the test turn-around time but also to deliver more consistent results. In this paper, we extend the conventional bag of word models from Euclidean space to non-Euclidean Riemannian manifolds and utilize them to classify the HEp-2 cells. The main motivation comes from the observation that HEp-2 cells can be efficiently described by symmetric positive definite matrices which lie on a Riemannian manifold. With this motivation, we first discuss an intrinsic bag of Riemannian words model. We then propose Fisher tensors which can in turn encode additional information about the distribution of the signatures in a bag of word model. Experiments on two challenging HEp-2 images datasets, namely ICPRContest and SNPHEp-2 show that the proposed methods obtain notable improvements in discrimination accuracy, in comparison to baseline and several state-of-the-art methods. The proposed framework, while hand-crafted towards cell classification, is a generic framework for object recognition. This is supported by assessing the performance of our proposal on a challenging texture classification task.

© 2013 Elsevier Ltd. All rights reserved.

1. Introduction

Indirect Immunofluorescence protocol applied on Human Epithelial Cell Type 2 (HEp-2 cells) is considered the standard way of detecting autoimmune diseases such as Systemic Lupus Erythematosus (SLE) and Rheumatoid Arthritis [1,2]. However, this protocol is labor intensive and time consuming as at least two scientists are required in the analytic step [1]. Recently, several studies have proposed cell classification systems to automatically classify HEp-2 cell images, not only to reduce the test turn-around time but also to deliver more consistent results (for example to adjudicate any discrepancy opinion between scientists) [3–12].

In this paper, we utilize COVariance Descriptor (CovD) in Bag of visual Words (BoW) [13] framework for identifying HEp-2 cell images. CovD is a generic and relatively novel image descriptor, introduced by Tuzel et al. [14] and has been successfully employed in a wide variety of visual recognition tasks including pedestrian detection [15], tracking non-rigid moving objects [16], analysing diffusion tensor images [17], face recognition [18], and action and gesture recognition [19].

Despite their appealing properties, CovDs lie on a special type of Riemannian manifold, the manifold of Symmetric Positive Definite (SPD) matrices or tensors. This in turn makes analyzing and developing inference methods on CovDs quite challenging. In recent years, attempts to generalize machineries originally designed for vector spaces to their Riemannian counterparts have received a surge of attention [15,17,19–25]. Following such trends, in this work we discuss how BoW models can be extended to their Riemannian versions.

To this end, we first elaborate on an intrinsic Bag of Riemannian Words (BoRW) model (see Section 4 for the definition of intrinsic methods). The intrinsic BoRW model takes into account the true geometry of tensors in obtaining its codebook and signatures.[1] We then introduce our main player in this work, Fisher Tensors (FT), which is the Riemannian version of Fisher vectors [26] and can encode additional information about the distribution of the signatures (as compared to BoRW). This is motivated by the recent success of Fisher vectors [26], especially on large-scale object recognition tasks, over conventional BoW models [26–28]. In FT,

* Corresponding author at: Australian National University, Canberra, ACT 0200, Australia. Tel.: +61 262676252.
E-mail addresses: m.faraki@ec.iut.ac.ir (M. Faraki), mehrtash.harandi@nicta.com.au (M.T. Harandi), a.wiliem@uq.edu.au (A. Wiliem), lovell@itee.uq.edu.au (B.C. Lovell).

[1] We use signature and histogram in BoW interchangeably in this paper.

0031-3203/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.patcog.2013.10.011

local descriptors encoded by CovD are soft-assigned to components of a Gaussian Mixture Model (GMM) and the image is represented using the gradient of the log-likelihood of the local CovDs with respect to the GMM. To achieve this, a probabilistic model (in this work GMM) should be defined over Riemannian manifolds. Our idea here is to utilize a diffeomorphism and define the probabilistic model by embedding the Riemannian manifold into a vector space. This leads to an elegant model with a significantly smaller codebook as compared to conventional BoW models. Rigorous tests on two challenging HEp-2 cell datasets, namely ICPR2012Contest [29] and SNPHEp-2 [12] show the efficiency and power of the proposed FT for classifying HEp-2 cells and texture images.

Contributions: In summary, there are five novelties in this work.

1. We propose the use of CovDs for analyzing indirect immunofluorescence images.
2. We elaborate on an intrinsic BoRW model.
3. We propose Fisher tensors, a mechanism for utilizing Fisher vectors on Riemannian manifolds.
4. We compare and contrast our proposed algorithms against several state-of-the-art methods (e.g. the methods proposed by Cordelli and Soda [9], Strandmark et al. [30], and Wiliem et al. [12]) on HEp-2 Cells classification task. Our empirical results show that the proposed approach obtains superior performance on two challenging and recently released HEp-2 Cells datasets, ICPR2012Contest and SNPHEp-2.
5. Finally, though in this work, FTs are hand-crafted towards classifying the HEp-2 cells, the proposed framework is a generic object classification. We therefore, consider the task of textures classification and show the superiority of FTs over the-state-of-the-art methods on KTH-TIPS [31] texture dataset.

We continue this paper as follows. Section 2 introduces recent advances in classifying HEp-2 cells and briefly reviews BoW models and Fisher vectors. Section 3 is dedicated to Riemannian geometry and serves as a grounding for follow-up sections. Section 4 discusses the intrinsic bag of Riemannian words model. Section 5 introduces the Fisher tensors, the Riemannian version of Fisher vectors. In Section 6 we compare the performance of the proposed method with previous approaches on several datasets. The main findings and possible future directions are summarized in Section 8.

2. Related work

In this section we discuss some of the most relevant work on cell classification systems for HEp-2 cell identification, bag of visual words, and Fisher vectors.

Systems for HEp-2 cell identification: Generally speaking, HEp-2 cells can be recognized either globally or locally. In the global methods, cell images are described by holistic features such as grey level statistics and textural features [7–11,30], whilst in the latter, statistics of local features (extracted from local patches) are used for recognition [12,32]. In addition, the holistic category can be further divided into two sub-categories, namely semi-automatic systems based on hand-picked features, and fully automatic systems that exploit feature selection methods. In the semi-automatic systems, expert knowledge and human interactions are required for selecting discriminative features (i.e., manual feature selection) [7–11,30]. One example is the work of Elbischger et al. where image statistics, cell shape and textural features, in conjunction with Nearest Neighbour (NN) classifier, are used to identify HEp-2 cells [7]. Recently, Strandmark et al. proposed ensemble learning on hand-picked features for recognising HEp-2 cells [30]. Methods classified within this category are prone to inter-laboratory settings (i.e., each laboratory may have different assays, sample preparation protocols, and acquisition hardware) [12].

To address this problem, automatic feature selection has been proposed [5,4,6]. The majority of algorithms in this category first extract a large set of features and then narrow them down to the most informative ones using feature selection methods. For example, Hiemann et al. extracted 1400 holistic features including textures and morphological features and designed an inference machinery by applying various feature selection methods [5]. Soda and Iannello proposed a multiple expert system (MES) where each system is designed and tuned to classify a specific pattern [4]. The confidence value of each sub-system is then combined to produce the final result, which can be understood as an indirect way of selecting the best features for each class. Compared to hand-picked feature systems, approaches based on automatic feature selection might be more robust to inter-laboratory settings since the system can always be retrained to a new setting by redoing the feature selection process. We refer interested readers to [33] which discusses several automatic HEp-2 cell classification systems.

Bag of visual words: Different from holistic descriptors which come with rigid spatial assumptions on images, it is possible to represent an image by a histogram of local features (i.e., BoW approach) [13]. The core idea of BoW methods is to learn a dictionary of visual words, or codebook, where visual words capture features or image patches that are more likely to occur or discriminative for the task in hand. After learning the dictionary, an image is represented by a histogram where each bin of the histogram shows the frequency of seeing the corresponding visual word in the image. Generally speaking, unlike holistic approaches, there is no need to hand pick or select features in BoW methods since robust and discriminative features can be learnt during the process [34–36]. Moreover, BoW methods have the reputation of being robust to image deformations such as image rotation and affine transformation compared to traditional holistic approaches [37,38].

Fisher vectors: Despite their tremendous simplicity and overwhelming power, conventional BoW models suffer from lack of order statistics in the signatures and a lossy quantization process. More specifically, in BoW model, each local descriptor is assigned to the nearest codeword. This can be understood as encoding the zero-order statistics from a probabilistic modeling perspective and has been studied in [26–28]. Thus, a curious mind might seek methods that bring higher-order statistics into the consideration. This school of thought leads us to the notion of Fisher Vector (FV) [26–28] which usually outperforms conventional BoW models. We note that while it is well established that large codebooks can improve the result in BoW models, assigning local descriptors to their nearest codewords is computationally expensive. Last but not least, in BoW model the descriptor quantization is a lossy process as studied for example in the work of Boiman et al. [39]. In FV, quantization will be replaced by computing the deviation of the vectors with respect to a generative probability model.

3. Background theory

In this section, we review Riemannian geometry of SPD matrices which in turn provides the grounding for techniques described in the following sections. More involved treatments on manifolds and related topics can be found in [40–42]. Throughout this paper we will use the following notation:

- $GL(d)$ is the general linear group, the group of real invertible $d \times d$ matrices.
- $Sym(d)$ is the space of real $d \times d$ symmetric matrices.
- $\mathcal{S}^d_{++}$ is the space of real $d \times d$ SPD matrices.

Download English Version:

https://daneshyari.com/en/article/530764

Download Persian Version:

https://daneshyari.com/article/530764

Daneshyari.com