ARTICLE IN PRESS
JID: PATREC [m5G;February 23, 2016;11:24]
Pattern Recognition Letters 000 (2016) 1–7
Contents lists available at ScienceDirect
Pattern Recognition Letters
journal homepage: www.elsevier.com/locate/patrec

Cell image classification by a scale and rotation invariant dense local descriptor☆

Diego Gragnaniello, Carlo Sansone*, Luisa Verdoliva
DIETI, University Federico II, Via Claudio 21, 80125 Naples, Italy

ARTICLE INFO

Article history:
Available online xxx

Keywords:
HEp-2000 cell classification
Indirect immunofluorescence images
Cell Image classification
Dense local descriptors
Bag-of-Words

ABSTRACT

This work tackles the problem of indirect immunofluorescence images classification. In particular, a dense local descriptor invariant both to scale changes and to rotations is proposed to classify six classes of staining patterns of the HEp-2 cells. In order to provide a compact and discriminative representation, the descriptor combines a log-polar sampling with spatially-varying gaussian smoothing applied on the gradients images in specific directions. Bag-of-Words is finally used to perform classification. Experimental results on the dataset provided in the recent contest hold in 2014 at ICPR show very good performance.

1. Introduction

Indirect Immunofluorescence (IIF) images are generated through the interaction of special sources of light with biological tissue. Specific patterns created by Anti-Nuclear Antibodies (ANAs) in the patient serum can be used to diagnose important autoimmune pathologies, such as systemic rheumatic diseases, multiple sclerosis and diabetes [3]. Due to the effectiveness of this methodology, the demand for IIF-based diagnostic tests targeting systemic autoimmune diseases has rapidly increased in the last few years.

Currently, the classification of the staining patterns is carried out by physicians through visual inspection. However, human evaluation is highly consuming of precious specialized-personnel time, and the result may vary strongly depending on the analyst's experience, workload, etc. [7]. In order to speed up the process, improve its accuracy, and make it fully reproducible, a computer-aided diagnosis system may be used. The main steps needed for IIF analysis are: cell segmentation [17,45], mitosis recognition [16,46], fluorescence intensity classification and, finally, staining pattern recognition. This last step has received a strong impulse thanks to recent contests [7,14] focusing on the development of effective and specific algorithms for the recognition of Human Epithelial type 2 (HEp-2) cell patterns. This is a very challenging task due to the strong within-class variability typical of these data, which can be appreciated even in the few examples of Fig. 1.

Several methods have been recently proposed in the literature to deal with this problem. They are all based on machine-learning, differing from one-another for the type of features extracted and/or the way classification is performed. The majority of methods rely on morphological or textural features. Some techniques also consider the combination of different attributes in order to take into account both global and local characteristics. Results reported in [6] show that the choice of the feature has a deeper impact on performance than the specific classification tool. In particular, all the top performing methods use textural descriptors. These features, extracted through the analysis of local neighborhood, and encoding slight variations in the underlying patterns, appear to be more discriminative than global features, based on shapes, and on the gray level distribution over the whole image.

Local descriptors are powerful tools for the analysis of images and have been successfully used for a number of tasks, such as texture classification, object recognition and image retrieval. The key idea is to analyze the image locally, in the neighborhood of each target pixel, capturing some discriminative information about the local signal behavior by means of a compact feature. Ideally, this representation should be able to characterize the region faithfully based on its intrinsic behavior, and irrespective of external factors, such as changes of illumination, scale, or view-angle. In this way, a distinctive trait of an image can be recognized even after various image transformations.

In this paper, we resort to a dense local descriptor called SID (Scale Invariant Descriptor), recently proposed by Kokkinos et al.[18]. Like the well-known SIFT (Scale-Invariant Feature

☆ This paper has been recommended for acceptance by Mario Vento.

* Corresponding author. Tel.: +39 81 768 3640; fax: +39 81 253 1726.
E-mail address: carlosan@unina.it, carlo.sansone@gmail.com, carlo.sansone@unina.it (C. Sansone).

http://dx.doi.org/10.1016/j.patrec.2016.01.007

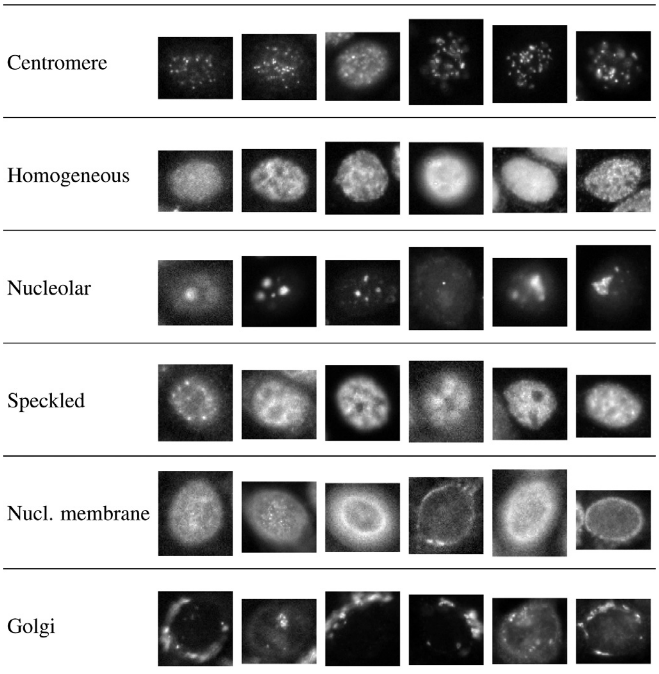

Fig. 1. Examples of cells with different staining patterns: Centromere, Homogeneous, Nucleolar, Speckled, Nuclear membrane and Golgi.

Transform) [27], it uses local image gradients as the main descriptive features. Differently from SIFT, however, it does not need to estimate the scale and the angle beforehand, since it is designed to be scale and rotation invariant by means of Fourier-domain analysis. It also shares some characteristics with the DAISY descriptor [44], like non-uniform sampling, and the convolutions of oriented gradients with spatially-varying filters. Nonetheless, thanks to a number of clever technical solutions, SID turns out to be more effective than DAISY in classification tasks [11]. It is worth underlining the interesting similarities between SID (and DAISY) and the biological mechanisms of vision. Indeed, as stated in [24], the scale-space representation which is built by convolution with a family of Gaussian kernels and derivatives of increasing width, closely resembles receptive field profiles observed in neurophysiological studies of the mammalian retina and visual cortex. This similarity is especially suggestive as we are looking for an automatic system which is able to analyze a scene and understand its content like our visual system does.

A preliminary version of the proposed technique was presented in [12], where we first analyzed the use of SID for cell classification. Here, we analyze it in more depth, adding new experiments which provide further insight into the performance and the potential of this approach.

The rest of the paper is organized as follows: Section 2 is devoted to the analysis of related work; Section 3 will describe SID while Section 4 provides a comprehensive performance assessment, also in comparison with recent state-of-the-art approaches. Finally, Section 5 draws conclusions.

2. Related work

In recent years, stimulated by some successful contests, many techniques have been proposed for the classification of cell images. Local features, describing the distribution of image micropatterns, have received a great deal of attention, especially the well-known local binary pattern (LBP) [33]. Indeed, the winner of the 2012 ICPR contest [7] used a technique based on CoALBP [32], that is, the co-occurrence of adjacent LBPs. A more discriminative descriptor is proposed in [49] by avoiding the binary quantization process needed to construct LBP. Further improvements were obtained by combining CoALBP with SIFT [43], or by using RiC-LBP [31], a rotation-invariant version of CoALBP. More in general rotation-invariance is a desirable property in a feature, as it allows us to automatically deal with different orientations of the cell images. This has motivated the work of Shen et al. [38], where a rotation invariant descriptor based on intensity order pooling is considered. Another interesting approach has been proposed by Larsen et al. [22], where shape index histograms are used for texture description, with rotation-invariant spatial pooling. An extension of this technique where the orientation of the second-order curvature is included in the histograms has been recently proposed by the same authors [21].

Unlike local features, global features are extracted by analyzing the whole image at once. They have been used as well for cell classification, aiming both at texture and shape description. In this case, some form of pre-processing (i.e. normalization, rescaling) is necessary to guarantee invariance to illumination, a problem solved automatically with local features through the use of local differences. Widespread global features are those proposed by Haralick et al. [13], consisting of 14 textural characteristics derived from the image Grey-Level Co-occurrence Matrix (GLCM). They are used both in [40] and in [2] combined with local features. In [39], instead, noting that some distinctions were visible at a specific scale, it was proposed to examine the DCT power spectrum to capture the scale of maximum textural variation, and to use information about pixel difference statistics at different scales. Shape is another important characteristic in cell images, as it is clear from Fig. 1. Therefore, many techniques [2,35,39,40] include information about the morphological properties of stained-cell regions, like the number, size, localization and shape of the cells. Some features are extracted directly from the original image, while others come from the binary image obtained after a thresholding operation.

A different approach to cell classification is followed in [50] and [8], where features are learnt automatically from the data. In the first case, a model based on the Independent Component Analysis is used, while the second technique resorts to a Convolutional Neural Network. Both methods are inspired to the multi-stage processes of the visual cortex and show promising results.

Some of the best performing methods use the concatenation of several features, which can be expected to capture different characteristics of the cell images. This is the case, for example, of the technique [28] that ranked first in the 2014 ICPR contest [26]. It uses a combination of non-binarized LBP, a variant of SIFT, Random Projections and Intensity histograms. Local and global textural descriptors are also combined with some ad hoc features in [40] and [42]. Instead, only textural local features (MR8, LBP and HOG) are considered in [20]. It is worth noting that for positive intensity HEp-2 IIF images it has been shown that using simple statistical and histogram features allows to obtain a classification accuracy comparable to more complex methods [10].

Although the choice of the features turns to be the fundamental step, the classification phase plays an important role too. When a high-dimensionality feature vectors is extracted for each pixel (i.e. SIFT) it is necessary to encode the local features in global image statistics. The Bag-of-Words paradigm is typically used for this task [4]. The baseline method (hard assignment) can be improved by using alternative encodings that retain more information about the original image features [1]. Different approaches have been considered, such as soft assignment and Fisher encoding [5]. Spatial pyramid matching [23] has also been used to take into account the correlation between feature vectors [38], and combined with a Dual-Region structure for cell classification [47]. Discriminative sparse representation for classification is instead used in [43].

Please cite this article as: D. Gragnaniello et al., Cell image classification by a scale and rotation invariant dense local descriptor, Pattern Recognition Letters (2016), http://dx.doi.org/10.1016/j.patrec.2016.01.007