Adaptive Co-occurrence Differential Texton Space for HEp-2 Cells Classification

Xiang Xu[1], Feng Lin[1], Carol Ng[2], and Khai Pang Leong[2]

[1] School of Computer Engineering, Nanyang Technological University, S-639798
[2] Department of Rheumatology, Allergy and Immunology, Tan Tock Seng Hospital, S-308433
{xxu4,asflin}@ntu.edu.sg, {carol_ng,khai_pang_leong}@ttsh.com.sg

Abstract. The occurrence of antinuclear antibodies in patient serum has significant relation to autoimmune disease. But identification of them suffers from serious problems due to human subjective evaluation. In this study, we propose an automatic classification system for HEp-2 cells. Within this system, a Co-occurrence Differential Texton (CoDT) feature is designed to represent the local image patches, and a generative model is built to adaptively characterize the CoDT feature space. We further exploit a more discriminant representation for the HEp-2 cell images based on the adaptive partitioned feature space, then feed the representation into a linear Support Vector Machine classifier for identifying the staining patterns. The experimental results on two benchmark datasets: ICPR12 dataset and ICIP2013 training dataset, verified that our method remarkably outperforms the other contemporary approaches for HEp-2 cells classification.

Keywords: HEp-2 cells, co-occurrence differential texton, generative model, adaptive partitioned feature space.

1 Introduction

Indirect-immunofluorescence (IIF) is the most recommended technique for detecting antinuclear antibodies (ANAs) in patient serum, which can reveal the occurrence of specific autoimmune diseases such as rheumatoid arthritis and multiple sclerosis. In the current clinical practices, IIF slide images are manually inspected by physicians with a fluorescence microscope, therefore it suffers from some intrinsic limitations due to subjective evaluation. Computer Aided Diagnostic (CAD) systems are proposed for automatically supporting the IIF diagnosis. The main technologies investigated in these CAD systems are automated preparation of slides with robotic devices, image acquisition, image segmentation, mitotic cell recognition, fluorescence intensity classification and staining pattern recognition. Staining pattern recognition is proven to be the most challenging task in the research community. In this study, we investigate into the approaches for automatic staining pattern classification of HEp-2 cells.

As a benchmark to evaluate and compare the new methods, a publicly available HEp-2 cells dataset was released at the first edition of the *HEp-2 Cells Classification Contest*. Nosaka et al. [1] propose an extension of Local Binary Pattern (LBP) descriptor (named CoALBP) to extract textural features and win the first prize of the contest. All the participated methods are reported in [2]. Inspired by the contest, Xu et al. [3] improve the coding method of Bag-of-Words (BoW) framework to reduce information loss

N. Navab et al. (Eds.): MICCAI 2015, Part III, LNCS 9351, pp. 260–267, 2015.
DOI: 10.1007/978-3-319-24574-4_31

caused by feature quantization. Shen et al. [4] adopt the BoW framework on intensity order pooling based gradient feature. Theodorakopoulos et al. [5] fuse the distribution of SIFT features [6] and gradient-oriented co-occurrence of LBPs into a dissimilarity space. Then they use a sparse representation-based classification mechanism for classification.

Although considerable progress has been made, research in HEp-2 cell image analysis is still in its early stage. In this study, we first propose a Co-occurrence Differential Texton (CoDT) feature to represent local patches of the HEp-2 cell images. LBP related features have been applied successfully in the HEp-2 cells classification [1], [7]. However, some important information is lost since the LBP represents the local structures with only two quantized levels. Our proposed CoDT feature reduces the information loss by ignoring the quantization. Furthermore, it captures the spatial relations among the differential micro-texton features to increase the discriminative power of features.

Then, we apply a Gaussian Mixture Model (GMM) to adaptively approximate the distribution of the proposed CoDT features. Thus the parameters, adjusted from the training cell images, are better fitting the CoDT feature space.

Last, we utilize the *Fisher Kernel* (FK) principal [8] to improve the BoW framework. The BoW framework is one of the most popular approaches for image classification. However, it suffers from some problems: (i) Information loss in feature quantization process is inevitable [9]; (ii) The cost of histogram computations depends on the number of visual words. Since better performance is always obtained with larger vocabulary, the computational cost is high. The FK based methods can handle these problems. The output image representation is fed into a linear SVM for final classification.

Our proposed framework (AdaCoDT) can exploit the advantages of both generative and discriminative approaches for image classification. Experimental results verify that AdaCoDT can provide remarkable classification performance for HEp-2 cells.

2 Method

2.1 Co-occurrence Differential Texton

LBP [10] can be obtained by thresholding the gray value of the circularly symmetric surrounding pixels with that of the center pixel within a local patch (micro-texton). The LBP at location (x, y) is defined as

$$LBP_{P,R}(x,y) = \sum_{i=1}^{P} 2^{i-1} sign(I(x_i, y_i) - I(x,y)), \tag{1}$$

$$sign(x) = \begin{cases} 1\,, & \text{if} \quad x \geq 0 \\ 0\,, & \text{otherwise.} \end{cases} \tag{2}$$

where $I(x, y)$ is a gray value at location (x, y) in image I and $I(x_i, y_i)$ denotes the gray value of P equal spaced pixels on a circle of radius R around center pixel (x, y).

Recently, some improved LBP features [1] have been applied in HEp-2 cells and shown superior performances compared with the conventional LBP. However, one major drawback of the LBP related methods is that they will lose some discriminant

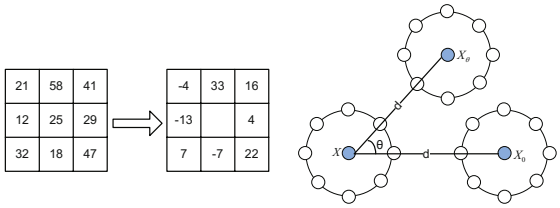

Fig. 1. Illustration of CoDT. (a) an example of Differential Vector (3×3 micro-texton). (b) two pairs of DVs with rotation angles 0° and θ respectively.

information since they represent the microstructure with only two quantized levels (i.e. 0 and 1). To reserve more discriminant information, we propose to use *Differential Vectors* (DV) to describe the micro-texton. A DV is a microstructural feature based on the differential domain skipping the quantization procedure, which is formulated as

$$DV_{P,R}(x,y) = (I(x_1,y_1) - I(x,y), ..., I(x_P,y_P) - I(x,y)) \tag{3}$$

To enhance the discriminative power, we further propose a Co-occurrence Differential Texton (CoDT) feature capturing the spatial relation between differential micro-textons. The CoDT feature can provide more information than individual DV since it characterizes more subtle and complex structure. The CoDT feature with one pair of DVs is illustrated in Fig 1 and formulated as

$$CoDT_{P,R,d}(\mathbf{x}) = [DV_{P,R}(\mathbf{x}), DV^{\theta}_{P,R}(\mathbf{x}_\theta)] = [DV_{P,R}(\mathbf{x}), DV_{P,R}(\mathbf{x} + \Delta\mathbf{x}_\theta)] \tag{4}$$

where $\mathbf{x} = (x,y)$ is the position vector in I and $\Delta\mathbf{x}_\theta = (d\cos\theta, d\sin\theta)$ is a replacement vector between a DV pair with interval d and rotation angle θ.

In this study, following the commonly practiced rules, we extract four pairs of DVs, that is $\theta = 0^\circ, 45^\circ, 90^\circ, 135^\circ$ (we find that the classification performance improves very little by using more pairs of DVs). The CoDT feature is a $5P$-dimensional feature vector.

2.2 HEp-2 Cell Image Representation in the Adaptive CoDT Feature Space

To combine the strengths of both generative and discriminative approaches for image classification, we characterize the proposed CoDT features of a HEp-2 cell image by a gradient vector derived from a generative model, then we feed the output image representations into a discriminative classifier for the identification of HEp-2 cells.

Let $X = \{x_n, n = 1,2,...,N\}$ be a set of samples from the CoDT feature space of one HEp-2 cell image. The probability density distribution of the CoDT feature is described by a GMM with parameters $\lambda = \{w_t, \mu_t, \Sigma_t, t = 1,2,...,T\}$, where w_t, μ_t and Σ_t are respectively the mixture weight, mean vector and covariance matrix of Gaussian t. Then we can formulate

$$p(x_n|\lambda) = \sum_{t=1}^{T} w_t p_t(x_n|\lambda), \quad s.t. \quad \sum_{t=1}^{T} w_t = 1 \tag{5}$$

where $p_t(x_n|\lambda) = \frac{exp\{-\frac{1}{2}(x_n-\mu_t)\Sigma_t^{-1}(x_n-\mu_t)\}}{(2\pi)^{D/2}|\Sigma_t|^{1/2}}$. The parameters of GMM can be adaptively estimated by Expectation Maximization (EM) algorithm [11]. Note that we assume that the covariance matrices are diagonal and denoted by $\sigma_t = diag(\Sigma_t)$.

The samples X can be characterized by $G_\lambda(X) = \bigtriangledown_\lambda \log p(X|\lambda)$. The gradient describes how the parameters λ should be modified to best fit X. To measure the similarity between two HEp-2 cell images, a *Fisher Kernel* (FK) [8] is calculated as

$$K_F(X,Y) = G_\lambda^T(X)F_\lambda^{-1}G_\lambda(Y) \tag{6}$$

where $F_\lambda = E_X[G_\lambda(X)G_\lambda^T(X)]$ is the *Fisher Information Matrix* (FIM).

As F_λ is symmetric and positive semi-definite, and F_λ^{-1} has the Cholesky decomposition $F_\lambda^{-1} = L_\lambda^T L_\lambda$, the FK can be re-defined as $K_F(X,Y) = \mathscr{G}_\lambda^T(X)\mathscr{G}_\lambda(Y)$, where $\mathscr{G}_\lambda(X) = L_\lambda \bigtriangledown_\lambda \log p(X|\lambda)$. We only consider the gradients with respect to the mean and covariance since the gradient with respect to the weights brings little additional information [12].

Let $\zeta(t)$ be the occupancy probability of the CoDT feature x_n for the t-th Gaussian:

$$\zeta_n(t) = \frac{w_t p_t(x_n|\lambda)}{\sum_{k=1}^{T} w_k p_k(x_n|\lambda)} \tag{7}$$

The normalized gradients are finally computed as

$$\mathscr{G}_{\mu_t^d}(X) = \frac{1}{\sqrt{w_t}} \sum_{n=1}^{N} \zeta_n(t)\left(\frac{x_n^d - \mu_t^d}{\sigma_t^d}\right) \tag{8}$$

$$\mathscr{G}_{\sigma_t^d}(X) = \frac{1}{\sqrt{w_t}} \sum_{n=1}^{N} \zeta_n(t)\frac{1}{\sqrt{2}}\left[\frac{(x_n^d - \mu_t^d)^2}{(\sigma_t^d)^2} - 1\right] \tag{9}$$

The Fisher representation is the concatenation of all the gradients $\mathscr{G}_{\mu_t^d}(X)$ and $\mathscr{G}_{\sigma_t^d}(X)$ for $d = 1, 2, ..., D$ dimension of the CoDT feature and for T Gaussians.

The proposed Adaptive CoDT (AdaCoDT) method has several advantages over the BoW framework. First, as a generalization of the BoW framework, the resulting representation is not limited to the number of occurrences of each visual word. It also encodes additional information about the distribution of features. Secondly, it reduces the information loss raised by the coding procedure of the BoW framework. Thirdly, it can be computed from much smaller codebooks therefore it reduces the computational cost. Lastly, with the same size of codebook, it is much larger than the BoW representation. Hence, it assures an excellent performance with a simple linear classifier [12].

3 Experiments and Comparisons

3.1 Datasets

We evaluate the performance of the proposed method for HEp-2 cells classification on two HEp-2 cell datasets. **The ICPR2012 dataset** was released on the ICPR'12 *HEp-2*

Table 1. Parameters for comparative algorithms.

Algorithm	$(\mathbf{P}, \mathbf{R})$ or $(\mathbf{P}, \mathbf{R}, \mathbf{d})$	T
FK-SIFT	/	128
CoALBP	(4,1,2),(4,2,4),(4,4,8)	/
RICLBP	(4,1,2),(4,2,4),(4,4,8)	/
LBP	(8,1),(12,2),(16,3)	/

cell classification contest. It consists of 1455 cells, which are manually segmented from 28 slide images. Each cell is identified as one of six patterns: centromere (ce), coarse speckled (cs), cytoplasmic (cy), fine speckled (fs), homogeneous (ho) and nucleolar (nu). According to the experimental protocol of the contest, the dataset is divided into a training set with 721 cells from 14 slide images and a test set with 734 cells from 14 slide images. **The ICIP2013 training dataset** was used as the training dataset in the ICIP'13 *Competition on cells classification by fluorescent image analysis*. It contains 13596 cells which are categorized into six classes: homogeneous (ho), speckled (sp), nucleolar (nu), centromere (ce), nuclear membrane (nm) and golgi (go). We partition the dataset into a training set containing 6842 cells from 42 slide images and a test set containing 6754 cells from 41 slide images. Some examples are shown in Fig 2.

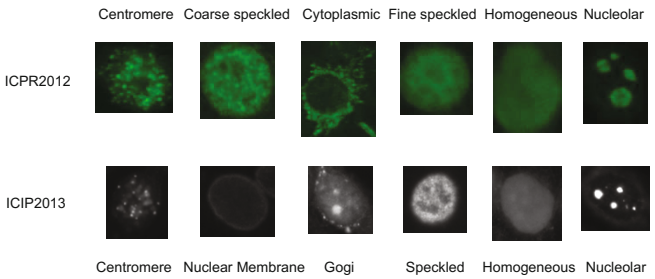

Fig. 2. Sample images of ICPR2012 and ICIP2013 datasets.

3.2 Experimental Results

We quantitatively compare the classification performance achieved by our proposed AdaCoDT framework, with LBP [10], CoALBP [2] (the winner of ICPR'12 contest), RICLBP [1], LSC based on dense SIFT(LSC-SIFT) [13], LSC based on CoDT (LSC-CoDT) and the FK based on dense SIFT (FK-SIFT). The parameters for each method are set as Table 1 which are optimized manually via several trials. With respect to LBP-related features, P is the number of neighbor pixels, R is the the radius and d is the interval between the LBP pair. The number of GMM components T is another parameter to be considered for the FK based methods. The codebook size of LSC method is 1024 due to the trade-off between classification accuracy and computational cost.

Performance Assessment on the ICPR2012 Dataset: We choose parameters $(P, R, d) = (24, 4, 8)$ and $T = 256$ for the proposed AdaCoDT via several trials. We utilize linear SVM due to its effectiveness and efficiency. It is trained using the training set by 10-fold cross validation strategy. The *one-vs-all* approach is used to handle our multi-class problem (the same strategy is used for ICIP2013 training dataset). Table 2(a) shows the classification performance of each method at the cell level. The AdaCoDT method outperforms all the other methods, achieving 75.2% of classification accuracy. The obtained accuracy is even higher than that of a human expert and significantly outperforms CoALBP which is the winner of the contest.

Table 3(a) illustrates the confusion matrices presenting the classification performance for each staining pattern at the cell level. It is obvious that cytoplasmic, centromere and homogeneous patterns are classified more accurately than the others. More particularly, cytoplasmic can achieve 100% of classification accuracy. To evaluate the classification performance at the image level, we report the corresponding confusion matrix in Table 3(b). The prediction for staining pattern of each image is decided by the most frequently assigned pattern of the cells within that image. The proposed AdaCoDT method obtains the classification accuracy of 85.7%, which indicates that 12 images are correctly classified while there are 14 images in the test set.

Table 2. Classification performance at the cell level.

(a) ICPR2012 dataset

Algorithm	Accuracy	Sensitivity
AdaCoDT	**75.2%**	**77.1%**
Human [2]	73.3%	/
LSC-SIFT	68.1%	69.4%
LSC-CoDT	66.9%	66.5%
FK-SIFT	66.6%	66.7%
CoALBP [2]	70.4%	68.7%
RICLBP [1]	68.5%	67.5%
LBP	58.9%	59.2%

(b) ICIP2013 training dataset

Algorithm	Accuracy	Sensitivity
AdaCoDT	**75.8%**	**72.9%**
LSC-SIFT	73.2%	71.9%
LSC-CoDT	70.6%	69.8%
FK-SIFT	69.7%	68.3%
CoALBP	67.1%	65.5%
RICLBP	66.4%	64.4%
LBP	60.7%	54.5%

Table 3. Confusion matrix for the ICPR2012 dataset via our proposed AdaCoDT method.

(a) The cell level (%)

	ce	cs	cy	fs	ho	nu
ce	**85.9**	8.1	0.0	0.0	0.0	6.0
cs	4.0	**75.3**	2.9	17.8	0.0	0.0
cy	0.0	0.0	**100.0**	0.0	0.0	0.0
fs	20.2	3.5	6.2	**52.6**	17.5	0.0
ho	8.3	2.8	0.6	11.1	**73.9**	3.3
nu	2.2	0.0	11.5	3.6	7.9	**74.8**

(b) The image level (%)

	ce	cs	cy	fs	ho	nu
ce	**100.0**	0.0	0.0	0.0	0.0	0.0
cs	0.0	**66.7**	0.0	33.3	0.0	0.0
cy	0.0	0.0	**100.0**	0.0	0.0	0.0
fs	**50.0**	0.0	0.0	**50.0**	0.0	0.0
ho	0.0	0.0	0.0	0.0	**100.0**	0.0
nu	0.0	0.0	0.0	0.0	0.0	**100.0**

Performance assessment on the ICIP2013 training dataset: We choose parameters $(P, R, d) = (16, 5, 10)$ and $T = 128$, which are optimized via several trials, for the

proposed AdaCoDT. The classification performance of different methods at the cell level are shown in Table 2(b). Our proposed AdaCoDT method achieves the best performance again. It is also worth noting that the size of codebook for the BoW framework is 1024 while the number of GMM components for the AdaCoDT method is only 128. With the same codebook size (the number of GMM components can be seen as the codebook size), the AdaCoDT method significantly outperforms the BoW framework.

Table 4(a) shows the confusion matrix of the AdaCoDT method at the cell level. Homogeneous pattern gets the highest classification accuracy rate of 89.5%, followed by nuclear membrane as they have distinguished characteristic compared with other patterns. Table 4(b) illustrates the confusion matrix at the image level. The AdaCoDT method obtains the classification accuracy of 87.8% at image level, which means that 36 images are correctly identified while there are 41 images in the test set.

Table 4. Confusion matrix for the ICIP2013 training dataset via our proposed AdaCoDT method.

(a) The cell level (%)

	ho	**sp**	**nu**	**ce**	**nm**	**go**
ho	**89.5**	3.4	4.1	0.3	2.4	0.3
sp	11.6	**66.7**	5.6	15.1	0.7	0.3
nu	0.8	8.1	**74.2**	11.7	2.6	2.6
ce	0.5	22.2	2.4	**74.7**	0.0	0.2
nm	1.1	2.4	1.1	0.4	**88.7**	6.3
go	6.6	5.0	38.2	0.8	5.8	**43.6**

(b) The image level (%)

	ho	**sp**	**nu**	**ce**	**nm**	**go**
ho	**100.0**	0.0	0.0	0.0	0.0	0.0
sp	12.5	**75.0**	0.0	12.5	0.0	0.0
nu	0.0	0.0	**100.0**	0.0	0.0	0.0
ce	0.0	25.0	0.0	**75.0**	0.0	0.0
nm	0.0	0.0	0.0	0.0	**100.0**	0.0
go	0.0	0.0	**50.0**	0.0	0.0	**50.0**

4 Conclusion

In this study, we have presented a promising framework, AdaCoDT, for automatic staining pattern classification of HEp-2 cells. First, we propose a CoDT feature which directly uses the differential vectors of micro-texton and its neighborhoods to reserve more discriminative information. It further captures the spatial information between neighboring micro-textons to provide strong discriminative and descriptive capability. Then, we approximate the distribution of CoDT feature as a GMM which can adaptively partition the CoDT feature space for the classification task of HEp-2 cells. Finally, we obtain a high discriminative and powerful descriptive HEp-2 cell image representation based on the adaptive CoDT feature space using FK principle. We feed the image representation into a linear SVM classifier to predict staining patterns of the HEp-2 cells. The AdaCoDT method combines the strengths of generative and discriminative approaches for image classification, therefore it can achieve excellent classification performance. Experimental results validate that the proposed AdaCoDT method can provide superior performance for HEp-2 cells classification, compared with the traditional LBP and its extensions. The new feature encoding method also significantly improves the classification performance in comparison of the BoW representation.

Acknowledgment. This work is partially supported by two research grants, MOE2011-T2-2-037 and RG139/14 from Ministry of Education, Singapore.

References

1. Ryusuke, N., Kazuhiro, F.: HEp-2 cell classification using rotation invariant co-occurrence among local binary patterns. Pattern Recogn. 47(7), 2428-2436 (2014)
2. Foggia, P., Percannella, G., Soda, P., Vento, M.: Benchmarking HEp-2 cells classification methods. IEEE Trans. Med. Imaging 32(10), 1878-1889 (2013)
3. Xu, X., Lin, F., Carol, Ng, Leong, K.P.: Linear local distance coding for classification of HEp-2 staining patterns. In: IEEE Winter Conference on Applications of Computer Vision (WACV), pp. 393-400 (2014)
4. Shen, L.L., Lin, J.M., Wu, S.Y., Yu, S.Q.: HEp-2 image classification using intensity order pooling based features and bag of words. Pattern Recogn. 47(7), 2419-2427 (2014)
5. Theodorakopoulos, I., Kastaniotis, D., Economou, G., Fotopoulos, S.: Hep-2 cells classification via sparse representation of textural features fused into dissimilarity space. Pattern Recogn. 47(7), 2367-2378 (2014)
6. Xu, X., Lin, F., Carol, Ng, Leong, K.P.: Staining Pattern Classification of ANA-IIF Based on SIFT Features. J. Medical Imag. Heal. Inform. (2)4, 419-424 (2012)
7. Nanni, L., Paci, M., Brahnam, S.: Indirect immunofluorescence image classification using texture descriptors. Expert Syst. Appl. 41(5), 2463-2471 (2014)
8. Perronnin, F., Dance, C.: Fisher kernels on visual vocabularies for image categorization. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1-8 (2007)
9. Boiman, O., Shechtman, E., Irani, M.: In defense of nearest-neighbor based image classification. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1-8 (2008)
10. Ojala, T., Pietikainen, M., Maenpaa, T.: Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Trans. Pattern Anal. 24(7), 971-987(2002)
11. Dempster, A.P., Laird, N.M., Rubin, D.B.: Maximum likelihood from incomplete data via the EM algorithm. J. R. Stat. Soc. Series B (Methodological) 1-38 (1977)
12. Sánchez, J., Perronnin, F., Mensink, T., Verbeek, J.: Image classification with the Fisher vector: Theory and practice. Int. J. Comput. Vision 105(3), 222-245 (2013)
13. Liu, L.Q., Wang, L., Liu, X.W.: In defense of soft-assignment coding. In: IEEE International Conference on Computer Vision (ICCV), pp. 2486-2493 (2011)