21st International Conference on Pattern Recognition (ICPR 2012)
November 11-15, 2012. Tsukuba, Japan

HEp-2 Cell Classification in IIF Images Using ShareBoost

I. Ersoy[1], F. Bunyak[1], J. Peng[2], K. Palaniappan[1]
[1]*Department of Computer Science, University of Missouri-Columbia, Columbia, MO 65211 USA*
[2]*Department of Computer Science, Montclair State University, Montclair, NJ 07043 USA*

Abstract

Indirect immunofluorescence (IIF) imaging is a method used for detection of antinuclear autoantibodies (ANA) for the diagnosis of autoimmune diseases. We present a feature extraction and classification scheme to classify the fluorescence staining patterns of HEp-2 cells in IIF images. We propose a set of complementary features that are sensitive to staining pattern variations among classes. Our feature set utilizes local shape measures via Hessian matrix, gradient features using our adaptive robust structure tensors and texture features. We apply our multi-view ShareBoost algorithm to this set using each feature descriptor as a separate view. ShareBoost utilizes a single re-sampling distribution for all views that helps the classifier to exploit the interplay between subspaces and is robust to noisy labels. Our experimental results show an average of over 90 percent accuracy in classification of six HEp-2 cell types.

1. Introduction

The distinct fluorescence patterns of HEp-2 cells in indirect immunofluorescence imaging serve as markers for a broad range of autoantibodies that are related to autoimmune disorders. HEp-2 cells are analyzed by specialists in order to report fluorescence intensity and several different ANA staining patterns that can be used to identify specific autoimmune diseases. Recently, there are studies in the literature to develop automated expert systems in order to support computer-aided diagnosis using a reliable and standardized HEp-2 cell classification to reduce observer variability [3, 5, 15]. Although there are many different visual staining patterns, they are classified into one of the following six groups: (i) centromere (CE): several discrete speckles (40-60) distributed throughout the interphase nuclei; (ii) coarse speckled (CS): a coarse granular nuclear staining of the interphase cell nuclei; (iii) cytoplasmatic (CY): fine fluorescent fibers running the length of the cell; (iv) fine speckled (FS): a fine granular nuclear staining of the interphase cell nuclei; (v) homogeneous (HO): diffuse staining of the interphase nuclei and staining of the chromatin of mitotic cells; (vi) nucleolar (NU): large coarse speckled staining within the nucleus, less than six in number per cell. Examples of these patterns are shown in Figure 1. In order to reliably classify cell images into these six classes, features that exploit the recurring blob structures and textures need to be extracted. These features have to complement each other in a way to increase the interclass separability, and also provide robustness to variations in noise and intensity levels that are common in fluorescence microscopy. We propose the use of eight feature descriptors that capture different complementary qualities of these six classes, and fit nicely within a multi-view boosting framework for reliable classification. Section 2 describes the proposed feature descriptors and their extraction, Section 3 describes our multi-view boosting algorithm that is robust to feature and label noise, Section 4 gives experimental results on the ICPR 2012 HEp-2 cell classification contest, and Section 5 concludes the paper.

2 Feature Descriptors

Characteristics of the data set, namely low resolution, lack of color, image noise, and small support maps limit the use of more complex multiscale features. Thus we rely on a set of simpler complementary feature descriptors accounting for varying distributions of intensity, texture and structural information (i.e. ridges, blobs) among different cell classes. We use binned histograms for descriptors to represent all features similar to our previous work on cell cycle phase classification [2]. For example, instead of deriving statistics from intensity (e.g. mean, variance, skewness, higher order moments etc.) we directly use the intensity histogram, since it models the intensity probability distribution for a local image region. This has two advantages: the classifier can decide which bins in the histograms have more discriminative power, and no prior parametric model (e.g. Gaussian) is imposed. Besides the intensity his-

togram, features used in our system can be grouped into edge-based, local shape-based, and texture-based features similar to [6, 7, 10]. This set of features distinguish between pairs of classes by utilizing homogeneity of intensities, distribution of blob structures, and texture differences.

Edge-based features capture both gradient magnitude and orientation information. Gradient magnitude histograms are computed similar to their intensity counterparts. Gradient orientation information is used through histogram of oriented gradients (HOG) descriptor that has been successfully used in many recent object recognition applications. HOG computes the histogram of gradient orientations weighted by corresponding gradient magnitudes over a patch or small region of an image. In this case, the cell mask is used as the support region. Robust orientation estimation is important for HOG-like descriptors. We also compute ARST-HOG, our extension to HOG descriptors, that combines the descriptive power of HOG with our earlier work on adaptive robust structure tensor (ARST) for improved orientation estimation [8]. Structure tensors are a useful tool for reliably estimating oriented structures within a neighborhood and in the presence of noise. Our ARST approach is an iterative scheme that adaptively and simultaneously, changes the size, orientation and weighting of the neighborhood used to estimate the local structure tensor. ARST-HOG outperforms standard HOG in some applications [10].

Local shape-based features are incorporated through three measures obtained from the Hessian matrix. Hessian matrix $\mathcal{H}$ describes the second order structure of local intensity variations around each point of the image $L(x, y)$.

$$\mathcal{H} = \begin{bmatrix} \frac{\partial^2 L}{\partial x^2} & \frac{\partial^2 L}{\partial x \partial y} \\ \frac{\partial^2 L}{\partial x \partial y} & \frac{\partial^2 L}{\partial y^2} \end{bmatrix} = \begin{bmatrix} L_{xx} & L_{xy} \\ L_{xy} & L_{yy} \end{bmatrix} \quad (1)$$

Different shape patterns (e.g. ridges, valleys, dark and bright blobs) can be described based on the value of the eigenvalues λ_1 and λ_2 of the Hessian matrix $\mathcal{H}$.

$$\lambda_{1,2} = \frac{1}{2}(L_{xx} + L_{yy} \pm \sqrt{(L_{xx} - L_{yy})^2 + (2L_{xy})^2})$$

First two local shape features are shape index (Eq.2) and normalized curvature index (Eq.3) derived from eigenvalues of intensity image Hessian matrix.

$$\theta(x, y) = tan^{-1} \frac{\lambda_1(x, y)}{\lambda_2(x, y)} \quad (2)$$

$$\phi(x, y) = \tan^{-1} \frac{\sqrt{\lambda_1^2 + \lambda_2^2}}{1 + I(x, y)} \quad (3)$$

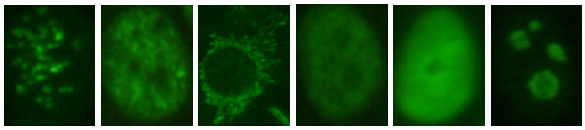

Figure 1: Six HEp-2 cell classes. Left to right: centromere(CE), coarse speckled(CS), cytoplasmatic(CY), fine speckled(FS), homogeneous(HO), nucleolar(NU).

The third measure is magnitude weighted histogram of Hessian matrix eigenvector orientations (HEV). This is a descriptor similar to HOG but the orientations correspond not to gradient vectors but to Hessian matrix eigenvectors.

Texture information can be incorporated through LBP or MBP histograms [4, 9]. Local binary pattern (LBP) operator is defined as a gray-scale invariant texture measure, derived from a general definition of texture in a local neighborhood, that unifies the statistical and structural models of texture analysis. We use uniform rotation-invariant LBP consisting of 18 unique patterns.

3. Classification with ShareBoost

Classifiers employed in real world applications must deal with various adversities such as noise in features as well as labels, and intra-class variations. It is often helpful to develop classifiers that rely on features from various sources (views) for classification. Such classifiers require an effective way of fusing the various sources of information. Earlier, we presented a novel shared sampling approach to boosting for learning from multiple representations (views) of data [11]. In this context, a view corresponds to a specific type of feature or attribute such as texture, edge and shape features. The proposed algorithm builds base classifiers independently from each view. Different from AdaBoost, where each view has its own re-sampling weight, our algorithm uses a single re-sampling distribution for all views at each boosting round. This distribution is determined by the view whose training error is minimal. This shared sampling mechanism restricts noise to individual views, thereby reducing sensitivity to noise. In addition, since the final strong classifier contains classifiers that are trained to focus on different views of the data, better generalization performance can be expected. Here, we briefly describe the shared sampling (ShareBoost) algorithm.

We are given a set of training samples: $X = \{(\mathbf{x}_i, y_i)\}_{i=1}^n$, and M disjoint views (features) for each example $x_i = \{x_i^1, \cdots, x_i^M\}$, where $x_i^j \in \mathbb{R}^{q_j}$, and $y_i \in \mathcal{Y} = \{-1, +1\}$. Each member x_i^j is known as a

view of sample x_i. We assume that samples (x_i, y_i) are drawn randomly and independently according to a fixed but unknown probability distribution D over $\mathcal{X} \times \mathcal{Y}$. Here the input space $\mathcal{X}$ is $\mathbb{R}^q$, where $q = \sum_{j=1}^{M} q_j$. The algorithm builds weak classifiers independently from each view. However, all data types share the same sampling distribution computed from the view having the smallest error rate. In this sense, ShareBoost possesses the ability to decide at each iteration which view to influence its final decision. This ability goes beyond simple subspace selection. It empowers ShareBoost not only to exploit the interplay between subspaces, but also to be more robust against noise. The key steps of the algorithm are shown below, where $\mathbb{I}(\cdot)$ is the indicator function that returns 0 or 1, T is number of base classifiers (i.e. iterations) and α_t is used to boost misclassified sample weights.

ShareBoost $(\{(x_i^j, y_i)\}_{i=1}^n)$

1. **Initialization:** $w_1(i) = \frac{1}{n}, i = 1, \cdots, n.$
2. **For** $t = 1$ **to** T
   (a) Compute base classifier h_t^j using distribution w_t
   (b) Calculate:$\epsilon_t^j = \sum_i w_t(i)\mathbb{I}(h_t^j(x_i^j) \neq y_i)$ and $\alpha_t^* = \frac{1}{2}\ln(\frac{1-\epsilon_t^*}{\epsilon_t^*})$, where $\epsilon_t^* = \min_j\{\epsilon_t^j\}$ with corresponding h_t^*
   (c) Update $w_{t+1}(i) = \frac{w_t(i)}{Z_t^*} \times \exp(-y_i h_t^*(x_i^*)\alpha_t^*)$, where Z_t^* is a normalization factor.
3. **Output:** $H(x) = sign(\sum_{t=1}^{T} \alpha_t^* h_t^*(x))$

---

4. Experimental Results

We use the data set reported in [3] and made available with the ground truth for the ICPR 2012 HEp-2 cell classification contest[1] to evaluate our feature extraction and classification algorithms for HEp-2 staining pattern classification. HEp-2 images that were acquired from a fluorescence microscope (40-fold magnification) coupled with a 50W mercury vapor lamp and with a digital camera (CCD with square pixel of 6.45 m) were manually segmented and annotated by specialists. Out of 28 image frames, 14 are made available as training set that contains 721 individual grayscale cell images and corresponding cell masks. A biomedical engineer manually segmented the cells and subsequently, each image was verified and annotated by a medical doctor specialized in immunology to report fluorescence intensity (negative, intermediate, positive) and staining pattern (homogeneous, fine speckled, coarse speckled, nucleolar, cytoplasmatic or centromere). First, we apply anisotropic diffusion for noise reduction and normalize each grayscale cell image independently, thus we do not utilize fluorescence intensity annotations as part of our feature vector. Our feature vector consists of 8 feature descriptors as described in Section 2, namely binned histograms of intensity, gradient magnitude, HOG, ARST-HOG, HEV, θ, ϕ and LBP. Each histogram except LBP (18 bins) is binned into 16 bins, that results in a 130 dimensional vector. These eight descriptors are used as eight views in our ShareBoost algorithm that uses a regression tree as weak base classifier. We run 150 iterations of boosting in a 5-fold cross validation scenario with stratified sampling to obtain average classification accuracies. Experimentation with different number of bins and different number of folds and training/testing set sizes showed no significant difference in classification accuracies. Since the ShareBoost algorithm is given for binary classification problem, we apply two schemes to create a multi-class classifier, namely one-per-class and error correcting output codes (ECOC) decoding. In one-per-class scheme, we train one classifier for each of the six classes. In ECOC scheme, we use six 6-bit codes that represent class labels and have at least a Hamming distance of three in between them to tolerate single bit errors. Both schemes obtain comparable accuracies, only former is shown in Table 1. We also give a comparison to Wndchrm [14], a recent classification tool for biological image analysis. Wndchrm computes a 1025-D feature vector comprising of numerous complex features in order to capture different qualities among a wide variety of biological images. We do not modify its default parameters and the results are shown as averages of a 5-split training and testing scheme where top 20 percent of the features are selected. Our method outperforms Wndchrm significantly as shown in Table 2.

5. Conclusion

We present a feature set and classification scheme for HEp-2 cell classification that captures the appearance variability of HEp-2 cells. The eight feature descriptors extracted from the training data are utilized as eight views in our multi-view ShareBoost algorithm. ShareBoost uses a single re-sampling distribution which is determined by the view whose training error is minimal. This scheme reduces sensitivity to feature and label noise as shown in [11] and the final strong classifier has increased generalization performance. Differ-

[1] http://mivia.unisa.it/hep2contest/index.shtml

Table 1: Confusion matrix for proposed features and one-per-class ShareBoost (rounded percentages). Overall average accuracy is 92.8%.

	CE	CS	CY	FS	HO	NU
CE	**94**	2	0	0	1	1
CS	1	**97**	0	2	0	0
CY	5	0	**95**	0	0	0
FS	3	0	0	**85**	12	0
HO	1	0	1	5	**93**	0
NU	6	2	0	1	0	**91**

Table 2: Confusion matrix for Wndchrm (rounded percentages). Overall average accuracy is 69.1%.

	CE	CS	CY	FS	HO	NU
CE	**61**	8	12	7	3	9
CS	4	**74**	13	2	0	7
CY	0	0	**100**	0	0	0
FS	2	2	12	**65**	15	4
HO	0	5	7	31	**49**	8
NU	4	5	4	1	2	**84**

ent studies in the literature [1, 3, 5, 12, 13, 15] with varying cell type taxonomies, imaging conditions and data sets make a direct comparison hard, but our experiments show very promising results of over 90 percent accuracy that is among the best results reported so far. As stated in [3], IIF imaging demands for highly specialized personnel for accurately performed and correctly reported laboratory determinations. IIF readings show inter-expert variability resulting in noisy training data for automated expert systems. ShareBoost's robustness to label noise can be a mitigating factor in the disagreement of cell labeling. As future directions, the ECOC multi-class scheme can be improved by mapping binary codes to actual clusters in the data as opposed to using arbitrary codes, or the ShareBoost algorithm itself can be extended to a multi-class framework. Analysis of the proposed feature descriptors in a multi-scale fashion is being studied. Also, the effects of different base classifiers and view selections using larger data sets need to be validated for computer-aided diagnosis.

References

[1] E. Cordelli and P. Soda. Color to grayscale staining pattern representation in iif. In *Computer-Based Medical Systems (CBMS), 2011 24th International Symposium on*, pages 1–6, june 2011.

[2] I. Ersoy, F. Bunyak, V. Chagin, C. Cardoso, and K. Palaniappan. Segmentation and classification of cell cycle phases in fluorescence imaging. *Lecture Notes in Computer Science (MICCAI)*, 5762:617–624, 2009.

[3] P. Foggia, G. Percannella, P. Soda, and M. Vento. Early experiences in mitotic cells recognition on hep-2 slides. In *IEEE Symp Computer Based Medical Systems*, pages 38–43, Oct. 2010.

[4] A. Hafiane, G. Seetharaman, K. Palaniappan, and B. Zavidovique. Rotationally invariant hashing of median binary patterns for texture classification. In *ICIAR*, pages 619–629, 2008.

[5] T.-Y. Hsieh, Y.-C. Huang, C.-W. Chung, and Y.-L. Huang. Hep-2 cell classification in indirect immunofluorescence image. In *Proc. Int. Conf. Info., Comm. Sig. Processing*, ICICS'09, pages 211–214, 2009.

[6] S. Jaeger, K. Palaniappan, C. Casas-Delucchi, and M. Cardoso. Classification of cell cycle phases in 3d confocal microscopy using PCNA and chromocenter features. In *Proc. of 7th Indian Conf. on Computer Vision, Graphics and Image Proc.*, pages 412–418, 2010.

[7] S. Jaeger, K. Palaniappan, C. Casas-Delucchi, and M. Cardoso. Dual channel colocalization for cell cycle analysis using 3d confocal microscopy. In *IEEE Int. Conf. Pattern Recognition*, 2010.

[8] S. K. Nath and K. Palaniappan. Adaptive robust structure tensors for orientation estimation and image segmentation. In *Lecture Notes in Computer Science (ISVC)*, volume 3804 of *ISVC'05*, pages 445–453, 2005.

[9] T. Ojala, M. Pietikainen, and T. Maenpaa. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Trans. on PAMI*, 24(7):971–987, jul 2002.

[10] K. Palaniappan, F. Bunyak, P. Kumar, I. Ersoy, S. Jaeger, K. Ganguli, A. Haridas, J. Fraser, R. Rao, and G. Seetharaman. Efficient feature extraction and likelihood fusion for vehicle tracking in low frame rate airborne video. In *Proceedings of the 13th Int. Conf. Information Fusion*, Jul. 2010.

[11] J. Peng, C. Barbu, G. Seetharaman, W. Fan, X. Wu, and K. Palaniappan. Shareboost: boosting for multi-view learning with performance guarantees. In *Lecture Notes in Artificial Intelligence (ECML PKDD)*, volume 6912 of *ECML PKDD'11*, pages 597–612, 2011.

[12] P. Perner, H. Perner, and B. Muller. Mining knowledge for hep-2 cell image classification. *Artificial Intelligence in Medicine*, 26(1-2):161–173, 2002.

[13] U. Sack, S. Knoechner, H. Warschkau, U. Pigla, F. Emmrich, and M. Kamprad. Computer-assisted classification of hep-2 immunofluorescence patterns in autoimmune diagnostics. *Autoimmunity Reviews*, 2(5):298–304, 2003.

[14] L. Shamir, N. Orlov, D. M. Eckley, T. Macura, J. Johnston, and I. G. Goldberg. Wndchrm - an open source utility for biological image analysis. *Source code for biology and medicine*, 3, 2008.

[15] P. Soda. Early experiences in the staining pattern classification of hep-2 slides. In *20th Int. Symp. Computer Based Medical Systems*, pages 219–224, june 2007.