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

Feature Analysis for Automatic Classification of HEp-2 Florescence Patterns : Computer-Aided Diagnosis of Auto-Immune Diseases

Subarna Ghosh and Vipin Chaudhary
Department of Computer Science and Engineering,
State University of New York (SUNY) at Buffalo, NY 14260
(sghosh7, vipin)@buffalo.edu

Abstract

Indirect ImmunoFluorescence (IIF) is currently the recommended method for the detection of antinuclear autoantibodies(ANA). It is an effective technique to reveal the presence of auto immune diseases; however, it is a subjective method and hence dependent on the experience and expertise of the physician. Moreover, inter-observer variability limits the reproducibility of IIF reading. To this end, we propose feature extraction methods for automatic recognition of staining patterns of HEp-2 images (provided as a part of the ICPR 2012 HEp-2 Cells Classification Contest) to develop a Computer-Aided Diagnosis system and support the specialists' decision . We compare the performances of various individual and combined features and show that a combination of HOG(Histogram of Oriented Gradients), Texture and ROI(Region of Interest) features are best suited for our task achieving an overall accuracy of 91.13% *using a Support Vector Machine as classifier.*

1. Introduction

A large percentage of the world's population suffers from auto-immune diseases(AD) like diabetes mellitus(type 1), multiple sclerosis and rheumatoid arthritis. The National Institutes of Health (NIH) estimates up to 23.5 million Americans suffer from AD. Currently, Indirect ImmunoFluorescence (IIF) is the recommended method for the detection of antinuclear autoantibodies(ANA) [1]. It is an effective technique to reveal the presence of auto immune diseases, however in standard practice the patterns are interpreted manually which might result in errors due to subjective misinterpretation, physical fatigue of the physician and less experience in the field. Moreover, inter-observer variability limits the reproducibility of IIF reading. Hence Computer-Aided Diagnosis (CAD) systems comprising of automatic recognition of staining patterns of HEp-2 images, can not only support the specialists' decision, it can also reduce variability and help to screen patients reliably.

2. Previous Research

In the past decade, there has been a lot of work in the direction of CAD systems for detection of ANA. Researchers have concentrated on cell stain pattern recognition [9, 10, 7, 5], mitotic cell detection [8], HEp-2 well pattern classification [11] and automation of the entire IIF process [6]. Petra Perner [9] showed the feasibility of an automated inspection system in 2001 using a dataset of 105 samples. Subsequently Perner et al. [10] experimented on 1041 cells and achieved an error rate of 16.9% for a 6-class stain pattern classification using a decision tree. Hsieh at al. [7] evaluated 1036 autoantibody fluorescence patterns from 44 IIF images that were divided into six pattern categories (diffuse, peripheral, coarse speckled, fine speckled, discrete speckled and nucleolar), achieving an accuracy of 80.3%. Elbischger et al. [5], achieved an accuracy of 90.25% using 9 hand-crafted features and a Mahalanobis distance classifier for 17 pattern classes. We however note that features like perimeter ratio and area ratio, take into account the overall mean ROI area and perimeter and hence the training and testing dataset are not totally independent.

Literature survey shows that in previous research, the HEp-2 florescence image dataset and the staining pattern types have been variable. Hence, to make a fair comparison possible, we elaborate on various feature extraction methods for the automatic classification of 6 florescence patterns (Fig. 1) on the dataset provided for the ICPR 2012 HEp-2 Cells Classification Contest.

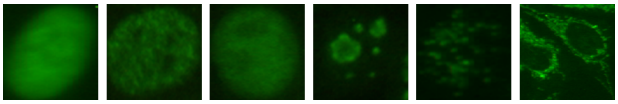
978-4-9906441-1-6

3. Our Approach

3.1. HEp-2 Cell Image Data

The contest HEp-2 images were acquired by means of a fluorescence microscope (40-fold magnification) coupled with a 50W mercury vapor lamp and with a digital camera (SLIM system by Das srl). The images have a resolution of 1388x1038 pixels and a color depth of 24 bits. Specialists manually segmented and annotated each cell at a workstation monitor, and reported data on fluorescence intensity and staining pattern. The dataset is constituted by 28 images, which contain 1457 cells.

Figure 1: This figure illustrates the six types of cell staining patterns: (from left to right) Homogeneous, coarse-speckled, fine-speckled, nucleolar, centromere and cytoplasmatic



Participants received 14 out of the 28 images to be used as the training set which contains 721 HEp-2 cell images. There are six staining patterns (see Fig. 1), the distribution being 208 Centromere, 109 Coarse Speckled, 58 Cytoplasmatic, 94 Fine Speckled, 150 Homogeneous and 102 Nucleolar patterns. This dataset contains 325 intermediate and 396 positive intensity images. Since the test set will be released after the conference, we conduct 10-fold cross-validation experiments on the training set and report our results.

3.2. Feature Extraction

In this section, we explore the performance of various features like SURF (Speeded-Up Robust Features) in a bag of words (BoW) model, texture- and ROI (Region of Interest)-based features and Histogram of Gradient (HOG) features. We convert the given RGB images into grayscale by eliminating the hue and saturation information while retaining the luminance. This is our only pre-processing step. In the subsequent subsections we discuss each feature extraction method in detail and finally present our classification results in Section 4.

1) ***Speeded-Up Robust Features (SURF)*** : SURF is a robust image detector and descriptor [2] that has been widely used in computer vision tasks like object recognition or 3D reconstruction. We use the openCV [3] implementation of SURF to detect key-points (setting Hessian threshhold of 1) and extract descriptors from each image. We create a vocabulary of features by clustering the training descriptors into 1000 bins and then create a histogram of responses for each test image to words in the vocabulary, resulting in a feature vector of length 1000. We use the Bag of Words (BoW) classes provided in openCV2.x to extract these features.

2) ROI-based Features: The dataset contains a binary mask for each cell, which corresponds to the Region Of Interest(ROI). We extract region based features namely area, eccentricity, major and minor axis length and perimeter of the region. We also extract the standard deviation of the gray values in the ROI, the 30th (P_{30}) and 60th (P_{60}) percentiles of the gray values in the ROI (the P-th percentile of N ordered values arranged in ascending order is the $n-$th value where $n = round(\frac{P}{100} * N + \frac{1}{2})$), the percentile range ($P_{range} = P_{60} - P_{30}$), and the roundness of the ROI ($r = 4 * \pi * area/perimeter$). Some of these features are inspired by [5]. We scale all the features between 0 and 1, making sure to use the training set scaling factor for the corresponding test set in each of the 10 cross-validation folds, thus ensuring that the training phase is completely unbiased by the test set.

3) Texture-based Features: We extract four well established texture features : contrast, homogeneity, correlation and energy from the GLCM (Gray-Level Co-occurrence Matrix) matrices constructed in 8 directions, using 32 bins/levels to scale the images. Mathematically, a GLCM matrix G is defined over an n x m image I, parameterized by an offset $(\Delta x, \Delta y)$ as :

$$G_{\Delta x \Delta y}(i,j) = \sum_{p=1}^{n}\sum_{q=1}^{m} \begin{cases} 1, & \text{if } I(p,q) = i \text{ and} \\ & I(p+\Delta x, q+\Delta y) = j \\ 0, & otherwise \end{cases} \quad (1)$$

We also extract the image entropy, the maximum and minimum values of local entropy, along with the maximum and minimum values of local standard deviation of the gray-levels. We scale all the features between 0 and 1, similar to that of the ROI-based features.

4) Normalized HOG Features: Histogram of Oriented Gradients (HOG) are feature descriptors used extensively in computer vision and image processing. This technique counts occurrences of gradient orientation in localized portions (windows) of an image. In our approach, given a masked cell image, we divide it into 6x6 = 36 sub-windows and fix the bin size to 36, resulting in a vector of length 1296.

4. Experimental Results and Discussion

To test the performance of the features extracted, we model multi-class SVM (Support Vector Machine)

Table 1: Confusion matrices for combined features using SVM as classifier (rows represent ground truth)

(a) SURF + Texture features

	HO	FS	CS	NU	CY	CE
HO	124	17	3	1	0	5
FS	27	58	5	0	1	3
CS	1	1	96	1	5	5
NU	0	2	1	86	3	10
CY	0	1	0	2	53	2
CE	3	7	8	7	4	179

(b) HOG+Texture features

	HO	FS	CS	NU	CY	CE
HO	135	10	3	0	1	1
FS	21	70	0	1	0	2
CS	0	1	106	0	2	0
NU	0	0	3	87	0	12
CY	0	0	7	0	50	1
CE	1	2	0	7	0	198

(c) Texture+ROI features

	HO	FS	CS	NU	CY	CE
HO	139	8	3	0	0	0
FS	25	66	0	0	0	3
CS	0	1	106	0	0	2
NU	1	0	1	90	0	10
CY	0	1	3	2	52	0
CE	2	2	4	3	0	197

(d) HOG+Texture+ROI features

	HO	FS	CS	NU	CY	CE
HO	136	11	2	0	1	0
FS	20	70	0	2	0	2
CS	0	1	108	0	0	0
NU	0	0	1	90	0	11
CY	0	0	6	0	52	0
CE	0	2	1	4	0	201

classifiers [4] for each of the 10-folds. Throughout our experiments, the kernel used is linear and the six classes are Homogeneous(HO), Fine Speckled(FS), Coarse Speckled(CS), Nucleolar(NU), Cytoplasmatic(CY) and Centromere(CE).

The confusion matrices for the combined features are shown in Table 1 and the detailed accuracy results of our 10-fold cross-validation experiments are tabulated in Table 2. We observe that the coarse-speckled pattern is the simplest to distinguish while the fine speckled is the hardest and is very often wrongly classified as homogeneous. Upon close inspection, we see that the fine-speckled patterns are indeed very similar to the homogeneous patterns. Moreover, we can also see that positive intensity images show higher accuracies than the intermediate ones, with HOG features being the most robust to intensity variation. Hence, combining the HOG, Texture and ROI features we obtain a robust pattern recognition system, that shows a high overall accuracy of 91.13%, where the coarse speckled pattern shows the highest accuracy of 99.08% and the fine-speckled pattern shows the lowest accuracy of 74.47%.

To complete the assessment of our features, we input our combined feature-set to various classifiers and illustrate the results in Table 3. The first two classifiers are 3-layer Artificial Neural Networks. ANN_600_12 has 600 and 12 nodes in the first and second hidden layers respectively, while ANN_1000_20 has 1000 and 20. The third is an SVM, which is the same as explained earlier. The next four are k-Nearest Neighbor classifers with k=3 or k=4, and the distance measure as cosine or Euclidean. The final classifier is a simple Naive Bayes. The high accuracies shown by all the classifiers reiterates the utility of our combined feature set for the task of cell staining pattern recognition.

5. Conclusion

We have proposed a composite feature set for the automatic classification of HEp-2 florescence patterns. Our features are a concatenation of HOG(Histogram of Oriented Gradients), Texture- and ROI(Region of Interest)-based features, effectively achieving an overall accuracy of 91.13% using a Support Vector Machine as classifier. In this paper we not only elaborately discuss the feature extraction process, we also experiment on a wide range of classifiers to prove the utility of these features. Moreover, the florescence dataset we have used, is the one provided for an ICPR 2012 contest, making fair comparison between methods possible.

References

[1] Center for disease control: Quality assurance for the indirect immunofluorescence test for autoantibodies to nuclear antigen (if-ana): Approved guideline. *NCCLS I/LA2-A*, 16(11), 1996.

[2] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool. Speeded-up robust features (surf). *Comput. Vis. Image Underst.*, 110, June 2008.

[3] G. Bradski. The OpenCV Library. *Dr. Dobb's Journal of Software Tools*, 2000.

Table 2: This table compares the 10-fold cross-validation results of various features for HEp-2 cell staining pattern classification. It shows the intensity-wise, class-wise and overall accuracy (in percentage) for individual and combined features, using SVM as classifier.

Feature Type	Percent Accuracy								
	Intensity-wise		Class-wise						Overall
	Positive	Inter	HO	FS	CS	NU	CY	CE	
HOG	83.84	80.31	82.67	64.89	93.58	70.59	82.76	89.42	82.25
Texture	90.40	81.54	87.33	68.09	96.33	86.27	68.97	93.75	86.39
ROI	80.56	50.78	87.33	0	88.99	58.82	86.21	70.19	67.13
SURF	77.53	57.85	68.67	44.68	59.63	76.47	84.48	75.96	68.67
SURF+Texture	87.88	76.31	82.67	61.70	88.07	84.31	91.38	86.06	82.69
HOG+Texture	91.41	87.38	90	74.47	97.25	85.30	86.21	95.19	89.60
Texture+ROI	93.43	86.15	92.67	70.21	97.25	88.24	89.66	94.71	90.15
HOG+Texture+ROI	92.17	89.85	90.67	74.47	99.08	88.24	89.66	96.63	91.13

Table 3: This table compares the 10-fold cross-validation results of various classifiers for HEp-2 cell staining pattern classification. It shows the intensity-wise, class-wise and overall accuracy (in percentage) for 7 classifiers: two ANNs(Artificial Neural Networks), one SVM(Support Vector machine),four k-NNs(k - Nearest Neighbor) and one Naive-Bayes.

Classifer Type	Percent Accuracy								
	Intensity-wise		Class-wise						Overall
	Positive	Inter	HO	FS	CS	NU	CY	CE	
ANN_600_12	89.14	88.62	91.33	73.40	92.66	86.27	86.21	94.23	88.90
ANN_1000_20	92.42	88.31	91.33	75.53	93.58	94.12	86.21	94.71	90.57
SVM	92.17	89.85	90.67	74.47	99.08	88.24	89.66	96.63	91.13
3-NN (Euclidean)	83.84	81.85	84.67	56.38	84.40	75.49	87.93	95.19	82.94
5-NN (Euclidean)	79.80	82.46	82.0	57.45	82.57	65.69	84.48	96.63	81.0
3-NN (Cosine)	92.42	75.69	92.67	52.13	94.50	89.22	96.55	83.65	84.88
5-NN (Cosine)	93.18	75.08	93.33	53.19	96.33	88.24	93.10	83.65	85.02
Naive-Bayes	77.78	77.23	81.33	53.19	77.98	84.31	79.31	81.73	77.53

[4] C.-C. Chang and C.-J. Lin. *LIBSVM: a library for support vector machines*, 2001.

[5] P. Elbischger, S. Geerts, K. Sander, G. Ziervogel-Lukas, and P. Sinah. Algorithmic framework for hep-2 fluorescence pattern classification to aid auto-immune diseases diagnosis. In *Biomedical Imaging: From Nano to Macro, 2009. ISBI '09. IEEE International Symposium on Biomedical Imaging*, 2009.

[6] R. Hiemann, N. Hilger, J. Michel, J. Nitschke, A. Bohm, U. Anderer, M. Weigert, and U. Sack. Automatic analysis of immunofluorescence patterns of hep-2 cells. *Annals of the New York Academy of Sciences*, 1109, 2007.

[7] T.-Y. Hsieh, Y.-C. Huang, C.-W. Chung, and Y.-L. Huang. Hep-2 cell classification in indirect immunofluorescence images. In *Information, Communications and Signal Processing, 2009. ICICS 2009. 7th International Conference on*, 2009.

[8] G. Percannella, P. Soda, and M. Vento. Mitotic hep-2 cells recognition under class skew. In *ICIAP (2)*, pages 353–362, 2011.

[9] P. Perner. Classification of hep-2 cells using fluorescent image analysis and data mining. In *Proceedings of the Second International Symposium on Medical Data Analysis*, ISMDA '01, 2001.

[10] P. Perner, H. Perner, and B. Müller. Mining knowledge for hep-2 cell image classification. *Artificial Intelligence in Medicine*, 26(1-2), 2002.

[11] P. Soda and G. Iannello. Aggregation of classifiers for staining pattern recognition in antinuclear autoantibodies analysis. *IEEE Transactions on Information Technology in Biomedicine*, 13(3), May 2009.