HEp-2 Cell Image Classification: A Comparative Analysis

Praful Agrawal, Mayank Vatsa, Richa Singh

Indraprastha Institute of Information Technology, Delhi, India

Abstract. HEp-2 cell image classification is an important and relatively unexplored area of research. This paper presents an experimental analysis of five different categories of feature sets with four different classifiers to determine the best performing combination of features and classifiers. The analysis is performed on the ICIP 2013 Cell Classification Contest Training dataset comprising over 13,000 cell images pertaining to six cell classes. The results computed with 10 fold cross validation show that texture features perform the best among all the explored feature sets and the combination of Laws features with SVM yields the highest accuracy.

1 Introduction

Human immune system creates antibodies to fight against infections whereas antinuclear antibodies affect healthy tissues (cell nucleus). Antinuclear Autoantibodies (ANA) test is widely used to determine whether the immune system is developing antibodies or not [1]. Indirect Immunofluorescence (IIF) based ANA test is state-of-the-art due to its high specificity and ability to discriminate the samples belonging to positive, intermediate, and negative classes. In this test, ANAs are detected by a specific pattern among 30 different fluorescence patterns which can be recognized via HEp-2 cells [2]. The process involves manually identifying fluorescence patterns which require visual inspection of slides under fluorescence microscope by highly qualified physicians. Manual evaluation suffers from some limitations such as inter-observer variability and scarcity of highly specialized personnel. The need for automation has been well accepted by the researchers and therefore, recent attempts are made to share HEp-2 cell image classification databases and develop automated algorithms.

HEp-2 cell image classification can be considered as a classical pattern recognition problem where cell patterns can be modeled using an efficient representation and classified using multi-class classification algorithms. Recent research has focused on image based features and standard classification algorithms. Table 1 summarizes some recent techniques for HEp-2 cell image classification. Majority of existing approaches have evaluated their performance using the ICPR 2012 Contest Training Dataset[1] containing approximately 1,400 cell images with positive and intermediate fluorescence intensity images.

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

Table 1. Some recent techniques on HEp-2 cell image classification.

	Features	**Classifier**
Wiliem *et al.* [3], 2013	Dual-region codebook based descriptor	Nearest Convex Hull
Ersoy *et al.* [4], 2012	Shape, gradient, and texture features	Multiview Shareboost
Ghosh & Chaudhary [5], 2012	HOG, ROI, SURF, and texture features	Multiclass SVM
Li *et al.* [6], 2012	LBP, Gabor, DCT, and statistical features	Multiclass Boosting SVM
Iannello *et al.* [7], 2012	SIFT features	Bag Of Visual Words
Ali *et al.* [8], 2012	Contrast based features	kNN
Theodorakopoulos *et al.* [9], 2012	Morphological and LBP features	SVM
Cordelli and Soda [10], 2011	Texture features	MLP, kNN, SVM, and Adaboost
Foggia *et. al.* [11], 2010	Morphological, texture and rectangle features	MLP, Naïve Bayes, kNN, SVM, and Adaboost
Soda & Iannello [12], 2009	Different features specific to each class	Multi-Expert System

The main focus of the paper is to analyze different features and classification approaches for HEp-2 cell image classification. In this research, we broadly categorize the features (corresponding to cell structures and appearances) used in literature under five different categories - Boundary, Descriptor, Shape and Size, Statistical, and Texture Features. Using this categorization, a comparative analysis is performed to understand their underlying discriminative ability across different types of cell patterns. The ability of these feature categories to distinguish among various types of cell patterns is evaluated using four different classifiers[2]. Unlike existing literature where researchers have reported overall classification accuracy, this research reports the performance on *positive* and *intermediate* intensity cell images to compare the complexity of the classification task within each intensity class. For each feature-classifier combination, three classifier models are trained - two pertaining to each intensity class and one for the combined set. The experiments are performed using the ICIP 2013 HEp-2 Cell Image Classification Contest training dataset [13] comprising more than $13,000$ cell images. Experimental results suggest that the combination of Laws features and SVM classifier yields a very high classification accuracy.

2 Features and Classifiers for Analysis

A Computer Aided Diagnosis (CAD) system designed for the HEp-2 cell classification from Indirect Immunofluorescence images is based on a well defined

[2] To the best of our knowledge, such a categorization and comparative analysis of different types of feature sets is not available in the HEp-2 cell image classification literature.

procedural model illustrated in Fig. 1. The inputs to a CAD system for HEp-2 classification are IIF images and the corresponding fluorescence intensity class label. Individual HEp-2 cell images are segmented from the IIF image and further categorized into different cell patterns using pattern recognition and machine learning algorithms. The database used and the features and classifiers compared and evaluated in this study are listed in the following subsections.

Fig. 1. Framework of a CAD system for HEp-2 cell classification.

2.1 HEp-2 Cell Database

ICIP 2013 HEp-2 cell image classification contest training dataset [13] contains 13,596 cell images pertaining to six cell patterns namely *Centromere, Golgi, Homogeneous, Nucleolar, Nuclear Membrane*, and *Speckled*. Cell images are segmented from IIF images pertaining to 83 subjects. For every cell image, there is a mask image of the same size describing the cell boundary in the corresponding cell image. In addition to cell pattern type, the intensity class is also provided for each cell image. In practice, intensity class can either be *Negative, Positive* or *Intermediate*, however the dataset contains images belonging to only *Positive* and *Intermediate* classes. Images from the intermediate intensity class are generally lower in contrast compared to the positive class. The dataset also includes bounding box information for each cell image. Sample images from the dataset are illustrated in Fig. 2 and Table 2 provides the quantitative summary of the dataset.

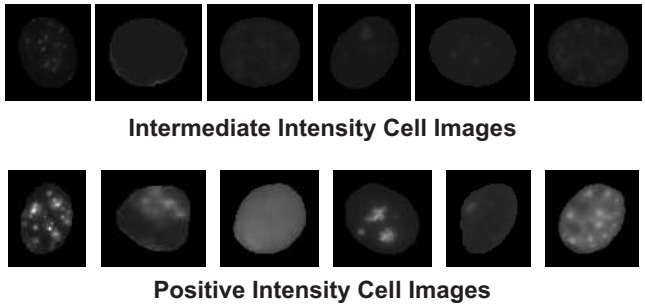

Fig. 2. Sample images from the ICIP 2013 cell image classification contest training dataset. The cell pattern types of images from left to right - Centromere, Golgi, Homogeneous, Nucleolar, Nuclear Membrane, and Speckled.

2.2 Features and Classifiers

Cell images are first segmented using the given mask images by assigning background intensity as zero in the resulting image. The features widely used in

Table 2. Summary of the ICIP 2013 cell classification contest training dataset.

Classes	Centromere	Golgi	Homogeneous	Nucleolar	Nuclear Membrane	Speckled
Intermediate	1363	375	1407	1664	1265	1374
Positive	1378	349	1087	934	943	1457
Total	**2741**	**724**	**2494**	**2598**	**2208**	**2831**

literature are categorized into five general categories of image and object based features: boundary, descriptor, shape and size, statistical, and texture features. Such a categorization represents distinctive characteristics of cell images and also helps in discriminating among different cell patterns. The features extracted and their categorization are presented in Table 3.

Table 3. Features extracted from cell images.

Category (length)	Features
Boundary (38)	Perimeter, mean Sobel gradient of boundary pixels, bending energy and 35 spectral energy features derived from FFT of radii vector [14].
HoDesc (221)	LBP and HOG histograms of length 59 and 162 respectively.
Shape and Size (15)	Area, major and minor axis length, eccentricity, orientation, convex area, convex deficiency, solidity, extent, aspect ratio, equivalent diameter, sphericity, compactness, inertiashape, and deviation in centre of mass [14].
Statistical (4)	Mean, standard deviation, skewness, and kurtosis of pixel values.
Texture (211)	65 SGLD [15], 44 GLRL [16], and 102 Laws features [17].

The performance of different feature categories are evaluated with four different classifiers for six class classification. The classifiers used in this research are listed below:

- **Naïve Bayes:** Probability distributions are estimated from the training dataset and Bayes decision rule is applied on the test data for classification.
- k **Nearest Neighbor:** The first nearest neighbors of probe instances are derived from the training data using Euclidean distance as the distance measure.
- **Support Vector Machine:** SVM with linear and non-linear (RBF) kernels is used for classification. The optimal values of parameters such as gamma and cost are estimated in a grid search manner using LIBSVM library [18].
- **Random Decision Forest:** The parameters such as number of trees and forest depth are estimated in a grid search manner.

3 Results and Analysis

Using the ICIP 2013 HEp-2 Cell Image Classification Contest training dataset [13], three subsets of the HEp-2 cell image dataset are considered for experiments: Positive (P), Intermediate (I), and Combined (C). The subsets are formed based

on the fluorescence intensity information available with the dataset. The first two subsets correspond to images from the positive and intermediate intensity class and the third set is the entire dataset containing images with both types of fluorescence intensity. The performance of five different feature categories and four different classifiers are evaluated with 10 fold cross validation on these three data subsets. Mean and standard deviation of accuracy values obtained in 10 folds of cross validation are reported in Table 4 and the key observations are explained below.

Table 4. Comparative results for five feature categories with four different classifiers on three datasets.

Classifier		Boundary	Statistical	Shape_Size	Texture	HoDesc
	P	29.65 ± 1.96	86.04 ± 0.89	34.35 ± 1.54	94.89 ± 1.16	70.59 ± 0.78
k**NN**	**I**	26.83 ± 1.45	76.70 ± 1.36	27.90 ± 1.96	69.12 ± 1.75	50.03 ± 1.38
	C	25.93 ± 1.13	76.29 ± 1.14	28.94 ± 1.35	77.33 ± 1.46	55.08 ± 1.37
	P	26.56 ± 1.88	70.05 ± 2.20	38.27 ± 1.76	56.77 ± 1.48	77.62 ± 2.23
Naïve Bayes	**I**	23.17 ± 1.20	49.76 ± 1.45	26.24 ± 0.81	33.23 ± 1.91	48.32 ± 1.22
	C	25.47 ± 0.98	54.24 ± 0.97	28.84 ± 1.12	37.24 ± 0.77	44.76 ± 0.98
	P	46.80 ± 1.81	91.38 ± 1.20	51.72 ± 1.99	96.34 ± 0.73	86.56 ± 0.66
RDF	**I**	45.54 ± 1.47	81.36 ± 1.04	48.28 ± 1.51	86.83 ± 1.48	75.23 ± 1.76
	C	39.75 ± 1.68	83.16 ± 1.11	45.27 ± 1.07	89.02 ± 0.64	77.58 ± 0.99
	P	32.73 ± 4.43	71.31 ± 1.55	33.60 ± 3.37	96.39 ± 0.79	85.95 ± 1.52
SVM-Linear	**I**	25.44 ± 2.39	41.61 ± 2.07	20.92 ± 2.01	87.96 ± 0.87	76.62 ± 0.80
	C	23.84 ± 1.39	45.56 ± 4.20	20.21 ± 3.70	87.84 ± 1.05	73.89 ± 0.64
	P	40.13 ± 1.47	93.48 ± 1.68	46.60 ± 1.59	$\mathbf{98.06 \pm 0.46}$	88.26 ± 1.23
SVM-RBF	**I**	35.11 ± 1.27	75.47 ± 1.38	34.05 ± 1.40	$\mathbf{90.72 \pm 1.35}$	77.46 ± 1.49
	C	30.17 ± 1.13	79.56 ± 1.22	35.69 ± 0.99	$\mathbf{92.85 \pm 0.63}$	79.40 ± 0.69

- The intermediate intensity images are generally lower in contrast as compared to positive intensity images, The experiments performed on the three datasets (positive, intermediate, and combined) clearly validate our assertion that the experiments should be conducted independently on images of two intensity, positive and intermediate. For instance, the classification accuracy of the best performing combination of texture features and SVM-RBF classifier differs by a significant 8% on positive and intermediate subsets when the classifier is trained independently on both the subsets. Overall results from Table 4 indicate that cell image classification for images pertaining to the intermediate intensity set is difficult as compared to the positive intensity set.
- The results further indicate that cells across different types of cell patterns have similar shape and size thereby increasing the inter-class similarity. On the other hand, grayscale distribution within different cell classes vary significantly thus increasing the inter-class variability. Therefore, the features based on grayscale values such as statistical moments of pixel values, texture features, and histograms of HOG and LBP representations yield better performance for cell image classification than boundary, shape, and size features.

– Texture features outperform all other feature categories by at least 5% for the positive intensity set and 13% for the intermediate intensity set. The second best results are provided by Statistical and HoDesc features using SVM-RBF classifier for positive and intermediate sets respectively. As discussed earlier, though the images pertaining to intermediate intensity set are difficult to categorize, texture features yield significant improvement in classifying these images.
– Among the four classification techniques used for experiments, SVM with RBF kernel yields the highest accuracy for HEp-2 cell image classification. The results of RDF and SVM-RBF are comparable with RDF providing the second best classification accuracy.
– To analyze the performance according to cell classes, the classification accuracy of texture features among different types of cell patterns is analyzed in Fig. 3. It can be observed that the classification performance for positive intensity images is higher than intermediate intensity images among all six cell patterns. Fig. 3 also shows that the cells of Golgi type are the most difficult to identify.

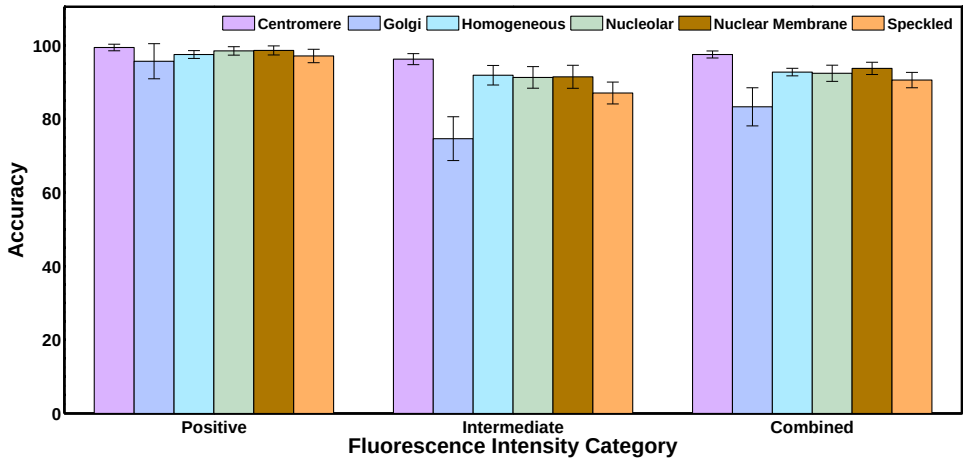

Fig. 3. Accuracies for different cell classes using texture features in combination with SVM classifier and RBF kernel.

The above analysis illustrates the impact of different kinds of features for HEp-2 cell image classification. It can be observed from Table 3 that the size of texture and HoDesc features is large. Therefore, the constituent elements of these two features are further analyzed to determine whether any of the elements alone is sufficient for classification. In this research, among several available texture representations such as SGLD (Spatial Gray Level Dependence), GLRL (Gray Level Run Length), wavelet, and Laws features, SGLD, GLRL, and Laws features are used to derive the texture properties of cell images. Among the HoDesc features, HoG and LBP features are analyzed. As shown in Table 5, in case of descriptor based features, HOG and LBP features show similar performance. A

feature level concatenation of HOG and LBP features helps in further enhancing the classification performance for both positive as well as intermediate intensity classes. On the other hand, comparative results suggest that the accuracy of texture features is primarily attributed to Laws features which not only outperform all other texture features, their performance is significantly higher than any other category of features. The overall analysis shows that the combination of Laws texture features and SVM with RBF kernel yield the best performance to identify the cell patterns in HEp-2 cell images. The combination achieves a classification accuracy of 97.90% and 90.49% for positive and intermediate intensity classes respectively.

Table 5. Comparative analysis of different feature sets used among texture and descriptor based feature categories.

Classifier		HoDesc		Texture		
		HOG	**LBP**	**GLRL**	**Laws**	**SGLD**
k**NN**	**P**	70.48 ± 0.82	71.70 ± 2.10	75.16 ± 1.75	93.82 ± 0.67	82.81 ± 1.30
	I	49.38 ± 1.41	56.74 ± 2.02	44.72 ± 1.62	68.21 ± 2.72	45.03 ± 1.32
	C	54.75 ± 1.28	59.79 ± 1.44	54.05 ± 1.58	76.20 ± 0.74	58.02 ± 0.93
Naïve Bayes	**P**	73.88 ± 1.78	70.15 ± 1.32	50.23 ± 1.54	58.49 ± 1.66	45.67 ± 1.51
	I	47.70 ± 1.57	42.25 ± 1.48	32.88 ± 1.22	34.81 ± 2.99	20.13 ± 1.09
	C	45.60 ± 0.91	38.27 ± 1.09	34.61 ± 0.67	37.83 ± 0.52	28.80 ± 0.72
RDF	**P**	83.21 ± 1.13	80.22 ± 1.35	82.27 ± 1.87	96.23 ± 0.53	85.33 ± 0.97
	I	69.47 ± 1.13	67.80 ± 1.71	52.32 ± 1.59	85.77 ± 1.35	54.54 ± 2.45
	C	72.36 ± 1.27	70.51 ± 1.00	61.42 ± 1.26	88.81 ± 0.57	63.53 ± 1.77
SVM-Linear	**P**	81.39 ± 1.51	79.15 ± 1.40	77.23 ± 1.57	94.68 ± 0.75	84.37 ± 1.33
	I	67.83 ± 1.01	62.96 ± 1.52	39.33 ± 1.18	81.06 ± 1.12	49.75 ± 2.98
	C	66.84 ± 1.28	61.24 ± 1.13	47.34 ± 0.81	83.29 ± 1.10	59.66 ± 0.98
SVM-RBF	**P**	86.30 ± 1.22	84.21 ± 1.47	84.91 ± 1.20	$\mathbf{97.90 \pm 0.71}$	90.81 ± 1.00
	I	73.19 ± 1.50	71.86 ± 1.51	49.96 ± 1.95	$\mathbf{90.49 \pm 1.30}$	58.45 ± 1.28
	C	76.23 ± 1.24	74.39 ± 1.37	61.96 ± 1.24	$\mathbf{92.42 \pm 0.59}$	69.23 ± 1.65

4 Conclusion and Future Work

In this research, the features for HEp-2 cell image classification are classified among five broad categories: boundary, shape and size, statistical, texture, and descriptor. This categorization helps in understanding the discriminating properties of HEp-2 cells among six pattern classes, namely, Centromere, Golgi, Homogeneous, Nucleolar, Nuclear Membrane, and Speckled. The performance of feature categories are evaluated using four different classifiers and 10 fold cross validation. The results on the ICIP 2013 HEp-2 Cell Image Classification Contest training dataset show that the texture features yield the best classification performance. The results further demonstrate that within texture features, Laws features alone are sufficient for classification. The comparative results also validate our assertion that the classification results of positive and intermediate intensity cell images should be reported independently. Currently, we are exploring (1) other feature extraction algorithms such as Gabor filters and wavelet

transforms and (2) fusion and feature selection paradigms to further enhance the classification accuracy.

References

1. Wiik, A.S., Høier-Madsen, M., Forslid, J., Charles, P., Meyrowitsch, J.: Antinuclear antibodies: a contemporary nomenclature using HEp-2 cells. Journal of Autoimmunity **35**(3) (2010) 276–290
2. NCCLS: Center for disease control - quality assurance for the indirect immunofluorescence test for autoantibodies to nuclear antigen (IF-ANA): Approved guideline. LA2-A **16**(11) (1996)
3. Wiliem, A., Wong, Y., Sanderson, C., Hobson, P., Chen, S., Lovell, B.: Classification of human epithelial type 2 cell indirect immunofluoresence images via codebook based descriptors. In: WACV. (2013) 95–102
4. Ersoy, I., Bunyak, F., Peng, J., Palaniappan, K.: HEp-2 cell classification in IIF images using shareboost. In: ICPR. (2012) 3362–3365
5. Ghosh, S., Chaudhary, V.: Feature analysis for automatic classification of HEp-2 florescence patterns: Computer-aided diagnosis of auto-immune diseases. In: ICPR. (2012) 174–177
6. Li, K., Yin, J., Lu, Z., Kong, X., Zhang, R., Liu, W.: Multiclass boosting SVM using different texture features in HEp-2 cell staining pattern classification. In: ICPR. (2012) 170–173
7. Iannello, G., Onofri, L., Soda, P.: A bag of visual words approach for centromere and cytoplasmic staining pattern classification on HEp-2 images. In: CBMS. (2012) 1–6
8. Ali, W., Piro, P., Giampaglia, D., Pourcher, T., Barlaud, M.: Biological cells classification using bio-inspired descriptor in a boosting k-NN framework. In: CBMS. (2012) 1–6
9. Theodorakopoulos, I., Kastaniotis, D., Economou, G., Fotopoulos, S.: HEp-2 cells classification via fusion of morphological and textural features. In: BIBE. (2012) 689–694
10. Cordelli, E., Soda, P.: Color to grayscale staining pattern representation in IIF. In: CBMS. (2011) 1–6
11. Foggia, P., Percannella, G., Soda, P., Vento, M.: Early experiences in mitotic cells recognition on HEp-2 slides. In: CBMS. (2010) 38–43
12. Soda, P., Iannello, G.: Aggregation of classifiers for staining pattern recognition in antinuclear autoantibodies analysis. IEEE TITB **13**(3) (2009) 322–329
13. Hobson, P., Percannella, G., Vento, M., Wiliem, A.: Competition on cells classification by fluorescent image analysis. In: ICIP. (2013) `http://nerone.diiie.unisa.it/contest-icip-2013/index.shtml`.
14. Boucheron, L.E.: Object- and Spatial-Level Quantitative Analysis of Multispectral Histopathology Images for Detection and Characterization of Cancer. PhD thesis, UCSB (2008)
15. Haralick, R.M., Shanmugam, K., Dinstein, I.: Textural features for image classification. IEEE T-SMC (6) (1973) 610–621
16. Tang, X.: Texture information in run-length matrices. IEEE TIP **7**(11) (1998) 1602–1609
17. Laws, K.I.: Textured image segmentation. Technical report, USC (1980)
18. Chang, C.C., Lin, C.J.: LIBSVM: A library for support vector machines. ACM T-IST **2**(3) (2011) 1–27