

Revisiting HEp-2 Cell Image Classification

ISHAN NIGAM, SHREYASI AGRAWAL, RICHA SINGH, (Senior Member, IEEE),
AND MAYANK VATSA, (Senior Member, IEEE)

Indraprastha Institute of Information Technology Delhi, New Delhi 110020, India

Corresponding author: M. Vatsa (mayank@iiitd.ac.in)

ABSTRACT The immune system in homo sapiens protects the body against diseases by identifying and attacking foreign pathogens. However, when the system misidentifies native cells as threats, it results in an auto-immune response. The auto-antibodies generated during this phenomenon may be identified through the indirect immunofluorescence test. An important constituent process of this test is the automated identification of antigen patterns in the cell images, which is the focus of this research. We perform a detailed literature review and present a framework to automate the identification of antigen patterns. The efficacy of the framework, demonstrated on the MIVIA ICPR 2012 HEp-2 Cell Contest and SNP HEp-2 Cell datasets, suggests that the algorithm is comparable with the state-of-the-art approaches.

INDEX TERMS Biomedical imaging, anti-nuclear antibody testing, indirect immunofluorescence test, HEp-2 cells, laws texture measure.

I. INTRODUCTION

The circulatory system in the human body transports micro-particles to facilitate a wide spectrum of functions. The immune system, a component of the circulatory system, defends the host by detecting foreign pathogens and attacking the invasions. Immunity in humans functions through two pathways. The body's *inherent* self-defense mechanism comprises of native micro-organisms, which counter pathogens without the presence of any external aid. On the other hand, humans also *acquire* the ability to defend against pathogens as the body learns to counter infections and develops antibodies against the pathogens. This acquired form of immunity, an imperfect process, might occasionally learn to incorrectly identify the body's native cells as pathogens and generate antibodies to defend against these perceived threats. Such agents are termed *auto-antibodies* and the conditions are identified as *auto-immune diseases*.

The processes which result in the generation of auto-antibodies are not completely understood. Rioux and Abbas [1] observe that individuals are tolerant of antigenic substances native to their systems and failure of such self-tolerance is likely to be a reason for autoimmunity. However, this is not the only cause for auto-immune diseases. For instance, the immune system responsible for counter-acting foreign infections may fail to act swiftly and entry of pathogens may indirectly affect the system; individuals with a genetic predisposition towards developing autoimmunity may develop a disease due to exposure to the pathogen. Auto-immune diseases, partly due to the difficulty in early

identification, have high mortality rates. The scale at which auto-immune diseases are prevalent in the general population is an important factor in understanding the significance of auto-immune diseases as a threat to the population. An extensive analysis of the National Health and Nutrition Examination Survey conducted by Satoh et al. [2] studied the prevalence of auto-immune diseases in the United States of America. The presence of auto-antibodies in the selected population was found to be more than 13% and the rate of auto-antibody affliction was found to rise with age. The study concluded that auto-antibodies are present in more than 32 million individuals in the United States of America. Fairweather et al. [3] observe that autoimmunity is more prevalent among women. It is suggested that this higher prevalence may be associated with a higher antibody production rate and a (Th)2-predominant immune response among the female population.

Antinuclear Antibodies (ANA) are auto-antibodies that affect the cell nucleus. The presence of ANA has been observed in conjunction with many autoimmune diseases. Therefore, the Antinuclear Antibody Test is conducted as a screening test for auto-immune diseases. Castro and Gourley [4] note that the associated laboratory investigation consists of a complete blood count which includes metabolic inflammatory markers, auto-antibodies, and flow cytometry. The common tests used for detecting and quantifying ANAs are indirect immunofluorescence (IIF) and enzyme-linked immunosorbent assay (ELISA) tests. Even though both methods achieve high sensitivity, the IIF

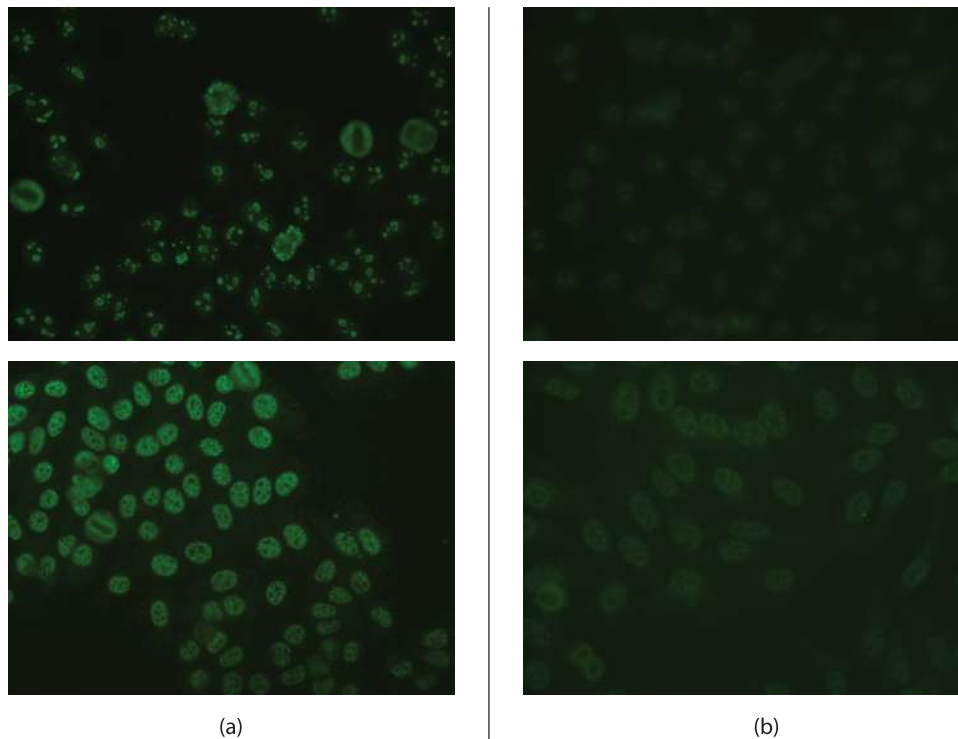


FIGURE 1. HEp-2 cell images of varying fluorescence intensity levels: (a) Positive, and (b) Intermediate. First row contains nucleolar antigen patterns, second row contains coarse speckled antigen patterns.

test is preferred and recommended as the ELISA test is limited to detection of a few specific patterns [5]. The IIF test achieves high sensitivity as well as high specificity for ANA detection and can be used to detect many antinuclear antigen patterns [6]. The IIF test is time consuming and requires human intervention. A medical expert categorizes the digital images as positive, intermediate, or negative based on the fluorescence intensity levels of the images. The images classified as negative are discarded, and positive and intermediate IIF image slides are used further in the diagnosis process. Figure 1 shows sample cell images belonging to positive and intermediate categories. The positive and intermediate images are further examined by experts to identify patterns being exhibited by the antinuclear antibodies as a result of reactions taking place within type-2 Human Epithelial (HEp-2) cells. The process is a time-consuming task and prone to errors at multiple stages. Thus, it becomes an important use-case for automation of the detection of autoimmune diseases.

As mentioned previously, the Indirect Immunofluorescence test, the most advanced technique for Antinuclear Antibodies, requires significant human intervention. Rigon et al. [7], Rigon et al. [8], and Bonroy et al. [9] report that the research community is actively exploring methods to automate the Indirect Immunofluorescence test. The most exigent aspect of the IIF test is the identification of patterns in HEp-2 cells from the blood sera obtained from patients. The organization of public classification challenges and recent availability of

public HEp-2 cell databases has significantly driven research in solving this problem. The key contributions of our research are as follows:

- We perform a review of existing literature and present comparative results of different approaches on popular public datasets.
- We propose a framework to perform identification of patterns from HEp-2 cells. The proposed framework is evaluated on multiple datasets and classification performance is compared to the current state-of-art in HEp-2 cell identification.

Section II presents a survey of the approaches used for HEp-2 cell classification. Section III presents the proposed methodology for recognizing HEp-2 cell patterns. Section IV presents an experimental evaluation of the proposed framework on the MIVIA ICPR 2012 Contest [10] and the SNP HEp-2 datasets [11]. Finally, Section V presents the conclusions of our research in the field of HEp-2 cell pattern classification.

II. LITERATURE SURVEY

Recognition of HEp-2 cell patterns through Indirect Immunofluorescence testing has become an important area of study for diagnosis of auto-immune diseases. In the past, efforts have been made towards automating the process by understanding how human experts classify HEp-2 cell images [12]. Soda and Iannello [13] present a multiple expert system that employs binary modules constituted by ensembles of classifiers. However, the framework is tested on

a small dataset and engenders the need for public datasets and benchmarking protocols for developing standardized recognition systems. Further, Cordelli and Soda [14] determine the optimal method to convert color images to greyscale images for the development of automated systems. The authors compare color-space conversions and analyze wide feature sets for each conversion method, applying several classification paradigms. Sack et al. [15] make recommendations for uniform processing and interpretation of HEp-2 cell tests for detection of antinuclear antibodies, though the research in HEp-2 cell classification has advanced beyond the recommendations made in the study. In 2013, Agrawal et al. [16] analyze a number of features and present comparative results of features and classifiers on several datasets. Further, Foggia et al. [17] provide a comprehensive summary of recognition approaches participating in the ICPR 2012 HEp-2 cell classification contest. The remainder of this section summarizes recent approaches to HEp-2 cell recognition.

Iannello *et al.* [18] extend the panel of detectable HEp-2 staining patterns, and introduced the centromere and cytoplasmic patterns to the problem of automated HEp-2 cell recognition. The approach extracts SIFT descriptors and classifies a stained image using a bag of visual words approach. The proposed approach achieves 98.3% accuracy on a dataset collected by the authors.

Ali et al. [19] propose a supervised learning method to distinguish between staining patterns based on local information extracted from images. The approach consists of two stages: the indexing stage is based on bio-inspired features that rely on contrast information distribution in segmented cells, and the supervised learning stage selects samples which optimally represent the cell categories. These samples are used in a k-Nearest Neighbor framework to predict the class of unlabeled cells. The proposed framework is tested on the MIVIA dataset [10] using 100-fold cross validation, randomly choosing 50% images for training, while testing on the remaining images. An average global precision of 96% is reported for the protocol.

Ghosh and Chaudhary [20] present a feature extraction method for automatic recognition of staining patterns of HEp-2 images. The authors propose a composite feature set - a concatenation of Histogram of Oriented Gradients (HOG), texture-based, and Region of Interest-based features. It is reported that an overall classification accuracy of 91.13% is achieved on the MIVIA dataset for 10-fold cross validation using Support Vector Machine (SVM) classifier.

Ersoy et al. [21] present a set of complementary features that are sensitive to staining pattern variations in HEp-2 cell images. The proposed features utilize local shape measures via Hessian matrices, gradient features using adaptive robust structure tensors, and texture features. The authors apply a multi-view ShareBoost algorithm using each feature descriptor as a separate view. The algorithm uses a re-sampling distribution which is determined by the view which presents the minimum training error. The authors

suggest that this scheme reduces sensitivity to features and label noise, and the final strong classifier has increased generalization performance. Experimental results on the MIVIA dataset show an average of over 90% classification accuracy for the various HEp-2 cell classes based on the ICPR 2012 Contest protocol.

Theodorakopoulos et al. [22] present a system for automatic classification of staining patterns on single-cell fluorescence images. The method consists of the extraction of morphological features from multiple binary images derived using multi-level thresholding of the fluorescence images. A form of Uniform Local Binary Pattern descriptor is employed to capture textural information in local neighborhoods followed by classification performed using non-linear SVMs. The proposed method is evaluated using 10-fold cross-validation on the MIVIA dataset, and achieves 95.9% overall classification accuracy.

Li et al. [23] present four image descriptors for HEp-2 cell pattern classification: Local Binary Pattern (LBP), Gabor transform, Discrete Cosine Transform, and a global appearance statistical descriptor. A multi-class boosting SVM algorithm is proposed to combine these descriptors - within each boosting round, four multi-class SVMs are trained corresponding to the four descriptors, and then combined into an integrated classifier. Experimental results on the MIVIA dataset using five-fold cross-validation show the efficacy of the proposed system.

Yang et al. [24] propose to learn image statistics based filters for HEp-2 cell recognition. The authors train a filter bank from unlabeled cell images by using Independent Component Analysis (ICA). The filter bank extracts filter responses from the images. The set of extracted responses is stacked into cubic regions. Average filter responses from multiple regions are stored in feature collection matrices. A SVM classifier in conjunction with histogram correlation kernel is used to classify the images. The specimen images are divided into approximately equivalent training and testing sets. Five-fold validation is performed by randomly selecting the training and test images. Experimental analysis performed on the SNP HEp-2 dataset and the MIVIA dataset shows the efficacy of the proposed approach.

Wiliem et al. [25] present a framework focusing on the specimen image classification problem, instead of traditional HEp-2 cell-level classification tasks. A specimen-level image descriptor is proposed that is highly discriminative and is semantically meaningful at the cell level. The authors propose two maximum-margin based learning schemes to discover cell attributes while maintaining the discrimination of specimen image descriptors. The learning scheme primarily focuses on discovering image-level attributes. The authors collect a novel HEp-2 cell dataset which is specifically proposed for specimen-level classification, and demonstrate the effectiveness of the proposed framework on the dataset using five-fold validation.

Nanni et al. [26] propose a system of multiple texture descriptors for HEp-2 cell recognition. The framework

uses an ensemble consisting of approaches proposed by Das and Sengur [27], and Das et al. [28] to perform cell classification. Since cells are classified into positive and intermediate intensity groups in IIF testing (Figure 1), the authors apply histogram equalization to reduce the discrete grayscale levels in the images. For each image, a set of descriptors are extracted and used to train Support Vector Machines, whose decisions are combined using the sum rule. The images are represented using a pyramidal multi-scale representation as described by Qian et al. [29] coupled with a multi-resolution LBP. In addition, the proposed approach is fused with the methods followed by Nanni et al. [30] for handling non-uniform bins. Classification results are combined by weighted sum rule using the approach proposed in Strandmark et al. [31]. Experiments performed on the MIVIA dataset using leave-one-out validation outperform the results obtained during the MIVIA ICPR 2012 Contest.

Larsen et al. [32] present a novel method for automatic classification of HEp-2 cells. A texture index histogram that captures second-order image structures is utilized. The authors use a spatial decomposition scheme which is radially symmetrical and suitable for cell images. Spatial decomposition is performed using donut-shaped pooling regions of varying sizes when gathering histogram contributions. The proposed method is evaluated using the ICIIP 2013 and ICPR 2012 datasets, and follows the standard protocols. Experiments show that shape index histogram based approach outperforms other popular HEp-2 cell texture descriptors.

Theodorakopoulos et al. [33] propose a system for automatic classification of staining patterns on HEp-2 fluorescent images. The method encodes gradient and textural characteristics. The Scale Invariant Feature Transform (SIFT) descriptor is used in conjunction with the Gradient-oriented Co-occurrence of LBP (GoC-LBP) descriptor proposed by the authors. The GoC-LBP descriptor is based on co-occurrences of uniform Local Binary Patterns along directions determined by the orientation of the local gradient. The two descriptors are combined in the dissimilarity space using a non-linear dissimilarity function. Cell classification is performed by employing a sparse representation-based mechanism. Experiments performed using the ICPR 2012 Contest protocol as well as leave-one-out protocol show that the proposed method provides an accuracy of 75.1% for cell-level classification. A classification accuracy of 85.7% is achieved for image-level classification.

Ponomarev et al. [34] present a fully automatic method to utilize the morphological properties of stained-cell regions for automatic classification. The authors utilize the number, size, localization, and shape of the stained-cell domains to extract seventeen features. A few features are extracted from the original image and the rest are computed after the image is converted to grayscale and image thresholding is performed using the Otsu binarization method. A SVM classifier is used for recognition. Experiments performed using the benchmarking protocol achieve a 95.56% classification accuracy for image-level classification and 70.57% accuracy

for cell-level classification using 10-fold cross-validation on the training set.

Shen et al. [35] propose to integrate an intensity order pooling based feature, the Multi-support Region Order-based Gradient Histogram (MROGH), into the Bag of Words framework for HEp-2 cell image classification. Since the proposed approach does not require orientation estimation, the authors suggest that the descriptor is more robust against large variations in rotation than classic object descriptors in the literature. Experimental results show that the approach achieves 74.39% cell level accuracy and 85.71% image level accuracy on the ICPR 2012 Contest protocol.

Cataldo et al. [36] propose a classification approach based on Subclass Discriminant Analysis, a dimensionality reduction technique that provides an effective representation of the cells in the feature space. The proposed approach is robust towards the high intra-class variance typical of HEp-2 cell patterns. The individual and combined contributions of morphological, global, and local image attributes are studied to generate an adequate characterization of the fluorescence patterns. The proposed approach provides a classification accuracy of 92.9% at the image using and 72.2% at the cell level using leave-one-out protocol.

Faraki et al. [37] suggest that HEp-2 cells can be efficiently described by symmetric positive definite matrices which lie on a Riemannian manifold; the authors also extend the Bag of Word (BoW) models from Euclidean space to non-Euclidean Riemannian manifolds. The Region Covariance descriptor, originally proposed by Tuzel et al. [38], is utilized in the BoW framework. Further, Fisher tensors are proposed to encode additional information about the distribution of the signatures in the BoW model. Experiments are performed on the ICPR 2012 Contest dataset and a classification accuracy of 70% is obtained for cell-level classification using leave-one-out protocol.

Nosaka and Fukui [39] propose to use the Rotation Invariant Co-occurrence among adjacent Local Binary Pattern (RIC-LBP) image feature and a linear SVM classifier to recognize HEp-2 cells. RIC-LBP provides high descriptive ability and robustness against local rotations of the cell images. In order to create a system that is robust to rotational variations, additional training images are synthesized by rotating the images in the training dataset. The authors suggest that the proposed method is robust towards uniform changes in intensity, and invariant to local and global rotations of the image. The method is evaluated on the benchmarking protocol and achieves 68.53% accuracy for cell-level recognition using the ICPR 2012 Contest protocol and 70.65% cell-level recognition rate for the leave-one-out protocol.

Kong et al. [40] present a supervised discriminative dictionary learning algorithm for classifying HEp-2 cell patterns. The proposed algorithm utilizes the K Singular Value Decomposition algorithm; during training, it takes into account the discriminative power of the dictionary atoms and reduces their intra-class reconstruction error during

each update. The authors posit that though the proposed approach is computationally intensive, it is viable for HEp-2 cell classification due to the small number of class categories. Experiments performed using the benchmarking protocol on the MIVIA ICPR 2012 Contest dataset achieve an accuracy of 67% for cell-level recognition.

Liu and Wang [41] explore the possibility of automatically learning discriminatory information described by a set of linear projections performed on the pixel values of HEp-2 cell images. The authors propose a multi-projection-multi-codebook scheme. Projection descriptors are created and multiple image representation channels are formed where each channel corresponds to one descriptor. The image representation obtained by combining the channels is reported to be more discriminative than single-projection schemes. The proposed system achieves 66.6% cell-level classification accuracy on the MIVIA ICPR 2012 Contest protocol.

Wiliem et al. [42] propose a recognition system composed of a Cell Pyramid Matching (CPM) descriptor fused with Multiple Kernel Learning. The CPM descriptor, an adaptation of the Spatial Pyramid Matching descriptor [43], is composed of regional histograms of visual words. The authors study multiple bag-of-words descriptor variants and various spatial structures and report that Discrete Cosine Transform patch-level features along with probabilistic encoding of histograms lead to the best performance. A classification accuracy of 67% is reported for cell-level classification on the MIVIA ICPR 2012 Contest protocol.

Iannello et al. [44] study the problem of HEp-2 cell classification from a holistic standpoint. The authors present a cascade system that first discriminates between mitotic and interphase cells, and subsequently recognizes the staining pattern of the interphase cells. The system uses morphological features (mean of multivariate Gaussian distribution, standard deviation, and fraction of cell area with mean intensity higher than Otsu's threshold), texture features (mean and skewness of intensity histogram, covariance and inertia of grey-level co-occurrence matrix), and Local Binary Pattern features (autocorrelation, covariance, energy around the absolute maximum of the output image second-order histogram) for classification. The approach is evaluated using eleven different classification paradigms and four classifiers; a 3-Nearest Neighbor classifier achieves 94.3% cell classification accuracy on the MIVIA HEp-2 cell dataset.

Ensafi et al. [45] present an automatic cell image classification technique that utilizes spatial scaled image representations and sparse codings of the SIFT descriptor. The algorithm improves its prediction rate by performing sparse coding at different scales. After sparse coding, the maximal pooling of scaled images is computed at three scales. A multi-class linear SVM is used to perform classification. The proposed method is tested on the MIVIA dataset using the ICPR 2012 Contest protocol. At cell level, the classification accuracy is 72.8%, and at the image level, an accuracy of 85.8% is achieved.

Hobson et al. [46] propose a benchmarking platform for the ANA IIF HEp-2 image classification problem. The authors suggest that the HEp-2 cell classification methods in the literature are not sufficient to achieve optimal performance. Instead, a CAD system is proposed which uses object bank representations [47]. The proposed system is evaluated on a database of ANA images, collected at the Sullivan Nicolaides Pathology laboratory, acquired between 2011 and 2013 from 1,001 patients' sera. The proposed CAD system is compared to a baseline approach where the dominant cell pattern is used to determine the pattern of a given HEp-2 image, and the MES system [13]. The Cell Bank approach significantly outperforms the other methods with 79.3% mean class accuracy on the dataset.

Theodorakopoulos et al. [48] present a framework for pre-processing HEp-2 images, extracting features, and performing classification. In the pre-processing stage, a sparse representation-based technique is used to denoise and normalize the images. Morphological descriptors are extracted using multi-level thresholding and combined with local gradient descriptors to encode textural and structural information at multiple scales. The proposed method is evaluated using the ICIP 2013 contest dataset and achieves 89.20% classification accuracy following a two-fold validation protocol, performing ten iterations with random permutations of the data.

Majtner et al. [49] propose a texture-based image descriptor for HEp-2 images. The two-dimensional greyscale image is treated as a topographic surface consisting of hills and valleys. The pixel intensities are used to represent the elevation of surfaces. The descriptor computes the properties of such surfaces in multiple orientations. The proposed descriptor is compared with several information descriptors on the MIVIA HEp-2 Dataset. A Nearest-Neighbor classifier based on the proposed descriptor achieves 91.1% classification accuracy using leave-one-out protocol.

Schaefer et al. [50] present a method for classifying HEp-2 cells using texture information. The proposed approach extracts multi-dimensional LBP texture features (MD-LBP) to characterize the cell area. A margin distribution based bagging pruning classifier ensemble is used to perform recognition. The algorithm is evaluated on the MIVIA ICPR 2012 Contest protocol and 71.39% classification accuracy is achieved on the test data.

Wiliem et al. [51] propose a system that can be scaled and has competitive accuracy towards classification of HEp-2 images. The system adapts a bag of visual words approach for 256 dictionary elements. A specimen image is considered as a visual document comprising of visual vocabularies represented by cells. A specimen image is represented by histograms of vocabulary occurrences. The performance of the proposed system is studied on a set of images taken from 262 ANA positive patient sera at the Sullivan Nicolaides Pathology laboratory. A SVM classifier with histogram intersection kernel is used to perform recognition. Experiments conducted on an

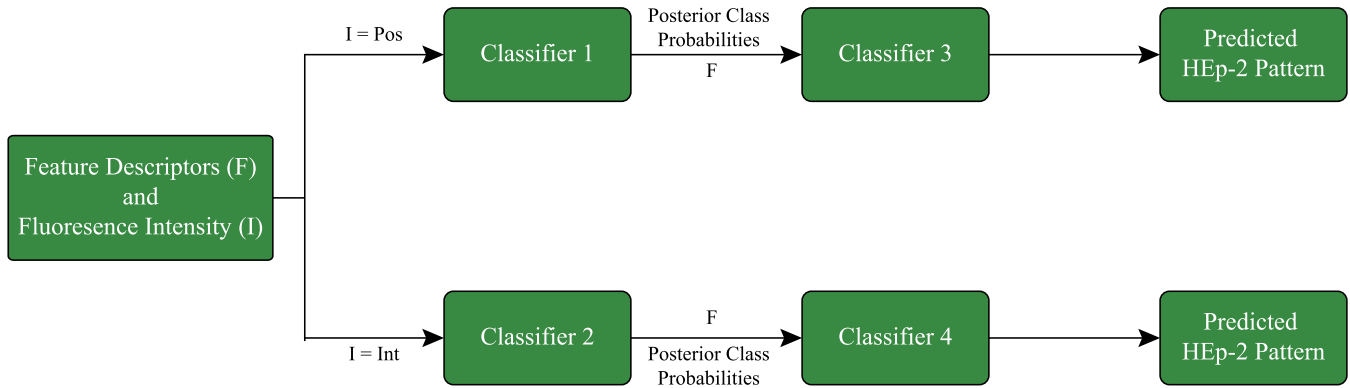


FIGURE 2. Proposed framework for HEp-2 cell classification. Images with distinct Fluorescence Intensity levels are treated individually. HEp-2 cell datasets typically consist of images represented by two Fluorescence Intensity levels - Positive and Intermediate.

in-house dataset demonstrate the efficacy of the proposed system.

The automated detection of Anti-Nuclear Antibodies using Indirect Immunofluorescence testing has significantly evolved from its humble beginnings in being able to only detect a few cell patterns with limited accuracy. Current state-of-art techniques have led to automated detection of up to six different types of HEp-2 cells with improved recognition rates. However, it is imperative that a standard is established to benchmark proposed approaches to be able to understand the relative merits and shortcomings of these approaches. Due to the availability of a constrained form of stained HEp-2 cell images, it is possible to create a community-wide benchmarking database and an accompanying protocol in order to analyze the efficacy of existing approaches as well as to evaluate proposed algorithms in the future. The ICPR 2012 Contest, the ICIP 2013 contest, and the SNP-HEp-2 dataset are significant efforts in this regard. However, the HEp-2 cell classification problem is no longer in a nascent stage of development and performance of approaches in the literature have reached a point where more comprehensive datasets with standardized protocols are necessary in order to push the envelope in automated HEp-2 cell classification.

III. PROPOSED FRAMEWORK

The research community has made extensive efforts to perform automated recognition of HEp-2 cells. However, these efforts have majorly focused on finding optimal feature representations and have not completely explored the problem of classification. In this research, we utilize the well-documented hypothesis in HEp-2 cell literature that for the same class of cells, the information encoded in cell images is distinct for positive and intermediate fluorescence intensity images. We propose a framework to recognize HEp-2 cell images using information extracted from the images as well as the quality of fluorescence intensity level of the image. Figure 2 illustrates the steps involved in the proposed framework.

A. FEATURE REPRESENTATION

The cells are extracted from the cell images according to the information provided in the database. Individual cell images are then used to extract efficient representation followed by classification. Since there is no fixed shape or geometry of different kinds of cell images, geometric features do not provide a good representation. However, it has been observed that individual cell classes have distinct texture patterns [52]. Texture is an ubiquitous and effective low-level feature for describing characteristic information associated with objects. As shown in Figure 3 and Figure 4, homogeneous pattern has an even diffusion pattern, fine speckled pattern has a distinct speckled pattern, whereas centromere has 40-60 discrete speckles. It is our assertion that encoding these characteristics can provide efficient representation and classification of HEp-2 cell classes.

Laws [53] has proposed a texture feature descriptor that measures the variations in local regions by tessellating the image in local windows of a fixed size. The descriptor uses multiple kernels to encode *texture energy*. Laws proposed three primary kernels that capture texture information - [1, 2, 1], [-1, 0, -1], [-1, 2, -1]. Further, five kernels are proposed emulating various forms of low-level structures in images - [1, 4, 6, 4, 1], [-1, 0, 2, 0, -1], [1, -4, 6, -4, 1], [-1, -2, 0, 2, 1], and [-1, 2, 0, -2, 1] representing edges, ripples, waves, lines, and spots, respectively. Texture filters are generated by computing outer products of two kernels of the same dimensions. Thus, 9 filters of size 3×3 and 25 filters of size 5×5 are formed. The 34 filters convolved with cell images represent texture images. The mean, absolute mean, and standard deviation of the texture images are computed as feature values. Thus, 102 features are calculated for an image. Agrawal et al. [16] suggest that these features can be used to model the characteristics of HEp-2 cell classes.

B. CLASSIFICATION

Once the Laws features are extracted from the individual cells, the next step is classification. Since the intensity













Cell Class →	Homogeneous	Centromere	Nucleolar	Fine Speckled	Coarse Speckled	Cytoplasmic
Positive Fluorescence Intensity Image						
Intermediate Fluorescence Intensity Image						

FIGURE 3. MIVIA ICPR 2012 Contest Dataset [10]: Examples of HEP-2 cells with various staining patterns.











Cell Class →	Homogeneous	Coarse Speckled	Fine Speckled	Nucleolar	Centromere
Positive Fluorescence Intensity Image					
Intermediate Fluorescence Intensity Image					

FIGURE 4. SNP HEP-2 Cell Dataset [11]: Examples of HEP-2 cells with various staining patterns.

levels of positive and intermediate images are different, the proposed HEP-2 cell recognition framework (Figure 2) attempts to learn independent classifiers for positive and intermediate fluorescence intensity images. For every distinct fluorescence intensity level, an independent Support Vector Machine [54] with polynomial kernel of 4th degree is learned from the Laws texture descriptors.

A standard SVM classifier provides the distance from the hyperplane along with the class label. Wu et al. [55] propose an extension of SVM classifier which provides labels along with the probability of occurrence for each class. Let \mathbf{x} represent the Laws features, the goal is to estimate the probability of occurrence p_i ,

$$p_i = P(y = i|\mathbf{x}), \quad \forall i = 1, \dots, c. \quad (1)$$

where c is the number of classes. The pair-wise class probabilities are estimated as:

$$r_{ij} \approx P(y = i|y = i \text{ or } j, \mathbf{x}) \quad (2)$$

If \hat{x} is the decision value at \mathbf{x} , the pair-wise class probability is estimated as:

$$r_{ij} \approx \frac{1}{1 + e^{P\hat{x}+Q}}, \quad (3)$$

where, P and Q are estimated by minimizing the negative log likelihood of the training data based on the labels and corresponding decision values. The above relationship is used to compute all possible pair-wise class probabilities and the following optimization is used to estimate the values of all p_i .

$$\min_{\mathbf{p}} \left(\frac{1}{2} \sum_{i=1}^n \sum_{j:j \neq i} (r_{ji}p_i - r_{ij}p_j)^2 \right) \quad (4)$$

subject to

$$p_i \geq 0, \quad \forall i \quad \text{and} \quad \sum_{i=1}^n p_i = 1 \quad (5)$$

Thus, two probabilistic models $P - SVM_1$ and $P - SVM_2$ are trained for positive and intermediate cells, respectively. The output probabilities for all six classes are concatenated

TABLE 1. Characteristics of the MIVIA and SNP HEP-2 cell datasets.

Characteristics	MIVIA	SNP HEP-2
No. of Specimen Images	28	40
No. of Images	1457	1884
No. of HEP-2 cell classes	6	5
HEP-2 cell classes	Homogeneous, Nucleolar, Centromere, Fine Speckled, Coarse Speckled, Cytoplasmic	Homogeneous, Nucleolar, Centromere, Fine Speckled, Coarse Speckled

with the d-dimensional Laws features and two new SVMs are trained - SVM_3 for positive cell images and SVM_4 for intermediate cell images.

$$p_1 = P - SVM_1(Laws_{positive}, y) \quad (6)$$

$$p_2 = P - SVM_2(Laws_{intermediate}, y) \quad (7)$$

$$p_3 = SVM_3((Laws_{positive}, p_1), y) \quad (8)$$

$$p_4 = SVM_4((Laws_{intermediate}, p_2), y) \quad (9)$$

During testing, the Laws features are extracted and depending on whether the features belong to a positive or intermediate image, the features are given as input to $P - SVM_1$ or $P - SVM_2$, respectively. $P - SVM$ outputs the probabilities for each class, which are combined with the original Laws features and input to SVM_3 or SVM_4 . The output class of SVM_3 or SVM_4 is considered as the final class of the cell image.

IV. EXPERIMENTAL ANALYSIS

The proposed algorithm is evaluated on two large publicly available databases with pre-defined protocols. We have performed multiple experiments to evaluate the performance and also to justify the selection of individual steps of the framework. The details of the databases and experimental protocols are provided in the next subsection followed by results in Sections IV-C and IV-D.

A. DATASETS AND EXPERIMENTAL PROTOCOLS

The HEP-2 cell recognition research community has introduced several HEP-2 cell datasets over the last five years. The MIVIA ICPR 2012 Contest Dataset [10] and the SNP HEP-2 Cell Dataset [11] are two of the most popular publicly available datasets. Table 1 summarizes the characteristics of the two databases.

- The MIVIA ICPR 2012 Contest dataset images are acquired by means of a fluorescence microscope coupled with a mercury vapor lamp and a digital camera. The dataset contains images pertaining to six different cell classes: centromere, coarse speckled, fine speckled, homogeneous, nucleolar, and cytoplasmic. A total of 28 specimen images, consisting of 1457 cell images are present in the MIVIA dataset. Figure 3 presents sample cell images from the dataset.

Foggia et al. [56] survey automated approaches towards classifying HEP-2 cells. Experiments are performed using the leave-one-out technique over all 28 images in the dataset: for each image in the dataset, a classifier instance is trained using 27 images; this classifier is used

TABLE 2. Number of correctly classified images in the MIVIA dataset: comparison of texture descriptors, Support Vector Machine kernels, and the effect of re-learning classifiers.

Feature	Retraining	Linear Kernel	Polynomial Kernel (deg=4)	RBF Kernel (g=3)
Laws	No	14	15	13
	Yes	20	24	20
DSIFT	No	14	13	13
	Yes	16	18	17
TP-LBP	No	15	13	12
	Yes	18	18	16

to classify the cells of the 28th image. We follow this protocol to be able to benchmark our proposed framework against state-of-the-art approaches in the literature.

- The SNP HEP-2 Cell Dataset specimen images are captured using a monochrome high dynamic range microscopy camera and an LED illumination source. The dataset has images pertaining to five cell classes: centromere, coarse speckled, fine speckled, homogeneous, and nucleolar, and consists of 1,884 cell images extracted from 40 specimen images. Figure 4 presents sample cell images from the dataset. For performance evaluation, pre-defined five-fold validation training and testing splits prepared by random selection are utilized.

We perform two sets of experiments on both the databases. The first experiment is performed to evaluate the effectiveness of individual components of the proposed algorithm whereas the second experiment is performed using the pre-defined protocols on both the databases.

B. EFFECTIVENESS OF INDIVIDUAL COMPONENTS

To evaluate the performance of Laws features, we compared its performance with two different feature descriptors: dense Scale Invariant Feature Transform, and three-patch Local Binary Pattern. The effectiveness of the two level SVM retraining is also compared with the performance obtained without retraining. The results are discussed below:

1) EFFECTIVENESS OF FEATURES

The Scale Invariant Feature Transform [57] is computed at each pixel in the HEP-2 cell image, and the resultant dense computation is referred to as DSIFT.¹ The classic Local Binary Pattern descriptor uses binary strings to encode

¹Open source VLFeat Toolbox is used to extract DSIFT features: <http://www.vlfeat.org>.

TABLE 3. Confusion matrix for cell-level classification on the MIVIA dataset using leave-one-out protocol.

Ground Truth	Prediction						
	Centromere	Homogeneous	Nucleolar	Coarse Speckled	Fine Speckled	Cytoplasmic	
Centromere	0.85	0.01	0.04	0.05	0.05	0.00	
Homogeneous	0.05	0.73	0.04	0.07	0.11	0.00	
Nucleolar	0.13	0.12	0.61	0.12	0.01	0.01	
Coarse Speckled	0.07	0.03	0.00	0.68	0.21	0.01	
Fine Speckled	0.07	0.18	0.00	0.26	0.49	0.00	
Cytoplasmic	0.00	0.00	0.00	0.12	0.00	0.88	

TABLE 4. Confusion matrix for image-level classification on the MIVIA dataset using leave-one-out protocol.

Ground Truth	Prediction						
	Centromere	Homogeneous	Nucleolar	Coarse Speckled	Fine Speckled	Cytoplasmic	
Centromere	1.00	0.00	0.00	0.00	0.00	0.00	
Homogeneous	0.00	1.00	0.00	0.00	0.00	0.00	
Nucleolar	0.00	0.00	0.75	0.25	0.00	0.00	
Coarse Speckled	0.00	0.00	0.00	0.80	0.20	0.00	
Fine Speckled	0.00	0.25	0.00	0.25	0.50	0.00	
Cytoplasmic	0.00	0.00	0.00	0.00	0.00	1.00	

TABLE 5. Image level classification results on the MIVIA dataset using leave-one-out protocol.

ID	True Class	Centromere	Homogeneous	Nucleolar	Coarse Speckled	Fine Speckled	Cytoplasmic
1	Homogeneous	0.00	0.98	0.00	0.00	0.02	0.00
2	Fine Speckled	0.02	0.48	0.00	0.21	0.29	0.00
3	Centromere	0.99	0.00	0.01	0.00	0.00	0.00
4	Nucleolar	0.10	0.12	0.38	0.39	0.01	0.00
5	Homogeneous	0.02	0.79	0.08	0.00	0.11	0.00
6	Coarse Speckled	0.05	0.00	0.00	0.88	0.07	0.00
7	Centromere	0.91	0.00	0.09	0.00	0.00	0.00
8	Nucleolar	0.36	0.2	0.62	0.00	0.00	0.00
9	Fine Speckled	0.00	0.02	0.00	0.48	0.50	0.00
10	Coarse Speckled	0.33	0.00	0.00	0.30	0.37	0.00
11	Coarse Speckled	0.00	0.07	0.00	0.73	0.20	0.00
12	Coarse Speckled	0.00	0.00	0.00	0.71	0.25	0.04
13	Centromere	0.96	0.02	0.00	0.02	0.00	0.00
14	Centromere	0.41	0.03	0.00	0.27	0.29	0.00
15	Fine Speckled	0.22	0.03	0.00	0.02	0.73	0.00
16	Centromere	1.00	0.00	0.00	0.00	0.00	0.00
17	Coarse Speckled	0.00	0.21	0.00	0.42	0.37	0.00
18	Homogeneous	0.00	0.71	0.00	0.23	0.03	0.03
19	Centromere	0.85	0.00	0.15	0.00	0.00	0.00
20	Nucleolar	0.11	0.35	0.50	0.00	0.00	0.04
21	Homogeneous	0.26	0.36	0.00	0.08	0.30	0.00
22	Homogeneous	0.00	0.76	0.09	0.06	0.09	0.00
23	Fine Speckled	0.00	0.21	0.00	0.42	0.37	0.00
24	Nucleolar	0.00	0.07	0.89	0.04	0.00	0.00
25	Cytoplasmic	0.00	0.00	0.00	0.13	0.00	0.87
26	Cytoplasmic	0.00	0.00	0.00	0.16	0.00	0.84
27	Cytoplasmic	0.00	0.00	0.00	0.11	0.00	0.89
28	Cytoplasmic	0.00	0.00	0.00	0.00	0.00	1.00

local texture around a pixel; the three-patch LBP² (TP-LBP) descriptor [58] is computed at each pixel by comparing values of three patches around a small patch centered on the pixel. SVM classifiers are independently trained on texture features, Laws, D-SIFT, and TP-LBP, extracted from HEp-2 cell images. For brevity and ease of objective analysis, we only present the comparative results for the texture descriptors in the form of image-level classification accuracy on the MIVIA database.

²Open source code available at <http://www.openu.ac.il/home/hassner/projects/Patchlbp/TP-LBP.m> is used to compute TP-LBP features.

The number of correctly classified images with different texture descriptors is presented in Table 2. It is observed that the Laws texture descriptor performs either at par with, or better than three-patch LBP and dense SIFT. With retraining, in all the cases Laws yields significantly better results compared to other two features.

2) EFFECTIVENESS OF RETRAINING

The proposed classification framework is devised based on inspiration from cascaded SVM classifiers. The features extracted from HEp-2 cell images may not provide sufficient discriminatory information. In such cases, the class

TABLE 6. Confusion matrix for cell-level classification on the SNP HEP-2 dataset.

Ground Truth	Prediction					
	Homogenous	Nucleolar	Centromere	Coarse Speckled	Fine Speckled	
Homogenous	1.00	0.00	0.00	0.00	0.00	0.00
Nucleolar	0.00	0.82	0.10	0.08	0.00	0.00
Centromere	0.03	0.03	0.73	0.21	0.00	0.00
Coarse Speckled	0.00	0.03	0.16	0.60	0.21	0.00
Fine Speckled	0.15	0.03	0.00	0.07	0.75	0.00

probabilities act as a complimentary set of features. The effectiveness of retraining is evaluated by comparing the performance with and without retraining. Laws features of positive HEP-2 images are classified with SVM_1 and features of intermediate images are classified with SVM_2 . The results of both cases are shown in Table 2. For every feature, the results are computed with three kernels (linear, polynomial, and Radial Basis Function (RBF)) with and without retraining on the MIVIA database. Agrawal et al. [16] have shown that for HEP-2 classification, the performance of Random Decision Forest is comparative to SVM. Therefore, we have only evaluated the performance of SVM with multiple kernels. With multiple kernels as well as features, the performance of retraining is significantly better than without retraining. Among all the kernels, polynomial kernel with degree four is observed to yield the best results.

C. EVALUATION ON THE MIVIA ICPR 2012 CONTEST DATASET

We present the performance of our proposed framework on the MIVIA database with leave-one-out protocol. The pre-defined protocol requires reporting the confusion matrix for both cell level and image level classification. Table 3 shows the cell-level confusion matrix which is obtained by adding the classification results of the 28 leave-one-out trials. The class of the image is reported as the most frequently occurring cell class in the given image (Table 4). As per the protocol, in Table 5, we also summarize the results for every test image in the leave-one-out protocol.

As shown in Table 3, centromere, homogeneous, and cytoplasmic are very well classified with the correct classification rate lying between 0.73 and 0.88. The minimum accuracy is obtained for fine speckled cells and it is often misclassified as coarse-speckled (0.26) and homogeneous (0.18). Similar to cell level classification, all the images belonging to centromere, homogeneous, and cytoplasmic are correctly classified and fine speckled images are most often misclassified. The overall image-level classification performance of the proposed algorithm on the MIVIA ICPR 2012 Contest dataset is 85.71%, whereas the cell-level classification accuracy is 70.89%.

D. EVALUATION ON THE SNP HEP-2 CELL DATASET

SNP HEP-2 protocol suggests five-fold cross validation with 905 cell images for training and 979 cell images for testing. The proposed framework is trained using the pre-defined

TABLE 7. Comparative performance of state-of-the-art approaches in automated HEP-2 cell recognition in terms of correct classification accuracy.

Authors	MIVIA Protocol 2		SNP HEP-2
	Cell	Image	
Di Cataldo et al. [36]	89.55%	96.43%	-
Faraki et al. [37]	71.70%	75.00%	74.7%
Theodorakopoulos et al. [33]	64.90%	75.00%	-
Kong et al. [40]	63.16%	71.43%	-
Nosaka and Fukui [39]	70.65%	85.71%	-
Wiliem et al. [42]	-	-	82.5%
Yang et al. [24]	-	-	82.1%
Proposed method	70.89%	85.71%	80.9%

training database and the results are reported on the testing database. Table 6 represents the cell-level confusion matrix for each of the five classes present in the database. The average accuracy of the proposed framework across five-folds is 80.9% with standard deviation of 1.03%. Here, homogeneous and nucleolar have the maximum correct classification rate and unlike the results on the MIVIA database, coarse speckled is often misclassified. It is worth mentioning that the results for the SNP HEP-2 cell dataset indicate cell-level recognition and the performance is likely to be higher for automated HEP-2 cell recognition at the specimen image level. However, unlike MIVIA ICPR 2012 contest, specimen image information is not provided in the SNP dataset.

E. COMPARISON WITH EXISTING ALGORITHMS

As discussed in literature review, several researchers have proposed models for automation of HEP-2 cell classification. Since the protocols are pre-defined, we can directly compare the performance of the proposed framework with existing results. Table 7 summarizes the results on both MIVIA and SNP-HEP2 databases. The results show that the classification performance of the proposed framework on both the databases is among the top-3 algorithms on both the databases. Only Faraki et al. [37] report results on both the databases, and it is observed that their method yields classification accuracies ranging between 71.70% and 75% on the two databases. The proposed algorithm yields 70.9% accuracy on the MIVIA cell-level protocol, 85.71% accuracy on MIVIA image-level protocol, and 80.9% accuracy on SNP HEP-2 database.

V. CONCLUSION

Recently, a number of methods have explored sophisticated information descriptors and classification paradigms to study

the problem of automated HEP-2 cell classification. This paper presents a detailed review of the literature pertaining to HEP-2 cell classification and discusses the challenges associated with this research problem. We also present a framework which combines Laws features with two level SVM classifier coupled with posterior class probabilities during the classification stage. Experiments on two publicly available databases show the effectiveness of the proposed algorithm which yields high classification accuracies.

ACKNOWLEDGMENT

Ishan Nigam and Shreyasi Agrawal equally contributed to this work. The authors thank the reviewers and associate editor for their feedback and constructive comments.

REFERENCES

- J. D. Rioux and A. K. Abbas, "Paths to understanding the genetic basis of autoimmune disease," *Nature*, vol. 435, no. 7042, pp. 584–589, 2005.
- M. Satoh et al., "Prevalence and sociodemographic correlates of antinuclear antibodies in the United States," *Arthritis Rheumatism*, vol. 64, no. 7, pp. 2319–2327, Jul. 2012.
- D. L. Fairweather, S. Frisnacho-Kiss, and N. R. Rose, "Sex differences in autoimmune disease from a pathological perspective," *Amer. J. Pathol.*, vol. 173, no. 3, pp. 600–609, Sep. 2008.
- C. Castro and M. Gourley, "Diagnostic testing and interpretation of tests for autoimmunity," *J. Allergy Clin. Immunol.*, vol. 125, no. 2, pp. S238–S247, Feb. 2010.
- P. L. Meroni and P. H. Schur, "ANA screening: An old test with new recommendations," *Ann. Rheumatic Diseases*, vol. 69, no. 8, pp. 1420–1422, 2010.
- M. J. Fritzler, "The antinuclear antibody test: Last or lasting gasp?" *Arthritis Rheumatism*, vol. 63, no. 1, pp. 19–22, Jan. 2011.
- A. Rigon, P. Soda, D. Zennaro, G. Iannello, and A. Afeltra, "Indirect immunofluorescence in autoimmune diseases: Assessment of digital images for diagnostic purpose," *Cytometry B, Clin. Cytometry*, vol. 72B, no. 6, pp. 472–477, Nov. 2007.
- A. Rigon et al., "Novel opportunities in automated classification of antinuclear antibodies on HEP-2 cells," *Autoimmunity Rev.*, vol. 10, no. 10, pp. 647–652, Aug. 2011.
- C. Bonroy et al., "Automated indirect immunofluorescence antinuclear antibody analysis is a standardized alternative for visual microscope interpretation," *Clin. Chem. Lab. Med.*, vol. 51, no. 9, pp. 1771–1779, 2013.
- P. Foggia, G. Percannella, P. Soda, and M. Vento, "Early experiences in mitotic cells recognition on HEP-2 slides," in *Proc. 23rd IEEE Int. Symp. Comput.-Based Med. Syst.*, Oct. 2010, pp. 38–43.
- A. Wiliem, Y. Wong, C. Sanderson, P. Hobson, S. Chen, and B. C. Lovell, "Classification of human epithelial type 2 cell indirect immunofluorescence images via codebook based descriptors," in *Proc. IEEE Workshop Appl. Comput. Vis.*, Jan. 2013, pp. 95–102.
- P. Perner, H. Perner, and B. Müller, "Mining knowledge for HEP-2 cell image classification," *Artif. Intell. Med.*, vol. 26, nos. 1–2, pp. 161–173, Sep./Oct. 2002.
- P. Soda and G. Iannello, "Aggregation of classifiers for staining pattern recognition in antinuclear autoantibodies analysis," *IEEE Trans. Inf. Technol. Biomed.*, vol. 13, no. 3, pp. 322–329, May 2009.
- E. Cordelli and P. Soda, "Color to grayscale staining pattern representation in IIF," in *Proc. 24th Int. Symp. Comput. Based Med. Syst.*, Jun. 2011, pp. 1–6.
- U. Sack et al., "Autoantibody detection using indirect immunofluorescence on HEP-2 cells," *Ann. New York Acad. Sci.*, vol. 1173, no. 1, pp. 166–173, Sep. 2009.
- P. Agrawal, M. Vatsa, and R. Singh, "HEP-2 cell image classification: A comparative analysis," in *Machine Learning in Medical Imaging Lecture Notes in Computer Science*, vol. 8184. Springer International Publishing, 2013, pp. 195–202.
- P. Foggia, G. Percannella, P. Soda, and M. Vento, "Benchmarking HEP-2 cells classification methods," *IEEE Trans. Med. Imag.*, vol. 32, no. 10, pp. 1878–1889, Oct. 2013.
- G. Iannello, L. Onofri, and P. Soda, "A bag of visual words approach for centromere and cytoplasmic staining pattern classification on HEP-2 images," in *Proc. 25th Int. Symp. Comput.-Based Med. Syst.*, Jun. 2012, pp. 1–6.
- W. B. H. Ali, P. Piro, D. Giampaglia, T. Pourcher, and M. Barlaud, "Biological cells classification using bio-inspired descriptor in a boosting k-NN framework," in *Proc. 25th Int. Symp. Comput.-Based Med. Syst.*, Jun. 2012, pp. 1–6.
- S. Ghosh and V. Chaudhary, "Feature analysis for automatic classification of HEP-2 fluorescence patterns: Computer-aided diagnosis of autoimmune diseases," in *Proc. 21st Int. Conf. Pattern Recognit.*, Nov. 2012, pp. 174–177.
- I. Ersoy, F. Bunyak, J. Peng, and K. Palaniappan, "HEP-2 cell classification in IIF images using Shareboost," in *Proc. 21st Int. Conf. Pattern Recognit.*, Nov. 2012, pp. 3362–3365.
- I. Theodorakopoulos, D. Kastaniotis, G. Economou, and S. Fotopoulos, "HEP-2 cells classification via fusion of morphological and textural features," in *Proc. 12th IEEE Int. Conf. Bioinform. Bioeng.*, Nov. 2012, pp. 689–694.
- K. Li, J. Yin, Z. Lu, X. Kong, R. Zhang, and W. Liu, "Multiclass boosting SVM using different texture features in HEP-2 cell staining pattern classification," in *Proc. 21st Int. Conf. Pattern Recognit.*, Nov. 2012, pp. 170–173.
- Y. Yang, A. Wiliem, A. Alavi, and P. Hobson, "Classification of human epithelial type 2 cell images using independent component analysis," in *Proc. 20th IEEE Int. Conf. Image Process.*, Sep. 2013, pp. 733–737.
- A. Wiliem, P. Hobson, and B. C. Lovell, "Discovering discriminative cell attributes for HEP-2 specimen image classification," in *Proc. IEEE Winter Conf. Appl. Comput. Vis.*, Mar. 2014, pp. 423–430.
- L. Nanni, M. Paci, and S. Brahmam, "Indirect immunofluorescence image classification using texture descriptors," *Expert Syst. Appl.*, vol. 41, no. 5, pp. 2463–2471, Apr. 2014.
- R. Das and A. Sengur, "Evaluation of ensemble methods for diagnosing of valvular heart disease," *Expert Syst. Appl.*, vol. 37, no. 7, pp. 5110–5115, Jul. 2010.
- R. Das, I. Türkoglu, and A. Sengur, "Diagnosis of valvular heart disease through neural networks ensembles," *Comput. Methods Programs Biomed.*, vol. 93, no. 2, pp. 185–191, Feb. 2009.
- X. Qian, X.-S. Hua, P. Chen, and L. Ke, "PLBP: An effective local binary patterns texture descriptor with pyramid representation," *Pattern Recognit.*, vol. 44, nos. 10–11, pp. 2502–2515, Oct./Nov. 2011.
- L. Nanni, S. Brahmam, and A. Lumini, "A simple method for improving local binary patterns by considering non-uniform patterns," *Pattern Recognit.*, vol. 45, no. 10, pp. 3844–3852, Oct. 2012.
- P. Strandmark, J. Ulén, and F. Kahl, "HEP-2 staining pattern classification," in *Proc. 21st IEEE Int. Conf. Pattern Recognit.*, Nov. 2012, pp. 33–36.
- A. B. L. Larsen, J. S. Vestergaard, and R. Larsen, "HEP-2 cell classification using shape index histograms with donut-shaped spatial pooling," *IEEE Trans. Med. Imag.*, vol. 33, no. 7, pp. 1573–1580, Jul. 2014.
- I. Theodorakopoulos, D. Kastaniotis, G. Economou, and S. Fotopoulos, "HEP-2 cells classification via sparse representation of textural features fused into dissimilarity space," *Pattern Recognit.*, vol. 47, no. 7, pp. 2367–2378, Jul. 2014.
- G. V. Ponomarev, V. L. Arlazarov, M. S. Gelfand, and M. D. Kazanov, "ANA HEP-2 cells image classification using number, size, shape and localization of targeted cell regions," *Pattern Recognit.*, vol. 47, no. 7, pp. 2360–2366, Jul. 2014.
- L. Shen, J. Lin, S. Wu, and S. Yu, "HEP-2 image classification using intensity order pooling based features and bag of words," *Pattern Recognit.*, vol. 47, no. 7, pp. 2419–2427, Jul. 2014.
- S. Di Cataldo, A. Bottino, I. U. Islam, T. Vieira, and E. Ficarra, "Sub-class Discriminant Analysis of morphological and textural features for HEP-2 staining pattern classification," *Pattern Recognit.*, vol. 47, no. 7, pp. 2389–2399, Jul. 2014.
- M. Faraki, M. T. Harandi, A. Wiliem, and B. C. Lovell, "Fisher tensors for classifying human epithelial cells," *Pattern Recognit.*, vol. 47, no. 7, pp. 2348–2359, Jul. 2014.
- O. Tuzel, F. Porikli, and P. Meer, "Region covariance: A fast descriptor for detection and classification," in *Proc. Eur. Conf. Comput. Vis.*, vol. 3952. 2006, pp. 589–600.
- R. Nosaka and K. Fukui, "HEP-2 cell classification using rotation invariant co-occurrence among local binary patterns," *Pattern Recognit.*, vol. 47, no. 7, pp. 2428–2436, Jul. 2014.

- [40] X. Kong, K. Li, J. Cao, Q. Yang, and L. Wenyin, "HEP-2 cell pattern classification with discriminative dictionary learning," *Pattern Recognit.*, vol. 47, no. 7, pp. 2379–2388, Jul. 2014.
- [41] L. Liu and L. Wang, "HEP-2 cell image classification with multiple linear descriptors," *Pattern Recognit.*, vol. 47, no. 7, pp. 2400–2408, Jul. 2014.
- [42] A. Wiliem, C. Sanderson, Y. Wong, P. Hobson, R. F. Minchin, and B. C. Lovell, "Automatic classification of human epithelial type 2 cell indirect immunofluorescence images using cell pyramid matching," *Pattern Recognit.*, vol. 47, no. 7, pp. 2315–2324, Jul. 2014.
- [43] S. Lazebnik, C. Schmid, and J. Ponce, "Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2006, pp. 2169–2178.
- [44] G. Iannello, G. Percannella, P. Soda, and M. Vento, "Mitotic cells recognition in HEP-2 images," *Pattern Recognit. Lett.*, vol. 45, pp. 136–144, Aug. 2014.
- [45] S. Ensafi, S. Lu, A. A. Kassim, and C. L. Tan, "Automatic CAD system for HEP-2 cell image classification," in *Proc. 22nd Int. Conf. Pattern Recognit.*, Aug. 2014, pp. 3321–3326.
- [46] P. Hobson, B. C. Lovell, G. Percannella, M. Vento, and A. Wiliem, "Classifying anti-nuclear antibodies HEP-2 images: A benchmarking platform," in *Proc. 22nd Int. Conf. Pattern Recognit.*, Aug. 2014, pp. 3233–3238.
- [47] L.-J. Li, H. Su, L. Fei-fei, and E. P. Xing, "Object bank: A high-level image representation for scene classification & semantic feature sparsification," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 23, 2010, pp. 1378–1386.
- [48] I. Theodorakopoulos, D. Kastaniotis, G. Economou, and S. Fotopoulos, "HEP-2 cells classification using morphological features and a bundle of local gradient descriptors," in *Proc. 1st Workshop Pattern Recognit. Techn. Indirect Immunofluorescence Images*, Aug. 2014, pp. 33–36.
- [49] T. Majtner, R. Stoklasa, and D. Svoboda, "RSURF—The efficient texture-based descriptor for fluorescence microscopy images of HEP-2 cells," in *Proc. 22nd Int. Conf. Pattern Recognit.*, Aug. 2014, pp. 1194–1199.
- [50] G. Schaefer, N. P. Doshi, S. Y. Zhu, and Q. Hu, "Analysis of HEP-2 images using MD-LBP and MAD-bagging," in *Proc. 36th Annu. Int. Conf. Eng. Med. Biol. Soc.*, Aug. 2014, pp. 4248–4251.
- [51] A. Wiliem, P. Hobson, R. F. Minchin, and B. C. Lovell, "A bag of cells approach for antinuclear antibodies HEP-2 image classification," *Cytometry A*, vol. 87, no. 6, pp. 549–557, Jun. 2014.
- [52] C. Buchner, C. Bryant, A. Eslami, and G. Lakos, "Anti-nuclear antibody screening using HEP-2 cells," *J. Vis. Experim.*, vol. 88, p. e51211, May 2014.
- [53] K. I. Laws, "Textured image segmentation," Signal Image Process. Inst., Univ. Southern California, Los Angeles, CA, USA, Tech. Rep. USCIP-940, 1980.
- [54] V. N. Vapnik, *Statistical Learning Theory*, vol. 1. New York, NY, USA: Wiley, 1998.
- [55] T.-F. Wu, C.-J. Lin, and R. C. Weng, "Probability estimates for multi-class classification by pairwise coupling," *J. Mach. Learn. Res.*, vol. 5, pp. 975–1005, Dec. 2004.
- [56] P. Foggia, G. Percannella, A. Saggese, and M. Vento, "Pattern recognition in stained HEP-2 cells: Where are we now?" *Pattern Recognit.*, vol. 47, no. 7, pp. 2305–2314, Jul. 2014.
- [57] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, Nov. 2004.
- [58] L. Wolf, T. Hassner, and Y. Taigman, "Descriptor Based Methods in the Wild," in *Proc. Workshop Faces 'Real-Life' Images, Detect., Alignment, Recognit.*, Oct. 2008, pp. 1–14.



SHREYASI AGRAWAL received the B.Tech. degree in computer science from the Indraprastha Institute of Information Technology Delhi, India, in 2014. She is currently a member of the Technical Staff with Adobe Systems, India. Her research interests include medical imaging and pattern recognition.



RICHA SINGH (S'04–M'09–SM'14) received the M.S. and Ph.D. degrees in computer science from West Virginia University, Morgantown, USA, in 2005 and 2008, respectively. She is currently an Associate Professor and Kusum and Mohandas Pai Faculty Research Fellow with the Indraprastha Institute of Information Technology Delhi, India. Her research has been funded by the UIDAI and DeitY, India. She is a recipient of the FAST Award by DST, India. Her areas of interest are biometrics, pattern recognition, and machine learning. She is also an Editorial Board Member of *Information Fusion* (Elsevier) and the *EURASIP Journal of Image and Video Processing*. She is a member of the Computer Society and the Association for Computing Machinery. She has co-authored over 150 research papers and received several best paper and best poster awards in international conferences. She was a recipient of the NVIDIA Innovations Award 2015 and the Best Reviewer Award at the IAPR International Conference on Biometrics 2013. She serves as the PC Co-Chair of the IEEE International Conference on Biometrics: Theory, Applications, and Systems 2016.



MAYANK VATSA (S'04–M'09–SM'14) received the M.S. and Ph.D. degrees in computer science from West Virginia University, Morgantown, USA, in 2005 and 2008, respectively. He is currently an Associate Professor and AR Krishnaswamy Faculty Research Fellow with the Indraprastha Institute of Information Technology Delhi, India. His research has been funded by the UIDAI, DST, and DeitY. He has authored over 150 research papers and received several best paper and best poster awards, including the NVIDIA Innovations Award 2015.

His areas of interest are biometrics, image processing, computer vision, and information fusion. He is a member of the Computer Society and the Association for Computing Machinery. He is a recipient of the FAST Award by DST, India. He is the Vice President (Publications) of IEEE Biometrics Council and an Associate Editor of the IEEE ACCESS. He is an Area Editor of the IEEE BIOMETRIC COMPENDIUM AND INFORMATION FUSION JOURNAL (Elsevier). He was also the Program Committee Co-Chair of the IAPR International Conference on Biometrics 2013 and the International Joint Conference on Biometrics 2014.

• • •



ISHAN NIGAM received the B.E. degree in information technology from the University of Delhi, India, in 2014. He is currently a Research Associate with the Indraprastha Institute of Information Technology Delhi, India. His research interests include computer vision, pattern recognition, and deep learning. He has served as a Reviewer of the *Information Fusion* (Elsevier) journal. He received the Best Poster Award at the IEEE BTAS 2015 Conference.