A Slightly Supervised Approach for Positive/Negative Classification of Fluorescence Intensity in HEp-2 Images

Giulio Iannello, Leonardo Onofri, and Paolo Soda

Computer Science & Bioinformatics Laboratory
Università Campus Bio-Medico di Roma, Integrated Research Centre
{g.iannello,l.onofri,p.soda}@unicampus.it

Abstract. Indirect Immunofluorescence on HEp-2 slides is the recommended technique to detect antinuclear autoantibodies in patient serum. Such slides are read at the fluorescence microscope by experts of IIF, who classify the fluorescence intensity, recognize mitotic cells and classify the staining patterns for each well. The crucial need of accurately performed and correctly reported laboratory determinations has motivated recent research on computer-aided diagnosis tools in IIF to support the HEp-2 image classification. Such systems adopt a fully supervised classification approach and, hence, their chance of success depends on the quality of ground truth used to train the classification algorithms. Besides being expensive and time consuming, collecting a large and reliable ground truth in IIF is intrinsically hard due to the inter- and intra-observer variability. In order to overcome such limitations, this paper presents a slightly supervised approach for positive/negative fluorescence intensity classification. The classification phase consists in matching parts of interest automatically detected in the test image with a Gaussian mixture model built over few control images. The approach, whose operating configuration can be adapted to the cost of misclassifications, has been tested over a database with 914 images acquired from 304 different wells, achieving remarkable results on positive/negative screening task.

1 Introduction

Antinuclear autoantibodies (ANAs) are important markers to diagnose autoimmune diseases, and the indirect immunofluorescence (IIF) assay on HEp-2 slides is the recommended method for their detection [2]. IIF makes use of a substrate containing a specific antigen that can bond with serum antibodies, forming a molecular complex. Then, this complex reacts with human immunoglobulin conjugated with a fluorochrome, making the complex observable at the fluorescence microscope. Each well, which is the part of the slide containing the serum of one patient, is read by a physician and the diagnostic procedure consists of fluorescence intensity classification, mitotic cell classification and staining pattern recognition.

With reference to fluorescence intensity classification, the guidelines [2] suggest scoring it semi-quantitatively using a scale whose scores range from 0 (negative samples) up to 4+ (brilliant green, maximal fluorescence), passing through 1+ (very subdued fluorescence), 2+ (defined pattern but diminished fluorescence) and 3+ (less brilliant green fluorescence). Since technical problems can affect test sensitivity and

A. Petrosino (Ed.): ICIAP 2013, Part II, LNCS 8157, pp. 319–328, 2013.

specificity, the same guidelines suggest comparing the sample with a positive and a negative control. The former allows the physician to check the correctness of the preparation process, whereas the latter represents the auto-fluorescence level of the slide under examination. Recently, Rigon *et al.* [11] statistically analyzed the variability between a set of physician's fluorescence intensity classifications and then proposed to classify the fluorescence intensity into three classes, corresponding to images with a *high*, *intermediate* and *low* fluorescence intensity. Although the detailed description of these classes is out of the scope of this paper, it is useful to report that high and intermediate fluorescence intensities correspond to sera positive to ANAs, i.e. they contain autoantibodies and need further investigation to set the final diagnosis (e.g. mitotic cell detection and staining pattern recognition). Conversely, a low fluorescence intensity implies that the patient serum does not contain autoantibodies, i.e. the result of the test is negative.

In autoimmune diseases, the availability of accurately performed and correctly reported laboratory determinations is crucial for the clinicians, thus motivating recent research on computer-aided diagnosis (CAD) tools in IIF to support the various phases of HEp-2 image classification, e.g. [3,5,13,15,16] to cite a few. It is worth noting that existing proposal adopt a fully supervised classification approach and, hence, their chance of success depends on the quality of ground truth used to train the classification algorithms. Besides being expensive and time consuming, collecting a large and reliable ground truth is intrinsically hard due to the inter- and intra-observer variability.

Within this framework, we present here an approach for fluorescence intensity classification which aims at distinguishing between positive and negative samples, i.e. those containing and not containing autoantibodies at 1:80 titer or higher, regardless of their degree of positiveness (i.e. high or intermediate intensity). This capability corresponds to perform a pre-selection of the cases that have to be examined, making easier to carry out mass screening campaign. Our research is motivated by recent literature and medical demands, where there is a general agreement on the fact that this capability is the first goal of a CAD system in IIF.

2 Approach

In this section we present the approach we use to classify the fluorescence intensity of a well as positive or negative, from its motivations up to its description.

2.1 Motivations

In machine learning and pattern recognition it is well known that a crucial issue of supervised classification approaches is the quality and reliability of the ground truth. This consideration holds also for IIF image classification, where existing CAD approaches use a ground truth constituted by a set of images labeled by physicians. However, IIF suffers from inter- and intra-observer variability [1,11], which could affect the robustness of the ground truth.

We notice that most of the disagreements between physicians during HEp-2 image annotation occur when they have to distinguish between a weak high intensity from an intermediate intensity and between an intermediate intensity from a low intensity [11,13]. This observation suggests us to develop a slightly supervised approach

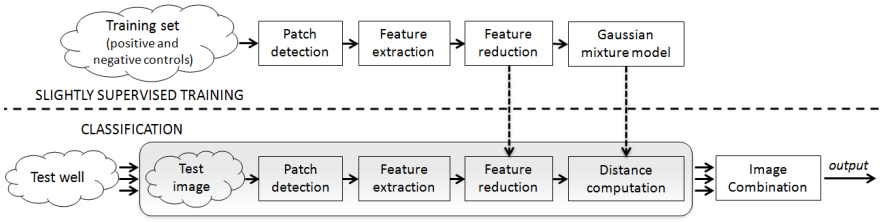

Fig. 1. Schematic of the proposed approach

which does not suffer from uncertainty introduced by ground truth collection and annotation. It requires only a set of positive and negative controls, whose classification does not introduce any ambiguity. Indeed, medical guidelines requires to use two known sera, one positive and one negative to prepare the control wells which will exhibit a high and a low intensity, respectively. In this way no expertise is required to develop the reference set of positive and negative controls and the approach does not introduce the burden of intermediate samples in the classification process.

Let us now report some further considerations on existing CAD approaches for fluorescence intensity classification of HEp-2 image. First, we observe that they can be differentiated by the dilution ratio used to prepare the samples [4,5,10,12,13]. Indeed, while in [13] it is used the 1:80 dilution ratio, the other works make use of a titer larger than or equal to 1:160. In the former case, which is the one recommended by the guidelines [2], the sample fluorescence intensity distributes over the three aforementioned classes, i.e. high, intermediate and low. In the latter case, at dilution ratios larger than or equal to 1:160, the intermediate class disappears, reducing the variability in fluorescence intensity and making the classification task easier than at 1:80 dilution. Second, we note that such works extract the features considering the whole image [13], or segmenting out the cells [4,5,10]. However, both these approaches present some open issues. In the former case, the noise introduced by background information may lead to poor performance. In the latter case, the performance strongly depend on the segmentation procedure adopted, which is typically threshold-based because it is assumed that the whole cells' body is brighter than the background. However, cell segmentation in IIF is a hard task, as proven in [6,7,9] which showed that, when HEp-2 cells exhibit either a low fluorescence intensity or some particular kind of staining patterns, the application of a thresholding procedure is not sufficient to obtain a proper segmentation of the cells.

2.2 Our Proposal

Fig. 1 schematically depicts both training and classification phases of our approach that aims at distinguishing between positive and negative samples. For this reason, we introduce the superclass of *positive* (p) samples, which is composed of samples with high and intermediate fluorescence intensity. On the other side, samples with a low intensity will constitute the negative class (n). Furthermore, our proposal is suited to classify samples collected at 1:80 titer, i.e. a dilution ratio producing images with a

variability in fluorescence intensity larger than the one given by higher titers (1:160 or even more), thus making the classification task more complex.

The upper panel of Fig. 1 represents the slightly supervised training phase. It consists of the four main steps, namely: (i) automatic detection of the parts of interest (also referred to as *patches* in the following) in the positive and negative control images, (ii) feature extraction from these patches, (iii) feature space reduction, and (iv) construction of two models representing patches of positive and negative controls, respectively.

The bottom panel represents the testing phase, which consists of steps similar to those of the training phase. The shaded area shows that patches computed from the image at hand are classified using the models derived from positive and negative controls during the training phase. Furthermore, there is also a combination step due to the fact that a single image acquired with the microscope magnification typically used in IIF does not cover all well surface. For this reason, several images are always acquired, thus permitting us to exploit a certain degree of redundancy through the integration of information extracted from different images of the same well. Hence, the classifications of images belonging to the same well are then combined by a given rule.

The following paragraphs detail the aforementioned steps of both training and testing phases.

Patch Detection. Both approaches extracting the features from the whole image and after a segmentation step suffer from the limitations reported in section 2.1. To overcome such issues, we apply the SIFT algorithm [8] that returns a set of point locations in that regions of the image with significant local intensity changes in spatial dimension, thus detecting the regions of the image where the cells are located. Since several SIFT points are usually detected in a region corresponding to a single cell, we aggregate in a single patch the points close to each other. This operation is performed using a hierarchical clustering procedure. Initializing each point as a single cluster, the algorithm iteratively merge the cluster r and s having the minimum distance according to Ward's distance definition:

$$d(r,s) = \sqrt{\frac{2n_r n_s}{(n_r + n_s)}} \|\bar{x}_r - \bar{x}_s\|_2, \tag{1}$$

where $\| \ \|_2$ is Euclidean distance; $\bar{x}_r$ and $\bar{x}_s$ are the centroids of clusters r and s, respectively; n_r and n_s are the number of elements in clusters r and s, respectively.

The high variability of fluorescence intensity and the number of cells in the images may introduce two cases affecting the effectiveness of patch extraction procedure. First, a small number of SIFT points may be detected when the intensity is very low, which in turn leads to few patches extracted from the image. Second, when there is a high density of cells, the algorithm may return a single patch covering many cells that are grouped together. In the former case, we force the algorithm to cover a minimum area by extracting patches in random positions when too few SIFT points are detected. In the latter case, we apply an additional stage which increases the size of the patches whose area is smaller than a threshold A_{min} and splits the patches whose area exceeds the threshold A_{max}. Considering the average dimension of the cells, we set the values of A_{min} and A_{max} to 400 and $A/16$ pixels, respectively, where A is the area of the whole image.

Feature Extraction and Reduction. The SIFT algorithm includes the extraction of local descriptors, i.e. histogram of oriented gradient (HoG), computed in a neighborhood of the detected interest points. However, we deem that SIFT descriptors are not suited for our intensity classification task for the following two reasons. First, the aggregation of several SIFT points reported in previous paragraph produces interest regions that are considerably larger than the neighborhood's size normally considered for SIFT descriptors. Second, HoG encodes gradient directions on small square blocks of subcellular regions, whereas gradient directions are not relevant features for describing fluorescence intensity. On these grounds, we extract from each detected patch a set of 19 features based on first and second-order gray-level histograms. The rationale lies in the meaning of these histograms: the former describes gray-level distributions, whereas the latter generally provides a good representation of the overall nature of the texture. Furthermore, it is worth noting that similar descriptors have already been used in previous works on IIF image classification [13,14].

Finally, we perform a dimensionality reduction of the feature space to achieve a more compact and discriminant description. In this step, we explored the principal component analysis (PCA) and linear discriminant analysis (LDA).

Model Construction. Our approach relies upon the definition of models describing the distribution of patches belonging to different fluorescence intensity. To this end, we consider patches belonging to the positive and negative controls that unequivocally correspond to high (h) and low (l) fluorescence intensity, respectively. In the feature space induced by the dimensionality reduction, we fit a two component Gaussian mixture (GM) model providing us the parameters $(\boldsymbol{\mu}_h, \boldsymbol{\Sigma}_h)$ and $(\boldsymbol{\mu}_l, \boldsymbol{\Sigma}_l)$ of the normal distributions $\mathcal{N}_h(\boldsymbol{\mu}_h, \boldsymbol{\Sigma}_h)$ and $\mathcal{N}_l(\boldsymbol{\mu}_l, \boldsymbol{\Sigma}_l)$.

Let us now recall that the positive superclass aggregates samples with high and intermediate fluorescence intensity (section 1). Since $\mathcal{N}_h(\boldsymbol{\mu}_h, \boldsymbol{\Sigma}_h)$ is computed over positive controls, we would take into account the characteristics of the images exhibiting an intermediate fluorescence intensity. To this aim, we add a third component to the GM model. Rather than using real images manually labeled as intermediate, which should introduce the disagreements described in section 2.1, the parameters of this GM component are synthetically derived as follows:

$$\boldsymbol{\mu}_i = w_1\boldsymbol{\mu}_l + (1 - w_1)\boldsymbol{\mu}_h, \quad \boldsymbol{\Sigma}_i = w_2\boldsymbol{\Sigma}_l + (1 - w_2)\boldsymbol{\Sigma}_h, \tag{2}$$

where $(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)$ are the parameters of intermediate (i) component and (w_1, w_2) are weights determining how much these parameters are similar to those computed from patches extracted from positive and negative controls.

In order to set the value of weights (w_1, w_2) we consider that misclassifications have dramatic differences when screening HEp-2 images for positive/negative classifications. Indeed, when a false positive classification occurs, the patient is subject to further investigations, increasing both his/her stress and national health service expense. Conversely, a false negative classification results in a lack of treatment for the patient with regrettable consequences. On these motivations, we would like to minimize false negative misclassifications, even if this could introduce more false positive. For this reason we introduce the loss matrix L, where the (k, j) entry L_{kj} corresponds to the cost of assigning a sample to the class j, when its true class is k:

$$L = \begin{bmatrix} 0 & C_{\text{FP}} \\ C_{\text{FN}} & 0 \end{bmatrix}. \tag{3}$$

The loss matrix L says that a correct decision has no cost, whereas: (i) there is a loss of C_{FP} in case of false positive classification, (ii) there is a loss of C_{FN} in case of false negative classification. Afterwards, we find the optimal values of (w_1, w_2) by minimizing the following loss function:

$$\mathscr{L}(w_1, w_2) = \sum_{k=1}^{2} \sum_{j=1}^{2} L_{kj} \cdot CM_{kj}^{(w_1,w_2)}, \tag{4}$$

where $CM_{kj}^{(w_1,w_2)}$ denotes the confusion matrix estimated on a validation set using the pairs (w_1, w_2).

Image Classification. In order to label the image as belonging to the positive or the negative class, we match patches extracted from the sample image at hand (s) with each of the three components of the GM model. To this aim, we assume that patches extracted from s are randomly drawn from a normal distribution $\mathcal{N}_s(\boldsymbol{\mu}_s, \boldsymbol{\Sigma}_s)$. Then we measure the distance between s and the j-th GM component using the Kullbach-Liebler divergence:

$$D_{KL}(\mathcal{N}_s, \mathcal{N}_j) = \frac{1}{2}\left(\text{tr}\left(\boldsymbol{\Sigma}_j^{-1}\boldsymbol{\Sigma}_s\right) + (\boldsymbol{\mu}_j - \boldsymbol{\mu}_s)^\top \boldsymbol{\Sigma}_j^{-1}(\boldsymbol{\mu}_j - \boldsymbol{\mu}_s) - k - \ln\left(\frac{\det\boldsymbol{\Sigma}_s}{\det\boldsymbol{\Sigma}_j}\right)\right). \tag{5}$$

The intensity assigned to the image will be the one whose component in the GM returns the minimum value of D_{KL}:

$$O(I) = \arg\min_j D_{KL}(\mathcal{N}_s, \mathcal{N}_j), \text{ with } j = h, l, i \tag{6}$$

where $O(I)$ is the output intensity assigned to the image I; h, l, and i represents the high, low and intermediate intensity, respectively. Finally, we label as positive the images labeled either with high or intermediate intensity, whereas we label as negative the images with a low fluorescence intensity.

Well Classification. Well classification is the final goal of our system since the well contains the serum of the patient. To combine the labels of images belonging to the same well, we use a conservative strategy which assigns the well to the positive class if at least one of its images is classified as positive. Otherwise, the well is labeled as negative. Formally:

$$O(W) = \begin{cases} p & \text{if } n_p > 0 \\ n & \text{otherwise} \end{cases} \tag{7}$$

where $O(W)$ is the output class assigned to the well W; p and n represent the positive and negative class, respectively; n_p is the number of images extracted from the well W labeled as positive.

Before detailing the dataset and the results, let us summarize the novelties introduced by this approach: (i) it extracts features by automatically detecting the parts of interest

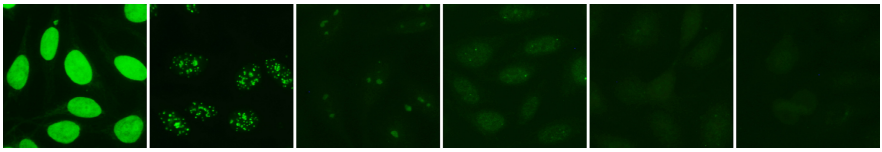

Fig. 2. HEp-2 example images. From left to right, we reported two example images for each of the three intensities (high, intermediate and low).

in the image, thus avoiding the burden of image segmentation; (ii) it does not need for a training set since it uses only the positive and negative controls to build the intensity model; (iii) the ground truth does not need to be manually labeled, because the positive and negative controls are prepared using already known sera; (iv) it introduces a novel representation of the intermediate intensity that is based on features extracted from the controls.

3 Dataset

We used a database of 914 HEp-2 labeled images acquired from 304 different wells at 1:80 titer, as recommended by the guidelines [2]. The wells were collected by consecutive patient sera, so that sample distribution is representative of the actual occurrence in the clinical routine. Two specialists took HEp-2 images with an acquisition unit consisting of a led fluorescence microscope coupled with a digital camera with CCD squared pixel of equal side to 6.45 μm. The images were independently labeled by the two physicians and they reached consensus on the cases where they disagreed. The samples are distributed among 83, 178 and 43 positive, negative and intermediate wells and, therefore, the a priori distribution of the wells is 27.3%, 58.6% and 14.1% for high, low and intermediate intensities, respectively. In our recognition task, this corresponds to 41.4% and 58.6% of samples belonging to positive and negative classes, respectively. The GM model is built over 24 images of controls, twelve for positive and twelve for negative ones. Fig. 2 shows some examples of Hep-2 images. Note that images with a high and an intermediate intensity vary also their staining pattern, resulting in different textures. This increases the intra-class variability and makes harder the classification task.

4 Results

This section first discusses the values of parameters of the system, and then presents the classification results.

According to the workflow depicted in Fig. 1, a first parameter of our approach when using PCA for feature reduction is the number of principal components to be used. We evaluated the number of components for which the normal distributions of positive and negative controls do not overlap, considering the value of 5σ as the cut-off point for the normal distributions. In this respect, we found that already the first component, which explains 61.1% of the overall variance, verifies this property. We therefore deem that

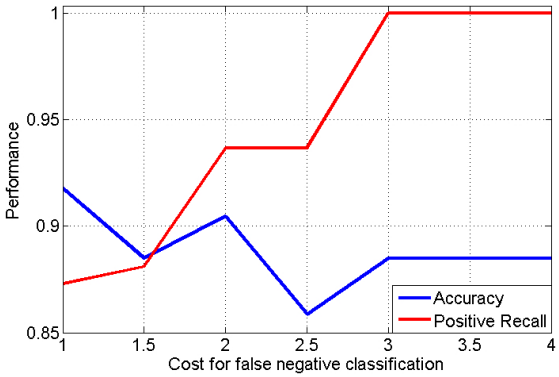

Fig. 3. Influence of C_{FN} value

it has a satisfactory discriminative power without complicating the gaussian model and thus permitting to better estimate the GM parameters. It is worth noting that we also ran the experiments adding components until they explained the 95% of the variance, achieving lower performance. In our opinion, although accounting for more variance of the data, a larger number of components increases the dimension of the feature space providing a less robust GM model. We also explored the performance achievable projecting the feature space over the single dimension obtained by the LDA. However, the performance achieved with the first component of PCA turned out to be better and we then report in the following the results achieved using only the first principal component.

Let us now turn the attention to the weights used to set the values of parameters of Gaussian distribution $\mathcal{N}_i(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)$. As reported in section 2.2, the pair (w_1, w_2) determines how much the distribution representing the intermediate fluorescence intensity is similar to the high or low intensity distributions given by positive and negative controls. The value of such pair is set minimizing a loss function taking into account the different costs of false positive (C_{FP}) and false negative (C_{FN}) classifications. For each test well, this optimization is composed of the following four steps. First, build a validation set using all images of other wells. Second, perform the whole classification process on such a validation set with different values of the pair (w_1, w_2). Since in IIF medical context C_{FN} is larger than C_{FP}, we set the value of C_{FP} to 1 and we vary the value of C_{FN} in [1;10] with step of 0.5. Third, find the pair (w_1, w_2) corresponding to the minimum of the loss function $\mathscr{L}(w_1, w_2)$. Fourth, classify the test well applying the approach described in section 2 using the weight values set in previous step. In order to deepen the influence of cost coefficients over the performance on the whole test set, Fig. 3 shows how performance vary as C_{FN} varies. It is worth observing that all positive samples are correctly classified when $C_{\text{FN}} \geq 3$, i.e. when the cost of a false negative classification is three times the cost of a false positive classification. Obviously, a larger accuracy is achieved when the two costs are equal. Due to the medical constrains which impose to minimize the number of false negative classification in the screening phase, in the following we report the results when the values of C_{FP} and C_{FN} are set to 1 and 3, respectively.

Table 1. Confusion matrices of positive/negative classifications for image (a) and well (b) recognition

(a)

		Hypothesized Class	
		Positive	*Negative*
True Class	*Positive*	373	7
	Negative	89	445

(b)

		Hypothesized Class	
		Positive	*Negative*
True Class	*Positive*	126	0
	Negative	35	143

The first phase of the classification consists in assigning a class to the images using matching algorithm. In this respect, Table 1a shows the confusion matrix at image classification level. Since the different images of a well will be combined to get the final label of the well, this is not the final output of our system. However, we can already observe that the 7 false negative outputs are due to misclassifications between intermediate and low fluorescence intensity. This is an important result since it prevents that a well with a strong positiveness might be classified as negative.

Once we get the intensity of the images, we use the combination rule reported in equation 7 to classify the wells. The result of this procedure is shown in Table 1b. We note that no positive well was classified as negative, i.e. the recall is 100%, even if this inevitably implies a reduced recall on the negative class (80.34%). However, this is not a relevant drawback since achieving a low false negative rate is the essential requirement of the medical domain. In conclusion, we deem that the performance of our system are suited to perform a pre-selection of the cases that have to be examined, enabling the physician to focus his/her attention only on relevant cases.

5 Conclusions

In this paper we have presented a slightly supervised classification approach for positive/negative fluorescence intensity classification. It permits to overcome the need of collecting a large and reliable ground truth, which is expensive, time consuming and intrinsically hard. Furthermore, the approach should be adapted to working scenario trough a cost matrix defining the cost of misclassifications. Our proposal can be used to perform a pre-selection of the cases that have to be examined, enabling the physician to focus his/her attention only on relevant cases, making also easier to carry out mass screening campaigns. Future works are directed towards the comparison of this approach with other proposal reported in the literature, and towards the test in daily routine.

Acknowledgements. This work has been carried out in the framework of the ITINERIS2 project, Codice CUP F87G10000080009, under the financial support of Regione Lazio (Programme 'Sviluppo dell'Innovazione Tecnologica nel Territorio Regionale', Art. 182, comma 4, lettera c), L.R. n. 4, 28 Aprile 2006).

References

1. Bizzaro, N., Tozzoli, R., et al.: Variability between methods to determine ana, anti-dsdna and anti-ena autoantibodies: a collaborative study with the biomedical industry. J. Immunol. Methods 219(1-2), 99–107 (1998)
2. 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)
3. Ersoy, I., Bunyak, F., Peng, J., Palaniappan, K.: HEp-2 cell classification in IIF images using shareboost. In: 2012 21st International Conference on Pattern Recognition (ICPR), pp. 3362–3365. IEEE (2012)
4. Hiemann, R., Hilger, N., et al.: Automatic analysis of immunofluorescence patterns of HEp-2 cells. Annals of the New York Academy of Sciences 1109(1), 358–371 (2007)
5. Hiemann, R., et al.: Challenges of automated screening and differentiation of non-organ specific autoantibodies on HEp-2 cells. Autoimmunity Reviews 9(1), 17–22 (2009)
6. Huang, Y.L., Jao, Y.L.: et al.: Adaptive automatic segmentation of HEp-2 cells in indirect immunofluorescence images. In: IEEE Int. Conf. on SUTC, pp. 418–422 (2008)
7. Huang, Y.L., et al.: Outline detection for the HEp-2 cell in indirect immunofluorescence images using watershed segmentation. In: IEEE Int. Conf. on SUTC, pp. 423–427 (2008)
8. Lowe, D.G.: An iterative image registration technique with an application to stereo vision. Int. J. Comput. Vision 60(2), 91–110 (2004)
9. Percannella, G., Soda, P., Vento, M.: A classification-based approach to segment HEp-2 cells. In: 25th IEEE Symp. on Computer-Based Medical Systems, pp. 1–5. IEEE (2012)
10. Perner, P., Perner, H., Muller, B.: Mining knowledge for HEp-2 cell image classification. Artif. Intell. Med. 26(1-2), 161–173 (2002)
11. Rigon, A., Soda, P., et al.: Indirect immunofluorescence in autoimmune diseases: Assessment of digital images for diagnostic purpose. Cytometry Part B: Clin. Cytometry 72B(6), 472–477 (2007)
12. Sack, U., Knoechner, S., et al.: Computer-assisted classification of HEp-2 immunofluorescence patterns in autoimmune diagnostics. Autoimmunity Reviews 2(5), 298–304 (2003)
13. Soda, P., Iannello, G., Vento, M.: A multiple experts system for classifying fluorescence intensity in antinuclear autoantibodies analysis. Pattern Anal. Appl. 12(3), 215–226 (2009)
14. Soda, P., Onofri, L., Iannello, G.: A decision support system for Crithidia Luciliae image classification. Artificial Intelligence in Medicine 51(1), 67–74 (2011)
15. Strandmark, P., Ulén, J., Kahl, F.: HEp-2 staining pattern classification. In: 2012 21st International Conference on Pattern Recognition (ICPR), pp. 33–36. IEEE (2012)
16. Theodorakopoulos, I., Kastaniotis, D., et al.: HEp-2 cells classification via fusion of morphological and textural features. In: 2012 IEEE 12th International Conference on Bioinformatics & Bioengineering (BIBE), pp. 689–694. IEEE (2012)