Analysis and Classification of Crithidia Luciliae Fluorescent Images

Paolo Soda[1], Leonardo Onofri[1], Amelia Rigon[2], and Giulio Iannello[1]

[1] University Campus Bio-Medico of Rome, Integrated Research Centre, Medical Informatics & Computer Science Laboratory, Rome, Italy
[2] University Campus Bio-Medico of Rome, Integrated Research Centre, Immunology, Rome, Italy
{p.soda,leonardo.onofri,a.rigon,g.iannello}@unicampus.it

Abstract. Autoantibody tests based on Crithidia Luciliae (CL) substrate are the recommended method to detect Systemic Lupus Erythematosus (SLE), a very serious sickness further to be classified as an invalidating chronic disease. CL is an unicellular organism containing a strongly tangled mass of circular dsDNA, named as kinetoplast, whose fluorescence determines the positiveness to the test. Conversely, the staining of other parts of cell body is not a disease marker, thus representing false positive fluorescence. Such readings are subjected to several issues limiting the reproducibility and reliability of the method, as the photo-bleaching effect and the inter-observer variability. Hence, Computer-Aided Diagnosis (CAD) tools can support physicians decision. In this paper we propose a system to classify CL wells based on a three stages recognition approach, which classify single cell, images and, finally, the well. The fusion of such different information permits to reduce the misclassifications effect. The approach has been successfully tested on an annotated dataset, proving its feasibility.

1 Introduction

Systemic Lupus Erythematosus (SLE) is a chronic inflammatory disease of unknown aetiology affecting multiple organ systems. In Europe, SLE incidence and prevalence every 100000 inhabitants ranges between $1.9-4.8$ cases and $12.5-68$ cases, respectively [1]. Recent study has reported that 5 and 10 years survival ranges between $91\%-97\%$ and $83\%-92\%$ [2], respectively. Although SLE should be tagged as a rare illness due to its incidence, it is considered a very serious sickness further to be classified as an invalidating chronic disease.

The numerous and different manifestations of SLE make its diagnosis a burdensome task. The criteria of American College of Rheumatology require to perform autoantibodies tests for SLE diagnosis [3]. Indeed, it suggests to detect the presence of autoantibodies directed against double strand DNA (anti-dsDNA) due to the high specificity of the test (92%).

The recommended method to detect autoantibodies disorders is the Indirect ImmunoFluorescence (IIF) assay. IIF makes use of a substrate containing a specific antigen that can bond with serum antibodies, forming a molecular complex.

P. Foggia, C. Sansone, and M. Vento (Eds.): ICIAP 2009, LNCS 5716, pp. 558–566, 2009.

Then, this complex reacts with human immunoglobulin conjugated with a fluorochrome, making the complex observable at the fluorescence microscope. In order to detect SLE autoantibodies, IIF is performed using the Crithidia Luciliae (CL) substrate [4,5], since it includes a compressed mass of double strand DNA (dsDNA) named as *kinetoplast*.

However, the readings in IIF are subjected to several issues that limits the reproducibility and reliability of the method, such as the photo-bleaching effect and the inter-observer variability [6,7]. To date, the highest level of automation in anti-dsDNA tests is the preparation of slides with robotic devices performing dilution, dispensation and washing operations [8,9]. Although such instruments helps in speeding up the slide preparation step, the development of Computer-Aided-Diagnosis (CAD) systems supporting IIF diagnostic procedure would be beneficial in many respects.

To our knowledge, only recent papers dealing with CAD tools for autoantibodies tests based on the HEp-2 substrate[1] can be found in the literature [10,11,12,13,14]. In this paper we focus on the development of a decision system able to classify anti-dsDNA images based on the Crithidia Luciliae substrate. The system is usable in practice, capable of performing a pre-selection of the cases to be examined and serving as a second reader. The system has been successfully tested on an annotated dataset.

2 Background and Motivations

Current guidelines for SLE diagnosis recommend to use CL substrate diluted at 1:10 titer, prepared according to IIF method [5].

CL is an unicellular organism containing both the nucleus and the kinetoplast. The latter is composed of a strongly tangled mass of circular dsDNA, making this test highly specific. Hence, the fluorescence of kinetoplast is the fundamental parameter to define the positiveness of a well, whereas the fluorescence of other parts of CL cells, e.g. the basal body or the flagellum, is not a marker of anti-dsDNA autoantibodies and, consequently, of SLE. Fig. 1 shows five stylized cells representative of different cases.

Panels A and B depict positive cases. The former shows a cell where only the kinetoplast exhibits a fluorescence staining higher than cell body, while in the latter also the nucleus is highly fluorescent. The other three panels show negative cells. In panel C, the cell is clearly negative since all cell body has a weak and quite uniform fluorescence. Panels D and E depict cells where regions different from the kinetoplast exhibit strong fluorescence staining. Indeed, in panel D the basal body, which is similar to kinetoplast in size and type of fluorescence staining, is lighter than cell body. In panel E, one or more parts different from the kinetoplast has a strong staining. Finally, notice also that occasionally fluorescence objects, i.e. artifacts, may be observed outside cell body.

[1] IIF tests performed using HEp-2 substrate permits to detect antinuclear autoantibodies (ANA).

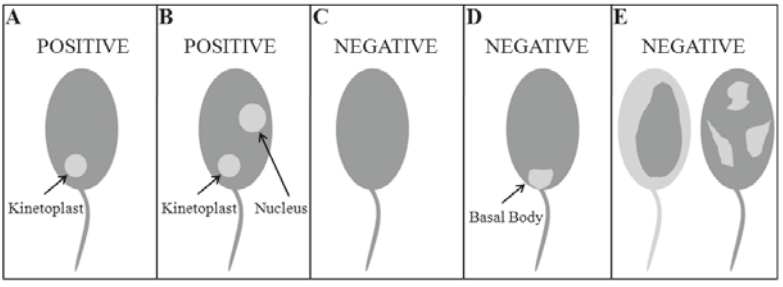

Fig. 1. Positive and negative cases depicted through stylized cells. Dark and light grey represent low and high fluorescence, respectively.

These observations suggest that several and different reasons of uncertainty can be observed in CL images, making the right determination of kinetoplast staining a demanding task. Such motivations are at the basis of CL tests inter-observer variability. Other reasons are the photo-bleaching effect that bleaches significantly the cells over a short period of time [7], as well as the lack of quantitative information supplied to physicians.

Developments in computer vision and artificial intelligence in medical image interpretation have shown that CAD systems can effectively support the specialists in various medical fields. In particular, in the area of autoantibodies testing some recent papers have presented CAD tools for the classification of HEp-2 images [10,11,12,13,14], which are used to detect other autoimmune diseases. To our knowledge no works in the literature propose solutions for automatic classification of CL slides. This paper therefore discusses a solution in this direction.

3 Methods

We aim at recognizing the positiveness or negativeness of a well, since it contains the serum of one patient. Due to microscope magnification typically used during the acquisition process, one digital image does not cover all the surface of one well. For this reason, several images are always collected. This feature permits us to exploit a certain degree of redundancy, integrating together the information extracted from different images of the same well. To recognize the staining of each well we initially classify the images acquired from the well under examination, and then label the well on the strength of the classification of its images (panel A of Fig. 2).

Panel B of Fig. 2 details the steps applied to label single images. The first one relies upon the observation that the kinetoplast fluorescence implies the presence in the image of a compact set of pixels brighter than other regions. Conversely, the absence of such a set suggests that the image is negative. For instance, the cell shown in panel C of Fig. 1 is classified as negative while the cells in the other panels contain regions candidate to be a kinetoplast and proceed in the next classification steps.

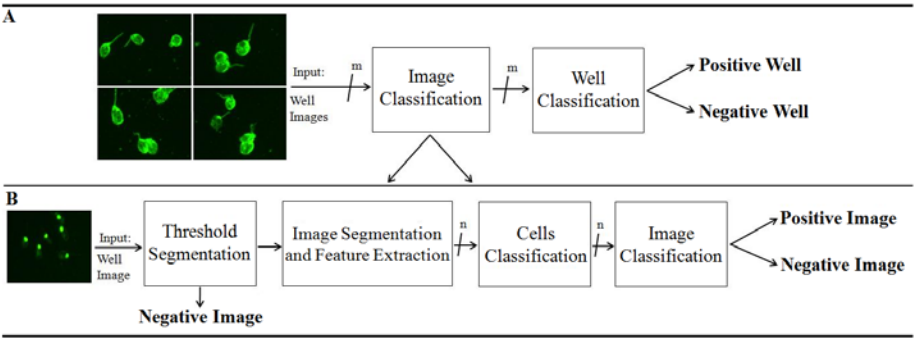

Fig. 2. Description of the classification approach

As a second step, we are interested in locating and classifying single cells. Hence, we segment only the images containing at least one set of pixels candidate to be a kinetoplast and, then, the segmented cells are classified. Finally, the staining of the image is computed on the basis of the labels of all cells.

In our opinion, the overall approach provides two benefits. First, the initial threshold segmentation allows a rapid classification of several images ($\approx 0.1s$ per image). Second, it is tolerant with respect to misclassifications in cells recognition: if enough cells per image are available, it is reasonable that misclassified cells, if limited, does not affect image classification. This permits to lower the effect of both erroneously segmented cells or artifacts. Notice also that a similar consideration holds also for well classification, where multiple images are available.

Following paragraphs focus on cell, image and well classification blocks, respectively.

Cell classification. The cells have been segmented using Otsu's algorithm [15]. Then, morphological operations, such as filling and connection analysis, output a binary mask for cutting out the cells from the image. Cells connected with the image border have been suppressed. From each segmented cell we extract a set of features belonging to different categories: (i) intensity histogram and co-occurence matrix, (ii) angular bin of Fourier Transform (FT) spectrum, (iii) radial bin of FT spectrum, (iv) wavelet transform, (v) Zernike moments, (vi) rectangle features, (vii) Local Binary Pattern both standard and circular as well as circular rotation invariant [16], (viii) morphological descriptors. The interested readers may refer to [17] for further details. Then, discriminant analysis permits us to identify the subset of most discriminant features.

Image classification. Each image contains several cells whose kinetoplasts can be positive or negative. According to previous considerations, as a first step we are interested in identifying clearly negative images, i.e. images that do not contain any compact fluorescent set of pixels. To this aim, we perform a threshold-based

segmentation that looks for connected regions satisfying conditions on both intensity and dimension.

When the image contains regions candidate to be a kinetoplast, we have to determine if it is a true positive image. For istance, false positive images occur when there is a fluorescent mass inside cell body, as shown in panel D and E of Fig. 1. To this end, we classify single cells as reported above and then, on their basis, we apply the Majority Voting (MV) rule which assigns to the image the label corresponding to the class receiving the relative majority of votes. Ties lead to reject the image.

Well Classification. As reported at the beginning of the section, the class of the whole well is determined on the basis of the classification of its images. To this aim, we apply again the MV rule where the votes summation is performed at image level. In case of tie the system does not cast a decision on the well asking for specialist classification. Indeed, due to the high specificity of CL tests it is preferable to reject the well rather than risking a misclassification.

4 Data Set

Since, to our knowledge, there are not reference databases of CL images publicly available, we acquired and annotated 342 images, 74 positive (21.6%) and 268 negative (78.4%), belonging to 63 wells, 15 positive (23.8%) and 48 negative (76.2%). Furthermore, specialists label a set of cells since our approach requires their labels to train the corresponding classifier. This operation has been performed only on images where the threshold-based segmentation identifies fluorescent regions. At the end of such a procedure, the cells data set consists of 1487 cells belonging to 34 different wells. The cells are therefore labelled: 928 as positive (62.4%) and 559 as negative (37.6%).

5 Experimental Evaluation

In order to classify CL cells, we investigate the performance that could be achieved appealing to popular classifiers belonging to different paradigms, as reported below.

As a neural network, we use a Multi-Layer Perceptron (MLP) with one hidden layer. The number of neurons in input and output layers are equal to features number and two (i.e. we deal with a two-classes recognition task), respectively. Preliminary tests have permitted to determine the best configuration of the MLP in terms of neurons number in the hidden layer: specifically, configurations from 1 up to 50 neurons were tested.

As a bayesian classifier, we apply the Naïve Bayes classifier, which does not demand for specific set-up.

Support Vector Machine (SVM) is adopted as a kernel machine, using a Radial Basis Function kernel. We have performed a grid search to determine the best

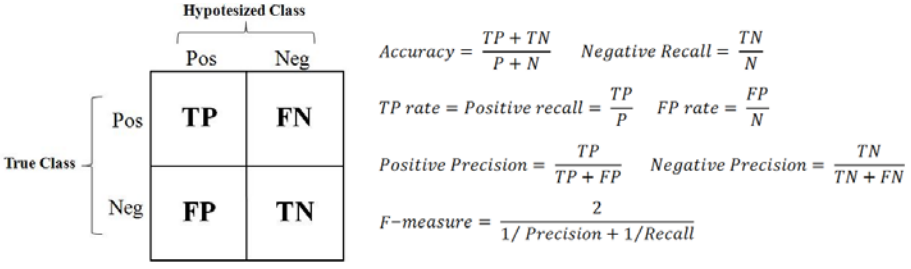

Fig. 3. Confusion matrix and derived performance metrics calculated for the two-classes recognition task. P and N represent the total number of positive and negative samples, respectively.

setup of the SVM in terms of Gaussian width, σ, and of regularisation parameter, C in the ranges $[0.01; 30]$ and $[0.1; 60]$, respectively.

AdaBoost is tested as ensemble of classifier, using decision stumps as base hypotheses and exploring different numbers of iterations (in the interval $[10; 100]$) to determine the best configuration.

In order to estimate the performance we measure the following common metrics: accuracy, recall, precision and f-measure, whose definition is reported in Fig. 3 together with the confusion matrix of a two-classes recognition problem. We measure also the area under ROC curve (AUC), since it permits to reduce the ROC performance to a single scalar value in $[0, 1]$: the more it approaches one, the better the performance [18].

5.1 Results and Discussion

In this subsection we present the results achieved in the classification of individual cells, images and whole wells.

Cell Classification. Table 1 reports the performance measures of the four classification paradigms tested in the recognition of positive and negative CL cells, estimated using 10-fold cross validation and randomly choosing training and test set elements.

These data point out that SVM classifier outperforms the others. They also show that accuracy is balanced over the two classes. This finding confirms previous papers reporting that SVMs have very good performance on binary classification task, since they have been originally defined for this type of problem [19].

Although these results are interesting, they do not accurately estimate the performance of the classifiers when the samples are outside the training set. Indeed, random cross validation does not take into account that cells belonging to the same well, and thus with not independent features, can be both in training and test sets. To overcome such a limitation, we divided the set of 1487 cells into 34 subsets, one for each well. We then perform a 34-folds cross validation using only SVMs, since they are the most performing classifier. Accuracy, negative

Table 1. Performance of the classifier recognizing positive and negative CL cells, estimated using 10-fold cross validation. In parenthesis we report the best configuration of each classifier: in case of MLP the number represents the hidden layer neurons, in case of SVMs the values are C and σ, and in case of Adaboost it is the iterations number.

		Recall(%)		Precision(%)		F-measure(%)		
Classifier	**Accuracy(%)**	**Pos**	**Neg**	**Pos**	**Neg**	**Pos**	**Neg**	**AUC**
Naive Bayes	89.4	88.3	91.2	94.4	82.4	91.2	86.6	0.938
MLP (40)	95.7	97.1	93.4	96.1	95.1	96.6	94.2	0.989
AdaBoost (50)	96.4	97.1	95.2	97.1	95.2	97.1	95.2	0.993
SVMs (17,1)	97.0	97.8	95.7	97.4	96.4	97.6	96.1	0.993

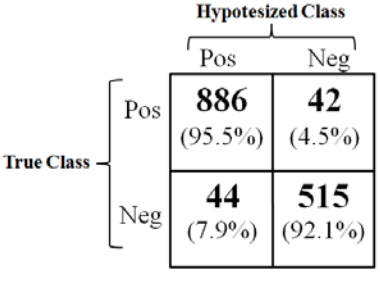

Fig. 4. Confusion Matrix of cells classification using SVM through 34-folds cross validation

and positive recall, negative and positive precision, negative and positive F-measure and AUC are now 94.2%, 92.1%, 95.5%, 92.5%, 95.3%, 92.3%, 95.4%, 0.984, respectively. Notice that the variation of such performance will not affect the results overall system results since the final classification depends on the classification of several cells.

Image Classification. The tests show that threshold-based segmentation does not exhibits FN, while 64.6% of images are correctly classified as negative (panel A of Fig. 5). The remaining 35.4% of images are then labelled having recourse to cell labels provided by 34-folds classification.

Panel B of Fig. 5 reports the confusion matrix achieved in the classification of such images, showing that only two of them have been misclassified. In conclusion, 340 images (99.4%) are correctly classified, attaining a positive and negative recalls of 98.6% and 99.6%, respectively.

Well Classification. With regard to well classification, we correctly classify 62 wells (98.4%) whereas only one well has been rejected since the same number of images have been labelled as positive and negative.

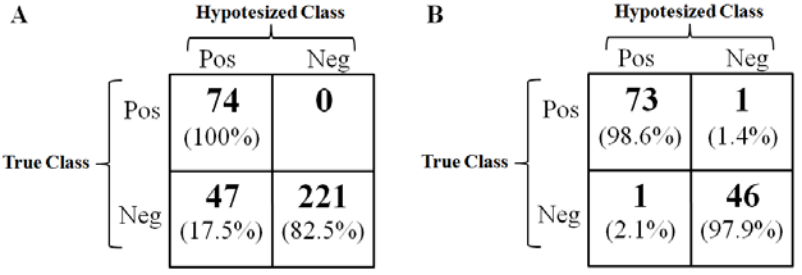

Fig. 5. Confusion Matrix of images classification. Panel A shows the result of threshold-based segmentation. Panel B shows the recognition result of the remaining images achieved using individual cells labels.

6 Conclusion

In this paper we have presented a system that can support the specialists in SLE diagnosis. As a novel contribution, this paper aims aims at classifying Chritidia Luciliae wells, since such a substrate is one of the most specific test to detect the considered disease. The proposed system is based on a three stages recognition approach, namely cells, images and well classification. It integrates several pieces of information providing a degree of redundancy that lowers the effect of misclassifications and permits us to achieve promising performance, even if improving tie breaking is the current challenge to make the system flexible to work in different operating scenario.We are currently engaged in populating a larger annotated database and testing the system in real-world tests to validate its use. The ultimate goal is the development of a CAD system providing the chance to automatically classify a panel of autoantibodies.

Acknowledgment

The authors thank Dario Malosti for his precious advices. This work has been partially supported by "Thought in Action" (TACT) project, part of the European Union NEST-Adventure Program, and by DAS s.r.l. of Palombara Sabina (`www.dasitaly.com`).

References

1. Lee, P.P.W., Lee, T.L., Ho, M.H.K., Wong, W.H.S., Lau, Y.L.: Recurrent major infections in juvenile-onset systemic lupus erythematosus–a close link with long-term disease damage. Rheumatology 46, 1290–1296 (2007)
2. Mok, C.C., Lee, K.W., Ho, C.T.K., et al.: A prospective study of survival and prognostic indicators of systemic lupus erythematosus in a southern chinese population. Rheumatology 39, 399–406 (2000)

3. Hochberg, M.C.: Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus. Arthritis and Rheumatism 40(9), 1725 (1997)
4. Piazza, A., Manoni, F., Ghirardello, A., Bassetti, D., Villalta, D., Pradella, M., Rizzotti, P.: Variability between methods to determine ANA, anti-dsDNA and anti-ENA autoantibodies: a collaborative study with the biomedical industry. Journal of Immunological methods 219, 99–107 (1998)
5. Tozzoli, R., Bizzaro, N., Tonutti, E., Villalta, D., Bassetti, D., Manoni, F., Piazza, A., Pradella, M., Rizzotti, P.: Guidelines for the laboratory use of autoantibody tests in the diagnosis and monitoring of autoimmune rheumatic diseases. American Journal of Clinical Pathology 117(2), 316–324 (2002)
6. Song, L., Hennink, E.J., Young, I.T., Tanke, H.J.: Photobleaching kinetics of fluorescein in quantitative fluorescence microscopy. Biophysical Journal 68(6), 2588–2600 (1995)
7. Rigon, A., Soda, P., Zennaro, D., Iannello, G., Afeltra, A.: Indirect immunofluorescence in autoimmune diseases: assessment of digital images for diagnostic purpose. Cytometry 72(6), 472–477 (2007)
8. Das s.r.l.: Service Manual AP16 IF Plus. Palombara Sabina (RI) (March 2004)
9. Bio-Rad Laboratories Inc.: PhD System (2004), `http://www.bio-rad.com`
10. Perner, P., Perner, H., Muller, B.: Mining knowledge for HEp-2 cell image classification. Artificial Intelligence in Medicine 26(1-2), 161–173 (2002)
11. Sack, U., Knoechner, S., Warschkau, H., Pigla, U., Emmerich, F., Kamprad, M.: Computer-assisted classification of HEp-2 immunofluorescence patterns in autoimmune diagnostics. Autoimmunity Reviews 2(5), 298–304 (2003)
12. Hiemann, R., Hilger, N., Michel, J., Anderer, U., Weigert, M., Sack, U.: Principles, methods and algorithms for automatic analysis of immunofluorescence patterns on HEp-2 cells. Autoimmunity Reviews, 86 (2006)
13. Soda, P., Iannello, G., Vento, M.: A multiple experts system for classifying fluorescence intensity in antinuclear autoantibodies analysis. Pattern Analysis & Applications (2008)
14. Soda, P., Iannello, G.: Aggregation of classifiers for staining pattern recognition in antinuclear autoantibodies analysis. IEEE Transactions on Information Technology in Biomedicine 13(3), 322–329 (2009)
15. Otsu, N.: A threshold selection method from gray-level histograms. IEEE Transactions on Systems, Man, and Cybernetics 9(1), 62–66 (1970)
16. Ojala, T., Pietikäinen, M., Mäenpää, T.: Multiresolution gray-scale and rotation invariant texture classification with local binary pattern. IEEE Transactions on Pattern Analysis and Machine Intelligence 24(7), 971–987 (2002)
17. Onofri, L., Soda, P.: Feature extraction and selection for Chritidia Luciliae discrimination. Technical report, Università Campus Bio-Medico di Roma (December 2008)
18. Fawcett, T.: Roc graphs: Notes and practical considerations for researchers. HP Laboratories (2004)
19. Vapnik, V.N.: The nature of statistical learning theory. Springer, Heidelberg (1995)