POLITECNICO DI TORINO
Repository ISTITUZIONALE

ANAlyte: a modular image analysis tool for ANA testing with Indirect Immunofluorescence

Availability:
This version is available at: 11583/2636782 since: 2020-02-12T12:15:08Z

Terms of use:

Publisher copyright

(Article begins on next page)

04 May 2024

ANAlyte: a modular image analysis tool for ANA testing with Indirect Immunofluorescence.

Santa Di Cataldo, Simone Tonti, Andrea Bottino, Elisa Ficarra

Dept. of Computer and Control Engineering, Politecnico di Torino, Cso Duca degli Abruzzi 24, 10129 Torino, Italy

Abstract

Background and objectives. The automated analysis of Indirect Immunofluorescence images for Anti-Nuclear Autoantibody (ANA) testing is a fairly recent field that is receiving ever-growing interest from the research community. ANA testing leverages on the categorization of intensity level and fluorescent pattern of IIF images of HEp-2 cells to perform a differential diagnosis of important autoimmune diseases. Nevertheless, it suffers from tremendous lack of repeatability due to subjectivity in the visual interpretation of the images. The automatization of the analysis is seen as the only valid solution to this problem. Several works in literature address individual steps of the work-flow, nonetheless integrating such steps and assessing their effectiveness as a whole is still an open challenge.
Methods. We present a modular tool, ANAlyte, able to characterize a IIF image in terms of fluorescent intensity level and fluorescent pattern without any user-interactions. For this purpose, ANAlyte integrates the following: (i) intensity classifier module, that categorizes the intensity level of the input slide based on multi-scale contrast assessment (ii) cell segmenter module, that splits the input slide into individual HEp-2 cells; (iii) pattern classifier module, that determines the fluorescent pattern of the slide based on the pattern of the individual cells.
Results. To demonstrate the accuracy and robustness of our tool, we experimentally validated ANAlyte on two different public benchmarks of IIF HEp-2 images with rigorous leave-one-out cross-validation strategy. We obtained overall accuracy of fluorescent intensity and pattern classification respectively around 85% and above 90%. We assessed all results by comparisons with some of the most representative state of the art works.
Conclusions. Unlike most of the other works in the recent literature, ANAlyte aims at the automatization of all the major steps of ANA image analysis. Results on public benchmarks demonstrate that the tool can characterize HEp-2 slides in terms of intensity and fluorescent pattern with accuracy better or comparable with the state of the art techniques, even when such techniques are run on manually segmented cells. Hence, ANAlyte can be proposed as a valid solution to the problem of ANA testing automatization.

Keywords: Automated characterization of HEp-2 slides, Indirect immunofluorescence, ANA testing, Microscopy image processing

1. Introduction

The analysis of indirect immunofluorescent (IIF) images is of paramount importance to a large number of clinical applications. For example, the Antinuclear Autoantibody (ANA) test leverages on the analysis of IIF images to diagnose autoimmune disorders falling into the category of so-called connective tissue diseases (CTDs), which affect a remarkable percentage of the population. Among the others, rheumatoid arthritis, psoriatic arthritis and scleroderma.

CTDs are characterized by a spontaneous overactivity of the immune system, which will introduce extra autoantibodies into the circulatory system of the patient. Thus,

the presence of such autoantibodies in the blood is a reliable indication of disease.

ANA testing is a widespread blood exam that uses a slide of cultured HEp-2 (Human epithelial type 2) cells as a substrate to reveal the presence of said antibodies in the serum of a patient. The HEp-2 cells have antigens that specifically bind the targeted autoantibodies which, in turn, create a link with secondary antibodies conjugated with fluorophores (see Figure 1). When visualized under the microscope, the antigen-antibody reaction will establish a particular fluorescent pattern on the HEp-2 cells, which is specific of the autoantibody type. Hence, identifying the pattern of the IIF image (see Figure 2 for examples) allows a differential diagnosis of the CTDs.

The analysis is commonly performed as follows. First, the physicians identify the positive specimens. This requires classifying the IIF images into a number of categories, based on the intensity of the fluorescent signal.

---

*Corresponding author: Santa Di Cataldo Tel.: +39-011-090-7020; fax: +39-011-090-7099;
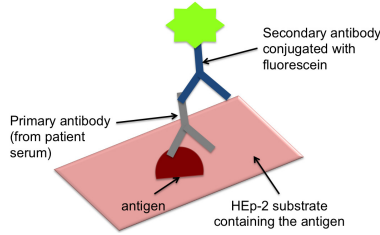   Email address:* santa.dicataldo@polito.it (Santa Di Cataldo)

Figure 1: Principle of Indirect Immunofluorescence using HEp-2 as a substrate.

Although there is no general consensus among the laboratories about the specific number of categories that should be identified, recent studies show that three intensity levels allow to obtain the highest discrimination capability and at the same time to minimize variability between operators. Such levels are negative (i.e. no fluorescence at all), intermediate and positive [1]. Second, the physicians categorize the intermediate and positive specimens based on the fluorescent pattern revealed on the HEp-2 cells (a few examples are reported in Figure 2). This ultimately allows to identify the type of CTD affecting the patient.

While ANA testing per se is widely spread and universally recognized as a valid diagnostic technique, the significance and reliability of its results are critically affected by the subjectivity of the human interpretation of the IIF slides [2, 3]. On top of that, employing highly skilled and trained physicians for the analysis of massive amounts of images often translates into unsustainable costs for the health-care system.

In the attempt to seek valid solutions to this problem, the automatization of IIF image analysis for ANA testing is now receiving an ever-growing attention. In the following, we provide a short overview of latest research directions in this field and we introduce the main contributions of our work.
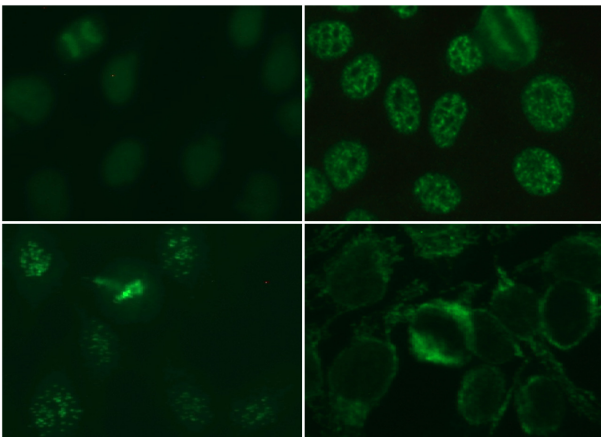

Figure 2: Examples of IIF images with different fluorescent patterns.

1.1. Related works and contributions of the paper

To be useful to a clinical setting, a computer-aided image analysis system for ANA testing should be able to automatically characterize a IIF specimen in terms of both fluorescent intensity and fluorescent pattern. Such characterization can be either performed on a per-image basis (leveraging on global features of the input image) or on a per-cell basis. In most of the available literature, fluorescent intensity categorization is obtained with a per-image approach, while fluorescent pattern classification is preferably performed with a per-cell approach. In the latter case, the system first applies cell segmentation techniques to split the IIF slide into individual HEp-2 cells, and then it determines the fluorescent pattern of each cell, which ultimately allows to establish the pattern of the whole slide. This is based on the assumption that the pattern of the specimen coincides with the pattern of the large majority of the HEp-2 cells in the slide, which is valid for most of the ANAs used in diagnosis.

In the last few years, literature has proposed several automated techniques tackling IIF analysis. Nevertheless, most of the publications focus on individual tasks of the ANA testing chain, overlooking the other steps of the analysis (e.g. HEp-2 cell segmentation [4, 5, 6], fluorescent intensity classification [1, 7, 8], HEp-2 pattern recognition [9, 10, 11, 12]). Hence, even though the proposed techniques show promising results per se, there is still little evidence on how effective they would be when they are integrated into a complete analysis work-flow and applied to a real clinical setting. On the other hand, first attempts of commercial systems for computer-aided IIF analysis are validated on private databases of HEp-2 images, which prevents a direct comparison with other techniques [13, 14, 15].

With the aim of sorting out the ever-growing activity of the research community in this field, since 2012 international contests on HEp-2 pattern classification have been regularly hosted by some of the major conferences on pattern recognition and image processing (respectively ICPR 2012, ICIP 2013 and ICPR 2014 [16, 17, 18]), as well as by a special issue of the journal Pattern Recognition in 2014 [11]. While classification paradigms presented in such contests are often well-established machine learning approaches (e.g. Support Vector Machines, boosting algorithms, neural networks and random forest techniques), most of the original contributions are in the design of suitable image attributes to characterise the HEp-2 patterns. Proposed descriptors include global statistics of grey-level distributions (e.g. grey-level co-occurrence matrices [19]), morphological measures of shape or topology [20], or combination of the two [21, 22], as well as various formulations of local binary patterns. In particular, a novel variant of local binary patterns, namely Rotation Invariant Co-occurrence among adjacent LBPs (RIC-LBPs), was proposed by the winner of ICPR 2012 contest [23]. Other top performing approaches of the same contest focused on classic textural descriptors incorporated with feature encoding

and dictionary learning, either bag of words (BoW, earlier proposed for HEp-2 classification by [9]), or specifically designed sparse representations of the cell images [24]. Feature encoding approaches were very popular also in the next contest hosted by ICPR 2014 (I3A), but with a more sophisticated characterisation of the HEp-2 patterns: for example, the biologically-inspired dense local descriptors proposed by [25], the combined textural and structural information extracted by [22], or the multi-scale descriptors with cell pyramids designed by the winner of the contest [26]. Conversely, another top-performing approach replaced the use of hand-crafted features with convolutional neural networks taking the HEp-2 cells directly as input [27].

Ultimately, the outcome of the IIF image analysis contests is two-fold. First, a very interesting survey and assessment of the recent research efforts towards automated ANA testing, albeit solely focused on the HEp-2 pattern classification task. Second, the release of public benchmarks of IIF images, on which the automated techniques can now be tested and compared.

In our previous works we addressed two major steps of automated IIF analysis independently from each other, and we used the public HEp-2 benchmarks to demonstrate the accuracy of our proposed techniques compared to the most representative works in recent literature. In [28] and [29] we tackled the problem of fluorescent pattern classification. We investigated several types and combinations of image attributes, showing that the integration of different types of descriptors (morphological, local texture and global texture features) is the most suited for this purpose. Furthermore, we proposed a classification scheme to cope with the high within-class variance typical of HEp-2 cell images. When tested on manually segmented cells, this approach was found to be among the top pattern classification techniques proposed by recent literature [11]. Then, in [30] we focused on HEp-2 cell segmentation, which is a challenging step that should ideally precede HEp-2 pattern classification. We presented a cell segmentation approach that automatically adapts to images with different intensity levels or fluorescent patterns and we experimentally validated our technique based on the concordance with manual segmentations performed by human experts.

Although our previous works obtained promising results, they do not say much on on how the automatization and integration of the segmentation and classification steps would impact on the final ANA testing outcome.

In this paper, we improve and extend our work as follows. (i) We tackle the task of automated intensity level classification, which is the IIF image analysis step that we had not addressed yet. In order to handle intensity variations that are inherent of fluorescence imaging, we propose a technique that leverages on multi-scale image contrast assessment and machine learning. (ii) We integrate intensity level classification, cell segmentation and pattern classification into a single modular tool, ANAlyte. After being trained on a suitable set of pre-labelled images, AN-

Alyte is able to receive a IIF slide as input and provide as output its intensity level and fluorescent pattern without requiring user-interaction. (iii) We experimentally validate ANAlyte on two different public benchmarks of IIF HEp-2 images to demonstrate the performance and robustness of our tool in a realistic clinical scenario. Unlike our previous works, here we focus our experimental analysis not on the validation of the individual techniques (the interested reader can refer to the publications mentioned before), but on demonstrating the effectiveness of our tool in the context of ANA testing.

1.2. Structure of the paper

This paper is structured as follows. In Section 2 we introduce ANAlyte, our proposed tool, and we describe its main modules in detail. In Section 3 we characterize the datasets used in our experiments and we present and discuss the experimental results. Finally, Section 4 concludes the paper.

2. ANAlyte image analysis tool

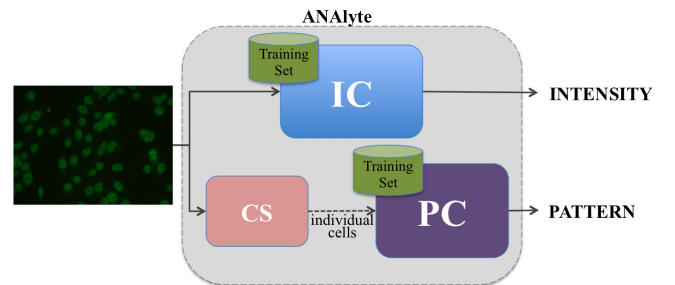The architecture of ANAlyte is summarized in Figure 3, where its main modules are represented with their connections:

Figure 3: ANAlyte architecture: Intensity Classifier (IC), Cell Segmenter (CS), Pattern Classifier (PC).

Intensity Classifier (IC) receives a IIF image as input and categorizes such image into one of the intensity levels represented by a pre-labelled Training Set.

Cell Segmenter (CS) extracts individual HEp-2 cells from the input IIF image.

Pattern Classifier (PR) receives the individual cells as obtained from CS and provides a categorization of the input IIF image into one of the fluorescent patterns represented by a pre-labelled Training Set.

ANAlyte provides as output the intensity level and the fluorescent pattern of the input slide. As anticipated in Section 1, such information is critical for the differential diagnosis of autoimmune diseases.

In the following, we describe the main modules of ANAlyte with more details.

2.1. INTENSITY CLASSIFIER (IC)

In the clinical practice, IIF samples are categorized into a specific number of levels based on the visual assessment of their fluorescent intensity compared to a set of negative and positive controls. The aim of IC module is the automatization of such practice, as follows: first, a number of significant sub-regions are automatically selected from the image (*Sub-regions Selection*, Section 2.1.1). Based on these sub-regions, a set of features are extracted at different scales and locations of the image (*Multi-scale Contrast Assessment*, Section 2.1.2). Finally, based on the computed features, the input image is automatically assigned to a specific intensity level by means of machine learning techniques (*Automatic Intensity Classification*, Section 2.1.3).

2.1.1. Sub-regions Selection

The intensity level of a IIF image is the perceivable strength of its fluorescent signal. On account of this, the categorization of IIF intensity levels can be interpreted as a problem of image contrast quantification, where contrast is traditionally defined as the difference of brightness between objects and background. Nevertheless, measuring the contrast of a IIF image is complicated by the extreme imbalance of foreground and background and by the presence of intensity variations across the image. In order to overcome these problems we perform contrast assessment on multiple sub-regions of the IIF image, taking advantage of multi-scale local information.

Given an input image I, sub-regions are chosen as follows:

1. we perform a rough binarization of I by means of Otsu thresholding;
2. we compute the connected regions of the foreground;
3. we calculate the centroids of the connected regions, spatially sorted from left to right:

$$\bar{c} = \begin{pmatrix} c_1 & c_2 & \cdots & c_{N/2} & \cdots & c_{N-1} & c_N \end{pmatrix}, \quad (1)$$

where N is the total number of connected regions (see example in Figure 4-b);

4. we select n equally-spaced elements from vector $\bar{c}$ (details are provided later on this Section), obtaining a vector $\bar{c}_s$:

$$\bar{c}_s = \begin{pmatrix} c_{s,1} & c_{s,2} & \cdots & c_{s,n} \end{pmatrix} \quad (2)$$

5. we generate $n+1$ rectangular windows:

$$\bar{W} = \begin{pmatrix} W_0 & W_1 & W_2 & \cdots & W_n \end{pmatrix}, \quad (3)$$

where W_0 is a global window that includes the whole image I, while $W_i(i > 0)$ are local windows centered on $c_{s,i}$ and with size gradually decreasing at the increase of i.

All the local windows are created starting from foreground objects. This ensures that the corresponding sub-images will contain both background and foreground, and thereby that measuring local contrast within such regions is significant.

Vectors $\bar{c}_s$ and $\bar{W}$ are generated by an *iterative windowing* process reported in Algorithm 1, with the following characteristics: (i) at each iteration, few elements of vector $\bar{c}$ (Equation 1) are selected with a certain sampling interval, and the corresponding local windows are generated; (ii) the sampling interval of the centroids as well as the size of the generated windows are decreased at each iteration. (iii) the procedure terminates when the generated windows are too small to include an adequate proportion of foreground and background. (iv) the final $\bar{c}_s$ and $\bar{W}$ vectors (Equations 2 and 3, respectively) are obtained by concatenating the output of each iteration.

Algorithm 1 Iterative windowing

1: $N_{ITER} \leftarrow 1$
2: $[d_{W1}, d_{W2}] = size(I)$ {I: input image}
3: $S_{INT} \leftarrow N$ {N: number of connected regions}
4: **while** $(d_{W1} \cdot d_{W2} \geq TH)$ **do**
5:    **if** $N_{ITER} = 1$ **then**
6:       Generate a window of size $[d_{W1}, d_{W2}]$ centered on a point with coordinates $(d_{W1}/2, d_{W2}/2)$
7:    **else**
8:       Perform a selection of S_{INT}-spaced elements from $\bar{c}$, symmetrically with respect to its middle.
9:       Center a $[d_{W1}, d_{W2}]$-sized window on each of the selected elements.
10:   **end if**
11:   $N_{ITER} \leftarrow N_{ITER} + 1$
12:   $S_{INT} \leftarrow S_{INT}/2$
13:   $[d_{W1}, d_{W2}] = size(I)/2^{N_{ITER}}$
14: **end while**

With reference to the pseudo-code in Algorithm 1 and to the example displayed in Figure 4-c, the *iterative windowing* proceeds as follows:

$N_{ITER} = 1$ generates one window W_0 that coincides with the bounding box of the whole image (red-colored)

$N_{ITER} = 2$ generates two windows W_1 and W_2 of $size(I)/4$ (green-colored), whose centers are obtained selecting $N/2$-spaced symmetric elements from vector $\bar{c}$ (Equation 1):

$$\begin{pmatrix} c_1 & \cdots & \boxed{c_{\frac{N}{4}}} & \cdots & \boxed{c_{\frac{3}{4}N}} & \cdots & c_N \end{pmatrix}$$

$N_{ITER} = 3$ generates four windows $W_3 \cdots W_6$ of $size(I)/8$ (yellow-colored), whose centers are obtained selecting $N/4$-spaced symmetric elements from vector $\bar{c}$:

$$\begin{pmatrix} \cdots & \boxed{c_{\frac{N}{8}}} & \cdots & \boxed{c_{\frac{3}{8}N}} & \cdots & \boxed{c_{\frac{5}{8}N}} & \cdots & \boxed{c_{\frac{7}{8}N}} & \cdots \end{pmatrix}$$

As reported in line 4 of Algorithm 1, the iterations end when the area of the local window is lesser than a threshold TH. Such threshold is set during the training phase of
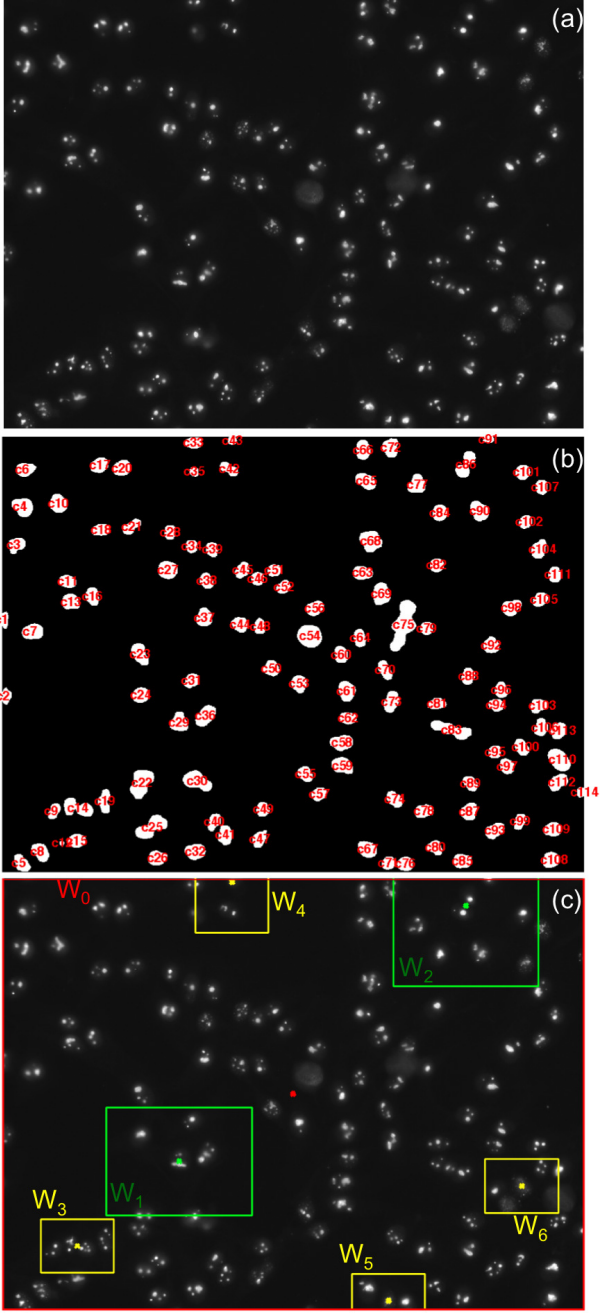
4

Figure 4: Phases of sub-regions selection: (a) input image; (b) connected regions of the binarized image, with spatial left-to-right sorting of the centroids; (c) windows generated by *iterative windowing* (Algorithm 1).

the algorithm, so that the sub-regions always contain a reasonable amount of background[1].

The ending result of the *iterative windowing* process is

---

[1]During the training phase, the operator draws the bounding box of a random HEp-2 cell in a IIF image. Such procedure is repeated for a set of training images of different fluorescent pattern. Hence, TH is estimated as twice the maximum bounding boxes area.
N.B. The training set must include images acquired with same size, imaging conditions and magnification of the input image.

the vector $\bar{W}$, obtained by concatenating all the windows generated at each iteration of the procedure. Hence, the total number of local windows, n, depends on the number of iterations, and ultimately on the threshold TH.

2.1.2. Multi-scale Contrast Assessment

By computing image contrast within each window in $\bar{W}$, we obtain a contrast features vector $\bar{C}t$:

$$\bar{C}t = \begin{pmatrix} Ct_{,0} & Ct_{,1} & \cdots & Ct_{,n-1} & Ct_{,n} \end{pmatrix}, \qquad (4)$$

where $Ct_{,0}$ is a measure of the global contrast of the whole image and $Ct_{,1} \cdots Ct_{,n}$ are measures of local contrast within sub-images of gradually decreasing size.

Literature provides several descriptors, which are best suited for quantifying image contrast at different scales [31]. Thanks to its robustness to intensity variations across the image, *RMS Contrast* (C_{RMS}), defined as the standard deviation of pixel intensities, is the most used metrics for the quantification of global contrast:

$$C_{RMS} = \sqrt{\frac{1}{d_{W1} \cdot d_{W2}} \sum_{i=0}^{d_{W1}-1} \sum_{j=0}^{d_{W2}-1} (I_{ij} - I_{avg})^2}, \qquad (5)$$

where I_{ij} is the $[0, 1]$ intensity of the (i, j)-th pixel of the image, (d_{W1}, d_{W2}) is the image size, and I_{avg} is the average image intensity.

Conversely, *Michelson Contrast* (C_M) is defined as the ratio between the spread and the sum of maximum and minimum intensity values over the image window, as follows:

$$C_M = \frac{I_{max} - I_{min}}{I_{max} + I_{min}} \qquad (6)$$

C_M is most significant when bright and dark portions take up similar fractions. Hence, in our context we use this metrics to compute contrast in the sub-images, where such pre-condition most likely applies.

Summarizing, with reference to Equation 4, the contrast features vector is calculated as follows:

$$Ct_{,i} = \begin{cases} C_{RMS} & \text{if } i = 0 \\ C_M & \text{if } i > 0 \end{cases} \qquad (7)$$

2.1.3. Automatic Intensity Classification

The contrast features vector $\bar{C}t$ is fed into a k-NN classifier, equipped with a set of training examples, i.e. images with a-priori known intensity labels. As introduced in Section 1, the intensity labels will be in our case three (either positive, intermediate or negative). Nevertheless, the classification technique can be generalised to any number of intensity levels, provided that such labels are adequately represented in the training set. The input image and the training examples are mapped into a $(n + 1)$ dimensional feature space, where the coordinates are represented by the $n + 1$ elements of $\bar{C}t$. Hence, the input image is assigned to the most represented intensity level among its k nearest neighbors.[2]

---

[2]$k = 7$, heuristically set

2.2. CELL SEGMENTER (CS)

CS module performs the automated segmentation of the individual HEp-2 cells. Such task is one of the most challenging of automated IIF analysis, because a single segmentation algorithm has to cope with a large heterogeneity of shapes and textures (see some examples in Figure 2). In order to address this problem, CS is implemented as an *adaptive marker-controlled watershed* algorithm, where both the preprocessing and the marker selection strategy self-adapt to the peculiar characteristics of the input image. First, the *Image Normalization* step is performed in order to soften textural differences of the input images. Then, the *Automatic Marker Selection* extracts a set of internal and external markers defining the position of the foreground and background, respectively. Finally, such markers are used to drive a *Marker-Controlled Watershed* technique towards the boundaries of the individual cells.

IIF images with different textural characteristics demand specific segmentation strategies. Hence, to enforce the adaptiveness of the algorithm, the incoming image is automatically categorized into one of the following:

(i) Smooth textured images, with bright cell bodies and dark background (e.g. first image of Figure 2).

(ii) Rough textured images, with cell bodies partly bright and partly dark (e.g. last three images of Figure 2).

Such categorization is based on the average area of the connected regions returned by Otsu's thresholding algorithm, as proposed in [30] (large area $\Rightarrow$ smooth textured, small area $\Rightarrow$ rough textured).

2.2.1. Image Normalization

For both image categories, histogram equalization is performed in to enhance the cell bodies with respect to the background, followed by gray-scale morphological opening. As *Rough textures* are characterized by large dark areas within the cell body, this category of images are also preprocessed with a combination of Top-Hat filtering and morphological greyscale reconstruction [32].

2.2.2. Automatic Marker Selection

The automatic marker selection consists of the following steps. First, an *adaptive fuzzy c-means clustering* roughly separates the foreground regions from the background. Then, cell clusters (i.e. the regions including two or more touching cells) are split into separate objects by *Randomized Hough Transform*. At the end of this procedure, a collection of *internal* and *external markers* is obtained, defining the position of the individual cells and of the background, respectively.

As observed by [6], pixel intensities in IIF images can be roughly grouped into three bands: low, medium and high intensity. Pixels in the low and high intensity bands belong to, respectively, background and foreground. The content of the medium intensity band, on the other hand,

strictly relates to the textural characteristics of the image, and hence to the image category: in the *smooth textured* images it contains background pixels only, while in *rough textured* images it contains both background and foreground pixels.

Upon such observation, the *adaptive identification of foreground regions* is implemented as follows. A three-classes *fuzzy c-means clustering* (FCM [33]) is applied to separate the image pixels into the three intensity bands (low, medium and high-intensity). FCM assigns each pixel value a [0,1] degree of membership to each of the three clusters, summarized by the *membership functions* M_{low}, M_{med} and M_{high} reported in Figure 5. Then, the ultimate separation of foreground and background is obtained by image thresholding, where the threshold I_{TH} is computed as follows:

$$I_{TH} = \begin{cases} \text{intersection}(M_{med}, M_{high}) & \text{if } Smooth \\ \text{barycenter}(M_{med}) & \text{if } Rough \end{cases} \quad (8)$$
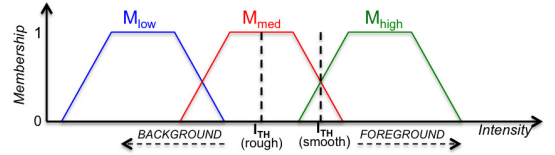

Figure 5: Image thresholding based on FCM membership functions.

To cope with uneven illumination problems typical of fluorescence microscopy, the clustering algorithm is implemented with a sliding window approach. Hence, each pixel is assigned a local I_{TH} value, based on the intensities of its neighbors.

The foreground objects identified with FCM may either contain (i) one individual cell or (ii) multiple touching cells (i.e. cell clusters), which need to be separated further into individual objects. These two categories are identified by imposing a threshold I_E on the *ellipticity value* of each object:

$$\begin{cases} \text{ellipticity} \geq I_E \Rightarrow \text{individual cell} \\ \text{ellipticity} < I_E \Rightarrow \text{cell cluster,} \end{cases} \quad (9)$$

where *ellipticity* is the ratio between the area of a connected component and the area of its best-fitted ellipse, and I_E is the average ellipticity of all the objects in the image. The rationale of this approach is that individual cells have a more regular shape than cell clusters, and hence higher ellipticity.

The identified cell clusters are then split into elliptical sub-regions by means of a *Randomized Hough Transform*, RHT [34], a probabilistic variant of classical Hough Transform. The resulting ellipses are a very rough approximation of the cell boundaries. However, their centers indicate the most likely position of the individual cells within the cell clusters.

Ultimately, we obtain a set of *internal markers* defining the position of the individual cells in the image and a set of

external markers defining the background. Internal markers include: (i) the pixels of the individual cells as obtained from Equation 9 (ii) the centers of the ellipses returned by the RHT-based decomposition of the cell clusters. Conversely, the external markers include the edge points of a Voronoi diagram built using the internal markers as seeds.

2.2.3. Marker-Controlled Watershed

Cell segmentation is obtained with a *marker-controlled watershed* algorithm [35]. In our implementation, watershed technique is applied to the gradient of the image as obtained after *Image Normalization*.

As touching cells as a rule do not have clear edges between each other, a certain degree of imprecision in detecting the exact boundaries can still be expected. Nevertheless, as it will be shown in Section 3, this does not have critical impact on pattern classification accuracy.

2.3. PATTERN CLASSIFIER (PC)

The segmented cells returned by CS are fed into PC module, which performs the automatic categorization of the fluorescent pattern.

Prior to the actual pattern classification algorithm, the cell images undergo a preprocessing stage. More specifically: (i) *image normalization* is applied in order to reduce variabilities of cell size and intensity. As for the size, all the bounding boxes of the input cells are resized to a standard 128x128 pixels dimension. As for the intensity, pixel values are linearly remapped so that the bottom 1% and the top 1% of the intensities are saturated. Thanks to intensity normalization, the same classifier can be applied irrespective of the intensity level of the input image. Hence, the PC module is completely independent from the IC module, in that it does not need to receive intensity as an input parameter. (ii) *mitotic cell detection* as reported in [36] is applied in order to identify and remove the mitotic cells (i.e. cells undergoing cellular division). Mitotic cells are in minority in a HEp-2 slide, and are characterized by a very distinct morphology compared to the other cells. Hence, selectively removing such cells facilitates the pattern recognition process.

Then, the following steps are performed: (i) *Feature Extraction* computes a feature vector representing relevant textural and morphological attributes of the cells (ii) a *Subclass Discriminant Analysis (SDA)* approach remaps the cells' representation into a novel feature space providing a better class separation. (iii) Based on the remapped features of the individual cells, *Image Classification* provides as ultimate output the fluorescent pattern of the IIF image.

2.3.1. Feature Extraction

Three types of image descriptors are extracted:

1. *morphological features*, focusing on shape attributes of the fluorescent signal;

2. *global texture descriptors*, summarizing the overall distribution of the gray-levels in the cell;

3. *local texture descriptors*, representing isolated contribution of small regions or pixel neighborhoods.

As demonstrated in [29], the integration of these three types of attributes provides a more comprehensive characterization of the fluorescence patterns.

As for *morphological features*, the computed descriptors are the circularity of the cell (defined as Area/Perimeter^2), as well as a set of shape attributes extracted from the binarization of the cell image at seven increasing intensity thresholds, as defined in [29].

As for *global texture descriptors*, a set of first and second order statistical measures are derived from the grey-level co-occurrence matrix (GLCM), representing the distribution of co-occurring values between neighboring pixels according to different distances and directions.

Local texture descriptors, on the other hand, are based on Local Binary Patterns (LBPs), i.e. binary patterns representing the intensity relations between a pixel and its neighbours. More specifically, features are extracted applying a modified formulation of LBP which provides information about the co-occurrence of multiple adjacent LBPs (Co-occurrence of Adjacent LBPs, CoALBP).

Morphological, GLCM and CoALBP features are computed and concatenated into a single feature vector. Then, in order to extract the subset of features that is most relevant for the pattern classification task, a robust feature selection method based on minimum-Redundancy-Maximum-Relevance (mRMR) algorithm is applied. mRMR sorts the features that are most relevant for the characterization of the classification variable, pointing at the minimization of their mutual similarity and maximization of their correlation with the classification label, as provided by the training set. The ultimate size of the the selected features vector is hence 150, as defined in [29].

2.3.2. Subclass Discriminant Analysis (SDA)

After feature selection, fluorescent pattern classification is obtained with a machine learning technique based on Subclass Discriminant Analysis (SDA) that has been already introduced in [29].

To cope with the inherent non-linearity and within class variability of IIF patterns, SDA approximates the underlying distribution of each class by a mixture of Gaussians. Thereby, the classes are divided into a set of subclasses, working towards the maximization of the distance among means of classes and, at the same time, of the distance among means of subclasses within each class. Hence, the ultimate problem to be solved is finding the optimal number of subclasses for each class. In our work, the optimal subdivision is the one able to maximize the recognition rate of a *leave-one-out* process (details in Section 3) in a dataset independent from the one used for validation purposes.

Our implementation has two major differences compared to original SDA formulation. First, while classical SDA partitions the samples into a set of equal-sized subclasses sorted with a Nearest-Neighbor algorithm, we apply a K-means clustering algorithm. This allows to model more effectively the non-linearity in the data. Second, instead of increasing at each iteration the number of subclasses of all groups of a same amount, as in [37], we apply a heuristic approach. We first set a maximal number of subclasses for each class, based on the number of per-class samples. Then, all possible permutations of class subdivisions are tested, halting the subdivision of a specific class when the minimal number of samples in any sub-cluster is reached.

2.3.3. Image Classification

Once the feature vectors have been projected on the sub-space defined by SDA, the classification is performed with k-NN algorithm.[3]. More specifically, k-NN is applied to categorize each individual cell into one of the fluorescent patterns represented by the training set, treating the HEp-2 cells as independent samples. Ultimately, the class of the input IIF image is obtained based on the majority class of its individual cells.

3. Experimental Evaluation

3.1. Validation strategy

All the experiments in this work are based on a *leave-one-out cross-validation* (LOOCV) protocol. At each iteration (one per each image of the validation dataset), all the IIF slides except one are used for training the tool, using the left out slide for testing purposes. This strategy ensures a complete separability and independence of training and testing, and maximizes the number of independent samples (i.e. of specimens belonging to different patients) used for training, which strengthens the learning process of the classification algorithm. Furthermore, the execution of several training/test runs allows to assess the generalization capabilities of the tool.

The two outputs of ANAlyte (i.e. fluorescent intensity level and fluorescent pattern) were evaluated independently, in terms of rate of correct classifications. In order to provide a better insight into the tool's behaviour, fluorescent pattern accuracy was evaluated not only *image-wise* (i.e. in terms of IIF slides correctly classified, which is the information that is relevant for clinical purposes) but also *cell-wise*, in terms of individual cells correctly classified.

In all cases the percentage accuracy value was obtained by adding the classification counts of each iteration and

---
[3]$k = 8$, heuristically set

then calculating a percentage over all runs, as follows:

$$LOOCVacc(\%) = \frac{\sum_{i=1}^{n_i} \# \text{ correct classifications}}{N} \cdot 100,$$

(10)

where n_i is the number of LOOCV iterations (which coincides with the number of IIF slides of the testing dataset) and N is the total number of classification counts over all iterations.

3.2. HEp-2 cell benchmarks

In our experiments we used two different datasets, here referenced as *MIVIA* and *I3Asel*. The full characterization of these datasets is reported in Table 1.

MIVIA dataset is an annotated database of Indirect Immunofluorescence (IIF) HEp-2 images publicly available at [38]. The database was firstly described in [39] and later used for the *HEp-2 Cells Classification Contest* hosted by the conference ICPR 2012 and for a Special Issue of the journal Pattern Recognition on the *Analysis and Recognition of IIF Images* [40]. Ever since it has been largely adopted as a benchmark for IIF image analysis tools. It contains 1388x1038 sized 24 bit color depth images of 28 different patients (one image per patient). HEp-2 substrates were prepared with fixed dilution of 1:80, as per recommendations of the guidelines, and images were acquired with a fluorescence microscope (40-fold magnification) coupled with a 50W mercury vapor lamp and a digital CCD camera with 6.45 μm size square pixels. The dataset provides manual segmentations and annotations of all the HEp-2 cells, which are 1582 in total. Physicians specialized in immunology reviewed all the segmentations and reported on the fluorescence intensity (either intermediate or positive level), mitotic phase and fluorescent pattern of each cell (respectively, one among six: homogeneous, fine speckled, coarse speckled, nucleolar, centromere and cytoplasmic). The images are quite evenly distributed among different intensity levels and fluorescent patterns (details in Table 1).

I3Asel dataset originates from the *First Workshop on Pattern Recognition Techniques for IIF images (I3A)* of the conference ICPR 2014 [18]. 1001 sera of patients positive to ANA testing were diluted to 1:80 and HEp-2 specimens were photographed using a monochrome camera fitted on the fluorescent microscope, obtaining 1388x1040 sized IIF images with two different intensity levels (again, intermediate and positive) and seven different pattern classes (homogeneous, speckled, nucleolar, centromere, nuclear membrane, golgi and mitotic spindle), with large majority of images belonging to the first four patterns. Unlike *MIVIA dataset*, neither manual segmentation nor ground truth about the mitotic phase of the cells were provided, but only their fluorescent intensity level and pattern. 252 out of 1001 specimens were made publicly available, while the rest was used to test the participants' results and remained private.

I3Asel dataset contains 71 images belonging to the public portion of the dataset [4], chosen as follows: first, we picked all the specimens of the least represented pattern (i.e. golgi). Then, we randomly selected the specimens of the other classes ensuring (i) the independence of the images (i.e. each image derives from a different specimen) and (ii) a good balance of the different patterns and fluorescent levels (details in Table 1), except for mitotic spindle patterns, that were excluded from the study. Indeed, unlike regular ANAs, mitotic spindle autoantibodies (MSAs) can be observed on only a small portion of the specimen cells. As the per-cell classification approach followed by ANAlyte assumes that the pattern of the specimen coincides with the pattern of the majority of the cells in the image, the identification of MSAs is out of our scope.

Table 1: Characterization of the HEp-2 cell datasets used as benchmarks.

	MIVIA		I3Asel	
	images	cells	images	cells
Pattern				
Homogeneous	5	345	12	1626
Fine speckled	4	225	–	–
Coarse speckled	5	239	–	–
Centromere	6	388	12	1603
Nucleolar	4	257	12	1308
Cytoplasmic	4	128	–	–
Speckled	–	–	12	1556
Golgi	–	–	11	988
Nuclear membrane	–	–	12	1262
Intensity				
Positive	15	857	35	4292
Intermediate	13	725	36	4051
tot	28	1582	71	8343

3.3. Results and Discussion

Table 2 reports the LOOCV accuracy of fluorescent intensity and fluorescent pattern classification (for the latter, both image-wise and cell-wise).

The first row of the table shows values obtained with a semi-automated version of ANAlyte, where the pattern classification is performed on manually segmented cells, using the cell segmentation ground truth provided along with the benchmark (the experiments were run only on *MIVIA* dataset in this case, because *I3Asel* does not provide such ground truth). The second row of the table shows results obtained with the fully-automated version of ANAlyte (i.e. including the automated segmentation of

the HEp-2 cells, as depicted in Figure 3), on both *MIVIA* and *I3Asel* datasets.

Table 2: LOOCV accuracy values (%).

ANAlyte	dataset	INTENSITY LEVEL	FLUOR. PATTERN	
			image-wise	cell-wise
Semi-auto[(*)]	MIVIA	85.71	96.43	89.55
Fully-auto	MIVIA	85.71	92.86	73.44
	I3Asel	84.51	90.14	76.91

[(*)]Run on manually pre-segmented cells.

In the following, we analyze in depth the results on intensity level and fluorescent pattern classification.

3.3.1. Intensity level classification

The accuracy of intensity level classification was above 80% in all our experiments. More specifically, as shown in the third column of Table 2, the rate of IIF slides correctly classified was respectively 85.71% for *MIVIA* and 84.51% for *I3Asel* dataset. As the Intensity Classifier module works on the full slide and not on the individual HEp-2 cells (see Figure 3), the image classification accuracy does not depend on cell segmentation. Hence, values reported in the second and third row of the table are the same.

For a deeper insight into the obtained results, in Table 3 we report the confusion matrices of both datasets, where columns represent the instances (i.e. the number of IIF slides) in a predicted intensity level, and rows the instances in the actual intensity level. To improve the readability of the matrices, the darkness of the background of each cell is proportional to its content.

As can be seen from the table, the classification accuracy for intermediate and positive images was about 85% and 87% in *MIVIA dataset*, 92% and 77% in *I3Asel dataset*.

Table 3: *Confusion matrices for intensity level classification.*

MIVIA dataset

		Predicted	
	level	Intermediate	Positive
Actual	Intermediate	11 (84.62%)	2 (15.38%)
	Positive	2 (13.33%)	13 (86.67%)

I3Asel dataset

		Predicted	
	level	Intermediate	Positive
Actual	Intermediate	33 (91.67%)	3 (8.33%)
	Positive	8 (22.86%)	27 (77.14%)

Ultimately, we report the accuracy of intensity classification obtained in both benchmarks, grouped by fluorescent pattern. As it is visible from Table 4, the only patterns that obtained accuracy below 75% were the cytoplasmic in MIVIA dataset and golgi in *I3Asel* dataset (respectively, 50% and 63.64%).

3.3.2. Fluorescent pattern classification

The fully-automated version of ANAlyte generally obtained a good fluorescent pattern classification accuracy:

---

[4]As anticipated in Section 3, LOOCV strategy is best suited to evaluate the robustness and generalization capabilities of a classifier. Nevertheless, as SDA training is based on repeating a cross-validation per each permutation of the sub-classes, in our case LOOCV imposes a cross-validation within the cross-validation, which is computationally very intensive. Hence, in our experiments we run LOOCV on a sub-set of the I3A images, randomly selected. This is different from the strategy used by I3A contest, where tests were run on a private set of images with same characteristics of the training set.

Table 4: *Accuracy of intensity classification grouped by pattern.*

Pattern	MIVIA dataset	I3Asel dataset
Homogeneous	100%	80%
Fine speckled	75%	–
Coarse speckled	100%	–
Centromere	100%	83.33%
Nucleolar	75%	91.67%
Cytoplasmic	50%	–
Speckled	–	100%
Golgi	–	63.64%
Nuclear membrane	–	83.33%

Table 5: *Fluorescent pattern classification (image-wise): confusion matrices.*

(a) MIVIA dataset (manually pre-segmented cells)

class	Cent.	Hom.	Nucl.	C. Sp.	F. Sp.	Cyt.
Cent.	5 (83.33%)	1 (16.67%)	0	0	0	0
Hom.	0	5 (100%)	0	0	0	0
Nucl.	0	0	4 (100%)	0	0	0
C. Sp.	0	0	0	5 (100%)	0	0
F. Sp.	0	0	0	0	4 (100%)	0
Cyt.	0	0	0	0	0	4 (100%)

(b) MIVIA dataset (fully-automated)

class	Cent.	Hom.	Nucl.	C. Sp.	F. Sp.	Cyt.
Cent.	5 (83.33%)	0	0	0	1 (16.67%)	0
Hom.	0	5 (100%)	0	0	0	0
Nucl.	0	0	4 (100%)	0	0	0
C. Sp.	0	0	0	5 (100%)	0	0
F. Sp.	0	0	0	0	4 (100%)	0
Cyt.	0	0	0	1 (25%)	0	3 (75%)

(c) I3Asel dataset (fully-automated)

class	Cent.	Hom.	Nucl.	Sp.	Gol.	Num.
Cent.	11 (91.67%)	0	0	1 (8.33%)	0	0
Hom.	0	12 (100%)	0	0	0	0
Nucl.	0	0	12 (100%)	0	0	0
Sp.	1 (8.33%)	1 (8.33%)	1 (8.33%)	8 (66.67%)	0	1 (8.33%)
Gol.	0	0	1 (9.09%)	0	10 (90.91%)	0
Num.	0	1 (8.33%)	0	0	0	11 (91.67%)

the rate of images correctly classified was 92.86% for *MIVIA* dataset and of 90.14% for *I3Asel* dataset, respectively.

By comparing the results of the two experiments run on *MIVIA* dataset, we can also quantify the impact of automated cell segmentation on the final outcome of the tool. When considering cell-wise classification results (last column of Table 2), the accuracy drop from the semi-automated to the fully-automated version of ANAlyte was quite significant: from 89.55 to 73.44%, respectively. This is, indeed, not surprising, if we consider that the textural and morphological features used for pattern classification are highly dependent on the region of interest they are extracted from (see Section 2.3.1). As a consequence, small errors in the identification of the cellular borders can easily mislead textural analysis and hence translate into a wrong categorization of the cellular pattern. Nevertheless, in order to assess the real significance of these numbers in a realistic clinical context, we need to compare them with the level of accuracy of a human specialist working in the same conditions as the automated tool. As reported by [40], a physician was asked to manually classify each cell of *MIVIA* dataset, obtaining a cell classification accuracy of 73.3%. This value is indeed comparable to the 73.44% obtained by ANAlyte. Hence, we can conclude that the cell segmentation error does not affect pattern classification in a way that is critical for its clinical application.

On the other hand, if we analyze results on image-wise classification, we can observe that the accuracy drop due to cell segmentation errors was much smaller. As shown in Table 2, the accuracy was 96.43% with the semi-automated version of ANAlyte and 92.86% with the fully-automated version, while the accuracy reported for a human specialist in the same dataset is 87% [40]. If we consider that the fluorescent pattern of the IIF slide (and not of the individual cells) is the output that is relevant for diagnostic purposes, this result further supports the convenience of using ANAlyte in a clinical context.

Ultimately, in order to assess the behavior of the tool in the different pattern classes, we report the confusion matrices of all performed experiments (respectively, in terms of image-wise classification in Table 5 and cell-wise classification in Table 6). Again, the columns represent the predicted classes and the rows the actual classes as reported by the corresponding benchmarks.

From the analysis of the confusion matrices, we can draw the following considerations:

(i) in general, image-wise (Table 5) and cel-wise (Table 6) accuracies do not have a direct correspondence, even when considered within the same experiment. In other words, the patterns with the lowest (or highest) cell-wise accuracy do not necessarily maintain the same trait in terms of image-wise accuracy. For example, nucleolar and speckled patterns in *MIVIA* dataset had relatively low cell-wise accuracy values (lower than 70%), but the corresponding image-wise accuracy was 100%. Likewise, golgi pattern in *I3Asel* dataset (around 61% cell-wise accuracy, against 92% rate of IIF slides correctly classified). To this end, we have to consider the different type of information carried by these two metrics. Image-wise accuracy (as already explained before in this Section) is the most significant descriptor for assessing the overall performance of a pattern classification tool in the context of ANA testing. On the other hand, the rate of cells correctly classified supplies additional information on the robustness of classification, but only provided that it is considered jointly with the former metrics.

(ii) by comparing the results of the fully-automated ANAlyte with the ones on manually segmented cells (see Table 5(a) and (b)), we can observe that the only pattern which was somehow affected by cell segmentation errors is the cytoplasmic. Indeed, the cytoplasmic images are peculiar compared to all the others (i.e. the fluorescence arises from the cytoplasm rather than the nucleus of the cell), which translates into a much higher probability of cell segmentation errors. Nevertheless, the impact on pattern classification accuracy was reasonably low (one misclassified slide in Table 5(b) against none in Table 5(a)). All other patterns registered a drop of cell-wise classification accuracy, but their image-wise accuracy was left

Table 6: *Fluorescent pattern classification (cell-wise): confusion matrices (%).*

(a) MIVIA dataset (manually pre-segmented cells)

class	Cent.	Hom.	Nucl.	C. Sp.	F. Sp.	Cyt.
Cent.	83.47	5.60	2.24	3.08	5.60	0
Hom.	0.91	93.03	0.61	0.91	4.24	0.30
Nucl.	1.25	1.24	93.78	3.32	0	0.41
C. Sp.	2.38	0.48	1.43	93.33	2.38	0
F. Sp.	1.44	12.5	0	4.33	81.73	0
Cyt.	0	0	0	2.75	0	97.25

(b) MIVIA dataset (fully-automated)

class	Cent.	Hom.	Nucl.	C. Sp.	F. Sp.	Cyt.
Cent.	75.64	2.09	3.48	4.64	12.06	2.09
Hom.	2.96	79.57	0.81	4.30	9.68	2.69
Nucl.	14.84	3.91	67.97	6.25	0.39	6.64
C. Sp.	5.24	1.75	8.39	68.88	3.50	12.24
F. Sp.	7.06	12.55	1.96	4.31	68.63	5.49
Cyt.	3.91	1.74	0.87	14.78	2.17	76.52

(c) I3Asel dataset (fully-automated)

class	Cent.	Hom.	Nucl.	Sp.	Gol.	Num.
Cent.	82.78	0.06	1.93	11.29	3.18	0.75
Hom.	0	92.42	0.56	4.91	0.75	1.37
Nucl.	2.51	0.73	78.88	7.36	7.69	2.83
Sp.	11.55	8.66	6.18	64.94	2.42	6.25
Gol.	4.15	3.34	17.41	8.91	60.63	5.57
Num.	0.79	9.11	2.46	9.51	3.49	74.64

unchanged.

(iii) When considering the results on *I3Asel* datasets shown in Table 5(c), ANAlyte's performance was very good (higher than 90%) for all patterns, except for speckled (around 67%). Furthermore, the four (out of a total of twelve) misclassified speckled images were not categorized into one specific pattern class, but respectively as centromere, homogeneous, nucleolar and nuclear membrane. To clarify this behaviour, it is worth noting that the *I3Asel* benchmark is characterized by very high within-class variability of the speckled images, which makes the interpretation of this pattern particularly ambiguous compared to the others. For example, with reference to Figure 6, sample (b), speckled, has textural characteristics that are in between those of (a), speckled, and (c), centromere.
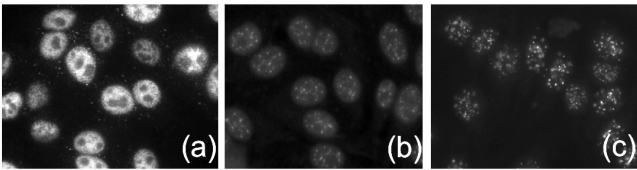

Figure 6: I3Asel benchmark. (a),(b) examples of IIF slides with speckled pattern. (c) example of IIF slide with centromere pattern.

This consideration does not apply to *MIVIA* dataset, where the speckled pattern is split into two distinct categories (coarse and fine, respectively). Indeed, the classification accuracy of the speckled patterns in *MIVIA* dataset was 100% (see Table 5(b)).

In order to assess the results obtained in our experiments, it is useful to compare them with the published results of recent pattern classification approaches. This assessment is summarized in Figure 7, where we show:

(i) on the left, a boxplot of the results of the works published in the Special Issue on *Analysis and Recogni-*

tion of Indirect Immunofluorescence Images of the journal Pattern Recognition [11]. Such numbers were obtained on *MIVIA* dataset, using the manually segmented cells as input, and computing fluorescent pattern accuracy in terms of rate of images correctly classified in a LOOCV protocol, as explained in Section 3.1. The magenta and green dots represent, respectively, the accuracy of the semi-automated version of ANAlyte (i.e. run on manually segmented cells, same as for the boxplot), and the accuracy of the fully-automated version.

(ii) on the right, a boxplot of the results obtained by the participants of I3A contest [18] (specimen classification task), with a green dot representing ANAlyte's accuracy on the *I3Asel* dataset. As anticipated in Section 3.2, the test database of I3A contest is not publicly available. Hence, a direct comparison between ANAlyte and the other methods is not possible in this case. However, in order to increase the consistency between I3A contestants and ANAlyte, both sets of results are shown for the same six patterns (i.e. with exclusion of MSAs). Accuracy is computed in terms of average per-class correct rate, same as it was done by the contest's organisers.
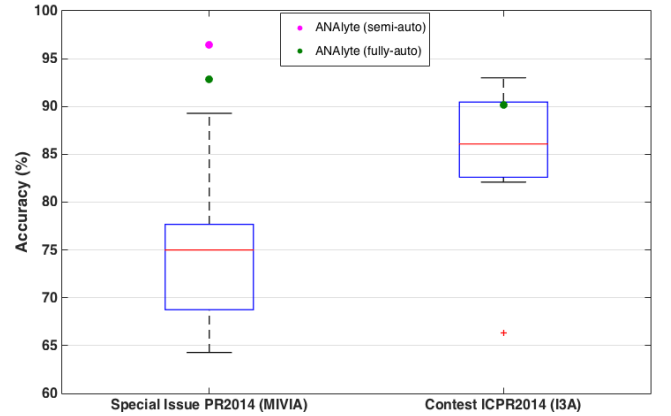

Figure 7: Fluorescent pattern classification (image-wise). Left: comparison between ANAlyte and methods published in Pattern Recognition 47(7) 2014. Right: comparison between ANAlyte and participants of the HEp-2 specimen classification contest at I3A Workshop 2014.
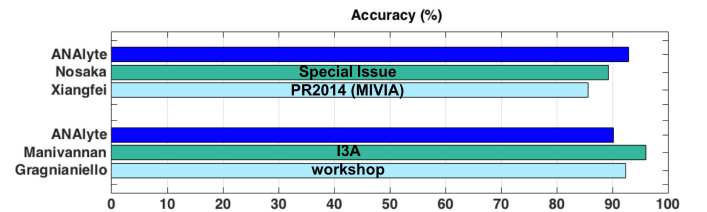

Figure 8: Fluorescent pattern classification (image-wise). Top: comparison between ANAlyte and top two methods from Pattern Recognition 47(7) 2014. Bottom: comparison between ANAlyte and top two methods from proc. of I3A Workshop 2014.

Ultimately, in Figure 8 we show the LOOCV accuracy of ANAlyte (fully-automated version) and of the best per-

forming techniques of Figure 7. More specifically, the top and bottom parts of the figure show:

(i) the top two methods from the special issue Pattern Recognition 47 (7) 2014, respectively by Nosaka et al. and Xiangfei et al. In [23], Nosaka et al. applied a framework combining a novel type of LBP descriptor, namely Rotation Invariant Co-Occurrence among adjacent LBPs (RIC-LBP), with a traditional multi-class SVM classifier. In [24], Xiangfei et al. proposed a novel dictionary learning strategy for the sparse representation of the cell images, leveraging on textural and gradient features.

(ii) the top two methods from the proceedings of I3A Workshop 2014[5], respectively by Manivannan et al. and Gragnaniello et al. In [26], Manivannan et al. applied sparse coding with max-pooling to aggregate texture features with cell pyramids at two different sets of scales. In [25], Gragnaniello et al. proposed a Bag of Words representation of a novel dense local descriptor combining log-polar transformation with multi scale gaussian smoothing applied to the gradient images.

From the analysis of Figure 7 and 8, we can observe that ANAlyte obtained the best results among all the other methods of Pattern Recognition 47 (7) 2014 and it was in line with the top-quartile performance methods of I3A workshop, albeit outperformed by the top ranking techniques proposed by Manivannan and Gragnaniello et al. (respectively, by 6% and 2.25%).

When interpreting such results, it must be taken into account that most of the other methods were assessed on pre-segmented cells, while ANAlyte addresses the automatisation of the whole analysis, including cell segmentation. Sure indeed, in spite of the accuracy drop due to automated cell segmentation, we can conclude that ANAlyte is comparable in terms of fluorescent pattern classification with the top approaches proposed by recent literature.

4. Conclusions

The automatization of the analysis of IIF HEp-2 specimens is widely recognized as the only valid solution to avoid the subjectivity of the human interpretation of the images and to make the results of ANA testing completely repeatable. For this purpose, in this paper we proposed a fully-automated tool, ANAlyte, that provides the information required by the immunologists (i.e. the intensity level and fluorescent pattern of the IIF image) without needing user interaction.

To this effect, ANAlyte combines the following modules:

(i) the INTENSITY CLASSIFIER (IC) module performs a categorization of the fluorescent intensity into a given number of levels, based on a training set of labelled images. IC extracts a set of features aimed at quantifying image contrast at different scales and in several locations of the input slide, in order to handle intensity variations that are inherent of fluorescence imaging. Finally, this set of contrast features is fed into a machine learning algorithm that provides the intensity level of the input image

(ii) the CELL SEGMENTER (CS) module splits the input image into individual HEp-2 cells without any a-priori knowledge about its intensity level or pattern. Since cells with different patterns require different processing steps, CS automatically adapts the segmentation procedure based on a pre-classification step. Then, an adaptive marker-controlled watershed approach is applied to obtain the segmentation of the individual HEp-cells.

(iii) the PATTERN CLASSIFIER (PR) module receives the individual HEp-2 cells from the CS module and categorizes them into a set of fluorescent patterns represented in its training set. The intra-class variability that is inherent of this type of images is softened by subclass discriminant analysis, that splits each pattern into an optimal number of sub-classes. Finally, a voting on the fluorescent pattern of the individual cells allows to determine the fluorescent pattern of the input image.

Extensive experimental results run on two public benchmarks of HEp-2 images allow to:

(i) show that the integration of CELL SEGMENTER and PATTERN CLASSIFIER modules does not impact in a critical way on the effectiveness of the tool when compared with human evaluation. Hence, a full automatization of pattern classification is feasible and can be applied to a clinical setting.

(ii) demonstrate the good level of accuracy and robustness of ANAlyte in analyzing the fluorescent intensity and pattern of IIF images without user-interaction.

In the future, we plan to make our tool available for use in a clinical laboratory and run more tests on real patient data.

References

[1] P. Soda, G. Iannello, M. Vento, A multiple expert system for classifying fluorescent intensity in antinuclear autoantibodies analysis, Pattern Analysis and Applications 12 (3) (2009) 215–226.

[2] A. Rigon, P. Soda, D. Zennaro, G. Iannello, A. Afeltra, Indirect immunofluorescence in autoimmune diseases: Assessment of digital images for diagnostic purpose, Cytometry Part B: Clinical Cytometry 72B (6) (2007) 472–477.

[3] A. Melegari, C. Bonaguri, A. Russo, B. Luisita, T. Trenti, G. Lippi, A comparative study on the reliability of an automated system for the evaluation of cell-based indirect immunofluorescence, Autoimmunity Reviews 11 (10) (2012) 713 – 716.

[4] C. Creemers, K. Guerti, S. Geerts, K. Van Cotthem, A. Ledda, V. Spruyt, HEp-2 cell pattern segmentation for the support of autoimmune disease diagnosis, in: Proceedings of the 4th International Symposium on Applied Sciences in Biomedical and Communication Technologies, ISABEL '11, 2011, pp. 28:1–28:5.

[5] B. Divya, H. Nanjundaswamy, N. Muniraj, Segmentatation of iif Hep-2 ANA cells based on thresholding and connectivity,

---

[5] As for the second boxplot of Figure 7, it should be taken into consideration that ANAlyte was assessed on a subset of the I3A images.

International Journal of Microsystems Technology and its Applications (IJMTA) 1 (1) (2012) 26–30.

[6] G. Percannella, P. Soda, M. Vento, A classification-based approach to segment HEp-2 cells, in: Computer-Based Medical Systems (CBMS), 2012 25th International Symposium on, 2012, pp. 1–5.

[7] D. Bertin, N. Jourde-Chiche, P. Bongrand, N. Bardin, Original approach for automated quantification of antinuclear autoantibodies by indirect immunofluorescence, Clinical and Developmental Immunology (2013) 1–8.

[8] G. Iannello, L. Onofri, P. Soda, A slightly supervised approach for positive/negative classification of fluorescence intensity in hep-2 images, in: Image Analysis and Processing ICIAP 2013, Vol. 8157 of Lecture Notes in Computer Science, Springer Berlin Heidelberg, 2013, pp. 319–328.

[9] G. Iannello, L. Onofri, P. Soda, A bag of visual words approach for centromere and cytoplasmic staining pattern classification on hep-2 images, in: Computer-Based Medical Systems (CBMS), 2012 25th International Symposium on, 2012, pp. 1–6.

[10] P. Agrawal, M. Vatsa, R. Singh, Hep-2 cell image classification: A comparative analysis, in: Machine Learning in Medical Imaging, Vol. 8184 of Lecture Notes in Computer Science, 2013, pp. 195–202.

[11] P. Foggia, G. Percannella, A. Saggese, M. Vento, Pattern recognition in stained HEp-2 cells: Where are we now?, Pattern Recognition 47 (7) (2014) 2305 – 2314.

[12] A. Larsen, J. Vestergaard, R. Larsen, HEp-2 cell classification using shape index histograms with donut-shaped spatial pooling, Medical Imaging, IEEE Transactions on 33 (7) (2014) 1573–1580.

[13] K. Egerer, D. Roggenbuck, R. Hiemann, M.-G. Weyer, T. Buttner, B. Radau, R. Krause, B. Lehmann, E. Feist, G.-R. Burmester, Automated evaluation of autoantibodies on human epithelial-2 cells as an approach to standardize cell-based immunofluorescence tests, Arthritis Research and Therapy 12 (2) (2010) R40.

[14] A. Rigon, F. Buzzulini, P. Soda, L. Onofri, L. Arcarese, G. Iannello, A. Afeltra, Novel opportunities in automated classification of antinuclear antibodies on hep-2 cells, Autoimmunity Reviews 10 (10) (2011) 647 – 652.

[15] J. Voigt, C. Krause, E. Rohwader, S. Saschenbrecker, M. Hahn, M. Danckwardt, C. Feirer, K. Ens, K. Fechner, E. Barth, T. Martinetz, W. Stocker, Automated indirect immunofluorescence evaluation of antinuclear autoantibodies on HEp-2 cells, Clinical and Developmental Immunology (2012) 1–7.

[16] ICPR 2012, contest on HEp-2 cells classification (nov 2012).
URL http://mivia.unisa.it/contest-hep-2/

[17] ICIP 2013, contest on HEp-2 cells classification (sep 2013).
URL http://mivia.unisa.it/icip-2013-contest-on-hep-2-cells-classification/

[18] B. Lovell, G. Percannella, M. Vento, A. Wiliem, First workshop on pattern recognition techniques for iif images (Aug. 2014).
URL http://i3a2014.unisa.it/

[19] R. Stoklasa, T. Majtner, D. Svoboda, Efficient k-nn based HEp-2 cells classifier, Pattern Recognition 47 (7) (2014) 2409 – 2418.

[20] G. V. Ponomarev, V. L. Arlazarov, M. S. Gelfand, M. D. Kazanov, ANA HEp-2 cells image classification using number, size, shape and localization of targeted cell regions, Pattern Recognition 47 (7) (2014) 2360 – 2366.

[21] V. Snell, W. Christmas, J. Kittler, Hep-2 fluorescence pattern classification, Pattern Recognition 47 (7) (2014) 2338 – 2347.

[22] I. Theodorakopoulos, D. Kastaniotis, G. Economou, S. Fotopoulos, HEp-2 cells classification using morphological features and a bundle of local gradient descriptors, in: Pattern Recognition Techniques for Indirect Immunofluorescence Images (I3A), 2014 1st Workshop on, 2014, pp. 33–36.

[23] R. Nosaka, K. Fukui, Hep-2 cell classification using rotation invariant co-occurrence among local binary patterns, Pattern Recognition 47 (7) (2014) 2428 – 2436.

[24] X. Kong, K. Li, J. Cao, Q. Yang, L. Wenyin, HEp-2 cell pattern classification with discriminative dictionary learning, Pattern Recognition 47 (7) (2014) 2379 – 2388.

[25] D. Gragnaniello, C. Sansone, L. Verdoliva, Biologically-inspired dense local descriptor for indirect immunofluorescence image classification, in: Pattern Recognition Techniques for Indirect Immunofluorescence Images (I3A), 2014 1st Workshop on, 2014, pp. 1–5.

[26] S. Manivannan, W. Li, S. Akbar, R. Wang, J. Zhang, S. McKenna, HEp-2 specimen classification using multiresolution local patterns and svm, in: Pattern Recognition Techniques for Indirect Immunofluorescence Images (I3A), 2014 1st Workshop on, 2014, pp. 41–44.

[27] Z. Gao, J. Zhang, L. Zhou, L. Wang, HEp-2 cell image classification with convolutional neural networks, in: Pattern Recognition Techniques for Indirect Immunofluorescence Images (I3A), 2014 1st Workshop on, 2014, pp. 24–28.

[28] I. Ul Islam, S. Di Cataldo, A. Bottino, E. Macii, E. Ficarra, A preliminary analysis on hep-2 pattern classification: Evaluating strategies based on support vector machines and subclass discriminant analysis, in: Biomedical Engineering Systems and Technologies, Vol. 452 of Communications in Computer and Information Science, Springer Berlin Heidelberg, 2014, pp. 176–190.

[29] S. Di Cataldo, A. Bottino, I. U. Islam, T. F. Vieira, E. Ficarra, Subclass discriminant analysis of morphological and textural features for HEp-2 staining pattern classification, Pattern Recognition 47 (7) (2014) 2389 – 2399.

[30] S. Tonti, S. Di Cataldo, A. Bottino, E. Ficarra, An automated approach to the segmentation of HEp-2 cells for the indirect immunofluorescence ANA test, Computerized Medical Imaging and Graphics 40 (2015) 62 – 69.

[31] M. Olsen, D. Hartung, C. Busch, R. Larsen, Contrast enhancement and metrics for biometric vein pattern recognition., in: Advanced Intelligent Computing Theories and Applications - 6th International Conference on Intelligent Computing, ICIC, 2010, pp. 425–434.

[32] L. Vincent, Morphological grayscale reconstruction in image analysis: applications and efficient algorithms, Image Processing, IEEE Transactions on 2 (2) (1993) 176–201.

[33] J. C. Bezdek, Pattern Recognition with Fuzzy Objective Function Algorithms, Kluwer Academic Publishers, Norwell, MA, USA, 1981.

[34] Y. Xie, Q. Ji, A new efficient ellipse detection method, in: Pattern Recognition, 2002. Proceedings. 16th International Conference on, Vol. 2, 2002, pp. 957–960 vol.2.

[35] X. Han, Y. Fu, H. Zhang, A fast two-step marker-controlled watershed image segmentation method, in: Mechatronics and Automation (ICMA), 2012 International Conference on, 2012, pp. 1375–1380.

[36] S. Tonti, S. Di Cataldo, E. Macii, E. Ficarra, Unsupervised HEp-2 mitosis recognition in indirect immunofluorescence imaging, in: Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE, 2015, pp. 8135–8138.

[37] M. Zhu, A. M. Martinez, Subclass discriminant analysis, IEEE Trans. Pattern Anal. Mach. Intell. 28 (8).

[38] MIVIA Lab, HEp-2 images dataset (Jan. 2015).
URL http://mivia.unisa.it

[39] P. Foggia, G. Percannella, P. Soda, M. Vento, Early experiences in mitotic cells recognition on HEp-2 slides, in: Computer-Based Medical Systems (CBMS), 2010 IEEE 23rd International Symposium on, 2010, pp. 38–43.

[40] P. Foggia, G. Percannella, P. Soda, M. Vento, Benchmarking HEp-2 cells classification methods, Medical Imaging, IEEE Transactions on 32 (10) (2013) 1878–1889.

13