

# Combined Hierarchical Watershed Segmentation and SVM Classification for Pap Smear Cell Nucleus Extraction

Maykel Orozco-Monteagudo<sup>1</sup>, Cosmin Mihai<sup>2</sup>, Hichem Sahli<sup>2,3</sup>, and Alberto Taboada-Crispi<sup>1</sup>

<sup>1</sup> Centro de Estudios de Electrónica y Tecnologías de la Información (CEETI),  
Universidad Central de Las Villas, Santa Clara,  
Cuba

<sup>2</sup> Department of Electronics and Informatics (ETRO),  
Vrije Universiteit Brussel, Brussels,  
Belgium

<sup>3</sup> Interuniversity Microelectronics Centre (IMEC), Leuven,  
Belgium

{morozco, ataboada}@uclv.edu.cu, {cmihai, hsahli}@etro.vub.ac.be

**Abstract.** In this paper, we propose a two-phase approach to nuclei segmentation/classification in Pap smear test images. The first phase, the segmentation phase, includes a morphological algorithm (watershed) and a hierarchical merging algorithm (waterfall). In the merging step, waterfall uses spectral and shape information as well as the class information. In the second phase, classification, the goal is to obtain nucleus regions and cytoplasm areas by classifying the regions resulting from the first phase based on their spectral and shape features, merging of the adjacent regions belonging to the same class. Between the two phases, three unsupervised segmentation quality criteria were tested in order to determine the best one selecting the best level after merging. The classification of individual regions is obtained using a Support Vector Machine (SVM) classifier. The segmentation and classification results are compared to the segmentation provided by expert pathologists and demonstrate the efficacy of the proposed method.

**Keywords.** Microscopic images, cell segmentation, watershed, SVM.

## ***Extracción de núcleos de células en imágenes de la prueba de Papanicolaou usando watershed jerárquico y máquinas de vectores soporte***

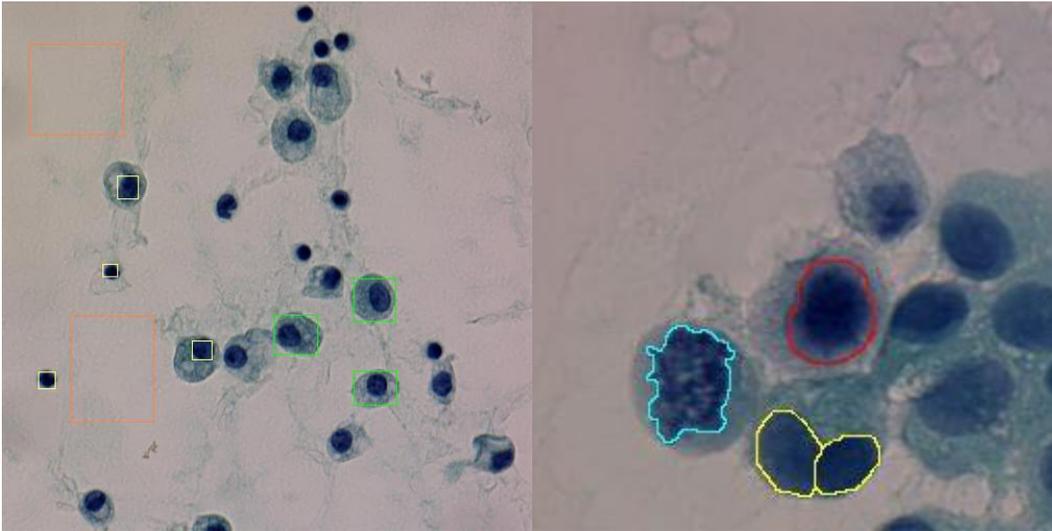
**Resumen.** En el presente trabajo se presenta un método en dos etapas para la segmentación y clasificación de núcleos de células en imágenes

tomadas de la prueba de Papanicolaou. La primera etapa, la etapa de segmentación, está formada por un algoritmo morfológico (watershed o marcas de agua) y un algoritmo jerárquico de mezclado (waterfall o salto de agua). Para realizar el mezclado de regiones, waterfall usa información espectral, de forma y de las regiones que se separarán. En la segunda etapa, la etapa de clasificación, el objetivo es obtener los núcleos a partir de las clasificaciones de las regiones obtenidas en la primera etapa. Antes de realizar la clasificación, fueron probadas tres medidas no supervisadas de calidad de la segmentación para determinar el mejor resultado de la mezcla de regiones. La clasificación de las regiones se realizó usando Máquinas de Vector Soporte. Los resultados fueron comparados con las segmentaciones realizadas por patólogos demostrando se la eficacia del método propuesto.

**Palabras clave.** Segmentación, imágenes microscópicas, segmentación de células, marcas de agua, máquinas de vector soporte.

## **1 Introduction**

Cervical cancer, currently associated with the Human Papilloma Virus (HPV) as one of the major risk factors, affects thousands of women each year. The Papanicolaou test (known as the Pap test) is used to detect pre-malignant and malignant changes in the cervix [22]. In the majority of cases, cervical cancer can be prevented by early detection of abnormal cells in smear tests [22]. The Pap smear test is the most



**Fig. 1.** Pap-smear cell images. (a) Dark blue parts (yellow squares) represent the nucleus. Pale blue parts (green squares) are the cytoplasm. Magenta parts (orange squares) are the background. (b) Touching cells (within the yellow borders). Color ranges from pale to dark blue (within the red border). Nuclei with a non-uniform shape are within the green borders

common method of cervical cancer screening. A pathologist should examine each slide of the Pap smear accurately. The procedure takes away most of the time and energy of pathologists while they review thousands of cells each day. One of the main steps to help the pathologist is to develop an automatic system that can segment and classify cells of the smear's scene in normal and abnormal groups.

Since the cervix is wiped out with a swap, the Pap test is classified as an invasive method. It is used only for screening purposes and not for diagnosis. These cells are examined under a microscope for abnormalities [22]. Trained biologists are required to evaluate these tests. In underdeveloped countries, the death rate due to cervical cancer is significantly higher because of the lack of personnel trained in this field and repeated follow-up tests. As a result, women in developed countries have less than 0.1% chance of developing cervical cancer while their counterparts in underdeveloped countries have a 3-5% chance of developing cervical cancer.

Nucleus extraction from the Pap smears is the first step for further detection of abnormalities or pre-cancerous cells. Two classes of regions are considered: nucleus regions and other regions

that include cytoplasm and background (Figure 1). The overall proportion of the nucleus pixels is approximately between 7% and 10%. Cell nuclei are blue (dark to pale) and cytoplasm is blue-green (Figure 1a). Red blood corpuscles are colored reddish. The spatial configuration and the color of the cells are extremely variable. Isolated or touching cells as well as clustered or overlapping cells can be found (Figure 1b). The automated segmentation of the cell nucleus in Pap smear images is one of the most interesting fields in cytological image analysis [21]. In the last years, cell nuclei segmentation has been extensively studied by several researchers. In [24], frequency domain features are used to detect abnormal cervical cell images. In [27], statistical geometric features, which are computed from several binary thresholded versions of texture images, are used to classify among normal and abnormal cervical cells. Lezoray and Cardot [17] extract the nucleus of the cervical cells using a combination of a color pixel classification scheme  $k$ -means and a Bayesian classification with a color watershed segmentation algorithm. In [2], a segmentation scheme and its performance are evaluated using Pap smear samples in the presence of heavy additive noise.

Developing automated algorithms for segmenting nuclei continues to pose interesting challenges. Much of the difficulty arises from the inherent color and shape variability. The goal of the present work is to develop automated and computationally efficient algorithms that improve upon previous methods using watershed approach [1]. In this work, we propose a hybrid two-step approach to cell segmentation in Pap smear images. The first phase consists of creating a nested hierarchy of partitions, which produces a hierarchical segmentation that uses the spectral and shape information as well as the class information. After constructing the hierarchy of nested partitions by performing waterfall algorithm [13] on the image, we automatically select the most meaningful hierarchical level by using a segmentation quality criterion. The second phase aims at identifying the nucleus and cytoplasm areas by classifying the segments (regions) resulting from the first phase using multiple spectral and shape features, and further merging the neighboring regions belonging to the same class. The selection of individual regions is obtained using a Support Vector Machine (SVM) classifier, based on spectral and shape features.

The remainder of the paper is organized as follows. Section 2.2 describes the segmentation algorithm used to segment the images and produce a hierarchy of nested partitions. Section 2.3 proposes an unsupervised segmentation quality criterion to select a set of hierarchical levels, on which SVM classification is applied in Section 2.4 to classify the segmented region in nucleus/non-nucleus and pruning the segmentation errors by merging the adjacent segmented regions which may have been over-segmented. Section 3 presents and discusses the obtained results. Finally, conclusions are presented in Section 4.

## 2 Materials and Methods

### 2.1 Pap Smear Cells Databases

In order to test the proposed algorithm, images from two Pap-smear cells databases were used. The first database is the image database used in [17], referred to as Lezoray database. In these

images, cells are colored using the international standard of coloration of Papanicolaou [22], and pixels in the ground truth are divided in two classes: nucleus pixels and other pixels. From this database, we used 20 images that contain approximately 160 nuclei.

The second database is the Herlev image database [14]. In these images, cells are also colored using the international standard of coloration of Papanicolaou, and pixels in the ground truth are divided in four classes: nucleus pixels, two kinds of cytoplasm pixels, and the background pixels. In our experiments, we consider only two classes: nucleus pixels and other pixels. The Herlev database contains seven kinds of images as described in Table 1. In our experiment, we used 140 images randomly selected, 20 per type.

### 2.2 Hierarchy of Partitions

The waterfall algorithm [4] is used here for producing a nested hierarchy of partitions,  $P^h = \{r_1^h, r_2^h, \dots, r_{m_h}^h\}; h = 1, \dots, n$ , which preserves the inclusion relationship  $P^h \supseteq P^{h-1}$ , implying that each atom of the set  $P^h$  is a disjoint union of atoms from the set  $P^{h-1}$ . For successively creating hierarchical partitions, the waterfall algorithm removes from the current partition (hierarchical level) all the boundaries completely surrounded by higher boundaries. The starting partition is obtained using watershed transform [25], being a morphological segmentation applied on the gradient magnitude of an image in order to guide the watershed lines to follow the crest lines and the real boundaries of the regions. In our implementation, we use the DiZenzo gradient [11], which calculates the maximum rate of change in one pixel based on partial derivatives in RGB color space.

For producing the nested hierarchy, in this work, we use the approach proposed in [13], where the saliency measure,  $E(\tilde{r} = r_i \cup_j | r_i, r_j)$ , of a boundary between two neighboring segments  $r_i$  and  $r_j$  (being the cost of merging the regions  $r_i$  and  $r_j$ ), is based on a collection of energy

**Table 1.** Herlev database. Classes 1-3 (242 cells) are normal cells and 4-7 (675 cells) are abnormal ones

Class	Category	Cell Type	Count
1	Normal	Superficial squamous epithelial	74
2	Normal	Intermediate squamous epithelial	70
3	Normal	Columnar epithelial	98
4	Abnormal	Mild squamous non-keratinizing dysplasia	182
5	Abnormal	Moderate squamous non-keratinizing dysplasia	146
6	Abnormal	Severe squamous non-keratinizing dysplasia	197
7	Abnormal	Squamous cell carcinoma in situ intermediate	150
Total			917

functions used to characterize the desired single-segment properties and the pair-wise segment properties [13].

$$E(\tilde{r} = r_i \cup r_j | r_i, r_j) = E(\tilde{r}) + E(r_i, r_j) \quad (1)$$

The single segment properties Eq. (1),  $E(\tilde{r})$ , includes segment homogeneity ( $E_{\text{hom}}$ ), segment convexity ( $E_{\text{conv}}$ ), segment compactness ( $E_{\text{comp}}$ ), and color variances ( $E_{\text{var}_c}$ ) within the segment.

$$E(\tilde{r}) = \frac{1}{E_{\text{hom}}(\tilde{r})} \sum_c E_{\text{var}_c}(\tilde{r}) \cdot (1 + |E_{\text{conv}}(\tilde{r})|)^{\text{sign}(E_{\text{conv}}(\tilde{r}))} \cdot (1 + |E_{\text{comp}}(\tilde{r})|)^{\text{sign}(E_{\text{comp}}(\tilde{r}))} \quad (2)$$

Taking into account the type of images we are dealing with, in this work we propose the following merging criterion:

$$E(\tilde{r} = r_i \cup r_j | r_i, r_j) = \phi(c_i = c_j | r_i, r_j) \cdot (E(\tilde{r}) + E(r_i, r_j)) \quad (3)$$

**Table 2.** Region-Based Features

Feature Type	Description
Morphometrics (69)	Generals (9): Major Axis Length, Minor Axis Length, Equivalent Diameter, Orientation, Area, Perimeter, Eccentricity, Convex Area, Solidity. Zernike moments (49): $Z_{i,j}$ , $i, j = 1..12$ , $i + j$ even. Skeleton features (4): skeleton length, skeleton length-convex hull area ratio, skeleton length-region area ratio, and skeleton branch length-skeleton length ratio. Convex hull features (3): eccentricity, shape factor, and fraction of overlap of the convex hull.
Textural-based (19)	Grey Level Co-occurrence Matrix based feature or Haralick features (13). Gabor features (6): $G_{f,\theta}$ , frequency $f \in \{0.06, 0.12, 0.24\}$ and angle $\theta \in \{0, \pi/2\}$
Edge-based (5)	Edge fraction of above-threshold pixels along the edge, measure of edge gradient intensity homogeneity, two measures of edge direction homogeneity, measure of edge direction difference.
Color-based (27)	Nine statistical measures (mean, 0.1-trimmean, max, min, median, standard deviation, interquartile range, skewness, and kurtosis) in the three channels of the RGB color space.

where  $\phi(c_i = c_j | r_i, r_j)$  is a factor favoring the merging of regions with similar classes [19];  $E(\tilde{r})$  is the merged region property as defined in [13], it depends on homogeneity, color variance, convexity, and compactness;  $E(r_i, r_j)$ , the pairwise region property, is defined as:

$$E(r_i, r_j) = -\log \left( \sum_{k=1}^b \sqrt{P_{r_i}^{(k)} \cdot P_{r_j}^{(k)}} \right) \quad (4)$$

being the Bhattacharyya merging criterion proposed in [7], with the number of bins used  $b=32$ .

Different from [19], the parameter  $\phi(c_i = c_j | r_i, r_j)$ , representing the potential of having neighboring regions with similar class membership, is here defined as follows:

$$\phi(c_i = c_j | r_i, r_j) = \frac{1}{1 + p(c_i = c_j | \mathbf{f}(r_i), \mathbf{f}(r_j))} \quad (5)$$

where

$c_i, c_j \in \Omega = \{\omega_1 = \text{nucleus}, \omega_2 = \text{no-nucleus}\}$ , are the classes of  $r_i$  and  $r_j$ , respectively, and  $p(c_i = c_j | \mathbf{f}(r_i), \mathbf{f}(r_j))$  is the probability that  $r_i$  and  $r_j$  belong to the same class, given the feature vectors  $\mathbf{f}(r_i)$  and  $\mathbf{f}(r_j)$ . In this work, this probability is estimated using Support Vector Machine (SVM) [10] classification trained using as feature vector  $\mathbf{f}(r)$ , the mean of the  $L, a$ , and  $b$  channels of the regions in the Lab color space. Unfortunately, we cannot directly use the output of the SVM as a probability measure. The output  $h(\mathbf{f}(r))$  of the SVM is a distance measure between a test pattern and the separating hyper plane defined by the support vectors. There is no clear relationship with the posterior class probability  $p(\text{class} = \omega | \mathbf{f}(r))$  that the pattern  $\mathbf{f}(r)$  belongs to the class  $\omega$ . Platt [23] proposed an estimate for this probability by fitting the SVM output  $h(\mathbf{f}(r))$  with a sigmoid function:

$$p(\text{class} = \omega_k | \mathbf{f}(r)) = \frac{1}{1 + \exp(A \cdot h(\mathbf{f}(r)) + B)} \quad (6)$$

The parameters  $A$  and  $B$  are found using maximum likelihood estimation from a training set. Finally, the probability in Eq. (5) is estimated as follows:

$$p(c_i = c_j | \mathbf{f}(r_i), \mathbf{f}(r_j)) = \frac{p_1 p_2}{p_1 p_2 + (1 - p_1)(1 - p_2)} \quad (7)$$

with  $p_1 = p(c_i = \omega_1 | \mathbf{f}(r_i))$  and  $p_2 = p(c_j = \omega_1 | \mathbf{f}(r_j))$  are estimated using Eq. (6).

The parameters of the SVM classifier have been selected as follows. A linear kernel SVM and Gaussian kernel SVMs (with different values for  $\sigma$ ) were trained using a 10-fold cross-validation. A grid search method was used to select the best parameters of the SVM. The penalty parameter of the error  $C$  was tested in  $C = \{2^i : i = -1..14, \infty\}$ , as well as the parameter of the Gaussian kernel  $\sigma$  in  $\sigma = \{2^i : i = -3..4\}$ . The best performance was obtained for  $C = 1024$  and Gaussian kernel SVM with  $\sigma = 0.5$ .

### 2.3 Segmentation Level Selection

In order to judge the quality of the image segmentation, three unsupervised criteria were tested: color homogeneity (CH) [12], the method of Borsotti (BOR) [6], and the method of Rosenberger (ROS) [26].

CH combines a region homogeneity measure with an inter-region contrast measure:

$$\text{CH}(P^h) = \frac{\sum_{j=1}^{m_h} C(r_j^h) / (m_h - 1)}{\sum_{j=1}^{m_h} H(r_j^h) / (\text{Card}(I) - m_h)} \quad (8)$$

where, for an image  $I$ ,  $\text{Card}(I)$  denotes the area of the full image (number of pixels).  $H(r_j^h)$  is the region homogeneity measure (i.e., local color error) of the region  $r_j^h$ , while  $C(r_j^h)$  is the inter-region separability measure (color difference between adjacent regions), it expresses a trade-off between the border accuracy of a region and the difference between the region and its neighbors.

$CH^h$  penalizes those regions with a large color error and a low border contrast. If the color error inside the region is low and the adjacent regions are significantly different (high border contrast), its value is high: the higher the value of  $CH^h$ , the better the segmentation result is. Using this criterion, the selection of the first hierarchy, often being oversegmented, is excluded.

The method of Borsotti (BOR) arose as an improvement of the method of Liu and Yang [18] in order to penalize the phenomenon of oversegmentation.

$$\text{BOR}(P^h) = \frac{\sqrt{m_h}}{10^4 \cdot \text{Card}(I)} \sum_{k=1}^{m_h} \left[ \frac{E_k^2}{1 + \log(\text{Card}(r_k^h))} + \left( \frac{\chi(\text{Card}(r_k^h))}{\text{Card}(r_k^h)} \right)^2 \right] \quad (9)$$

where  $\text{Card}()$  is the size (area) of a region  $r_k^h$  or the image  $I$ ;  $\chi(\text{Card}(r_k^h))$  is the number of regions having the same size (area) as region  $r_k^h$ ; and  $E_k$  is the sum of the Euclidean distances between the RGB color vector of the pixels of  $r_k$

and the color vector attributed to the region  $r_k$  in the segmentation result. Lower values of the Borsotti criterion mean a good segmentation. In order to harmonize with the other criteria, we used the value of  $1 - \text{BOR}(P^h)$  as a segmentation quality criterion.

The method of Rosenberger (ROS) combines intra- and inter-regions disparities. The intra-

region disparity is computed by the normalized standard deviation of gray levels in each region.

The inter-region disparity computes the dissimilarity of the average gray level of two neighbour regions in the segmentation result.

$$\text{ROS}(P^h) = \frac{1 - \frac{4}{255^2} m_h \cdot \sum_{i=1}^{m_h} \sigma^2(r_i^h)}{2} + \frac{2}{m_h(m_h - 1)} \sum_{\substack{i,j=1 \\ i \neq j}}^{m_h} \left( \frac{|\bar{g}_I(r_i^h) - \bar{g}_I(r_j^h)|}{512} \right) \quad (10)$$

where  $\bar{g}_I(r_j^h)$  is the average of the gray level of  $r_j^h$  and  $\sigma^2(r_j^h)$  is the variance of the gray-level of  $r_j^h$ . Higher values of the Rosenberger criterion mean a good segmentation.

Applied on the hierarchy of partitions  $P^h = \{r_1^h, r_2^h, \dots, r_{m_h}^h\}; h = 1..n$ , the above criteria produce, in an unsupervised manner, the best segmentation level. However, most of the time the produced partition does not correspond to the best segmentation level according to the biologists' criteria. Moreover, applying all the criteria together, they do not often agree on the best segmentation. A suitable approach is to prune the segmentation of the selected partition by merging adjacent regions belonging to the same class. Indeed, as depicted in Figure 2, the selected level shows a cell with two regions, after region-based classification and extra merging, the final segmentation/classification result has been refined.

## 2.4 SVM Region Classification

Support vector machines (SVM) have been proven to be powerful and robust tools for tackling classification tasks [10]. Mainly, SVM classification requires the following optimization problem

$$\max_{\lambda} = \left( \sum_{i=1}^N \lambda_i + \frac{1}{2} \cdot \sum_{1 \leq i, j \leq N} \lambda_i \lambda_j y_i y_j K(\mathbf{x}_i, \mathbf{x}_j) \right) \quad (11)$$

subject to

$$0 \leq \lambda_i \leq C, \text{ with } i=1..N \quad (12)$$

and

$$\sum_{i=1}^N \lambda_i y_i = 0 \quad (13)$$

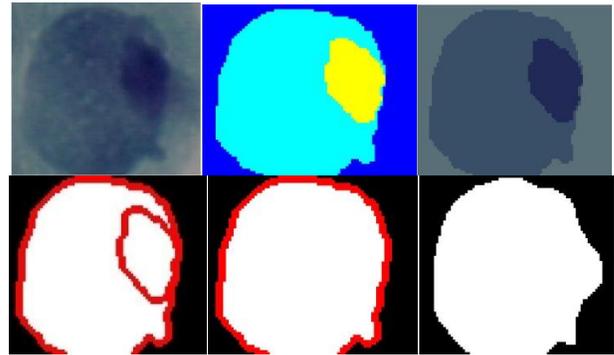
where  $N$  is the number of cases,  $y_i$  and  $y_j$  are 1 or -1 if  $\mathbf{x}_i$  and  $\mathbf{x}_j$  belong to one of the two classes  $\omega_1$  or  $\omega_2$ , respectively,  $K$  is the kernel, and  $C$  is a non-negative real number. The best selection of the kernel function is yet a challenging problem in the SVM community. In our work, we used a linear kernel, Gaussian kernels, and polynomial kernels [10].

Unlike mostly used SVM pixel-based classification, we propose to apply SVM on region (segment) features. In order to classify the regions obtained from the segmentation process, a wide set of region features were calculated (Table 2).

In an attempt to optimize the dimensionality of the feature set, a subset of features was selected from the feature set (Table 2) via stepwise discriminant analysis [15]. This method uses Wilks'  $\lambda$  statistic to iteratively determine which features are best able to separate the classes from one another in the feature space. Since it is not possible to identify a subset of features that are optimal for classification without training and testing classifiers for all combinations of the input features, optimization of Wilks'  $\lambda$  is a good choice.

The nine (out of 116) features that were statistically most significant in terms of their ability to separate the two classes identified using this approach are listed in Table 3. This set was used in the classification phase of the work.

Let  $\mathbf{r} = \{\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_n\}$  be a given region formed by the set of pixels  $\{p_{x_i, y_i}(r_i, g_i, b_i),$



**Fig. 2.** Merging after classification. (1) Original image. (2) Label image of the result after waterfall merging. (3) Mosaic image of the result after waterfall merging. (4) Classification result. White parts mean nucleus. (5) Neighbours regions that belong to the same class are merged into one. (6) Manually segmented ground truth

**Table 3.** Selected features using stepwise discriminant analysis

Feature	Description
$F_1$	Mean of the green channel
$F_2$	0.1-trim mean of the blue channel
$F_3$	Solidity
$F_4$	Maximum value of the red channel
$F_5$	Edge fraction of pixels along the edges
$F_6$	Edge gradient intensity homogeneity
$F_7$	Edge direction difference
$F_8$	Shape factor of the convex hull
$F_9$	Region area

$\{p_{x_2, y_2}(r_2, g_2, b_2), \dots, p_{x_n, y_n}(r_n, g_n, b_n)\}$ . The above features are estimated as follows:

**Mean of the Green Channel**

$$F_1(\mathbf{r}) = \frac{1}{n} \sum_{i=1}^n g_i \quad (14)$$

**Trimmed mean of the blue channel**

$$F_2(\mathbf{r}) = \frac{1}{0.9 \cdot n} \sum_{p_1 \leq b_i \leq p_2} b_i \quad (15)$$

where  $p_1 = p_b^{.05}$  and  $p_2 = p_b^{.95}$  are the .05 and .95 percentile of the blue channel, respectively.

**Solidity or convexity**

$$F_3(\mathbf{r}) = \frac{\text{Area}((\mathbf{r}))}{\text{Area}(\text{ConvexHull}((\mathbf{r})))} \quad (16)$$

**Max value of the red channel**

$$F_4(\mathbf{r}) = \max_{1 \leq i \leq n} \{r_i\} \quad (17)$$

**Edge fraction of pixels along the edges**

$$F_5(\mathbf{r}) = \frac{|\{p_i : p_i \in \text{Edge}\}|}{\text{Area}((\mathbf{r}))} \quad (18)$$

**Edge gradient intensity homogeneity**

Each region is convolved separately with the

kernels  $N = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{pmatrix}$  and  $W = N'$  to find

the intensity gradients in the two orthogonal directions of  $N$  ( $G_N$ ) and  $W$  ( $G_W$ ). The intensity of the gradient at all points in the image was calculated using  $A(x, y) = \sqrt{G_N^2 + G_W^2}$ , and a four-bin histogram was calculated for the values in this edge intensity image. The final feature is the fraction of all values that are in the first bin of this histogram.

**Edge direction difference**

For the eight-bin histogram used for measuring the edge direction homogeneity [5], the difference between the bins for an angle and for that angle plus  $\pi$  is calculated by summing bins 1-4 and subtracting the sum of bins 5-8. This difference was normalized by the sum of all eight bins.

**Convex hull shape factor or convex hull compactness**

$$F_8(\mathbf{r}) = \frac{4 \cdot \pi \cdot \text{Area}(\text{ConvexHull}((\mathbf{r})))}{\text{Perimeter}(\text{ConvexHull}((\mathbf{r})))^2} \quad (19)$$

**Region Area**

$$F_9(\mathbf{r}) = n \quad (20)$$

where  $n$  is the number of pixels in the region.

**3 Results and Discussion**

**3.1 Selection of the Segmentation Quality Criterion**

First, we evaluated the output of the three segmentation selection criteria of Section 2.3 for deciding the best one to be used for the type of images we are dealing with. The evaluation of the segmentation results has been made by using the Vinet measure proposed in [9]. Vinet is a supervised measure widely used to quantify the distance between two segmentations  $A$  and  $B$ , with  $m$  and  $n$  regions, respectively, of an image with  $N$  pixels. First, a label superposition table is computed:  $T_{ij} = |A_i \cap B_j|$  with  $0 \leq i \leq m$  and  $0 \leq j \leq n$ . The maximum of this matrix gives the two most similar regions extracted from  $A$  and  $B$ , respectively. The similarity criterion is defined by  $C_0 = \max(T_{ij})$  with  $0 \leq i \leq m$  and  $0 \leq j \leq n$ .

The research of the second maximum (without taking into account the two last regions) gives the similarity criterion  $C_1$  and so on to  $C_{k-1}$ , where  $k = \min(m, n)$ . The dissimilarity measure between the two segmentations  $A$  and  $B$  is given by

$$D(A, B) = 1 - \frac{1}{N} \sum_{i=0}^{k-1} C_i \quad (21)$$

When one of the segmentations ( $B$ , for example) is a ground truth, then lower values of the Vinet measure mean a good segmentation (close to the reference). Instead of the greedy approach to calculate the values of  $C_i$ ,  $i = 1.. \min(n, m)$ , the Hungarian algorithm [20] can be used. The Hungarian algorithm increases the computational cost but the accuracy of the measure will increase too.

Table 4 shows the evaluation results of the segmentation evaluation criteria applied on the images from the Lezoray database. The selected hierarchical level, provided by each of the evaluation criterions, has been compared to a manually delineated segmentation using the Vinet measure. From the segmentation point of view, the best results were obtained using the Borsotti criterion.

**Table 4.** Results for different level selection criteria

Level Selection Criterion	Vinet Measure
Method of Rosenberger	0.0281
Method of Borsotti	0.0223
Color Homogeneity	0.0252

### 3.2 Evaluation on the Lezoray Database

The proposed approach was applied to the twenty images that contain approximately 160 nuclei. The evaluation of the segmentations was done using a leave-one-out cross-validation approach with the Vinet distance with respect to the manually obtained segmentation. Training of the SVM was done using the SVM-KM toolbox [8].

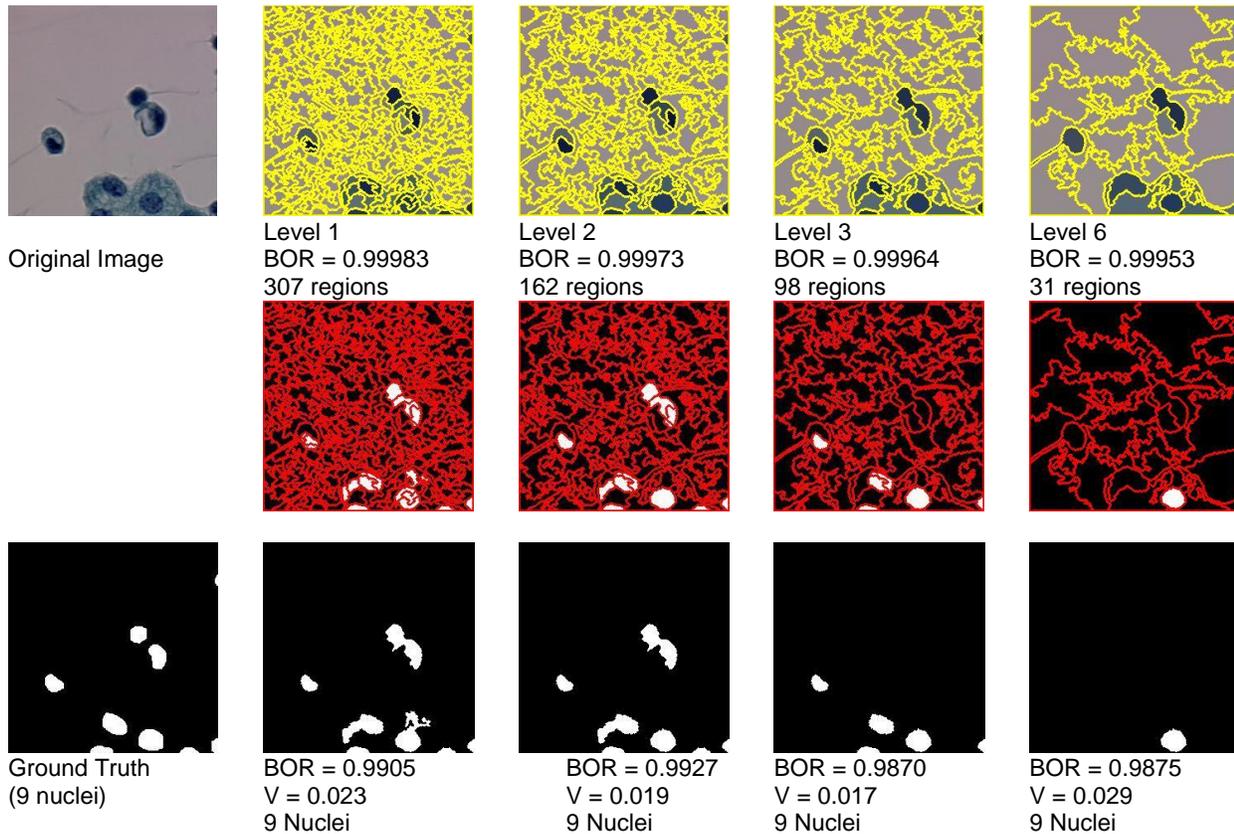
Figure 3 illustrates the proposed approach in one of the tested images. The first row depicts some hierarchical levels along with their BOR criterion and the number of regions. As it can be noticed, the hierarchical Level 1 is the best according to the BOR criterion. Moreover, the BOR criterion between the first three levels is almost identical. After SVM classification and

**Table 5.** Confusion matrices - SVM classification of the regions shown in Figure 3

Level		Non-Nucleus	Nucleus
1	Non-Nucleus	290 (100.0 %)	0 (0.00 %)
	Nucleus	4 (23.53 %)	13 (76.47 %)
2	Non-Nucleus	151 (100.0 %)	0 (0.00 %)
	Nucleus	1 (9.09 %)	10 (90.91 %)
3	Non-Nucleus	91 (98.91 %)	1 (1.09 %)
	Nucleus	0 (0.00 %)	6 (100.0 %)
6	Non-Nucleus	29 (96.67 %)	1 (3.33 %)
	Nucleus	0 (0.00 %)	1 (100.0 %)

**Table 6.** Lezoray database - Overall

SVM Classifier	Vinet Measure	Accuracy	F-measure
Linear Kernel	0.0223	0.9733	0.9853
Gaussian Kernel $\sigma = 0.5$	0.0494	0.9109	0.9587
Gaussian Kernel $\sigma = 1$	0.0456	0.9235	0.9644
Gaussian Kernel $\sigma = 2$	0.0407	0.9448	0.9704
Gaussian Kernel $\sigma = 4$	0.0323	0.9592	0.9781
Gaussian Kernel $\sigma = 8$	0.0274	0.9668	0.9821
Linear Kernel $d = 2$	0.0375	0.9546	0.9750
Linear Kernel $d = 3$	0.0528	0.9552	0.9758
CCW	0.0455	0.9571	0.6445
GEE	0.0243	0.9784	0.9880
SVMP	0.0356	0.9743	0.8601



**Fig. 3.** Illustration of the approach. (Upper-Left) Original Image. (Bottom-Left) Ground Truth. (Upper-Right) Four Hierarchical Levels. (Bottom-Right) Results after classification and second merging

merging of the neighboring regions belonging to the same class (the second and third rows of Figure 3), the hierarchical Level 2 gives better segmentation results with respect to the BOR criterion and the Vinet distance,  $V$ . Table 5 gives quantitative results of the segmentations of Figure 3 by the confusion matrixes. In Level 1, both nucleus and non-nucleus regions are oversegmented and the classification accuracy is the lowest. On the opposite, Level 6 is undersegmented but the classification accuracy is the best. The best compromise between segmentation and classification results is obtained in Level 3.

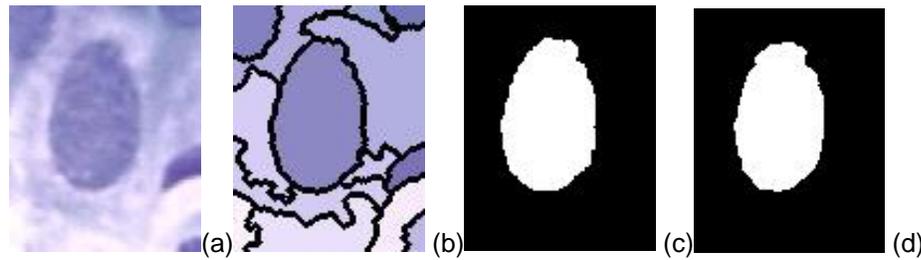
The averages of the Vinet measure, classification accuracy and F-measure over all the Lezoray images are summarized in Table 6. Following the Vinet measure, the best results were obtained using the SVM classifier with a

linear kernel. The segmentation and classification accuracy can be considered good due to the intra-expert distance between 0.01 and 0.04 [3]. Small and in-border nuclei are the principal sources of errors.

Table 7 shows the confusion matrix for the SVM classification performance using a linear kernel. From the object classification point of view, the obtained specificity (98.49%) greater than the sensibility (90.84%) means that the proposed classifier tends to classify as non-nucleus regions a nucleus regions but not the opposite.

**Table 7.** Lezoray database - Confusion Matrix

	Non-Nucleus	Nucleus
Non-Nucleus	98.49%	1.51%
Nucleus	9.16%	90.84%



**Fig. 4.** Segmentation and classification results for an image from the Herlev database. (a) Original image. (b) Mosaic image after segmentation. (c) Nucleus region classification. (d) Ground truth

To further assess our results, we give in the last rows of Table 6 the results obtained using three state of art methods, namely, the Cooperative color watershed proposed in [17] (CCW), the Hierarchical Segmentation of [13] (GEE), and a Pixel-based SVM classification (SVMP) [10]. As it can be seen from Table 6, the proposed approach produces good results.

### 3.3 Evaluation on the Herlev Database

The evaluation of the approach using different SVM kernels has been also made on the Herlev database as shown in Table 8.

In order to compare with the results reported in [16] (KAL), the Zijdenbos measure [28] was used to evaluate the segmentation quality with respect to the ground truth. The Zijdenbos measure is defined as the ratio of twice the common area between two regions to the sum of the individual areas. When the Zijdenbos measure is greater than 0.7, there is an excellent agreement between

**Table 8.** Herlev database - Overall evaluation

SVM Classifier	Zijdenbos Measure	Accuracy	F-measure
Linear Kernel	0.7152	0.9385	0.9663
Gaussian Kernel $\sigma = 1$	0.9325	0.9978	0.9988
Gaussian Kernel $\sigma = 10$	0.8046	0.9655	0.9812
KAL	0.9232		

both segmentations. As it can be seen, we obtain comparable results to the ones applied to the Lezoray database. Good results were obtained in all the approaches but very good results were achieved by using a Gaussian kernel with unitary variance.

## 4 Conclusions

In this work, we tested a hybrid segmentation-classification approach to automatically select and improve the segmentation of nucleus of cells in images of the Papanicolaou test using nested hierarchical partitioning, segmentation level selection, and an SVM classifier for the final merging to avoid over-segmentation. The segmentation was done by a morphological algorithm (watershed) and a hierarchical merging algorithm (waterfall) based on spectral information and shape information, as well as the class information. A SVM classifier was used to separate two classes of regions, nucleus and not nucleus regions (cytoplasm and background) using appropriate set of features (morphometric, edge-based, and convex hull-based). The proposed approach allows pruning most of the wrongly segmented cells avoiding over/under segmentation and produces segmentation closer to what is expected by expert pathologists.

## References

1. **Angulo, J. & Serra, J. (2005).** Segmentación de imágenes en color utilizando histogramas bi-variables en espacios color polares

- luminancia/saturación/matiz. *Computación y Sistemas*, 8(4), 303–316.
2. **Bak, E., Najarian, K., & Brockway, J.P. (2004).** Efficient segmentation framework of cell images in noise environments. *26<sup>th</sup> Annual International Conference of the IEEE Engineering in Medicine and Biology Society (IEMBS '04)*, San Francisco, California, USA, 1802–1805.
  3. **Bergen, T., Steckhan, D., Wittenberg, T., & Zerfass, T. (2008).** Segmentation of leukocytes and erythrocytes in blood smear images. *30<sup>th</sup> Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS 2008)*, Vancouver, Canada, 3075–3078.
  4. **Beucher, S. & Meyer, F. (1993).** The morphological approach to segmentation: the watershed transformation. *Mathematical Morphology in Image Processing* (433–482). New York: M. Dekker.
  5. **Boland, M.V. & Murphy, R.F. (2001).** A neural network classifier capable of recognizing the patterns of all major subcellular structures in fluorescence microscope images of HeLa cells. *Bioinformatics*, 17(12), 1213–1223.
  6. **Borsotti, M., Campadelli, P., & Schettini, R. (1998).** Quantitative evaluation of color image segmentation results. *Pattern Recognition Letters*, 19(8), 741–747.
  7. **Calderero, F. & Marques, F. (2008).** General region merging approaches based on information theory statistical measures. *15<sup>th</sup> IEEE International Conference on Image Processing (ICIP 2008)*, San Diego, CA, USA, 3016–3019.
  8. **Canu, S., Grandvalet, Y., & Rakotomamonjy, A. (2003).** SVM and Kernel Methods MATLAB Toolbox. *Perception de Systemes et Information*, INSA de Rouen, France.
  9. **Cohen, L., Vinet, L., Sander, P.T., & Gagalowicz, A. (1989).** Hierarchical regional based stereo matching. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'89)*, San Diego, CA, USA, 416–421.
  10. **Cristianini, N. & Shawe-Taylor, J. (2000).** *An Introduction to Support Vector Machines and other kernel-based learning methods*. Cambridge; New York: Cambridge University Press.
  11. **Geerinck, T., Mihai, C., Jabloun, M., Vanhamel, I., & Sahli, H. (2009).** Multi-scale image analysis of satellite data using perceptual grouping. *IEEE International Geoscience and Remote Sensing Symposium*.
  12. **Geerinck, T., Sahli, H., Henderickx, D., Vanhamel, I., & Enescu, V. (2009).** Modeling Attention and Perceptual Grouping to Salient Objects. *Attention in Cognitive. Lecture Notes in Computer Science*, 5395, 166–182.
  13. **Jantzen, J., Norup, J., Dounias, G., & Bjerregaard, B. (2005).** Pap-smear Benchmark Data for Pattern Classification. *Nature inspired Smart Information Systems (NiSIS 2005)*, Albufeira, Portugal, 1–9.
  14. **Jennrich, R.I. & Sampson, P. (1960).** Stepwise discriminant analysis. *Mathematical methods for digital computers*, 339–358.
  15. **Kale, A. & Aksoy, S. (2010).** Segmentation of Cervical Cell Images. *20<sup>th</sup> International Conference on Pattern Recognition (ICPR)*. Istanbul, Turkey, 2399–2402.
  16. **Lezoray, O. & Cardot, H. (2002).** Cooperation of color pixel classification schemes and color watershed: a study for microscopic images. *IEEE Transactions on Image Processing*, 11(7), 783–789.
  17. **Liu, J. & Yang, Y.H. (1994).** Multiresolution color image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(7), 689–700.
  18. **Lucchi, A., Smith, K., Achanta, R., Lepetit, V., & Fua, P. (2010).** A Fully Automated Approach to Segmentation of Irregularly Shaped Cellular Structures in EM Images. *13<sup>th</sup> International Conference on Medical Image Computing and Computer-Assisted Intervention, Part II (MICCAI'10)*, Beijing, China, 463–471.
  19. **Munkres, J. (1957).** Algorithms for the assignment and transportation problems. *Journal of the Society for Industrial and Applied Mathematics*, 5(1), 32–38.
  20. **Pantanowitz, L., Hornish, M., & Goulart, R.A. (2010).** The impact of digital imaging in the field of cytopathology. *Cytojournal*, 6(6).
  21. **Papanicolaou, G.N. & Traut, H.F. (1943).** Diagnosis of uterine cancer by the vaginal smear. *Yale Journal of Biology and medicine*, 15(6), 924.
  22. **Platt, J.C. (2000).** Probabilities for SV Machines. *Advances in large margin classifiers* (61–74), Cambridge, Mass.: MIT Press.

23. **Ricketts, I.W., Banda-Gamboa, H., Cairns, A.Y., & Hussein, K. (1992).** Automatic classification of cervical cells-using the frequency domain. *IEE Colloquium on Applications of Image Processing in Mass Health Screening*, London, UK, 9/1–9/4.
24. **Roerdink, J.B.T.M. & Meijster, A. (2000).** The watershed transform: Definitions, algorithms and parallelization strategies. *Fundamenta Informaticae- Special issue on Mathematical morphology*, 41(1-2), 187–228.
25. **Rosenberger, C. (1999).** Mise en oeuvre d'un système adaptatif de segmentation d'images. Thèse de doctorat, Université de Rennes 1, Rennes, Francia.
26. **Walker, R.F. & Jackway, P.T. (1996).** Statistical geometric features-extensions for cytological texture analysis. *13<sup>th</sup> International Conference on Pattern Recognition*, Vienna, Austria, 2, 790–794.
27. **Zenzo, S.D. (1986).** A note on the gradient of a multi-image. *Computer Vision, Graphics, Image Processing*, 33(1), 116–125.
28. **Zijdenbos, A.P., Dawant, B.M., Margolin, R.A., & Palmer, A.C. (1994).** Morphometric analysis of white matter lesions in MR images: method and validation. *IEEE Transactions on Medical Imaging*, 13(4), 716–724.



**Mykel Orozco Monteagudo**

received his B.S. and M. Sc. degrees in Computer Sciences from Universidad Central de Las Villas (UCLV), Cuba, in 2002 and 2006, respectively. Since 2005, he

has been in the Center of Studies on Electronic and Information Technologies (CEETI) as an Auxiliary Professor. His research interests are image processing and pattern recognition for biomedical applications.



**Cosmin Mihai**

received his Software Engineering degree from Universitatea Politehnica din București in 2003. At present, he works as a Software Design Engineer at Honeywell Aerospace, Brno, Czech Republic. His main

research interests include computational vision,

image and motion segmentation, and machine learning.



**Hichem Sahli**

received his Ph.D. degree in Computer Sciences from the Ecole Nationale Supérieure de Physique Strasbourg, Strasbourg, France, in 1991. He joined the Department of CAD and Robotics of the Ecole

des Mines de Paris in 1992. In 1999, he moved to the Vrije Universiteit Brussel, Brussels, Belgium, where currently he is a Professor at the Department of Electronics and Informatics. He coordinates the research team in Computer Vision and Audio Visual Signal Processing. His main research interests include computational vision, image and motion segmentation, machine learning, visual reconstruction, and optimization. Dr. Sahli is a member of IEEE and ACM.



**Alberto Taboada-Crispi**

is a graduate of Electronic Engineering (Universidad Central de Las Villas, UCLV, 1985), Master in Electronics (UCLV, 1997), and Ph.D. (University of New Brunswick, Canada, 2002). He worked as

a Specialist for Medical and Laboratory Equipment in the Villa Clara Provincial Hospital from 1985 to 1988. Since then, he has been with the Faculty of Electrical Engineering (FIE) of the UCLV. Currently, he is a Professor and Researcher, and the Director of the Center for Studies on Electronics and Information Technologies (CEETI), FIE, UCLV. His topics of interest include instrumentation, analog and digital processing of signals and images, and biomedical applications.

*Article received on 30/04/2011; accepted on 29/02/2012.*