Title
Radial feature descriptors for cell classification and recommendation

Authors
Silva, Romuere RV
Araujo, Flavio HD
Ushizima, Daniela M
et al.

Publication Date
2019-07-01

DOI
10.1016/j.jvcir.2019.04.012

Peer reviewed

Radial Feature Descriptors for Cell Classification and Recommendation

Romuere R. V. Silva[a,b], Flavio H. D. Araujo[a,b], Daniela M. Ushizima[c,d,*], Andrea G. C. Bianchi[e], Claudia M. Carneiro[e], Fatima N. S. Medeiros[a]

[a] *Federal University of Ceará, Fortaleza, CE, Brazil*
[b] *Federal University of Piauí, Picos, PI, Brazil*
[c] *University of California, Berkeley, CA, USA*
[d] *Lawrence Berkeley National Laboratory, Berkeley, CA, USA*
[e] *Federal University of Ouro Preto, Ouro Preto, MG, Brazil*

Abstract

This paper introduces computational tools for cell classification into normal and abnormal, as well as content-based-image-retrieval (CBIR) for cell recommendation. It also proposes the radial feature descriptors (RFD), which define evenly interspaced segments around the nucleus, and proportional to the convexity of the nuclear boundary. Experiments consider Herlev and CRIC image databases as input to classification via Random Forest and bootstrap; we compare 14 different feature sets by means of False Negative Rate (FNR) and Kappa (κ), obtaining FNR= 0.02 and $\kappa = 0.89$ for Herlev, and FNR= 0.14 and $\kappa = 0.78$ for CRIC. Next, we sort and rank cell images using convolutional neural networks and evaluate performance with the Mean Average Precision (MAP), achieving MAP= 0.84 and MAP= 0.82 for Herlev and CRIC, respectively. Cell classification show encouraging results regarding RFD, including its sensitivity to intensity variation around the nuclear membrane as it bypasses cytoplasm segmentation.

Keywords: radial feature descriptors, cell classification, image retrieval, convolutional neural networks

*Corresponding author: dani.lbnl@berkeley.edu; phone: +1 510 486-4061
Email addresses: romuere@ufpi.edu.br (Romuere R. V. Silva), flavio86@ufpi.edu.br (Flavio H. D. Araujo), dani.lbnl@berkeley.edu (Daniela M. Ushizima), andrea@iceb.ufop.br (Andrea G. C. Bianchi), carneirocm@gmail.com (Claudia M. Carneiro), fsombra@ufc.br (Fatima N. S. Medeiros)

1. Introduction

Cervical cancer affects women globally and it is one of the leading causes of female death by cancer in developing countries. In the 1970s, cervical cancer was one of the most common causes of American women death by cancer. Recently, the death rate has gone down by more than 50%, with the Pap test being the main reason for this improvement [1].

Worldwide, the Pap test is the most commonly used method for the early identification [2] of precancerous lesions during primary screening. The examination relies on a cytologist who visually searches for abnormal cells, focusing on features associated with morphological alterations, such as the cytoplasm and nucleus sizes, chromatin distribution in the cell nuclei, the shape of cell clumps, and the nucleus/cytoplasm ratio. Cell screening remains the most common approach worldwide, however the reliance upon manual examination by different pathologists hinders the ability of public health programs to scale to the population growth. The major challenge in automating cervical cancer screening has been the use of large image databases from conventional Pap tests as input to segmentation algorithms. Particularly, the cytoplasm segmentation of overlapping cells poses substantial uncertainty during cell analysis.

This paper proposes a computational approach to address limitations of previous cell recognition systems, and Figure 1 summarizes the proposed workflow for cervical cell image analysis. We perform classification and retrieval tests on digitized microscopy from conventional Pap smears coming from two datasets: (a) Herlev [3], a public database of cervical cells that includes masks for nuclear and cytoplasmic areas; (b) the Cell Recognition for the Inspection of the Cervix (CRIC) database, which contains high-resolution labeled digitized micrographs.

Our cell recognition algorithms provide color-based cell segmentation, as well as cell classification and image retrieval. In order to detect the nuclear area for CRIC images, we propose a segmentation algorithm for nuclei detection, which combines 2 unsupervised algorithms: mean shift and k-means. We also

developed a new algorithm for radial feature descriptions (RFD), as illustrated in Figure 1(c); this cell feature can bypass the cytoplasm boundary detection in order to capture information inside and around the nucleus for cytoplasm characterization. The advantage of this approach is that it uses narrow bands around the nucleus, which targets the cytoplasm texture while remaining independent of its boundaries. Because the cytoplasm edge detection of overlapping cells is an ill-defined and computationally intense task [4, 5], alternatives to such a step can enable real-time analysis, and improve classification accuracy.

We compare RFD with other 13 different feature sets, based on previously published work [6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18]; among them, 7 feature sets were designed specifically for cell classification [6, 7, 8, 9, 10, 11, 12]. Ultimately, features sets consist of the concatenation of descriptors provided by each feature extraction method. Using images from both databases, each feature set is input to a Random Forest algorithm [19] that classifies cells into normal or abnormal. In addition, we also perform Content-Based Image Retrieval (CBIR) experiments to sort and rank cells under cosine similarity as a key step into enabling cell recommendation systems.

The main contributions of our work are: a) design RFD as a set of descriptors to quickly and accurately profile cells, without prior cytoplasm segmentation; b) catalog digital micrographs of high-resolution together with ground-truth; c) classification, sorting and ranking results using 14 different recognition approaches, which includes two convolutional neural networks.

Section 2 discusses the related studies, focusing on reviewing algorithms for cell segmentation and descriptors. Section 3 introduces the cell databases used to test the algorithms. Section 4 describes the proposed color-based nuclei segmentation technique to generate the ground truth (masks) for the CRIC database. In Section 5, we introduce the radial feature descriptors for cervical cell images. Meanwhile, Section 6 describes the methodology that we used to classify abnormal and normal nuclei from both cell image data sets. Section 7 discusses the classification and CBIR experiments and compares the results from different cell descriptors. Finally, Section 8 summarizes the main findings and

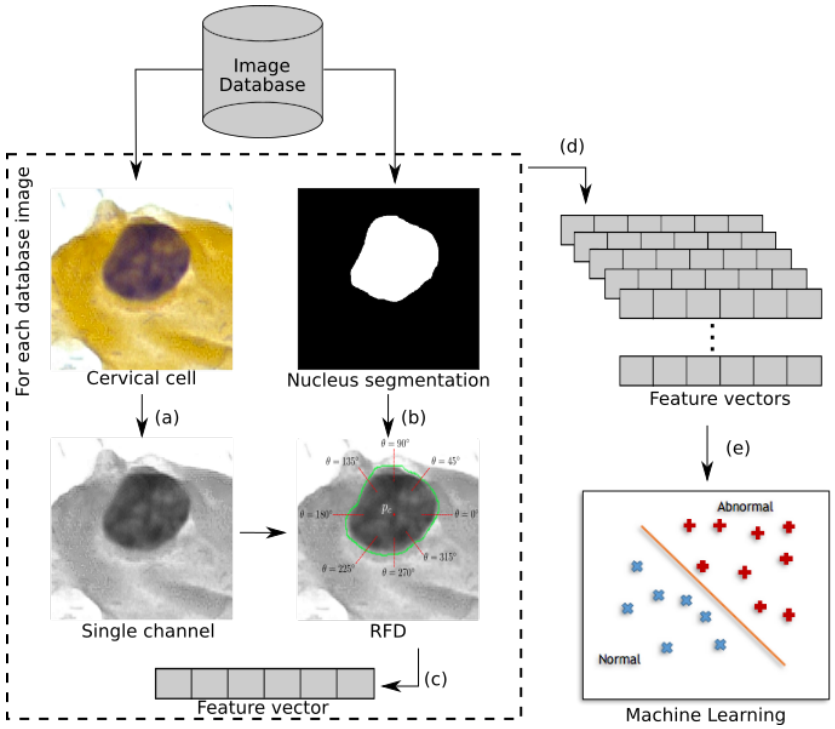


Figure 1: From raw micrographs to cell classification: (a) color transformation, (b) border detection, (c) intensity spectrum, (d) batch processing, and (e) Random Forest classification.

conclusions of this paper.

2. Related Work

2.1. Automating Pap Smear Analysis

Automated detection systems for Pap smear images either aim to separate cells into two classes, normal and abnormal, or they perform multiclass classification of the samples, according to the lesion levels. There is a vast literature on the methods used for automatic cervical cell analysis [20], mostly focusing on cell segmentation and classification tasks, detecting nucleus and cytoplasm. However, highly accurate results of cytoplasm segmentation have been mostly restricted to synthetic images [4, 21, 22, 23] and real non-overlapping cells [5].

Improvements on real cells segmentation [5, 24] have shown new potential to analyze cervical cell using computer-aided systems for Pap tests.

Even when manual or semi-automatic delineation drives the cell segmentation, shape descriptors can be automatically calculated to quantify the cell size, nucleus-cell area ratio, roundness, and elongation. Shape features have been used to classify cells into normal and abnormal patterns [11]; however, these methods often behave accurately for only a small number of cells, and they rely on the segmentation of both the nucleus and cytoplasm boundaries. In addition, texture features are useful to represent chromatin distribution patterns in cell nuclei where atypical concentrations are often associated with cancer cells [25].

2.2. Nucleus Segmentation

Most description methods for Pap smear images have focused on nucleus segmentation [5, 20, 26]. Li et al. [5] introduced a method that depends on converting cell images to the CIELAB color space. Next, it extracts the L component, and then applies the non-local mean filter to reduce background noise, which provides enhanced input data for the k-means clustering algorithm to extract the initial contours of the nucleus. These rough contours initialize the Radiating Gradient Vector Flow (RGVF) Snake [5] to estimates more accurate nucleus boundaries. Experimental results showed highly accurate pixel classification, which achieved 0.9197% of Zijdenbos similarity index (ZSI) [27] for the Herlev database.

Alternative unsupervised machine learning schemes in [21, 26] proposed graph-based algorithms to segment the nucleus and cytoplasm of cervical cells with different degrees of overlap in seconds. Despite those methods continuing to be among the best algorithms, as discussed in [24], accuracies drop drastically when applied to conventional Pap smears. Section 4 presents unsupervised machine learning schemes that evolved from this previous work and addresses digitized Pap tests from real-world scenarios.

2.3. Cell Description

Previous research on cervical cell analysis consider information about nuclei [11], cytoplasm or both [25, 10]. Promising algorithms to identify cells using shape [11, 10] and texture features [28, 29] continue to be published, with more recent proposals constructing feature vectors using hybrid feature sets [24, 21, 25, 30, 9].

Marinakis et al. [10] extracted a set of 20 features from the nucleus and cytoplasm for classification experiments using the Herlev database. This set of features comprises shape attributes, such as area, diameter, and elongation, in addition to intensity attributes, such as brightness, maxima and minima; these are calculated using the maximum/minimum intensity value within a 3x3 neighborhood of a specific area. The feature selection in [10] used a genetic algorithm to search for the best performing subset, which achieved FN and FP of 2.66% and 10.74%, respectively, when labeling cells into two classes using the 10-fold cross-validation and Nearest Neighbor.

In contrast, Plissiti and Nikou [11] extracted a set of 9 shape and intensity features from the nucleus region, such as area, diameter, elongation, brightness, maxima, and minima. These authors achieved a harmonic mean (H-mean) of the sensitivity and the specificity of 0.74 in classifying cells into two classes using Fuzzy C-Means on the Herlev database.

Bejnordi et al. [25] introduced a set of structural texture features to quantify nuclear chromatin patterns in cells from conventional Pap smears. The results of the feature selection showed that the structural texture feature was the most relevant in the classification experiments. When combining structural and conventional features, the best classification performance was 95.4% for normal and abnormal cell discrimination.

The combined use of shape and texture features was also reported by Mariarputham et al. [9], who used 7 features sets that included the relative size of the nucleus and cytoplasm, the dynamic range, and the first 4 moments of the intensities, relative displacement of nucleus within the cytoplasm, Gray Level Co-occurrence Matrix (GLCM) [13], Local Binary Pattern histogram (LBP) [16],

Tamura features [31], and Edge Orientation Histogram (EOH) [32]. Next, they applied support vector machines (SVM) to the images from the Herlev database and achieved a precision of 97.38% for normal squamous, 93.89% for intermediate squamous, 86.90% for columnar, 87.33% for mild dysplasia, 58.52% for severe dysplasia, 84.72% for carcinoma, and 83.62% for moderate dysplasia.

Previous investigations show that textural information plays a key role in the description of biological data, with applications ranging from face recognition [33, 34, 35] to cell classification [4, 36, 37] of cervical cells from Pap smears, which motivates our proposal of a new descriptor based on the nucleus edge; it computes the texture information around the nuclear membrane using images from the Herlev and CRIC databases. Despite the presence of many different scientific publications, they lack a common metric to compare the cell analysis performance, therefore we propose the evaluation of 14 different cell description methods using the same classifier and the same performance metrics: false negative rate and the κ index.

3. Materials

3.1. Herlev Database

Researchers from the Herlev University Hospital, Denmark, created a public database known as Herlev [3], which consists of 917 labeled single cells from Pap tests, acquired at a magnification of 0.201 μm/pixel. The Herlev database

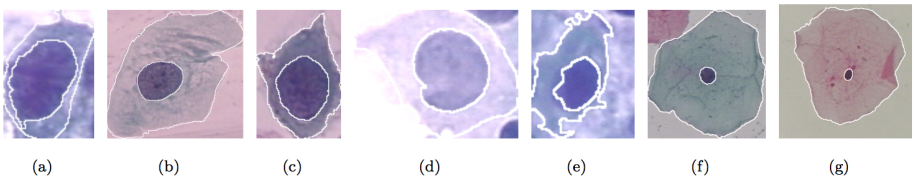


Figure 2: Cervical cell samples from the Herlev image database. (a) Intermediate squamous cell carcinoma in situ. (b) Mild squamous non-keratinizing dysplasia. (c) Moderate squamous non-keratinizing dysplasia. (d) Severe squamous non-keratinizing dysplasia. (e) Columnar epithelial. (f) Intermediate squamous epithelial. (g) Superficial squamous epithelial. White boundaries are masks that correspond to the nucleus and cytoplasm.

is divided into seven cell types and two classes: 675 abnormal cell images (squamous cell carcinoma in situ intermediate - 150, mild squamous non-keratinizing dysplasia - 182, moderate squamous non-keratinizing dysplasia - 146, and severe squamous non-keratinizing dysplasia - 197) and 242 normal cell images (columnar epithelial cells - 98, intermediate squamous epithelial cells - 70, and superficial squamous epithelial cells - 74).

This database has been widely used for the development and tests of new cervical cell classification systems [11, 10, 28, 9, 6, 12, 8]. Figure 2 shows image samples from the Herlev database: the white boundaries highlight the ground-truth images segmented by a specialist.

3.2. CRIC Database

The CRIC database[1] contains more than 2,000 cervical cells from 169 digitized Pap smear glass slides, each with 1,392×1,040 pixels, acquired with a Carl Zeiss microscope with a Zeiss AxioCam MRc camera at 40× magnification, generating images of 0.255 μm/pixel. Figure 3 displays a sample image from the CRIC database. These anonymized images are collected by providers of the Brazilian Universal Health-care System known as S.U.S., and made available to our team through the Science without Borders program.

Similar to the Herlev, the CRIC collection contains cells from routine conventional Pap smears. In addition, CRIC has unique and essential characteristics, such as including overlapping cells, debris, and other findings that are inherent to Pap tests. The CRIC database samples come from a broad racial diversity, which is a trace of the Brazilian population. Although the CRIC database has labels for each normal and abnormal cell, they lack nuclei segmentations. Consequently, this paper also includes our nucleus segmentation method (Figure 4), which automates the mask image generation for the CRIC.

[1]Original cervical cell images will be available upon paper acceptance at http://bit.ly/centercric.

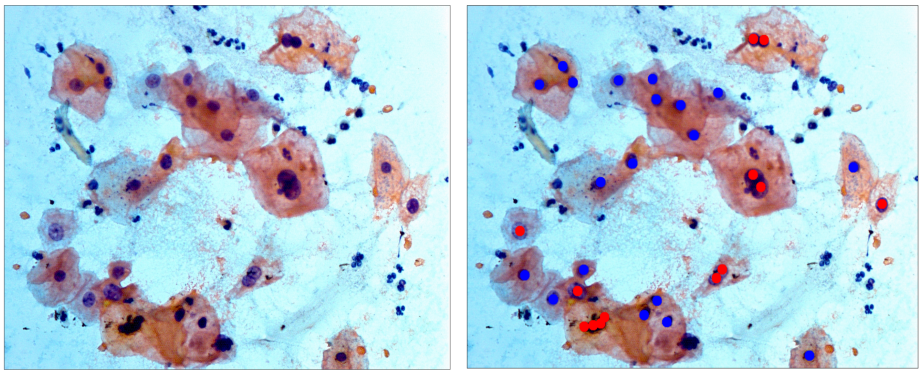


Figure 3: CRIC database image with cell samples: (a) original image, and (b) information about nuclei positions and classification. Blue points indicate the nucleus of normal cells and red spots indicate abnormal cells.

4. Proposed Cell Segmentation from Color Images

Our automated cell segmentation algorithm has three main steps: region clustering, ranking of cell regions, and cell individualization. Using the CRIC database, we extract the green channel from the RGB color images as this is the most correlated to the nuclear materials: the dying protocol tags nucleus in purple, which is a mixture of red and blue [38]. Our experiments include tests that handle several color models, including Lab, RGB, and HSV[39].

The first step relies on the mean shift pre-processing algorithm to promote pixel intra-class homogeneity, followed by a rough pixel grouping that uses the k-means algorithm ($k = 2$) for region clustering based on color features. The next image transformation uses a morphological opening operation for noise reduction associated to small debris with a circular structural element of size equals to 10 pixels.

The following step extracts shape features from pre-classified regions and then separates these regions into clumps (cell clusters), nucleus candidates, and artifacts (e.g., white blood cells such as neutrophils). The clumps are all regions with area greater than 4,000 pixels, which return to the region clustering step. The nuclei candidates are all regions with area and compactness greater than

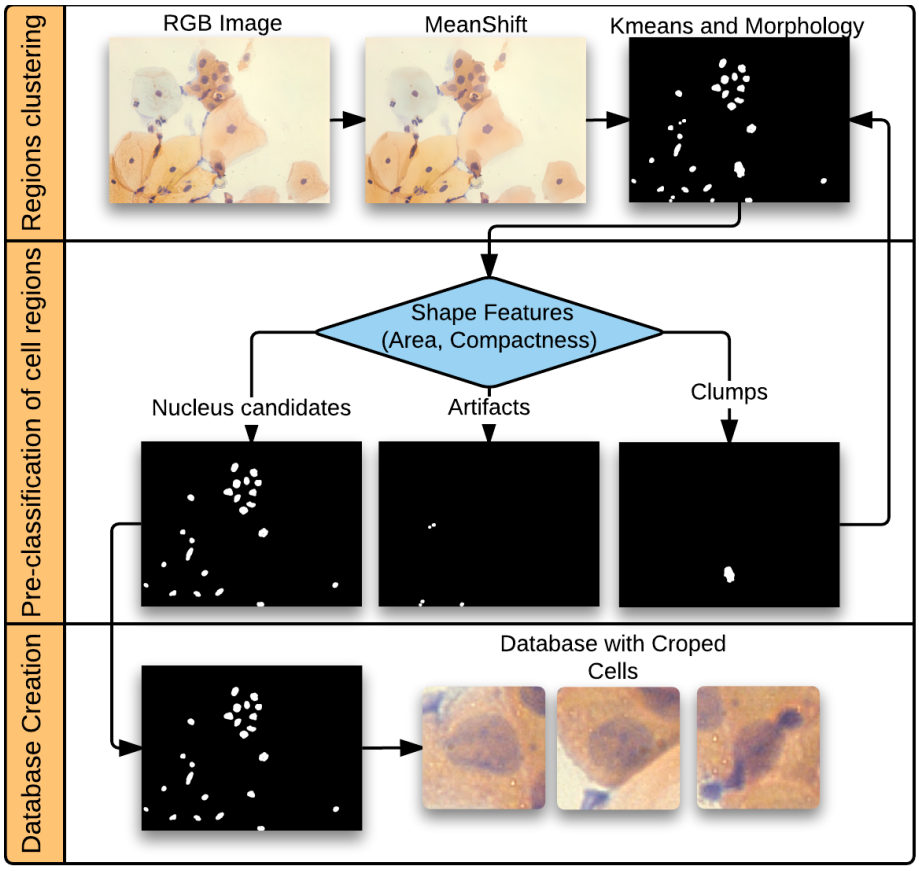


Figure 4: Flowchart of the proposed segmentation methodology. Our method has three steps: region clustering, pre-classification of region cells; and database creation.

600 pixels and 0.3, respectively; all nuclei candidates are inputs for the cell individualization step. The artifacts are the regions that do not follow the previous conditions, therefore they are removed.

The last step detects individual cells and creates a database of single cells, with a cropped tile centered around the corresponding nucleus.

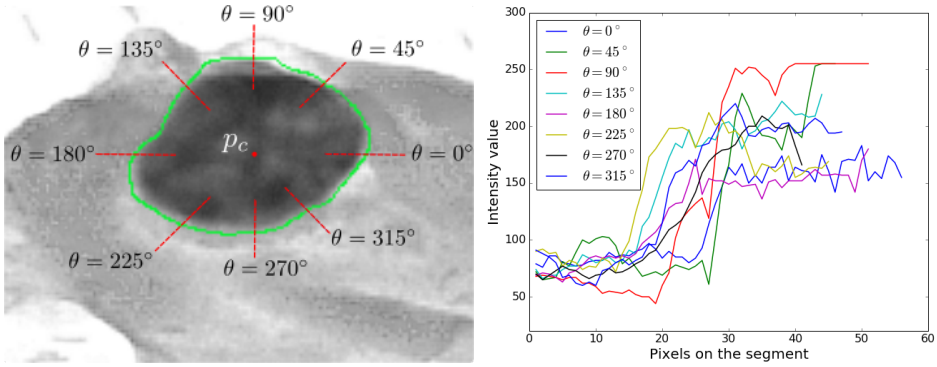
Figure 4 illustrates the main steps of the segmentation process, and the Algorithm 1 shows additional details of each step. After the segmentation, images are organized into a test set that comprises masks of 1,004 abnormal and 1,466 normal cell images from the CRIC database.

Algorithm 1 Proposed algorithm for nucleus cell segmentation.

```
1: procedure SEGMENTATION(Image)
2:     Extract green channel from RGB
3:     Compute mean shift algorithm
4:     Compute k-means
5:     Compute open morphology
6:     for each region do
7:         Compute area
8:         Compute compactness
9:         if area > 600 & area < 4000 & compactness > 0.3 then
10:            Crop region
11:        else if area > 4000 then
12:            go to line 3
13:        else
14:            Remove region
```

5. New Radial Feature Descriptors

According to previous work using the Herlev database [11, 10], features from the whole cell lead to higher cell classification accuracy than feature vectors extracted from the nucleus only. Accounting for such findings, we designed RFD to describe information that combines both the nucleus and the cytoplasm texture, while only relying on the nucleus segmentation. This descriptor associates the radial histogram (RH) and the gray-level run length matrix (GLRLM) [40] to deliver information about the intensity variation within the cytoplasmic area. Previous methods based on radial information [41, 42] report improved stability of features at different resolutions, more efficient computations, and robustness to noise; these are some of the motivations to extend radial-based descriptors to applications in cell analysis, as described in the next section.



(a) (b)

Figure 5: Intensity variation around the nucleus of an abnormal cell: (a) a grayscale cell image with the green contour corresponding to the nucleus mask, and (b) intensities of all pixels for each θ in (a).

5.1. Radial Histogram

Nucleus texture plays a major role in cervical cell description because the chromatin distribution in normal cells is more homogeneous than those in abnormal cells, which presents high intensity variation. Therefore, the loosely packed form of chromatin (euchromatin) is often associated to normal cells [43]. In contrast, abnormal cells usually have a more dense chromatin distribution, which visually translates into high pixel intensity inside the nucleus while pixels near the nucleus boundary look blurry and indistinct.

Figure 5(a) illustrates the significant intensity variation from the nucleus to cytoplasm. The curves in Figure 5(b) show such variation along the segments at different angles; segments start inside the nucleus with ramifications towards the cell cytoplasm. This sample illustrates that the intensity variation across the abnormal cell is not constant.

The RH defines evenly interspaced segments around the cell nucleus, and proportional to the convexity of the nuclear boundary; it depends on lines around the nucleus edges at various angles to compute the intensity variation along radial lines. A crucial parameter of this descriptor is the number of angles, which

is proportional to the number of points at the nuclear perimeter. To speed up the calculations, we insert the parameter n, where $0 < n \leq 1$, to control the number of edge points in the computation of RH. The angle between two lines arising from two consecutive points at the boundary defines the $angle = \frac{2\pi}{n}$. The number of pixels in each line varies according to the nucleus size, which defines the parameter d, which is the distance between the nucleus center of mass and the nucleus boundary.

We compute the descriptor using the nucleus masks of cell images from both databases. From the nucleus mask, we extract a set of edge points P_e and we then compute the center of mass (P_c) of the nucleus region. Our method selects n points from the set P_e, i.e., $\{p_1, p_2, \cdots, p_n\}$, following the criteria:

$$p_j = p \in P_e \mid \theta_j = \arctan(\Delta y/\Delta x) \simeq j \times \frac{2\pi}{n}, \tag{1}$$

where, $\Delta x = x_{p_j} - x_{p_c}$, $\Delta y = y_{p_j} - y_{p_c}$, $j = 1 \cdots n$ and p is a point in the P_e. The values x_{p_j}, y_{p_j} are coordinates of a particular point in the P_e set and x_{p_c}, y_{p_c} are the coordinates of P_c. Our algorithm searches for n contour points separated by the same angle related to the center of mass of the nucleus contour. We then define an external point (p_j^+) and an internal point (p_j^-) for each p_j previously computed. These points are used to calculate the intensity variation into the cervical cell. They are defined as the points at a distance of D_j from p_j towards the center (p_j^-) and away from it (p_j^+). The distance D_j is given by

$$D_j = |P_c - p_j| * d, \tag{2}$$

where $|\cdot|$ denotes the modulus, d is a parameter that controls the size of D_j, and $*$ stands for multiplication. Ultimately, we calculate the histogram of all pixels within the interval $[p_+^j, p_-^j]$ for $j = 1 \cdots n$.

We notice that the computational complexity for RH was $\mathcal{O}(n * d)$, where n is the number of points used from the nucleus edge.

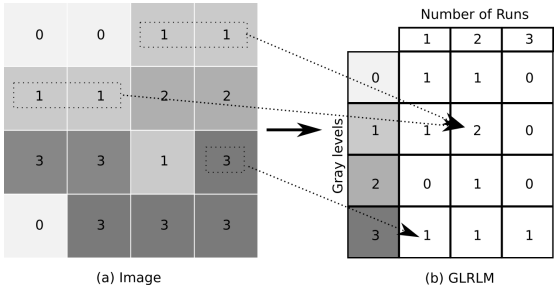


Figure 6: Gray level run length matrix computation: (a) gray level image with intensity values between 0 and 3, and (b) GLRLM histogram.

5.2. Gray-Level Run Length Matrix

The Gray-Level Run Length Matrix (GLRLM) is a two-dimensional histogram H of elements, where each element $h(i, j)$ contains the total number of occurrences of runs (i.e., a set of pixels with the same value) in a given direction $\theta = \{0°, 45°, 90°, 135°\}$ [40] [44]. Run-length statistics capture the coarseness of a textured region in a specific direction, which has applications ranging from brain tumor detection [14] to soil image analysis [45].

The GLRLM captures the chromatin distribution, which is critical for separating normal and abnormal cells. Figure 6 shows how to compute the GLRLM for a general purpose application; notice these calculations occur only inside the nucleus region. From H, we empirically selected and extracted the following features: Short Run Emphasis, Long Run Emphasis, Gray Level Non-uniformity, Run Length Non-uniformity, Run Percentage, Low Gray Level Run Emphasis, High Gray Level Run Emphasis, Short Run Low Gray Level Emphasis, Short Run High Gray Level Emphasis, Long Run Low Gray Level Emphasis, and Long Run High Gray Level Emphasis [40, 44]. Notice that we use four different directions $(0°, 45°, 90°, 135°)$ for each of these features.

The concatenation of both RH and GLRLM into a single feature vector gives rise to RFD, with RH comprising 256 features derived from the histogram of an 8-bit intensity image, and GLRLM leads to a feature vector of 44 attributes. Therefore, RFD has dimensionality equals to 300.

6. Metrology for Quantitative Microscopy

Here, we classify images from the Herlev and CRIC databases as normal or abnormal cells, and in seven lesion levels. We carried out the classification of both databases running the Random Forest [19] algorithm. We also carried out sorting and ranking experiments using CBIR for cell images from both databases. In this process, we used the cosine distance metric to compute the similarity between feature vectors [46].

6.1. Parameter Estimation

We estimate parameters for the Random Forest classifier by computing the features for each image database, then obtaining a feature vector database. Next, we split data into two subsets: the first one corresponds to 30% of the samples and it is used for parameter estimation while the second subset, i.e. the classification experiment uses 70% of the data.

The following variables of the Random Forest algorithm are part of the parameter estimation: the number of trees in the forest, the function to measure the quality of a split using Gini index [19], the maximum depth of the tree, the maximum number of features to consider when looking for the best split, the minimum number of samples required to split an internal node, and the minimum number of samples required to be at a leaf node. Table 1 shows the range of parameter values used for the Random Forest.

6.2. Performance Evaluation and Classification

Both Herlev and CRIC databases are unbalanced, therefore the Kappa index (κ) [47], which does not depend on data balancing, is a better choice to evaluate the classification performance than the Receiver Operating Characteristic Curve (ROC) [48]. The κ also allows to establish the accuracy level according to: Bad, Reasonable, Good, Very Good, and Excellent. Related papers [6, 8, 49] on cell classification have adopted κ for performance evaluation.

In addition, we calculate the False Negative Rate (FNR) to identify the best feature-set/classifier combination since FNR is a critical information for

Table 1: Parameters of the Random Forest Classifier where s is the size of the feature vector (number of features), and $range(\alpha, \beta, \delta)$ is a function that returns values between α and β with steps of δ.

Parameters	Range
Number of trees	$range(10, 1000, 50)$
Quality of a split	Gini, entropy
Maximum depth	$range(1, 100, 1)$
Max features	$range(1, s, 1)$
Min samples (internal node)	$range(1, s, 1)$
Min samples (leaf node)	$range(1, s, 1)$

health care systems; FNR indicates the ratio between the number of abnormal cells, which are classified as normal cells. In order to evaluate and compare the CBIR experiments, we compute the Mean Average Precision (MAP) [50], that measures the ranking quality and is a well-known performance metric for CBIR.

6.3. Simulations with 0.632 Bootstrap

The bootstrap method [51] creates several subsets from the original dataset without using prior information. Given a dataset $\mathbf{x} = (x_1, x_2, ..., x_N)$ of size N, we generate M randomly distributed subsets. Each bootstrap sample, defined by $\mathbf{x}^* = x_1^*, x_2^*, ..., x_N^*$, is composed of N feature vectors from the original dataset.

Assume that $\varepsilon(\mathbf{x}_{train}^{*^m}, \mathbf{x}_{test}^{*^m})$ is the evaluation rate for training with $\mathbf{x}_{train}^{*^m}$ and testing with $\mathbf{x}_{test}^{*^m}$. Using the bootstrap technique, we generate M training sets $(\mathbf{x}_{train}^{*^1}, \mathbf{x}_{train}^{*^2}, ..., \mathbf{x}_{train}^{*^M})$ where each $\mathbf{x}_{train}^{*^m} = x_1^{*^m}, x_2^{*^m}, ..., x_N^{*^m}$ is obtained by choosing N attribute vectors, with replacement, from the original dataset $\mathbf{x}$. We define the testing set by $\mathbf{x}_{test}^{*^m}$, this set consists of attribute vectors that do not appear in $\mathbf{x}_{train}^{*^m}$.

$$\varepsilon_{0.632} = \varepsilon(\mathbf{x}, \mathbf{x}) - \hat{w}_{0.632}^m \tag{3}$$

where $\varepsilon(\mathbf{x}, \mathbf{x})$ is the index for training/testing with the original dataset $\mathbf{x}$ and:

$$\hat{w}_{0.632}^m = 0.632[\varepsilon(\mathbf{x}, \mathbf{x}) - \varepsilon(\mathbf{x}_{train}^{*^m}, \mathbf{x}_{test}^{*^m})]. \tag{4}$$

The estimate for the 0.632 method is calculated by averaging Eq. 4 over M bootstrap samples. The $\varepsilon_{0.632}$ value estimate is then given by

$$\varepsilon_{0.632} = 0.368\varepsilon(\mathbf{x}, \mathbf{x}) + \frac{0.632}{M} \sum_{m=1}^{M} \varepsilon(\mathbf{x}_{train}^{*^m}, \mathbf{x}_{test}^{*^m}). \quad (5)$$

Further details about the 0.632 bootstrap estimator can be found in [51]

7. Results and Discussion

We analyzed cell images from Herlev and CRIC data sets, focusing on separating abnormal from normal cells. We ran our unsupervised segmentation algorithm to samples from CRIC to obtain masks, and used the Herlev masks as made available in [3]. Our experiments considered fourteen sets of image descriptors, which are input as separate feature vectors to the Random Forest algorithm to classify cervical cells. These vectors are also used as signatures to retrieve cells by similarity, which depend on the cosine distance measure as part of our CBIR experiments.

In all experiments, we normalize the features in the $[0, 1]$ interval. After the parameter estimation procedure, the optimum values obtained for the Random Forest classifier were: number of trees = 910, quality of split = gini, maximum depth = 95, max features = 9, min samples (internal nodes) = 3, min samples (leaf node) = 2.

RFD-based features depend on n and d parameters ranging over the interval $[0.1, 1.0]$ with steps of 0.2. We compared RFD results with thirteen other feature extraction methods from the literature. In addition, we also assessed separately six sets of image descriptors with different dimensionality (dim):

1. GLCM [13] derived metrics, such as contrast, dissimilarity, energy, homogeneity, correlation, second angular moment ($dim = 6$);
2. Histogram features, such entropy, mean, energy, variance, roughness, skewness and kurtosis from the image histogram ($dim = 7$);

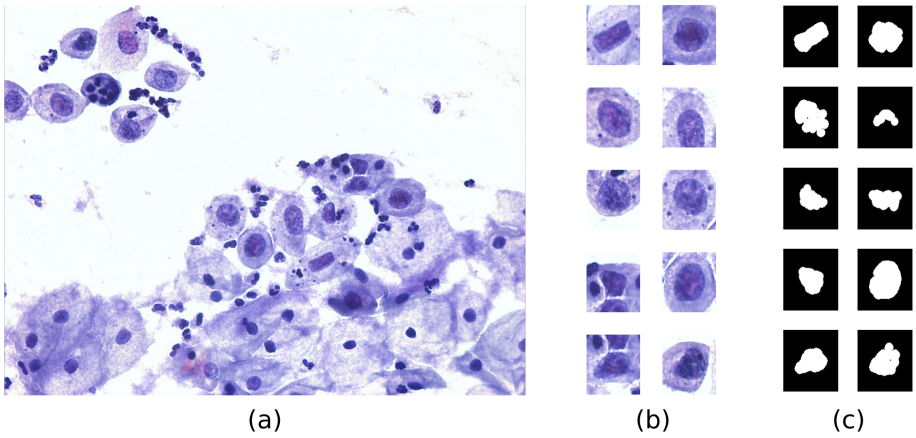


Figure 7: Image segmentation results for the CRIC database: (a) original image, (b)cropped samples from the original image, and (c) the corresponding nuclei segmentation results.

3. Histogram of oriented gradients (HOG) [15] with features whose dimensionality is proportional to HOG parameters, here 20 blocks, 12 sub-blocks using 12 directions ($dim = 2,880$);
4. LBP [16] with features proportional to the pixel depth, here 8-bit images ($dim = 256$);
5. Two Convolutional Neural Networks (CNNs): the Inception-Resnet-v2, a CNN trained with the ImageNet [52] database ($dim = 1,536$), and the LeNet [18], a shallower CNN than the Inception-Resnet-v2 but suitable deep learning algorithm to describe and classify image databases with a small amount of samples ($dim = 192$).

Figure 7a shows an original image from the CRIC database, Figure 7b shows samples of cropped cells from this image, and Figure 7c shows the corresponding nuclei segmentation results from our unsupervised algorithm.

7.1. Binary Class Labeling

Table 2 shows the indexes FNR and κ obtained with the bootstrap method using the Herlev and CRIC databases, respectively. We report the result of the best set of parameters for RH and RFD, which are $n = 0.7$ and $d = 0.5$. In

addition, we observed that lower values of n and d resulted in lower classification rates. For example, if we use $n = 0.1$, then the descriptor will consider only 10% of the nucleus edge pixels, which is seldom enough information for RFD to ascertain that normal and abnormal cells are different. Similar analysis applies to the parameter d that manages the number of pixels of each line: if d is too small (e.g., $d = 0.1$), then fewer pixels will be analyzed across the nucleus and cytoplasm regions.

Table 2: Cell classification with RFD and its individual components using GT (Herlev) and the proposed segmentation method (CRIC): accuracy in terms of FNR and κ.

	Herlev		CRIC	
	FNR	κ	FNR	κ
GLRLM	**0.02±0.01**	0.86±0.04	0.19±0.02	0.73±0.02
RH	**0.02±0.01**	0.86±0.04	**0.14±0.03**	**0.78±0.03**
RFD	**0.02±0.01**	**0.89±0.04**	0.15±0.03	0.77±0.03

Bold numbers indicate the best results.

For higher values of these parameters, information redundancy dominates the metric, which is detrimental to the cell classification; higher values of n lead to the inclusion of repeated data, as illustrated in the lower resolution images with less than 360 points at the nucleus edge. Analogously, higher values of d will add outliers, such as points outside the cytoplasm. The best result that the proposed method and its individual components achieved for the Herlev images was $\kappa = 0.89$. Our proposed RH descriptor reached the best results both for Herlev and CRIC, while our other proposal, RFD, reached slightly inferior results in comparison with GLRLM and RH (FNR$= 0.02 \pm 0.01$) for the CRIC database. For CRIC, the best result was achieved using only RH, while RFD attained the second best result. Our hypothesis is that the lower accuracy rate of the GLRLM affects RFD negatively. Because the CRIC masks are the result of an unsupervised segmentation algorithm (Section 4), the nucleus edge detection was rougher than the ground truth of Herlev, which also impacted the accuracy rate.

Figure 8 presents both the segmentation and classification results, with cor-

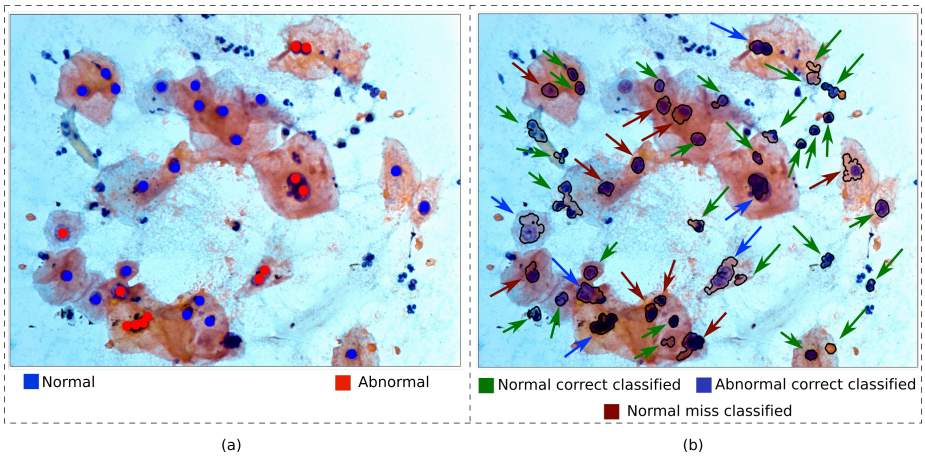


Figure 8: The result for both segmentation and classification algorithms applied to an image from the CRIC database: (a) original image with GT for each nucleus obtained from the classification algorithm, and (b) classification of each segmented region. Black edges correspond to the segmentation results.

rect classification of all of the abnormal cells and no false negatives. Although there are a few non-nucleus regions detected as potential cells, our method did not classify any of them as abnormal cells. Consequently, poor segmentation of some regions did not impact the FNR. Using only nuclear regions, the challenge is the description of chromatin texture and engineering feature vectors. Therefore, we compared the RFD with seven different other methods that were available in the literature. To produce a fair comparison, we applied the same classification methodology to all feature sets. Table 3 shows a comparative analysis using the Herlev and CRIC datasets.

Methods from the literature, with the exception of [11], relied on nucleus and cytoplasm segmentations. Because the CRIC database contains only the nucleus segmentation, we compared the results for the CRIC database with the result of the method developed by [11], which uses nuclei regions as input, and the general purpose methods applied in the nucleus region. Plissiti et al. [11] achieved $\kappa = 0.74 \pm 0.02$ and FNR$= 0.18 \pm 0.02$ and, therefore, it underperformed RFD in terms of both Kappa ($\kappa = 0.78 \pm 0.03$) and False

Table 3: Comparative analysis for classification experiments: FNR and κ using Herlev and CRIC database.

Methods	Herlev		CRIC	
	FNR	κ	FNR	κ
[6]	0.02±0.02	0.84±0.04	-	-
[7]	0.02±0.02	0.86±0.04	-	-
[8]	0.03±0.02	0.83±0.04	-	-
[9]	**0.01±0.01**	0.82±0.05	-	-
[10]	0.03±0.02	0.82±0.04	-	-
[11]	0.05±0.01	0.76±0.03	0.18±0.02	0.74±0.02
[12]	0.02±0.02	0.88±0.04	-	-
[13]	0.07±0.03	0.56±0.06	0.22±0.05	0.50±0.03
[14]	0.02±0.01	0.83±0.04	0.32±0.06	0.50±0.03
[15]	0.26±0.04	0.77±0.04	0.50±0.05	0.43±0.04
[16]	0.21±0.06	0.75±0.05	0.27±0.04	0.65±0.03
[17]	**0.01±0.01**	0.77±0.05	**0.04±0.02**	0.74±0.05
[18]	**0.01±0.01**	0.78±0.04	0.10±0.02	0.72±0.03
RFD	0.02±0.01	**0.89±0.04**	0.14±0.03	**0.78±0.03**

− represents the methods that rely on the cytoplasm segmentation. Bold numbers indicate the best results.

Negative Rate (FNR= 0.14±0.03) for CRIC. The lower FNR may be due to the quality of the cell image segmentation. RFD yielded $\kappa = 0.89$, outperforming the other methods for Herlev. Although the best rate for the FNR was obtained with the method introduced by Mariarputham et al. [9], it reached the lowest κ.

7.2. Multiclass Labeling

Table 4 presents results for seven types of cells from the Herlev database. We run the one-against-all classification using all feature extraction methods, which performed poorly due to the small number of images per cell class. Moreover, the seven classes are unbalanced and there is a great diversity of nuclear structures within abnormal classes.

Table 4: Comparative analysis of FNR and κ using GT for seven classes in Herlev database.

	CI	LD	MD	NC	NI	NS	SD
				[6]			
κ	0.58±0.07	0.70±0.05	0.49±0.06	0.65±0.08	0.89±0.05	0.90±0.05	0.47±0.06
FNR	0.36±0.08	0.24±0.07	0.45±0.07	0.33±0.09	0.11±0.08	0.10±0.07	0.43±0.06
				[7]			
κ	**0.66±0.05**	0.72±0.04	**0.53±0.05**	**0.76±0.08**	0.88±0.05	**0.92±0.04**	0.55±0.06
FNR	**0.32±0.07**	**0.23±0.06**	0.46±0.05	**0.26±0.09**	0.10±0.08	0.08±0.07	**0.40±0.06**
				[8]			
κ	0.64±0.06	0.72±0.05	0.50±0.07	0.73±0.07	0.88±0.07	0.91±0.05	0.50±0.06
FNR	0.33±0.08	0.24±0.05	0.47±0.07	0.27±0.09	0.12±0.08	0.08±0.07	0.44±0.06
				[9]			
κ	0.51±0.07	0.65±0.05	0.41±0.04	0.44±0.07	0.69±0.10	0.76±0.10	0.46±0.05
FNR	0.53±0.05	0.39±0.05	**0.61±0.03**	0.58±0.05	0.37±0.11	0.31±0.11	0.55±0.05
				[10]			
κ	0.64±0.05	0.74±0.05	0.52±0.06	0.71±0.06	0.89±0.04	0.91±0.05	**0.51±0.05**
FNR	0.34±0.07	**0.23±0.06**	0.47±0.06	0.30±0.07	0.11±0.07	**0.08±0.06**	0.45±0.06
				[11]			
κ	0.46±0.06	0.55±0.06	0.49±0.05	0.61±0.07	0.66±0.08	0.86±0.04	0.43±0.06
FNR	0.47±0.06	0.37±0.07	0.44±0.06	0.37±0.08	0.32±0.11	0.13±0.07	0.47±0.06
				[12]			
κ	0.65±0.07	**0.75±0.04**	0.52±0.07	0.72±0.06	**0.90±0.05**	0.91±0.06	0.50±0.06
FNR	0.32±0.08	**0.23±0.06**	0.48±0.07	0.30±0.08	**0.10±0.07**	0.10±0.07	0.45±0.06
				RFD			
κ	0.51±0.05	0.66±0.05	0.44±0.05	0.55±0.08	0.74±0.09	0.87±0.06	0.50±0.05
FNR	0.53±0.04	0.34±0.05	0.58±0.03	0.50±0.06	0.33±0.10	0.16±0.08	0.52±0.04

CI: carcinoma in situ, LD: light dysplastic, MD: moderate dysplastic. NC: normal columnar, NI: normal intermediate, NS: normal superficial, SD: severe dysplastic. Bold numbers indicate the best results.

Using the CRIC database, we compare our algorithm results to those obtained with the algorithm proposed by Plissiti et al. [11]. Our approach achieved the best κ value for five classes. For FNR, the Plissiti et al. [11] algorithm pro-

duced the best result among the other methods. This may happen because these methods relied on cytoplasm segmentation, which affected the final result.

7.3. Content-Based Image Retrieval for Cell Recommendation

A Content-Based Image Retrieval (CBIR) system is a search engine that uses similarity metrics to retrieve and rank images from a database by comparing their feature vectors to an input query [36, 37]. Besides the classification procedure, we also performed the CBIR experiments to explore the RFD generalization for image recommendation tasks. We computed the MAP [50] metric for performance evaluation, considering: (a) the whole database (MAP), (b) only images of normal cells (MAP_n), and (c) only images of abnormal cells (MAP_{an}). Table 5 shows the results using the Herlev and CRIC databases for the proposed descriptor and its individual components. RFD achieved MAP values equal to 0.84 and 0.82 for Herlev and CRIC, respectively.

Table 5: Cell-image retrieval with RFD and its individual components using GT from Herlev and from CRIC (proposed segmentation method): accuracy in terms of MAP.

	Herlev			CRIC		
	MAP	MAP_n	MAP_{an}	MAP	MAP_n	MAP_{an}
GLRLM	0.76±0.16	0.68±0.20	0.84±0.12	0.71±0.18	0.79±0.13	0.63±0.23
RH	0.79±0.16	0.76±0.18	0.81±0.13	0.70±0.16	0.80±0.12	0.61±0.20
RFD	**0.84±0.14**	**0.81±0.18**	**0.88±0.10**	**0.82±0.15**	**0.86±0.12**	**0.77±0.17**

Bold numbers indicate the best results.

We assessed the CBIR experiments using the same feature vectors employed for the classification experiments. Table 6 shows results for Herlev and CRIC databases, using the features based on the proposed descriptor set. We achieved the best MAP and MAP_n values for both databases, which led us to conclude that RFD outperformed the other features for normal cell retrieval. Concerning MAP_ab, the method introduced in [11] attained the best result for the Herlev database while Lenet [18] outperformed the others for the CRIC database. Although those methods outperformed RFD, they did not succeed in both

Figure 9: Graphical result for a CBIR experiment using RFD. The first eight rows of the first column contains abnormal cell queries and the other rows contains normal cell queries. Green and red edges corresponds to images correctly and incorrectly recommended, respectively.

databases, simultaneously, as RFD did. Figure 9 illustrates the CBIR result using RFD and cell samples from the CRIC image database.

7.4. Processing Time

We compared the processing time of our algorithm to the one proposed by Sarwar et al. [12], which achieved the second best κ and FNR values. The

Table 6: Comparative analysis for CBIR experiments: MAP using Herlev and CRIC database.

Methods	Herlev			CRIC		
	MAP	MAP_n	MAP_{ab}	MAP	MAP_n	MAP_{ab}
[6]	0.71±0.21	0.68±0.36	0.85±0.07	-	-	-
[7]	0.79±0.18	0.67±0.28	0.91±0.07	-	-	-
[8]	0.77±0.21	0.69±0.36	0.85±0.07	-	-	-
[9]	0.72±0.18	0.62±0.30	0.82±0.06	-	-	-
[10]	0.81±0.19	0.74±0.31	0.88±0.07	-	-	-
[11]	0.81±0.17	0.69±0.25	**0.93±0.10**	0.70±0.20	0.73±0.19	0.66±0.20
[12]	0.81±0.19	0.74±0.31	0.88±0.07	-	-	-
[13]	0.57±0.08	0.34±0.13	0.79±0.03	0.58±0.14	0.64±0.18	0.51±0.10
[14]	0.63±0.14	0.45±0.18	0.81±0.10	0.57±0.13	0.64±0.11	0.50±0.15
[15]	0.60±0.06	0.40±0.09	0.79±0.02	0.61±0.08	0.79±0.06	0.44±0.10
[16]	0.55±0.06	0.23±0.08	0.86±0.03	0.56±0.08	0.52±0.07	0.60±0.09
[17]	0.70±0.16	0.56±0.27	0.85±0.06	0.69±0.14	0.62±0.17	0.77±0.10
[18]	0.69±0.15	0.53±0.26	0.86±0.05	0.81±0.21	0.75±0.25	**0.87±0.18**
RFD	**0.84±0.14**	**0.81±0.18**	0.88±0.10	**0.82±0.15**	**0.86±0.12**	0.77±0.17

− indicates methods dependent on cytoplasm segmentation. Bold numbers indicate the best results.

average nucleus computation time for the algorithm introduced by Sarwar et al. [12] during feature extraction was 0.035s; however, it required an average time of 5.68s to remove the background and segment the cytoplasm region because they apply the algorithm proposed by Li et al.'s [5].

The overall processing time of the RFD to extract the features using our non-optimized Python code took 0.06s on average for Herlev images and 0.04s for CRIC images on a PC with a 3.1 GHz Intel Core i7 processor and 16 GB RAM.

Our five implemented modules, the nucleus segmentation, the feature extraction, the classification, the retrieval and the evaluation tasks, are Python code that leverages important scientific packages, such as `numpy`, `scipy`, `scikit-image`, `scikit-learn`, and `tensorflow`.

8. Conclusion

The main biological motivation of characterizing cells by texture are: (a) the level of chromatin distribution inside the nucleus indicates abnormality in the cell; (b) normal cells present less variation of intensity inside the nucleus; (c) the transition between nucleus and cytoplasm is smoother in abnormal cells; (d) it is possible to classify and retrieve overlapping cells with a high degree of accuracy.

In this paper, we proposed an unsupervised nuclei segmentation method that was designed to obtain masks for a new cell database (CRIC) with high image resolutions, combining mean shift and k-means clustering. We also introduced the radial feature descriptors (RFD), which characterize cells through a mechanism of by-passing the cytoplasm segmentation in favor of texture feature extraction at the cytoplasmic zone. Our experiments included a public database (Herlev) and its respective ground truth, as well as the new database. RFD showed promising results when tested against both image databases.

Testing all these methods enabled us to deliver a compact indicator of the intensity variation of the nuclei, enabling fast cell classification. We investigated several parameters and metrics and reached a processing time of $\approx$0.04s/cell in the worst case scenarios, i.e., near real-time feedback. We also compared RFD with other thirteen sets of well-known descriptors, performing cell classification and CBIR experiments, and RFD proved competitive. Our results in terms of the performance evaluation metrics are encouraging because the κ and MAP reached high values while the FNR achieved low values for both databases.

Future work will include investigation of larger datasets and new metrics coming from the RFD histogram, such as entropy, skewness, energy, and kurtosis, which may allow more compact representation to speed-up the characterization process.

Acknowledgments

This work was supported by CNPq (304673/2011-0, 472565/2011-7, 401120/2013-9, 306600/2016-1), CAPES/CNPq-PVE (401442/2014-4, 207307/2015-6, 207306/2015-0) and Fapemig. This research is also funded in part by the Gordon and Betty Moore Foundation through Grant GBMF3834 and by the Alfred P. Sloan Foundation through Grant 2013-10-27 to the University of California, Berkeley. Algorithmic work is partially supported by the Office of Science of the US Department of Energy (DOE) under Contract No. DE-AC02-05CH11231, Advanced Scientific Computing Research (ASCR) Early Career Award. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and they do not necessarily reflect the views of DOE or the University of California. We are grateful to the cytopathologists, Alessandra Tobias and Mariana Trevisan, who classified the cervical cells manually for the CRIC database, and to the BIDS data scientists and staff for encouraging the exploration of Python packages in deploying open-source tools.

References

[1] American Cancer Society, `https://www.cancer.org/cancer/cervical-cancer/about/key-statistics.html` (2016).

[2] K. R. Castleman, B. S. White, Optimizing cervical specimen classifiers, IEEE Transactions on Pattern Analysis and Machine Intelligence PAMI-2 (5) (1980) 451–457. `doi:10.1109/TPAMI.1980.6592366`.

[3] J. Jantzen, J. Norup, G. Dounias, B. Bjerregaard, Pap-smear benchmark data for pattern classification technical University of Denmark, Nature inspired Smart Information Systems (2005) 1–9.

[4] Z. Lu, G. Carneiro, A. P. Bradley, An improved joint optimization of multiple level set functions for the segmentation of overlapping cervical cells, IEEE Transactions on Image Processing 24 (4) (2015) 1261–1272.

[5] K. Li, Z. Lu, W. Liu, J. Yin, Cytoplasm and nucleus segmentation in cervical smear images using radiating gvf snake, Pattern Recognition 45 (4) (2012) 1255 – 1264.

[6] T. Chankong, N. Theera-Umpon, S. Auephanwiriyakul, Automatic cervical cell segmentation and classification in pap smears, Comput. Methods Prog. Biomed. 113 (2) (2014) 539–556.

[7] Y.-F. Chen, P.-C. Huang, K.-C. Lin, H.-H. Lin, L.-E. Wang, C.-C. Cheng, T.-P. Chen, Y.-K. Chan, J. Chiang, Semi-automatic segmentation and classification of pap smear cells, IEEE J. Biomed. Health Inform. 18 (1) (2014) 94–108.

[8] A. Gençtav, S. Aksoy, S. Onder, Unsupervised segmentation and classification of cervical cell images, Pattern Recognition 45 (12) (2012) 4151–4168.

[9] M. E. J. Mariarputham, A. Stephen, Nominated texture based cervical cancer classification, Comput Math Methods Med. 2015 (2015) 1 – 10.

[10] Y. Marinakis, G. Dounias, J. Jantzen, Pap smear diagnosis using a hybrid intelligent scheme focusing on genetic algorithm based feature selection and nearest neighbor classification, Computers in Biology and Medicine 39 (1) (2009) 69 – 78.

[11] M. E. Plissiti, C. Nikou, Cervical Cell Classification Based Exclusively on Nucleus Features, Springer Berlin Heidelberg, Berlin, Heidelberg, 2012, pp. 483–490.

[12] A. Sarwar, V. Sharma, R. Gupta, Hybrid ensemble learning technique for screening of cervical cancer using papanicolaou smear image analysis, Personalized Medicine Universe 4 (2015) 54 – 62.

[13] R. Haralick, K. Shanmugam, I. Dinstein, Textural features for image classification, IEEE Trans Syst Man Cybern 3 (6) (1973) 610–621.

[14] N. Nabizadeh, M. Kubat, Brain tumors detection and segmentation in {MR} images: Gabor wavelet vs. statistical features, Computers & Electrical Engineering 45 (2015) 286 – 301.

[15] N. Dalal, B. Triggs, Histograms of oriented gradients for human detection, in: Computer Vision and Pattern Recognition, 2005, pp. 886–893.

[16] T. Ojala, M. Pietikäinen, D. Harwood, A comparative study of texture measures with classification based on featured distributions, Pattern Recognition 29 (1) (1996) 51 – 59.

[17] C. Szegedy, S. Ioffe, V. Vanhoucke, Inception-v4, inception-resnet and the impact of residual connections on learning, Computing Research Repository abs/1602.07261.

[18] Y. Lecun, L. Bottou, Y. Bengio, P. Haffner, Gradient-based learning applied to document recognition, in: Proceedings of the IEEE, 1998, pp. 2278–2324.

[19] L. Breiman, Random forests, Mach. Learn. 45 (1) (2001) 5–32.

[20] H. Irshad, A. Veillard, L. Roux, D. Racoceanu, Methods for nuclei detection, segmentation, and classification in digital histopathology: A review-current status and future potential, IEEE Reviews in Biomedical Engineering 7 (2014) 97–114.

[21] D. Ushizima, A. Bianchi, C. Carneiro, Segmentation of subcellular compartiments combining superpixel representation with voronoi diagrams, in: Overlapping Cervical Cytology Image Segmentation Challenge - IEEE ISBI, 2014, pp. 1–2.

[22] H. A. Phoulady, D. Goldgof, L. O. Hall, P. R. Mouton, A framework for nucleus and overlapping cytoplasm segmentation in cervical cytology extended depth of field and volume images, Computerized Medical Imaging and Graphics 59 (2017) 38 – 49.

[23] A. Tareef, Y. Song, H. Huang, D. Feng, M. Chen, Y. Wang, W. Cai, Multi-pass fast watershed for accurate segmentation of overlapping cervical cells, IEEE Transactions on Medical Imaging (2018) 1–1.

[24] Z. Lu, G. Carneiro, A. Bradley, D. Ushizima, M. S. Nosrati, A. Bianchi, C. Carneiro, G. Hamarneh, Evaluation of three algorithms for the segmentation of overlapping cervical cells, IEEE J. Biomed. Health Inform. (2016) 1–11.

[25] B. E. Bejnordi, R. Moshavegh, K. Sujathan, P. Malm, E. Bengtsson, A. Mehnert, Novel chromatin texture features for the classification of pap smears, in: Proc. SPIE, Vol. 8676, 2013, pp. 867608–867608–8.

[26] G. Ramalho, D. Ferreira, D. Ushizima, A. Bianchi, C. Carneiro, Cell reconstruction under voronoi and enclosing ellipses from 3d microscopy, in: Overlapping Cervical Cytology Image Segmentation Challenge - IEEE ISBI, 2015, pp. 1–2.

[27] A. P. Zijdenbos, B. M. Dawant, R. A. Margolin, A. C. Palmer, Morphometric analysis of white matter lesions in mr images: method and validation, IEEE Transactions on Medical Imaging 13 (4) (1994) 716–724.

[28] J. Sá, A. Backes, A color texture analysis method based on a gravitational approach for classification of the pap-smear database, in: IEEE International Conference on Image Processing, 2014, pp. 2280–2284.

[29] N. Noroozi, A. Zakerolhosseini, Computer assisted diagnosis of basal cell carcinoma using z-transform features, Journal of Visual Communication and Image Representation 40 (2016) 128 – 148.

[30] A. Kale, S. Aksoy, Segmentation of cervical cell images, in: 20th International Conference on Pattern Recognition (ICPR), 2010, pp. 2399–2402.

[31] H. Tamura, S. Mori, T. Yamawaki, Textural features corresponding to visual perception, IEEE Trans Syst Man Cybern 8 (6) (1978) 460–473.

[32] W. T. Freeman, W. T. Freeman, M. Roth, M. Roth, Orientation histograms for hand gesture recognition, in: In International Workshop on Automatic Face and Gesture Recognition, 1994, pp. 296–301.

[33] L. Li, H. Ge, J. Gao, Maximum–minimum–median average msd-based approach for face recognition, AEU - International Journal of Electronics and Communications 70 (7) (2016) 920 – 927.

[34] J. qiang Gao, L. ya Fan, L. zhong Xu, Median null(sw)-based method for face feature recognition, Applied Mathematics and Computation 219 (12) (2013) 6410 – 6419.

[35] L. Li, J. Gao, H. Ge, A new face recognition method via semi-discrete decomposition for one sample problem, Optik 127 (19) (2016) 7408 – 7417.

[36] D. Ushizima, F. Araujo, R. Silva, Searchable datasets in python: images across domains, experiments, algorithms and learning, in: Proc. of PyData, Vol. 1, San Francisco, CA, 2016, pp. 1–2.

[37] F. H. Araujo, R. R. Silva, F. N. Medeiros, D. D. Parkinson, A. Hexemer, C. M. Carneiro, D. M. Ushizima, Reverse image search for scientific data within and beyond the visible spectrum, Expert Systems with Applications 109 (2018) 35 – 48.

[38] D. M. Ushizima-Sabino, L. d. F. da Costa, E. G. Rizzatti, M. A. Zago, A texture approach to leukocyte recognition, Real-Time Imaging 10 (4) (2004) 205–216.

[39] R. C. Gonzalez, R. E. Woods, Digital Image Processing (3rd Edition), Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 2006.

[40] M. M. Galloway, Texture analysis using gray level run lengths, Computer Graphics and Image Processing 4 (2) (1975) 172 – 179.

[41] C. Ding, J. Choi, D. Tao, L. S. Davis, Multi-directional multi-level dual-cross patterns for robust face recognition, IEEE Transactions on Pattern Analysis and Machine Intelligence 38 (3) (2016) 518–531.

[42] H. Pei, S. Yanqing, T. Chaowei, Z. Siman, Center-symmetric local binary pattern based on weighted neighbor contribution, Optik 127 (23) (2016) 11599 – 11606.

[43] S. Watanabe, T. Iwasaka, M. Yokoyama, M. Uchiyama, T. Kaku, T. Matsuyama, Analysis of nuclear chromatin distribution in cervical glandular abnormalities, Acta cytologica 48 (4) (2004) 505—513.

[44] X. Tang, Texture information in run-length matrices, IEEE Transactions on Image Processing 7 (11) (1998) 1602–1609.

[45] F. R. Ajdadi, Y. A. Gilandeh, K. Mollazade, R. P. Hasanzadeh, Application of machine vision for classification of soil aggregate size, Soil and Tillage Research 162 (2016) 8 – 17.

[46] C. Singh, K. P. Kaur, A fast and efficient image retrieval system based on color and texture features, Journal of Visual Communication and Image Representation 41 (2016) 225 – 238.

[47] J. Landis, G. Koch, The measurement of observer agreement for categorical data., Biometrics 33 (1) (1977) 159–174.

[48] A. P. Bradley, The use of the area under the ROC curve in the evaluation of machine learning algorithms, Pattern Recognition 30 (7) (1997) 1145–1159.

[49] X. Wang, B. Zheng, R. R. Zhang, S. Li, X. Chen, J. J. Mulvihill, X. Lu, H. Pang, H. Liu, Automated analysis of fluorescent in situ hybridization (fish) labeled genetic biomarkers in assisting cervical cancer diagnosis, Technology in cancer research and treatment 9 (3) (2010) 231–242.

[50] B. Wang, D. Brown, Y. Gao, J. L. Salle, March: Multiscale-arch-height description for mobile retrieval of leaf images, Information Sciences 302 (2015) 132 – 148.

[51] B. Efron, Estimating the error rate of a prediction rule: improvement on cross-validation, Journal of the American Statistical Association 78 (382) (1983) 316–331.

[52] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, L. Fei-Fei, ImageNet Large Scale Visual Recognition Challenge, International Journal of Computer Vision 115 (3) (2015) 211–252.