Published in final edited form as:

Ann Biomed Eng. 2012 May ; 40(5): 1192–1204. doi:10.1007/s10439-011-0498-8.

Automated detection of dual p16/Ki67 nuclear immunoreactivity in liquid-based Pap tests for improved cervical cancer risk stratification

Arkadiusz Gertych^{1,*}, Anika O. Joseph², Ann E. Walts³, and Shikha Bose³

¹ Bioinformatics Laboratory, Department of Surgery, Cedars-Sinai Medical Center, Los Angeles CA

² Department of Biomedical Engineering, University of Southern California, Los Angeles CA

³ Department of Pathology and Laboratory Medicine, Cedars-Sinai Medical Center, Los Angeles CA.

Abstract

The Papanicolaou (Pap) test is a routine cytological procedure for early detection of dysplastic lesions in cervical epithelium. A reliable screening method is crucial for triage of women at risk; however manual screening and interpretation are associated with relatively low sensitivity and substantial interobserver diagnostic variability. P16 and Ki67 biomarkers have been recently proposed as adjunctive tools in the diagnosis of high-risk human papillomavirus (hrHPV) associated dysplasias to supplement the morphological characteristics of cells by additional colorimetric features. In this study, an automated technique for the evaluation of dual p16/Ki67 immunoreactivity in cervical cell nuclei is introduced. Smears stained with p16 and Ki67 antibodies were digitized, and analyzed by algorithms we developed. Gradient-based radial symmetry operator and adaptive processing of symmetry image were employed to obtain the nuclear mask. This step was followed by the extraction of features including pixel data and immunoreactivity signature from each nucleus. The features were analyzed by two support vector machine classifiers to assign a nucleus into one of four types of immunoreactivity: p16 positive (p16⁺/Ki67⁻), Ki67 positive (p16⁻/Ki67⁺), dual p16/Ki67 positive (p16⁺/Ki67⁺) and negative (p16⁻/Ki67⁻) respectively. Results obtained by our method correlated well with readings by two cytopathologists (n=18068 cells); p16⁺/Ki67⁺ nuclei were classified with respective precisions of 77.1% and 82.6%. Specificity in identification of p16⁻/Ki67⁻ nuclei was better than 99.5%, and the sensitivity in detection of all immunopositive nuclei was 86.3% and 89.4% respectively. We found that the quantitative characterization of immunoreactivity provided by the additional highlighting of classified nuclei can positively impact the efficacy and screening outcome of the Pap test.

Keywords

Pap test; cervical cancer screening; immunocytochemistry; computer analysis; nuclei segmentation; quantification

*Corresponding author: Arkadiusz Gertych PhD Assistant Professor Bioinformatics Lab, Department of Surgery, Cedars-Sinai Medical Center, 8797 Beverly Boulevard, Suite #214 Los Angeles, CA 90048 Phone: 310-423-2964, Arkadiusz.Gertych@cshs.org, agertych@gmail.com.

Conflict of interest statement

The authors declare that they have no conflict of interest.

Introduction

The Pap test, originally developed in the 1960s, is an effective method for screening women at risk for cervical cancer. The primary goals are to identify neoplastic and pre-neoplastic epithelial cells, to stratify women for the risk of cervical cancer, and to isolate cases for additional definitive studies. In 2000, approximately 61 million women in the United States had at least one Pap test performed. Of these 3.1 million (~5%) were interpreted as abnormal and required medical follow-up [1, 2].

Currently, Pap smear screening is almost exclusively performed via transmitted light microscopy and requires significant manual interaction. A trained cytotechnologist or pathologist applies qualitative criteria to identify cells with altered morphology. [3]. Characteristic changes in nuclear size and texture that highly correlate with hrHPV infection serve as diagnostic features. However, this process is subjective, labor-intensive, costly, and is associated with false positive and false negative error rates that negatively affect the efficacy of manual screening [4].

Imaging-based routines are well suited for automated processing of Pap tests. Current scientific efforts in this field have mainly focused on the development of systems that can distinguish abnormal cells by extracting morphological and optical characteristics from the nucleus and cytoplasm. Two analytical components are required for consistent processing of Pap smear images: automated segmentation and classification of cellular features. Several approaches have recently been proposed for the segmentation of nuclei that can support automated analysis of images acquired from Papanicolaou stained smears. In [5] an unsupervised edge enhancement was introduced to outline nuclear and cytoplasm contours in monochromatic images of single cells and in [6] a region growing technique combined with moving k-means clustering implemented further as a diagnosis support system was proposed. Unfortunately, these techniques were able to outline nuclei and cell membranes only in non-overlapping cells. Moreover, the contours obtained by [6] were sensitive to local image contrast and resulted in split outlines of hypochromatic nuclei. An approach presented in [7] utilized color images of smears in combination with fuzzy logic and *a priori* human knowledge for nuclear segmentation. Although the detection efficacy was generally high, it was significantly decreased for nuclei with low optical density. A more reliable technique that employs morphological analysis and clustering for nuclei localization was presented in [8]. Its sensitivity and specificity varied depending on the type of clustering that was applied to refine nuclear center prototypes and reached 90.57% and 75.28%, respectively, for fuzzy clustering, and 69.86% and 92.02%, respectively, for SVM when confronted with visual evaluation by an expert. The same authors compared also the advantages and limitations of other available methods dedicated to nuclear segmentation. The use of hyperspectral imaging to distinguish normal cells from low- and high- grade dysplastic cells in Papnicolau stained smears was evaluated in [9]. Spectral profiles of pixels from the nuclear region were fitted to the reference spectra by the least-square technique to find the best match. Yet, this method relied heavily on pixel-based analysis, and the output consisted of a blend of differently matched pixels. Such non-homogenous results can be difficult in visual interpretation and lead to indefinite reading outcomes. Finally, in [10] an infrared spectral characterization technique was employed to characterize hrHPV infected cells. Unstained normal and low-grade dysplastic cells (with confirmed hrHPV infection) were distinguished by principal component analysis of spectral characteristics from normal cells that were not infected with hrHPV. The experimental setup required that the infrared imaging session be followed by Pap staining and re-imaging in the visible range in order to visually correlate spectral findings with cytomorphology.

There are at least two possible ways to improve the efficacies of manual and automated screening routines: develop more reliable Pap smear image analysis algorithms and reduce Pap test result ambiguity by inclusion of sensitive staining compounds. In fact, the lack of nuclear and cellular cytochemical reporters in the Pap-based stains has already led to an ongoing effort in cytopathology to incorporate p16 and Ki67 biomarkers into the diagnostic and screening workflows [11-13]. P16 is a tumor suppressor gene product that is overexpressed in most hrHPV associated cervical carcinomas and dysplasias and can be detected in early stages of cervical neoplasia [14, 15]. Ki67, a nuclear protein that is present during mitosis and the S, G1, and G2 phases of the cell cycle but not during G0, has been shown to be a good marker for cellular proliferation that is applicable in routine liquid-based cytology [16, 17]. The CINtec dual staining reagent kit detects over-expression of Ki67 as red/nuclear and p16 as a brown appearance of the whole cell with stronger nuclear manifestation. Cells that stain negative for p16 and Ki67 exhibit blue tones in both the cytoplasm and nucleus (Figure 1). Essentially, the co-detection of p16/Ki67 in cytological tests can serve as an indicator of cell cycle de-regulation, which occurs during hrHPV-induced oncogenic transformations [18-20].

Further evaluation of p16/Ki67 dual immunostaining in Pap smears by quantitation of digitized images could improve the diagnostic accuracy and efficacy of the Pap test and provide useful information for clinical management. In this study we introduce a novel nuclei segmentation technique based on radial symmetry transform to localize and classify colorimetric signals in high-resolution images of smears. Immunoresponse was determined by classification of signatures obtained from pixels located within the detected nuclei. K-nearest neighbor (kNN), and support vector machines (SVM) classifiers were tested for best performance. We investigated if the combined method termed quantitative profiling of immunoreactivity (QPI) has the potential to improve Pap test screening efficacies. Validation was performed in cells from clinical smears. Excluding slide digitization and classifier training steps, the entire analytical process is fully automated. To the authors' best knowledge no research has been undertaken so far in the field of automated quantification of immunostaining in liquid-based Pap tests.

Materials and Methods

Specimen preparation and image acquisition

The experimental data was approved for research by the Institutional Review Board and consisted of 24 liquid-based (SurePath™) cervical Pap smears: 14 diagnosed with abnormal cells and 10 interpreted as negative for malignant and dysplastic cells. The smears were destained and then restained utilizing the CINtec®PLUS (mtm laboratories, Westborough, MA) kit that detects overexpression of p16 and Ki67. Automated staining of slides was followed by manual screening performed by experienced cytopathologists who manually annotated (dotted/ink marked) the abnormal cells.

Entire slides were imaged using a high-resolution slide scanner (iScan2.0, BioImagene, CA) dedicated for anatomic pathology work. The slides were randomly split into two sets: twelve for training and twelve for testing such that each set contained five negative and seven positive slides respectively. Several non-overlapping fields of view (I_{FOV}) were selected in each of the slides. In the negative slides the selection of I_{FOV} s was arbitrary whereas in the positive slides the annotation markers were used as a guide. At the time of acquisition a 20× objective (0.5 NA) was set to approximate typical settings used for manual slide screening. Image focus was automatically adjusted by the scanner. 60 I_{FOV} s were collected in total. Each I_{FOV} was exported to the uncompressed TIFF format as a red-green-blue (RGB) image of approximate size of $1.5k \times 1.5k$ pixels with spatial and intensity resolutions of $k=0.46\mu m$ /pixel and 8bit/pixel per color channel.

Analytical workflow

Output files were analyzed through the analytical workflow (Fig.2) consisting of three steps: a) preprocessing with nuclei segmentation, b) extraction of colorimetric signatures, and c) supervised nuclei classification, with the ultimate goal to assign a nucleus to one of the following four categories: p16⁺/Ki67⁻, p16⁻/Ki67⁺, p16⁺/Ki67⁺, and p16⁻/Ki67⁻.

Obtained classification results were pseudo-colored and superimposed onto I_{FOV} to allow for comparison with the original image and performance evaluation. To train classifiers immunopositive and -negative cells randomly selected from the training images were labeled by the cytopathologist (assigned to one of the four above categories). We collected N=100 pattern nuclei: 20 for each type of positive immunoreactivity and 40 for the dual negative nuclei (p16⁻/Ki67⁻) to accommodate for large variability in distribution of characteristic colorimetric features ascribed to individual slides. For training set formation the features extracted from pattern nuclei were linked with the labels and stored in a repository.

Nuclei Segmentation

The proposed segmentation of nuclei employs monochromatic I_{FOV_s} that are run through anisotropic smoothing, gradient-based radial symmetry transform, and adaptive binarization of the radial symmetry image, to enhance and separate gradients of optically dense and round objects and then convert them to a mask of nuclei.

In Perona-Malik anisotropic diffusion [21] significant image edges are preserved whereas the inhomogeneous areas in-between the edges are smoothened. The filter can be implemented as:

$$I^{(p,t+1)} = I^{(p,t)} + \Delta t \sum_{s \in \eta_p} g(|\nabla I_{p,s}^{(t)}|) \nabla I_{p,s}^{(t)}, \quad (1)$$

where: p is the image pixel, s is the closest neighbor of p , η_s denotes the spatial neighborhood of p , and $|\eta_s|$ (four in 2-D images) represents the number of neighbors of p . $g(|\nabla I|)$ is the gradient magnitude-dependent edge stopping function, employing $\nabla I_{p,s}^{(t)}$ as the difference between the value of pixel p and each pixel in η_s and Δt is the scale-space

interval. To maintain stability Δt must satisfy $\Delta t \leq \frac{1}{|\eta_{sp}|}$, and is set to 0.2.

In the first iteration $I^{(p,0)} = I_{FOV}$. The edge stopping function was $g(|\nabla I|) = e^{-\left(\frac{|\nabla I|}{K}\right)^2}$. The gradient threshold K was assessed by Canny noise estimator and set to 5 [21]. The number of iterations t was determined and limited to 10 using 5% error margin between two consecutive runs. Last iteration image is denoted as $\bar{I}_{FOV}$.

Gradient-based analysis was employed as an intermediate step towards finding nuclear object candidates in the pre-processed image $\bar{I}_{FOV}$. Specifically, the circularity in the underlying image content was targeted. To highlight localization of structures that fulfill this criterion the radial symmetry transform [22] was applied first. It is a circularity-sensitive point of interest operator that for each pixel p and its proximity defined by radius r accumulates contribution of image gradients pointing at or away from p in orientation projection O_r and a magnitude projection M_r images. The transform is a function of r and its output maps the degree of roundness at p considered as an object center. Our goal is to detect dark objects over a light background with corresponding pixels evaluated as follows:

$$p^{(o)}(p) = p - \text{round} \left(r \frac{\nabla \bar{I}_{FOV}(p)}{\|\nabla \bar{I}_{FOV}(p)\|} \right) \quad (2)$$

where: *round* is the nearest integer rounding operator, $\|\nabla \bar{I}_{FOV}(p)\|$ is the gradient magnitude and $\nabla \bar{I}_{FOV}(p) / \|\nabla \bar{I}_{FOV}(p)\|$ is the unit gradient at p .

O_r collects the respective number of gradient components, whereas the contrast discrepancy of these components is reflected in M_r . The accumulation is performed as below:

$$O_r(p^{(o)}(p)) = O_r(p^{(o)}(p)) - 1 \quad (3)$$

$$M_r(p^{(o)}(p)) = M_r(p^{(o)}(p)) - \|\nabla \bar{I}_{FOV}(p)\| \quad (4)$$

The symmetry image is obtained by convolving O_r and M_r with a Gaussian mask:

$$S_r = F_r * G_r \quad (5)$$

where: $F_r(p) = |M_r(p) O_r(p)|^\alpha$ is the composite contribution image, α is the radial strictness parameter, and G_r is the isotropic two-dimensional Gaussian kernel.

Unwanted components from non-circular objects and line segments can be partly suppressed by rising $|O_r(p)|$ to an exponential power. Gaussian filter smoothens low amplitude gradients remaining in $M_r(p)$. Following [22] $\alpha = 2$, while the smoothing factor σ in G_r is set to $0.5r$.

Ultimately, the output image S_r is a gray-level image in which pixel intensity is quasi-proportional to the degree of radial symmetry of an object in the input image.

Given the selective characteristic of the radial symmetry operator we further investigated its properties and added knowledge-based morphological criteria to arrive at a robust detector of nuclei. First, we created an artificial image (Fig.3a) with three dark circular objects of different radii ($r_d=3, 5$ and 7) placed over a bright background, and examined three respective radial symmetry responses: S_3 , S_5 and S_7 (Fig.3b). One can note that the local increase of intensity in S_r coincides with the corresponding object's location. Furthermore, for different r , the fluctuation of signal amplitude in S_r follows a certain pattern. We interpreted these observations in the following way: a) for a fixed r the amplitude of peaks in S_r changes monotonically according to the increase or decrease of r_d and b) if $r = r_d$ then the width of peak in S_r allows for an exclusive distinguishing of the object with radius r_d from the other objects by intensity thresholding (Fig.3b). For example, setting up the threshold at a specified value T leads to segmentation of the object with radius $r_d=5$ from S_5 . The object with $r_d=5$ can also be segmented from S_3 or S_7 , however, by applying much lower levels, and the two remaining objects (with $r_d=3$ and 7) can be extracted using the same principle. These findings suggest that applicability of the radial symmetry operator can be extended on tasks going beyond just the point of interest identification.

Next, we assumed that it is possible to segment out the object with radius r_d by substituting specific values of $r = r_d$ to and applying adaptive thresholding of $S_r=r_d$ afterwards. Since the amplitude in the output symmetry image is a function of local gradients, the maximal threshold at which an object can be successfully detected depends also on relative contrast between the object and the background. To better understand these dependencies we created

a threshold map ($\mathbf{T}_{r_d}$) for each of the three objects to be detected. All combinations of intensities typical for an 8-bit gray level image were substituted to the object and background in the artificial image and for each such pair an individual threshold was determined (Fig. 3c). Interestingly, the map turned out to be specifically organized. Isothreshold lines parallel to the diagonal of the map were formed for these pairs of object-background intensity combinations for which the intensity difference was constant (Fig. 3c).

In a real case nuclei segmentation scenario, this property can be useful in seeking the most relevant T in $\mathbf{T}_{r_d}$ especially when the object-background contrast is not known in advance. In our application T is iteratively tested in the range of $[\min(\mathbf{T}_{r=r_d}), \max(\mathbf{T}_{r=r_d})]$. Such approach results in a number of binary objects that need to be analyzed to single out the most pertinent candidates and control segmentation performances. According to [23] in 99% of normal cells and the bulk of abnormal cells, the nuclei are primarily round or oval and the majority of nuclear radii range between $1\mu\text{m}$ - $8\mu\text{m}$ and $4.5\mu\text{m}$ - $11\mu\text{m}$ with the most frequent occurrences at $r_{\mu\text{m}}=2.5\mu\text{m}$ and $r_{\mu\text{m}}=7.5\mu\text{m}$ in normal and dysplastic cells respectively. Thus, if for a given symmetry image $S_{r=r_d}$ the area of a binary object closely matches the area of a circle with radius r_d and object's circularity is high, it is assumed that the binary object's contour will closely correspond to the boundary of the object in the analyzed image. To isolate such objects circular area pertinent to radius r_d and shape determined by a ratio of major and minor axis lengths were implemented as morphological criteria. The set of applicable radii was established based on [23]. Converting the two most frequent radii by $r_d = \text{round}(r_{\mu\text{m}}/k)$ leads to radii of $r_d=5$ and $r_d=16$ pixels. Since the width of peaks in the symmetry image allows also for the detection of objects with radii larger than those substituted to (Fig. 3b) the final range of applicable radii was delimited to $R_d = \{3, 14\}$ pixels. Shape criterion was applied to deal with spatial configuration and separation of binary objects. Typically, for low thresholds, a close localization of two or more nuclei may lead to superposition of contributing symmetry components, and yield a single large object with poor circularity. Such objects are erased until a higher threshold causes an appearance of objects with both criteria satisfied.

Lastly, we took the above experiential and morphological considerations into account and implemented them as a nuclei segmentation procedure. Its pseudo code is presented below. For one pass of the outer loop one radius from R_d is substituted and the $S_{r=r_d}$ is obtained. In the inner loop consecutive thresholds T from the range of $[\min(\mathbf{T}_{r=r_d}), \max(\mathbf{T}_{r=r_d})]$ are applied, and the shape features are evaluated for each binary object found. Objects with poor circularity or excessive area are removed. For all radii in the predefined range the circularity was arbitrary set to 1.5. The threshold margins were $T_{\max} = 5$, and $T_{\min} = 1$, with the increment of $\Delta = 0.1$. The final mask of nuclei $\hat{S}$ is returned a logical union of sub-masks obtained for processing of $S_{r=r_d} \forall r_d \in R_d$.

START: Determine nuclear mask $\hat{S}$

inititalize empty $\hat{S}$;

for all $r_d \in R_d$ do

calculate $S_{r=r_d}$;

inititalize empty $\bar{S}_r$;

$T_{\max} = \max(\mathbf{T}_{r=r_d}); T_{\min} = \min(\mathbf{T}_{r=r_d});$

$T = T_{\min};$

repeat

$$\bar{S}_r = \begin{cases} 1 & \text{if } S_{r=r_d} \geq T \\ 0 & \text{otherwise} \end{cases};$$

$$Objects = \text{Labeling}(\bar{S}_r);$$

$$O = Objects \text{ with } (area_{obj} \geq \pi r_d^2) \cup (circularity_{obj} \geq c);$$

$$\bar{S}_r = \bar{S}_r - O;$$

$$\bar{S}_r = \bar{S}_r \cup \bar{S}_r;$$

$$T = T + \Delta;$$

until $T > T_{max}$

$$\widehat{S} = \widehat{S} \cup \bar{S}_r;$$

end for

STOP where: Labeling is a procedure for detecting connected regions in binary images [24], $area_{Obj}$ and $circularity_{Obj}$ are the variables storing the value of area and circularity of binary objects, and c is the circularity constraint. An object with c equal to 1 is circular, whereas $c > 1$ can represent an ellipse or a line segment. $\widehat{S}$ is the final mask of nuclei.

p16/Ki67 features of nuclear immunoreactivity

According to the workflow in Figure 2 the availability of nuclear mask allows for the extraction of colorimetric signals of immunoreactivity. Since no specific guidelines exist, we chose original RGB color space to characterize the expression of Ki67 and p16 markers. First, two sets of features in the training set of nuclei were formed: a) kernel density estimates (kde) that can model the multimodality, and probability density function in sparse data samples (Fig. 4a), and b) colorimetric signatures (s) reflecting contribution of particular immunoreactivities in the nucleus (Fig.4c). A nonparametric kde approach [25] applied to estimate main modes of RGB intensities yielded a $kde_i = [kde_R, kde_G, kde_B]$, vector for i -th nucleus $i \in [1, N]$. Each vector received the same immunoreactivity label as the pattern nucleus it originated from (Fig.4a). The second feature required nuclear pixels to be labeled in advance. In this case pixels of i -th pattern nucleus were classified against all labeled kde vectors excluding i -th via kNN ($n=1$) classification (Fig.4b). As a result a four element signature $s_i = [q_{p16+/Ki67-}, q_{p16-/Ki67+}, q_{p16+/Ki67+}, q_{p16-/Ki67-}]$ was obtained for i -th pattern nucleus: where q is the normalized quantity of pixels assigned to respective immunoreactivity classes (Fig.4c). One component of s_i represents the fraction of pixels that share the same label of immunoreactivity as assigned by kNN. The normalization was carried out by dividing individual quantities by the total number of pixels within the nucleus area. All s_i signatures were labeled in the same manner as kernel density estimates. In the testing set of nuclei the kde vectors were obtained in an identical manner whereas the signatures were obtained via classification of individual pixels against all labeled kde_i vectors from the training set (Fig.4d).

An absolute nuclear immunoreactivity as interpreted by a human observer can fall into one of the four predefined classes. Mimicking this interpretation task by an imaging-based classification presents a challenge in recognizing the respective class from the bulk of pixels of different color and intensity. A signature can represent this bulk of signals in a systematic way. In our case kde is first used to approximate the modes of basic color intensities, whereas s provides quantities of initially classified pixels. Since the intensities attributed to different immunoreactivities vary across the nuclear area the signature seems to be a feature that is more suitable for discriminating the mixture of individual signals than the kde does. We tested discriminating power of the signatures and kernel density estimates in classification schemes involving SVM and kNN classifiers.

Classification of nuclei

We implemented two one-against-all multiclass SVM classifiers: first one (coarse) trained with $x_i = kde_i$ vectors, and second one (fine) trained with colorimetric signatures $x_i = s_i$ to arrive at individual nucleus classification (Fig.4d). A three-dimensional ($m=3$) and a four dimensional ($m=4$) feature spaces were used for training of the two respective SVM classifiers. The set of class labels consisted of expert defined classes: $y_i \in \{p16^-/Ki67^+, p16^+/Ki67^-, p16^+/Ki67^+, p16^-/Ki67^-\}$ and was used for training of SVMs. Prior to training the features of pattern nuclei were normalized by subtracting respective means and dividing by the standard deviations. To improve outcomes of the SVMs the features were mapped nonlinearly into a high-dimensional space by the use of the Gaussian kernel function:

$\phi(x_i, x_j) = \exp(-\|x_i - x_j\|^2 / \sigma^2)$, $x_{i,j} \in \mathbb{R}^m$. The optimal kernel width of $\sigma = 0.5$ was found by performing 15-fold cross-validation in the training set. More details on SVMs can be found in [26]. Normalization of features from tested nuclei was respectively performed using the means and standard deviations from the training set of data. Our four-class classification problem was decomposed into four one-against-all binary classifications, with a *max-win* voting tactic adopted for classification of the testing instances. Ultimately the coarse classification provides quantities to form a colorimetric signature of immunoreactivity, whereas the fine one classifies that signature by assigning a nucleus under consideration to one of the class label y_i . The combined method consisting of these two consecutive SVM classifiers is termed CF-SVM.

Performance evaluation

To facilitate the evaluation process nuclei processed by the proposed algorithms were automatically outlined and color-coded on the original image. Our experts manually estimated each outline as true positive (TP), false positive (FP) or false negative (FN) respectively. A FP detection was assigned if the outline was found in portions of cytoplasm or artifacts outside of nuclei, and a FN for a missed nucleus respectively. A TP finding was considered if the outline enclosed or was located within the nucleus. Segmentation accuracy was measured by metrics commonly used in pattern recognition tasks: precision = $TP / (TP + FP)$, and recall = $TP / (TP + FN)$ (sensitivity). Apart from this, the nucleus area agreement between automated results (A) and ground truth tracings (G) by two experts was evaluated

by means of the Jaccard index $J = \frac{G \cap A}{G \cup A}$, where $\cap$ and $\cup$ are the logical conjunction and disjunction operators of the binary areas to be compared. $J \in [0,1]$, and maximal agreement is reached if the ground truth and the automatically generated outlines are identical.

The proposed CF-SVM classification was compared to kNN_{kde} classifiers ($n=1$ and 5), and a SVM_{kde} classifier solely utilizing kde features extracted from the training set of nuclei. Evaluation of cell-based classification performances involved sensitivity and specificity analyses as well. Precision articulated the tradeoff between TP and FP classifications for individual classes, whereas recall (sensitivity) was used to compare expert's and machine-

based findings from all nuclei with positive (either single or dual) marker over-expression. The specificity = $TN/(TN + FP)$ reflected the detection rate of nuclei with negative expression marked as true negative (TN). All parameters involved in image pre-processing, nuclei segmentation, and classification were derived from the training data, artificial images or prior anatomical knowledge and fixed as described in text.

Results

Utilizing the methodology we developed, cell nuclei were first automatically localized in digitized slides and then two features were extracted to recognize nuclear immunostaining. For nuclei segmentation assessment several images from the test slides were singled out, preprocessed with and without anisotropic diffusion and shown to two cytopathologists who manually identified all cells ($n=4228$) and categorized the segmentation results into TP, FN and FP categories. Example segmentation outputs are shown in Figures 5a-5b, and respective findings are collected in Table 1. Figures 6 and 7 demonstrate the methods' performance in clustered cells.

For nucleus outline evaluation an I_{FOV} containing in total 384 cells with various expressions of immunoreactivity was selected. The I_{FOV} was pre-processed and the average $J_{\%}$ index from pooled individual Jaccard indices was derived. $J_{\%}$ was 81.4% for comparison of Observer 1 and Observer 2, whereas it was 76.5% and 72.1% for comparison of Observer 1 and Observer 2 vs. the automated processing respectively. Since an improvement was noted (Tab.1) we permanently embedded the anisotropic diffusion step into our analytical workflow

The major goal of screening in this setting is to correctly identify the presence or absence of nuclei with dual $p16^{+}/Ki67^{+}$ over-expression. Along with the detected immunoreactivities the nuclei were color-coded to facilitate the reading and assessment of the results. In this part of study 18068 nuclei were automatically analyzed. Classification rates by CF-SVM method evaluated by two observers are shown in a confusion matrix (Tab.2). Analogous matrices were formed for the remaining techniques and resulting performances are juxtaposed in Table 3.

Comparable sensitivities and specificities were obtained for the kNN_{kde} methods; however the precision in distinguishing individual immunoreactivities reached higher levels through the application of SVMs. Among these, the rates related to dual $p16^{+}/Ki67^{+}$ over-expression leveraged the precision to 77.1% and 82.6% respectively, whereas the sensitivity of all immunopositive nuclei detection was respectively 86.3% and 89.4% for the proposed CF-SVM classification. The specificities were found relatively high (better than 99%) regardless of the applied method.

Discussion

Recognition of abnormal cells in liquid-based cytological specimens is a challenging and complex task. Manual screening and interpretation which is currently a standard clinical practice is mostly qualitative and has limited prospects for quantitative evaluation of entire slides. Utilization of $p16/Ki67$ antibodies as surrogate biomarkers for hrHPV-related cervical neoplasias has prompted a search for tools to supplement and improve evaluation of cervicovaginal smears [11-13, 18]. Hence, when this dual immunostain became commercially available cell nuclei emerged as an attractive screening target. The main focus of this work is to establish and test the feasibility of QPI - a novel screening technique which utilizes quantification and classification of immunocytochemical features in images from liquid-based Pap smears. The proposed method analyzes features solely from cell nuclei that

constitute small, yet detectable and quantifiable fraction of the whole specimen. This distinguishes our approach from other techniques in which the delineation of cytoplasmic and nuclear areas are required [5-7, 10].

In a given cytological image the intensities of background and cellular components including the nuclei can vary depending on marker concentration and heterogeneity, and thereby negatively impact visual and automated screening routines. The newly proposed segmentation procedure does not make any assumptions about distribution of nuclei on the slide and to a large extent can deal with overlapping or isolated cells with different optical densities. This capability is enabled by adaptive searching for gradient-based radial symmetries coupled with anisotropic smoothing that enhanced relevant edges, boosted the segmentation performance, and thus lowered the rates of FP and FN detection. In the majority of cases, the FP detections were limited to dense and round corners of cell membranes, debris material, and other artifacts - all yielding p16⁺/Ki67⁺ expert classification. On the other hand approximately 85% of the FN detections (misses) were caused by weak nuclear staining - evaluated later by cytopathologist as non-immunoreactive.

The proposed approach involved a blind thresholding in the analytically determined range to preserve symmetry components of round objects and discard those originating from image noise and non-circular shapes. Intermediate binary images were iteratively refined by means of morphological features to arrive at a final mask of nuclei with predefined radii. The radii were implemented as a finite set of consecutive integers. The threshold increment and the circularity criterion were fixed as well. A consequence of such concept may be the abovementioned FN detection or a fluctuation of the output contour. In the latter case the resultant contour may be positioned at the nucleus boundary, exceed it, or belong to its interior. Of these, the cases with contours detected outside the nuclear territory were least frequent. Figures and comparisons with ground truth reflect both types of inaccuracies. Nevertheless, for our screening application it seems to be more acceptable to quantify immunoreactivity from a reduced nuclear area vs. an area corrupted by the adjacent cytoplasm or background. In case of nonimmunoreactive cells the border discrepancy may even be tolerable with a wider margin.

Currently the system can adequately resolve clumped cells of various nuclear sizes, intensities, and types of immunoresponses (Fig.5 and 6); however a higher cell density can pose limitations. Two or more nuclei forming a cluster can be separated only if the gradients located between their borders are strong enough. An overlap or close adjacency of two or more nuclei may weaken the individual overall symmetry signals (Fig. 6b and 7) and lead to capturing immunoreactivity from compromised nuclear areas. The nuclear contour error in handling these challenging cases may be larger than that encountered in non-overlapping cells. Major factors contributing to this effect are differences in shape and intensity between the adjacent nuclei. Despite several shortcomings our nuclei segmentation procedure appears to be well suited for the aforesaid purpose due to the following reasons: a) we extensively evaluated a much larger number of nuclei compared to other studies [5], [6] and [9], b) it maintains a high detection rate; with precision >93.3% and recall >92.8% regardless of the staining intensity compared to the 90% average effectiveness shown in [7] and recall of 90.57% or 69.86% in [8], and c) it shows good nuclear area agreement when measured against two human observers. Although the segmentation provides accurate outlines for the majority of nuclei, there were instances with partially encircled cytoplasm. It may skew colorimetric signatures in Ki67 positive cells because it is the cytoplasm rather than the nucleus that expresses distinct immunoreactivity. However, in all remaining types of cells analyzed, this effect will have a much smaller influence since the expression of markers in the cytoplasm will correspond to those found in the nucleus.

As there has been no better color space defined yet for this type of quantitative cytology, we selected RGB intensities and this selection led to satisfactory recognition rates by CF-SVM. Two approaches were tested: *kde*-based estimation of main intensity modes in RGB channels alone, and in conjunction with signatures derived from intermediate pixel-level classifications. Since nuclei that stain negative for p16/Ki67 show colorimetric patterns that are very distinct from the immunoreactive ones, their classification was very accurate in all tested methods. The performances of all techniques decreased as the complexity of patterns to be classified increased (Fig. 5, 6 and Tab. 3). While purely blue and red nuclear patterns were classified by reference techniques with relatively high confidence, the recognition of brown and reddish-brown nuclei required a more sophisticated approach. The solution involving colorimetric signatures proved to be more effective than the two tested kNN_{kde} and the SVM_{kde} classifications because quantification of signals was achieved at a higher level of scrutiny. The CF-SVM yielded lowest FP and FN rates and resulted in the highest precision and sensitivity indicating that the colorimetric signatures are more suitable than the main modes of RGB colors for quantitative dissection of complex immunocytochemical patterns in Pap smears. It may also be useful for the evaluation of other biomarkers utilized in cyto- and histopathology or cancer research. Along with the advantages described above, the proposed classification has several limitations. It may inherit some of the drawbacks attributed to supervised learning such as output dependence on training patterns' sparseness, redundancy, heterogeneity etc. On the other hand, manual selection of patterns has been frequently used in digital pathology to adjust for the non-uniformity of staining across specimens and optical characteristics of imaging modalities. The proposed system is also unable to distinguish nuclei of squamous epithelium from artifacts such as debris particles. Nevertheless, it can highlight cells with a high likelihood of hrHPV-related abnormality thus immediately drawing an expert's attention to the cells of interest and the need for further examination while reducing the screening effort that might otherwise have been spent on noninformative cells.

Conclusions

In this paper we address the problem of automated recognition of immunoreactivity in cervical smears through radial symmetry-based segmentation of nuclei followed by extraction, formation and classification of signatures - a derivative of hrHPV-specific over-expression. Experimental outcomes suggest that our proof-of-concept automated profiling tool can aid in human and machine-based analytical workflows of screening. Encouraged by satisfactory performance we plan to seek opportunities in improving robustness, extending evaluation studies on larger number of samples, and move towards high-throughput whole slide analysis to reach the consistency of expert human observers.

Acknowledgments

This work was supported in part by a grant from the Department of Surgery at Cedars-Sinai Medical Center, and in part by a NIH grant 5R21CA143618-02 (to AG). We also thank Dr Hunter Hardy M.D. for technical help in specimen imaging.

References

1. Byers T, Mouchawar J, Marks J, Cady B, Lins N, Swanson GM, Bal DG, Eyre H. The American Cancer Society challenge goals. How far can cancer rates decline in the U.S. by the year 2015? *Cancer*. 1999; 86:715–727. [PubMed: 10440701]
2. Sirovich BE, Welch HG. The frequency of Pap smear screening in the United States. *J Gen Intern Med*. 2004; 19:243–250. [PubMed: 15009779]

3. Nayar R, Solomon D. Second edition of 'The Bethesda System for reporting cervical cytology' - atlas, website, and Bethesda interobserver reproducibility project. *Cytojournal*. 2004; 1:4. [PubMed: 15504231]
4. Nanda K, McCrory DC, Myers ER, Bastian LA, Hasselblad V, Hickey JD, Matchar DB. Accuracy of the Papanicolaou test in screening for and follow-up of cervical cytologic abnormalities: a systematic review. *Ann Intern Med*. 2000; 132:810–819. [PubMed: 10819705]
5. Yang-Mao SF, Chan YK, Chu YP. Edge enhancement nucleus and cytoplasm contour detector of cervical smear images. *IEEE Trans Syst Man Cybern B Cybern*. 2008; 38:353–366. [PubMed: 18348920]
6. Mat-Isa NA. Automated Edge Detection Technique for Pap Smear Images Using Moving K-Means Clustering and Modified Seed Based Region Growing Algorithm. *International Journal of The Computer, the Internet and Management*. 2005; 13:45–59.
7. Sobervilla P, Montseny E, Vaschetto F, Lerma E. Fuzzy-Based Analysis of Microscopic Color Cervical Pap Smear Images: Nuclei Detection. *International Journal of Computational Intelligence and Applications*. 2010; 9:187–206.
8. Plissiti ME, Nikou C, Charchanti A. Automated detection of cell nuclei in pap smear images using morphological reconstruction and clustering. *IEEE Trans Inf Technol Biomed*. 2011; 15:233–241. [PubMed: 20952343]
9. Siddiqi AM, Li H, Faruque F, Williams W, Lai K, Hughson M, Bigler S, Beach J, Johnson W. Use of hyperspectral imaging to distinguish normal, precancerous, and cancerous cells. *Cancer*. 2008; 114:13–21. [PubMed: 18213691]
10. Schubert JM, Bird B, Papamarkakis K, Miljkovic M, Bedrossian K, Laver N, Diem M. Spectral cytopathology of cervical samples: detecting cellular abnormalities in cytologically normal cells. *Lab Invest*. 2010; 90:1068–1077. [PubMed: 20368702]
11. Longatto Filho A, Utgawa ML, Shirata NK, Pereira SM, Namiyama GM, Kanamura CT, Santos Gda C, de Oliveira MA, Wakamatsu A, Nonogaki S, Roteli-Martins C, di Loreto C, Mattosinho de Castro Ferraz Mda G, Maeda MY, Alves VA, Syrjanen K. Immunocytochemical expression of p16INK4A and Ki-67 in cytologically negative and equivocal pap smears positive for oncogenic human papillomavirus. *Int J Gynecol Pathol*. 2005; 24:118–124. [PubMed: 15782067]
12. Sahebali S, Depuydt CE, Boulet GA, Arbyn M, Moeneclaey LM, Vereecken AJ, Van Marck EA, Bogers JJ. Immunocytochemistry in liquid-based cervical cytology: analysis of clinical use following a cross-sectional study. *Int J Cancer*. 2006; 118:1254–1260. [PubMed: 16161049]
13. Walts AE, Bose S. p16, Ki-67, and BD ProExC immunostaining: a practical approach for diagnosis of cervical intraepithelial neoplasia. *Hum Pathol*. 2009; 40:957–964. [PubMed: 19269017]
14. Juric D, Mahovic V, Rajhvajn S, Ovanin-Rakic A, Skopljanac-Macina L, Barisic A, Projic IS, Babic D, Susa M, Corusic A, Oreskovic S. Liquid-based cytology--new possibilities in the diagnosis of cervical lesions. *Coll Antropol*. 2010; 34:19–24. [PubMed: 20432728]
15. Tsoumpou I, Arbyn M, Kyrgiou M, Wentzensen N, Koliopoulos G, Martin-Hirsch P, Malamou-Mitsi V, Paraskevidis E. p16(INK4a) immunostaining in cytological and histological specimens from the uterine cervix: a systematic review and meta-analysis. *Cancer Treat Rev*. 2009; 35:210–220. [PubMed: 19261387]
16. Dunton CJ, van Hoven KH, Kovatich AJ, Oliver RE, Scacheri RQ, Cater JR, Carlson JA Jr. Ki-67 antigen staining as an adjunct to identifying cervical intraepithelial neoplasia. *Gynecol Oncol*. 1997; 64:451–455. [PubMed: 9062149]
17. Scholzen T, Gerdes J. The Ki-67 protein: from the known and the unknown. *J Cell Physiol*. 2000; 182:311–322. [PubMed: 10653597]
18. Bose S, Evans H, Lantzy L, Scharre K, Youssef E. p16(INK4A) is a surrogate biomarker for a subset of human papilloma virus-associated dysplasias of the uterine cervix as determined on the Pap smear. *Diagn Cytopathol*. 2005; 32:21–24. [PubMed: 15584047]
19. Meyer JL, Hanlon DW, Andersen BT, Rasmussen OF, Bisgaard K. Evaluation of p16INK4a expression in ThinPrep cervical specimens with the CINtec p16INK4a assay: correlation with biopsy follow-up results. *Cancer*. 2007; 111:83–92. [PubMed: 17334990]

20. Sahebali S, Depuydt CE, Segers K, Vereecken AJ, Van Marck E, Bogers JJ. Ki-67 immunocytochemistry in liquid based cervical cytology: useful as an adjunctive tool? *J Clin Pathol.* 2003; 56:681–686. [PubMed: 12944552]
21. Perona P, Malik J. Scale-Space and Edge Detection Using Anisotropic Diffusion. *IEEE Trans. Pattern Anal. Mach. Intell.* 1990; 12:629–639.
22. Loy G, Zelinsky A. Fast Radial Symmetry for Detecting Points of Interest. *IEEE Trans. Pattern Anal. Mach. Intell.* 2003; 25:959–973.
23. Reagan JW, Hamonic MJ. The cellular pathology in carcinoma in situ; a cytohistopathological correlation. *Cancer.* 1956; 9:385–402. [PubMed: 13304855]
24. Haralick, RM.; Shapiro, LG. Computer and robot vision. Addison-Wesley Pub. Co.; Reading, Mass.: 1992.
25. Botev ZI, Grotowski JF, Kroese DP. Kernel Density Estimation Via Diffusion. *Ann Stat.* 2010; 38:2916–2957.
26. Schölkopf, B.; Burges, CJC.; Smola, AJ. Advances in kernel methods : support vector learning. MIT Press; Cambridge, Mass.: 1999.

Figure 1. Images of Pap smears with applied dual p16/Ki67 immunocytochemical staining. Single positive immunoreactive cells: p16⁻/Ki67⁺ or p16⁺/Ki67⁻ and dual positive Ki67⁺/p16⁺ are pointed by arrows. Relevant colorimetric signals are located predominantly in the nuclei.

Figure 2.

Pap smear image analysis workflow. Main steps involve segmentation of nuclear mask, extraction of colorimetric signatures and classification leading to a quantitative readout. The classification step utilizes features from cells manually labeled by cytopathologist.

Figure 3.

Object detection in example radial symmetry images. An artificial image with circular objects darker than the surrounding background a) is subjected to symmetry transform S_r with different radius parameter ($r=3, 5$ and 7). b) Profiles in individual images demonstrating circular object detection using a intensity thresholding. Threshold T_5 signifies that the object with $r_d=5$ pixels can be detected in S_5 . T_3 , T_5 and T_7 are the uppermost possible levels to detect the respective objects in symmetry images. All thresholds detecting an object of radius r_d in S_r in a wide range of object-background intensities can be determined analytically and collected in the matrix $\mathbf{T}_r$ - figure c).

Figure 4.

p16/Ki67 immunoreactivity features and SVM-based classification scheme: a) kde vectors in RGB color space, b) feature formation process in the training set of nuclei, c) signatures of nuclei with over-expression of markers: Ki67 (1-2), p16 (3-4), dual p16/Ki67 (5-6), and no expression (7-8), d) coarse-fine classification for testing nuclei.

Figure 5.

Example nuclei quantification results: a) original image, b) original image with superimposed perimeters of segmented nuclei, c) color coded classification of immunoreactivity: red for p16⁻/Ki67⁺, brown for p16⁺/Ki67⁻, yellow for dual p16⁺/Ki67⁺ and blue for p16⁻/ki67⁻.

Figure 6.
Nuclei segmentation performance in overlapping cells: a) original images, b) segmentation results.

Figure 7. Nuclei segmentation performance in clusters of cells: a) original images, b) segmentation results.

Table 1

Nuclei segmentation performance in immunostained Pap smears.

Preprocessing algorithm	# of detected nuclei	Cytopathologist evaluation				Precision [%]	Recall [%]
			TP	FP	FN		
None	4053	Observer 1	3426	627	802	84.52	81.03
		Observer 2	3391	662	802	83.66	80.87
Anisotropic smoothing	4174	Observer 1	3933	241	295	94.22	93.02
		Observer 2	3896	278	295	93.33	92.86

Evaluation of the proposed CF-SVM nuclei classification method by two observers. Respective fields show number of nuclei assigned to different immunoreactivity categories.

Table 2

Algorithm-based classification	Cytopathologist evaluation											
	p16-/Ki67+		p16+/Ki67-		p16-/Ki67+		p16+/Ki67-		p16-/Ki67-		p16+/Ki67-	
	Obs.1	Obs.2	Obs.1	Obs.2	Obs.1	Obs.2	Obs.1	Obs.2	Obs.1	Obs.2	Obs.1	Obs.2
	No. of nuclei		Obs.1, Obs.2		Obs.1, Obs.2		Obs.1, Obs.2		Obs.1, Obs.2		Obs.1, Obs.2	
p16-/Ki67+	738	733	10	12	4	3	34	38	34	38	34	38
p16+/Ki67-	2	6	172	168	3	2	31	31	31	31	31	31
p16+/Ki67+	4	7	7	8	57	54	1	1	1	1	1	1
p16-/Ki67-	37	62	47	50	0	1	16921	16892	16921	16892	16921	17005
Total	781	808	236	238	64	60	16987	16962	16987	16962	16987	17005

Table 3
Comparison of pooled nuclei classification performances from different techniques.

Automated analysis vs. Observer 1				
	KNN _{kde} (n=1)	KNN _{kde} (n=5)	SVM _{kde}	CF-SVM
p16 ⁺ /Ki67 ⁺	78.4	89.8	88.3	93.8
p16 ⁺ /Ki67 ⁻	62.1	58.3	74.3	82.7
p16 ⁻ /Ki67 ⁺	35.3	45.5	70.2	82.6
Sensitivity [%]	53.6	56.2	71.9	89.4
Specificity [%]	99.4	99.5	99.5	99.6

Automated analysis vs. Observer 2				
	KNN _{kde} (n=1)	KNN _{kde} (n=5)	SVM _{kde}	CF-SVM
p16 ⁺ /Ki67 ⁺	77.9	89.1	87.9	93.2
p16 ⁺ /Ki67 ⁻	61.6	57.7	72.5	81.2
p16 ⁻ /Ki67 ⁺	34.6	42.3	67.4	77.1
Sensitivity [%]	51.5	58.4	71.1	86.3
Specificity [%]	99.4	99.4	99.5	99.6