Chalmers Publication Library

Novel chromatin texture features for the classification of Pap smears

Citation for the published paper:
Ehteshami Bejnordi, B. ; Moshavegh, R. ; Sujathan, K. (2013) "Novel chromatin texture features for the classification of Pap smears". Proceedings of SPIE. Medical Imaging 2013: Digital Pathology. 9th Conference on Digital Pathology at the SPIE Medical Imaging Symposium. Lake Buena Vista, FL. FEB 10-11, vol. 8676

http://dx.doi.org/10.1117/12.2007185

(article starts on next page)

Novel Chromatin Texture Features for the Classification of Pap Smears

Babak Ehteshami Bejnordi^{ac}, Ramin Moshavegh^{ac}, K. Sujathan^{b},
Patrik Malm^{d}, Ewert Bengtsson^{d}, and Andrew Mehnert^{ac}

^{a}Department of Signals and Systems, Chalmers University of Technology, Gothenburg, Sweden;
^{b} Department of Pathology, Regional Cancer Centre (RCC), Thiruvananthapuram, India;
^{c} MedTech West, Sahlgrenska University Hospital, Gothenburg, Sweden;
^{d} Centre for Image Analysis, Uppsala University, Uppsala, Sweden

ABSTRACT

This paper presents a set of novel structural texture features for quantifying nuclear chromatin patterns in cells on a conventional Pap smear. The features are derived from an initial segmentation of the chromatin into blob-like texture primitives. The results of a comprehensive feature selection experiment, including the set of proposed structural texture features and a range of different cytology features drawn from the literature, show that two of the four top ranking features are structural texture features. They also show that a combination of structural and conventional features yields a classification performance of 0.954±0.019 (AUC±SE) for the discrimination of normal (NILM) and abnormal (LSIL and HSIL) slides. The results of a second classification experiment, using only normal-appearing cells from both normal and abnormal slides, demonstrates that a single structural texture feature measuring chromatin margination yields a classification performance of 0.815±0.019. Overall the results demonstrate the efficacy of the proposed structural approach and that it is possible to detect malignancy associated changes (MACs) in Papanicoloau stain.

Keywords: Cervical cancer, feature extraction, feature selection, classification, chromatin texture, malignancy associated changes, structural texture features, Pap smear classification

1. INTRODUCTION

The Papanicolaou (Pap) test is the primary screening test for cervical cancer.[1] It involves the microscopic examination of cells sampled from the cervix. It is laborious and time-consuming involving the review of possibly hundreds of thousands of cells for visual signs of cancerous or precancerous changes.[2] One in every 10 to 20 positive cases is missed in routine screening.[1] Two major factors affect the accuracy of the Pap test. The first is sampling error wherein no diagnostic cells make it on to the slide. The other is interpretation error for reasons including fatigue, inexperience, and habituation. Computer-assisted interpretation can potentially address the issue of interpretation error. The malignancy-associated change (MAC) phenomenon may potentially address sampling error. Several studies have demonstrated the phenomenon.[3–5] MACs are subtle sub-visual, i.e. visually imperceptible, changes in the appearance of otherwise normal-looking cells from an abnormal Pap slide. MACs cannot be detected on a cell-by-cell basis but rather in a population of cells. The MAC approach obviates the need to perform an exhaustive review of all of the cellular material to identify diagnostic cells.

The MAC features that appear to have the most discriminatory power are those characterizing nuclear texture or more specifically the chromatin pattern in the nucleus. Traditionally MAC features have been based on a stochastic approach to defining texture. However such features do not correspond well to the terms used by cytopathologists to describe chromatin texture such as heterogeneity, granularity, margination, condensation, compaction, clumping, diffuse, blobs and particles.[2] This motivates interest in a structural approach to chromatin texture analysis. This paper presents a set of novel structural texture features for quantifying chromatin texture.

B. Ehteshami Bejnordi: (✉) ehteshami@live.se
R. Moshavegh: (✉) raminmoshavegh@gmail.com
A. Mehnert: (✉) andrew.mehnert@chalmers.se

The features are derived from an initial segmentation of the chromatin into blob-like texture primitives. The paper also presents an investigation of the most discriminatory subset of features, from among the proposed features and a wide range of features drawn from the literature, for discriminating between normal and abnormal Pap smears using the MAC approach.

2. METHODS

The types of features that can be extracted from cytology images can be broadly classified into four classes: morphometric, densitometric, textural, and contextual.[6] Whilst each of these can conceptually be computed for both nuclei and chromatin blobs/particles, in this study contextual features were excluded for nuclei, and textural and densitometric features were excluded for chromatin blobs. Morphometric features describe the geometry (shape, size, position and boundary) of the object and are computed from its binary mask. Densitometric features describe the intensity or optical density of the object and are computed from the histogram of gray-values within the object. Contextual features are derived from the spatial relationship between objects; e.g. the number of neighbors an object has, and statistics of the distances to neighboring objects. Texture features describe the spatial variation of gray-levels within an object. Such features can be broadly classified into statistical features and structural/syntactic features. Statistical texture features used in this study include fractal texture features, run-length features, histogram based features, gray-level co-occurrence matrix (GLCM) features, and complex Daubechies wavelets features. The wavelet features used in this study to characterize nuclear texture were computed from the first level of the wavelet decomposition.[7] A down-sampled binary mask of the original image is used to identify and retain only the nuclei regions in the decomposed images for the purpose of feature extraction.

The structural texture features used in this study to characterize nuclear texture are derived from a segmentation of the nuclear chromatin into blob-like texture primitives. In particular they are derived from morphometric and contextual features computed for these blobs. The approach is described in the next section.

2.1 Novel structural chromatin texture features

Our proposed structural chromatin texture features are based on an initial segmentation of the chromatin into blob-like texture primitives.[2] For a given nucleus this involves: (i) applying a 3×3 median filter (to attenuate impulse noise); (ii) up-sampling by a factor of 3 (to facilitate the rendering of watershed lines in a subsequent step); (iii) locating the regional minima (inner markers); (iv) computing the watershed transform of the filtered image with respect to the inner markers (to produce an outer marker that delineates a zone of influence (ZOI) around each regional minimum); (v) computing the magnitude of gradient image for the filtered image; and (vi) applying the watershed transform to the gradient image with respect to both the inner markers and the outer marker (to delineate each dark chromatin blob). Our structural features (see Table 1) in essence describe the attributes and arrangement of these resulting chromatin blobs. They fall into two broad categories. The first consists of statistics of morphometric features computed for individual particles in a nucleus. The second consists of contextual features computed from the particles in a nucleus. These contextual features can be further subdivided into the following classes of features: margination, clustering, blob features, discrete texture features, and features derived from the Voronoi diagram and its associated graphs.

The margination features characterize the distances of segmented chromatin blobs to the nucleus boundary. They are computed from the cookie-cutting distances[2] illustrated in Figure 1. In this study a 6th order polynomial was fitted to the counts in the cumulative frequency histogram of these cookie-cutting distances, and the slopes at the 25th and 37th percentiles were used as margination features $M1$ and $M2$ respectively for each nucleus. The mean and standard deviation of these cookie-cutting distances were used as features $M3$ and $M4$ respectively. The four features implicitly include information about the size of the nucleus. For this reason it was decided in this study to include features $M5$ and $M6$ defined to be $M3$ and $M4$, respectively, each normalized to the maximum value of the distance transform of the nucleus. The last margination feature, $M7$, is defined to be the minimum distance between the centroids (geometrical centers) of the chromatin particles and the nucleus boundary.

Table 1: Summary of the extracted nucleus features

Morphometric Features
Area, perimeter, mean-radius, elliptic variance[8]

Densitometric Features
Photometric features:[6] Integrated optical density (IOD), mean optical density (MOD), variance of optical density, skewness of optical density, kurtosis of optical density

Texture Features	
Statistical	**Fractal texture features:[9]** Fractal area 1, fractal area 2, fractal dimension
	Run length features:[3] SRE, LRE, GLN, RLN, RP, LGRE, HGRE, SRLGE, SRHGE, LRLGE, LRHGE
	Histogram based features:[9] Mean, standard deviation, and skewness of the gray level histogram
	GLCM features:[6] Contrast, correlation, energy, entropy, local homogeneity, maximum probability, cluster shade, cluster prominence and information measure of correlation (H-IMC1, H-IMC2)
	Complex Daubechies wavelet features of nuclei:[7] Mean, standard deviation, and skewness of the gray level histogram. GLCM features – contrast, correlation, energy, entropy, local homogeneity, maximum probability, cluster shade, cluster prominence and information measure of correlation (W-IMC1, W-IMC2)

Statistics of the chromatin particle morphometric features	
Structural	Mean of chromatin particle areas and perimeters, mean of chromatin particle areas normalized to the nucleus area, mean of chromatin particle perimeters normalized to the nucleus perimeter, chromatin particle compactness (P2A)

Contextual chromatin particle features

Margination: M1, M2, M3, M4, M5, M6, M7 (See section 2.1)

Clustering: C1, C2, C3, C4 (See section 2.1)

Blob features: Heterogeneity (the ratio between the area of the segmented dark and light regions, and the nucleus area),[10] the ratio between the area of the segmented dark regions and the nucleus area, number of segmented dark particles, number of segmented light particles, area ratio of each chromatin particle to its watershed ZOI, average distance between the geometrical center of the nucleus and pixels of all chromatin particles, distance between the geometrical center of the nucleus and center of mass of the chromatin particles

Discrete texture features:[9, 11] Medium and high DNA amount, medium and high DNA area, medium and high DNA compactness, medium-high DNA compactness, center of gravity (the distance from the geometrical center of the blob to the center of gravity of the optical density function)

Area-Voronoi diagram: Mean and standard deviation (SD) of the areas of the Voronoi regions, mean and SD of the areas of the Voronoi regions normalized to the nucleus area, area disorder,[3] Voronoi regions roundness,[3] mean of area ratio of each chromatin particle to its Voronoi region[3]

Delaunay graph and generalized Delaunay graph: Mean of the Delaunay triangle areas, mean and SD of the Delaunay edge lengths, average of the mean and SD of the edge lengths connected to each vertex (chromatin particle), maximum Delaunay edge length, mean of the number of connections per chromatin particle, number of vertices (chromatin particles) on the graph boundary

Gabriel graph and relative neighborhood graph: Mean and SD of the graph edge lengths, maximum edge length, average of the mean and SD of the graph edge lengths connected to each vertex (chromatin blob), mean of the number of connections per vertex (chromatin blob)

Minimum spanning tree: Mean and SD of the edge lengths, total edge length, edge disorder,[3] minimum to maximum edge ratio, mean of the number of connections per vertex, percentage of vertices connected to one vertex (MST1), percentage of vertices connected to two vertices (MST2), percentage of vertices connected to more than two vertices (MST3)
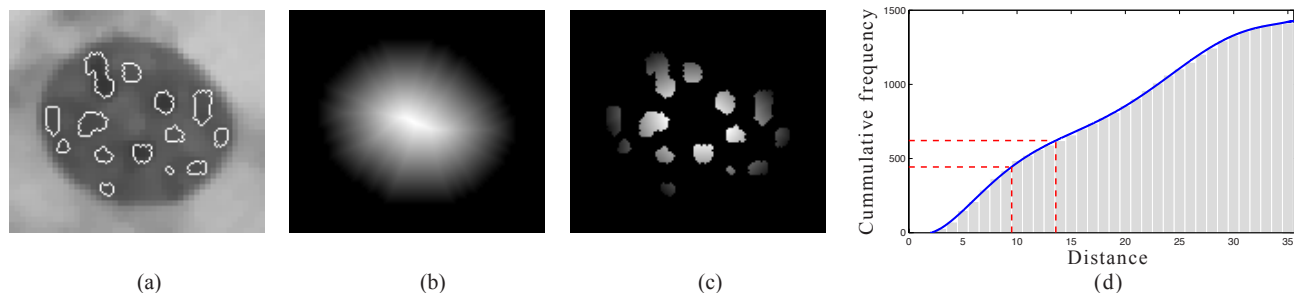
Figure 1: Illustration of how the features $M1$ and $M2$ are computed. (a) Segmented chromatin particles. (b) Euclidian distance transform of the nucleus. (c) Portions of *(b)* cut out by the segmented chromatin particle masks. (d) Cumulative frequency histogram of these cut-out distances with a polynomial fit through the counts superimposed. Features $M1$ and $M2$ are the slopes at the 25th and 37th percentiles respectively.

A subset of the contextual features are those derived from the Voronoi diagram and a generalization called the area Voronoi diagram. In this study the Delaunay graph (dual of the Voronoi diagram) is computed for the centroids of the particles in the nucleus and several features are then directly computed from it. The area Voronoi diagram is computed from the watershed transform of the distance transform of the set complement of the union of the chromatin particle masks. The region adjacency graph defined on the resulting watershed regions yields a generalization of the Delaunay graph as shown in Figure 2. Several sub-graphs—Gabriel graph, relative neighborhood graph, and the minimum spanning tree[12–14]—are also computed. Each type of graph defines a different adjacency relationship between the particles. Details concerning the explicit features computed from the Voronoi diagram and associated graphs are given in Table 1.
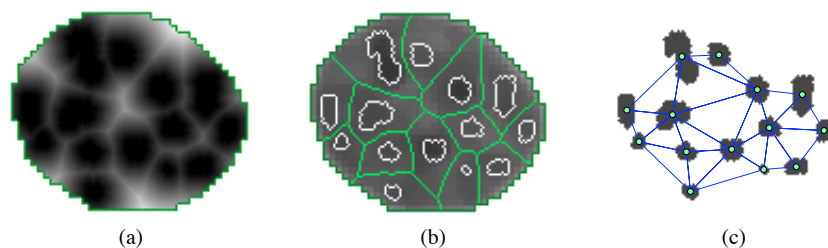

Figure 2: Illustration of how the area Voronoi diagram and generalized Delaunay graph are computed. (a) Distance transform of the set complement of the union of the chromatin particle masks in Figure 1a. (b) Area Voronoi diagram and segmented chromatin particles. (c) The region adjacency graph defined on the Voronoi regions then defines a generalized Delaunay graph.

The clustering features characterize the way the chromatin particles are clustered in the nucleus. They are computed from the area Voronoi diagram. Features $C1$ and $C2$ are computed similarly to features $M1$ and $M2$ except that the cookie-cutting distances are obtained from the distance transform of the Voronoi regions as shown in Figure 3. Features $C3$ and $C4$ are computed from the area Voronoi diagram and the distance transform image used in its construction (see Figure 2a and 2b) as follows. Firstly for every pair of Voronoi regions that share a common edge, the convex hull is computed for the union of the corresponding pairs of particles. Secondly the union of these convex hulls is filled to form a binary mask (Figure 3b). Thirdly this mask is intersected with the divide lines of the area Voronoi diagram. Finally features $C3$ and $C4$ are the mean and standard deviation, respectively, of the distance transform values corresponding to the residual divide line pixels in this intersection (Figure 3c).

Details concerning the blob features and the discrete texture features are given in Table 1.
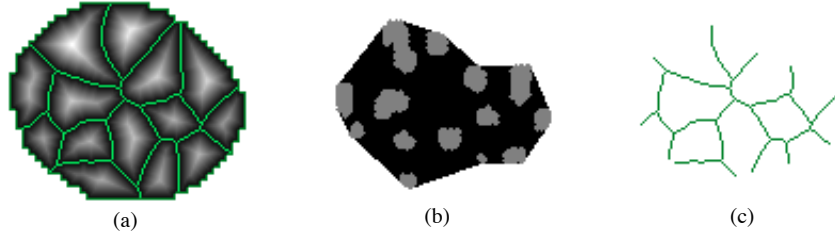
Figure 3: Illustration of how the features $C1$, $C2$, $C3$, and $C4$ are computed. (a) Distance transform of the Voronoi regions in Figure 2b. (b) Binary mask (with particles superimposed) formed from the union of the convex hulls (see the text for details). (c) Voronoi region divide lines clipped according to the mask in *(b)*.

2.2 Pap smear images and ground truth

The image data used in this study originate from a set of 68 conventional Pap smear slides. Each slide was reviewed by a cytopathologist and assigned a cytological diagnosis according to the Bethesda system.[15] The cytopathologist subsequently acquired representative FOVs from each slide. In the case of abnormal smears this included FOVs with and without diagnostic cells. The cytopathologist also labeled individual cells in each FOV accordingly. The FOV images were acquired using a monochrome CCD camera mounted on a light microscope with a 40× objective lens. Each image has a gray-scale resolution of 8 bits per pixel and is of size 1024×1344 pixels with square pixels of size 0.25μm.

2.3 Experiments

To evaluate the performance of the proposed structural chromatin texture features two experiments were performed. The aim of the first experiment was to determine the most discriminatory subset of features, from among the proposed features and a wide range of features drawn from the literature, for discriminating between normal and abnormal Pap smear slides. A normal smear is deemed to be one that has been labeled as NILM (Negative for Intraepithelial Lesion for Malignancy). An abnormal smear is one that has been labeled as either LSIL (Low grade squamous intraepithelial lesion) or HSIL (High grade squamous intraepithelial lesion). The aim of the second experiment was to evaluate the performance of a variety of classifiers built using the feature subset obtained in the first experiment to discriminate between the normal and abnormal slides.

In the first experiment, the free-lying nuclei in each digitized FOV were automatically segmented using the algorithm we have previously described.[16] For each nucleus we computed the 159 features listed in Table 1. The mean and standard deviation of each feature over all segmented nuclei for a slide yielded a total of 318 slide-based features. An additional random feature was included to gauge the efficacy of feature selection. As a compromise between having a reasonable number of cells per slide for MAC analysis and having a reasonable number of slides for our study we decided to impose the following two criteria. Firstly each slide should have a minimum of 80 non-superficial cervical cells with a NILM label. Secondly, 10% to 40% of the cells in the abnormal slides should be diagnostic cells. This left a total of 44 slides: 25 NILM, 10 LSIL, and 9 HSIL. The following three feature selection methods were investigated: state-of-the-art multiple support vector machines with recursive feature elimination (MSVM-RFE),[17] L1-regularization path for generalized linear models,[18] and the recently introduced guided regularized random forest (GRRF).[19] To optimize classification performance it is necessary to couple each classification approach with the most relevant feature selection method. In this experiment we paired the MSVM-RFE with a linear kernel SVM classifier, the GRRF with an ordinary RRF[20] classifier, and the L1-regularization procedure with a generalized linear model. Given the limited number of slides available in this study, all the slides were used for feature selection. However an internal cross validation (in-loop feature selection) technique[21] was employed to avoid overfitting and overoptimistic estimation of the performance. Feature selection by the SVM-RFE was achieved using 100 iterations of double 5-fold cross validation, each iteration with random combinations of samples in the training and the test set. The inner cross validation loop was used for feature selection and tuning the optimal parameters for the SVM classifier with a linear kernel. The outer cross validation loop was used to evaluate the classifier model performance in terms of the AUC (area under the receiver operating

characteristic curve). For GRRF feature selection and the L1-regularization path for generalized linear models, 100 repeated 5-fold cross validation was employed.

In the second experiment only the cells marked by the cytologist as NILM on normal and abnormal slides were segmented and slides with a minimum of 80 such cells retained. This left a total of 44 slides: 26 NILM, 9 LSIL, and 9 HSIL. To avoid the curse of dimensionality, only feature subsets with one or two features identified in experiment 1 were considered. Two classifiers were evaluated using 100 iterations of 5-fold cross validation: SVM with a linear kernel and a logistic regression (LR) classifier.

3. RESULTS

Experiment 1: The classification performance results together with the four top-ranked features for the three feature selection approaches are presented in Table 2. "SD of IOD" and "Mean of $M2$" are the top two features in all three feature selection (FS) methods. The top four features appearing most frequently among the feature selection methods are: SD of IOD, Mean of $M2$, SD of $MST3$ and Mean of W-$IMC1$. Figure 4a shows boxplots for each of these features for both the normal and abnormal slides. The best AUC (0.954±0.019) was achieved using the L1-regularization procedure with a generalized linear model.
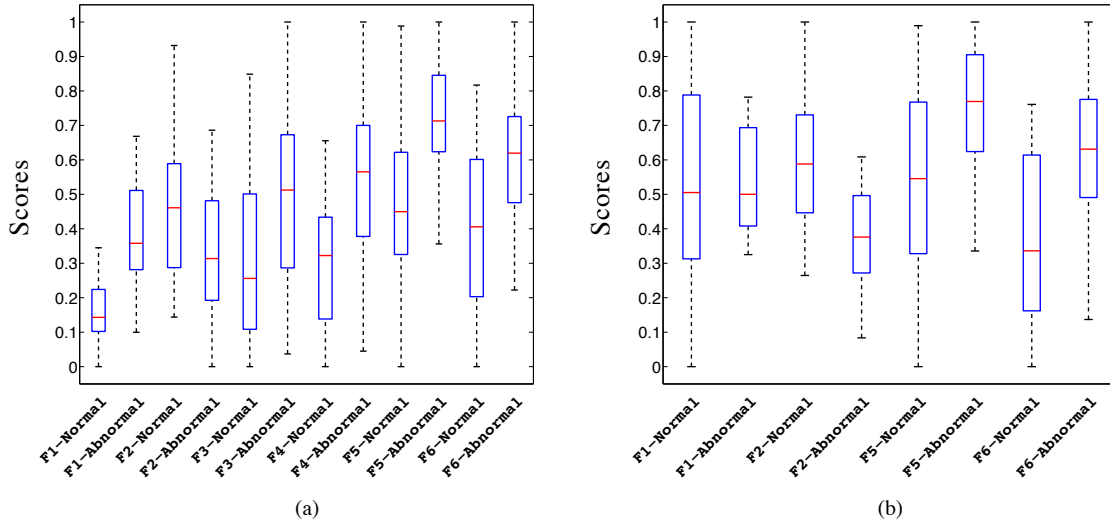

Figure 4: (a) Boxplots of the features $F1$, $F2$, $F3$, $F4$, $F5$, $F6$ for experiment 1. (b) Boxplots of the features $F1$, $F2$, $F5$, $F6$ for experiment 2.

Experiment 2: The classification performance results are presented in Table 3. Boxplots for each feature for both the normal and abnormal slides are shown in Figure 4b. The feature "Mean of $M2$", which is a measure of chromatin margination inside the nucleus, achieved the best classification performance in both classifiers for the detection of MACs. This is also supported visually in Figure 4b. The feature "SD of IOD", whilst quite discriminatory in the first experiment, does not show any discriminatory power in the second experiment. This is not surprising given that the Pap stain is not stoichiometric and that none of the cells in experiment 2 were diagnostic.

Table 2: Classification results for experiment 1.

FS Method	Top 4 features	AUC (Mean±SE)
MSVM-RFE	$F1$, $F2$, $F3$, $F4$	0.934±0.038
L1-Reg	$F1$, $F2$, $F5$, $F6$	0.954±0.019
GRRF	$F1$, $F2$, $F5$, $F6$	0.935±0.029
$F1$: SD of IOD, $F2$: Mean of $M2$, $F3$: SD of $C1$, $F4$: SD of nucleus area, $F5$: SD of $MST3$, $F6$: Mean of $W\text{-}IMC1$		

Table 3: Classification results for experiment 2.

Classifier	Features	AUC (Mean±SE)
SVM	$F2$	0.815±0.015
	$F5$	0.706±0.053
	$F6$	0.792±0.026
	$F2$, $F6$	0.822±0.024
Logistic Regression	$F2$	0.803±0.026
	$F5$	0.591±0.054
	$F6$	0.784±0.027
	$F2$, $F6$	0.752±0.038

4. CONCLUSIONS AND DISCUSSION

A set of novel structural texture features for quantifying nuclear chromatin patterns in cells on a conventional Pap smear was presented. The experimental results demonstrate the efficacy of this structural approach and that a combination of these structural texture features and conventional features can be used to discriminate between normal (NILM) and abnormal (LSIL and HSIL) slides with high accuracy. They also demonstrate that it is possible to detect malignancy associated changes (MACs) in Papanicoloau stain. The most discriminating single feature, Mean of $M2$, for detecting malignancy associated changes (MACs) is a structural feature measuring chromatin margination. This in turn suggests the possibility of developing a fully automated Pap smear screener based on MACs.

Most of the proposed contextual features for describing chromatin arrangement can be applied at the nucleus level to quantitatively characterize the spatial arrangement of cells on a slide. This will be the subject of future work.

REFERENCES

[1] DeMay, R. M., "Common problems in Papanicolaou smear interpretation," *Archives of Pathology Laboratory Medicine* **121**(3), 229–38 (1997).

[2] Mehnert, A., "Image analysis for the study of chromatin distribution in cell nuclei with application to cervical cancer screening," PhD Thesis, School of Information Technology and Electrical Engineering, The University of Queensland (2004).

[3] Guillaud, M., Cox, D., Adler-Storthz, K., Malpica, A., Staerkel, G., Matisic, J., van Niekerk, D., Poulin, N., Follen, M., and MacAulay, C., "Exploratory analysis of quantitative histopathology of cervical intraepithelial neoplasia: Objectivity, reproducibility, malignancy-associated changes, and human papillomavirus," *Cytometry Part A* **60A**(1), 81–89 (2004).

[4] Mommers, E., Poulin, N., Meijer, C. M., Baak, J. A., and van Diest, P. J., "Malignancy-associated changes in breast tissue detected by image cytometry," *Analytical Cellular Pathology* **20**, 187–195 (Jan 2000).

[5] Sacile, R., Montaldo, E., Ruggiero, C., Nieburgs, H., and Nicolo, G., "A decision support system to detect morphologic changes of chromatin arrangement in normal-appearing cells," *IEEE Transactions on NanoBioscience* **2**, 118–123 (June 2003).

[6] Rodenacker, K. and Bengtsson, E., "A feature set for cytometry on digitized microscopic images," *Analytical Cellular Pathology* **25**(1), 1–36 (2003).

[7] Niwas, S., Palanisamy, P., Sujathan, K., and Bengtsson, E., "Analysis of nuclei textures of fine needle aspirated cytology images for breast cancer diagnosis using complex Daubechies wavelets," *Signal Processing* (2012).

[8] Peura, M. and Iivarinen, J., [*Efficiency of Simple Shape Descriptors*], World Scientific (1997).

[9] Doudkine, A., MacAulay, C., Poulia, N., and Palcic, B., "Nuclear texture measurement in image cytometry," *Pathologica* **87**, 286–299 (1995).

[10] Young, I. T., Verbeek, P. W., and Mayall, B. H., "Characterization of chromatin distribution in cell nuclei," *Cytometry* **7**(5), 467–74 (1986).

[11] Sameti, M., Ward, R., Morgan, J., and Palcic, B., "Image feature extraction in the last screening mammograms prior to detection of breast cancer," *IEEE J Selected Topics Signal Processing* **3**(1), 46–52 (2009).

[12] Okabe, A., Boots, B., Sugihara, K., Chiu, S. N., and Kendall, D. G., [*Definitions and Basic Properties of Voronoi Diagrams*], John Wiley and Sons, Inc. (2008).

[13] Toussaint, G. T., "The relative neighbourhood graph of a finite planar set," *Pattern Recognition* **12**(4), 261–268 (1980).

[14] Jaromczyk, J. and Toussaint, G., "Relative neighborhood graphs and their relatives," *Proceedings of the IEEE* **80**, 1502–1517 (Sep 1992).

[15] Diane, S., Ritu, N., Kurman, R., Davey, D., and Wilbur, D., [*The Bethesda System for Reporting Cervical Cytology: Definitions, Criteria, and Explanatory Notes*], 2nd edition NewYork, Springer (2004).

[16] Moshavegh, R., Ehteshami Bejnordi, B., Mehnert, A., Sujathan, K., Malm, P., and Bengtsson, E., "Automated segmentation of free-lying cell nuclei in Pap smears for malignancy-associated change analysis," *Proceedings of the 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC 2012)* , 5372–5375 (Sep 2012).

[17] Kai, B. D., Rajapakse, J., Haiying, W., and Azuaje, F., "Multiple SVM-RFE for gene selection in cancer classification with expression data," *IEEE Transactions on NanoBioscience* **4**, 228–234 (Sep 2005).

[18] Park, M. Y. and Hastie, T., "L1-regularization path algorithm for generalized linear models," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **69**(4), 659–677 (2007).

[19] Deng, H. and Runger, G., "Gene selection with guided regularized random forest," *CoRR* , arxiv, 28 Sep 2012, http://arxiv.org/abs/1209.6425 (30 Sep 2012).

[20] Deng., H. and Runger, G., "Feature selection via regularized trees," *CoRR* , arxiv, 21 March 2012, http://arxiv.org/abs/1201.1587 (12 April 2012).

[21] Markowetz, F. and Spang, R., "Molecular diagnosis. classification, model selection and performance evaluation," *Methods of Information in Medicine* **44**(3), 438–443 (2005).