

Cervical Cell Classification Using Features Related to Morphometry and Texture of Nuclei

Juan Valentín Lorenzo-Ginori¹, Wendelin Curbelo-Jardines¹,
José Daniel López-Cabrera¹, and Sergio B. Huergo-Suárez¹

¹Center for Studies on Electronics and Information Technologies,
Universidad Central “Marta Abreu” de Las Villas (UCLV), Cuba
juanl@uclv.edu.cu

Abstract. The Papanicolaou test is used for early prediction of cervical cancer. Computer vision techniques for automating the microscopy analysis of cervical cells in this test have received great attention. Cell segmentation is needed here in order to obtain appropriate features for classification of abnormal cells. However, accurate segmentation of the cell cytoplasm is difficult, due to cell overlapping and variability of color and intensity. This has determined a growing interest in classifying cells using only features from the nuclei, which are easier to segment. In this work, we classified cells in the pap-smear test using a combination of morphometric and Haralick texture features, obtained from the nucleus gray-level co-occurrence matrix. A comparison was made among various classifiers using these features and data dimensionality reduction through PCA. The results obtained showed that this combination can be a promising alternative in order to automate the analysis of cervical cells.

Keywords: Papanicolaou test, features, texture, dimensionality reduction, classifiers.

1 Introduction

Cervical cancer is, after breast cancer, the most common form of this disease among the female population. Early detection of this has contributed to a considerable reduction of the associated mortality rate. The Papanicolaou test [1] is the standard procedure currently used for early prediction of cervical cancer. In this test, microscopy analysis of the so-called pap-smear, a sample of cervical cells appropriately stained, is analyzed in a microscope to detect abnormal cells, which can be considered precursors of the disease. However, there are some drawbacks associated to microscopy analysis of the pap-smear by human experts: some rate of false negatives appears due to subjectivity, routine and tiredness of the analysts. This has determined a growing interest in developing automated analysis procedures using computer vision techniques. A typical image from the standard Papanicolaou test is shown in Figure 1.

The analysis of pap-smears through computer vision requires prior image segmentation, in order to extract appropriate features to classify the cells. A set of 20 morphometric features, half of them related to cells' cytoplasm have been used to classify

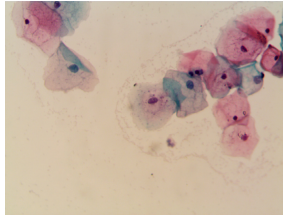


Fig. 1. A typical pap-smear image. The nuclei appear as dark spots within the clearer, coloured cytoplasm.

the cervical cells in seven well defined classes, from which a benchmark database has been built [2]. Three out of the seven cervical cells classes are considered as normal and four of them are abnormal, thus raising two problems: classification in seven classes, and binary classification in normal-abnormal. This work addresses the latter.

A formal attempt to classify cervical cells using different classifiers and the previously mentioned features from nuclei and cytoplasm, with feature selection using genetic algorithms, is shown in [3]. Other classification approaches can be found in [4] and [5], using features that are mostly associated to cells' morphometry, and the latter using also four textural features, in this case applied to a technique called liquid-based cervical smears, that differ from the standard pap test to which the present work is devoted. In regard to segmentation, some works address the problem of finding an appropriate way to segment the cytoplasm [6], however, it has been found that the cell cytoplasm is very difficult to segment with good accuracy, due to overlapping of cells and their variability of color and intensity. On the other hand, the cell nuclei usually appear better defined in the images, and this allows improving segmentation accuracy. Therefore, some works have been devoted to nuclei segmentation in pap-smears [7], [8]. This situation has led to conduct research on the possibility of classifying cells using only features obtained from their nuclei as in [9], where nine nuclei morphometric features were used.

In this work, we explore the possibility of improving classification of cells in the pap-smear using information from the nuclei only, but additionally including texture features. This was motivated by the fact that staining in pap smears makes visible the chromatin textural patterns in the nuclei and this information is used by cytopathologists to classify the cells [10].

Here a method using morphological image processing [11] was employed to calculate, from a given nucleus image, its gray-level co-occurrence matrix (GLCM) and the associated Haralick features associated to texture [12]. These features were used, together with the previously known morphometric ones, to construct a combined feature matrix. Dimensionality reduction using principal components analysis (PCA) was also employed, given the relatively large amount of features obtained. Four classifiers were used: linear, Mahalanobis distance, K-nearest-neighbors (KNN) and support vector machines (SVM). Comparison among classifiers' performance was made through statistical methods [13] and results showed that using this combination of features, the binary classification results can be improved.

This article is organized as follows: in section 2, the main characteristics of the cells used in the experiments are described, as well as the methods used to calculate the texture features and the experiments to test the classifiers' performance. We summarize and discuss the main results in section 3, and conclusions are exposed in section 4.

2 Materials and Methods

The cervical microscopy cell images used in this work were obtained from the Herlev database [2], which contains 917 annotated images, each with a manually segmented version that can be used as ground-truth. The different classes to which they belong are shown in Table 1.

Table 1. Cervical cells in the Herlev image database

Class	Category	Cell type	Number of cells	Sub-totals
1	Normal	Superficial squamous epithelial	74	
2	Normal	Intermediate squamous epithelial	70	
3	Normal	Columnar squamous epithelial	98	242
4	Abnormal	Mild squamous non-keratinizing dysplasia	182	
5	Abnormal	Moderate squamous non-keratinizing dysplasia	146	
6	Abnormal	Severe squamous non-keratinizing dysplasia	197	
7	Abnormal	Squamous cell carcinoma in situ intermediate	150	675

2.1 Texture Analysis

The set of features employed for classification is shown in Table 2, which included both morphometric and Haralick texture features. In order to obtain the latter, the GLCM from the cell nuclei was calculated. Once the nuclei had been segmented (an operation that is out of the scope of this paper), an image represented by a square matrix containing the texture pattern of the nucleus region was obtained. The problem associated to acquiring this matrix was formulated in terms of obtaining the largest x - y oriented square that can be inscribed in the nucleus region. This operation was performed through morphological image processing in the following way:

1. Obtain a binary mask which represents the location of the segmented nucleus.
2. Successively erode the nucleus mask with a square structuring element (SE) increasing its side length one pixel per iteration, until the nucleus mask disappears.
3. Go back one step and obtain the square SE that, when its side is enlarged just one pixel, completely erodes the nucleus mask. Using the erosion of the nucleus mask with this SE, pick one pixel from the resulting binary image.
4. With the same SE used in step (3), perform a dilation of this pixel. The result will be a largest inscribed square in the selected nucleus.
5. Use this binary square as a mask and perform an array multiplication with the nucleus image to obtain a square matrix in grayscale containing the nucleus texture.
6. Expand linearly the intensity of this image to the maximum interval allowed, in order to enhance its contrast.

This process is illustrated in Figure 2. The co-occurrence matrix of this image [12] was calculated for three pixel offset values (1, 4, 7) and four spatial orientations (0, $\pi/4$, $\pi/2$, $3\pi/4$), forming a three dimensional array in which the spatial orientation corresponds to dimension 3. A final co-occurrence matrix was obtained by selecting the maximum along dimension 3, as we are interested in the most significant co-occurrence values independently of their associated spatial orientation.

Table 2. Morphometric and texture features

Nuclei features/morphometric	
Mean intensity	Area
Maximum intensity	Perimeter
Minimum intensity	$(Area)^{1/2}/Perimeter$
Solidity	Entropy (grayscale image)
Eccentricity	
Nuclei features/Haralick coefficients	
Autocorrelation	Sum of squares: Variance
Contrast	Sum average
Correlation	Sum variance
Cluster prominence	Sum entropy
Cluster shade	Difference variance
Dissimilarity	Difference entropy
Energy	Information measure of correlation 1
Entropy	Information measure of correlation 2
Homogeneity 1	Inverse difference (INV)
Homogeneity 2	Inverse difference normalized (INN)
Maximum probability	Inverse difference moment norm.

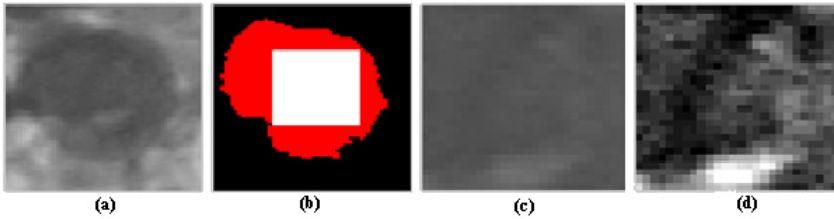


Fig. 2. A sample is obtained from a cell's nucleus. (a) Grayscale image from the Herlev database, (b) the corresponding nucleus mask (red) and the inscribed square (white), (c) the grayscale sample obtained using the mask and (d) the sample after enhancing its contrast.

After this, the set of Haralick features was calculated, finally forming a feature matrix of size $N_c \times (22+9)$, where N_c is the amount of cells contained in each class in Table 1.

2.2 Cell Classification

The cell classification process had two purposes: determining if using texture features meant a statistically significant improvement in classification accuracy and comparing the performance of various classifiers. Classes 3 and 4 from the Herlev database (see Table 1) were used to perform an abridged evaluation. The indexes of classifiers' effectiveness used were: sensitivity (se), specificity (sp), positive and negative predictive values (pp and np) and the F-measure (harmonic mean of se and sp) with emphasis in the last two. The classifiers evaluated were: linear, Mahalanobis distance, k nearest neighbors (KNN) and (after testing some SVM options), a Gaussian radial basis function kernel SVM with $\sigma=2$. In all cases, dimensionality reduction (DR) by principal components analysis (PCA) was employed. Several values of DR were employed and the best among them was used when comparing the classifiers.

An m -fold cross-validation ($m=20$) was performed in which the indexes of effectiveness were calculated. The series of indexes values were used for determining, using the Friedman test [13], if there was a statistically significant difference in two situations: (1) among the alternatives of features used: morphometric only, texture only or both, with various alternatives of dimensionality reduction, and (2) among the various classifier algorithms employed.

3 Results and Discussion

3.1 Classification Performance for Different Alternatives of Features

After performing several tests, we determined the most favorable values for PCA data reduction as well as the classifier with best performance. Table 3 shows the values of the indexes of effectiveness for the Mahalanobis classifier, with morphometric features only, without PCA data reduction and with DR to seven features, as well as using texture only and all the features, the latter two with 17 features after DR. Notice that the highest pn and F-measure values were obtained using all features with DR to

17 features. The Friedman test was realized using results from a 20-fold cross-validation, with the four data reduction alternatives as related samples. The result for the Mahalanobis classifier is shown in Table 4, in which the higher rank was obtained for the all-features case, DR to 17, for the F-measure ($p < 0.05$). Further pair-wise analysis using the Wilcoxon signed rank test tended to confirm the superiority of this alternative, although this result is somewhat limited due to correlation in the training data. Results were similar for the np index, and for the rest of the classifiers, the best performance in most cases appeared when using all features with DR to 17, regarding F-measure and np. Image resolution and offset values could also affect these results.

3.2 Determining the Best Classifier

After performing the previous experiments, we made a comparative evaluation of the various classifiers, again with an m-fold cross-validation with $m=20$, now using the best alternative, i. e., both morphometric and texture features with PCA and 17 features after DR. The Friedman test was used again, for which we made the classification using for all the classifiers the same grouping of cells in training and test sets in the cross-validation. The corresponding results are shown in Table 5. The ranks obtained suggest that the SVM classifier was the best. However, further pair-wise Wilcoxon test showed that the statistically significant differences are among SVM-Mahalanobis and KNN-linear, the former pair being better in comparison to the second, with no statistically significant difference between the paired methods. However, there is an important difference in terms of computer time, as shown in Table 6. A binary classification for the whole Herlev database was made, and its results, although inferior, were consistent with those from the abridged experiment.

Table 3. Results for the Mahalanobis classifier

	Morphometric features		Texture features	Combined features
	without DR	DR to 7	DR to 17	DR to 17
np	0,895	0,823	0,677	0,910
F_measure	0,916	0,917	0,854	0,971

Table 4. Results of the Friedman test for the for the Mahalanobis classifier, $p < 0.05$

Alternative	Mean rank, F	Mean rank, np
Morphometric features without DR	2, 80	3,10
Morphometric, 7 features after DR	2,23	2,03
Haralick, 17 features after DR	1,00	1,00
Combining all features, DR to 17	3,98	3,88

Table 5. Results of the Friedman test for the F-measure obtained with the different classifiers using the combined features with DR, $p < 0.05$

Classifier	Mean rank, F	Mean rank, np
Linear, combined, DR	2,03	2,33
Mahalanobis, combined, DR	4,00	3,25
KNN, morphometric, no DR	1,00	1,00
SVM, combined, DR	2,98	3,43

Table 6. Computer time for the various classifiers

Classifier	Time, s
Linear	0,152
Mahalanobis	0,151
SVM	14,619
KNN	0,155

4 Conclusions

In this work, binary classification of cells in the Papanicolaou test was performed using features from the cells' nuclei only. Morphometric features of the nuclei were calculated firstly. Then a square sub-image from each cell nucleus was extracted using morphological image processing, and its GLCM and the associated Haralick features were calculated. These were combined with the morphometric data to build a feature matrix, whose dimensionality was reduced through PCA. An abridged m-fold cross-validation experiment using classes 3 and 4 described in Table 1 was made. Results showed a more accurate classification in terms of the negative predictive value and the F-measure in comparison to using morphometric data only. From the classifiers tested: linear, KNN, Mahalanobis and SVM, the latter two showed better results. Classification results using the whole database, although inferior compared to the abridged experiment, tended to confirm the advantages of using also nuclei texture features. Evaluation of the classifiers was made using statistical hypothesis testing.

The results obtained showed advantages from using also texture features when classifying cells in the Papanicolaou test using data from the nuclei only. This suggests a number of alternatives to be evaluated in future work, for example: exhaustive search for the best offsets in the GLCM, new methods to extract texture features like morphological granulometry, using kernel PCA or feature selection methods for dimensionality reduction, and trying other classifying algorithms.

Acknowledgments. The authors acknowledge M.D. Irene Rodríguez-Santos for her collaboration and Ph.D. Julián L. Cárdenas-Barreras for reviewing the manuscript.

This work received support from the Canadian International Development Agency (CIDA) Project “Biomedical Engineering Education in Cuba”, from the Institutional University Co-operation (IUC) program of the Flemish University council (VLIR) and from the Center for Studies in Electronics and Information Technologies at UCLV.

References

1. Papanicolaou, G.N.: A new procedure for staining vaginal smears. *Science* 95, 438–439 (1942)
2. Jantzen, J., Norup, J., Dounias, G., Bjerregaard, B.: Pap-smear Benchmark Data For Pattern Classification. Presented at the Nature Inspired Smart Information Systems, NiSIS (2005)
3. Marinakis, Y., Dounias, G., Jantzen, J.: Pap smear diagnosis using a hybrid intelligent scheme focusing on genetic algorithm based feature selection and nearest neighbor classification. *Comput. Biol. Med.* 39, 69–78 (2009)
4. Mat-Isa, N.A., Mashor, M.Y., Othman, N.H.: An automated cervical pre-cancerous diagnostic system. *Artif. Intell. Med.* 42, 1–11 (2008)
5. Huang, P.-C., Chan, Y.-K., Chan, P.-C., Chen, Y.-F., Chen, R.-C., Huang, Y.-R.: Quantitative Assessment of Pap Smear Cells by PC-Based Cytopathologic Image Analysis System and Support Vector Machine. In: Zhang, D. (ed.) ICMB 2008. LNCS, vol. 4901, pp. 192–199. Springer, Heidelberg (2007)
6. Kale, A., Aksoy, S.: Segmentation of Cervical Cell Images. Presented at the IEEE 2010 International Conference on Pattern Recognition, August 23 (2010)
7. Sobrevilla, P., Montseny, E., Vaschetto, F., Lerma, E.: Fuzzy-based analysis of microscopic color cervical pap smear images: nuclei detection. *Int. J. Comput. Intell. Appl.* 9, 187–206 (2010)
8. Plissiti, M.E., Nikou, C., Charchanti, A.: Automated Detection of Cell Nuclei in Pap Smear Images Using Morphological Reconstruction and Clustering. *IEEE Trans. Inf. Technol. Biomed.* 15, 233–241
9. Plissiti, M.E., Nikou, C.: Cervical Cell Classification Based Exclusively on Nucleus Features. In: Campilho, A., Kamel, M. (eds.) ICIAR 2012, Part II. LNCS, vol. 7325, pp. 483–490. Springer, Heidelberg (2012)
10. Patten, S.F.: *Diagnostic cytopathology of the uterine cervix*. S. Karger AG, Basel (1978)
11. Soille, P.: *Morphological Image Analysis: Principles and Applications*. Springer (2004)
12. Haralick, R.M., Shanmugam, K., Dinstein, I.: Textural features for image classification. *IEEE Trans. Syst. Man Cybern.* 3, 610–621 (1973)
13. Demsar, J.: Statistical Comparisons of Classifiers over Multiple Data Sets. *J. Mach. Learn. Res.* 7, 1–30 (2006)