

Automated image analysis of digital colposcopy for the detection of cervical neoplasia

Sun Young Park

University of Texas
Department of Biomedical Engineering
Austin, Texas 78712

Michele Follen

University of Texas
M. D. Anderson Cancer Center
Department of Biostatistics and Applied Mathematics
and Department of Gynecologic Oncology
Center for Biomedical Engineering
Houston, Texas 77030

Andrea Milbourne**Helen Rhodes**

University of Texas Health Science Center
Department of Gynecology, Obstetrics and
Reproductive Sciences
Houston, Texas 77030

Anais Malpica

University of Texas
M. D. Anderson Cancer Center
Department of Pathology
Houston, Texas 77030

Nick MacKinnon**Calum MacAulay**

British Columbia Cancer Center
Department of Optical Imaging and Gynecologic
Oncology
Vancouver, British Columbia V5Z 1L3
Canada

Mia K. Markey

University of Texas
Department of Biomedical Engineering
Austin, Texas 78712

Rebecca Richards-Kortum

Rice University
Department of Bioengineering
Houston, Texas 77005

1 Introduction

Cervical cancer is the second most common cancer in women worldwide, and 80% of cases occur in the developing world where the fewest resources exist for management.¹ Most cases of cervical cancer can be prevented through screening programs aimed at detecting precancerous lesions. Screening for

Abstract. Digital colposcopy is a promising technology for the detection of cervical intraepithelial neoplasia. Automated analysis of colposcopic images could provide an inexpensive alternative to existing screening tools. Our goal is to develop a diagnostic tool that can automatically identify neoplastic tissue from digital images. A multispectral digital colposcope (MDC) is used to acquire reflectance images of the cervix with white light before and after acetic-acid application in 29 patients. A diagnostic image analysis tool is developed to identify neoplasia in the digital images. The digital image analysis is performed in two steps. First, similar optical patterns are clustered together. Second, classification algorithms are used to determine the probability that these regions contain neoplastic tissue. The classification results of each patient's images are assessed relative to the gold standard of histopathology. Acetic acid induces changes in the intensity of reflected light as well as the ratio of green to red reflected light. These changes are used to differentiate high-grade squamous intraepithelial (HGSIL) and cancerous lesions from normal or low-grade squamous intraepithelial (LGSIL) tissue. We report diagnostic performance with a sensitivity of 79% and a specificity of 88%. We show that diagnostically useful digital images of the cervix can be obtained using a simple and inexpensive device, and that automated image analysis algorithms show a potential to identify histologically neoplastic tissue areas. © 2008 Society of Photo-Optical Instrumentation Engineers. [DOI: 10.1117/1.2830654]

Keywords: digital colposcopy; image analysis; early detection; cervical cancer.

Paper 07071RR received Feb. 28, 2007; revised manuscript received Oct. 21, 2007; accepted for publication Oct. 23, 2007; published online Jan. 16, 2008.

cervical neoplasia using the Papanicolaou (Pap) smear, followed by colposcopy, biopsy, and treatment of neoplastic lesions has dramatically reduced the incidence and mortality of cervical cancer in every country in which organized programs have been established.² This is due to the fact that the precursors, denoted as cervical intraepithelial neoplasia (CIN) or squamous intraepithelial lesions (SIL) can take 3 to 20 yr to develop into cancer. However, due to the lack of resources

Address all correspondence to Rebecca Richards-Kortum, PhD, Rice University, Bioengineering Department, 6100 Main, Houston, TX 77005; Tel: 713-348-5869; Fax: 713-348-5877; E-mail: rkortum@rice.edu

Table 1 Results of pooled analyses from Sankaranarayanan et al.¹¹

Screening Methodology	Number of Studies Reviewed	Number of Patients	Percent of Patients with HG and Greater	Number of Patients within HG Range	Sensitivity Pooled	Sensitivity Range	Specificity Pooled	Specificity Range
VIA	11	54,981	16	7–27	0.77	0.56–0.94	0.86	0.74–0.94
VIAM	3	16,900	14	11–18	0.64	0.61–0.71	0.87	0.83–0.90
VILI	10	49,080	16	9–29	0.92	0.76–0.97	0.85	0.73–0.91
Cytology	5	22,663	6	2–14	0.58	0.29–0.77	0.95	0.89–0.99
HPV HC II	4	18,085	7	6–9	0.67	0.46–0.81	0.94	0.92–0.95

VIA is visual inspection with acetic acid, VIAM is visual inspection with acetic acid and magnification, VILI is visual inspection with Lugol's iodine, cytology means conventional cytology, HG is high grade and HPV HC II is human papillomavirus hybrid capture II test (Digene, Gaithersburg, Maryland).

and infrastructure, 238,000 women die each year of cervical cancer; more than 80% of these deaths occur in developing countries.^{1,3} We are interested in applying optical technologies to replace expensive infrastructure for cervical cancer screening in the developing world.

The use of direct visual inspection (DVI), visual inspection with acetic acid (VIA), and visual inspection with Lugol's iodine are being explored as alternatives to Pap smear and colposcopic examination in many developing countries.^{4–7} Recent reviews of the performance of these methods found that they have sufficient sensitivity and specificity, when performed by trained professionals, to serve as viable alternatives to Pap screening in low-resource settings.^{8–10} Table 1 shows the results of a review by Sankaranarayanan et al.,¹¹ which favorably compares these methods to human papillomavirus (HPV) testing and conventional cytology.

Because DVI relies on visual interpretation, it is crucial to define objective criteria for the positive identification of a lesion and to train personnel to correctly implement a program based on these criteria. Denny et al. noted that restricting the definition of a positive VIA test to a well-defined acetowhite lesion significantly improved specificity, while reducing sensitivity.¹⁰ In a series of 1921 women screened in Peru, Jeronimo et al. found that the DVI positivity rate dropped from 13.5% in the first months to 4% during subsequent months of a 2-yr study; the drop in positivity rate was hypothesized to be due to a learning curve for the evaluator.¹² Bomfim-Hyppolito et al. investigated the use of cervicography as an adjunct to DVI. A simple Sony digital camera was used to photograph the cervix before and after the application of acetic acid.¹³ Photographs were later interpreted by an expert colposcopist. The addition of cervicography improved both sensitivity and specificity. However, this approach prevents the implementation of see-and-treat strategies in low-resource settings because of the need for expert review.

Recently, optical techniques have been investigated as an alternative detection method in a quantitative and objective manner. Several studies have demonstrated that optical spectroscopy has the potential to improve the screening and diagnosis of neoplasia. Ferris et al. studied multimodal hyperspectral imaging for the noninvasive diagnosis of cervical neoplasia.¹⁴ They reported a sensitivity of 97% and a specificity of 70%. Huh et al.¹⁵ measured the performance of opti-

cal detection of HGSIL using fluorescence and reflectance spectroscopy, finding a sensitivity of 90% and a specificity of 70%.

As another promising application of optical techniques for cervical cancer screening, a number of studies investigated whether digital image processing techniques could be used to automate the interpretation of colposcopic images.^{16–20} Craine and Craine¹⁷ introduced a digital colposcopy system for archiving images and visually assessing features in the images. Shafi et al.¹⁸ and Cristoforoni et al.¹⁹ used a digital imaging system for colposcopy, which enables image capture and simple processing. To assess the various colposcopic features, the acquired images were manually analyzed by an expert. By examining the relationship between colposcopic features and histology outcomes, Shafi et al.'s study provided information about which features are most useful to the expert observer. Image interpretation in these early studies mainly relied on experts' qualitative assessment of colposcopic images and provided limited quantitative analysis. Li et al. developed a computer-aided diagnostic system using colposcopic features such as acetowhitening changes, lesion margin, and blood vessel structures.²⁰ They prototyped image processing algorithms for detection of those features and showed promising preliminary results. However, the diagnostic performance of the system has not been reported.

Recently, advances in consumer electronics have led to inexpensive, high-dynamic-range charge-coupled device (CCD) cameras with excellent low light sensitivity. At the same time, advances in vision chip technology enable high-quality image processing in real time. These advances may enable the acquisition of diagnostically useful digital images of the cervix in a relatively inexpensive way, with or without magnification. Moreover, automated analysis algorithms based on modern image processing techniques have the potential to replace clinical expertise, which may reduce the cost of screening. The purpose of this study is to explore whether digital colposcopy, combined with recent advances in camera technology and automated image processing, could provide an inexpensive alternative to Pap screening and conventional colposcopy.

Several challenges for diagnostic digital colposcopic image analysis remain. First, previous studies have investigated only a few features and have not taken advantage of the ma-

ture field of image analysis and computer-automated techniques. Second, previous studies have compared image features of selected normal and abnormal areas of the cervix, but have not applied the approach to an entire image to identify whether lesions are present. Finally, previous studies have used biopsies from selected areas as the gold standard. A gold standard is necessary for the entire field of view to address the issue of lesion localization.

In this pilot study, we explored the potential of an approach to automated image analysis of white-light colposcopy images that addresses the three challenges just discussed. First, our approach employed automated image analysis techniques including image registration, pattern recognition, clustering, and classification. Second, our algorithm can identify high-grade precancerous tissue areas from an entire image. Third, a gold standard for the entire cervical image was constructed using a whole cervix specimen acquired from a loop electrosurgical excision procedure (LEEP), which was intensively sectioned.

2 Materials and Methods

2.1 Instrumentation

A multispectral digital colposcope (MDC) was developed to acquire reflectance images of the entire cervix with white-light illumination. The MDC consists of a commercially available, tilt-stand colposcope (Model 1DL, Leisegang, Germany) with a video-rate color CCD camera and frame grabber for image acquisition. Details of this instrument can be found in papers by Benavides et al.²¹ and Park et al.²² The colposcope produces stereoscopic vision at $7.5\times$ magnification. Reflectance images were captured using an inexpensive, commercially available, video rate, color CCD camera (CV S3200 Rev. B, JAI, Japan).

2.2 Clinical Study

The study protocol was reviewed and approved by the institutional review boards at the University of Texas M.D. Anderson Cancer Center, the University of Texas Health Science Center in Houston, the Lyndon B. Johnson County Hospital and the Harris County Health District, the British Columbia Cancer Agency, Rice University, and the University of Texas at Austin. Eligible patients were at least 18 yr old, not pregnant, and were referred to the colposcopy clinic with an abnormal Pap smear. Written informed consent was obtained from each participant, and all patients underwent a history, complete physical exam, Pap smear, cultures for gonorrhea and chlamydia, pancolposcopy (including cervix, vagina, vulva, and perianal area), and colposcopically directed biopsies. Only patients with HG lesions who were scheduled at a previous visit were eligible for the study. Following colposcopic examination, but prior to the treatment with an LEEP, white-light reflectance images were acquired with the MDC from each patient at baseline. A second set of reflectance images was acquired following the application of acetic acid (6%) on the cervix for 2 min. Acetic acid enhances the differences in appearance between normal and dysplastic tissue. The patients then underwent an LEEP procedure. After the LEEP, for histopathology map construction, the margin boundary of LEEP was carefully marked on one of the acquired MDC images. The specimen was inked and cut by the

study pathologist. Each ectocervix was cut into 12 pieces. At each trial site, only the expert study pathologists reviewed all the hematoxylin and eosin (H&E) stained slides jointly, noting areas of CIN, squamous epithelium, columnar epithelium, and transformation zone. Standard diagnostic criteria were used as the gold standard by the study pathologists to classify histopathologic samples as HG SIL, low-grade SIL, or negative for dysplasia.²³ The histopathologic slides were scanned, preserving real tissue size, and then the images were uniformly enlarged by 15% to account for shrinkage by formalin,²⁴ under the assumption of uniform shrinkage. The histopathology images were then reconstructed into a 3-D "pathology map" delineating areas of intraepithelial neoplasia using 3-D image Visualization Software® (Able Software, Lexington, Massachusetts). When the map was constructed, the ratio between the actual size of each pixel of MDC-measured images and that of the scanned images was calculated to establish the appropriate mapping between the MDC image and the histopathology map. The position of the LEEP specimen boundary and the transformation zone were also incorporated to correlate between MDC images and the histopathology images. After construction, pathology maps were reviewed by the study clinicians and pathologists and compared to the white light reflectance MDC images.

2.3 Data Processing

Image preprocessing was performed to remove impulse noise spikes, specular reflection, and regions obscured by the presence of blood on the tissue surface. To properly register images obtained before and after acetic acid application, an automatic registration algorithm was developed. In the registration algorithm, the location of the cervical os was used as a reference point.

Our approach to automated digital image analysis for screening for cervical neoplasia can be divided into two major steps. First, tissue regions with similar optical properties are clustered together. Second, classification algorithms are used to determine whether these regions contain neoplastic tissue. These steps are now described in detail.

2.4 Classification of Localized Image Areas

To identify optical parameters that are most useful to determine whether localized image areas contain neoplastic tissue, the original red-green-blue (RGB) color intensity feature space was extended and features for image analysis were selected from this extended feature space. The extended feature space is composed of five categories of features: (1) raw RGB intensity values, (2) ratios of individual RGB intensity values, (3) differences in individual RGB intensity values, (4) differences in RGB intensity ratios, and (5) changes in gray-scale intensity values following application of acetic acid. For each of the pre- and post-acetic-acid images, the ratios of all possible pairs of features were calculated. Acetic-acid-induced differences in RGB intensity values, ratios, and gray-scale values were also calculated. Five features were selected as potentially useful for classification: intensity values for red, green, and blue channels; the ratio of intensity in the green channel to that in the red channel; and the changes in gray-scale intensity values.

Given these five features, we next developed classifiers to estimate the tissue type of a localized region. We designed an ensemble classifier, wherein a number of different classifiers are applied and then the final classification is based on the most highly predictive of the individual classifiers. The four classifiers in the ensemble are a linear classifier with Euclidean distance, a linear classifier with Mahalanobis distance, a K -nearest neighbor (KNN) classifier with eight neighbors, and a support vector machine with a linear kernel.

The classifiers were trained using leave-one-patient-out cross-validation. In this method, the algorithm is trained on all data excluding one patient and then applied to the held-out patient; then the process is repeated successively for all patients in the data set.²⁵

This is very useful in small studies, but may overtrain the algorithm. For each patient, image regions were identified corresponding to histopathologically proven areas of normal epithelium or low-grade squamous intraepithelial lesions (LGSILs) and HGSILs. Because an HGSIL is treated and an LGSIL is followed, HGSILs were considered to be abnormal, while areas of normal tissue and LGSIL were considered to be normal. Training data were obtained using a window-based approach wherein 20×20 -pixel windows were selected from areas of normal tissue or minor atypia, LGSIL, and HGSIL within an image. For each window, the five features were calculated for each pixel. Then the mean, the standard deviation, the 95th percentile, and entropy of each feature were calculated for the 20×20 -pixel window.

The 20×20 -pixel windows forming the training set were sampled in three different ways to assess the effect of training data on the resulting classifier performance. In the first approach (training data set 1) abnormal areas were selected from regions that were both pathologically HGSIL and in which the acetic acid intensity change was visually obvious. In the second approach (training data set 2), abnormal areas were selected from pathologically HGSIL regions in which the acetic acid intensity change was visually less apparent. In the third approach (training data set 3), abnormal data were selected based only the pathology map and without reference to the colposcopic images. The classifier performance was tested using the method of cross-validation and evaluated using receiver-operator characteristic curve (ROC) analysis. The classifier designed with training set data was then used to classify entire images as described below.

2.5 Image Analysis

The previous section described how features useful for classification were identified, and how classifiers were developed for local image regions. However, for truly automatic image interpretation, procedures must be developed to automatically segment the image into locally similar regions.

Local regions for analysis were segmented using a variant of the K -means clustering algorithm. After segmenting the image into local regions using the cluster algorithm, image classification was performed using the algorithm developed in the previous section. For each cluster, windows of pixels were randomly selected from the cluster and, based on the classification results, a probability of abnormality was calculated and assigned to the cluster. The abnormal regions were then defined as those clusters for which the probability of abnormal-

ity exceeds a preset threshold value; the threshold was chosen in such a way that the selected probability yields a Pareto optimal point from the ROC curve analysis. The ROC curve analysis is described in Sec. 3.

2.6 Performance Assessment

To assess the classification results of each image, a grid-based approach for calculating sensitivity and specificity was used. In this approach, an image is divided into grids and the classification results of the pixels in each grid of an image are compared with the corresponding pathology map. If the number of misclassified pixels is less than 20% of the number of pixels in the grid, then the grid point is considered to be correctly classified. Using this approach, the sensitivity Se_{image} and the specificity Sp_{image} of an image are defined as follows:

$$Se_{\text{image}} = \frac{\text{number of grids correctly identified as positive grids}}{\text{number of positive grids}},$$

$$Sp_{\text{image}} = \frac{\text{number of grids correctly identified as negative grids}}{\text{number of negative grids}}.$$

To assess the overall sensitivity and specificity of the automated image analysis procedure, the threshold probability resulting in the most desirable sensitivity and specificity values was determined in the Pareto sense.²⁶ This threshold probability was applied to all patients' image classification results. If an image contains an area of HGSIL whose radius is greater than 1 mm, then the patient is classified as a HG patient. Using this approach, the sensitivity Se_{MDC} and specificity Sp_{MDC} are defined as follows:

$$Se_{\text{MDC}} = \frac{\text{number of patients correctly identified as HGSIL}}{\text{number of HGSIL patients}},$$

$$Sp_{\text{MDC}} = \frac{\text{number of patients correctly identified as LGSIL}}{\text{number of normal or LGSIL patients}}.$$

3 Results

Pre- and post-acetic-acid white-light images were measured from 29 patients. Histopathology examination revealed 21 patients (73%) with CIN 2 or 3 lesions in the ectocervix, 2 patients (6%) with CIN 1 lesions, and 6 patients (21%) with no evidence of CIN. These eight patients each had a previous biopsy showing HGSIL.

Figure 1 illustrates the training data acquisition. As illustrated in the figure, abnormal data for the training set were sampled from pixels in HGSIL regions, and normal data for the training set were sampled from histopathologically normal or LGSIL regions.

Figure 2 illustrates the change in the ratio of the intensity in the green channel to that of the red channel induced by acetic acid for the training set data identified as histologically normal and abnormal. The application of acetic acid does not substantially change the green-to-red intensity ratio measured from normal tissue, as shown in both the histogram and the box plot. However, the green-to-red intensity ratio increases for abnormal samples following application of acetic acid, as

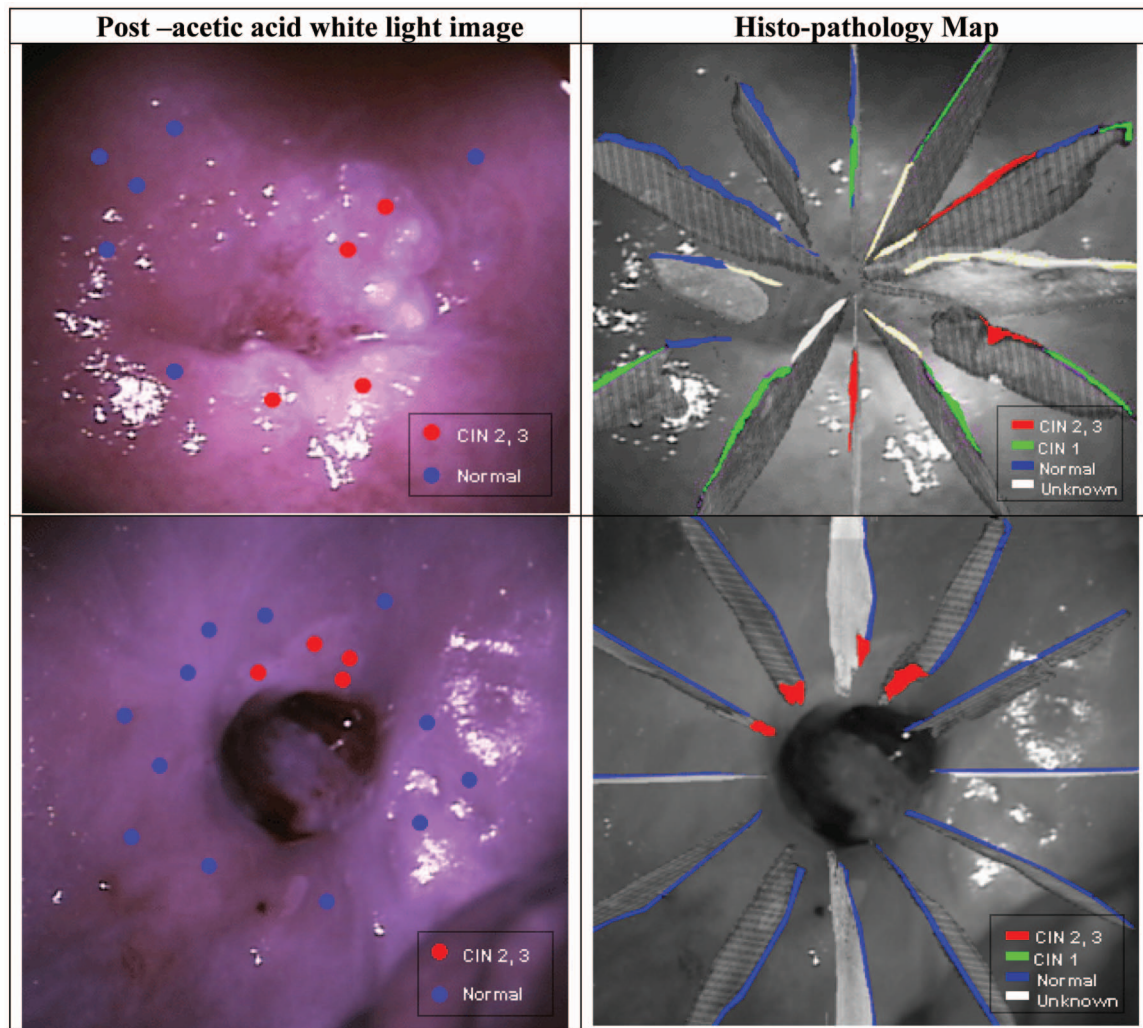


Fig. 1 Post-acetic-acid white-light images from two patients (left column) with superimposed histopathology map (right column). The dots in the left column indicate the regions randomly selected for use in the localized image area classification. The tissue type at these points was determined according to the histopathology map. The histopathology map includes an image of an H&E-stained histopathology section of the LEEP specimen obtained at each clock position of the LEEP specimen. The epithelium within these sections is color coded according to the histopathologic diagnosis, with normal epithelium indicated in blue, CIN 1 in green, and CIN 2 and CIN 3 in red. Areas of missing epithelium are colored in white.

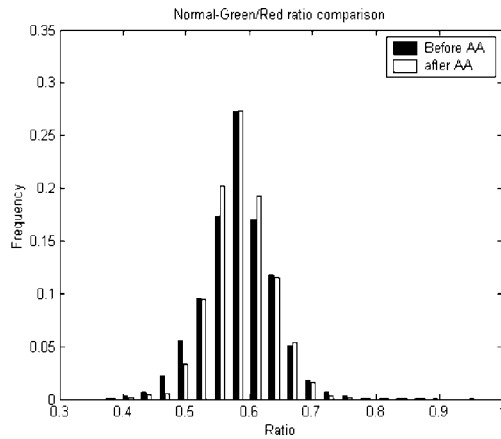
shown in both the histogram and the box plot. Following application of acetic acid, this parameter shows considerable separation between the normal and abnormal samples, and this intensity ratio was one of the features selected for use in the classification algorithm.

As shown the box-and-whisker plots in Fig. 2, for both normal and abnormal samples, the data dispersion is large and does not change much before and after acetic acid application. However, for normal samples, the median values before and after acetic acid application are similar, while abnormal samples show a considerable increase in the median value after application of acetic acid. Based on this observation, the 95th percentile values are used as a diagnostic feature for image classification.

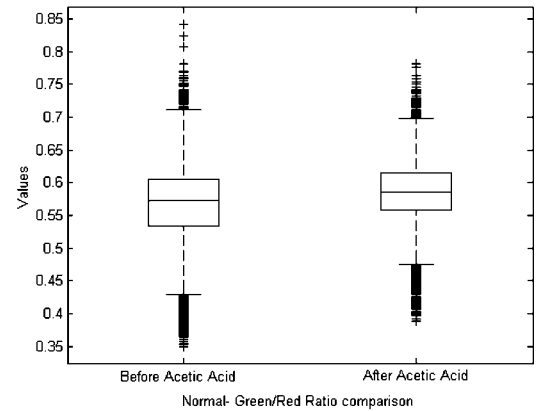
The performance of the proposed multiclassifier for localized image areas with three different training data sets was analyzed and the resulting ROC curves are shown in Fig. 3. As shown in Fig. 3, training data set 1 yielded the best results; these data were sampled from those regions where both the

pathology map indicated HG disease and the colposcopic images showed significant changes after addition of acetic acid. The performance of training data set 3, where image regions were chosen based only on the pathology map, is better than that of training data set 2, where abnormal areas were selected from pathologically abnormal regions in which the acetic acid intensity change was visually less apparent. For each ROC curve, the point with the most desirable sensitivity and specificity values was selected among Pareto-optimum points.²⁶ The selected point for training data set 1 corresponds to a sensitivity of 91% and a specificity of 80%. The optimum point for training data set 2 provided a sensitivity of 82% and a specificity of 69%, and that of training data set 3 corresponds to a sensitivity of 87% and a specificity of 83%.

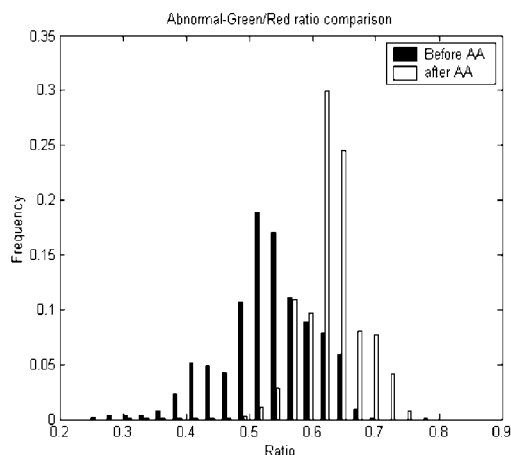
Figure 4 shows before and after acetic acid images for a patient with CIN 1, 2, and 3, together with the pathology map and the image classification result. The high-probability regions in the disease probability map correlate well with regions of CIN 2 and 3 in the histopathology map. Figure 5

Normal
Samples

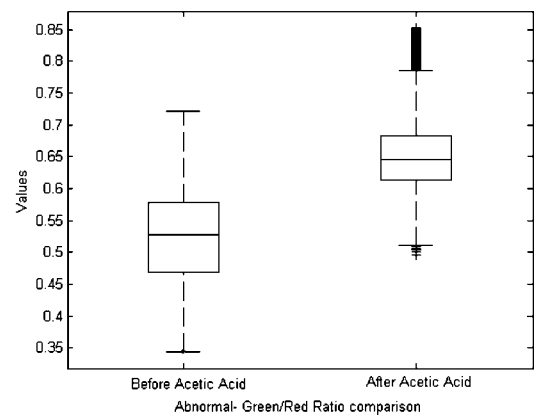
(a)



(b)

Abnormal
Samples

(c)



(d)

Fig. 2 Histogram and box plot showing the ratio of the reflectance in the green channel to that in the red channel before and after acetic acid: (a) histogram and (b) box plot of the histologically normal samples indicates little change in this feature and (c) histogram and (d) box plot of this feature increase dramatically after acetic acid for histologically abnormal specimens.

shows similar data for a patient with a completely normal ectocervix. The highest probability of abnormality in the probability map is 17%, and for any threshold value greater than 17%, all regions are diagnosed as normal. Figure 6 shows the performance of image classification for 29 patients using the ROC curve. The classification result of each patient was assessed using a grid-based approach. Using the average ROC curve, the threshold probability (0.68) with the most desirable average sensitivity (82%) and specificity (73%) values was determined in the Pareto sense.²⁶

Figure 7 shows the results of the automated image diagnosis after applying the threshold probability (0.68) for 29 patients with a normal or LGSIL or HGSIL histopathology map. Patients 1 to 8 had a normal or LGSIL histopathology map (eight patients). For patients 1 to 7, the diagnosis based on automated image analysis matches the pathology report. However, for patient 8, there is a discrepancy between the automated image analysis result and the pathology report; the automated image analysis identifies a considerable abnormal area, but the pathology map indicates that there is no abnormal tissue present.

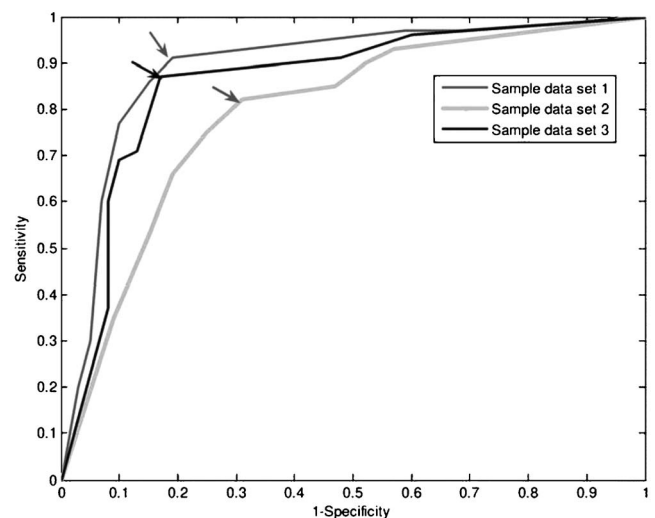


Fig. 3 ROC curves for multiclassifier discrimination of HGSIL versus normal or LGSIL sample for three different sample data sets.

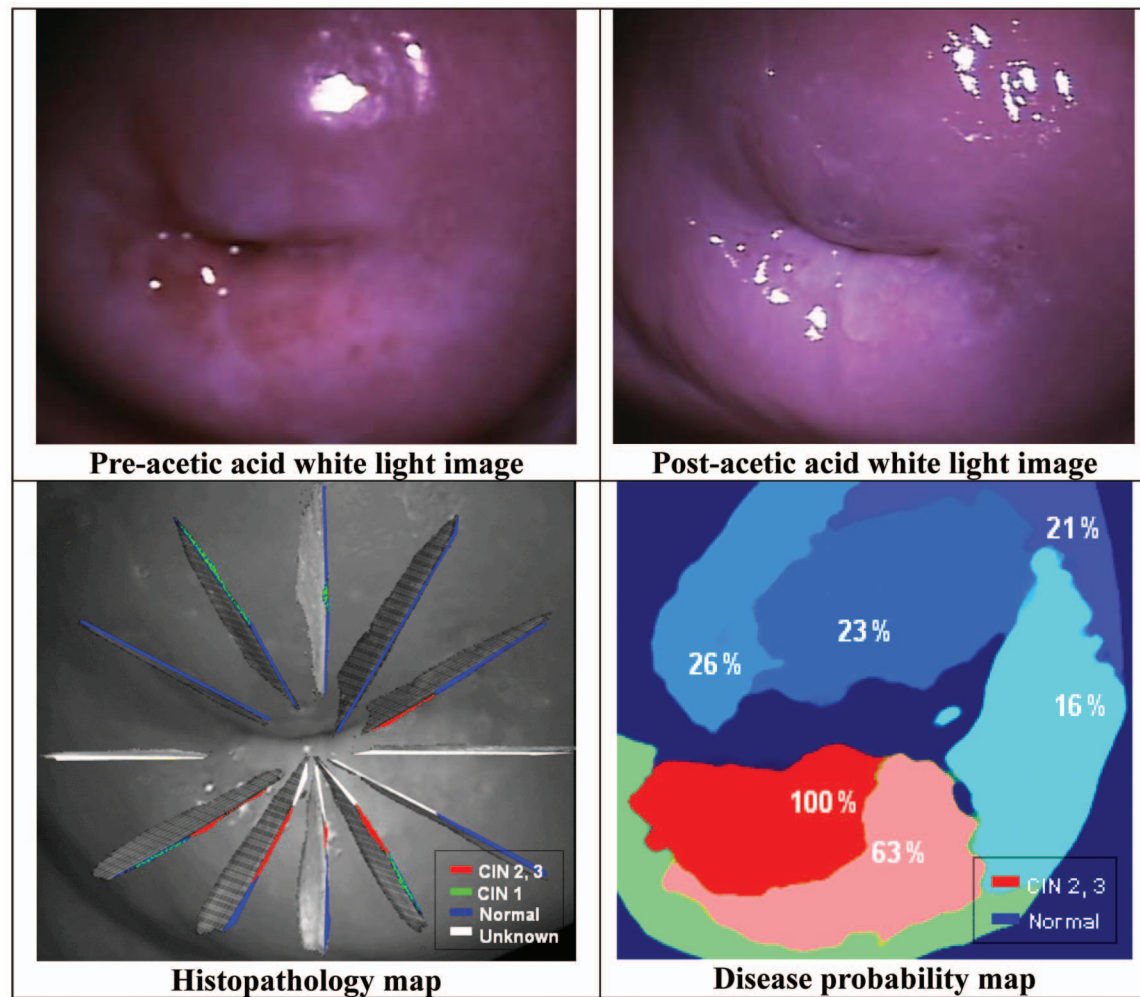


Fig. 4 Image classification results for a patient with CIN 1, 2, and 3 lesions using the proposed image classification approach.

Patients 9 to 29 shown in Fig. 7 demonstrated to have CIN 2 or 3 by LEEP (21 patients). For 15 patients where the classification algorithm yielded a true positive result, the area of HGSIL identified from the classification algorithm correlates well with that from the histopathology map (patients 9 to 23). Had the algorithm been used to determine whether treatment was indicated, it would have resulted in a correct decision. For six patients, the classification algorithm yielded an incorrect result (patients 24 to 27) or inconclusive result (patients 28 and 29). In four of these patients, the automated image analysis procedure did not identify areas of HGSIL later identified in the LEEP. In two patients, areas of HGSIL were identified in the image analysis procedure, but they did not correspond to the locations where HGSIL was later documented in the histopathology map. The resulting sensitivity and specificity values are 79% and 88, respectively, for 27 patients, not including the two inconclusive results.

4 Discussion

This pilot study shows promising agreement between our white-light image analysis results and histopathology with a sensitivity of 79% and a specificity of 88%. This is comparable to the performance of conventional colposcopy, where

the sensitivity and specificity have been reported^{27–29} to range from 87 to 96% and 34 to 85%. The reported sensitivity and specificity values were obtained without considering the inconclusive results of two patients. If they are considered to be false negative, the sensitivity is diminished to 71%. If they are considered to be true positives, the sensitivity increases to 81%.

Several other image processing approaches have been explored to aid in colposcopic image interpretation. Intensity change due to acetic acid application has been utilized in many diagnostic image analysis studies. Pogue et al.³⁰ analyzed acetowhitening of HGSIL lesions of six patients with biopsy confirmed CIN 2/3. They showed the temporal and spatial changes of green to red intensity ratio as an important acetowhitening feature to separate cervical lesions from normal regions using time-course reflectance images. However, they did not report diagnostic performance using this feature. Balas³¹ assessed the alteration in the light-scattering properties of the cervix using the maximum contrast between acetic acid responsive and nonresponsive areas for 16 patients. The time course of the intensity of backscattered light was adopted as an important diagnostic feature. The study showed the diagnostic potential of this technique; however, the quantitative

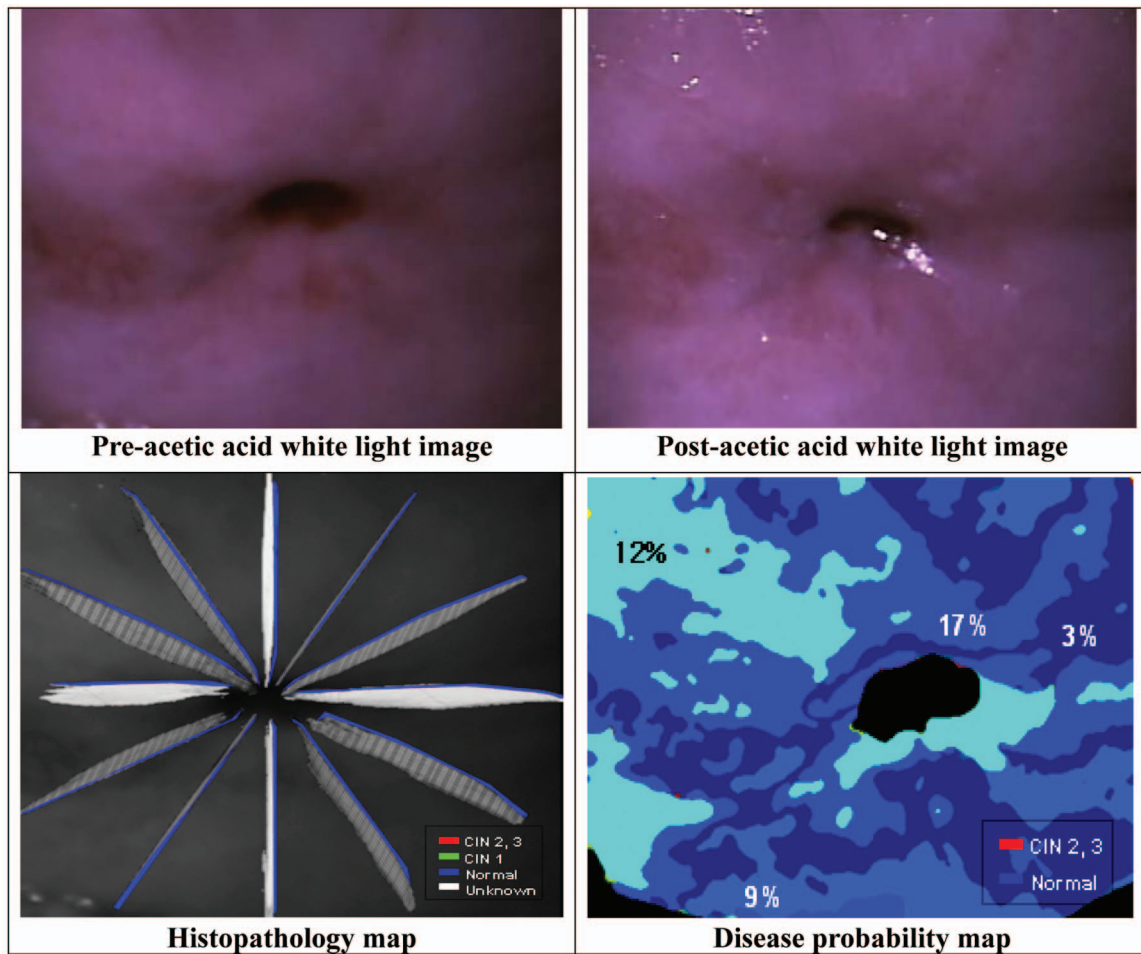


Fig. 5 Image classification results for a patient with normal ectocervix using the proposed image classification approach.

diagnostic performance of the method was not reported. In our study, we developed a gold-standard based approach to rigorously assess sensitivity and specificity in the image classification context. This enabled us to quantitatively assess the performance of our system.

A potential weakness of this study is that it was performed on patients with HG lesions; the predictive value may be lower in a screening population. In addition, our approach is limited to the detection of neoplastic lesions on the ectocervix only. To find lesions located within the endocervix, additional information is required. Moreover, our proposed approach may fail to detect precancerous lesions, which do not show acetowhitening. There may be lesions that do not have keratin and thus are not acetowhite. We suspect that glandular lesions, both *in situ* and invasive, may not be detected using this algorithm. Early and advanced invasive squamous cancers are often erythematous and not acetowhite. We are continuing to explore the addition of fluorescent light to enable detection of those lesions that are not keratinized and thus not acetowhite.

A major advantage of digital colposcopy with automated image analysis is that it requires only limited resources, so clinical application of this technique could provide an inexpensive alternative to Pap screening and conventional colposcopic assessment. Furthermore, this technique could also provide a map showing where to perform biopsies for provid-

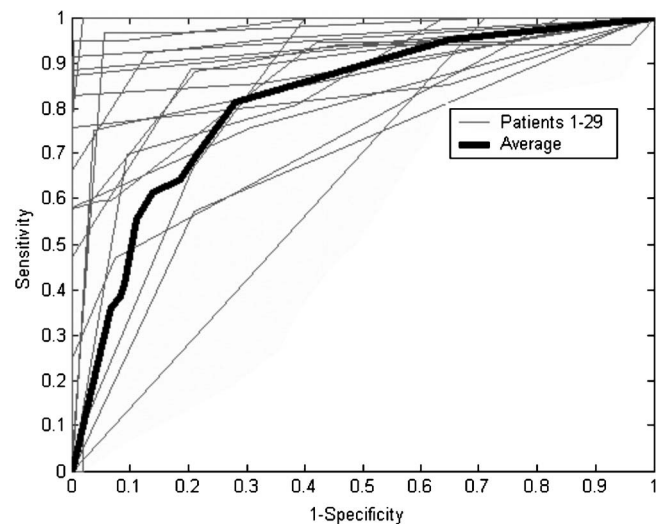


Fig. 6 ROC curves for 29 patient image classification. The best threshold value (0.68) was assessed from the point of the average ROC curve corresponding to 82% of sensitivity and 73% of specificity.






























				
Pt 1: True Negative	Pt 2: True Negative	Pt 3: True Negative	Pt 4: True Negative	Pt 5: True Negative
				
Pt 6: True Negative	Pt 7: True Negative	Pt 8: False Positive	Pt 9: True Positive	Pt 10: True Positive
				
Pt 11: True Positive	Pt 12: True Positive	Pt 13: True Positive	Pt 14: True Positive	Pt 15: True Positive
				
Pt 16: True Positive	Pt 17: True Positive	Pt 18: True Positive	Pt 19: True Positive	Pt 20: True Positive
				
Pt 21: True Positive	Pt 22: True Positive	Pt 23: True Positive	Pt 24: False Negative	Pt 25: False Negative
				
Pt 26: False Negative	Pt 27: False Negative	Pt 28: False Negative (Correct Decision)	Pt 29: False Negative (Correct Decision)	

Fig. 7 Automated image-based diagnosis results for 29 patients with a normal epithelium or LGSIL or HGSIL within the ectocervix by LEEP. A threshold was applied to the disease probability map, and those regions diagnosed as normal are represented in black, while the areas diagnosed as HGSIL are shaded in gray. The area enclosed by a white line represents the area reported as CIN 1 or 2 from the pathology map. The algorithm yielded 7 true negative results, 1 false positive result, 15 true positive results, and 6 false negative results.

ers with less experience. In the future, with further research to help connect point spectroscopic based detection and multi-spectral imaging, this technique may provide an inexpensive, *in vivo* screening solution in an automated manner. For practical applications of this technique, its cost effectiveness will need to be further analyzed and compared with those of other screening alternatives in low-resource settings.

Digital colposcopy, if the equipment can be produced inexpensively, may provide an alternative to DVI, VIA, and VILI for screening in low-resource settings. While DVI is less expensive in terms of instrumentation and methods, it is an entirely provider-dependent screening test. Its success depends on extensive formal training and experience. In the developing world, where health care resources are limited and there are few trained health care providers, screening tools with minimal training requirements are essential.³² Digital colposcopy with automated image analysis has the potential to

reduce the cost and time associated with training, leading to faster development and deployment of effective screening programs for cervical cancer throughout the developing world.

Acknowledgments

Support from the National Cancer Institute, Program Project Grant PO1 CA 82710-04 is gratefully acknowledged. We would also like to acknowledge the contributions of the research staff (Sylvia Au, Olya Shuhatovich, and Trey Kell) and nurse colposcopists (Judy Sandella, Alma Sbach, and Karen Rabel).

References

1. Cancer Facts & Figures 2006, available at: <http://www.cancer.org>.
2. L. G. Koss, "The Papanicolaou test for cervical cancer detection. A triumph and a tragedy," *J. Am. Med. Assoc.* **261**, 737–743 (1989).

3. D. M. Parkin, F. Bray, J. Ferlay, and P. Pisani, "Global cancer statistics," *Ca-Cancer J. Clin.* **55**, 74–108 (2005).
4. S. J. Goldie, L. Gaffikin, J. D. Goldhaber-Fiebert, A. Gordillo-Tobar, C. Levin, C. Mahé, and T. C. Wright, "Cost-effectiveness of cervical-cancer screening in five developing countries," *N. Engl. J. Med.* **353**, 2158–2168 (2005).
5. B. M. Nene et al., "Early detection of cervical cancer by visual inspection: a population-based study in rural India," *Int. J. Cancer* **68**, 770–773 (1996).
6. R. Wesley, R. Sankaranarayanan, B. Mathew, B. Chandralekha, A. Aysha-Beegum, N. S. Amma, and M. K. Nair, "Evaluation of visual inspection as a screening test for cervical cancer," *Br. J. Cancer* **75**, 436–440 (1997).
7. P. Basu et al., "Evaluation of downstaging in the detection of cervical neoplasia in Kolkata, India," *Int. J. Cancer* **100**, 92–96 (2002).
8. M. J. Germar and M. Meriardi, "Visual inspection with acetic acid as a cervical cancer screening tool for developing countries. Review for postgraduate training course in reproductive health/chronic disease 2003," available at <http://www.gfmer.ch/> (Sep. 14, 2006).
9. R. Sankaranarayanan et al., "The role of low-level magnification in visual inspection with acetic acid for the early detection of cervical neoplasia," *Cancer Detect. Prev.* **28**, 345–351 (2004).
10. L. Denny, L. Kuhn, A. Pollack, and T. Wright, "Direct visual inspection for cervical cancer screening: an analysis of factors influencing test performance," *Cancer* **94**, 1699–1707 (2002).
11. R. Sankaranarayanan, R. Rajkumar, R. Theresa, P. O. Esmy, C. Mahe, K. R. Bagyalakshmi, S. Thara, L. Frappart, E. Lucas, R. Muwonge, S. Shanthakumari, D. Jeevan, T. M. Subbarao, D. M. Parkin, and J. Cherian, "Initial results from a randomized trial of cervical visual screening in rural south India," *Int. J. Cancer* **109**, 461–467 (2004).
12. J. Jeronimo, O. Morales, J. Horna, J. Pariona, J. Manrique, J. Rubinos, and R. Takahashi, "Visual inspection with acetic acid for cervical cancer screening outside of low-resource settings," *Rev. Panam Salud Publica* **17**(1), 1–5 (2005).
13. S. Bomfim-Hyppolito, E. S. Franco, R. G. Franco, C. M. de Albuquerque, and G. C. Nunes, "Cervicography as an adjunctive test to visual inspection with acetic acid in cervical cancer detection screening," *Int. J. Gynaecol. Obstet.* **92**, 58–63 (2006).
14. D. G. Ferris, R. A. Lawhead, E. D. Dickman, N. Holtzapfel, J. A. Miller, S. Grogan, S. Bambot, A. Agrawal, and M. L. Faupel, "Multimodal hyperspectral imaging for the noninvasive diagnosis of cervical neoplasia," *J. Low. Genit. Tract Dis.* **5**(2), 65–72 (2001).
15. W. K. Huh, R. M. Cestero, F. A. Garcia, M. A. Gold, R. M. Guido, K. McIntyre-Seltman, D. M. Harper, L. Burke, S. T. Sum, R. F. Flewelling, and R. D. Alvarez, "Optical detection of high-grade cervical neoplasia *in vivo*: results of a 604 patient study," *Am. J. Obstet. Gynecol.* **190**, 1249–1257 (2004).
16. W. E. Crisp, B. L. Craine, and E. A. Craine, "The computerized digital imaging colposcope: future directions," *Am. J. Obstet. Gynecol.* **162**, 1491–1497 (1990).
17. B. L. Craine and E. R. Craine, "Digital imaging colposcopy: basic concepts and applications," *Obstet. Gynecol. Clin. North Am.* **82**, 869–873 (1993).
18. M. I. Shafi, J. A. Dunn, R. Chenoy, E. J. Buxton, C. Williams, and D. M. Luesley, "Digital imaging colposcopy, image analysis and quantification of the colposcopic image," *Br. J. Obstet. Gynaecol.* **101**, 234–238 (1994).
19. P. M. Cristoforoni, D. Gerbaldo, A. Perino, R. Piccoli, F. J. Montz, and G. L. Capitanio, "Computerized colposcopy: results of a pilot study and analysis of its clinical relevance," *Obstet. Gynecol. (N.Y., U. S.)* **85**, 1011–1016 (1995).
20. W. Li, V. Van Raad, J. Gu, U. Hansson, J. Hakansson, H. Lange, and D. Ferris, "Computer-aided diagnosis (CAD) for cervical cancer screening and diagnosis: a new system design in medical image processing," in *Lect. Notes Comput. Sci.*, **3765**, 240–250 (2005).
21. J. M. Benavides et al., "Multispectral digital colposcope for *in vivo* detection of cervical cancer," *Opt. Express* **11**, 1223–1236 (2003).
22. S. Park et al., "Multispectral digital microscopy for *in vivo* monitoring of oral neoplasia in the hamster cheek pouch model of carcinogenesis," *Opt. Express* **13**, 749–762 (2005).
23. "The Future II Study Group, Effect of prophylactic human papillomavirus L1 virus-like-particle vaccine on risk of cervical intraepithelial neoplasia grade 2, grade 3, and adenocarcinoma *in situ*: a combined analysis of four randomised clinical trials," *Lancet* **369**, 1861–1868 (2007).
24. H. Boonstra, J. W. Oosterhuis, A. M. Oosterhuis, and G. J. Fleuren, "Cervical tissue shrinkage by formaldehyde fixation, paraffin wax embedding, section cutting and mounting," *Virchows Arch. A: Pathol. Anat. Histopathol.* **402**, 195–201 (1983).
25. R. Kohavi, "A study of cross-validation and bootstrap for accuracy estimation and model selection," in *Proc. Int. Joint Conf. on Artificial Intelligence*, pp. 1137–1143, San Mateo, CA (1995).
26. J. P. Aubin, *Optima and Equilibria: An Introduction to Nonlinear Analysis*, 2nd ed., Springer-Verlag, Berlin (1998).
27. M. Follen, D. Schottenfeld, G. Tortolero-Luna, S. B. Cantor, and R. Richard-Kortum, "Colposcopy for the diagnosis of squamous intraepithelial lesions: a meta-analysis," *Obstet. Gynecol. (N.Y., NY, U. S.)* **91**, 626–631 (1998).
28. J. L. Benedet, G. H. Anderson, J. P. Mastic, and D. M. Mille, "A quality-control program for colposcopic practice," *Obstet. Gynecol. (N.Y., NY, U. S.)* **78**, 872–875 (1991).
29. L. Seshadri, P. Jairaj, and H. Krishnaswami, "Colposcopy in the diagnosis of cervical neoplasia," *Indian J. Cancer* **27**, 180–186 (1990).
30. B. W. Pogue, H. B. Kaufman, A. Zelenchuk, W. Harper, G. C. Burke, E. E. Burke, and D. M. Harper, "Analysis of acetic acid-induced whitening of high-grade squamous intraepithelial lesions," *J. Biomed. Opt.* **6**, 397–403 (2001).
31. C. Balas, "A novel optical imaging method for the early detection, quantitative grading, and mapping of cancerous and precancerous lesions of cervix," *IEEE Trans. Biomed. Eng.* **48**, 96–104 (2001).
32. M. F. Parker, "Emerging technology in cervical cancer screening: spectroscopy," *Clin. Obstet. Gynecol.* **48**, 209–217 (2005).