Computer-Assisted Diagnosis in Colposcopy: Results of a Preliminary Experiment?

Grit Mehlhorn a  Christian Münzenmayer b  Michaela Benz b  Andreas Kage b  Matthias W. Beckmann a  Thomas Wittenberg b

a Department of Gynecology, Erlangen University Hospital, and b Fraunhofer Institute for Integrated Circuits IIS, Erlangen, Germany

Abstract

Purpose: Diagnosis of cervical intraepithelial neoplasia (CIN) is currently based on the histological result of an aiming biopsy. This preliminary study investigated whether diagnostics for CIN can potentially be improved using semiautomated colposcopic image analysis. Methods: 198 women with unremarkable or abnormal smears underwent colposcopy examinations. 375 regions of interest (ROIs) were manually marked on digital screen shots of the streaming documentation, which we provided during our colposcopic examinations (39 normal findings, 41 CIN I, and 118 CIN II–III). These ROIs were classified into two groups (211 regions with normal findings and CIN I, and 164 regions with CIN II–III). We developed a prototypical computer-assisted diagnostic (CAD) device based on image-processing methods to automatically characterize the color, texture, and granulation of the ROIs. Results: Using n-fold cross-validation, the CAD system achieved a maximum diagnostic accuracy of 80% (sensitivity 85% and specificity 75%) corresponding to a correct assignment of abnormal or unremarkable findings. Conclusion: The CAD system may be able to play a supportive role in the further diagnosis of CIN, potentially paving the way for new and enhanced developments in colposcopy-based diagnosis.

Key Words
Cervical intraepithelial neoplasia  •  Colposcopy  •  Computer-assisted diagnosis  •  Automated tissue characterization  •  Image analysis

Introduction

In Germany, a substantial reduction of the incidence of cervical cancer as well as of the resulting morbidity and mortality could be achieved by the introduction of a cytological public screening system in the early 1970s [1]. The aim of the Pap smear is to identify precancerous lesions in the genitalia as early as possible, clarify them in a differentiated way, and treat them when appropriate. Once an abnormal cytological finding leads us to believe that a precancerous lesion or even a carcinoma could be present, the diagnostic guidelines require a histological analysis to be carried out for confirmation. To enhance the diagnostic accuracy of the assessment, colposcopic examination is an absolute requirement [2–4]. With much-improved visual conditions and the use of various contrast-enhancing solutions (such as acetic acid), better diagnostic clarification is possible on the basis of a targeted biopsy [5]. This type of examination requires exten-
sive knowledge, and it is also highly dependent on the examiner’s objective level of experience [6]. Colposcopy had a sensitivity of 86% and a specificity of 30% in distinguishing a healthy cervix from one with cervical intraepithelial neoplasia (CIN)/carcinoma. In distinguishing healthy cervices with low-grade lesions (CIN 1) from those with high-grade lesions (CIN 2/3)/carcinoma, colposcopy had a sensitivity of 61% and a specificity of 94% [7]. Despite this, differential colposcopy of this type determines whether a targeted biopsy should be taken and whether histological diagnosis and relevant treatment are carried out [8, 9]. In addition to the development and evaluation of this type of enhanced assessment method, advances in modern information technology also need to be taken into account. These can provide interactive or automatic support for diagnosis and reporting, and they may also be helpful for training and quality control purposes. Computer-assisted diagnosis (CAD) systems of this sort are already becoming accepted in various medical domains, particularly in preclinical applied research. For example, for the detection of nodules in the lung by computed tomograms [10] and lesions in mammograms [11–13], assessment of endoscopic images in gastroenterology [14–16], or in the field of otorhinolaryngology for analyzing so-called visual biopsies of leukoplakia on the vocal folds [17], preclinical studies have shown that this type of CAD system can be quite helpful for physicians for reporting purposes and for establishing the final diagnosis.

**Materials and Methods**

A total of 198 women with either unremarkable or abnormal smears were retrospectively analyzed. The findings were reexamined using a prototypic CAD system for colposcopy developed by the Fraunhofer IIS [18, 19]. As the automated analysis of colposcopic images is exclusively based on the visual characteristics of the images, only images that were taken with the same magnification level (×0.6) and in which the complete vaginal part of the patient’s cervix was visible after acetic acid application (fig. 1, 2) were analyzed. In addition, the images were selected relative to the focus, contrast and brightness quality of the picture, and in relation to the size and shape of dysplastic changes and their demarcation from the surroundings. The colposcopic criteria for cervical dysplasia in accordance with the new nomenclature had to be appreciable and assignable [5]. Furthermore, as a selection criteria of dysplasia in colposcopic images, the related histological findings had to correspond. During the dysplasia consultation, the internal and external genitalia were initially examined using a cytological smear and colposcopic visualization with acetic acid

![Fig. 1. Colposcopic nomenclature RIO 2011.](image-url)
(5%). The capture of the image was taken 1–2 min after application of acetic acid. The colposcopist was accredited by the German Society of Cervical Pathology and Colposcopy. The visual findings were assessed and documented in accordance with RIO 2011 nomenclature, and in case of abnormal findings the relevant histological analysis was carried out.

**Image Capture and Image Data**

The visual colposcopic examination was carried out with a binocular colposcope. All images were captured immediately after the acetic acid test using a video camera (Zeiss) mounted on the colposcope, digitized with a frame-grabber card (EYE-2p, RGB), and stored for documentation purposes in the compressed JPEG image format in the ViewPoint image archiving system (GE Healthcare/ViewPoint Bildverarbeitung Ltd., Wessling, Germany). The images for the cases were retrospectively classified into two groups, with one image for each patient. Group A included images from 39 patients with normal findings and 41 patients with low-grade intraepithelial lesions (CIN I). In group B, images from 118 patients with higher-grade intraepithelial lesions were used, corresponding to moderate to severe intraepithelial neoplasia (CIN II and CIN III). An experienced examiner retrospectively outlined and marked 375 representative regions of interest (ROIs) from the collection of 198 colposcopy images and assigned them manually to the two groups using an image annotation tool developed for the study. A total of 211 ROIs (normal/CIN I) in group A and 164 ROIs (CIN II/III) in group B were marked. Higher-grade dysplastic and histologically confirmed lesions (group B) were marked in red and normal or only slightly dysplastic tissue (group A) was marked in green. Figure 1 shows an image with unremarkable findings, while figure 2 illustrates high-grade changes (CIN III) in the cervix.

**Fig. 2.** Green marking of unremarkable findings in the uterine cervix.

**Fig. 3.** Red marking of abnormal findings, with a CIN III lesion.
Features

Using color texture analysis methods, numerical features were calculated for each of the marked ROIs and used to characterize the marked surface structures on the basis of intensities, structures, textures, and color. The following methods in particular were used [20, 21]:

Color histograms (HSTs) show the intensity (or spectral) distributions for all three color channels in an image. In addition to using the three channels from color images in the red, green, and blue (RGB) color space, other color spaces, like for example the Ohta or HSI (hue – saturation – intensity) color spaces [21], can be used to calculate the spectral distributions. To describe the color distributions in an ROI, the HST can be used directly, or representative features such as moments, variances, and energies can be calculated for each sub-histogram.

Color extended statistical geometric features (CESGFs; [22]) describe geometrical micro-structures in an ROI based on statistical features (number, irregularity, mean, and variance) computed from a stack of binary images describing the micro-structures at various levels of intensity. In order to incorporate color into the analysis, the micro-structure statistics are either computed in each color channel alone or correlated with each other during the computation by pairwise combining micro-structures logically.

Color extended gray-level co-occurrence matrices (CEGLCMs) are the most familiar procedures in the field of texture-based tissue characterization. A gray-level co-occurrence matrix (GLCM) describes the frequencies of the co-occurrence of gray-level combinations of the pixels in a certain ROI. From the thus obtained GLCM, a set of 14 features based on statistical characteristics (mean, variance, energy, correlation, entropy, etc.) is computed and used to characterize the underlying tissue. To include color in the computation, the obtained features from the three color channels are fused or, during the calculation of the GLCMs, information from pairwise color channels is combined [21].

Color extended sum and difference histograms (CESDHs) are a fast version of GLCM in which the matrix calculation is approximated using two histograms.

Classification

To evaluate the CAD system, each of the 375 manually marked image regions (ROIs) was characterized using the groups of features described above (HST, CESGF, CESDH, and CEGLCM) and their variants (resolution reduction, image preprocessing, color correction, and color spaces). On the basis of the feature vectors obtained in this way, a k-nearest-neighbor (kNN) classifier [23] with n-fold cross-validation was used to evaluate the characteristics of these feature groups relative to the automatic separability of the two groups, i.e. CIN II–III versus CIN I/benign. The kNN classifier selects, for an unknown query vector \( q \) with characteristic features, those \( k \) vectors in a database of preclassified and histologically validated reference vectors that are most similar to the query vector \( q \) relative to a defined distance measure. The diagnostic group and assignment probability of the unknown query vector \( q \) are determined on the basis of the vectors most similar to \( q \). The kNN has been shown in a large number of studies to be a strong and powerful approach for tasks involving imaging-based distinctions and classifications for medical applications [23].

Results

Digital colposcopies for 198 patients, with the 376 manually marked image areas, were used to evaluate the proposed CAD system. The average age of the women examined was 31 years; the youngest patient was 16 and the oldest was 74 years old. The peak age range was 21–42 years.

Normal findings were noted in 39 of the 80 patients included in group A classified and documented as belonging to the current RIO 2011 colposcopic terminology [24]. Eight patients had atrophic findings, 11 patients had metaplastic findings, and 20 patients had ectopy of the cervix; they were well assessed using coloscopy and were therefore assigned to group A. Forty-one patients in examination group A had imaging findings with fine acetic acid-positive areas (low-grade CIN). In accordance with the current colposcopic nomenclature, this corresponds to grade 1 abnormal colposcopic findings and is assigned to ‘minor changes’. Minor changes on coloscopy involve fine acetic acid-positive findings, with or without a fine punctate or mosaic pattern [25, 26]. As there is no major difference in treatment implications between this and normal or low-grade lesions, these were assigned to examination group A.

Examination group B included 99 patients with high-grade cervical dysplasia (CIN III) and 19 patients with moderate dysplasia (CIN II). These findings correspond to grade 2 abnormal colposcopic findings in the current colposcopic nomenclature and are therefore assigned to ‘major changes’. Major changes describe coloscopically dense acetic acid-positive findings, with or without coarse punctate or coarse mosaic patterns [24]. All dysplastic lesions are histologically confirmed.

Technical Results

The diagnostic accuracy currently achievable using the automatic classification of colposcopic lesions – on the basis of the data presented above, in which the CAD system is based on a distinction between high-grade and low-grade lesions – was assessed using what is known as n-fold cross-validation in the four groups of features described on their variants. The highest overall classification rate of 80% (with a sensitivity of 85% and a specificity of 75%) was achieved with CESDHs in combination with image enhancement steps for color correction. The highest specificity of 88% for recognizing CIN II–III lesions was achieved with HST in the Ohta color space, with an overall classification rate of 75% and a sensitivity of 64%. A maximum sensitivity (for recognizing CIN I/
normal tissue) of 75% was achieved with the statistical geometric features, with an overall classification rate of 78% and a specificity of 82%.

Discussion

The CAD system for colposcopic images investigated here shows better accuracy in distinguishing between unremarkable or slightly abnormal cervical findings of very low clinical relevance and higher-grade cervical findings than previous routine diagnosis using a purely visual approach. In indirect comparison with this, Hammes et al. [7] were only able to achieve a maximum sensitivity of 61% with a specificity of 94% for enhanced colposcopy in 468 patients with colposcopically evaluated cervical abnormalities when comparing healthy and slightly altered tissue with high-grade and carcinomatous abnormal tissue. Similarly low sensitivity rates were reported by Soutter et al. [25, 26]. A high sensitivity rate of 86% for a colposcopic examination to distinguish between healthy and dysplastic cervical epithelium has been reported [7]. Colposcopy during statutory screening for early cancer detection is considered to be clearly inferior to cytology in routine diagnosis. For this reason, colposcopy is only intended for use in diagnostic clarification prior to cytology in routine diagnosis. For this reason, colposcopy is only intended for use in diagnostic clarification [24]. The additional use of CAD methods for enhanced colposcopy evidently makes it possible to achieve a marked improvement in the overall accuracy of enhanced diagnosis of clinically relevant precancerous lesions in the uterine cervix. Although these initial results have not been examined further statistically, they can nevertheless be regarded as an improvement on the results reported in the literature with conventional, purely visual colposcopy. In our view, further improvement in the overall recognition rate, sensitivity, and specificity, and in the ability to distinguish between low-grade and high-grade changes in the cervix, would be possible both by standardizing the image capture procedure and by developing specific imaging features. Standardizing the image capture procedure would help improve the CAD approach by producing colposcopic images at a constant image quality, with standard lighting and a standardized image section; and the development and use of new, dedicated features for characterizing tissue could allow the identification of additional differences between high-grade and low-grade lesions. In particular, the potential increase in sensitivity provided by supplementary usage of a CAD system suggests decisive clinical relevance with regard to biopsy clarification and subsequent surgical treatment, as the number of unnecessary biopsies could be reduced. Using digital colposcopy, Schädel et al. [27, 28] investigated similar approaches to automated ways of improving diagnostic processes and thereby increasing quality control. CAD of the type evaluated for the first time in the present study on a model basis is not capable of competing with an experienced examiner, far less replacing him or her. However, appropriate CAD systems have tremendous potential in a supportive capacity in the further diagnosis of CIN. CAD systems will not be able to compete with the diagnostic confirmation provided by histology, but they may be able to reduce the need for histological examinations. In principle, this type of CAD system can make it possible in cases of diagnostic uncertainty to use an ‘objective’ analysis of features in order to systematically improve colposcopy-based diagnosis. The use of standardized imaging techniques to acquire pictures, reproducible cases, and the availability of immediate feedback from the user to the system make the application of this type of CAD system appear feasible both for direct diagnosis and indirectly for training purposes. Further research on such systems is therefore justified. Even though we know that the described experiment is only a preliminary one, based on less than 200 cases, we strongly believe that the first results obtained show the potential of such a system compared to the results cited from the literature. We are currently working toward this end by collecting more data with the same image quality and image resolution, and we are planning such an experiment. Nevertheless, compared to the results in the literature, we believe that our experiments and findings are already showing the tendency and potential to support the assessment of colposcopic images and diagnostic findings.

Disclosure Statement

There are no conflicts of interest.

References


Mehlhorn/Münzenmayer/Benz/Kage/ Beckmann/Wittenberg
Computer-Assisted Colposcopy Diagnosis

Acta Cytologica 2012;56:554–559


