Medical & Biological Engineering & Computing (2021) 59:1545–1561
https://doi.org/10.1007/s11517-021-02388-w

ORIGINAL ARTICLE

A whole-slide image grading benchmark and tissue classification for cervical cancer precursor lesions with inter-observer variability

Abdulkadir Albayrak[1,6] · Asli Unlu Akhan[2] · Nurullah Calik[3] · Abdulkerim Capar[4] · Gokhan Bilgin[5,6] · Behcet Ugur Toreyin[4] · Bahar Muezzinoglu[7] · Ilknur Turkmen[8] · Lutfiye Durak-Ata[4]

Received: 25 March 2020 / Accepted: 3 June 2021 / Published online: 10 July 2021

Abstract
The cervical cancer developing from the precancerous lesions caused by the human papillomavirus (HPV) has been one of the preventable cancers with the help of periodic screening. Cervical intraepithelial neoplasia (CIN) and squamous intraepithelial lesion (SIL) are two types of grading conventions widely accepted by pathologists. On the other hand, inter-observer variability is an important issue for final diagnosis. In this paper, a whole-slide image grading benchmark for cervical cancer precursor lesions is created and the "Uterine Cervical Cancer Database" introduced in this article is the first publicly available cervical tissue microscopy image dataset. In addition, a morphological feature representing the angle between the basal membrane (BM) and the major axis of each nucleus in the tissue is proposed. The presence of papillae of the cervical epithelium and overlapping cell problems are also discussed. Besides that, the inter-observer variability is also evaluated by thorough comparisons among decisions of pathologists, as well as the final diagnosis.

Keywords Cervical cancer · Human papillomavirus · Cervical intraepithelial neoplasia (CIN) · Squamous intraepithelial lesion (SIL) · Digital pathology · Whole-slide imaging · Histopathological images · Morphological features · Inter-observer variability

1 Introduction

Cervical cancer is one of the most commonly seen cancer types in the world and the fourth most common cause of death, which develops from precursor lesions [35]. Studies indicate that, almost all cervical cancer cases develop by the effect of human papillomavirus (HPV), which reaches epithelial basal layer cells with the help of micro-injuries in the cervical epithelium. Carcinogenic effect of the virus occurs when HPV's genome integrates with the cell genome [33, 36, 40]. This effect, which requires a certain period of time, appears as morphological changes in the cervical epithelium. These precancerous lesions characterized by dysplastic changes are called squamous intraepithelial lesions (SIL).

Impact of HPV on the cervical epithelium varies throughout the life cycle of the virus, which in turn results in different morphological changes. Early diagnosis can be made possible by the analysis of these morphological structures [32, 35]. After being infected by HPV, basal cells proliferate and the epithelium loses its maturation. As well as the loss of maturation which results in polarity loss in the epithelium, cells show nuclear enlargement, nuclear irregularity, and hyperchromasia. Depending on the proliferation process, the number of mitoses also increases. The effect of viral proteins on the cyto-skeleton reveals halo cells with characteristic perinuclear halo named

✉ Gokhan Bilgin
gbilgin@yildiz.edu.tr

1 Department of Computer Engineering, Dicle University, Diyarbakır, 21280, Turkey
2 Selahaddin Eyyubi State Hospital, Diyarbakır, 21100, Turkey
3 Department of Biomedical Engineering, Istanbul Medeniyet University, Istanbul, 34700, Turkey
4 Informatics Institute, Istanbul Technical University, Istanbul, 34469, Turkey
5 Department of Computer Engineering, Yildiz Technical University (YTU), 34220, Istanbul, Turkey
6 Signal and Image Processing Laboratory (SIMPLAB), Yildiz Technical University, Istanbul, 34220, Turkey
7 Department of Pathology, Istanbul Medipol University Hospital, Istanbul, 34214, Turkey
8 Pathology Center of Memorial Healthcare Group, Istanbul, 34385, Turkey

"koilocytes" (it means hollow in Greek). These dysplastic changes are graded according to whether they are seen among the lower, middle, and upper parts of the epithelium [16, 28, 30]. They are reported according to Cervical Intraepithelial Neoplasia (CIN) 1–3 in the CIN-based grading and LSIL (low-grade squamous intraepithelial lesion) and HSIL (high-grade squamous intraepithelial lesion) in the SIL-based grading [4, 6, 25]. Currently, the use of SIL-based grading is recommended, yet the CIN-based grading is also used.

Pathological diagnosis of cervical biopsies varies depending on artifacts due to laboratory steps and pathological interpretation. Due to the spread of women health screening programs, the diagnosis of cervical biopsies is frequently encountered and this diagnosis variability has become a more important problem [14, 21]. Cervical biopsy interpretation has inter- or intra-observer variability, which means that a biopsy may have different diagnoses by different pathologists or by the same pathologist at different times, and it is accepted to an extent in the literature [22, 34]. Studies have been made to overcome this problem with classification systems suitable for the nature of HPV or with the help of immunohistochemical techniques [6, 8, 33].

Increasing role of the information technology (IT) on the area of medicine has a positive impact on the pathology. Digital pathology includes diagnosis, education, consultation, archiving, and also morphometric evaluation tools [2, 19]. Studies about morphometric analysis are available for different tissues and systems [12], as well as for cervical lesions [7, 11, 26, 37] in the literature. De et al. [7] studied image analysis methods on 62 digital images of cervical epithelial lesions labeled with Normal, CIN1, CIN2, and CIN3. The cervical regions are manually marked by the pathologist on selected epithelial images, and these regions are divided into vertical segments by calculating the medial axis. The obtained epithelial segments are examined in terms of structural, geometric, and profile-based properties. Contrast, energy levels, pixel correlation values, and neighborhood features of the pixels within the vertical segment are studied as the structural features. Geometric features include the distances between nuclei centers and Delaunay triangulation. In the profile-based feature extraction, correlation values and the brightness values of all pixels of each row of vertical segment are calculated. Linear discriminant analysis (LDA) and support vector machines (SVM) are utilized to classify feature vectors of vertical segments. First, each of the vertical segments is classified individually, then these decisions are fused to obtain a whole epithelium classification result. The effect of individual decisions of vertical segments on the whole epithelium classification result is also examined. One-to-one correspondence between the system result and pathological diagnosis is named as "Exact Class Label" (1st approach), only one class difference between system result and pathological diagnosis is named as "Windowed Class Label" (2nd approach), and bigger differences between system result and pathological diagnosis are named as Normal versus CIN (3rd approach). Different classification performances are calculated using different approaches and features. This study reported that using all the structural, geometric, and profile-based features, recognition rates of 62.3% on vertical segments and 39.3% on whole epithelium are reached.

Guo et al. [11] proposed enhanced image analysis methods on the cervix image dataset that was formed in De et al. [7]. They have increased the success of classification by adding structural features of the nucleus and cytoplasm in addition to the features extracted from similar vertical segments as in De et al. [7]. These features consist of nucleus, cytoplasm, and acellular areas and ratios, color scale (red, green, blue) brightness values, numbers of triangles obtained by Delaunay triangulation at upper, middle, and lower epithelium regions. The features are classified by the classification methods of the previous study. The name "Windowed Class Label" used in the second approach in the previous study is changed to "Off-By-One Class Label." In this study, as well as using the same dataset as De et al., they made a difference of examination by two different pathologists. The diagnostic success rates of the extracted features are determined by the Attribute Information Gain Ratio (AIGR) algorithm. They evaluated the success of features according to two different classification approaches. As a result of adding structural features of cervical regions, they have increased their classification success up to 82–88.5%.

Wang et al. studied morphometric analysis methods on 31 digital images of cervical biopsies [37]. Their study consists of two steps as the automated segmentation of squamous epithelium and the CIN classification. In the first step, the epithelium is segmented using the difference of the visual properties of five different regions consisting of squamous epithelium, columnar epithelium, stroma, background, and erythrocytes. The medial axis is drawn parallel to the basal membrane (BM) and upper membrane (UM) borders after the epithelial region is segmented. Square windows with 250 × 250 pixel dimensions are created on normal curves of medial axis. The feature vectors are obtained by calculating the average of nucleus area, the average area of the triangles obtained by Delaunay triangulation, and the average edge length analyzed within each window. This study reported that they have reached accuracy rates ranging from 60 to 95% by using different classification methods with the obtained feature vectors.

Keenan et al. proposed a study to analyze 230 digital cervix images consisting of normal, koilocytosis, and CIN1-, CIN2-, and CIN3-labeled lesions [15]. The feature vectors

which contain the nucleus area, the nucleus cytoplasm ratio, the ratio of nucleus area to cytoplasm, and the edges/areas of the Delaunay triangles are analyzed. The Kappa value for the observer difference between the two pathologists involved in the study is 0.415. The classification performance of the system in distinguishing normal and CIN lesions is 98.7%. This study reported that the overall success rate is 62.3%, where the worst performance is achieved on CIN2-labeled patterns.

Nagdhy et al. proposed a study to classify a total of 475 cervical biopsies with normal, CIN1, CIN2, CIN3, and invasive carcinomas using three different methods [26]. The nucleus area, core cytoplasm ratio, core boundary irregularity, and areas of Delaunay triangles are analyzed. They reached up to 97% with respect to specificity and 100% with respect to sensitivity using different methods including Gabor-based texture descriptor, gray-level co-occurrence matrix(GLCM) texture descriptor, and pre-trained convolutional neural network.

In this study, morphometric analysis methods for the cervical SIL diagnosis are investigated on a new digital cervical image dataset. The numerical values of the morphological features used by the pathologists in the diagnosis are extracted. The statistical significance of their contribution to diagnosis is examined. Within the scope of the study, a Computer Aided Diagnostic Auxiliary System (CADAS) is developed and its performance is evaluated.

The contributions of this study are as follows:

- A new whole-slide image grading benchmark for grading of cervical dysplasias is created and introduced to histopathological image analysis community. To the best of the authors' knowledge, the "Uterine Cervical Cancer Database" introduced in this article is the first publicly available cervical tissue microscopy image dataset.
- Images obtained from the dataset are labeled by two pathologists to investigate the inter-observer variability in cervical dysplasia grading.
- Pathologists diagnosed each image patch stained with hematoxylin and eosin (H&E) in the dataset independently. In the likely case of inconsistent diagnoses, the image patches that are stained with p16 and Ki67 immunhistochemical dyes are analyzed to decide a final diagnosis.
- A morphometric analysis method for cervical SIL diagnosis is proposed.
- The presence of papillaries in the dataset that leads to tangential sections is one of the important parameters that pathologists account for when diagnosing.

The remainder of the paper is organized as follows: in Section 2, the pathological pre-processing, whole-slide imaging, and the proposed method are explained. In Sections 3 and 4, experimental results and discussions are presented, respectively. Finally, conclusions are drawn in Section 5.

2 Materials and methods

This study is conducted by a group of scientists and medical researchers. Cervical tissue slide samples with diagnosis results are collected in the pathology laboratory of Istanbul Medipol University (IMU) Hospital, Istanbul, Turkey. Figure 1 shows the processing steps followed in this study for the proposed CADAS to grade cervical cancer precursor lesions.

2.1 Data collection and image acquisition

Within the scope of the study, 128 high-resolution slides from 54 patients are scanned at the pathology laboratory of IMU Hospital. Figure 2 represents the whole slide images obtained from the dataset. The images are stained with H&E dyes, and also treated with Ki67 and p16 biomarkers. Hematoxylin stains cell nuclei blue, and eosin stains the extracellular matrix and cytoplasm pink. With the help of staining, the general layout and distribution of cells can be seen, as well as an overview of the structure of a tissue sample. The Ki67 is used as a marker for a special stain of Ki67 protein to show the proliferative index of cells. The p16 immunostaining is used as a marker for a cell cycle regulatory protein that is overexpressed in cervical dysplasia for transcriptionally active HPV. All high-resolution images are then divided into 957 small epithelium pieces by the pathologist. Each slide in the dataset is diagnosed after splitting into smaller epithelial pieces. In total, 957 epithelial pieces are obtained from the whole slides. A total of 471 of the 957 images are diagnosed as normal, 240 of them are diagnosed as CIN1, 107 of them are diagnosed as CIN2 LSIL, and 57 of them are diagnosed as CIN3.

The comparison of the dataset proposed in this study with the datasets used in studies in the literature is presented in Table 1. As can be seen in Table 1, Wang et al [37] includes 31 samples as whole slide images, while our proposed dataset includes 128 whole slide images. The columns of the table indicate the publication year of the proposed studies, the presence of papillae in the images, the usage of different biomarkers, the number of images, observer variability, accessibility, and the annotation of basal and upper membranes of the epithelium. The main distinguishing importance of our dataset can be expressed as being a publicly available benchmark dataset, using different biomarkers while preparing the data, the presence of papillae, and annotation of basal and upper membranes, which shows the region of interest. Scanning the dataset as a whole slide and marking

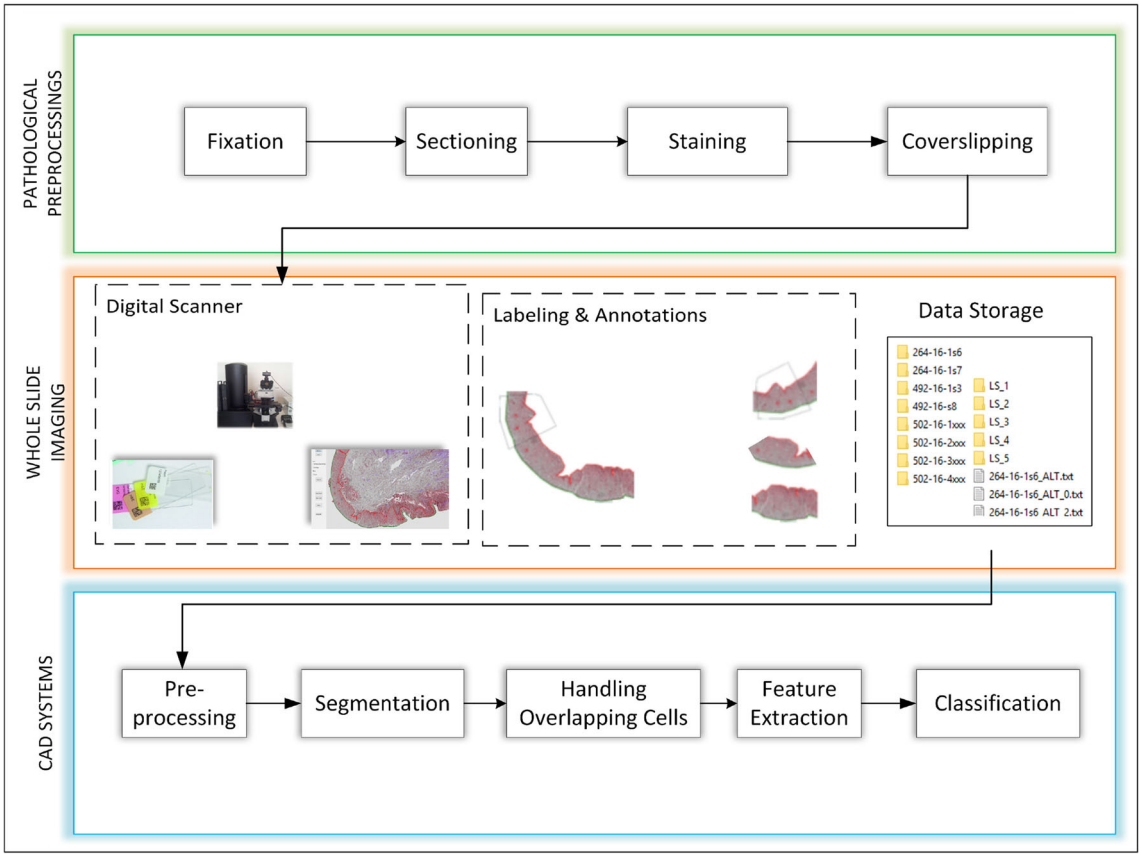


Fig. 1 Processing steps followed in the proposed study. The first step describes the pathological preprocess which is handled in the pathology laboratory. Whole-slide scanning and the filing process mentioned in the second row are done by the medical researcher and the computer scientist in collaboration. CADAS framework developed for grading the cervical cancer pre-cursor lesion is mentioned in the third row

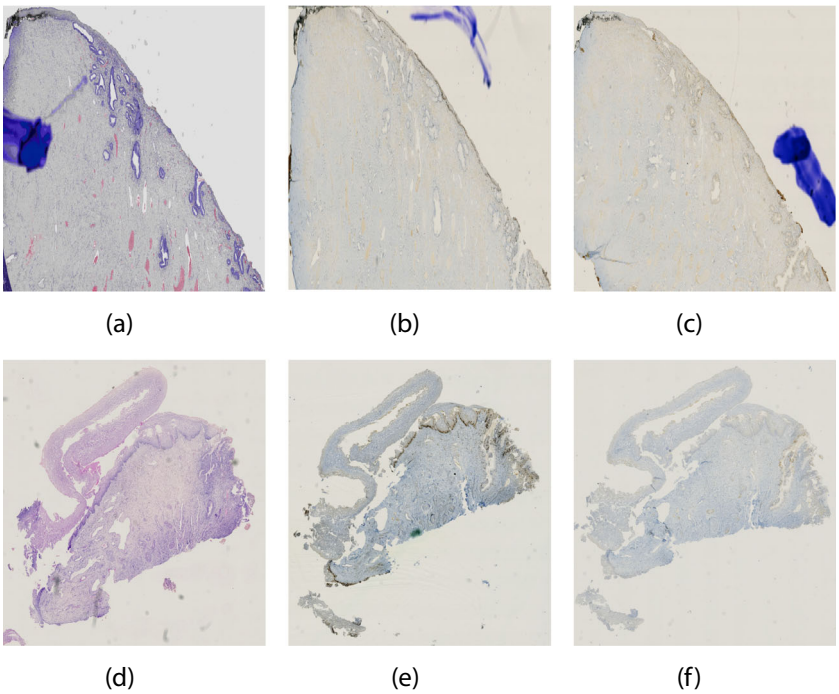


Fig. 2 Image samples obtained from the dataset. The sub-figures of **a** and **d** represent the whole-slide images which are stained with H&E; the sub-figures of **b** and **e** represent the same images stained with Ki67 immunohistochemical dye and the sub-figures **c** and **f** show the same images obtained with p16 immunohistochemical dye

Table 1 Datasets used in studies proposed to diagnose cervical squamous intraepithelial lesions

References	year	papillae	biomarkers	# of patients	# of images	inter-observer variablity	intra-observer variability	publicly available	squamous annotation	epithelium
Keenan et. al. [15]	2000	no	no	–	230 patches	yes	yes	no	no	
Wang et. al.[37]	2009	no	no	–	**31 whole slide images**	yes	yes	no	**yes**	
Miranda et. al. [24]	2012	no	no	–	160 patches	no	no	no	no	
Rahmadwati et. al. [26]	2012	no	no	–	475 patches	no	no	no	no	
De et. al. [7]	2013	no	no	–	61 patches	no	no	no	no	
Guo et. al. [11]	2016	no	no	–	61 patches	yes	yes	no	no	
Proposed	2020	**yes**	**yes**	**54**	**128 whole slide images**	yes	yes	**yes**	**yes**	

the membrane coordinates by expert pathologists will be a good reference for future studies aiming to automatically detect epithelial regions. Papillae and biomarkers have crucial importance for specialists to pay attention to in the diagnosis of cervical cancer. Squamous epithelium is located on fibrous connective tissue. The fibrous connective tissue makes finger-like projections, called papillae, into the epithelium (see Fig. 6). Papillae have blood vessels feeding the epithelium and rarely free nerve endings. Regardless which part of the epithelium the papillae is located, the cells around the papillae are assumed to be cells located in the basement membrane and analyzed accordingly by the specialist [5]. Automated approaches performed without marking the papillae regions will analyze cell morphology incorrectly and this will cause inaccurate grading.

The images of the H&E, p16, and Ki67 preparations are acquired by an off-the-shelf whole-slide scanner (see Fig. 3). The scanner has a capability of up to $\times$20optical and $\times$40 digital zoom. The whole-slide images obtained with the high-resolution scanner are transferred to the digital platform to be processed by several image processing techniques and also to be interpreted by the pathologists. The images are stored in lossless TIFF format without any loss. The sizes of the images obtained by the scanner are varied from 7500×7700 to $55,700 \times 165,000$ and there are more than one diagnosis in a single lesion.

2.2 Information about released cervix dataset

The dataset consists of 128 whole-slide images of 54 patients, where each image is placed in a separate folder. Each of these folders contains a whole-slide image and its subfolders divided into sub-lesions. Each subfolder contains an image representing the relevant sub-lesion and sub-level folders containing patch samples extracted from this sub-lesion. The reason why each sub-lesion is divided into several patch samples is that there is more than one diagnostic tissue in each sub-lesion. In each folder and its subfolders, there are two text files containing the pixel locations of the basal and upper epithelial membranes of the respective images. In addition, if there are papillae in the image, there is an extra text file containing the pixel locations of the papillae in each related folder and sub-folder. For each folder/sub-folder level, image/patch labels obtained by two experts are also included in a text file named as *label*. The "Uterine Cervical Cancer Database" introduced in this article is the first publicly available cervical tissue microscopy image dataset within the knowledge of the authors. The dataset is only allowed for academic studies. The entire dataset and explanations related to the dataset will be shared on request from http://spacing.itu.edu.tr/datasets.

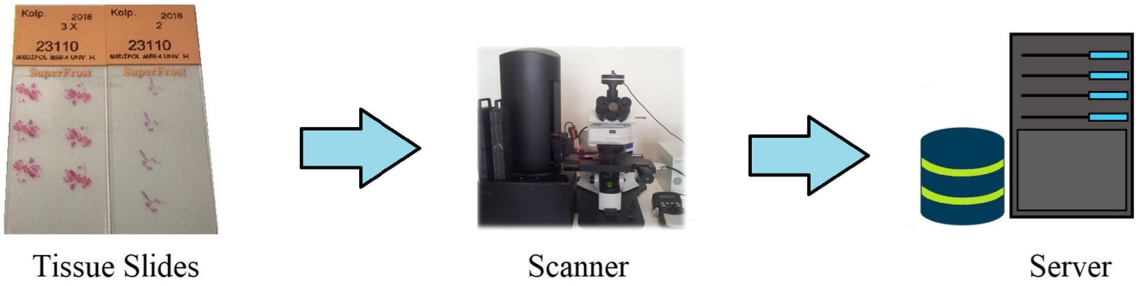


Fig. 3 Image acquisition system: Tissue slides are scanned using a high-resolution scanner. The scanned slides are then transferred to a server to store the images in a file system

2.3 Ethics statements

The authors confirm that all samples taken from patients were prepared in accordance with the legislation prepared by the Ministry of Health of Turkey and in accordance with international agreements and European Union standards. All experimental protocols were approved by the Istanbul Medipol University's licensing committee. Informed consent was obtained from all subjects whose tissue samples were used in experiments. In tissue sample collection for the dataset, there were no subjects under 18.

2.4 Annotation and image labeling

A graphical user program is developed within the scope of this study for pathologists to mark/label the BM and papillae of the cervical epithelium. After marking the membranes and the papillae, the program extracts the hot spot region from the background. Figure 4 represents a sample input image taken from the dataset and the image after marking the pixel locations of the epithelium.

At first, two pathologists made the diagnosis independently for each Small Epithelial Piece (SEP) image patch. A final diagnosis is then made by observing the same lesions stained with p16 and Ki67 immunohistochemical dyes in case of disagreement. According to the final diagnosis, 471 of SEP (49.2%) are labeled as normal, 240 of them (25.1%) are CIN1, 107 of them (11.2%) are CIN2, and 139 of them (14.5%) are CIN3. However, 150 of Large Epithelial Piece (LEP) (46.9%) are labeled as normal, 79 of LEP (24.7%) are CIN1, 34 of LEP (10.6%) are CIN2, and 57 of LEP (17.8%) are CIN3 (see Table 2). Diagnostic distributions of the SIL-based grading are shown in Table 3. Similarly, distribution of final diagnosis in SIL-based grading is as follows: 471 of SEP (49.2%) are normal, 240 of SEP (25.1%) are LSIL, and 246 of SEP (25.7%) are HSIL. Similarly, 150 of LEP (46.9%) are normal, 79 of LEP (24.7%) are LSIL, and 91 of LEP (28.4%) are HSIL.

2.4.1 Inter-observer variability

Interpretations of morphologic changes representing dysplasia may differ between physicians or for the same physician in different time intervals. This variety can be interpreted as inter-/intra-observer agreement/disagreement, respectively. Artifacts associated with the biopsy procedure and tangential sections in the microscopic examination are also effective on this variety. Inter- and intra-observer agreement rates are in the range of 0.20 and 0.47 in the literature [21, 22]. Regarding CIN-based grading, SIL-based grading provides higher inter-observer and intra-observer agreement rates. The highest diagnosis diversity is reported between the groups of CIN2, while the lowest is CIN3. The disagreement rates are smaller between normal and CIN1 groups. McCluggage et al. reported weak inter-observer agreement in the CIN-based grading with Kappa

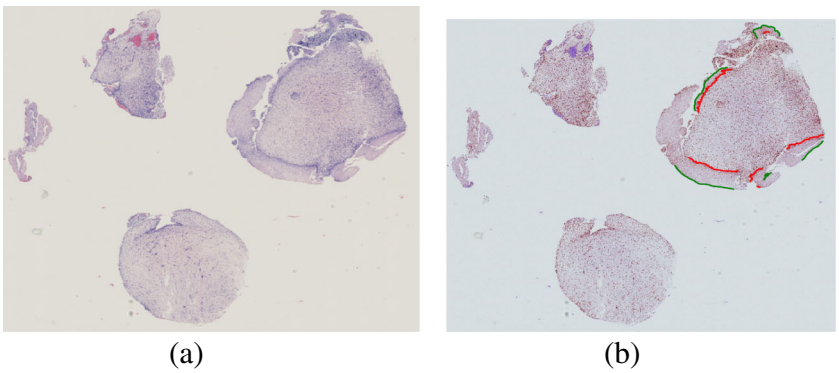


Fig. 4 Annotation and hotspot region extraction. **a** Input image obtained from the dataset; **b** extracted hotspot cervix region for further analysis. The red and green curves drawn around the lesion represent the BM and the upper membrane (UM), respectively

Table 2 Number of epithelium pieces in each class depending on the CIN-based grading

		Normal	CIN1	CIN2	CIN3	Total
CIN-Based Grading	SEP	471	240	107	139	957
	LEP	150	79	34	57	320

value of 0.2. Although the compatibility rates are reported low, Kappa value is found as 0.3 in the SIL-based grading. Failure to achieve the expected high agreement rates is interpreted by the pathologists involved in the study not being familiar with SIL-based grading. The same test is repeated with the observers who have experience in use of the SIL-based grading for 6 months or more and new Kappa values are calculated as 0.33 (intra-observer) and 0.47 (inter-observer). Galgano et al. tried to maximize the agreement rates between the observers with p16 and Ki67 immunohistochemical methods [9]. The Kappa value is found to be 0.68 by immunohistochemistry examination while standard H&E detection has the Kappa value of 0.47. In the study, it is stated that the low agreement rates associated with diagnostic differences can be increased by using SIL-based grading rather CIN-based grading, or utilizing some immunohistochemical methods aiding diagnosis.

2.4.2 Final diagnosis

Immunohistochemical examinations are used as an assistive method to obtain the diagnosis in case the morphological features are not clearly interpreted. p16, Ki67, and ProExC are the most widely used immunohistochemical studies for cervical precursor lesions [9, 10, 17, 23, 29]. Staining pattern with p16 is important in immunohistochemical evaluation; block-like and strong staining demonstrates HrHPV associaticon, with at least one-third of the epithelium. Ki67 is an indicator of proliferation. Positivity may also be seen in other proliferating cells, such as inflammatory cells as in keratinocytes. For this reason, it must be interpreted carefully in the presence of inflammation. ProExC is similar to Ki67 in terms of being a proliferation indication and its staining type. p16 and Ki67 are frequently used in routine practice. HrHPV-associated lesions show strong "nuclear" or "nuclear and cytoplasmic," block-like staining with p16. The squamous metaplasia, atrophy, and reactive regenerative changes that appear in the SIL discriminator pattern show a negative staining pattern. While Ki67 normally stains parabasal cells, positivity is also observed in higher epithelial sections in relation to the grade of dysplasia in SIL.

2.5 Morphometric feature extraction and tissue classification

In this study, a morphological analysis–based feature extraction method is used for the grading of cervical cancer precursor lesions. The processing steps followed in the study are represented in Fig. 1. The first row of the diagram describes the pathological pre-processes which are handled in the pathology laboratory. Whole-slide scanning and the filing process mentioned in the second row are done by the medical researcher and the computer scientist in collaboration. This section describes the CADAS framework which is mentioned in the third row.

2.5.1 Creating the small epithelial pieces

Small epithelium piece is the part of a large lesion which is cropped by a specialist for further analysis. Similar to the whole lesion, the membrane boundaries of the small epithelial piece are also annotated by the specialist. In Fig. 5, the red curve corresponds to the pixel locations of the basal membrane (BM), while the green curve corresponds to the pixel locations in the upper membrane (UM). Determining the BM and UM pixel locations allows us to localize cells in the epithelium. Figure 5a represents a high-resolution histopathological image obtained from the dataset; Fig. 5b and c represent SEP images cropped from the related image. After the region of interest has been obtained, an interface developed within the scope of the study is used to divide the whole epithelium to SEP which can be assumed equal in length (see Fig. 6).

The image patches which are analyzed in this study are presented in Fig. 6. The pixel locations of BM and UM of the epithelium are marked by the pathologists with the use of a graphical interface. Pixel locations data information of the papillae represented with yellow curves are also stored in separate files.

Table 3 Number of epithelium pieces in each class depending on the SIL-based grading

		Normal	LSIL	HSIL	Total
SIL-based Grading	SEP	471	240	246	957
	LEP	150	79	91	320

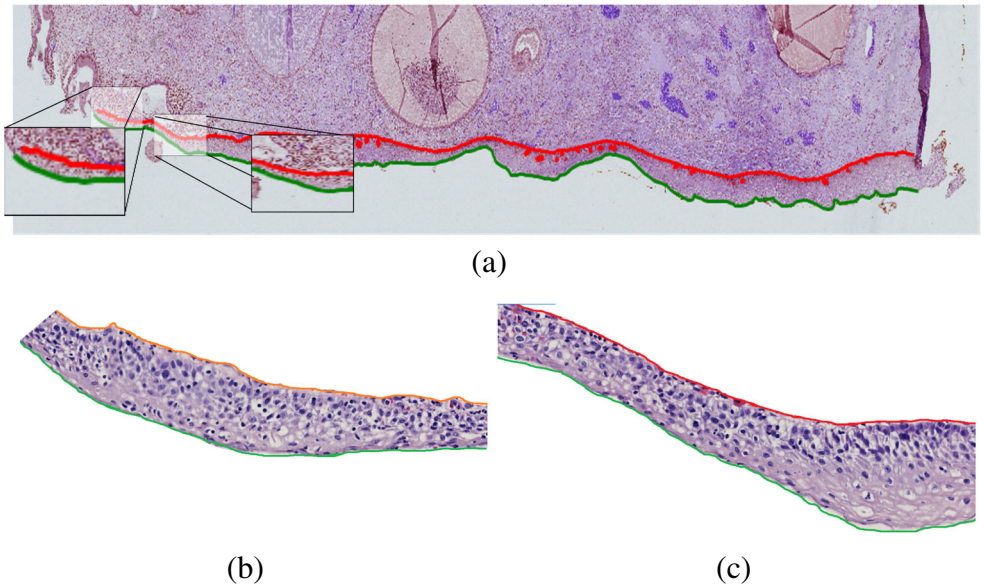


Fig. 5 **a** A high-resolution histopathological image example obtained from the dataset. **b** and **c** represent SEP images cropped from the same image

2.5.2 Obtaining cells by simple linear iterative clustering superpixels segmentation algorithm

After small epithelial pieces are obtained, the high-resolution histopathological images are ready for further analysis. First, a median filter of dimension 9×9 is applied to the image to remove the artifacts without affecting the boundaries. Then, cellular structures have been obtained by simple linear iterative clustering (SLIC) superpixels segmentation algorithm, which is one of the methods that have not been widely used in histopathological images yet. This method performs the segmentation process based on the color similarities and neighbor relations of the pixels in the image [1]. The grid size N is expressed as

$$N = \sqrt{\frac{w \times h}{k}} \tag{1}$$

where k is the number of superpixels for a given input image; w and h represent the width and height of the given image patch, respectively. The Euclidean distance, d_{rgb}, of the related pixel to the superpixel center is

$$d_{rgb} = \sqrt{(r_j - r_i)^2 + (g_j - g_i)^2 + (b_j - b_i)^2} \tag{2}$$

where i represents each pixel and j represents the center of related superpixel. Here, r, g, and b represent the brightness values of red, green, and blue colors of the respective pixels. RGB color space is used in this study instead of using Lab color space as mentioned in Achanta et al. [1]. The distance between the locations of each related pixel and related superpixel is calculated as mentioned in Eq. 3:

$$d_{xy} = \sqrt{(x_j - x_i)^2 + (y_j - y_i)^2} \tag{3}$$

where, x_j and y_j are the horizontal and vertical pixel locations of each center pixel, and x_i and y_i values are the locations of each pixel to be clustered:

$$d_s = d_{rgb} + m/(N \times d_{xy}) \tag{4}$$

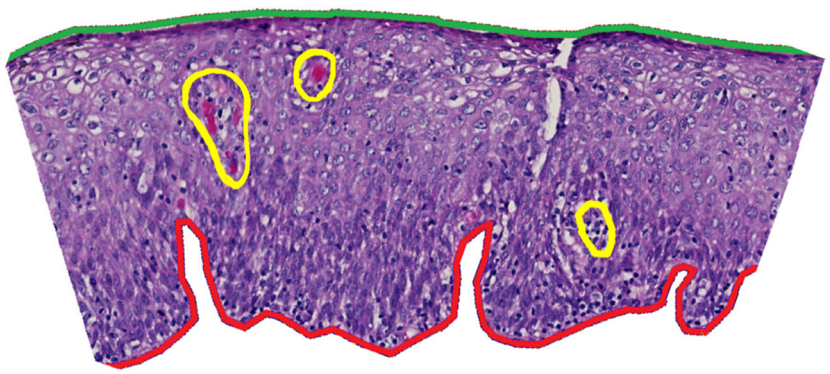


Fig. 6 Sample SEP image patch. BM (red), UM (green), and papillae (yellow) structures are considered by pathologists for grading

the value of d_s is the sum of the (x, y) plane distance normalized by the grid interval N and the RGB distance. Here, normalization is done so that the calculation of the pixel location does not directly affect the brightness interval. The value of m is defined to set the compactness of superpixels (Fig. 7).

Algorithm 1 Center estimation and distribution generation.

Input $\mathbf{I}_{dist} \in \mathbb{R}^{N\times M}$ ▷ Distance transform of binary image
Output $\mathbf{C} \in \mathbb{R}^{K\times 2}$ ▷ Cell center matrix

procedure LOCAL MAXIMA FINDING
  for $\forall(i, j) \in \mathbf{I}_{dist}$ **do**
    $\mathbb{S} = \{(u, v)|(i-u)^2 + (j-v)^2 \leq r^2\}$
    if $\arg\max\{\mathbf{I}_{dist}(\mathbb{S})\} == (i, j)$ **then**
      $\mathbf{C}_{k,:} \leftarrow [i, j]$
  return $\mathbf{C}$

Input $\mathbf{I}_{dist} \in \mathbb{R}^{N\times M}$
Output $\mathbb{H}$ ▷ Multiplexed Coordinates Set

procedure TEST C3
  $\mathbb{H} \leftarrow \emptyset$
  for $\forall(i, j) \in \mathbf{I}_{dist}$ **do**
    $\mathbb{H} \leftarrow \mathbb{H} \cup \alpha \otimes [i, j]$ ▷ $\otimes$: multiplexing operator
  return $\mathbb{H}$

According to the SLIC method applied in this study, the cellular structures become more compact and can be separated from the background when each obtained SEP image is expressed with 3000 superpixels. A cellular structure in cervical precursor lesion is approximately 20 × 20. A crucial point to note here is that the superpixels' sizes should not exceed the size of the cellular structure. As can be calculated from Eq. 1, choosing at least 2000 superpixels will guarantee most of the superpixels do not exceed the size of a cellular structure. A smaller number of superpixels cause overlapping cells. It can be quite difficult to distinguish cellular structures, especially those close to the BM. Superpixels which represent the cellular structures are darker than the superpixels which represent the background (fat-like tissues). To obtain the cellular structures, superpixels which have brightness above a certain threshold value are eliminated. At this stage, small artifacts similar to the cellular structures and some inflammation can remain with the cells as a foreground information. These structures can be eliminated with a morphological size operation that can be applied to the binary image after segmentation stage. Final segmentation result of a sample SEP image patch obtained from the dataset is shown in Fig. 8.

2.5.3 Handling the overlapping cells problem

Following the morphological operations, overlapping cells are separated. There are studies targeting the problem of overlapping nuclei in the literature [3, 31]. Because the presence of overlapping cell structures significantly reduces the success of CADAS, solving the problem of overlapping nucleus at this point is very crucial as a significant contribution in this area. In our study, it is observed that after the segmentation process, there are a large number of overlapping nucleus structures, especially around the BM.

Overlapped cells are intensively present on the SEP image patches. The problem of overlapping of cells should be handled in order to obtain the morphological characteristics of the cell nuclei. Binary images segmented by using SLIC algorithm usually consist of small cellular-like noisy parts. These unwanted small pixel groups are

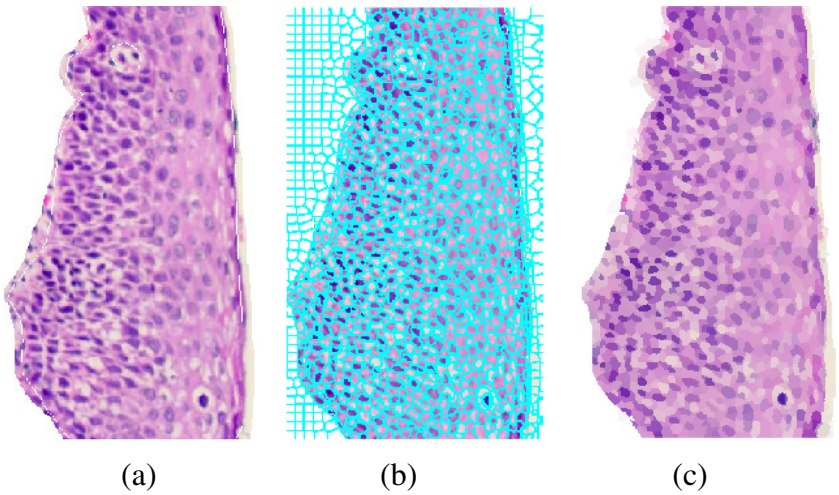


Fig. 7 Implementation of SLIC superpixels segmentation algorithm to a sample image patch **a** SEP image patch, **b** overlay of 3000 superpixels on the related image patch, **c** resulting pre-segmented image obtained after applying SLIC method

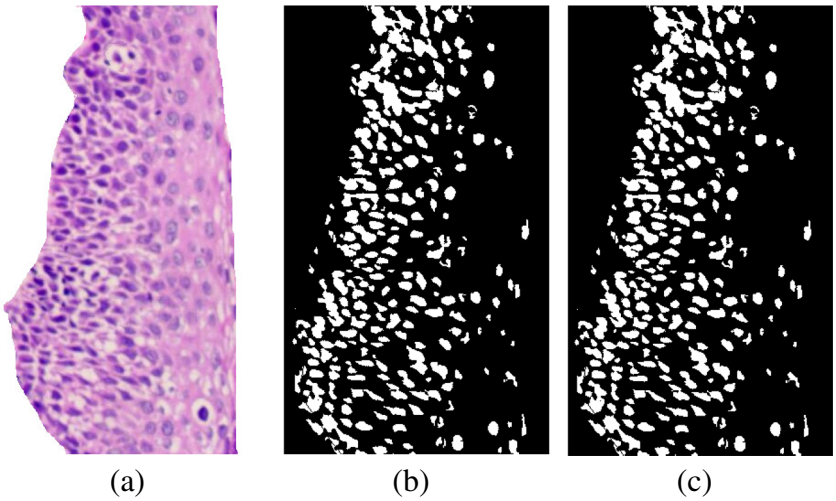


Fig. 8 Final segmentation result of a sample SEP image patch: **a** given input image, **b** segmentation of the image, **c** final binary image after post-processing

eliminated by an automatic method, which clears pixel groups smaller than 50 pixels. Therefore, the circumference of the cells is also quite rough after the segmentation process. As shown in Fig. 9a, the binary image processing operation "closing" has been applied to make the binary regions more compact. In order to obtain cell centers, the distance transform is applied to get local maxima as shown in Fig. 9c. Local maxima are estimated by using Algorithm 1.

The ellipse form is able to model the cell shapes mathematically well. Thereby, Gaussian Mixture Model (GMM) is one of the best candidates for ellipse fitting over cell heaps. GMM is one of the most common algorithms for statistical data modeling [27]. The main purpose of the algorithm is to express the distributions, p(x), as the sum of weighted Gaussian distributions.

$$p(\mathbf{x}) = \sum_{i=1}^{K} \phi_i \mathcal{N}(\mathbf{x}|\mu_i, \Sigma_i) \tag{5}$$

where $\mathcal{N}(\mathbf{x}|\mu_i, \Sigma_i)$ intends normal distribution which has mean μ_i and covariance matrix Σ_i.

$$\sum_{i=1}^{K} \phi_i = 1 \tag{6}$$

where ϕ_i represents the coefficients of normal distributions.

Figure 10 shows an example of overlapping nucleus and how these overlaps are resolved in a small patch of the image obtained from the dataset. The basic structure of the algorithm that determines the cell overlapping is based on the determination of local maxima from the distance of the cell centers to the boundaries. Once the cell centers are determined, the distance transformation yields the value of the multiplexing for each pixel. The processing steps applied for center estimation, distribution generation, and multiplexing are given in Algorithm 1. $\mathbf{C}$ and $\mathbb{H}$ refer to row-wise cell centers matrix, and multiplexed coordinate set, respectively.The α parameter obtained from the distance transform indicates how many times the corresponding coordinate will be repeated in the set $\mathbb{H}$. Thereby, the distance values of the pixels away from the border are higher, so the amount of these in $\mathbb{H}$ will be more. Pixel distribution becomes more suitable for GMM. For example, if each pixel is far away from a boundary, the location of the related pixel is multiplexed. By applying GMM on the obtained multiplexed coordinate distribution, suitable ellipses are obtained for each cell (see Fig. 11).

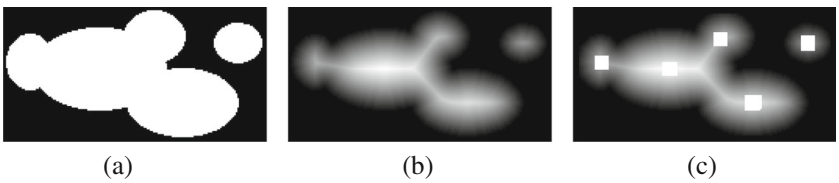


Fig. 9 Cells taken from the tissues are in often overlapped form. For the solution of this problem, it is important that the cell centers are firstly correctly estimated: **a** binary mask of overlapping cell heaps, **b** distance transform of binary mask, **c** finding local maxima

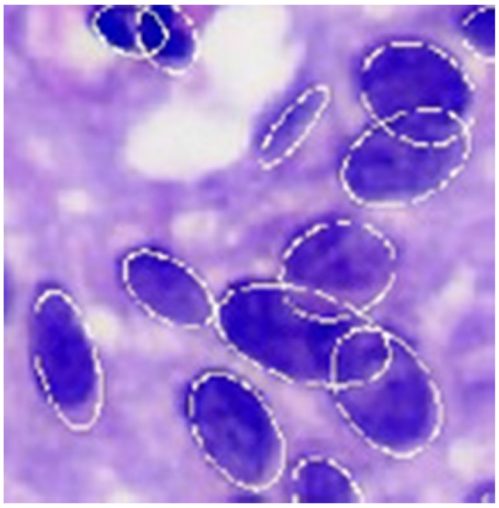


Fig. 10 Separating overlapping cells after the segmentation process

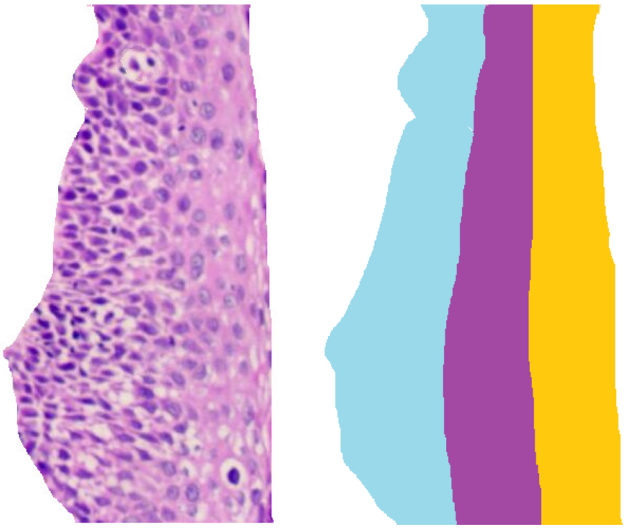


Fig. 12 **a** Sample image obtained from the dataset and **b** layered result of the same image into three sections from BM to UM. The distance of each pixel to the BM and UM is calculated using the pixel location. The distance from the pixel coordinate to the each membrane indicates the region it belongs to

2.5.4 Obtaining the morphological features of each SEP

After the cell segmentation and elimination of the cell overlap problem, several morphological features of each cell are extracted. The morphological features extracted for grading each SEP image patch are represented in Table 4. Average nucleus area (ANA), average cytoplasm area (ACA), nucleus–cytoplasm area ratio (NCR), nucleus perimeter (NP), border irregularity (BI), hyperkromasis index (HI), and polarity loss index (PLI) are the features represented from the first row to the end, respectively. ANA defines the average nucleus area while ACA is the average cytoplasm area. NCR describes the division results of nucleus ratio to the cytoplasm ratio. NP is the average length of the border pixels of nucleus. BI is the ratio of the surrounding length of each pixel to the ellipse that fits each nucleus. The value which represents the hyperkromasis of cell is calculated by taking the standard deviation of the pixel intensity values of the cellular structure. PLI is determined by calculating the magnitude of each cellular structure to the BM. All the features are extracted for each region shown in Fig. 12

Since the emphasized morphological features change depending on their distance to the BM and the UM, each image segment is divided into three main regions as presented in Fig. 12. Then, morphological features related to each region are extracted and stored for further analysis in grading the SEP image patch.

The dataset of feature vectors is imbalanced. Different classifiers have been proposed in the literature for imbalanced datasets [13, 20, 38]. The Weighted k-Nearest Neighbor (w-kNN) algorithm is one of these. In this study, w-kNN algorithm is preferred because of its fast operation and practical use [18, 39]. The w-kNN algorithm takes into account the k closest neighbors class similar with the k-NN algorithm. Differently, for each neighbor, the weight w defined in $w = \frac{1}{d(\mathbf{x}_i,\mathbf{x}_q)^2}$ is assigned to classify according to the weight of the classes. $d(.,.)$ is Euclidean distance function. If neighbor sample $\mathbf{x}_i$ is far from query sample $\mathbf{x}_q$, the effect on the classification is weak, and vice versa.

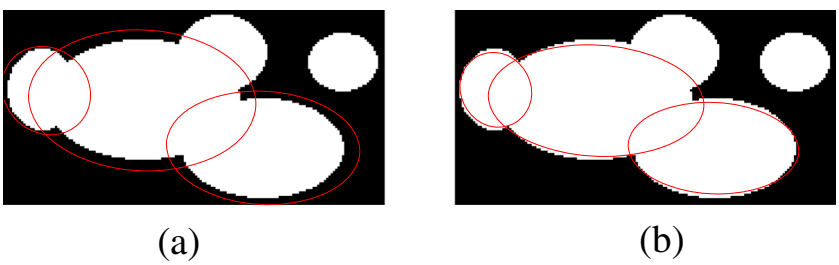


Fig. 11 Generation of ellipses based on uniform and normal distribution after the estimation of center locations: The cell population modeled as the normal distribution is more suitable for the GMM algorithm than the uniform distribution

Table 4 List of morphological features extracted in the proposed tissue classification method

Features	Description
Average nucleus area (ANA)	the average nucleus area of each region
Average cytoplasm area (ACA)	the region which represents the subtraction of total nucleus area from total area of each region
Nucleus-cytoplasm ratio (NCR)	the ratio of total nucleus area to the total the cytoplasm area in each region
Nucleus perimeter (NP)	the length of the curve which surrounds the nucleus in each region
Border irregularity (BI)	obtained by dividing length of the uniform ellipses that fit the nucleus to the circumference of the related nucleus
Hyperkromosis index (HI)	represents the standard deviation value of the parabasal cells with respect to the cells of the same lesion
Polarity lossindex (PI)	The average angle between the BM and the major axis of all nucleus

Table 5 Agreement/disagreement of the pathologists in diagnosis of SEP with respect to CIN-based grading

		Pathologist 2				
		Normal	CIN1	CIN2	CIN3	Total
Pathologist 1	Normal	354	37	0	0	391
	CIN1	88	156	8	0	252
	CIN2	13	52	71	15	151
	CIN3	4	9	22	128	163
	Total	459	254	101	143	957

Table 6 Agreement/disagreement of the Pathologist 1, Pathologist 2 and the proposed method with final diagnosis in diagnosing of SEP with respect to CIN-based grading

CIN Grading		Pathologist 1				Pathologist 2				Proposed			
		N	C1	C2	C3	N	C1	C2	C3	N	C1	C2	C3
Final Diagnosis	N	**377**	74	18	2	**424**	44	3	0	**386**	61	17	7
	C1	13	**176**	37	14	26	**193**	16	5	116	**91**	19	14
	C2	0	2	**87**	18	8	17	**79**	3	13	30	**47**	17
	C3	1	0	9	**129**	1	0	3	**135**	4	15	18	**102**
	Tot.	391	252	151	163	459	254	101	143	519	197	101	140

Table 7 Agreement/disagreement of the proposed method and Delaunay Triangulation (DT) [15] with final diagnosis in the diagnosis of SEP with respect to CIN-based grading

CIN Grading		Proposed				DT			
		N	**C1**	**C2**	**C3**	**N**	**C1**	**C2**	**C3**
Final Diagnosis	N	**386**	61	17	7	**471**	0	0	0
	C1	116	**91**	19	14	240	**0**	0	0
	C2	13	30	**47**	17	107	0	**0**	0
	C3	4	15	18	**102**	139	0	0	**0**
	Total	519	197	101	140	0	0	0	0

Table 8 Evaluation of pathologists and the proposed method with respect to f-measure, precision and recall metrics according to CIN-based grading system

Classes	Evaluation by	Final Diagnosis		
		f-measure	precision	recall
Normal	Pathologist 1	0.87	0.80	**0.96**
	Pathologist 2	**0.91**	**0.90**	0.92
	Proposed	0.78	0.82	0.74
CIN1	Pathologist 1	0.71	0.73	0.69
	Pathologist 2	**0.78**	**0.80**	**0.76**
	Proposed	0.42	0.38	0.46
CIN2	Pathologist 1	0.67	**0.81**	0.57
	Pathologist 2	**0.75**	0.73	**0.78**
	Proposed	0.45	0.44	0.47
CIN3	Pathologist 1	0.85	0.92	0.79
	Pathologist 2	**0.95**	**0.97**	**0.94**
	Proposed	0.73	0.73	0.73
weighted overall accuracy	Pathologist 1	0.80	0.80	0.82
	Pathologist 2	**0.87**	**0.87**	**0.87**
	Proposed	0.65	0.65	0.64

The weighted overall accuracies are obtained by normalizing the metrics of each class with the number of classes

3 Results

In this section, the classification results obtained by using the proposed study are compared with the diagnoses given by the pathologists. The similarities and differences of the diagnoses of SEP images given by the two pathologists with respect to CIN-based grading are represented in Table 5. Diagonal values refer to the number of images which have the same diagnosis of two pathologists. The agreement ratio of the pathologists with respect to CIN-based grading is shown in Table 5.

The agreement/disagreement of pathologist 1 and pathologist 2 to the final diagnosis of each SEP image patch are presented in Table 6. Final diagnosis is determined according to the disagreement of pathologists for an SEP image patch. The SEP that are not labeled as the same by the pathologists are then observed from the same tissue stained with p16 and Ki67 immunohistochemical dyes.

Table 6 represents the agreement between the final diagnosis and two pathologists with respect to CIN-based grading system. It can be observed that pathologist 2 has more compatible diagnosis result than pathologist 1 with final diagnosis. However, pathologist 1 has more consistent diagnosis in CIN2 SEP image patches. An important information to be drawn from the table is that the number of windowed classes (labelling CIN1 instead of Normal tissue; Normal or CIN2 instead of CIN1; CIN1 or CIN3 instead of CIN2 and CIN2 instead of CIN3) is high.

The agreement between the final diagnosis with the proposed method and the DT [15] method which is one

Table 9 Agreement/disagreement of the Pathologists with final diagnosis in the diagnosis of SEP with respect to SIL-based grading

SIL Grading		Pathologist1			Pathologist2		
		N	**LSIL**	**HSIL**	**N**	**LSIL**	**HSIL**
Final Diagnosis	N	**377**	74	20	**424**	44	3
	LSIL	13	**176**	51	26	**193**	21
	HSIL	1	2	**243**	9	17	**220**
	Total	391	252	314	459	254	244

Table 10 Agreement/disagreement of the proposed method with final diagnosis in diagnosis of SEP with respect to SIL-based grading

SIL Grading		Proposed System		
		N	**LSIL**	**HSIL**
Final Diagnosis	N	**381**	57	33
	LSIL	106	**91**	43
	HSIL	16	27	**203**
	Total	503	175	279

of the best known algorithms used in diagnosis of cervical cancer grading depending on CIN-based grading system is represented in Table 7.

Agreement between the final diagnosis and the proposed method with respect to CIN-based grading system is represented in Table 7. Normal and CIN3 SEP images are classified accurately. However, classifying CIN1 and CIN2 SEP images is less accurate while comparing with CIN1 and CIN3. The classification accuracy of the proposed method is 65.4%.

Tables 8 and 11 represent the classification performance of two pathologists and the proposed method with respect to CIN-based and SIL-based grading systems when tested on the dataset introduced in this study respectively. The columns of the tables represent f-measure, precision, and recall values of the proposed method and two specialists. Considering the overall evaluations, it can be observed that pathologist 2 has the most successful classification performance both in CIN-based and SIL-based grading systems. Pathologist 1 has the highest recall value in normal tissues. The results suggest that, when compared with the pathologists' performance, there is room for improvement for the proposed method. The data-driven approaches supported with morphological features are expected to yield better classification results for this challenging dataset.

Another system that pathologists pay attention to while diagnosing tissues is the SIL-based grading system. In this system, the CIN2 grade is assumed to be composed of two intermediate levels, namely, levels CIN3-like and CIN1-like. CIN2 lesions which resemble CIN3 and CIN3 are expressed as HSIL; CIN2 which resemble CIN1 and CIN1 are expressed as LSIL. The treatment of precursor lesions of cervical cancer varies according to LSIL and HSIL. The agreement between the final diagnosis and two pathologists with respect to SIL-based grading system is represented in Table 9. Pathologist1 has more accurate results than pathologist 2 in normal and LSIL, while pathologist 2 has a more accurate result in diagnosing HSIL. If the diagnosis agreements of the pathologists according to Tables 6 and 9 are compared, it can be observed that pathologists make a more consistent diagnosis in the SIL-based grading system.

Agreement between the final diagnosis and the proposed method with respect to SIL-based grading system is represented in Table 10. The results obtained from the SIL-based grading system of the proposed method is improved

Table 11 Evaluation of pathologists and the proposed method with respect to f-measure, precision and recall metrics according to SIL-based grading system

Classes	Evaluation by	Final Diagnosis		
		f-measure	precision	recall
Normal	Pathologist 1	0.87	0.80	**0.96**
	Pathologist 2	**0.91**	**0.90**	0.92
	Proposed	0.78	0.80	0.76
LSIL	Pathologist 1	0.71	0.73	0.70
	Pathologist 2	**0.78**	**0.80**	**0.76**
	Proposed	0.44	0.38	0.52
HSIL	Pathologist 1	0.87	**0.99**	0.77
	Pathologist 2	**0.89**	0.89	**0.90**
	Proposed	0.78	0.83	0.73
weighted overall accuracy	Pathologist 1	0.84	0.84	0.86
	Pathologist 2	**0.88**	**0.88**	**0.89**
	Proposed	0.72	0.73	0.71

to 70.5% compared to the CIN-based grading system as can be seen from Table 11.

4 Discussion

In this study, it is aimed to to develop a "Computer Assisted Diagnostic Auxiliary Systems (CADAS)" which will assist pathologists on grading of cervical cancer. Our study on cervical dysplasia has the largest dataset according to the similar studies available in the literature to the best of our knowledge. The size of the images scanned by the high-resolution histopathological image scanner ranges from 7500×7700 to $55,700 \times 165,000$.

It is clear that it is very difficult for a single limited-capacity computer (in terms of processor and RAM) to analyze these images in large sizes and high resolution. Besides, an important point to note is that there may be more than one diagnosis (NORMAL-CIN1, CIN1-CIN2, or CIN2-CIN3) in the different regions of the same lesion. Because of these reasons, dividing the entire epithelium into smaller lesions helps diagnosis and eases processing the lesions in a computer with limited resources. The lesions of each image in the dataset and all related whole-slide images are labelled by two pathologists.

Furthermore, the fact that the diagnoses are given by two pathologists, and the reassessment and determination of the definitive diagnosis during inconsistent cases increased the reliability of the CADAS training set. The developed CADAS promises to be used as an assistant system in the future because of numerical values that are found to be in parallel with the diagnostic parameters used by the pathologists (such as ratio of nucleus to cytoplasm, nucleus boundary irregularity, polarity loss, and hyperchromaticity) and statistically significant. The studies in the literature are mostly designed by engineers and the contribution of pathologists is very limited. For this reason, there are some shortcomings when viewed from the perspective of pathology and clinical approach. In the development of a CADAS to be used in pathology, the presence of pathologists at every step is a necessary requirement.

5 Conclusion

In this paper, we present a new benchmark dataset of cervical cancer precursor lesions, which we make available to the scientific community for grading the cervical intra-epithelial neoplasia. Each image in the dataset is labeled by two pathologists to reveal the inter-observer variability. In case of different diagnoses, p16 and Ki67 immunohistochemical dyes are used to decide a final diagnosis (ground truth). There are also papilla areas that seriously affect the performance of automated methods, which makes this study unique to the best of our knowledge.

A morphological analysis–based feature extraction method is also proposed in the study for the grading of cervical cancer precursor lesions. The result of the study is also compared with each pathologist and the ground truth. The results show that CAD systems could be used as a secondary decision system for experts with some improvement. It is aimed to improve the classification performance of our CAD system by developing up-to-date image processing and machine learning algorithms especially types of deep learning.

Acknowledgements The authors also would like to thank Argenit Company and Istanbul Medipol University Hospital for providing and annotating the whole-slide histopathological images of cervical cancer precursor lesions image dataset. The authors would like to thank the reviewers for all useful and instructive comments on our manuscript.

Funding This work is in part funded by ITU BAP MAB-2020-42314 project and also supported by the Scientific Research Projects Coordination Department, Yildiz Technical University, under Project 2014-04-01-KAP01.

Declarations

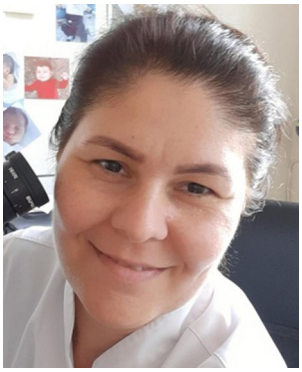
Conflict of interest The authors declare no competing interests.

References

1. Achanta R, Shaji A, Smith K, Lucchi A, Fua P, Süsstrunk S (2012) Slic superpixels compared to state-of-the-art superpixel methods. IEEE Trans Pattern Anal Mach Intell 34(11):2274–2282
2. Al-Janabi S, Huisman A, VDP J (2012) Digital pathology: current status and future perspectives. Histopathology 61(1):1–9
3. Arteta C, Lempitsky V, Noble JA, Zisserman A (2012) Learning to detect cells using non-overlapping extremal regions. In: International conference on medical image computing and computer-assisted intervention, Springer, pp 348–356
4. Cox JT, Wilkinson EJ, O'connor DM (2013) Historical perspective: terminology for lower anogenital tract pathology. AJSP: Rev Reports 18(4):158–167
5. Darragh TM, Colgan TJ, Cox JT, Heller DS, Henry MR, Luff RD, McCalmont T, Nayar R, Palefsky JM, Stoler MH et al (2013) The lower anogenital squamous terminology standardization project for hpv-associated lesions: background and consensus recommendations from the college of american pathologists and the american society for colposcopy and cervical pathology. Int J Gynecol Pathol 32(1):76–115
6. Darragh TM et al (2012) The lower anogenital squamous terminology standardization project for hpv-associated lesions: background and consensus recommendations from the college of american pathologists and the american society for colposcopy and cervical pathology. Arch Path Lab Med 136(10):1266–1297
7. De S, Stanley RJ, Lu C, Long R, Antani S, Thoma G, Zuna R (2013) A fusion-based approach for uterine cervical cancer histology image classification. Comput Med Imaging Graph 37(7):475–487
8. Doorbar J (2007) Papillomavirus life cycle organization and biomarker selection. Dis Markers 23(4):297–313

9. Galgano MT, Castle PE, Atkins KA, Brix WK, Nassau SR, Stoler MH (2010) Using biomarkers as objective standards in the diagnosis of cervical biopsies. Am J Surg Path 34(8):1077
10. Guo M, Baruch A, Silva E, Jan Y, Lin E, Sneige N, Deavers MT (2011) Efficacy of p16 and proexc immunostaining in the detection of high-grade cervical intraepithelial neoplasia and cervical carcinoma. Am J Clin Pathol 135(2):212–220
11. Guo P, Banerjee K, Stanley RJ, Long R, Antani S, Thoma G, Zuna R, Frazier SR, Moss RH, Stoecker WV (2016) Nuclei-based features for uterine cervical cancer histology image analysis with fusion-based classification. IEEE J Biomed Health Inform 20(6):1595–1607
12. He L, Long LR, Antani S, Thoma GR (2012) Histology image analysis for carcinoma detection and grading. Comput Methods Programs Biomed 107(3):538–556
13. Hong X, Chen S, Harris CJ (2007) A kernel-based two-class classifier for imbalanced data sets. IEEE Trans Neur Netw 18(1):28–41
14. Jordan J, Arbyn M, Martin-Hirsch P, Schenck U, Baldauf JJ, Da Silva D, Anttila A, Nieminen P, Prendiville W (2008) European guidelines for quality assurance in cervical cancer screening: recommendations for clinical management of abnormal cervical cytology, part 1. Cytopathology 19(6):342–354
15. Keenan S, Diamond J, McCluggage G, Bharucha H, Thompson D, Bartels PH, Hamilton PW (2000) An automated machine vision system for the histological grading of cervical intraepithelial neoplasia (cin). J Pathol 192(3):351–362
16. Kurman RJ (2013) Blaustein's pathology of the female genital tract. Springer Science & Business Media, Berlin
17. Lim S, Lee M, Cho I, Hong R, Lim S (2016) Efficacy of p16 and ki-67 immunostaining in the detection of squamous intraepithelial lesions in a high-risk hpv group. Oncol Lett 11(2):1447–1452
18. Liu W, Chawla S (2011) Class confidence weighted knn algorithms for imbalanced data sets. In: Pacific-asia conference on knowledge discovery and data mining, Springer, pp 345–356
19. Madabhushi A, Lee G (2016) Image analysis and machine learning in digital pathology: Challenges and opportunities. Med Image Anal 33:170–175
20. Mazurowski MA, Habas PA, Zurada JM, Lo JY, Baker JA, Tourassi GD (2008) Training neural network classifiers for medical decision making: The effects of imbalanced datasets on classification performance. Neural Netw 21(2-3):427–436
21. McCluggage W, Bharucha H, Caughley L, Date A, Hamilton P, Thornton C, Walsh M (1996) Interobserver variation in the reporting of cervical colposcopic biopsy specimens: Comparison of grading systems. J Clin Path 49(10):833–835
22. McCluggage W, Walsh M, Thornton C, Hamilton P, Caughley L, Bharucha H et al (1998) Inter-and intra-observer variation in the histopathological reporting of cervical squamous in traepithelial lesion susing a modified bethesda grading system. Br J Obstet Gynaecol 105(2):206–210
23. de Melo F, Lancellotti C, da Silva M (2016) 16 and ki-67 and their usefulness in the diagnosis of cervicalexpression of the immunohistochemical markers p intraepithelial neoplasms. Rev Bras Ginicol Obstet 38(2):82–87
24. Miranda GHB, Soares EG, Barrera J, Felipe JC (2012) Method to support diagnosis of cervical intraepithelial neoplasia (cin) based on structural analysis of histological images. In: 2012 25th IEEE International symposium on computer-based medical systems (CBMS) IEEE, pp 1–6
25. Mitra A, MacIntyre D, Lee Y, Smith A, Marchesi JR, Lehne B, Bhatia R, Lyons D, Paraskevaidis E, Li J et al (2015) Cervical intraepithelial neoplasia disease progression is associated with increased vaginal microbiome diversity. Sci Rep 5:16,865
26. Naghdy G, Ros MB, Todd C et al (2012) Computer aided decision support system for cervical cancer classification. In: Applications of Digital Image Processing XXXV, SPIE, vol 8499, p 849919
27. Najar F, Bourouis S, Bouguila N, Belghith S (2017) A comparison between different gaussian-based mixture models. In: 2017 IEEE/ACS 14th International conference on computer systems and applications, AICCSA'17, IEEE, pp 704–708
28. Nayar R, Wilbur DC (2015) The Bethesda system for reporting cervical cytology: definitions, criteria, and explanatory notes. Springer, Berlin
29. Ozaki S, Zen Y, Inoue M (2011) Biomarker expression in cervical intraepithelial neoplasia: potential progression predictive factors for low-grade lesions. Hum Pathol 42(7):1007–1012
30. Park KJ, Soslow RA (2009) Current concepts in cervical pathology. Arch Pathol Lab Med 133(5):729–738
31. Qi X, Xing F, Foran DJ, Yang L (2011) Robust segmentation of overlapping cells in histopathology specimens using parallel seed detection and repulsive level set. IEEE Trans Biomed Eng 59(3):754–765
32. Stoler MH (2000a) Advances in cervical screening technology. Mod Pathol 13(3):275–284
33. Stoler MH (2000b) Human papillomaviruses and cervical neoplasia: a model for carcinogenesis. Int J Gynecol Pathol 19(1):16–28
34. Stoler MH, Schiffman M et al (2001) Interobserver reproducibility of cervical cytologic and histologic interpretations: realistic estimates from the ascus-lsil triage study. Jama 285(11):1500–1505
35. Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A (2012) Global cancer statistics. CA Cancer J Clin 65(2):87–108
36. Van Zummeren M, Kremer WW, Leeman A, Bleeker MC, Jenkins D, van de Sandt M, Doorbar J, Heideman DA, Steenbergen RD, Snijders PJ et al (2018) Hpv e4 expression and dna hypermethylation of cadm1, mal, and mir124-2 genes in cervical cancer and precursor lesions. Mod Pathol 1
37. Crookes D, Wang Y, Eldin OS, Hamilton P, Wang S, Diamond J (2009) Assisted diagnosis of cervical intraepithelial neoplasia (cin). IEEE J Sel Top Signal Process 3(1):112–121
38. Zhuang L, Dai H (2006) Parameter optimization of kernel-based one-class classifier on imbalance learning. J Comput 1(7):32–40
39. Zuo W, Lu W, Wang K, Zhang H (2008) Diagnosis of cardiac arrhythmia using kernel difference weighted knn classifier. In: 2008 Computers in cardiology, IEEE, pp 253–256
40. Zur HH (2009) Papillomaviruses in the causation of human cancers—a brief historical account. Virol J 384(2):260–265

Abdulkadir Albayrak is a research assistant in Dept. of Computer Engineering at Dicle University. He studies machine learning, computer vision, and biomedical image engineering.



Asli Unlu Akhan is a specialist in pathology at Selahaddin Eyyubi State Hospital. She studies mainly gynecopathology and dermatopathology, and is interested in digital pathology.

Behcet Ugur Toreyin is a faculty member with the Informatics Institute at Istanbul Technical University. His research interests are in signal processing and pattern recognition.

Nurullah Calik received the Ph.D., B.Sc., M.Sc. degrees in Electronics and Communications Eng. from Yildiz Technical University (YTU) in 2010, 2013 and 2019. He is now with the Institute of Informatics at Istanbul Techinal University (ITU) as a postdoctoral researcher. He is an Assistant Professor in the Dept. of Biomedical Engineering at Istanbul Medeniyet University.

Bahar Muezzinoglu is a Professor in Dept. of Pathology at Istanbul Medipol University Hospital. Her interest areas are gynecopathology, uropathology and bond and soft tissue pathology.

Abdulkerim Capar is a faculty member with the Informatics Institute at Istanbul Technical University. He mainly studies image/video processing, computer vision, pattern recognition, and machine learning.

Ilknur Turkmen is a Professor in Pathology Center of Memorial Healthcare Group. She studies mainly gynecopathology, dermatopathology, and gastrointestinal pathology, and is interested in digital pathology.

Gokhan Bilgin is an Associate Professor in Dept. of Computer Engineering at Yildiz Technical University. He studies machine learning, image processing, biomedical engineering, and remote sensing.

Lutfiye Durak-Ata is a faculty member with the Informatics Institute of Istanbul Technical University, Istanbul, Turkey. Her research interests are in adaptive and statistical signal processing, and communications theory.