IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received March 16, 2018, accepted May 5, 2018, date of publication May 22, 2018, date of current version July 6, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2839338

Automated Methods for the Decision Support of Cervical Cancer Screening Using Digital Colposcopies

KELWIN FERNANDES[1,2], JAIME S. CARDOSO[1,2], (Senior Member, IEEE), AND JESSICA FERNANDES[3]

[1]Centre for Telecommunications and Multimedia, Instituto de Engenharia de Sistemas e Computadores Tecnologia e Ciência, 4200-072 Porto, Portugal
[2]Departamento de Engenharia EletrotÃľcnica e de Computadores, Faculdade de Engenharia, Universidade do Porto, 4200-465 Porto, Portugal
[3]Facultad de Medicina, Universidad Central de Venezuela, Caracas 1053, Venezuela

Corresponding author: Kelwin Fernandes (kafc@inesctec.pt)

This work was supported in part by the Project NanoSTIMA: Macro-to-Nano Human Sensing: Towards Integrated Multimodal Health Monitoring and Analytics through the North Portugal Regional Operational Program (NORTE 2020), under the PORTUGAL 2020 Partnership Agreement, under Grant NORTE-01-0145-FEDER-000016, in part by the European Regional Development Fund, and in part by Fundação para a Ciência e a Tecnologia within Ph.D. under Grant SFRH/BD/93012/2013.

ABSTRACT Cervical cancer remains a significant cause of mortality in low-income countries. However, it can often be cured by removing the affected tissues when detected in early stages. Therefore, it is relevant to provide universal and efficient access to cervical screening programs, being digital colposcopy an inexpensive technique with high potential of scalability. The development of computer-aided diagnosis systems for the automated processing of digital colposcopies has gained the attention of the computer vision and machine learning communities in the last decade, giving origin to a wide diversity of tasks and computational solutions. However, there is a lack of a unified framework to discuss the main tasks and to assess their performance. Thus, in this paper, we studied the core research lines surrounding the automated analysis of digital colposcopies and built a topology of problems and techniques, including their key properties, advantages, and limitations. Also, we discussed the open challenges in the area and released a database that serves as a common basis to evaluate such systems.

INDEX TERMS Cervical cancer, digital colposcopy, computer aided diagnosis, machine learning, computer vision.

I. INTRODUCTION

Cervical cancer remains a significant cause of mortality in low-income countries [54]. Despite the possibility of prevention with regular cytological screening, cervical cancer is the cause of more than 500,000 cases per year, and kills more than 250,000 patients in the same period, on world basis [21]. Cervical cancer can be prevented by means of the human papillomavirus infection (HPV) vaccine, and regular low-cost screening programs (e.g., cytology, digital colposcopy) [28]. Furthermore, cervical cancer can often be cured by removing the affected tissues when identified in early stages [21], [28]. The development of cervical cancer is usually slow and preceded by changes in the cervix (dysplasia). Despite the presence of symptoms on its later stages (e.g., postcoital bleeding, bleeding between periods, increased vaginal discharge, and pelvic pain), the absence of early-stage symptoms might incur in carelessness prevention. Additionally, in developing countries, resources to perform screening programs with universal access are scarce and insufficient. Also, patients usually have poor adherence to routine screening due to low problem awareness.

While improving the resection of lesions in the first visits has a direct impact on patients that attend the screening programs, the most vulnerable populations have difficult access to such programs' information and medical centers. Consequently, the individual risk estimation has a key role in this context to optimize the efficacy of these programs. Identifying patients with the highest risk of developing cervical cancer can improve the targeting efficacy of cervical cancer screening programs. Thus, recent attempts to address the predictive analysis of this problem have been proposed [22], including a competition sponsored by Genentech and Symphony Health Solutions [50].

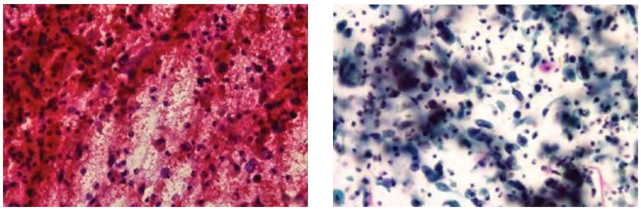


FIGURE 1. Samples of cytological screening [98]. Left: conventional cytology. Right: liquid based cytology.

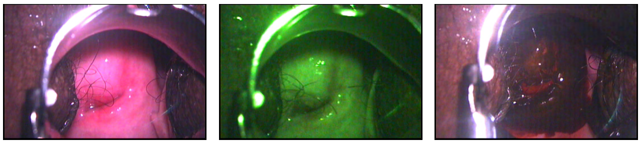


FIGURE 2. Modalities of the colposcopy examination. From left to right: Hinselmann, Green-filter, Schiller.

During the cervical cancer examination, cervical cancer screening programs cover the following stages:

- Cytology, either conventional or liquid (see Figure 1).
- Colposcopy, covering several modalities (see Figure 2).
- Biopsy.

These stages are often done in a cascade fashion, by moving towards the succeeding steps with the discovery of relevant indicators on the preliminary ones. Both cytology and colposcopy are image-based screening processes. The former focuses on the examination of vaginal and cervical cells under the microscope and the latter on the macroscopic examination with the naked eye (or with a magnifier lens).

The conventional cytological screening involves manual smearing and staining [10]. The complexity of the acquisition process for conventional cytology requires mobilizing expert teams to the field. Even when the acquisition is properly done, the uneven distribution of cells may induce dense regions where light cannot penetrate and empty regions of the slide [10]. Other artifacts such as blood may harm the effectiveness of this screening modality. In order to overcome these difficulties, liquid-based cytology (LBC) preparations have been delved. Liquid preparations help to uniformize the distribution of cells and to dilute the presence of external factors. Some common techniques for the preparation of LBC can be found in [84] and [104]. However, the increase of costs (e.g., about 5 to 10 times higher [10]) and technical difficulties to make these equipment available in remote locations appease the use of this technique in low-income countries.

On the other hand, digital colposcopy is a low-cost technology that complements cytology during screening and triage. Nowadays, portable and mobile devices have been introduced in the market as an alternative to traditional colposcopes [59], [60], [78], facilitating its scalability and portability to locations with vulnerable populations. The main drawback of the digital colposcopy is the high sensitivity variability when carried out by experts with different levels of expertise. Plenty efforts have been devoted in the last two decades to automate the analysis of colposcopy images to support the medical decision process and to provide a data-driven channel for communication of findings. These efforts aim to objectify the analysis of this modality.

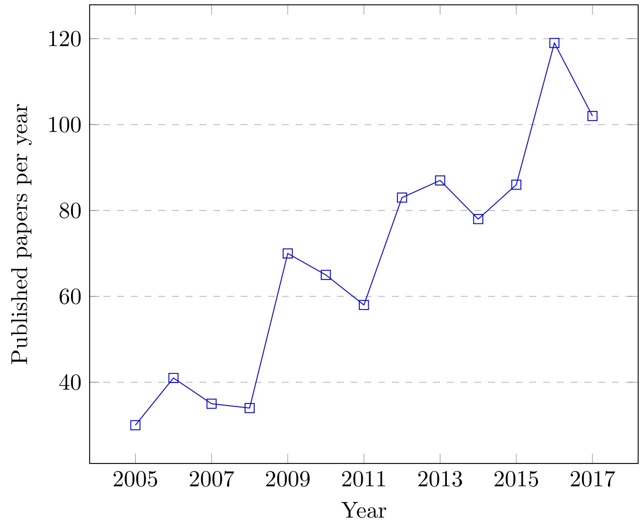


FIGURE 3. Number of papers reported by Google Scholar for the query (`computer vision` OR `image processing` OR `machine learning`) AND (`colposcopy` OR `cervigram`), not including patents nor citations.

The automated analysis of digital colposcopies using Machine Learning and Computer Vision techniques has grown over the last years. Figure 3 shows the number of published papers per year reported by Google Scholar in this area. In addition to the aforementioned competition on the analysis of vulnerable population, Intel and MobileODT organized a competition for the automatic analysis of digital colposcopies in 2017 [51]. This increasing interest has resulted in a stable community with well identified problems that range, from the quality assessment [22] and enhancement [37], [68] of digital colposcopies, to the segmentation of the anatomical parts of the cervix [13], [69], to the final diagnosis [93], [114], [115]. While the vast majority of databases that were used in the development of these papers are closed, as a result of these competitions, new public databases of considerable size and with new challenges were released [44], [51]. We are facing a possible turning point in the area, with the driving interest of governments and companies involved in the area, and the new advent of deep learning techniques that has been permeating all the areas of computer vision.

Therefore, it is relevant at this point to formalize the basis of the area, providing a comparative analysis of the main tasks involved in the area and the solutions that have been proposed in the last years. In this paper, we aim to provide such foundations. Also, we release a database that will be continuously updated with transverse annotations. Finally, we enumerate the open problems and challenges in the area.

The rest of the paper is organized as follows. Section II defines background medical concepts involved in the analysis of digital colposcopies. Section III describes the main tasks

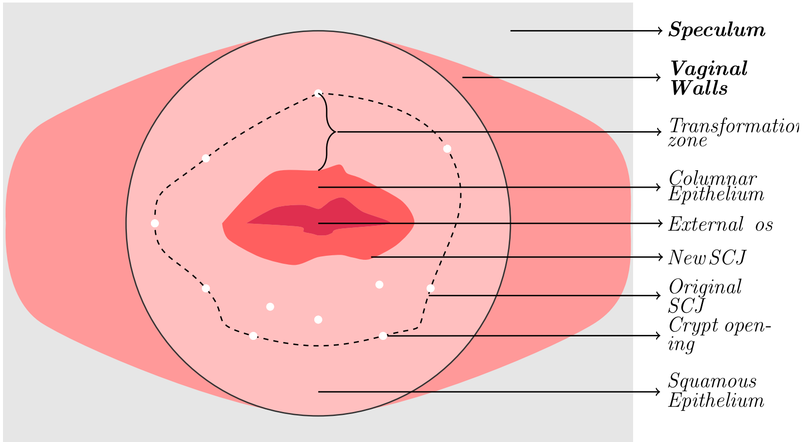


FIGURE 4. **Relevant parts of the Cervix Anatomy and external objects (in bold).**

involved in colposcopic image processing and the solutions that have been proposed in the literature to tackle each one of them. Section V describes the available databases and challenges associated to each one of them. Section VI describes the database and annotations provided in this work. Finally, section VII concludes the work and discuss the main open challenges in the area.

II. PRELIMINARY CONCEPTS

A. CERVIX ANATOMY

The main regions of interest in the analysis of colposcopy images include the external orifice (external os), the area of ectopy, the squamocolumnar junction (SCJ), the transformation zone and the area of squamous epithelium (exocervix). Figure 4 identifies the location of these regions.

Overall, the epithelium are the superficial cells of the cervix. The low environmental aggression in the internal orifice of the cervix makes the cells in that region of columnar type. Thereby, it is relatively easy to observe the vascularity in this region. Conversely, the aggressive environment in the external region, caused by external factors such as the acid pH levels and trauma during intercourse, makes the external cells are of the squamous type. In some cases, the columnar epithelium extends outside the external orifice and gets exposed. Being exposed to external stimuli, columnar epithelium turns into squamous epithelium, giving origin to the transformation zone (see Figure 4). The intersection of these two regions is the SCJ.

B. COLPOSCOPY EXAMINATION

The observation of the cervix following the recommended protocol for digital colposcopies covers four main stages [95].

First, observation of the squamous and columnar epithelium with a magnifier lens is performed after application of a normal saline solution. During this step, the squamous epithelium is observed to define landmarks of the transformation zone. The squamous epithelium is typically smooth with a pink tone. The main landmarks of interest constitute crypt openings and nabothian follicles. These artifacts define the external boundary of the transformation zone. The SQJ defines the inner border. The entire observation of the regions of interest is often unachievable from a single image since the SQJ may recede into the canal as the woman ages. Also, the columnar epithelium is observed at this stage. The common appearance of the columnar epithelium is dark red with complex patterns such as grape-like or sea-anemone tentacles-like or villous appearance [95].

To improve the visualization of the vasculature, a green filter is used on the colposcope to enhance the contrast of the vessels. The two most common vascular patterns observed in the squamous epithelium are reticular and hairpin-shaped capillaries [95]. These patterns are typically found on specific regions of the cervix.

The third stage of the colposcopy examination consists in the observation of the cervix tissues after application of 5% acetic acid solution. This step is known as Hinselmann. In this step, squamous and columnar epithelium should be observed again. The change of appearance of these tissues after the application of acetic acid improves the discriminability of these regions by a human expert. Precancerous lesions can be observed in this phase.

Finally, the physician applies Lugol's iodine solution to the cervix, a step that is known as the Schiller's test. The normal vaginal and cervical squamous epithelium stain and become mahogany brown or black [95], the immature squamous metaplastic epithelium does not stain or partially stain. Some abnormal patterns such as cervical polyps do not stain with iodine [95]. Thereby, the Schiller's test improves the discriminability of normal and abnormal regions in the transformation zone.

Cervical cancer is characterized by the abnormal growth of cells on the cervix. The wide spectrum of abnormal features associated with cervical intraepithelial neoplasia (CIN) may difficult the labor of a medical examiner. The high variance of appearance between women may difficult an objective assessment from unskilled examiners. Thereby, the characterization of these patterns and the identification of abnormal features in each part of the cervix anatomy have a direct impact on the expert decision.

III. MAIN TASKS

The applications surrounding the development of Computer-aided Diagnosis (CAD) systems for digital colposcopies cover a wide spectrum of tasks, from the analysis of the image quality, to the semantic segmentation of the image on its constituents parts, to the final diagnosis of the patients. Thus, computer vision and machine learning researchers have gathered around these tasks in the last decades.

The main source of data for this analysis comes from static color images directly captured from digital colposcopes. However, given that in some cases it is not possible to observe all the structures of the cervix in a single frame as well as its response to the acetic acid solution, some lines of work focused on the analysis of multiple views and even continuous videos.

In this section, we organize the literature into five main areas:

- Quality assessment and enhancement of digital colposcopies (section III-A).
- Segmentation of cervix tissues (section III-B).
- Image Registration (section III-C).
- Detection and characterization of abnormal tissues (section III-D).
- Classification of patient traits (section III-E).

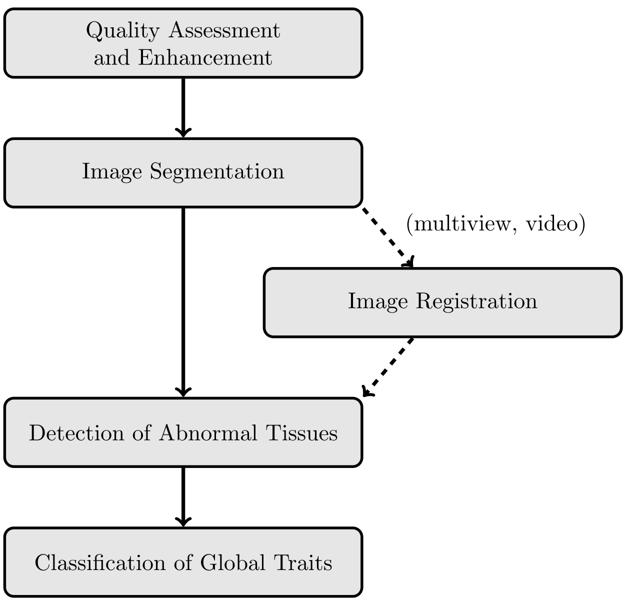


FIGURE 5. Pipeline of the main steps in the development of CAD systems for the automation of digital colposcopy analysis.

These tasks are typically applied in a cascade fashion as illustrated in Figure 5. However, some methods may ignore parts of the pipeline or even include additional dependencies between them. For instance, techniques focusing on static data would ignore the image registration step, and some strategies to address the image quality require to segment the cervix tissues. Thus, this pipeline serves as a general overview of the main tasks but can be adapted to the intrinsic properties of each automation strategy.

In the rest of this section, we do a comparative analysis of the main methodologies that have been applied to each problem, and we discuss their advantages and limitations.

A. QUALITY ASSESSMENT AND ENHANCEMENT

The concept of quality has attained a significant interest in the computer vision research community. Traditional methodologies focus on a low-level notion of quality, measuring distortions of the image at a signal level [29], [53]. In medical imaging, the idea of quality goes beyond low-level aspects of the images to semantic concepts such as visibility of the anatomical body parts, patient's pose, the absence of artifacts, among others.

Therefore, methodologies to address the assessment and enhancement of medical image quality are often application-specific and require extensive domain knowledge. In this section, we cover the main lines of research in this area for colposcopic image processing.

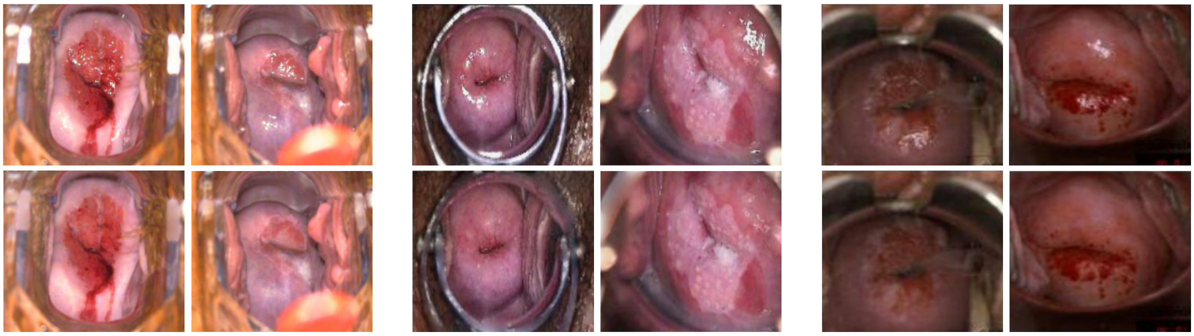
1) QUALITY ASSESSMENT

In the area of quality assessment (QA), Gu and Li [39] proposed a framework to validate the quality of uterine cervical imagery in an online scenario, so the physician may perform corrections to improve the acquisition of data in real time. In [39], the QA problem is modeled as a binary task where the program is required to decide if the image is good enough or not. Six types of issues were handled: zoom, position, foreign objects, contrast, blur, and contamination. These traits were quantified using different models and, using a thresholding operator, it is decided if there are features with low quality. The main disadvantage of this approach is the simplicity of the quality decision model (i.e., thresholding operators). Also, no quantitative assessment of the methodology is presented.

Fernandes *et al.* [22] proposed a learning methodology to tackle this problem. First, several features related to the image quality are extracted, including the area of the main parts of the cervix, the presence of specular reflections, observability of the entire cervix, and color statistics. Then a Support Vector Machine is used to learn the quality decision model on a set of images, covering several modalities (e.g., Hinselmann, Green light, and Schiller) and inter-expert annotations. Fernandes *et al.* proposed a transfer learning approach to improve the robustness of the learning process, where the knowledge acquired from the other modalities and experts is reused when a model for a new modality/physician is learned.

2) QUALITY ENHANCEMENT

Several works have been proposed in the area of quality enhancement of colposcopic images [12], [13], [37], [58], [62], [66], [68], [92], most of them focusing on the removal



(a) (b) (c)

FIGURE 6. **Illustration of the results for specular reflection removal proposed in [13], [37], and [62]. Top: original images. Bottom: corrected images. (a) Lange *et al.* [62]. (b) Das *et al.* [12], [13]. (c) Gordon *et al.* [37].**

of specular reflections (SR) [12], [13], [37], [58], [62], [66]. The remaining works, proposed by Li *et al.* [68] and Rouhbakhsh *et al.* [92] focused on the enhancement of images by means of color and contrast normalization. It is relevant to highlight that image enhancement can be done with two goals in mind, which may lead to different techniques and evaluation settings. First, this process may be done to boost the performance of automatic image processing algorithms. Second, image enhancement can be done for human visualization purposes. This can be done by several means, such as: highlighting relevant patterns of the image that are indistinguishable by the human eye, recovering damaged regions of the image, among others. All the aforementioned papers focused on the improvement for further automatic image analysis.

a: REMOVAL OF SPECULAR REFLECTIONS

Specular reflections or glares raise challenging problems in medical image analysis, as it degrades (partially or entirely) the information in the affected pixels [62]. Moreover, it can introduce artifacts in feature extraction algorithms [62] . The acquisition conditions and involved tissues in the colposcopic assessment are prone to generate this phenomenon.

Lange [62] proposed a method to remove this type of reflection using the green channel of the RGB color space, which classifies the types of glares that can be found in these images in two categories: large saturated regions (detected with adaptive thresholds), and small high contrast regions (detected with morphological operators and thresholding). Once these regions are identified, missing information is filled using interpolation using Laplace's equation and modifying the intensity component of the HSI color space in the transformed image. The method is validated using qualitative subjective inspection. Similarly, Das *et al.* [12], [13] proposed a similar approach to manage SR. First, the affected regions are detected using the intersection of a thresholding operator on the three RGB channels independently. Then, Laplace's equation is used to select the smoothest possible interpolant.

Gordon *et al.* [37] proposed a different approach in both, detection and removal of SR. In the detection subtask, fixed thresholds are used to detect high brightness and low color saturation areas. Then, pixels located in neighborhoods with high gradients are selected as SR candidates. These pixels are mapped to the Saturation-Value space from the HSV color space, and a mixture of two Gaussians is fitted. In the results, one of the Gaussians represent pixels with color information and the other contains merely white pixels. The pixels that belong to the second Gaussian are considered as damaged and are removed from the original image. To fill the damaged regions, a simple inpainting technique that propagates the color of the surrounding pixels is executed. This process is done under the assumption that the color underneath the SR regions is almost constant and similar to the neighboring pixels.

Although none of the aforementioned papers show objective assessment of their methods, visual inspection suggests similar results in all of them. Figure 6 shows sample images presented by each author. The methodology proposed by Das *et al.* shows some undercorrected areas, where residual specular reflections are observable. Also, despite the additional number of manually-defined parameters, the unsupervised learning stage proposed by Gordon *et al.* makes it more adaptable to new settings and datasets.

b: IMAGE NORMALIZATION

Li *et al.* [68] propose a color calibration system to map the color appearance of different colposcopes into one standard color space with normalized illumination. The process involves a preliminary calibration system where the physician presents a target color palette to the colposcope. The main disadvantage of this method is that it should be done before the acquisition of the images, which limits its applicability to already acquired datasets. Also, with the advent of mobile colposcopes, the acquisition conditions can change quickly, requiring continuous calibration.

Other attempts, such as the one proposed by Rouhbakhsh *et al.* [92], perform simple normalization using brightness and contrast equalization.

The actual impact of this step in the final pipeline will depend on the type of assumptions made by the following steps of the automatic analysis. The types of invariance (e.g., pose, illumination, etc.) that can be ensured at this stage will facilitate the job of the following methods. However, as we introduce additional constraints, the applicability of automatic methodologies for the analysis of digital colposcopies will be confined, especially on remote settings with inexperienced staff. In counterpart, a new trend in deep learning to induce robust models is augmenting the data by introducing simulated perturbations (e.g., rotations, flips, contrast stretching, etc.).

B. SEMANTIC IMAGE SEGMENTATION

Most efforts on the line of semantic image segmentation focused on the Hinselmann stage of the colposcopy protocol. Also, it is assumed that specular reflections have been removed from the image either during acquisition or as a preprocessing step.

The main trend in segmentation of the different regions of the cervix focus on the segmentation of the cervix from the outer parts (i.e., vaginal walls and speculum) and the segmentation of the acetowhite regions. Typical methodologies on this line belong to the class of unsupervised methods (e.g., clustering). The most common models are K-means [12], [38], [80], [83], [86], [105], [116], Gaussian Mixture Models [37], [38], [66], [74], [86], [90], [100], [116], [117], [122], and Mean shift [66]. Regarding the feature space, most methodologies use raw color information on different color spaces, being the Lab color space the most widely used [12], [37], [38], [45], [58], [86], [90], [100], [116], [117], [122], [122], followed by RGB [74], [83], CIE Luv [66] and K-L color spaces [66]). Some additional features such as color ratios [83], texture information [37], [74], [122], and spatial information (i.e., distance to the image center) [86], [90], [100], [116], [117], [122] are used.

Clustering algorithms at a pixelwise level do not guarantee spacial consistency of the segmented regions, even when spatial features are considered. Thereby, post-processing step was carried out in these works to decide the final segments that represent the areas of interest. Das *et al.* [12], [13] and Traversi *et al.* [105] use the largest contour as the cervix representative, Gordon *et al.* [37] select the cluster with the lowest mean distance to the image center and highest mean redness level as the cervix region, using size to solve ties. Gu and Li [39] used morphological operators to fill small holes in the final segmentation.

Since the core cervix structures have smooth and almost indistinguishable contours, the performance of these methodologies is limited, not being able to differentiate the cervix from other structures such as the vaginal walls. For instance, Figure 7 shows the results of the method proposed by Das *et al.*, where an oversegmentation of the cervix is done.

In order to counteract oversegmentations of the vaginal walls, some authors applied domain knowledge on the expected shape of the cervix. For instance, some works used active contours [38], [71], [88], [122] solely or as a post-processing technique. Lotenberg *et al.* [71] include shape-priors (e.g., circles and ellipses) to encourage this behavior.

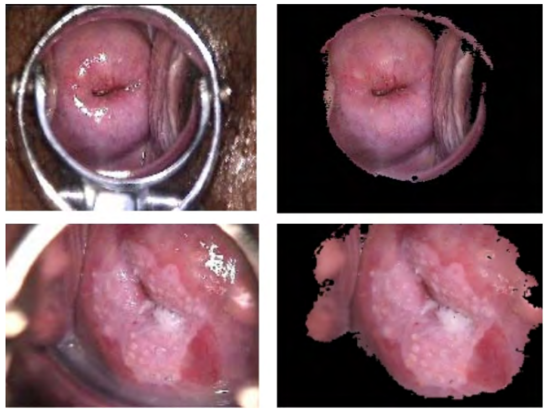


FIGURE 7. Das *et al.* [12] - input images (left), cervix segmentation (right).

Van Raad and Bradley [88] applied iterative multi-scale active contours by sequentially using the previous contour to reduce the initialization impact.

For the segmentation of other regions, such as the columnar epithelium, Gordon *et al.* [37] used a cascade of GMM, where the first level segments the cervix from the background using the redness level of the Lab color space, and the second level segments the columnar epithelium from the rest of the cervix using texture and contrast features. Li *et al.* [66] used a cascade of GMM on the K-L color space and Mean shift on CIE-Luv to segment the cervix from the background and the external orifice from the cervix respectively.

A different unsupervised approach was proposed by Lange [61] based on the watershed algorithm. First, cervix and vagina are segmented using a hue color classifier. Then, the watershed algorithm is applied to detect the low-intensity border around the cervix. Finally, they extract a feature related to the acetowhite response consisting in the product of the green channel in the RGB color space and the saturation value in the HSI color space. The watershed algorithm is applied iteratively on the gradient of the acetowhite feature to segment the cervix into a disjoint subset of coherent regions in terms of acetowhite response. This step addresses the separation of columnar en squamous epithelium, such that the columnar epithelium is identified as the resulting regions from the acetowhite watershed segmentation that have lower feature values than the surrounding regions (i.e., valleys). The same idea is applied over the gradient of the red channel to detect the external orifice by identifying the valleys. Figure 8 shows segmentations obtained by the methodology proposed by Lange [61]. While the core regions of the cervix are properly identified, some artifacts are observed such as the recognition of external objects as acetowhite regions and the disconnected appearance of the external orifice (external os).

Some authors refine the segmentation task by detecting the external orifice [38], [69], [117], [122]. This process is done by gradient analysis in order to find the largest concave region in the image.

In general, these works do not present any objective assessment of the attained performance in terms of segmentation

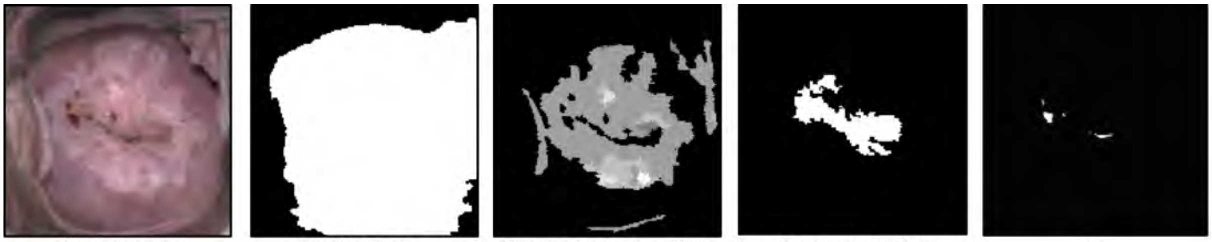


FIGURE 8. Lange and Ferris [61], [64] - cervix segmentation.

quality, providing in some cases a subjective notion of expert satisfaction [12], [13], [37], [39], [61], [64], [66].

The main drawback of these unsupervised strategies is the low semantic level at the decision process, working at a pixel or neighborhood level. Thereby, spatial coherence is unattained in most cases. Moreover, the lack of contours difficult the separability of the main regions without a global image representation. A common assumption of these techniques is that the cervix covers a significant portion of the image and that external objects (e.g., colposcope, gloves, swabs, etc.) are not present. Thereby, their robustness to unconstrained settings is limited.

To overcome the limitations of the unsupervised segmentation algorithms several supervised methodologies have been proposed using traditional segmentation-by-classification pipelines consisting of feature extraction and modeling with Support Vector Machines [47], [69], [117]. These techniques rely on color [47], [117] and texture information [117]. Huang *et al.* performed the recognition on superpixels resulting of a preliminary unsupervised clustering step [47]. Then, they use a one-vs-one SVM to classify the regions as acetowhite, columnar epithelium and squamous epithelium. While they present results for the pixelwise classification accuracy of cervix and non-cervix tissues, they do not show any quantitative results of the final multiclass segmentation.

Recent advances in semantic segmentation of digital colposcopies using deep learning techniques can be found in [19], [20], and [24]. The work of Fernandes and Cardoso [19] tackles the joint segmentation of several objects(i.e., colposcope, vaginal walls, cervix, transformation zone, and external orifice) in digital colposcopies. The proposed methodology extends the U-net deep architecture to improve the spatial ordinal consistency between objects. Namely, they induce segmentations where the objects of interest appear nested one inside the other. They validated the performance of their model on two databases covering all the colposcopy modalities and achieved a macro-average Dice's coefficient of 51.24% and 66.98% on the databases [22] and [51] respectively. Besides the capability of segmenting the entire set of objects in a global fashion, using deep neural networks enables segmentations with higher semantic level, where the segmentation of cervix tissues without edges is achieved by considering feature spaces with a high level of abstraction.

C. IMAGE REGISTRATION

According to Shapiro and Stockman [96], image registration defines the process whereby locations of two images from similar viewpoints of essentially the same scene are geometrically transformed in such a way that corresponding points of the two images have the same coordinates after transformation. The definition of Shapiro and Stockman might be relaxed when considering multimodal image registration by accepting a broad definition of *similar viewpoints of essentially the same scene*. Medical image registration is a challenging process, the intrinsic properties of each modality may distort the visual aspect of the objects in the image. We can think about medical image registration even in extreme cases where the images to align represent the external (e.g., RGB or depth image of the body) and internal structures (e.g., X-rays, ultrasound, etc.). The registration of body parts is complex, given the elastic deformations that occur in the body. For instance, the cervix is distorted in a non-rigid manner due to the patient breathing, muscular movements, etc. Even more, the modalities involved in the colposcopy may reveal and hide structures. For instance, the green light enhances vascularities, Hinselmann shows acetowhite regions and Schiller's test strongly dichotomizes the cervix into normal and abnormal areas.

In the literature, there are several works that have targeted the registration of colposcopies [1], [4], [5], [9], [30]–[33], [42], [43], [63], [65], [69], [76], [77], [82]. Three main lines of work have been proposed: global (either rigid or elastic), landmark-based and segmentation-based registration.

Since most of these works dealt with images from the same phase (typically Hinselmann), they were able to use standard (normalized) cross-correlation techniques [1], [4], [42], [43], commonly used when images belong to the same modality. In order to overcome the natural variations of the cervix, some works refine the rigid registration using local elastic registration techniques [31], [33], [63], [65].

Acosta-Mesa and his collaborators have a line of work in this area [1]–[5], [42], [43], either as the core focus of their work or as a preliminary step for the final classification of patients. Thus, Acosta-Mesa *et al.* [5] proposed a two-stage method to deal with local deformations. First, a phase correlation is applied in order to remove global translation difference between images. This method has some advantages when dealing with different contrast and brightness and with some simple intra-modal changes (i.e., acetowhite response), as can be observed in the acetowhite response [5]. Then, local deformations are removed using locally normalized cross-correlation. To accelerate the registration process, they proposed a method to register cervical images in grayscale [77], which performs a search of small local regions of the image in consecutive frames. The main challenge of colposcopy registration is the lack of distinctive landmarks in almost the entire cervix anatomy. In this sense, Acosta-Mesa *et al.* [1] proposed to use a manual stain landmark (at acquisition time) using Lugol solution and to use this landmark to simplify the registration process. While it is true that using such landmarks reduces the complexity of automatic methods for image registration, it adds complexity to the physician labor and could occlude relevant regions of the image with abnormal tissues.

Garcia-Arteaga *et al.* [31]–[33] proposed several methods for colposcopic image registration. In [33], an elastic registration algorithm was proposed, representing the problem as an optimization over a set of continuous deformation vector fields. Regularization was modeled by describing equilibrium in an elastic material using a linearized 2D elasticity operator (also used by Li *et al.* in [65]). The registration method is done in a multiscale fashion to speed up the process. No objective results are provided, but a mere visual inspection. Similarly to the two-stage approach used by Acosta-Mesa *et al.*, Garcia-Arteaga and collaborators [31] applied rigid registration with cross-correlation followed by elastic registration.

Given the challenges involved in global registration techniques, several attempts to address the problem as a landmark detection have been proposed [9], [27], [30], [69], [76]. These techniques take advantage of interest points such as Harris corner detector [9], [27], [76] that can be used to register images over time. Then, local descriptors such as SIFT [76], cross-correlation and distance [9] are used to identify matches. In posterior work, Garcia-Arteaga *et al.* [30] introduce geometric information about feasible deformations to remove false positives.

An alternative line of work use pre-segmented regions of the cervix to conduct registration [69], [82]. This kind of segmentation produces very coarse results, especially when the reference objects are of limited size such as the external orifice [82].

D. ABNORMAL TISSUE DETECTION AND CHARACTERIZATION

In the area of abnormal tissue detection and characterization, several methods have been proposed, some of them included hyper-spectral imaging [17], [25], [26], [40], [41], [108]. Since the current challenge in digital colposcopy is the scalability to remote healthcare centers with low resources, we will discuss methods that are able to work with traditional digital colposcopy that can be ported to current mobile devices. Namely, we focus on image processing techniques that handle RGB color images (and video).

In this section, we discuss works that tackled the localized recognition of these abnormalities. This could be considered as a midpoint between the previous section that addressed the pixelwise segmentation of the anatomic part and the next section that will cover the detection of relevant traits at a patient level (i.e., medical records, demographic data, etc.). In this sense, the following strategies address the problem of identifying and characterizing abnormal tissues at specific regions of the cervix. The main assumption of the works in this area is that the image constitutes the cervix regions (either by detection and cropping or by constrained acquisition) and that relevant anatomic parts have been segmented in a previous stage. Also, most works assume specular reflections (see III-A.2.a) have been removed. This last assumption is especially relevant since these artifacts could be easily recognized as positive acetowhite (AW) lesions. We can study these works from three different perspectives: lesion of interest, learning paradigm and type of data. Table 1 presents the main alternatives in these lines.

TABLE 1. Summary of the main categories of work on the detection of abnormal tissues.

Topic	Alternatives
Lesions	• Acetowhite lesions. • Vessels and mosaicism.
Learning paradigm	• Unsupervised. • Supervised.
Data	• Image-based. • Sequence of images (temporal).

Two main types of abnormal traits have been addressed in the literature: acetowhite lesions and abnormal vascularities/mosaicism.

For the detection of vascularities and mosaicism, most works relied on simple image processing techniques on static images. The main lines of research involve morphological operators and template matching [15], [48], [67], [100]–[102], [106], being the former of unsupervised nature and the latter of supervised nature with lazy learning (i.e., neighbor-based). Thereby, these techniques are highly sensitive to changes to the image resolution, scaling and illumination. This area is almost unexplored and has space for more robust techniques, able to cope with complex vessel patterns and with unconstrained settings.

As for segmentation, several works in the detection of acetowhite lesions applied unsupervised techniques, ranging from K-means [46], [68], Gaussian Mixture models [14], [16], [36], [37], [66], [100], [103], [107], [107], Mean Shift [65] and Watershed analysis [34], [35], [62],

[66], [110] to adaptive thresholding [11], [103]. The goal of watershed analysis techniques was mainly to over-segment the cervix according to the acetowhite response of the features [35]. Also, some works used deterministic annealing [106] and active contours [18] to detect lesions. The most widely used features for static images include color [34], [36], [37], [46], [62], [65], [66], [68], [87], [100], [101] texture [36], [68], edges [79] and spatial information [46], [100], also used by other supervised methods that will be explained below. The main limitation of unsupervised strategies is their low discriminative power to differentiate abnormal acetowhite regions from squamous epithelium since they have similar colors [37]. Other problems such as a high number of false negatives in regions with shadows and false positives on the vaginal walls are also typical [37]. In Figure 9, Gordon *et al.* illustrate these problems in the resulting images. This effect is present in general for methods that make predictions using local information (i.e., pixelwise data) without considering a global representation of the image.

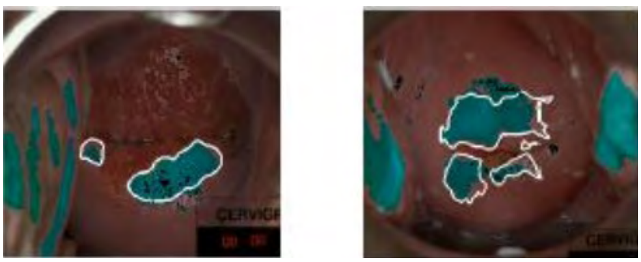


FIGURE 9. **Gordon *et al.* [37] - AW detected regions (green), manual annotations contours(white).**

Then, several methods addressed the problem from a supervised learning perspective. In general, this was done by extracting features from individual pixels, overlapping and non-overlapping tiles or by super-pixels obtained by segmentation techniques and applying a learning mechanism on the corresponding space. For classification, the most used method was KNN [1], [2], [4], [56], [72], [75], [85], [86], [89], [91], [92], [113], followed by Support Vector Machines [8], [56], [69], [85], [113], naive Bayes [1], [5], [85], [90] and multi layer perceptrons [92], [97], [113]. Other authors used Adaboost [112], [113], Conditional Random Fields [74], [83], among others [82], [92], [94], [113], [119]. The most common features for static images were color histograms at different scales [56], [69], [75], [85], [89], [113], oriented color gradients [56], [92], [112], [113], other color-based features [8], [75], [82], [83], [90], [92], [112], [119], [121], edges and texture [56], [69], [112], [113], [118], discrete wavelet transform [72], [86], [91], [118], and the amount of punctuation and vessels [74], [81], [83]. For sequence-based recognition, the features involve changes on the temporal acetowhite response either in a two image sequence (i.e., before and after application of acetic acid) [31], [31], [32], [65], [70], [81], [82] or at a fine-grained resolution level [1]–[5], [42], [94].

While these efforts addressed the problem from a pixelwise perspective, Alush *et al.* [6], [7] modeled the problem from a boundary-based approach, by classifying edges of superpixels as lesion or not. In this sense, a more global concept of the image is built. Superpixels are built using the watershed algorithm. The classification is performed by learning a dictionary of visual words and the problem is modeled using Markov Random Field (MRF), where each superpixel corresponds to a binary random variable indicating whether the region is part of the lesion. The final detection is done using belief propagation. Another dictionary-based algorithm was proposed by Zhang *et al.* [120], who used the K-Singular Value Decomposition method (K-SVD) to create positive and negative dictionaries of sparse representations. Finally, reconstructive errors of the sparse coefficients from the test images are calculated and compared for classification purposes.

A recent trend in the lesion detection combines different modalities for improving the final performance. In this sense, we have the work of Xu *et al.* [111] that combines text and image features in a late fusion and the work of Song *et al.* [99] that combines results from several modalities (e.g., cytology, HPV, colposcopy) and demographics (e.g., age) to train their model. Results on this direction seem promising.

Van Raad and her collaborators [109] proposed an automatic characterization of the lesions borders. After segmenting acetowhite regions using GMM, contours are characterized by detecting smoothness and irregularities. This type of characterization is relevant for medical teams as a way to introduce explanatory predictions for the decision support. Figure 10 shows the detected AW regions and the segment characterized as smooth (green) and irregular (black).

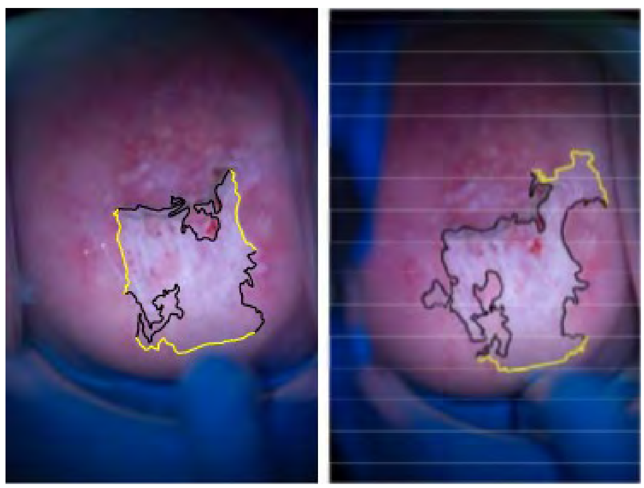


FIGURE 10. **Van Raad *et al.* [109] - yellow segments are characterized as smooth contours and black segments as irregular.**

For multi-image – temporal – approaches, where a sequence of images is presented to the model, the main lines of work were presented by the teams of Li *et al.* [65], Park *et al.* [82], Liu *et al.* [70], Acosta-Mesa [1], [4], [5], and Garcia-Arteaga [31], [32]. In these cases, the change in color before and after the application of acetic acid is used. The works conducted by Li *et al.* [65], Liu *et al.* [70], and Park *et al.* [82] focused on pairs of images (i.e., pre-acetic and post-acetic). The first two approaches used Mean Shift clustering and level sets respectively. The last approach, proposed by Park *et al.* [82] validated the performance of ensembles of supervised classification algorithms.

Acosta-Mesa and his team worked at a more fine-grained level [1], [4], [5] by extracting information from continuous frames at a pixelwise level to measure the acetowhite response. Their preliminary approach modeled the response using a parabola, which parameters are then used as features to classify the tissues using the naïve Bayes classifier. In successive works [1], [4], they explored discretization schemes to encode time series information, being able to surpass the human-level performance by 3% in terms of accuracy at a dataset with about 50 patients (76% and 73% respectively). This study was replicated for the assessment under green light in [42] and [43]. In a more recent work [73], they studied the performance of several classifiers on temporal data, achieving the best results with neural networks (89% of accuracy). Active contours were used as a post-processing step to identify suitable candidates for biopsy in [73] and [89].

Finally, Garcia-Arteaga *et al.* [31], [32] also considered time series analysis on the acetowhite response of the pixels. They focused on differentiating abnormal from normal tissues as a first task, achieving considerable performance (79.3% accuracy and 85% ROC AUC). Also, they present results for the classification of low-grade and high-grade lesion classification, achieving an accuracy of 92% and ROC AUC of 87%. While these results are satisfactory, the datasets are very limited in terms of the number of patients (3 and 10).

As a side application, Fernandes *et al.* [20] tackled the detection and characterization of lesions on the vagina using deep neural networks. While the images of study are from digital colposcopies, the application of interest is the forensic evaluation of sexual assault.

E. CLASSIFICATION OF GLOBAL TRAITS IN COLPOSCOPIES

Typical CAD systems involve the detection of global traits observed at images, from low-level tasks such as the modality recognition [21] to more semantic tasks like the identification of the cervix type [52], [55] and cancer detection [93], [114], [115].

Fernandes *et al.* [21] proposed a framework to recognize the acquisition modality of each frame in a video sequence. They propose a supervised learning scheme using color information and K-Nearest Neighbors. Global consistency between the predicted modalities and the colposcopy protocol is enforced using weighted finite automata. Also, a preprocessing step to filter noisy frames where the physician manipulates the cervix region is proposed.

Several works have addressed the problem of classifying a cervigram as cancer or non-cancer, being the line of research proposed by Huang and her team the most prominent [56], [99], [111]–[113], [115]. The standard scale to grade Cervical Intraepithelial Neoplasia (CIN) consists of three ordinal grades [115]: CIN1 (mild), CIN2 (moderate), and CIN3(severe). However, most works address the predictive task as a binary classification one by considering the classification of CIN1 from CIN2/3 or cancer (CIN2/3+) [115]. An alternative binary task is two classify the lesions as low-grade and high-grade squamous intraepithelial lesion, generally corresponding to CIN1 and CIN2/3+ respectively. After some preprocessing steps that cover removal of specular reflections and identification of the region of interest containing the cervix, these works extracted image features in a pyramidal fashion, including color histograms (typically on the Lab color space) [112], [113], [115], histogram of gradients [112], [113], [115], and Local Binary Patterns (LBP) [113], [115]. Several classifiers were used, including tree ensembles (Random Forest, Gradient Boosting, AdaBoost), neural networks, Logistic Regression, Support Vector Machines and K-Nearest Neighbors. Random Forests achieved a top performance of 84% ROC AUC in a dataset collected by the NCI (National Cancer Institute) in the Guanacaste project [44] with +1000 patients. In a more recent work, Xu *et al.* [115] compared the performance of deep features and the aforementioned pipeline based on traditional methodologies. In this sense, they extracted the features from the last dense layers from CaffeNet [49] trained on ImageNet and fine-tuned the last layer. While they achieved higher performance by using handcrafted features, further gains may be observed by training the network end-to-end instead of the final layer.

Xu *et al.* [114] proposed a multimodal approach to predict cervical cancer by merging deep features from AlexNet [57] and high-level information from medical records (e.g., age, HPV status, etc.). They were able to improve the performance obtained with image data from 88.77% ROC AUC to 94%.

Sato *et al.* [93] used Convolutional Neural Networks trained from scratch to predict cervical cancer on colposcopies with Hinselmann and Green filter modalities. As in the works mentioned above [114], [115], the architecture is considerably shallow, with three groups of convolutional-pooling layers and a couple of densely connected layers. They trained the architecture on a dataset with 485 images achieving 50% accuracy in recognizing three balanced classes. Further investigation of state-of-the-art architectures and regularization techniques (e.g., transfer learning, data augmentation) should be conducted in order to assess the actual capabilities of deep methodologies in this area.

In a recent competition about the categorization of cervix based on their transformation zones, deep learning methodologies achieved the best performance. The task was to characterize the cervix into three types depending on the transformation zone tissues type and observability [51].

- CAD for Digital Colposcopy
  - Quality
    - Enhancement
      - Color normalization [70]
      - Specular reflection removal [13], [39], [64]
    - Assessment [23], [41]
  - Segmentation
    - Methodologies
      - Unsupervised [14], [39], [63], [68], [88], [119]
      - Supervised [20], [49], [71], [119]
    - Features
      - color [13], [47], [68], [85], [88], [92], [119]
      - location [88], [92], [102], [118], [119], [124]
      - texture [39], [76], [124]
  - Registration
    - Global
      - Rigid [1], [4], [44], [45]
      - Elastic [33], [35], [65], [67]
    - Landmark-based [9], [10], [28], [32], [71], [78], [84]
  - Abnormality detection
    - Acetowhite regions
      - Image based [58], [58], [58], [71], [71], [77], [91], [120]
      - Sequence based [1], [34], [44], [45], [67], [75], [84]
    - Vessels and Mosaicism [16], [50], [69], [102]–[104], [108]
  - Classification
    - Modality Recognition [22]
    - Cancer vs. non Cancer
      - Deep
        - Image data [95], [117]
        - Multimodal data [116]
      - Traditional
        - color histograms [114], [115], [117]
        - gradients [114], [115], [117]
        - texture [115], [117]
    - Transformation Zone Type [54], [57]

FIGURE 11. Summary of the main research topics and selected works in the area.

The database consists of more than 1800 images from several modalities (Green, Hinselmann, and Schiller). The acquisition setting was unconstrained, having images with bad quality and images where the cervix was considerably small. Main pipelines to solve this problem involve the segmentation of the cervix and its transformation zone using the U-net architecture [55] and an ensemble of deep architectures to classify the images [52].

IV. SUMMARY

The research ecosystem on machine learning and computer vision techniques for the decision support of digital colposcopies has reached a sound point, with well-identified problems and paradigms to tackle them. Here, we will summarize the main traits of the aforementioned tasks and solutions. Figure 11 gathers the main features and works in each of these areas.

In the area of quality enhancement, the removal of specular reflections and the standardization of the color space are the main tasks of interest. The former has been tackled by a detection-inpainting scheme while the second one has been solved using camera calibration and simple image processing techniques. For quality assessment, the main features of interest are the complete observability of the cervix and the absence of disturbing artifacts such as specular reflections, bleeding, and external objects.

The semantic segmentation of the cervix tissues has been one of the areas that perceived more attention from the

TABLE 2. Summary of the datasets available databases.

Author	Multimodal	Multiview	Size		Annotations	
			Images	# Patients	Spatial	Global
Acosta-Mesa et al. [2]	No	Yes	10 videos	10	Yes*	No
Fernandes et al. [23]	Yes	No	287	100	Yes	Yes
Guanacaste Project (NCI/NIH)	Yes	Yes	+2k	387	No	Yes
Intel and MobileODT [53]	Yes	No	2k	-	Yes*	Yes

* Provided by us as part of this project.

research community. The vast majority of works addressed the problem using unsupervised clustering techniques. However, the hard assumptions on these works and the smooth boundaries between the cervix tissues demand more expressive models with high semantic power. Therefore, some supervised methodologies have appeared, including traditional machine learning and deep learning pipelines.

The registration of colposcopies was studied for unimodal settings. On the one hand, global image registration techniques that attempt to find a good global alignment of the images have been proposed. These techniques typically involve a rigid alignment of the main structures of the cervix, followed by an elastic registration to address eventual deformations of the body parts. On the other hand, landmark-based registration aims to detect and track points of interest.

For the spatial location and characterization of lesions, basic image processing techniques were used to detect vessels and mosaicism, including morphology operators and template matching. The recognition of acetowhite lesions received more attention, with methods covering both unsupervised and supervised techniques, and static and continuous acquisitions.

The final step of any CAD system is the diagnosis support. Therefore, providing a global decision per patient has been widely studied using machine learning techniques. Traditional methodologies include color and texture information while novel methods attempt to learn relevant features using deep methodologies. Some works have addressed the aggregation of multimodal data (e.g., medical records) achieving the best results in the literature.

V. DATABASES

As important as the methods delved to solve the aforementioned tasks are the databases used to validate their findings. Thus, the actual impact of any data-driven system relies on the similarity between the test database and the organic data acquired on a daily basis on medical facilities. Also, the diversity of acquisition settings and abnormalities is relevant. In this sense, we summarize the main aspects of the available datasets in the area (see Table 2).

A. Acosta-Mesa et al.

Acosta-Mesa and his team made available 10 videos with digital colposcopies of 10 patients after application of acetic acid [2]. The database does not contain manual annotations and the acquisition was very controlled. The duration of the sequences is 30 seconds (311 frames). The images have high quality and allow to study small patterns with high temporal and spatial resolution. This dataset can be used to validate (elastic) registration techniques and detection of acetowhite regions.

We made available as part of this project, manual annotations of 10 landmarks per video to validate the performance of image registration techniques.[1]

B. Fernandes et al.

The dataset was acquired by Fernandes *et al.* [22] in collaboration with *Hospital Universitario de Caracas* from Venezuela. The number of images is 287, including three modalities (Green light, Hinselmann, and Schiller). Several features were extracted from the dataset for the quality assessment task. The original subjective quality annotations were performed by six experts. The dataset also contains manual segmentation masks of the colposcope, vaginal walls, cervix, external orifice, and artifacts. It can be used to validate the performance of quality assessment methodologies and semantic image segmentation algorithms. The dataset can be accessed in the UCI Machine Learning repository.[2]

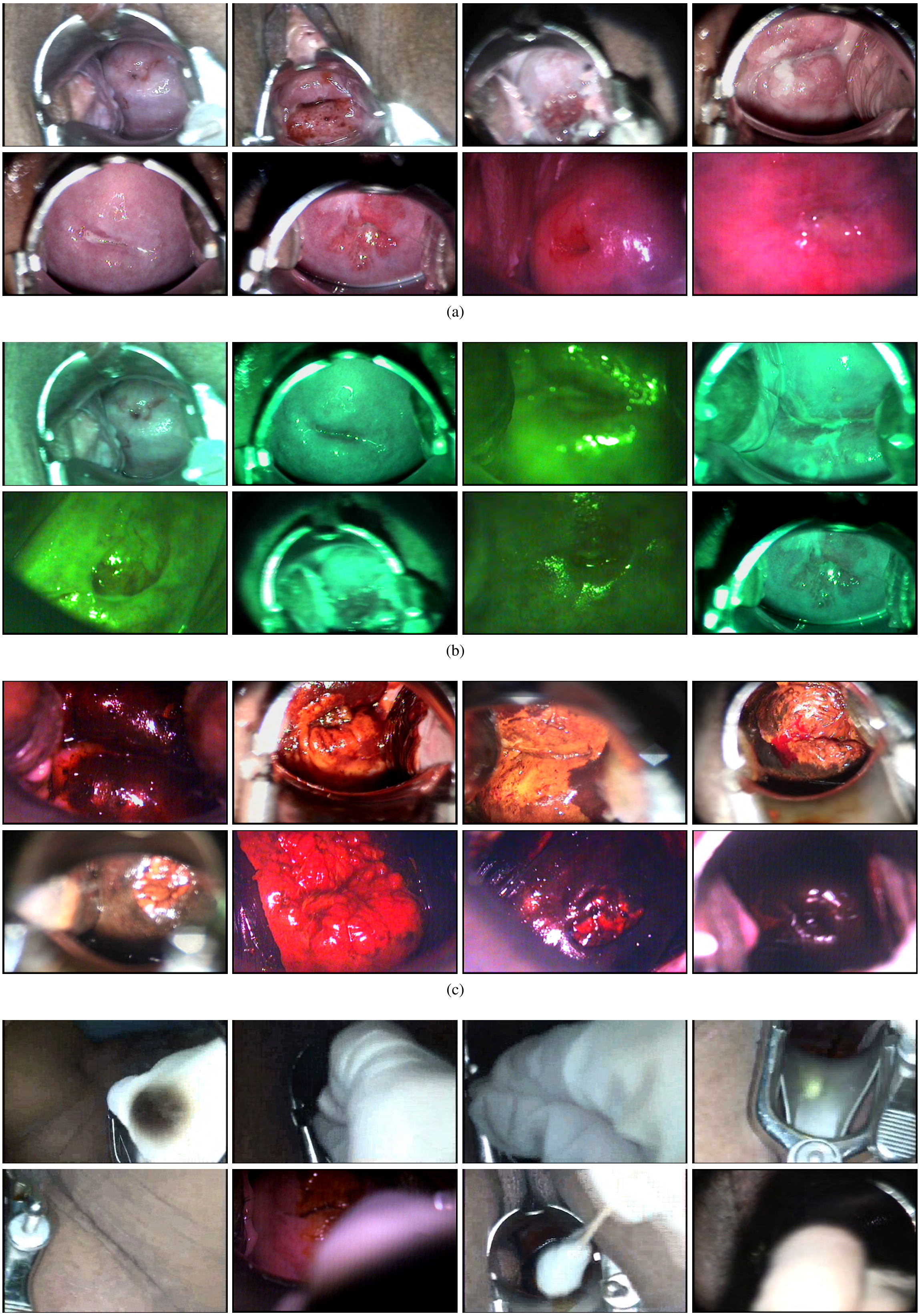
C. GUANACASTE PROJECT (NCI/NIH)

The dataset is made available by the National Cancer Institute (NCI)/National Institute of Health (NIH). The NCI collected the dataset in the Guanacaste project [44]. It consists of data from 10,000 anonymized women [115]. Technical works that use this database extract a subset of about 1,112 patient visits (767 visits in the CIN1 category and 345 visits in the CIN2/3+ category) [112], [113], [115]. The dataset contains multiple visits per patient, including multimodal information such as the age of the patient, HPV test, histology results, etc. The patients are annotated with the corresponding CIN progression level (i.e., normal, CIN1, CIN2, CIN3, and cancer). The presence of multiple images per patient in combination with other sources of data encourages the development of multi-view and multi-modal algorithms. Some precomputed visual features can be found in [113] and [115].

The dataset contains global information about the patient. Therefore, the dataset can be directly used to evaluate automatic methods for the detection of cervical intraepithelial neoplasia and cancer. It is also sensible to be used for

[1] https://github.com/kelwinfc/cervical-cancer-screening

[2] https://archive.ics.uci.edu/ml/datasets/Quality+Assessment+of+Digital+Colposcopies



(a)

(b)

(c)

(d)

FIGURE 12. **Sample images from the DCDB database. (a) Hinselmann. (b) Green filter. (c) Schiller. (d) Noisy frames where the physician manipulates the colposcope and the cervix region.**

assessment of semantic segmentation and registration techniques but it would require further annotations.

D. INTEL & MobileODT

Intel and MobileODT made available a database with about 2000 static images, covering the main modalities of the digital colposcopy. The dataset was released under the scope of a competition to identify the type of cervix among three types according to the location of the transformation zone [51]:

1) **Type I:** completely ectocervical, fully visible, small or large.
2) **Type II:** has endocervical component, fully visible, may have ectocervical component which may be small or large.
3) **Type III:** has endocervical component, is not fully visible, may have ectocervical component which may be small or large.

The dataset contains 1481 training images with annotations about the cervix type. The images distribution is unbalanced with 17%, 53%, and 30% respectively. All the cervix images in this dataset are considered normal (not cancerous) but the identification of the cervix type may require further testing [51]. The dataset has a large number of images that have not been curated but that can be used for the development of semi-supervised approaches.

As part of this project, we provide manual segmentation masks for this database, including the cervix region, transformation zone and external orifice.

VI. DCDB: DIGITAL COLPOSCOPY DATABASE

As was discussed in the previous section, several datasets have been acquired and made available by the research community. However, given the lack of a dataset that can be used on the assessment of all the aforementioned tasks, we collected a database with 129 digital colposcopies in video format. The videos were acquired between 2013 and 2015 at *Hospital Universitario de Caracas* in Caracas, Venezuela. The dataset covers the entire examination, including the main modalities of the colposcopy examination and intervals where the physician manipulates the cervix region. Thus, the dataset raises several challenges, from the multimodal and time-based integration of the decision to the identification of the proper frames to apply the models. In this sense, this dataset is close to a real-life scenario for the assessment of automated techniques. Figure 12 shows sample images from the database. Figure 13 summarizes some statistics about the videos.

Also, we make available the following annotations:

- **Modality Detection:** temporal annotations for the videos of the start and ending points of the modalities per frame. Also, we include annotations of the transition and noisy frames.
- **Quality Assessment:** annotations from 6 experts in an ordinal scale (i.e., poor, fair, good, excellent) for 287 images.

Minimum duration (seconds)	7
Maximum duration (seconds)	658
Average duration (seconds)	98
Median duration (seconds)	74

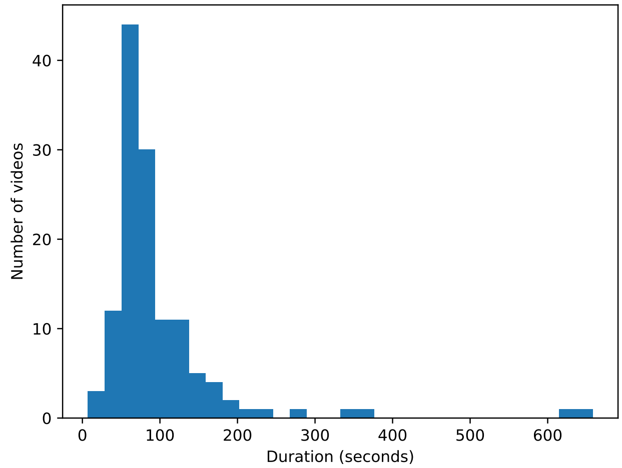


FIGURE 13. Summary of the database statistics and distribution of the video durations.

- **Semantic Segmentation:** annotations for 287 images of the colposcope, vaginal walls, cervix, external orifice, and artifacts.
- **Image Registration:** landmark annotations for image registration, including five points per video annotated every ten frames.
- **Abnormalities:** annotations about the lesions and abnormalities present in the image.

The videos and annotations will be continuously updated and improved. The database can be accessed online in [23].[3] Further details about the dataset, training/test partitions can be found in the project website.

VII. CONCLUSIONS AND CHALLENGES

The automated analysis of digital colposcopies has attained significant attention from the machine learning and computer vision research communities. We studied in this paper the main lines of research that have been conducted in this field and built a topology of tasks and approaches that encompass the area. While the area reached a certain level of maturity, the recent investment of companies and governments in the area and the recent publication of large databases open the possibility to include more advanced techniques in the development of CAD systems for digital colposcopies.

The main contributions of this work can be summarized as follows:

- We performed a review of the literature in the area.
- We established a common ground for the analysis of CAD systems for digital colposcopies.

[3]https://github.com/kelwinfc/cervical-cancer-screening If you find any problem with the source, please send an email to the corresponding authors of this paper.

- We released a video database with partial annotations that covers the main areas that were identified.
- We provide annotations for databases from third parties.
- We released source code and benchmarks for comparison on these databases.

Finally, despite the huge efforts that have been devoted to this area, several open challenges were identified. Below, we enumerate the main open problems in the area.

QUALITY ASSESSMENT AND ENHANCEMENT:
The notion of quality is a very subjective concept. Therefore, using a binary scale (i.e., bad, good) to define the quality of a digital colposcopy is too reductionist. Thus, a fully automated system should be able to identify, for each expert, the expected image quality in order to 1) suggest improvements during acquisition in real-time, and 2) retrieve the best frame to the human expert in order to maximize the confidence of his decision. While there is space on the normalization and enhancement of images without constrained acquisition settings, the appearance of deep learning methods that are robust to such variability may reduce the impact of these techniques.

SEGMENTATION OF CERVIX TISSUES:
In the area of semantic segmentation, the main limitation of the current strategies is the lack of adaptability to unconstrained settings. Due to the low semantic level of the techniques proposed in the literature, they are not able to segment objects with smooth transitions such as the cervix and the vaginal walls or the squamocolumnar junction. Thus, most of the published works focused on the segmentation of three entities: background, cervix and the external orifice. Moreover, current techniques are not able to cope with several modalities. Also, it is relevant to explore techniques to promote spatial consistency among the detected objects.

The development of deep learning architectures for semantic segmentation may be able to circumvent these problems, being able to represent global semantic properties of the image.

IMAGE REGISTRATION:
The main open challenge on the registration of digital colposcopies lies in the elastic registration of several modalities. Given the different signal statistics and disjoint observability of certain structures on the modalities, traditional registration techniques would not be able to cope with multimodal registration. Using segmented regions identified on each modality may drive a coarse alignment of the main cervix structures. However, a deformable alignment of the inner structures of the cervix would require additional complexity.

LESION DETECTION AND CHARACTERIZATION:

Besides the multimodal and temporal analysis that is intrinsic in all tasks, learning to detect and characterize lesions with cost-effective ways of annotations is a relevant problem. Traditional methodologies require a significant amount of manual labeling, including spatial localization of the lesions at an image level. Learning to detect lesions from weakly supervised annotations, where the expert identifies the presence of lesions in the video without explicitly identifying their boundaries, would directly impact the scalability of these frameworks.

CLASSIFICATION OF GLOBAL TRAITS:
Being the final stage of any CAD system, the amount of open problems in this area is prominent. We should look towards holistic frameworks able to extract knowledge from each modality (image and non-image data). Also, the inclusion of information from multiple visits from the same patient over the years should be addressed to identify long-term changes in the cervix.

While current strategies have simplified the prediction task to binary settings, developing predictive systems that can identify the progression of the lesions following the CIN ordinal scale would accelerate the acceptance of these systems.

Current systems work on a disjoint fashion by applying the aforementioned tasks in a cascade fashion. In this sense, the knowledge acquired from one task such as segmentation is not used when learning to solve another task such as quality assessment or cancer prediction. Thus, it is relevant to study transfer learning and multitask learning approaches to induce more robust and holistic decisions.

The final challenge–which is ubiquitous in all machine learning tasks for medicine– is the construction of interpretable and explanatory models. The proper support to the human expert must go beyond a simple categorical label. In order to facilitate and improve the work of the physicians and in order to have a tangible impact in the fight against the disease, CAD systems should be able to illustrate the human expert with similar examples from the past, to identify the factors that influenced the decision, to suggest treatment options with potential pros/cons for each case, among others.

REFERENCES

[1] H.-G. Acosta-Mesa, N. Cruz-Ramírez, K. Gutiérrez-Fragoso, R.-E. Barrientos-Martínez, and R. Hernández-Jiménez, "Assessing the possibility of identifying precancerous cervical lesions using aceto-white temporal patterns," in *Proc. Adv. Decis. Support Syst.*, 2010, pp. 107–116.

[2] H.-G. Acosta-Mesa, N. Cruz-Ramírez, and R. Hernández-Jiménez, "Aceto-white temporal pattern classification using k-NN to identify precancerous cervical lesion in colposcopic images," *Comput. Biol. Med.*, vol. 39, no. 9, pp. 778–784, 2009.

[3] H.-G. Acosta-Mesa, N. Cruz-Ramírez, R. Hernández-Jiménez, and D.-A. García-López, "Modeling aceto-white temporal patterns to segment colposcopic images," in *Proc. Iberian Conf. Pattern Recognit. Image Anal.*, 2007, pp. 548–555.

[4] H.-G. Acosta-Mesa, F. Rechy-Ramírez, E. Mezura-Montes, N. Cruz-Ramírez, and R. Hernández-Jiménez, "Application of time series discretization using evolutionary programming for classification of precancerous cervical lesions," *J. Biomed. Inform.*, vol. 49, pp. 73–83, Jun. 2014.

[5] H.-G. Acosta-Mesa *et al.*, "Cervical cancer detection using colposcopic images: A temporal approach," in *Proc. 6th Mexican Int. Conf. Comput. Sci. (ENC)*, Sep. 2005, pp. 158–164.

[6] A. Alush, H. Greenspan, and J. Goldberger, "Lesion detection and segmentation in uterine cervix images using an arc-level MRF," in *Proc. IEEE Int. Symp. Biomed. Imag., From Nano Macro*, Jun. 2009, pp. 474–477.

[7] A. Alush, H. Greenspan, and J. Goldberger, "Automated and interactive lesion detection and segmentation in uterine cervix images," *IEEE Trans. Med. Imag.*, vol. 29, no. 2, pp. 488–501, Feb. 2010.

[8] Y. Artan and X. Huang, "Combining multiple 2v-SVM classifiers for tissue segmentation," in *Proc. 5th IEEE Int. Symp. Biomed. Imag., From Nano Macro*, May 2008, pp. 488–491.

[9] J. D. García-Arteaga and J. Kybic, "Automatic landmark detection for cervical image registration validation," *Proc. SPIE*, vol. 6514, p. 65142S, Mar. 2007.

[10] E. Bengtsson and P. Malm, "Screening for cervical cancer using automated analysis of PAP-smears," *Comput. Math. Methods Med.*, vol. 2014, Mar. 2014, Art. no. 842037.

[11] I. Claude, R. Winzenrieth, P. Pouletaut, and J.-C. Boulanger, "Contour features for colposcopic image classification by artificial neural networks," in *Proc. 16th Int. Conf. Pattern Recognit.*, vol. 1, Aug. 2002, pp. 771–774.

[12] A. Das, A. Kar, and D. Bhattacharyya, "Elimination of specular reflection and identification of ROI: The first step in automated detection of cervical cancer using digital colposcopy," in *Proc. IEEE Int. Conf. Imag. Syst. Techn. (IST)*, May 2011, pp. 237–241.

[13] A. Das, A. Kar, and D. Bhattacharyya, "Preprocessing for automating early detection of cervical cancer," in *Proc. 15th Int. Conf. Inf. Vis. (IV)*, Jul. 2011, pp. 597–600.

[14] A. Das, A. Kar, and D. Bhattacharyya, "Implication of technology on society in asia: Automated detection of cervical cancer," in *Proc. IEEE Conf. Technol. Soc. Asia (TSA)*, Oct. 2012, pp. 1–4.

[15] A. Das, A. Kar, and D. Bhattacharyya, "Detection of abnormal regions of precancerous lesions in digitised uterine cervix images," in *Proc. Int. Elect. Eng. Congr. (iEECON)*, Mar. 2014, pp. 1–4.

[16] A. Das, A. Kar, and D. Bhattacharyya, "Early detection of cervical cancer using novel segmentation algorithms," *Invertis J. Sci. Technol.*, vol. 7, no. 2, pp. 91–95, 2014.

[17] M. J. DeWeert *et al.*, "Analysis of spatial variability in hyperspectral imagery of the uterine cervix *in vivo*," *Proc. SPIE*, vol. 4959, pp. 67–76, Jul. 2003.

[18] O. El Meslouhi, M. Kardouchi, H. Allali, and T. Gadi, "Semi-automatic cervical cancer segmentation using active contours without edges," in *Proc. 5th Int. Conf. Signal-Image Technol. Internet-Based Syst. (SITIS)*, Nov. 2009, pp. 54–58.

[19] K. Fernandes and J. S. Cardoso, "Ordinal image segmentation using deep neural networks," in *Proc. Int. Joint Conf. Neural Netw.*, Jul. 2018.

[20] K. Fernandes, J. S. Cardoso, and B. Astrup, "A deep learning approach for the forensic evaluation of sexual assault," in *Pattern Analysis and Applications*. London, U.K.: Springer, 2018.

[21] K. Fernandes, J. S. Cardoso, and J. Fernandes, "Temporal segmentation of digital colposcopies," in *Pattern Recognition and Image Analysis*. Cham, Switzerland: Springer, 2015, pp. 262–271.

[22] K. Fernandes, J. S. Cardoso, and J. Fernandes, "Transfer learning with partial observability applied to cervical cancer screening," in *Proc. Iberian Conf. Pattern Recognit. Image Anal.*, 2017, pp. 243–250.

[23] K. Fernandes, J. S. Cardoso, and J. Fernandes. (2018). *Cervical Cancer Screening*. Accessed: May 17, 2018. [Online]. Available: https://github.com/kelwinfc/cervical-cancer-screening

[24] K. Fernandes, R. Cruz, and J. S. Cardoso, "Deep image segmentation by quality inference," in *Proc. Int. Joint Conf. Neural Netw.*, Jul. 2018.

[25] D. G. Ferris *et al.*, "Colposcopy quality control: Establishing colposcopy criterion standards for the national cancer institute ALTS trial using cervigrams," *J. Lower Genital Tract Disease*, vol. 2, no. 4, pp. 195–203, 1998.

[26] D. G. Ferris *et al.*, "Multimodal hyperspectral imaging for the noninvasive diagnosis of cervical neoplasia," *J. Lower Genital Tract Disease*, vol. 5, no. 2, pp. 65–72, 2001.

[27] A. B. Flores, L. A. Robles, R. M. M. Tepalt, and J. D. C. Aragon, "Identifying precursory cancer lesions using temporal texture analysis," in *Proc. 2nd Can. Conf. Comput. Robot Vis. (CRV)*, May 2005, pp. 34–39.

[28] C. F. D. Control *et al.*, "Cervical cancer screening among women aged 18-30 years—United states, 2000-2010," *MMWR Morbidity Mortality Weekly Rep.*, vol. 61, nos. 51–52, p. 1038, 2013.

[29] J. Galbally, S. Marcel, and J. Fierrez, "Image quality assessment for fake biometric detection: Application to iris, fingerprint, and face recognition," *IEEE Trans. Image Process.*, vol. 23, no. 2, pp. 710–724, Feb. 2014.

[30] J. D. García-Arteaga, J. Kybic, J. Gu, and W. Li, "Geometric and information constraints for automatic landmark selection in colposcopy sequences," in *Proc. VISAPP*, 2007, pp. 333–338.

[31] J. D. García-Arteaga, J. Kybic, and W. Li, "Automatic colposcopy video tissue classification using higher order entropy-based image registration," *Comput. Biol. Med.*, vol. 41, no. 10, pp. 960–970, 2011.

[32] J. D. García-Arteaga, "Multichannel image information similarity measures: Applications to colposcopy image registration," Ph.D. dissertation, Faculty Elect. Eng., Czech Tech. Univ., Prague, Czech Republic, 2012.

[33] J. D. García-Arteaga, J. Kybic, and W. Li, "Elastic image registration for movement compensation in digital colposcopy," in *Proc. Biosignal Anal. Biomed. Signals Images*, Brno, Czech Republic, Jun. 2006, pp. 236–238.

[34] S. Gordon and H. Greenspan, "Segmentation of non-convex regions within uterine cervix images," in *Proc. 4th IEEE Int. Symp. Biomed. Imag., From Nano Macro*, Apr. 2007, pp. 312–315.

[35] S. Gordon and H. Greenspan, "An agglomerative segmentation framework for non-convex regions within uterine cervix images," *Image Vis. Comput.*, vol. 28, no. 12, pp. 1682–1701, 2010.

[36] S. Gordon, G. Zimmerman, and H. Greenspan, "Image segmentation of uterine cervix images for indexing in PACS," in *Proc. 17th IEEE Symp. Comput.-Based Med. Syst. (CBMS)*, Jun. 2004, p. 298.

[37] S. Gordon, G. Zimmerman, R. Long, S. Antani, J. Jeronimo, and H. Greenspan, "Content analysis of uterine cervix images: Initial steps toward content based indexing and retrieval of cervigrams," *Proc. SPIE*, vol. 6144, p. 61444U, Mar. 2006.

[38] H. Greenspan *et al.*, "Automatic detection of anatomical landmarks in uterine cervix images," *IEEE Trans. Med. Imag.*, vol. 28, no. 3, pp. 454–468, Mar. 2009.

[39] J. Gu and W. Li, "Automatic image quality assessment for uterine cervical imagery," *Proc. SPIE*, vol. 6146, p. 61461B, Mar. 2006.

[40] U. P. Gustafsson *et al.*, "*In-vivo* fluorescence and reflectance imaging of human cervical tissue," *Proc. SPIE*, vol. 5031, pp. 521–531, May 2003.

[41] U. P. Gustafsson *et al.*, "Fluorescence and reflectance monitoring of human cervical tissue *in vivo*: A case study," *Proc. SPIE*, vol. 4959, pp. 100–110, Jul. 2003.

[42] K. Gutiérrez-Fragoso, H. G. Acosta-Mesa, N. Cruz-Ramírez, and R. Hernández-Jiménez, "Automatic classification of acetowhite temporal patterns to identify precursor lesions of cervical cancer," *J. Phys., Conf. Ser.*, vol. 475, no. 1, p. 012004, 2013.

[43] K. Gutiérrez-Fragoso, H. G. Acosta-Mesa, N. Cruz-Ramírez, and R. Hernández-Jiménez, "Optimization of classification strategies of acetowhite temporal patterns towards improving diagnostic performance of colposcopy," *Comput. Math. Methods Med.*, vol. 2017, Jul. 2017, Art. no. 5989105.

[44] R. Herrero *et al.*, "Design and methods of a population-based natural history study of cervical neoplasia in a rural province of Costa Rica: The Guanacaste Project," *Rev. Panamericana Salud Pública*, vol. 1, no. 6, pp. 411–425, 1997.

[45] J. Hill, E. Corona, J. Ao, S. Mitra, and B. Nutter, "Information theoretic clustering for medical image segmentation," in *Proc. Adv. Comput. Approaches Biomed. Eng.*, 2014, pp. 47–70.

[46] S. Huang, M. Gao, D. Yang, X. Huang, A. Elgammal, and X. Zhang, "Unbalanced graph-based transduction on superpixels for automatic cervigram image segmentation," in *Proc. IEEE 12th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2015, pp. 1556–1559.

[47] X. Huang, W. Wang, Z. Xue, S. Antani, L. R. Long, and J. Jeronimo, "Tissue classification using cluster features for lesion detection in digital cervigrams," *Proc. SPIE*, vol. 6914, p. 69141Z, Mar. 2008.

[48] Q. Ji, J. Engel, and E. Craine, "Texture analysis for classification of cervix lesions," *IEEE Trans. Med. Imag.*, vol. 19, no. 11, pp. 1144–1149, Nov. 2000.

[49] Y. Jia *et al.*, "Caffe: Convolutional architecture for fast feature embedding," in *Proc. 22nd ACM Int. Conf. Multimedia*, 2014, pp. 675–678.

[50] Kaggle. (2015). *Cervical Cancer Screening*. Accessed: Mar. 7, 2018. [Online]. Available: https://www.kaggle.com/c/cervical-cancer-screening

[51] Kaggle. (2017). *Intel & MobileODT Cervical Cancer Screening*. Accessed: Mar. 5, 2018. [Online]. Available: https://www.kaggle.com/c/intel-mobileodt-cervical-cancer-screening

[52] Kaggle. (2017). *Intel & MobileODT Cervical Cancer Screening Competition, 1st Place Winner's Interview: Team 'Towards Empirically Stable Training'*. Accessed: Mar. 5, 2018. [Online]. Available: http://blog.kaggle.com/2017/08/22/intel-mobileodt-cervical-cancer-screening-competition-1st-place-winners-interview-team-towards-empirically-stable-training/

[53] L. Kang, P. Ye, Y. Li, and D. Doermann, "Convolutional neural networks for no-reference image quality assessment," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 1733–1740.

[54] R. P. Kauffman, S. J. Griffin, J. D. Lund, and P. E. Tullar, "Current recommendations for cervical cancer screening: Do they render the annual pelvic examination obsolete?" *Med. Princ. Pract.*, vol. 22, no. 4, pp. 313–322, 2013.

[55] N. Kaur, N. Panigrahi, and A. Mittal, "Automated cervical cancer screening using transfer learning," in *Proc. Int. Conf. Recent Adv. Eng. Sci. Manage.*, 2017, pp. 2110–2119.

[56] E. Kim and X. Huang, "A data driven approach to cervigram image analysis and classification," in *Color Medical Image Analysis*. Dordrecht, Netherlands: Springer, 2013, pp. 1–13.

[57] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2012, pp. 1097–1105.

[58] V. Kudva, K. Prasad, and S. Guruvare, "Detection of specular reflection and segmentation of cervix region in uterine cervix images for cervical cancer screening," *IRBM*, vol. 38, no. 5, pp. 281–291, 2017.

[59] C. T. Lam *et al.*, "Design of a novel low cost point of care tampon (POCkeT) colposcope for use in resource limited settings," *PLoS ONE*, vol. 10, no. 9, p. e0135869, 2015.

[60] C. T. Lam *et al.*, "An integrated strategy for improving contrast, durability, and portability of a pocket colposcope for cervical cancer screening and diagnosis," *PLoS ONE*, vol. 13, no. 2, p. e0192530, 2018.

[61] H. Lange, "Automatic detection of multi-level acetowhite regions in RGB color images of the uterine cervix," *Proc. SPIE*, vol. 5747, pp. 1004–1017, Apr. 2005.

[62] H. Lange, "Automatic glare removal in reflectance imagery of the uterine cervix," *Proc. SPIE*, vol. 5747, pp. 2183–2193, Apr. 2005.

[63] H. Lange, R. Baker, J. Hakansson, and U. P. Gustafsson, "Reflectance and fluorescence hyperspectral elastic image registration," *Proc. SPIE*, vol. 5370, pp. 335–345, May 2004.

[64] H. Lange and D. G. Ferris, "Computer-aided-diagnosis (CAD) for colposcopy," *Proc. SPIE*, vol. 5747, pp. 71–84, Apr. 2005.

[65] W. Li, D. G. Ferris, and R. W. Lieberman, "Computerized image analysis for acetic acid induced intraepithelial lesions," *Proc. SPIE*, vol. 6914, p. 69143A, Mar. 2008.

[66] W. Li, J. Gu, D. Ferris, and A. Poirson, "Automated image analysis of uterine cervical images," *Proc. SPIE*, vol. 6514, p. 65142P, Mar. 2007.

[67] W. Li and A. Poirson, "Detection and characterization of abnormal vascular patterns in automated cervical image analysis," in *Proc. Int. Symp. Vis. Comput.*, 2006, pp. 627–636.

[68] W. Li, M. Soto-Thompson, and U. Gustafsson, "A new image calibration system in digital colposcopy," *Opt. Express*, vol. 14, no. 26, pp. 12887–12901, 2006.

[69] M. Liang, G. Zheng, X. Huang, G. Milledge, and A. Tokuta, "Identification of abnormal cervical regions from colposcopy image sequences," in *Proc. Int. Conf. Comput. Graph., Vis. Comput. Vis.*, 2013, pp. 1–7.

[70] J. Liu, L. Li, and L. Wang, "Acetowhite region segmentation in uterine cervix images using a registered ratio image," *Comput. Biol. Med.*, vol. 93, pp. 47–55, Feb. 2018.

[71] S. Lotenberg, S. Gordon, and H. Greenspan, "Shape priors for segmentation of the cervix region within uterine cervix images," *J. Digit. Imag.*, vol. 22, no. 3, pp. 286–296, 2009.

[72] V. Margariti, M. Zervakis, and C. Balas, "Wavelet and physical parametric analysis of the acetowhitening optical effect: comparative evaluation of performances in non-invasive diagnosis of cervical neoplasia," in *Proc. 10th IEEE Int. Conf. Inf. Technol. Appl. Biomed.*, Nov. 2010, pp. 1–4.

[73] A. Marquez-Grajales, H.-G. Acosta-Mesa, E. Mezura-Montes, and R. Hernández-Jiménez, "Cervical image segmentation using active contours and evolutionary programming over temporary acetowhite patterns," in *Proc. IEEE Congr. Evol. Comput. (CEC)*, Jul. 2016, pp. 3863–3870.

[74] D. P. Mary, V. Anandan, and K. G. Srinivasagan, "An effective diagnosis of cervical cancer neoplasia by extracting the diagnostic features using CRF," in *Proc. Int. Conf. Comput., Electron. Elect. Technol. (ICCEET)*, Mar. 2012, pp. 563–570.

[75] G. Mehlhorn, C. Münzenmayer, M. Benz, A. Kage, M. W. Beckmann, and T. Wittenberg, "Computer-assisted diagnosis in colposcopy: Results of a preliminary experiment?" *Acta Cytol.*, vol. 56, no. 5, pp. 554–559, 2012.

[76] O. El Meslouhi, H. Allali, T. Gadi, M. Kardouchi, "Colposcopic image registration using opponentSIFT descriptor," in *Proc. Int. Symp. Commun. Mobile Netw.*, 2010, pp. 1–4.

[77] E. Mezura-Montes, H.-G. Acosta-Mesa, D.-S. Ramírez-Garcés, N. Cruz-Ramírez, and R. Hernández-Jiménez, "An image registration method for colposcopic images," *Comput. Math. Methods Med.*, vol. 2013, Aug. 2013, Art. no. 285962.

[78] MobileODT. (2018). *Eva Colposcope*. Accessed: Feb. 21, 2018. [Online]. Available: https://mobileodt.com/products/eva-colp/

[79] S. Mustafa, S. Adeshina, M. Dauda, and W. Soboyejo, "Classification of cervical cancer tissues using a novel low cost methodology for effective screening in rural settings," in *Proc. 11th Int. Conf. Electron., Comput. Comput. (ICECCO)*, Sep. 2014, pp. 1–4.

[80] V. Pallavi and K. Payal, "Automated analysis of cervix images to grade the severity of cancer," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2011, pp. 3439–3442.

[81] S. Y. Park, D. Sargent, R. Lieberman, and U. Gustafsson, "Domain-specific image analysis for cervical neoplasia detection based on conditional random fields," *IEEE Trans. Med. Imag.*, vol. 30, no. 3, pp. 867–878, Mar. 2011.

[82] S. Y. Park *et al.*, "Automated image analysis of digital colposcopy for the detection of cervical neoplasia," *J. Biomed. Opt.*, vol. 13, no. 1, p. 014029, 2008.

[83] S. Y. Park, D. Sargent, R. Wolters, and R. W. Lieberman, "Semantic image analysis for cervical neoplasia detection," in *Proc. IEEE 4th Int. Conf. Semantic Comput. (ICSC)*, Sep. 2010, pp. 160–165.

[84] BD Focal Point. (2018). *BD SurePath*. Accessed: Feb. 21, 2018. [Online]. Available: http://www.bd.com/tripath/labs/fp screening.asp

[85] P. R. Praba and H. Ranganathan, "Comparing different classifiers for automatic lesion detection in cervix based on colour histogram," *J. Comput. Appl.*, vol. 6, no. 1, pp. 15–18, 2013.

[86] P. R. Praba and H. Ranganathan, "Wavelet transform based automatic lesion detection in cervix images using active contour," *J. Comput. Sci.*, vol. 9, no. 1, p. 30, 2013.

[87] V. Prabitha *et al.*, "Multi-spectral diffuse reflectance imaging for detection of cervical lesions: A pilot study," *Int. J. Eng. Sci. Innov. Technol.*, vol. 3, no. 6, pp. 169–177, 2014.

[88] V. V. Raad and A. P. Bradley, "Active contour model based segmentation of colposcopy images of cervix uteri using Gaussian pyramids," in *Proc. 6th Int. Symp. Digit. Signal Process. Commun. Syst. (DSPCS)*, 2002, pp. 1–6.

[89] P. S. RamaPraba, M. P. Chitra, and P. Kumar, "Effective lesion detection of colposcopic images using active contour method," *Biomed. Res.*, vol. 28, Feb. 2017.

[90] P. S. RamaPraba and H. Ranganathan, "Automatic lesion detection in colposcopy cervix images based on statistical features," in *Global Trends in Information Systems and Software Applications*. Berlin, Germany: Springer, 2012, pp. 424–430.

[91] P. Ramapraba and H. Ranganathan, "Performance evolution of various wavelets in cervical lesion detection," *Indian J. Comput. Sci. Eng.*, vol. 4, no. 6, pp. 481–485, 2013.

[92] F. Rouhbakhsh, F. Farokhi, and K. Kangarloo, "Effective feature selection for pre-cancerous cervix lesions using artificial neural networks," *Int. J. Smart Elect. Eng.*, vol. 1, no. 3, pp. 199–204, 2012.

[93] M. Sato *et al.*, "Application of deep learning to the classification of images from colposcopy," *Oncol. Lett.*, vol. 15, no. 3, pp. 3518–3523, 2018.

[94] P. Schmid-Saugeon, J. D. Pitts, B. Kaufman-Howard, A. Zelenchuk, and D. M. Harper, "Time-resolved imaging of cervical acetowhitening," Spandidos Publ., Athens, Greece, Tech. Rep., 2004, vol. 15, no. 3.

[95] J. W. Sellors and R. Sankaranarayanan, *Colposcopy and Treatment of Cervical Intraepithelial Neoplasia: A Beginner's Manual*. New Delhi, India: Diamond Pocket, 2003.

[96] L. Shapiro and G. C. Stockman, *Computer Vision*. Upper Saddle River, NJ, USA: Prentice-Hall, 2001.

[97] P. W. Simões *et al.*, "Classification of images acquired with colposcopy using artificial neural networks," *Cancer Inform.*, vol. 13, p. 119, Jan. 2014.

[98] V. B. Singh, N. Gupta, R. Nijhawan, R. Srinivasan, V. Suri, and A. Rajwanshi, "Liquid-based cytology versus conventional cytology for evaluation of cervical pap smears: Experience from the first 1000 split samples," *Indian J. Pathol. Microbiol.*, vol. 58, no. 1, p. 17, 2015.

[99] D. Song *et al.*, "Multimodal entity coreference for cervical dysplasia diagnosis," *IEEE Trans. Med. Imag.*, vol. 34, no. 1, pp. 229–245, Jan. 2015.

[100] Y. Srinivasan, E. Corona, B. Nutter, S. Mitra, and S. Bhattacharya, "A unified model-based image analysis framework for automated detection of precancerous lesions in digitized uterine cervix images," *IEEE J. Sel. Topics Signal Process.*, vol. 3, no. 1, pp. 101–111, Feb. 2009.

[101] Y. Srinivasan *et al.*, "A probabilistic approach to segmentation and classification of neoplasia in uterine cervix images using color and geometric features," *Proc. SPIE*, vol. 5747, pp. 995–1003, Apr. 2005.

[102] Y. Srinivasan, B. Nutter, S. Mitra, B. Phillips, and E. Sinzinger, "Classification of cervix lesions using filter bank-based texture mode," in *Proc. 19th IEEE Symp. Comput.-Based Med. Syst. (CBMS)*, Jun. 2006, pp. 832–840.

[103] Y. Srinivasan, S. Yang, B. Nutter, S. Mitra, B. Phillips, and R. Long, "Challenges in automated detection of cervical intraepithelial neoplasia," *Proc. SPIE*, vol. 6514, p. 65140F, Mar. 2007.

[104] Thinprep. (2018). *THINPREP*. Accessed: Feb. 21, 2018. [Online]. Available: https://healthdxs.com/en/thinprep/

[105] M. Traversi, M. Falagario, and C. Guaragnella, "CADdy—Colposcopy learning machine for computer aided diagnosis," in *Proc. IEEE 3rd Int. Conf. Consum. Electron. Berlin (ICCE-Berlin)*, Sep. 2013, pp. 1–4.

[106] B. Tulpule, S. Yang, Y. Srinivasan, S. Mitra, and B. Nutter, "Segmentation and classification of cervix lesions by pattern and texture analysis," in *Proc. 14th IEEE Int. Conf. Fuzzy Syst. (FUZZ)*, May 2005, pp. 173–176.

[107] V. Van Raad, "Image analysis and segmentation of anatomical features of cervix uteri in color space," in *Proc. 13th Eur. Signal Process. Conf.*, Sep. 2005, pp. 1–4.

[108] V. Van Raad, "A new vision approach for local spectrum features in cervical images via 2D method of geometric restriction in frequency domain," in *Proc. Int. Comput. Vis. Biomed. Image Appl.*, 2005, pp. 125–134.

[109] V. V. Raad, Z. Xue, and H. Lange, "Lesion margin analysis for automated classification of cervical cancer lesions," *Proc. SPIE*, vol. 6144, p. 614454, Mar. 2006.

[110] J. Xiong, L. Wang, and J. Gu, "Image segmentation of the acetowhite region in cervix images based on chromaticity," in *Proc. 9th Int. Conf. Inf. Technol. Appl. Biomed.*, Nov. 2009, pp. 1–4.

[111] T. Xu, X. Huang, E. Kim, L. R. Long, and S. Antani, "Multi-test cervical cancer diagnosis with missing data estimation," *Proc. SPIE*, vol. 9414, p. 94140X, Mar. 2015.

[112] T. Xu, E. Kim, and X. Huang, "Adjustable adaboost classifier and pyramid features for image-based cervical cancer diagnosis," in *Proc. IEEE 12th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2015, pp. 281–285.

[113] T. Xu *et al.*, "A new image data set and benchmark for cervical dysplasia classification evaluation," in *Proc. Int. Workshop Mach. Learn. Med. Imag.*, 2015, pp. 26–35.

[114] T. Xu, H. Zhang, X. Huang, S. Zhang, and D. N. Metaxas, "Multimodal deep learning for cervical dysplasia diagnosis," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.*, 2016, pp. 115–123.

[115] T. Xu *et al.*, "Multi-feature based benchmark for cervical dysplasia classification evaluation," *Pattern Recognit.*, vol. 63, pp. 468–475, Mar. 2017.

[116] Z. Xue, S. Antani, L. R. Long, J. Jeronimo, and G. R. Thoma, "Comparative performance analysis of cervix ROI extraction and specular reflection removal algorithms for uterine cervix image analysis," *Proc. SPIE*, vol. 6512, p. 65124I, Mar. 2007.

[117] Z. Xue, S. Antani, L. R. Long, and G. R. Thoma, "An online segmentation tool for cervicographic image analysis," in *Proc. 1st ACM Int. Health Informat. Symp.*, 2010, pp. 425–429.

[118] Z. Xue, L. R. Long, S. Antani, and G. R. Thoma, "Automatic extraction of mosaic patterns in uterine cervix images," in *Proc. IEEE 23rd Int. Symp. Comput.-Based Med. Syst. (CBMS)*, Oct. 2010, pp. 273–278.

[119] Y. Yu, J. Huang, S. Zhang, C. Restif, X. Huang, and D. Metaxas, "Group sparsity based classification for cervigram segmentation," in *Proc. IEEE Int. Symp. Biomed. Imag., From Nano Macro*, Mar. 2011, pp. 1425–1429.

[120] S. Zhang, J. Huang, W. Wang, X. Huang, and D. Metaxas, "Cervigram image segmentation based on reconstructive sparse representations," *Proc. SPIE*, vol. 7623, p. 762313, Mar. 2010.

[121] Z.-Q. Zhang, T.-X. Wen, and J. Gu, "Cervical cancerous region detection using spectral matting technique: A pilot study on color space," *Bull. Adv. Technol. Res.*, vol. 5, no. 7, pp. 48–51, Jul. 2011.

[122] G. Zimmerman, S. Gordon, and H. Greenspan, "Automatic landmark detection in uterine cervix images for indexing in a content-retrieval system," in *Proc. 3rd IEEE Int. Symp. Biomed. Imag., Nano Macro*, Apr. 2006, pp. 1348–1351.

KELWIN FERNANDES received the B.Sc. degree *(summa cum laude)*in computer engineering from Universidad Simón Bolívar, Caracas, Venezuela, in 2012. He is currently pursuing the Ph.D. degree with the Universidade do Porto. He is also a Researcher with the Instituto de Engenharia de Sistemas e Computadores Tecnologia e Ciência, Porto, Portugal. His main research interests include machine learning, computer vision, and artificial intelligence.

JAIME S. CARDOSO received the Licenciatura degree in electrical engineering and the Ph.D. degree in computer vision from the Universidade do Porto, Porto, Portugal, in 1999 and 2006, respectively. He is currently an Associate Professor with habilitation with the Faculty of Engineering, Universidade do Porto, and a Senior Researcher of information processing and pattern recognition area with the Centre for Telecommunications and Multimedia, Instituto de Engenharia de Sistemas e Computadores Tecnologia e Ciência, Porto. His research focus on three major topics computer vision, machine learning, and decision support systems.

JESSICA FERNANDES received the Medical Surgeon degree *(magna cum laude)* from the Universidad Central de Venezuela (UCV), Caracas, Venezuela, in 2010, and the master's degree in contraception and sexual and reproductive health from the Universidad de Alcalá in 2018. She completed her specialization on obstetrics and gynecology with the Hospital Universitario de Caracas and UCV in 2014. She then did a specialization course on gynecological endocrinology with UCV in 2015. She was an intern in human reproduction with Unifertes–RedLara in 2015 and did the Post-Graduation Program with the Universidad de Alcalá, Madrid, Spain, in 2016.