Cytology Image Analysis Techniques Toward Automation: Systematically Revisited

SHYAMALI MITRA, NIBARAN DAS, and SOUMYAJYOTI DEY, Jadavpur University, India
SUKANTA CHAKRABORTY, Theism Medical Diagnostics Centre, India
MITA NASIPURI and MRINAL KANTI NASKAR, Jadavpur University, India

Cytology is a branch of pathology that deals with the microscopic examination of cells for diagnosis of carcinoma or inflammatory conditions. In the present work, the term cytology is used to indicate solid organ cytology. Automation in cytology started in the early 1950s with an aim to reduce manual efforts in the diagnosis of cancer. The influx of intelligent systems with high computational power and improved specimen collection techniques helped to achieve technological heights in the cytology automation process. In the present survey, we focus on image analysis techniques paving the way to automation in cytology. We take a short tour of 17 types of solid organ cytology to explore various segmentation and/or classification techniques that evolved during the past three decades to automate cytology image analysis. It is observed that most of the works are aligned toward three types of cytology: Cervical, Breast, and Respiratory tract cytology. These are discussed elaborately in the article. Commercial systems developed during the period are also summarized to comprehend the overall growth in respective domains. Finally, we discuss different state-of-the-art methods and related challenges to provide prolific and competent future research directions in bringing cytology-based commercial systems into the mainstream.

CCS Concepts: • Applied computing → Health informatics; • Computing methodologies → Machine learning;

Additional Key Words and Phrases: Cytology survey, image segmentation, image classification, malignant and benign, computer aided diagnosis

ACM Reference format:
Shyamali Mitra, Nibaran Das, Soumyajyoti Dey, Sukanta Chakraborty, Mita Nasipuri, and Mrinal Kanti Naskar. 2021. Cytology Image Analysis Techniques Toward Automation: Systematically Revisited. *ACM Comput. Surv.* 54, 3, Article 52 (April 2021), 41 pages. https://doi.org/10.1145/3447238

1 INTRODUCTION
Cancer, as of now, has become a soaring concern in individuals. Living organisms are composed of a large number of cells. Normal cells have a specific lifetime in which they reproduce and divide
to produce new cells, replacing old ones. Before the old cell is worn out, it passes its genetic information to new ones. These cell divisions are initiated by a control signal generated from a specific protein called cyclin. But cancer cells divide in a proliferative manner disobeying coordination among cells. Though cancer-causing genes are genetically inherited, various external factors like indulging in smoking and drinking, exposure to heavy metals, radiation, usage of plastics, and so on, are also responsible for growth of cancer cells. The cancer cells have several distinguishing features from benign cells that are extensively discussed in different types of cytology. More specifically, the study of tissues and cells using a microscope to detect malignancies or other inflammatory conditions is broadly categorized as cytology or cell biology.

Two major allied domains that need to be mentioned in this context are Hematology and Histopathology. Cytomorphology of hematological cells can be examined from (i) peripheral blood smear and (ii) bone marrow samples. The auto-analyzers in hematology are quite advanced without requiring blood smear preparation in the majority of cases. Blood smear analyzers, in addition to examining cellular details also investigate different cell lines such as Red Blood Cell (RBC), White Blood Cell (WBC), and platelets. In cell line diseases, a decrease or increase in standard reference of cell count with the change in cellular morphology (observed as abnormal cells) generally occur. A decrease in RBC count and size is usually called anaemia. Macrocytosis or an increase in RBC size occurs in megaloblastic anaemia. Poikilocytes and fragmented RBC are seen in hemolytic anaemia. An increase in WBC is usually indicative of any infection or hematological malignancies like leukaemia. Bone marrow examination is a very important part of hematology. As blood cells are manufactured in bone marrow, bone marrow examination gives a more detailed picture of the hematological disorder. There are definite criteria formed by the World Health Organization (WHO) for diagnosing different subtypes of hematological malignancies, e.g., blast cells should be more than 20% in peripheral blood smear or bone marrow for diagnosing acute leukaemia.

Histopathology, another well-known technique in detecting carcinoma, gives information of tissue for immuno-histological and molecular analysis for targeted cancer chemotherapy. Biopsy examination or histopathology, depends on many other factors like background architecture etc. In breast cancer, Fine Needle Aspiration Cytology (FNAC) is pretty good in diagnosing carcinoma and benign fibroadenomas, but histopathology is needed in some borderline and stromal lesions. In thyroid cell carcinoma, FNAC can diagnose almost all types of goitre, thyroiditis, papillary carcinoma, medullary carcinoma etc., and histopathology is necessary to differentiate follicular adenoma and carcinoma. In lung cancer, FNAC is suited to detect infections and a primary diagnosis of cancers whereas biopsy of lung is a difficult and risky procedure as there are chances of bleeding and haemorrhage in pneumothorax (entry of air in pleural space). Some solid organs are not recommended for FNAC, e.g., brain, heart, prostate gland, and so on. Prostate sampling using histopathology is dominant over FNAC, as cellular morphology is not markedly different in benign and malignant. In certain tests, cytology is sensitive to potentially high-grade tumors and when combined with other modalities of testing like histopathology, it can increase the sensitivity of detection to a greater extent. Several challenges associated with cytology images, more specifically solid organ cytology make the overall diagnosis process difficult even by experienced cytologists. The challenge percolates from the nuclear shape, size, density to diverse nature of cytoplasm, background debris, and so on. Throughout the article, we have used the term cytology to indicate solid organ cytology.

Primary screening of images from slides can be automated using image processing algorithms so that experts can concentrate only on areas of interest. Various techniques to develop automatic or semi-automatic systems for assisting cytologists in the diagnosis procedure can be found in the literature. But very few of them are commercially successful. Also, it is observed that meagre attention is given to cytology compared to histopathology.
Fig. 1. Highlights of the present survey.

despite its potential to classify a malignant cell in the least invasive manner. Some survey papers [28, 107, 140, 145, 153, 194, 197] address cytology image segmentation and classification methodologies for a single domain. This has led us to focus on different types of cytology-based research works to instantiate advances in cytology. So, in the present manuscript, we review a comprehensive list of 17 types of solid organ cytology based on their sites of origin. We consider diverse segmentation and classification algorithms with associated challenges in meeting higher performance requirements, outline techniques to address common bottlenecks, the correlation between them and provide meaningful insights on techniques for future developments in the automation of the diagnosis process. The research works over the past 30 years are systematically and extensively reviewed so that one can have a deeper insight into the evolution of methodologies.

The highlights of the present survey are given in Figure 1. In Section 2, we provide a brief introduction to cytology. In Section 3, we discuss automation in cytology and its underlying principles of image analysis. In Section 4, we discuss briefly about 17 types of cytology and undertake a concise and comprehensive survey on three domains viz. cervical, breast, and respiratory tract cytology, where extensive research works on advanced techniques of segmentation and classification of cytology images are reported. Rests are discussed in the miscellaneous section. In Section 5, we mention state-of-the-art systems to comprehend overall progress in respective domains. A separate section based on the authors’ view is presented to manifest overall progress in cytology image analysis techniques.

2 CYTOLOGY: A BRIEF INTRODUCTION

The body of an organism is composed of trillions of cells. Each cell possesses a cytoplasm and a nucleus. Cytoplasm, or the cell body, acts as an envelope to the nucleus containing chromosomes, the genetic material that undergoes mutation under certain changes. The changes are reflected in the morphology of the nucleus and cytoplasm. Thus, a micro examination of cells can unfold relevant information on any morphological alteration in response to a particular disease. These cells, which are important predictors of pre-malignant and malignant lesions can be sampled and examined under a microscope to diagnose different medical conditions except a few. FNAC is a procedure to primarily assess the cause of swelling by examining cells in it. The doctor who
performs the process usually fixes the swelling with one hand and introduces a 24-gauge needle with a 10-ml syringe attached to it. By creating a negative suction the needle is moved within the swelling. The cellular aspirate comes within the hub of the needle. It is then expressed on glass slides and smears are made with another glass slide called spreader. The slide containing material is then processed by two methods: Air drying and staining with May-Grunwald-Giemsa (MGG) stain and alcohol fixation using Papanicolaou or Pap stain. This process needs expertise as exact swelling is to be targeted without damaging the surrounding vital structure. Also, the cellular material should come out with less admixture of blood. FNAC is effective for the primary evaluation of any swelling as well as a safe procedure with very little pain. The three preparatory steps, i.e., Specimen collection, Slide Preparation, and Fixation are described in details in the supplementary section.

In benign cells, cellular materials are generally well-defined, exhibiting uniform chromatin distribution within the nucleus and a prominent cytoplasm. The nucleus boundary is usually regular and nearly elliptical in shape. Malignant cells, however, possess irregular nuclear boundaries, scanty cytoplasm and multiple nuclei with unusual sizes. To classify a specimen, these key characteristics are usually considered.

Two types of cytology tests are undertaken to determine the presence of carcinoma and if detected, figures out the extent of disease. A screening test may be performed at regular intervals to ensure the presence of disease. It is helpful to detect the disease at an infant stage when the response to the treatment is faster and as a result chances of fatality are reduced. Such screening is generally recommended for those who are highly prone to this disease. A diagnostic test is done when signs or symptoms are more common. This test is carried out to check the presence of disease and if so, aimed at precise identification of the specific condition of the disease.

### 3 PRINCIPLES OF CYTOLOGY IMAGE ANALYSIS

After a suitable slide is prepared, a screening or a diagnostic test can be performed for clinical analysis of the specimen. It can be done either manually or through an automated CAD-based system. A manual screening procedure directly involves an expert cytotechnologist to review/examine the slides. Whereas, in an automatic screening system, few steps are handled by a computer using some image processing protocols. For automated systems, the result is predicted by a locally or any remotely installed device without human counterpart. A pictorial representation of manual and automatic screening systems is given in Figure 2. Certain factors hinder the proper interpretation of slide leading to an erroneous result such as (a) Inexperience of pathologists about collection of the specimen from the exact location of the tumor causing sample inadequacy, (b) Poor preparation of slide that may contain excessive blood, mucus, inflammatory cells, background necrosis or certain foreign particles and obscuring materials giving an overall unclear and cloudy appearance, (c) Poor staining, and (d) Poor slide preservation. All these factors give rise to sampling error and purging of these errors is a prerequisite for the exact rendition of a sample.

#### 3.1 Automation as an Emerging Concept in Cytology

Since long back, image analysis using visual interpretation has been the primary essence in cytology image analysis. It was solely based on the examination of cells under a microscope by skilled practitioners. The objective behind the soaring trend of introducing CAD-based systems is to reduce time of the report generation process. Also, factors like reducing the workload of cytologists, human-induced errors, batch processing, and so on, bolstered motivations toward automation. Advantages of an automated screening system can be summarized as follows:
Fig. 2. A pictorial representation of manual and automatic malignancy screening systems for cytology images.

- It can act as a pre-screening system. To eradicate false-negative cases such systems exhibit high false positives so that, no single abnormal specimen that demands further investigation are disregarded. Thus, it acts as an adjunct to cytologists by eliminating the need to evaluate normal specimens, thus saving time and energy as well.
- It may run as a parallel system with the present manual screening system. Thus, the diagnosis process can be bi-directed to produce fewer chances of errors at specimen level.
- It promotes the development of web-based remote diagnosis system to facilitate low-resource clinical settings.
- It can be used as a rapid mass screening system for early detection of cancer.

The immediate impact of automation in cytology that has been realized by the entire community of doctors is that it allows batch processing and eliminates the manual tiring and time-consuming process. During the early 1950s, the idea to develop “Cytoanalyzer” \[34\], an automated scanning device, was the initial breakthrough to the concept of manual examination of slides. This could distinguish normal cells from abnormal cells using the information of nuclear size and optical density. However, the system suffered from two major disadvantages. First, the process was slow due to the low processing powers of computers at that time. Second, overlapping of cells and air drying factors of conventional pap smears complicated the identification procedure further with higher false-positive rates.

Atrophic changes in conventional smears of older patients added further hurdles to properly classify images by using automated analysis. So, there were several issues to conventional pap smear slides that hindered the correct analysis by CAD systems. To mitigate the issues, Zahniser et al. \[210\] came up with the idea of liquid cell suspension to make “a better pap test” of the cervical sample. They used Feulgen stain and the software developed in conjunction with liquid-based specimens gave better results with a dramatic decrease in false-negative rates to 1% and false-positive rates to 10%.

Another approach to the successful implementation of an automated screening system was to incorporate more shape and texture features of nuclei and cytoplasm to achieve more accuracy and robustness. This idea was given a shape by Toshiba with the name of CYBEST \[189\]. It was a prototype developed in 1972 with further development in a new desk size design in 1981. AutoPap 300 and SurePath were two automated screening systems that were marketed successfully.
Despite the limited success achieved so far with the existing systems, there were constant efforts to upgrade existing automated screening systems and design the new ones on a large scale. There were a handful of designs developed during the 1980s like CERVISCAN [182], LEYTAS [2], FAZYTAN [40], BioPEPR [128], and so on.

In cytology laboratories there are three major sections where automations are extremely pertinent and devices are currently accessible. They are summarized as follows:

- **Specimen preparation devices**: Two Food and Drug Administration (FDA) approved automated systems, namely, ThinPrep Processor and the AutoCyte Prep serve the purpose popularly. ThinPrep processor 5,000 uses thin prep technology, for cell dispersion, collection and transfer and is able to process approximately 35 slides/h. The preparation of monolayer slides in AutoCyte prep is done using liquid-based preparations.
- **Manual screening adjunctive devices**: These devices are used either independently or in conjunction with human for focusing the portion of slides having abnormalities. Thus, whole slide examination is not required. Some of the computerized FDA approved microscopes such as CompuCyte’s “M Pathfinder,” Accumed International’s “AcCell 2000” are widely used to mark abnormal cells. Zeiss Matscope is also an advanced integrated microscope that has a 360 degree rotating analyzer that uses extended-depth-of-field and multiple phase information to reduce technical errors. Hamamatsu has an automatic whole slide scanner that allows thick slides to be analysed using zoom adjusting technology giving sample clarity.
- **Automated screening devices**: AutoPap screener, a screening system is designed to select the portion of the slide containing abnormalities without going through whole slide. AutoCyte is designed to provide computer decisions thus reducing the workload of cytologists to a large extent.

3.1.1 **Commercially Available Systems.** The motive behind the development of smart CAD systems is that they should be capable of handling a large amount of samples to reduce the workload of cytologists. The evolution of the automated devices for analysis of cytology images started in 1950. With the improvement in slide preparation techniques, some systems became obsolete and some are modified accordingly. New systems evolved with better techniques and high computation powers. There are a number of systems for cervical cytology, but for breast and respiratory tract cytology, very few systems are found. Unfortunately, most of the systems are not commercially available. Some of the popular commercially available systems are recorded in Table 1.

3.2 **Image Analysis Steps**

The journey toward automation has lead researchers to enunciate image processing algorithms for better understanding and analysis of cytology images. During analysis of a cytology specimen, some basic steps need to be followed. In manual screening system of cytology images, such steps include preparation of cytology specimen, feeding, fixing, and adjusting the slide with required magnification in microscope to identify the region of interest. After proper examination based on the nature of specimen, opinion is declared by experts. Thus, each of these steps involves a human effort. In automatic screening system, all the above steps required prior to image analysis can be automated or semi-automated, including image digitization, enhancement [50], segmentation of images, features extraction and finally providing diagnostic information through judicious use of classifiers [137]. Each of these steps, as mentioned in Figure 3, give rise to a number of challenging research problems that have been studied in great detail over the past half century, since achievements in the field took place. If the goal is to create an interactive system, then some of the difficulties can be left to human operator for solving, but when fully automated system is a goal,
### Table 1. Different Commercially Successful Cytology-based Systems

| Manufacturer & Year of Manufacture | Cancer pre-screening systems | Description | Assessment |
|------------------------------------|------------------------------|-------------|------------|
| Airborne Instruments Laboratory, Inc., N. Y., 1950 | Cytoanalyser | An electronic optical machine to detect abnormal pap smear cells. The machine consists of a scanning microscope, a computer and analyzer, and a recorder. The scanner examines the significant area of the smear and converts the optical information into an electric beam, which is passed to the computer and analyzer. | Two sets of experiments were conducted between 1958-1960 and results reveal high false-negative cases, which hindered making it commercially available [156, 180]. |
| Watanabe and Toshiba, 1972 | CYBEST | The pap smear screening system that uses object extraction techniques like thresholding and differential approaches. It uses morphological features; nuclear size, N/C ratio, optical density, nuclear shape, chromatin pattern for image analysis. The model was upgraded to model 2 (1974), model 3 (1976) and model 4 (1981) accordingly with improved designs and reduced turnaround time. | Field tests of CYBEST model 4 shows occurrences of greater false positive cases. CYBEST model 4 takes nearly 3 min per specimen to evaluate [172]. |
| DJ Zahniser, 1979 | BioPEPR system | It was designed to pre-screen cervical smear based on cellular morphology nuclear area, nuclear optical density, nuclear texture, and N/C ratio. Additional programs were designed to recognize artifacts, overlapping nuclei and leukocytes. | A high false alarm rate of 24%, restricted it’s commercial feasibility and also it’s efficiency was highly dependent on quality of smears. Analysis rate was 4 min/smear [57]. |
| BD Diagnostics (Sparks, MD) | Focal Point Slide Profiler (FDA approval in 2006) | It is an FDA approved system that screens the entire slide. A Focal Point Score was derived from the slide using a model based on different features such as nuclear size, contour, ratio with cytoplasm, integrated optical density. A score less than a threshold is separated into “Further Review” category. | For glandular abnormalities, FocalPoint screened slides need to be reviewed exhaustively regardless of the quintile ranking. Also it was not cost effective and required heavy technical maintenance [131]. |
| Tripath (1998) formed with three companies Neopath, Neuromedical and Autocyte | Autoprep 300 (FDA approved in 1998) | A first FDA approved rescreening system that infused neural network using morphological features into low level programming. It used two resolution levels: low resolution to map the specimen and high resolution for selecting ROI. Only suspicious samples detected was labeled as ‘Required Visual Inspection’. The image processing system was developed using specific system board ASIC. | The system is confirmed for 25% cases without further screening as normal cases. Rest of the cases are categorized into five different stages of abnormality [185]. |
| Hologic, Inc., taken over from Cytyc Corp, Marlborough, MA | ThinPrep—Integrated Imager (T-3000) FDA approval in 2018 | Designed for pap tests, it consists of a single desktop system using an imaging station and microscope and directs technicians to examine only potentially abnormal areas. | Increased throughput by detecting only suspicious cells without complete manual slide review. It reviews a slide in ~90 seconds [60]. |
| Neuromedical Sciences Inc (NSI) Late 1990s | PAPNET Systems | A rescreening system that infused neural network into low level programming. It used two resolution levels: low resolution to map the specimen and high resolution for selecting ROI. Only suspicious samples detected was labeled as ‘Required Visual Inspection’. | The system is not cost effective and does not give hint of malignancy either. In the end of 1999, NSI was exhausted of capitals and declared an economic failure to run the project [87]. |
| C-DAC along with the RCC, Trivandrum | CerviSCAN | This pap smear screening system has a Piezo server controller connected to a microscope. It captures the information of a nucleus at various foci and stacks the information for every nuclei. Creating image stack required expert intervention and huge memory. Also, collecting information of nucleus at various focus was hindered by background debris [182]. | |
| W.N Street, 1990, University of Wisconsin Madison. | “Xcyt project” | A breast cancer diagnosis and prognosis system. | More false negative cases [165]. |
| Roger A. Kemp 2007 | LungSign test | LungSign is a fully automated system for analysis of sputum specimens slides automatically. Using cellular morphology it produces a score for individual specimen. | LungSign is an effective tool for detecting stage 1 cancer and needs to be upgraded for high stage cancers [82]. |

ACM Computing Surveys, Vol. 54, No. 3, Article 52. Publication date: April 2021.
Fig. 3. Image analysis steps of a typical segmentation-based image processing system, where the intermediate stages denote processing techniques.

all steps need to be addressed. In the next section, we focus only on automation techniques for image analysis.

4 DIFFERENT TYPES OF CYTOLOGY

There are different types of cytology based on their site of origin in human body. In Table 2, we have summarized 17 such types of cytology specifying their modalities of specimen collection. It has been observed that most of the automated approaches are devised for screening of Cervical, Breast and Respiratory tract cytology probably due to lack of standard data in other domains of cytology. Thus, three domains are discussed in separate sections. In the next section, we emphasized on works that throw light on succeeding works to evolve to better and robust algorithms. This will help the reader to gain a quick realization on research updates in respective domains. Images of different cytology are shown in Figures 4, 5, 6, and 7 to illustrate the diverse nature of specimens.

4.1 Cervical Cell Cytology

According to the latest report by WHO [192], cervical cancer ranks fourth among cancers frequently found in women. It originates body. The origin of cervical cancer is human papillomavirus (HPV) that causes abnormal growth of cells of cervix (cervical dysplasia). In earlier stages, there
Table 2. Different Types of Cytology [77, 143] with Their Modalities and Corresponding Overall Performances in Detecting Carcinoma

| SL# | Cytology | Used to detect | Modality of specimen collection | Accuracy (Acc)/Sensitivity(Sen)/Specificity(Spe) | Cytomorphology of benign cells (General characteristics irrespective of sub types) | Cytomorphology of malignant cells (General characteristics irrespective of sub types) | Limitations of cytology |
|-----|----------|----------------|---------------------------------|-----------------------------------------------|------------------------------------------------------------------|-----------------------------------------------------------------|-----------------------------------------------|
| 1   | Adrenal gland | To detect adenoma node | Adrenal FNA | Acc: 96%-98% Sen: 87% Spe: 100% | a) Many naked nuclei. b) Granular background. c) Intact cells with fuzzy cytoplasm. | a) Isolated cells with dense cytoplasm. b) Pronounced nuclear atypia. c) Mitoses. d) Background necrosis. | a) For nodules < 3cm diagnostic accuracy drops |
| 2   | Breast | To detect breast tumors | Nipple discharge(ND). a) Ductal lavage using FNA. | ND: Sen:41% - 60% FNA: Sen: > 97% Spe: > 97% | Round to oval monomorphic epithelial cell and adherent dark nucleus. (See Figure 7.i.a) | a) Large, angulated, eccentric, pleomorphic; nuclei with irregular boundaries. b) Irregular spacing. c) Background debris. (See Figure 7.i.b) | a) Nipple discharge cytology is not very effective. b) It is performed to very few patients who are asymptomatic. |
| 3   | Cerebrospinal fluid | To detect malignancy | Lumbar puncture by using needle in intervertebral space at L3 to L4 or L4 to L5. | Se:60% Spe: 100% | a) Sparsely cellular b) Round nuclei. c) Uniform nuclear contours. | a) Irregular nuclear contours b) Prominent nucleus c)Abnormal chromatin pattern | a) Additional processing is needed to preserve the specimen |
| 4   | Cervical cytology | Cervical cancer | Brush or stent cytology | Pap Smear Se:75.80% Sp:98.05% LBC (Liquid Based cytology) Se:76.67% 98.93% | Large polygonal cells with small round nuclei and large amount of eosinophilic cytoplasm. (See Figure 7.i.a) | Malignant melanoma: a) Clusters of small cells. b) Scarctyptoplasm c) Nuclear hyperchromasia. (See Figure 7.i.b) | a) False positive cases are more. b) Less useful for cystic and fibrotic lesion. |
| 5   | Gastrointestinal cytology | Gastrointestinal tract lesions | Brush cytology | Se:77%-94% Spe:95% | a) Uniformly arranged cellular structure. b) Clear nuclear boundary. | a) Isolated cells and crowded groups. b) Nuclear pleomorphism c) Hyperchromasia. | a) Sampling error b) Accuracy compromised in detecting Barrett’s esophagus. c) High false positive cases. |
| 6   | Kidney | Renal lesions | Renal FNA | Acc:73%-94% Sen:92% Spe:94% | a) Large dense globular structures. b) Proximal tubular cells are rare. | a) Large cells. b) Large nuclei. c) Abundant granular cytoplasm | a) FNAC has low sensitivity for small masses (<5cm) and complex cysts and large masses >5cm |
| 7   | Liver | Focal lesions of liver | FNA | Se:67%-100% Sp:100% Acc:90%-94% | Hepatocytes are large, polygonal, isolated cells with prominent nucleolus and binucleated. They are centrally placed and round to oval in shape. (See Figure 4.i.a) | a) Isolated cells or cells in nest. b) Increased nuclei to cytoplasm ratio. | a) Rare hemorrhage, pain, bile peritonitis. b) Also, cannot typify the tumors. |
| 8   | Lymph nodes | To confirm enlarged lymph nodes | FNA | Se:80% Spe: 90% Acc: 95% | a) Dispersed isolated cell pattern. b) Presence of lymphoglandular bodies. | a) Polymorphous population. b) Mant cells present. (See Figure 6.i.a) | a) Sampling error b) Cell vascular patterns are lost to some extent. |
| 9   | Ovarian lesions | Cystic ovarian masses | FNA | Self for borderline analysis: 84% - 93% Se: 26 - 40% | a) Sparsely to highly cellular. b) Cells possess round nuclei with coarsely granular chromatin with one or two small nucleoli. | a) Clusters and isolated cells. b) Large pleomorphic cells. c) Round nuclei with prominent nucleoli. (See Figure 6.i.a) | a) Higher false negative rates. b) Lesser reliability in distinguishing borderline tumor and carcinoma. c) Benign lesions not histologically categorized. |
| 10  | Pancreas and biliary tree | Pancreatic mass and a duct structure | FNA and Duct Brushings | FNA: Se: 86%-98% Spe: 100% | a) Round or oval nucleus. b) Evenly distributed nucleus. c) Finely granulated chromatin. d) Cytoplasm boundaries well defined. | a) Moderate to high cellularity. b) Shadowed sheets of disordered negative cases. d) Ductal cells. e) Irregular nuclear contours. f) Enlarged nucleoli. e) Irregular chromatin distribution. | a) Sampling error |
| 11  | Peritoneal washing | Spread of cancer in peritoneal surface | Washing cytology | F.N:51%-86% F.P: <3% | a) Isolated mesothelial cells in sheets, often folded. b) Pale chromatin, c) Small round or oval nuclei. | a) Cell clusters and isolated cells. b) Enlarged nuclei, coarse chromatin. c) Nuclear pleomorphism d) Scanty or abundant and vacuolated; cytoplasm. | a) Sampling error |

(Continued)
| SL# | Cytology | Used to detect | Modality of specimen collection | Accuracy (Area/Sensitivity/Specificity) | Cytomorphology of benign cells (General characteristics irrespective of sub types) | Cytomorphology of malignant cells (General characteristics irrespective of sub types) | Limitations of cytology |
|-----|----------|---------------|-------------------------------|-------------------------------------|---------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------|
| 12  | Salivary gland | To assess the fluid in diseased condition | FNA | Sp: High Se:71% F.P:1% FN: High | a) Numerous cells often dispersed. b) Round nucleus with one nucleolus. Dense cytoplasm with clear boundary. (See Figure 5.a.a) | a) Numerous large clusters. b) Presence of cell block sections. c) Second population (See Figure 5.a.b) | a) False findings are high. b) Immunocytochemistry can act as companion to improve diagnostic accuracy. |
| 13  | Respiratory tract | Abnormalities in respiratory tract | Sputum, Bronchial brushing, Transbronchial FNA, Trans esophageal FNA | Sputum: Sp:72% Se:75% Transbronchial FNA:bronchial brushing: Se:91% Sp: 96%–99% Acc:80%–85% | a) Intercalated flat sheets and tubules of duct cells arranged uniform, small–shaped consisting of dense sacc cytoplasm and uniform nuclei. b) Sparsely cellular with round clusters of serous cells. | a) Polygonal /round/ fibre like cells. b) Abundant dense and smooth cytoplasm filled with keratin. c) Small hyper–chromatic nuclei. d) Inconspicuous nucleoli. | a) Accuracy of sputum cytology is low and susceptible to location in malignancy due to less epithelial cells. b) Fails to subclassify malignancies in adenocarcinoma. |
| 14  | Salivary gland | Salivary gland lesions | FNA | Se: >90% Sp: >90% | a) Intergalated/flat sheets and tubules of duct cells arranged uniform, small–shaped consisting of dense sacc cytoplasm and uniform nuclei. b) Sparsely cellular with round clusters of serous cells. | a) Polygonal cells b) Abundant granular or vaculated cytoplasm. c) Presence of prominent nucleoli. d) Background necrosis prominent. (See Figure 6) | a) Malignancies cannot be distinguished. |
| 15  | Soft tissue masses | soft tissue masses | FNA | Se: 95% Sp: 54%–98% Acc: >90% F.P: 0.5%–5% FN: 2%–15% | a) Small bland nuclei without atypia. b) Uninvaculated nuclei of uniform size. | a) Hypercellular smear. b) Pleomorphism. c) Mitoses and background necrosis present. | a) Intermediary conditions between benign and malignant are tough to determine. |
| 16  | Thyroid cytology | Presence of thyroid nodule | FNA | Se: 98% Sp: 84% Acc: 97.5% | a) Sparsely cellular. b) Uniform and evenly spaced follicular cells. c) Coarse chromatin pattern. (See Figure 4.1.a) | a) Enlarged nucleus. b) Nucleus oval or elongated in shape with irregular contours, pale in color. Overlapped nuclei. (See Figure 4.1.b) | a) Sampling error |
| 17  | Urine bladder washing | Bladder cancer | Depends on urine specimen type: Vouled, catheterized bladder washing, ileal loop | Catherized urine: Sp:95%–100% Se: 75% F.P:1.3%–15% | a) Presence of crystals b) Enlarged nucleus with prominent nucleolus. c) Coarse, vaculated cytoplasm. (See Figure 5.1.a) | a) Round dark nuclei b) Background necrosis (See Figure 5.1.b) | a) Distinguishing characteristics between high grade lesions from low grade lesions are ambiguous. |

Table 2. Continued

is hardly any symptom. In advanced stage, symptoms like abdominal pain, vaginal bleeding, abnormal watery discharge, bleeding after menopause, and so on, are commonly observed. There are mainly four types of cervical cancer: (a) **Squamous cell carcinomas**: A commonly known form of malignant tumor of cervix and accounts for nearly 85% of the overall cervical cancer in women of age group 40–55 years. These cells lie on the outer surface of the cervix and are thin and flat shaped. (b) **Adenocarcinomas**: It originates in glandular cells that line the upper area of cervix and accounts to 10% of overall cases of cervical cancer. Mean age of patients is 55 years. (c) **Adenosquamous carcinomas**: It consists of both the squamous and glandular cells and accounts for nearly 2–3% of cervical cancers. (d) **Small cell carcinomas**: Generally very aggressive in nature and rarely found (2–3%), which corresponds to stage IV of the cervical cancer. The cytomorphology exhibits a nesting pattern and often arranged in sheets.

4.1.1 **Modalities of Cervical Specimen Collection.** Cervical smears are normally obtained using a spatula and brush and categorized under brush cytology. A plastic spatula is rotated 360 degrees and the samples are exfoliated from ectocervix and endocervix. The sample is smeared
on one half of the slide. The slide is immediately spray fixed to avoid any air drying artifact that can cause distortion of cells. This technique is known as Papanicolaou test or Pap smear test.

4.1.2 Differential Characteristics Associated with Cytomorphology of Malignant Cells. Malignant cells of cervix are generally found in clusters of small cells with scant cytoplasm as shown in Figure 7(ii). Nuclear hyperchromasia and nuclear membrane irregularity are also commonly observed. These characteristics are crucial in categorizing the specimen into benign and malignant using both manual and automated analysis.

4.1.3 Significant Works on Automated Screening of Cervical Cell. Malignant nuclei differ significantly from benign nuclei in shape, size, and textural pattern. Thus, nuclei have sufficient diagnostic information embedded and possess enormous discriminative power to classify various stages of cancer [7]. Thus, accurate segmentation of nuclei is of prime importance for researchers,
working on cytology images. We categorize the review of works based on two approaches as stated previously: (i) Segmentation-based approach, (ii) Segmentation-free approach.

(i) Segmentation-based Approach: A good segmentation algorithm is the necessary footstep for extracting nuclei features and subsequent classification of cytology images. Various approaches to segment cervical cytology images are discussed below.

Thresholding-based: This is one of the common and simplest techniques for extraction of foreground pixels. Walker [186] cited some morphological operations based on octagonal structuring element. Coarse segmentation of gray-scale images was done using global thresholding. A closing morphological structural element of size smaller than the smallest nucleus removed cytoplasm. The nuclear heterogeneity was adjusted by an opening operator. But the method could not be fully automated. Zhang et al. [216] proposed an adaptive thresholding method where the RGB image was converted into HSV image. V-channel was extracted for histogram stretching. In the preprocessing stage a median filter was used to remove noise in the image. The issue of non-uniform illumination was mitigated using adaptive thresholding. An optimized selection of size of the neighborhood and an offset value ($\theta = 15$) was the major hurdle to segment clustered nuclei. Therefore, a concave point-based method was adopted with significant reduction in computational time.

Region-based: Region-based segmentation techniques seek to divide the image into sub images based on some homogeneity conditions. Mat-Isa et al. [108] proposed a combination of moving k-Means and modified seed-based region growing (MSBRG) algorithm to extract the region of interest (ROI) from the image. Each pixel was assigned to the nearest cluster point using a suitable threshold using Moving k-Means algorithm. Based on calculated thresholds and dynamic seed points, nucleus and cytoplasm edges were delineated using MSBRG. Thin prep images are generally afflicted by various factors like low contrast, haziness, noise, and so on, resulting in increased false negative cases. Mustafa et al. [116] proposed to enhance the contrast of such images via three methods, i.e., linear contrast, non-linear bright and dark contrast. An MSBRG method was used to segment nuclei. The later two contrast enhancement methods were used to increase the contrast of the bright areas and dark areas, respectively. In another work, Mat-Isa et al. [109] proposed a CAD system on region-growing-based features extraction (RGBFE) technique. It is composed of (a) an automatic feature extraction system to extract features like nucleus size, grey level intensity of nucleus, size and grey level intensity of the cytoplasm using region growing and (b) a diagnostic system using hierarchical hybrid multilayered perceptron (HHMLP) model. An accuracy of 97.50% was reported on a dataset consisting of 550 reported cases; 211 normal, 143 Low-grade Squamous Intra-epithelial Lesion (LSIL), and 196 High-grade Squamous Intra-epithelial Neoplasia (HSIL).

Contour-based: Active contour or snake model (ACM) seek to segment the image using initial contour information. But ACM fails to properly segment poor quality images and is also sensitive to the detection of initial contour. Snake model is one of the popularly used contour models for segmentation of nuclei from cytoplasm and background. Yang-Mao et al. [209] proposed an edge enhancement nucleus and cytoplasm contour detector (EENCC) algorithm involving five steps: trim-meaning filter, bi-group enhancer, computation of gradient, Maximum Valued Difference (MVD), and contour extraction. EENCC performed better than snake models with least relative distance error (RDE) of 0.288. Bamford et al. [6] prescribed an improved snake model for overcoming the initialization problem of conventional snake. Harand et al. [58] adopted a geometric active contour algorithm followed by Sauvola thresholding and FIFO queue structure to identify boundary of nucleus in each cell. The contour obtained in low resolution is taken as the input. It showed better results compared to the EENCC detector. Guan et al. [53] attempted to segment partially overlapped cells in high resolution images. A dynamic sparse contour algorithm was used to detect weak contour points in the nucleus. Gradient Vector Flow (GVF) snake model was applied to find
out the exact contour of cell. However, the method could not handle more than two overlapped cells. Moreover, its performance degrades with low resolution images.

**Deformable Contour model-based:** The Active Contour Model (ACM) pioneered by Xu et al. [206] is extensively used to extract objects with irregular boundaries, contour with noisy interior and small breaches. Tsai et al. [181] developed a three-staged cytoplasm and Nucleus Contour (CNC) detection technique for segmentation of nuclei and cytoplasm. A bi-group enhancer in first stage was used to sharpen the contour of nucleus by increasing the contrast between nucleus and cytoplasm regions. K-means clustering to detect cytoplasm contour was applied in second stage. In the last stage, contour of the nucleus was detected by applying Maximal Color Difference (MCD) method. But a prior denoising by means of median filter is a prerequisite to acquire a satisfactory result. Li et al. [97] used spatial K-means clustering algorithm to segment the image into background, nucleus, and cytoplasm for free lying cells. A radiating gradient vector flow (RGVF)-based snake algorithm was applied to get fine segmentation. RGVF involved a stack-based refinement and an edge map-based technique to identify unclear and incomplete edges, but eventually failed to segment overlapped cells. Also, the performance was highly dependent on accuracy of initial contour extraction. Bergmeier et al. [12] proposed a web-based software module with a prior noise removal technique using mean-shift and median filtering. A randomized Hough transform was used to detect candidate nucleus that was segmented by level set methods. The experiment was done on 207 cervical images with an F-measure of 96.15% and TP rate of 95.63%. But ROI was selected manually by cytologists. Lu et al. [101] proposed segmentation of non-overlapped nuclei using Maximally Stable Extremal Regions (MSER), where cell clumps were taken as input. To segment overlapped cells they used joint level set optimization. But, considering shape of the nucleus as ellipse to model the segmentation problem, the accuracy level was compromised to some extent. Also, cytoplasm boundary for low contrast images could not be properly delineated. To overcome this problem, Nosrat et al. [124] assumed a star shape apriori instead of ellipse, for better approximation of nuclei shape, which was encoded in a vibrational framework having directional derivatives. To address overlapped cells, they included Voronoi energy term, which bounded the amount of overlapping in adjacent cytoplasm. Wang et al. [188] proposed an FCM-based cell-clump detection using depth-first strategy model exploiting depth information of cells on ISBI-2015 dataset. Later fine segmentation was done using level set method. But it is hard to generalize the method for other standard datasets that lack depth information.

**Texture-based:** Texture is an intrinsic property of an image that is used to extract useful contents from an image [5]. Texture analysis using Local Binary Pattern (LBP) and Gray Level Co-Occurrence Matrix (GLCM) were chosen among seven sets of texture features in Reference [105].

**Graph-based:** A graph-cut approach was advised by Zhang et al. [217] to segment cervical cells of manual liquid-based cytology. A multi-way graph-cut approach was used to segment cytoplasm. This technique grossly classified the image into background (with lowest mean intensity) and cytoplasm, nuclei, inflammatory cells and debris. For precise nuclei segmentation, they used graph-cut approach adaptively and locally. To split touching nuclei, two concave point-based algorithms were combined. But it failed to detect cytoplasm boundary. Also, nucleus having poor contrast could not be delineated completely. For segmentation of nucleus Zhang et al. [218] initially defined a center point and size of the nucleus. Graph search-based method was used where the shape information was plotted as a graph (Cartesian to polar plot) by designing a cost function based on edge term and region term. Optimal path of the graph was dictated by dynamic programming. To obtain closed contour of nucleus in Cartesian coordinate system, reverse mapping was done. Proposed method was tested on two standard datasets: Herlev and HEMLBC. It showed faster optimization but the accuracy depends greatly on the closeness of initially detected nucleus center point to the actual one.
Clustering-based: Kim et al. [86] suggested a Fuzzy C-Means (FCM)-based segmentation on uterine cervical images using HSIC color space. A patch-based FCM clustering was proposed by Chankong et al. [23] for segmentation of nucleus, cytoplasm and background. Six nuclei features were extracted to classify cervical cells using Bayesian classifier, Linear discriminant analysis (LDA), KNN, SVM, ANN. For a two-class problem, they bagged accuracies of 97.83%, 97.00% and 99.27%, respectively, on ERUDIT, LCH, and Herlev dataset using ANN. Zhao et al. [221] used a Markov Random Field (MRF)-based segmentation model after generating superpixels from cervical images. A gap search algorithm with reduced time complexity was used to label nucleus, cytoplasm and other components of the cell. Saha et al. [147] introduced a superpixel merging model for crisp segmentation of overlapped nucleus. The efficiency of the method is controlled by pairwise regional control threshold of SLIC (Simple Linear Iterative Clustering) superpixel. Also, missing nuclei markers lead to failures in superpixel generation process.

Deep Learning-based: Deep learning-based framework is one of the latest trends in many applications and being extensively used in segmentation of cytology images. Song et al. [163] developed a Convolutional Neural Network (CNN) combined with a superpixel-based segmentation framework. The Gaussian noise generated during image acquisition process was removed by trimmed mean filter. The cytoplasmic mask was extracted by high-dimensional Otsu thresholding method. A superpixel-based SLIC segmentation was applied on masked images. To extract features of superpixels, CNN was applied. Coarse nuclei segmentation was done to reduce overlapping of inflammatory cells. A new template was constructed to improve segmentation. The nuclei region detection accuracy was obtained as 94.50% using CNN. Liu et al. [100] proposed a segmentation method by Mask-Regional Convolutional Neural Network (Mask-RCNN) and Local Fully Connected Conditional Random Field (LFCCRF). The Mask-RCNN constructed by residual neural network-based Feature Pyramid Network (FPN) was modified using feature maps obtained from pixel level information of nuclei. A Mask-RCNN-based coarse segmentation of nuclei was done and ROI was obtained by increasing the bounding box. Further refinement was done with LFCCRF having intensity and position information of all pixels in that region. Phoulady et al. [132] proposed a CNN-based nucleus detection technique using iterative thresholding method. The seed point of nuclei was initially detected. CNN was trained with patches of nuclei. This approach showed a precision, recall and F-score of 0.861, 0.895, and 0.878, respectively, on the CERVIX93 dataset. Song et al. [162] proposed multi-scale convolutional network (MSCN) for initial segmentation of nuclei to extract multiscale feature vectors. The images was fine segmented using graph partitioning to get accurate result. When compared to raw pixel segmentation, superpixel-based segmentation of cytoplasm and nucleus gave an improved accuracy of 5.06% and 2.06%, respectively. Zhang et al. [214] developed a R-FCN embedded with custom sized modified ResNET (NET-22) for segmentation of cervical cells collected from Guangzhou LBP Medicine Science & Technology Co. Ltd.. Using the developed technique they achieved average precision of 93.2% with fourfold cross-validation on 2,400 cells after data-augmentation.

Other approaches: In Reference [174], Intersecting Cortical Model (ICM) was studied where parameters were optimized using Particle Swarm Optimization (PSO). Several methods like Otsu thresholding, Expectation Maximization (EM), Region Growing and Fuzzy C-Means clustering were evaluated on 250 cervical images. ICM was reported with the highest PSNR of 62.95 dB. An unsupervised segmentation of Pap smear cells was proposed by Happy et al. [56] in two stages. Cells were divided on homogeneity and circularity basis using multi-scale hierarchical segmentation. The nucleus and cytoplasm were classified using a binary classifier. Extended Depth of Field (EDF) images at different focal planes with various degrees of overlap were considered. Modified Otsu using class prior probability was used to select suitable threshold to segment nucleus. The work was validated on ISBI 2015 challenge dataset with a DICE score of 0.86 and a TP rate 0.88.
Lakshmi et al. [3] applied Haar wavelet to nullify the effect of uneven staining in pap-smear images followed by adaptive median filter to remove noise. ICM combined with cuckoo search algorithm segment the nucleus in an optimized way. However, initial parameters were chosen manually. Iliyasu et al. [69] investigated a hybrid approach of feature set selection. They combined quantum particle swarm optimization (QPSO) algorithm with Fuzzy k-nearest neighbours (Fuzzy KNN) algorithm to scale down 17 features relating color, geometry and texture of nucleus to 7 features. Selected features prior to hybrid approach-based classification showed better classification accuracy compared to All-feature-based approach (no feature selection).

(ii) Segmentation-free Approach: With the onset of deep learning-based techniques [64, 99], deep convolutional networks or fusion of different CNNs can directly classify pap-smear cells without requiring prior segmentation. To successfully train a CNN architecture, a large and diverse dataset is necessary. In Reference [14], VGG net-16 was used for feature extraction for subsequent classification using Least Square Support Vector. Zhang et al. [219] introduced DeepPap, a CNN model using transfer learning. A pre-train architecture of image net was used for classification of pap-smear cells. Recognition accuracies of 98.3% and 98.6% on Herlev and HEMLBC dataset, respectively, were observed using fivefold cross-validation. Hyeon et al. [67] used VGG-16 net, a pre-trained CNN architecture for feature extraction. 71,344 Pap smear images were collected from Seegene Medical Foundation to form a dataset consisting of 8373 samples of normal class and 8373 samples of abnormal class. Maximum F1-Score of 78.17% was observed using SVM on test samples. Khamparia et al. [83] developed a web-based application using CNN-based feature extractor showing maximum of 97.89% using Random Forest on features extracted using ResNet-50. A small-sized test data (116) was considered from training data of 934. In Reference [222], Zhao et al. extracted 160-dimensional features consisting of different properties of chromosome/nucleus morphology, chromatin pathology and region intensity. An optimal set of features was selected using a reinforced margin-based approach with heuristic knowledge. Classification was done on a combination of features with single-stage or two-stage classifiers and maximum accuracy of 97.88% on Herlev dataset was reported. A CNN fusion model was proposed by Hussain et al. [66] for a multi-class problem. A dataset combining the conventional and LBC along with Herlev dataset was developed to report maximum accuracy of 98.9%.

Classification: After segmentation of nucleus, different nucleus centric features are extracted to classify pap-smear cells into two classes benign and malignant. Malignant cells are further classified into different categories based on their stage. A single classifier or ensemble of classifiers are used. Among different classifiers, binary classifier [59], KNN [106], ANN and its modifications, [23, 109, 110], SVM [127, 215], FCM [86, 133], and so on, are popularly used. Deep learning techniques like [117, 167, 219] are recently gaining popularity for classification of cervical cells.

Binary Classification: Holmquist et al. [59] developed a binary classification framework based on Linear Discriminant Analysis (LDA) to distinguish cervical images using density, shape, and texture-oriented features of nuclei.

KNN: Marinakis et al. [106] proposed a genetic algorithm-based technique for selection of optimized nucleus features using K-NN. They obtained maximum accuracies of 98.14% and 96.95% for 2 class and 7 class problems, respectively, on Herlev dataset.

Hierarchical approach: Mehdi et al. [110] introduced a hierarchical approach to classify images into mild, moderate and severe cells using ANN with back propagation algorithm. Ramli et al. [142] applied a non-linear hybrid multi-layered perception (HMLP) model using least square algorithm to classify cervical cells into normal and low-grade squamous epithelium. Mat-Isa et al. [109] also recommended an HMLP network to classify single cell images using different feature information such as nucleus size and grey level, cytoplasm size and grey level, and so on.
Support Vector Machine (SVM): Zhang et al. [215] used SVM classifier to detect cancerous cells from multispectral pap smear images. During feature selection Sequential Backward Selection (SBS) was applied. They obtained better performance after using SVM-based screening with smaller number of features. Monteagudo et al. [127] proposed a combination of SVM and waterfall algorithm to classify cancerous and normal specimens. Dong et al. [38] used optimized cell features selected by Classification and Regression Trees (CART) with SVM whose parameters are optimized by PSO. An accuracy of 99.78% has been reported for tenfold cross-validation on Herlev dataset.

Fuzzy C-Means (FCM): FCM is a clustering approach, which is widely known to segment images, has more recently been used in classification of images [86, 195]. In Reference [133], different nucleus specific information such as circumference, different ratios of nucleus and cytoplasm, degree of roundness, and so on, were used to detect true nuclei points using different extrema points. FCM was preferred over SVM to classify cells due to the unavailability of large dataset.

Convolutional Neural Network: Gautam et al. [46] proposed a cell nuclei detection technique using a patch-based CNN. Segmented nuclei were classified using transfer learning on AlexNet. They also proposed a decision tree-based technique. Recognition accuracies of 99.3% in two-class problem and 93.75% on a seven class problem was reported. Wu et al. [203] proposed AlexNet for classification, which was trained using threefold cross-validation. For data augmentation flipping and rotation techniques were employed. Training and testing process was done on both original and augmented datasets consisting of keratinizing squamous, non-keratinizing squamous and basaloïd squamous. Recognition accuracies of 93.33% and 89.48% for original and augmented images, respectively. Meiquan et al. [111] proposed an automated cell detection technique using Resnet-101. After features were extracted, the data were enhanced by rotating patches at 90°, 180°, and 270°. Recognition accuracies of 0.91, 0.78, and 0.70 were reported on validation set, two-class and four-class problems, respectively. Kuko et al. [90] introduced a new cervical dataset consisting of 4,596 cells, which was first clustered based on morphological features and then classified using a deep learning-based approach. An accuracy of 91.63% was reported on fivefold cross-validation technique.

Some notable works on Cervical Cancer cytology are shown in Table 3.

4.2 BREAST Cancer Cytology

Breast cancer, though, mostly prevalent among women all over the world, occurrences found in men are rare (cases found are mostly invasive in nature). In women, the incidence and deaths due to breast cancer is the highest among all types of cancer, and thus posing a global burden irrespective of all levels of modernization. Breast cancers originate either in the lobules or in the ducts that connect lobules to the nipple. Breast cancer generally does not show symptoms (before stage III or IIIc) when the tumor is small enough to be felt and can be easily cured. It starts with a painless lump and slowly progresses with other symptoms like pain, breast heaviness, swelling, and redness of the skin. Nipple abnormalities such as spontaneous watery or bloody discharge or retraction are also common. Samples of cytology images of breast are shown in Figure 7(i). Breast cancer is generally categorized into (i) Carcinoma In Situ, (ii) Invasive.

(i) In Situ: Here, instances of abnormal cells are found locally and have not spread to nearby cells. These are mainly of two types: ductal carcinoma in situ (DCIS) and lobular carcinoma in situ (LCIS) or lobular neoplasia.

- Ductal carcinoma in situ. In this condition, abnormal cells take the place of normal epithelial cells that surround breast ducts. It does not necessarily progress toward invasive stage.
Table 3. Brief Overview of Different Works in Cervical Cancer Cytology Identification Where Accuracy (Acc), Jaccard Index (JI), Specificity (Sp), Sensitivity (Se), and Dice Similarity Coefficient (DSC) are Mentioned in the Corresponding Bucket

| SL# | Author’s name | Name/Source (N/S) and Number of samples (Size) in the Dataset | Method for segmentation image | Features of Nucleus and cytoplasm and classifier used | Quantitative results and Findings (Acc, JI, Sp, Se, DSC) |
|-----|---------------|-------------------------------------------------------------|-------------------------------|-----------------------------------------------------|-----------------------------------------------------|
| 1   | Holmquist et al. (1978) [59] | N/S: University Hospital (B.S.), Uppsala, Sweden Size: 729 single cells of eight different categories (4 normal and 4 malignant) | Nuclei of single cells are segmented using dual wavelength method | Features: Density, shape, and texture- oriented features of nuclei. Classifier: LDA-based binary classifier. | Acc: 97% in binary class. Findings: Algorithm does not deal with superimposed cells. |
| 2   | Mat-Isa et al. (2008) [109] | N/S: Kota Bharu Hospital and Universiti Sains Malaysia Hospital, Malaysia Size: 550 images of three class | Modified seed-based region-growing algorithm | Features: Nucleus size, nucleus grey level, cytoplasm size and cytoplasm grey level. Region-growing-based features extraction. Classifier: H2MLP. | Acc: 97.5 Sp: 100 Se: 96.67. Findings: Algorithm does not deal with superimposed cells. |
| 3   | Chankong et al. (2009) [21] | N/S: ERUDIT Dataset. Size: 276 single cell images. 138 normal cells (63 superficial cells, 75 intermediate cells), and 138 abnormal cells (34 light dysplastic cells, 34 moderate dysplastic cells, 35 severe dysplastic cells and 35 carcinoma in situ cells). | Fourier Transform | Features: Mean, variance, and entropy. Classifier: Bayesian classifier, linear discriminant analysis (LDA), KNN, ANN, and SVM | Acc: 92.65 (Highest SVM). False positives: 7.35%. False negatives: 7.35%. Findings: Experimented with single cells. Overlapping cells, which is the practical scenario, are not considered. |
| 4   | Li et al. 2012 [97] | N/S: Herlev dataset, Size: 917 images | Spatial K-means clustering followed by RGVF Snake incorporating a stack-based refinement using edge map. Classifier: interpretation visual | Zijdenbos similarity index: 94%. Findings: only segmentation of isolated cells. Overlapping of cells not considered. |
| 5   | Gencav et al. [47] | Herlev data set, Size: 917 images of Pap cell Hacettepe data set, Size: 82 Pap test images | Non-parametric hierarchical segmentation. | Features: 12 features of cells including spectral and shape. Classifier: Sum of posterior probabilities of Bayesian, Decision Tree and SVM. | Correct classification rate is 96 % Findings: Parameter-free segmentation. Clusters of overlapped regions not addressed. |
| 6   | Zhang et al. (2014) [217] | N/S: Shenzhen Sixth People’s Hospital Huazhong University of Science and Technology Union Shenzhen Hospital, Shenzhen, China, 2010. Size: 51 images (21 slides). | Local graph cut and global graph cut to delineate cytoplasm and nucleus. Concave point-based algorithms to segment overlapping nucleus. | Features: nucleus and cytoplasm texture, shape and size. Classifier: Visual interpretation (only segmentation is reported) | Acc: 93% for segmentation purpose. Findings: Touching nuclei are segmented. Poor contrast images are underestimated. |
| 7   | Nosrati and Hamarneh (2015) [124] | N/S: ISBI Overlapping Cervical Cytology Image Segmentation Challenge. Size: 135 images (45 images for training and 90 images for testing). | Star shape prior and Voronoi energy term. | Features: (HOG) features are used for training the nucleus to cytoplasm ratio. Classifier: Visual interpretation. | DSC: 0.88 Findings: In multiple overlapped regions the boundary of each cell is little ambiguous. Time to segment each image is 6.6 sec. (approx.). |
| 8   | Mariarputham and Stephen (2015) [105] | N/S: Herlev dataset Size: 917 images | Image inversion, binarization, morphological closing with structuring element of size 5 and filling operation. | Features: 24 features including nucleus related, moment, texture etc. Classifiers: SVM and Neural Network. | Acc: 81.85% (best SVM) for seven class problem on tenfold cross-validation. Findings: Dill with only single cell images. |

(Continued)
Table 3. Continued

| SL# | Author’s name | Name/Source (N/S) and Number of samples (Size) in the Dataset | Method for segmentation image | Features of Nucleus and cytoplasm and classifier used | Quantitative results and Findings (Acc, JI, Sp, Se, DSC) |
|-----|---------------|-------------------------------------------------------------|-------------------------------|-----------------------------------------------------|------------------------------------------------------|
| 9   | Zhang et al. (2017b) [219] | N/S: 1) Herlev dataset 2) H and E stained manual LBC Dataset Size: Herlev dataset (917), HEMLBC dataset (21) | No segmentation | Used previously trained CNN (trained on ImageNet) with 5 convolution layers and 3 forward connections layer. | On Herlev: Acc: 98.3%, Sp: 98.3% Se: 98.2% On HEMLBC: Acc: 98.6%, Sp: 99.0 % Se: 98.3% Findings: Multiple and overapper cells are not considered. Misclassification rate is higher. |

- Lobular carcinoma *in situ*. Under this situation, growth of abnormal cells extends to some lobules of the breast and often progresses to invasive cancer.

(ii) Invasive: It has two subtypes: (a) Regional stage: Abnormal cells have spread to neighbouring tissues and nearby lymph nodes (stage II or III cancers). (b) Distant stage: A condition in which abnormal cells have metastasized to different organs or lymph nodes above collarbone (stage IIIc and stage IV cancers).

4.2.1 Modalities of Breast Specimen Collection. Routine screening examination of breast cancer is always encouraged for early detection before symptoms actually start to divulge. This can be done by simply perceiving a lump. Two modalities of specimen collection are normally practised for breast cancer detection and diagnosis: (a) Discharge cytology and (b) FNAC. FNAC is used to evaluate palpable and non-palpable breast lesions. A fine-needle or wider core needle or a surgical incision picks requisite amount of mass from several points of the site for diagnosis. Then microscopic analysis of breast tissue is done by an expert to determine the extent of percolation of abnormal tissues. Nipple discharge cytology is usually performed to a few patients who are generally asymptomatic.

4.2.2 Differential Characteristics Associated with Cytomorphology of Malignant Cells.

- Large, angulated, eccentric, pleomorphic nuclei with irregular boundaries.
- Irregular spacing between adjacent nuclei.
- Background debris in large proportion.

4.2.3 Significant Works on Automated Screening of Breast Cell. New image processing techniques [123, 165] have evolved over decades to commensurate the complex nature of breast cytology images embedded with several degradations. We arrange related works in a similar fashion: (a) Segmentation-based approach, (b) Segmentation-free approach. Image preprocessing techniques such as color conversion, image normalization, contrast enhancement, and so on [145], are considered customary prior to any segmentation technique and not discussed under broad heading. *(i) Segmentation-based approach:* Nuclei are segmented to extract nucleus centric features to classify benign and malignant cells. But grabbing suitable image segmentation algorithms fit for the purpose is a challenge to the research community. Difficulty in segmentation arises mainly due to variable structural and textural pattern of the nucleus with various degradations of the...
specimen arising out of sampling error. Overlapping nuclei create another challenge to properly segment nucleus from the rest. In this section, we highlight some important and popularly used segmentation techniques with the aim to develop knowledge of the existing works under each subheading.

**Contour-based:** Active contours, since their introduction by Kass et al. [80] in 1988, have been used in subsequent studies [6, 200]. Wolberg et al. [199] used manual segmentation to define nucleus boundary. For precise nucleus contour detection, they invoked snake model that can be confined to the nucleus region using proper energy function. Based on 11 features of isolated nuclei like size, perimeter, area, compactness, radial variance, concavity, texture, and so on, images were classified using Multisurface Method (MSM)-Tree classifier. Wolberg et al. [201], reported an accuracy of 97.5% for non-overlapped nuclei segmentation on test dataset. Usage of compact Hough transform for nuclei segmentation [113] and generalized Hough transform with deformable models [95] are found in literature. Street et al. [165] used Hough transform to find circle like structures followed by an active contouring technique. They proposed an automatic diagnostic and prognosis system "Xcyt" [165] for screening of breast cancer. Hough transform was also adopted by Hrebien et al. [62] followed by an automatic nuclei localization method based on (1+1) search strategy. To segment nuclei, a combination of active contour model, watershed and grow-cut algorithm was used. But the technique proved to be ineffective for overlapping nuclei. Also, false circles were created that could not be resolved.

**Texture-based approach:** Wavelet-based decomposition has proved to be a powerful tool in analyzing texture or chromatin pattern of nucleus [74, 191]. Weyn et al. [191], used wavelet as chromatin pattern descriptor for semi-automated diagnosis and grading of breast tumor. However, determination of tumor stage is hindered by increased false negative cases. The effect of Log-Gabor wavelet filter on HSV color space were investigated in Reference [123]. Color wavelet features were deducted on extracted features and compared the relative performances of classifiers, viz. SVM, Naive Bayes, and ANN. Highest accuracy accorded by SVM as 98.3% with sensitivity and specificity of 98% and 98.6%, respectively. Since DWT lacks phase information, complex wavelet transform was grabbed in follow up studies [20]. In Reference [73], Niwas et al. analyzed nuclear chromatin pattern using complex Daubechis Wavelets. Wavelet co-occurrence matrix were used to calculate statistical features like cluster shade and prominence, contrast, entropy, energy, local homogeneity and maximum probability K-NN with standard Euclidean distance was used to classify images.

**Region-based:** Marker controlled watershed segmentation was studied by Yang et al. [208]. Hrebien et al. [62] proposed nucleus segmentation technique using watershed, active contour and cellular automata grow cut techniques. They reported a segmentation accuracy of 68.74% but taking an average of 4–5 min to segment an image. Also, fake circles that were created during nuclei detection stage using Hough Transform could not be removed completely. George et al. [48] suggested an automated method for nuclei segmentation of breast FNAC images. They extracted Y component of YCbCr color space for grey level conversion followed by Hough transform to detect circular shaped structures. To eliminate false circles, Otsu’s thresholding method was applied. To detect nuclei boundaries by avoiding over-segmentation marker controlled watershed transform was used. Twelve features were extracted for classification using MLP, PNN, LVQ, and SVM. Level set methods with geometric active contours [85], Multi-level set method [173] are found in literature. Again these methods could not properly address specimens containing overlapped nuclei.

**Clustering-based:** Seed-based region growing and moving k-means was propounded by Isa et al. [71] to determine the stages of cancer. Filipczuk et al. [43] approached with three level binarization algorithm by extracting the luminance component using $L = 0.2126R + 0.7152G + 0.0722B$. Initial
segmentation was done using adaptive thresholding. Second level involves clustering algorithms such as k-means, FCM, and Gaussian mixture models (GMM) to partition the image into nucleus, cytoplasm and background using different color channels as features. In the final level, they combined two segmented images using an AND operator to give precise definition of the boundaries of the image. But this method suffers from two major limitations. The need for determining optimal parameters and issues associated with unsupervised clustering restricted its use to practical purpose.

Deep Learning-based approach: Saikia et al. [148] developed a CNN-based deep learning classification framework where images were augmented using techniques like cropping, shearing, rotation, mirroring, skewing, inverting, zooming. In subsequent phases, channel identification, histogram equalization and Otsu’s thresholding were used to segment candidate nuclei. Maximum accuracy of 96.25% was achieved using GoogLeNet architecture. Kowal et al. [89] proposed a CNN and seeded watershed-based breast cytology segmentation model. CNN-based semantic segmentation model was first applied to differentiate between nuclei and background. The generated semantic mask was transformed into a nuclei mask to extract touching and overlapping nuclei. The clustered nuclei were detected by area and roundness. Nuclei seeds were identified using conditional erosion process. The overlapping nuclei were separated by seeded watershed algorithm. With this approach, 83.4% of benign nuclei were classified using Hausdorff distance.

(ii) Segmentation-free approach: In this approach, instead of segmenting an image, entire image or randomly selected sub-regions are used for feature extraction [30] and/or classification. Deep learning-based techniques like CNNs are successfully implemented to detect breast cancers [45, 212]. Garud et al. [45] used GoogLeNet on randomly selected regions from images during training and testing. An accuracy of 89.71% was reported on test data after voting of classified regions. Żejmo et al. [212] used AlexNet and GoogLeNet by selecting small patches of 256 × 256 from the large sized images of 200,000 × 100,000 pixels. They reported accuracies of 80% and 83% on two networks, respectively. It was noticed that the accuracy observed in CNN was still lagging behind traditional feature based models. The accuracy was improved by increasing the number of training samples as suggested by Khan et al. [84]. They proposed a transfer learning based classification technique using VGG net, GoogLeNet, and ResNet. For data augmentation translation, color processing, scaling, horizontal and/or vertical flipping, rotation and noise perturbation techniques were used. Features related to circularity, compactness and roundness were extracted using CNN architectures. The classification accuracy was obtained as 97.525%. Some notable works on breast cancer cytology are shown in Table 4.

Classification: Classifiers stand in the final lap of the image analysis system upon which the ultimate decision regarding nature of the specimen (whether malignant or not) is bestowed. The extent to which a classifier can correctly classify images defines the classification accuracy of the system. With the finest improvement in algorithmic complexity of the well-known classifiers, researchers are able to grasp the complex and diverse nature of images. A substantial number of scientific articles are published on classifiers. Several classifiers, e.g., MSM-T [165, 198], KNN [73, 88, 191], Kernel induced methods like SVM had been studied in References [20, 42, 48] to achieve good accuracy. Artificial neural network based approaches such as MLP [75] and its variant hybrid MLP [71], ensemble of classifiers and Decision trees [88] were also investigated. A comparative analysis of SVM, Learning Vector Quantization (LVQ), Probabilistic Neural Network, MLP was done in Reference [48]. Studies are also motivated toward spawning prediction of recurrence time of the disease. As this requires a large amount of statistical data, it is usually treated as a classification problem. Bradley et al. [17] emphasized on logistic regression for determining recurrence time of the disease. Street et al proposed a neural network trained with backpropagation algorithm [165] for predictive breast cancer risk model. These prospective models are still going through several modifications to make the system robust.

ACM Computing Surveys, Vol. 54, No. 3, Article 52. Publication date: April 2021.
## Table 4. Brief Overview of Different Works in Breast Cancer Cytology Identification Where Accuracy (Acc), Jaccard Index (JI), Specificity (Sp), Sensitivity (Se), and Dice Similarity Coefficient (DSC) Are Represented in Their Corresponding Bucket

| SL # | Author’s name | Name/Source (N/S) and Number of samples (Size) in the Dataset | Method for image segmentation | Features and classifier used | Quantitative results and Findings |
|------|---------------|-------------------------------------------------------------|------------------------------|-----------------------------|----------------------------------|
| 1    | Wolberg et al. (1993) [199] | N/S: University of Wisconsin Clinical Sciences Center Highland Avenue Madison, Wisconsin, USA. Size: 119 images. 68 benign, 51 malignant. | Manual segmentation followed by active contour detection. | Features: size, perimeter, area, compactness, radial variance, concavity, texture, size uniformity, worst size, worst shape. Classifier: MSM-Tree. | Acc: 90% Findings: Manual segmentation of nucleus boundary is cumbersome and time consuming. |
| 2    | Weyn et al. (1998) [191] | N/S: Antwerp University Hospital, Wilrijk, Belgium. Size: 83 images. 20 benign, 63 malignant. | Segmentation not done | Features: Wavelet, Denstimetric, GLCM, morphometric features. Classifier: KNN. | Acc: 76.1% Findings: Operator supervision is required and also false negatives present. |
| 3    | Street (2000a) [165] | N/S: University of Wisconsin Hospitals and Clinics beginning in 1984. Size: 569 images. 357 benign, 212 malignant. | Generalized Hough transform | Features: 30 numbers of nuclear morphometric features. Classifier: MSM-Tree | Acc: 97.5%. Se: 96.7%, Sp: 98.0%. Findings: Accurate closed boundaries, false findings persist, overlapped cells not taken care of. Not fully remote, and requires end user intervention. |
| 4    | Isa et al. (2007) [71] | N/S: Penang General Hospital and Hospital Universitiy Malaysia, Kelantan, Malaysia. Size: 1,300 images in 4 categories. | Segmentation not done | Features: 13 features of different cellular and nuclei properties. Classifier: Hybrid MLP (800 training and 500 test cases) | Acc:100% Findings: Intelligent system. Where stage of cancer can be determined. |
| 5    | Jeleń et al. (2008b) [75] | N/S: Medical University of Wroclaw, Poland. Size: 66 (intermediate malignancy), 44 (high malignancy) | Level set method | Features: area, perimeter, convexity, eccentricity and one texture feature. Classifier: MLP, Self-organizing maps (SOM), PCA, SVM. | Acc: Maximum 94.24% with SVM. Findings: Degree of malignancy can be determined. |
| 6    | Malek et al. (2009) [103] | N/S: Farhat Hached Hospital, Sousse, Tunisia. Size: 200 images in test data set. 120 benign, 80 malignant. | GVF Snake | Features: wavelet-based texture features. Classifier: Fuzzy C-means. | Acc: 95% Findings: Developed FPGA-based hardware system of software counterpart. |
| 7    | Kowal et al. (2011) [88] | N/S: Regional Hospital in Zielona Góra, Poland. Size: 500 images from 50 patients (10 images/patient). 25 benign, 25 malignant. | Adaptive thresholding and Gaussian mixture model-based segmentation. | Features: Different nuclei and cytoplasm centric features. Classifier(s): KNN, Naive Bayes classifier, DecisionTrees, Ensemble Classifier. | Acc: 98% (Maximum using KNN). Findings: Incorrect nuclei clusters are generated for lower no. of pixels. No standard dataset. |
| 8    | Filipczuk et al. (2013a) [42] | N/S: Regional Hospital in Zielona Góra, Poland. Size: 737 images from 67 patients (cases) | Circular Hough transform | Features: Nuclei and cytoplasm centric features with different intensities of RGB images. Correlation, Energy, Homogeneity, etc. Classifier(s): KNN, Naive Bayes, Decision Tree, SVM | Acc: 98.51% (Maximum using SVM). Findings: False circles are eliminated. No standard dataset. |
| 9    | Issac et al. (2013) [73] | N/S: Pathology lab of Regional Cancer Center, Thriruvananthapuram. Size: 334 (benign proliferative 31) (infiltrating) | K-means clustering technique in LAB color space. | Features: Textures and Nuclei centric features extraction after using complex Daubechies Wavelet Transform. Classifier: KNN | Acc: 93.9%, Se: 92.2%, Sp: 95.9%. Findings: Complex wavelet performed better than real wavelet. More morphometric features can improve results. |
| 10   | George et al. (2014) [48] | N/S: Ain Shams University Hospitals, Egypt Size: 92 images. 45 benign, 47 malignant having 11,502 cell nuclei. | Marker controlled watershed segmentation | Features: 12 statistical features, 10 texture features and 2 intensity-based features. Classifier(s): SVM, LVQ, PNN, MLP | Using PNN (max) Se: 96.32% Sp: 94.57%. Findings: Developed fully automated remote system. Grading of malignancy not done. |

(Continued)
| SL # | Author’s name | Name/Source (N/S) and Number of samples (Size) in the Dataset | Method for image segmentation | Features and classifier used | Quantitative results and Findings |
|------|---------------|-------------------------------------------------------------|------------------------------|-------------------------------|----------------------------------|
| 11   | Garud et al. (2017) [45] | N/S: Sub-divisional Hospital, Kharagpur, and Midnapur Medical College and Hospital, Midnapur, India. Size: 37 samples. 24 benign, 13 malignant. | Manually selected ROI | Features/classifier: Images are subdivided randomly to fit in GoogLeNet architecture and performances are reported both ROI level and sample level. | Acc: ROI: 80.76%. Samples: 89.71%. Findings: Nuclei centric feature-based approach perform better than GoogLeNet. |
| 12   | Zejmo et al. (2017) [212] | N/S: Regional Hospital in Zielona Góra, Poland. Size: 50 patients 25 benign, 25 malignant cases. | No segmentation | Features: 697 patches/ image were extracted. Classifier: CNN models AlexNet and GoogLeNet. Used randomly selected patches of 256 × 256 from large sized images. | Acc: 80% and 83% on patches using AlexNet and GoogLeNet. Findings: GoogLeNet perform better than AlexNet. |

Fig. 8. Samples of cytology images of Lung: (a) Benign case (MGG stain, 40× magnification), (b) Malignant case (Small cell Lung Carcinoma, MGG stain, 40× magnification), (c) Malignant case (Non-small cell Lung Carcinoma, MGG stain, 40× magnification).

### 4.3 RESPIRATORY TRACT CYTOLOGY

The respiratory tract is divided into upper tract from nasal space to larynx and lower tract from trachea to lungs. Lungs are a pair of internal organs located on either side of chest and exchange oxygen and carbon dioxide between the air we breathe and the blood. The inhaled air passes through main windpipe, trachea and conducts air into each lungs via left or right bronchus. The air passages divide into smaller tubes known as tracheobronchial tree and connect with tiny air sacs called alveoli. Lungs are protected with a thin tissue membrane known as pleura. Once lung cancer starts to binge to other parts of the body, particularly to lymph nodes, adrenal glands, liver, brain and bones, it becomes fatal. Causes of lung cancer include smoking, drinking or exposure to various air and water pollutants. Histological subtypes of lung cancer can originate from different locations of the tracheobronchial tree. There are two major types of lung cancer: Non-small cell lung cancer (NSCLC), which accounts for about 85%, and Small cell lung cancer (SCLC), which is more aggressive than NSCLC tumors, accounts for the rest. Samples of lung cytology images are shown in Figure 8. Nasal cytology [37], which is also a part of respiratory tract cytology is getting prominence nowadays to primarily diagnose rhino allergy. As the nasal cytology is not used popularly, therefore, we have mainly concentrated on lung cytology in the present article.

#### 4.3.1 Modalities of Specimen Collection

There are three major techniques to collect specimen from lung nodules: (i) Sputum collection, (ii) Bronchial Techniques, and (iii) FNA.
(i) \textit{Sputum collection}: Sputum samples are usually collected in a Cytolyt container during morning for three consecutive days when a person coughs up, known as "triple-morning test." As people under poverty line indulge at a greater rate in filter-less smoking, it is recommended to undergo screening at finite intervals and is extremely helpful for patients in low resource clinical settings. It has good sensitivity for central tumors compared to peripheral tumors.

(ii) \textit{Bronchial Techniques}: (a) \textit{Bronchial Brushing}: Cells are exfoliated using a brush from the periphery of the bronchial tree. A bronchoscope is used to guide the pathway. Cells obtained from brush are immediately fixed in alcohol. Its sensitivity and accuracy is relatively high compared to sputum cytology as it permits direct visualization of lesion.

(b) \textit{Bronchial Washing}: In this procedure, some amount of fluid is forced into lungs through bronchoscope and the washing is retrieved back. The washing contains requisite amount of cells required for cytological analysis.

(c) \textit{Brochoalveolar Lavage}: In this procedure, after an infusion of buffered saline solution in alveoli of lungs, the solution containing alveolar milieu is withdrawn back for further analysis.

(iii) \textit{FNA}: Lung nodules are also aspirated using computer tomography (CT) guided FNA. It has three sub types \cite{27}:

(a) \textit{Transbronchial FNA}: For lesions located in sub-bronchial regions, sensitivity is 56\% and specificity is 74\%.

(b) \textit{Transesophageal FNA}: It is done through endoscopy of esophagus. It is helpful in sampling of mediastinal lymph node.

(c) \textit{Percutaneous FNA}: It is rapid diagnosis of pulmonary cancer having sensitivity and specificity of 89\% and 96\%, respectively.

4.3.2 \textit{Differential Characteristics Associated with Cytomorphology of Lung Malignant Cells}.

- Polygonal/round/fibre-like cells.
- Abundant dense and smooth cytoplasm filled with keratin.
- Small hyperchromatic nucleus.
- Inconspicuous nucleolus.

4.3.3 \textit{Significant Works}. Compared to cervical and breast cytology, works in respiratory tract cytology are limited, which may be due to the lack of standard publicly available dataset.

\textbf{(i) Segmentation-based approach Region-based}: Kancherla et al. \cite{79} used a seeded region growing method to classify images of Biomada dataset. Seventy nine features were extracted related to shape, intensity, color, and wavelet-based features, and so on, to achieve a recognition accuracy of 87.8\% using bagging on Random Forest.

\textbf{Color-based}: Segmentation of sputum cells using color as the key discriminating factor has been studied in References \cite{39, 168}. Vineeth et al. \cite{92} explored the impact of color representation on classification of lung sputum images. Bayesian Classifier was used to classify sputum cells into three classes, i.e., nucleus, cytoplasm and background with zero mis-classification error rate using HSV components as feature vector.

\textbf{Clustering-based}: Rachid et al. \cite{138} proposed an unsupervised technique using HNN to cluster pixels of sputum images into nuclei, cytoplasm and background classes. To escalate the accuracy of segmented regions, Sammouda et al. \cite{150} inserted an energy function having a cost term. However, the method suffers from early local minimum of HNN due to which the superimposed cells could not be delineated. In Reference \cite{169}, sputum images were segmented into nucleus, cytoplasm, and background using HNN-based module. 8-connectivity was used to find connected component in nucleus. The proposed framework proved to be robust in systematic setting of the classification parameter. In Reference \cite{170}, a Bayesian classification framework was proposed to investigate whether a pixel in the image belongs to sputum cell or not. HNN segmented the
nucleus, cytoplasm and background region of 88 color sputum images. Segmentation accuracies using HNN and FCM were 88.62% and 64.91%, respectively, in HSV color space. Forseberg et al. [44] proposed an adaptive thresholding technique on whole-slide images of endobronchial ultrasound-guided Transbronchial Needle Aspiration (TBNA). K-means clustering was used to segment the nucleus. A processing time of 7.8 seconds/sample was reported. Sarvaiya et al. [152] guided a sputum image segmentation by classifying colors using k-means clustering. Kecheril et al. [81] proposed a localization technique of cellular region by scale space-based determinant of Hessian. Due to presence of non-cellular artifacts, Otsu’s threshold method was adopted. K-means clustering was used to segment nuclei. 32 scales were generated and at each scale saddle maxima and saddle minima catastrophe points were calculated to generate 64-dimensional feature vectors. SVM with RBF kernel was used with tenfold cross-validation to classify images.

**Mean shift-based:** Werghi et al. [190] proposed a robust Bayesian classifier using components of four different color spaces. The RGB and HSV color spaces showed consistency in all resolutions. For sputum cell segmentation, they used mean shift technique by a judicious choice of threshold parameter using both spatial and chromatic information. A reasonable accuracy of 85.51% was obtained, which overruled the traditional HNN framework.

**Deep-learning-based:** Baykal et al. [9] proposed a faster region-based CNN and single shot multi-box nucleus detector on pleural effusion cytology images. For feature extraction Resnet-50,101, InceptionNet-V2, mobileNet were used. Faster RCNN with Resnet-101 were used to detect nucleus with an F-measure of 98.12%.

(ii) Segmentation-free approach: Teramoto et al. [178] introduced CNN for classification of adenocarcinoma, squamous cell carcinoma, and small cell carcinoma with an accuracy of 71.1% without segmentation. To increase 298 samples to 5,000 data augmentation techniques like rotation, flipping, color correction, spatial filtering, and so on, were used. A similar approach was adopted by Reference [179] to obtain an accuracy of 79.3% for patch-based classification and 87% for case-based classification using a fine tuned CNN VGG-16. Baykal et al. [10] proposed transfer learning-based CNN for classification of serous cells where they achieved a maximum recognition accuracy of 93.44%. Significant works on respiratory tract cytology are shown in Table 5.

**Classification:** Bayesian classifiers [39, 190], ANN [33, 171], SVM [33, 81, 155, 171], Random forest [32], and so on, are some commonly used classifiers in this domain. Dimauro et al. [35] proposed a cell-based classification of nasal cytology into seven classes using three-layered CNN. Coarse segmentation was done using enhancement filters combined with Otsu’s method. It was fine tuned using watershed algorithm to achieve a segmentation accuracy of 97% and classification accuracies of 94% on validation set and 99% on test set.

### 4.4 Miscellaneous

So far, we presented different methodologies on three cytology domains where majority of automation techniques are applied. In this section, we have clubbed together the automated analysis of the rest in a single section due to lack of colossal works in each single domain.

Daskalakis et al. [31] proposed a pixel-based technique to segment 53 benign and 62 malignant images of thyroid nodules collected from the University hospital of Patras. Twenty-six morphological and textural features were extracted. Four classifiers were used and a maximum accuracy of 95.7% was reported using majority voting-based combination technique among KNN, PNN, and Bayesian network. Gopinath et al. [52] proposed a mathematical morphology-based approach for segmentation of thyroid cytology images. DWT-based features were extracted. Using classifier fusion technique they obtained a recognition accuracy of 96.66% on test data (30 images). Wu et al. [204] used level set methods, watershed algorithm, CRF, EM algorithm, for segmentation of multispectral cytology images. They exploited the idea of spectral data using a multi-CRF model with a
| SI# | Author’s name                  | Name/Source (N/S) and Number of samples (Size) in the Dataset | Method for image segmentation | Features and classifier used | Quantitative results and Findings |
|-----|--------------------------------|---------------------------------------------------------------|-------------------------------|------------------------------|----------------------------------|
| 1   | Taher et al. (2015) [171]      | N/S: Tokyo center of lung cancer in Japan. Size: 100 sputum color images. | Mean shift segmentation.      | Features: N/C ratio, curvature, circularity, Eigen ratio, density. Classifier: ANN and SVM | Acc: 97% Se: 97% Sp: 96% Findings: Single sputum cell. Cell boundary not properly delineated. |
| 2   | Shajy et al. (2015) [155]      | N/S: Regional Cancer center, Thiruvananthapuram, Kerala. Size: 32 training and 59 testing. | Image localization using maximization of determination of Hessian in scale and k-means used for clustering. | Features: 690 space scale catastrophic point based on features Classifier: SVM | Acc: 87.53% Se: 76.9% Sp: 92.82% Findings: Only glandular cells are considered for segmentation, while missing some actual cells during segmentation. |
| 3   | Kecheril et al. (2015) [81]    | N/S: Regional cancer centre, Thrubanthapuram, India. Size: 10 benign, 24 malignant | Random walk with K-means. Watershed was used to remove unwanted regions. | Features: 25 features. (10 geometric, 14 texture features and 1 color-based) Classifier: ANN and SVM | Using ANN (Max) Acc: 97.46% Se: 97.5% Sp: 97.6% Findings: Description of samples are misleading. |
| 4   | Teramoto et al. (2019) [179]   | N/S: Fujita Health University, Toyoake City, Japan. Size: 197 images of 46 cases. | 621 Patch images of size 224 × 224 are extracted. | Features and classifiers: Used VGG-16 with data augmentation. | Acc: 79.3% for patches Acc: 87% for cases Findings: Not explored other CNN architectures. |
| 5   | Dholey et al. (2018a) [32]     | N/S: Medical College Kolkata, and EKO centre Kolkata, India. Size: 600 images from 120 samples | GMM-based hidden MRF. Morphological filters used to remove unwanted regions. | Features: SIFT, Bag of word and visual dictionary after clustering. Classifiers: Random Forest | Acc: 98.88% Se: 97.31% Sp: 99.54% Findings: Description of samples are misleading. |
| 6   | Dholey et al. (2018b) [33]     | N/S: Medical College Kolkata, and EKO centre. Kolkata, India. Size: 500 images from 100 samples. | Mean shift filter followed by Otsu thresholding. | Classifier: CNN | Result of test set on 7 class problem Acc: 98.6% Se: 99.7% Sp: 99.4% Findings: Manual acquisition of fields. No standard dataset. |
| 7   | Dimauro et al. (2019) [35]     | N/S: Rhinology Clinic of the Otolaryngology Department of the University of Bari. Size: 14 Nasal cytology images. | Otsu thresholding algorithm. Morphological opening operation followed by labelling, marking objects. | Classifiers: CNN(ConvNet) 1st approach: Data augmentation: geometric transformations (reflection, rotation, translation) 2nd approach: hyperparameter-optimization | Result: Acc: 94.0% Se: 96.4% Sp: 82.5% Findings: Low cost system for preparing report of rhino-cytogram. |
| 8   | Dimauro et al. (2018) [36]     | N/S: superficial cells from the nasal mucosa collected from Rhinology Clinic of the Otolaryngology Department of the University of Bari | | | |
bottom-up approach using local probabilistic model and an unsupervised top-down approach with a probabilistic latent semantic analysis (PLSA) representation. PLSA was used to assign a latent topic to a pixel determined by Expectation-Maximization algorithm, which was further upgraded by local mapping. This approach showed better accuracy in low contrast and noisy environment.

To identify the appropriate type of ovarian cancer, Wu et al. [202] proposed a CNN-based approach using AlexNet. Eighty-five labelled specimens in four categories, namely, serous carcinoma, mucinous carcinoma, endometrioid carcinoma, and clear cell carcinoma, were collected from Xingjiang Medical University, and 1,848 ovarian cytological images of 1,360×1,024 dimensions were considered. They used rotation-based augmentation technique to increase the size of the dataset to 20,328 samples. Images were divided into four sub-images and resized them to 227×227. They reported an accuracy of 78.20% using AlexNet on tenfold cross-validation.

Scalbert et al. [154] developed a doodle-style transfer-based isolated cell image generator for urine cytology. But generation of whole slide could not be achieved using the proposed scheme.

Chatterjee et al. [24] used a combination of statistical features such as morphology, intensity, color, texture, and histogram for diagnosis of oral malignancy. They reported maximum recall of 95.55% using random forest classifier. Some notable works on Miscellaneous cytology are shown in Table 6.

5 AUTHORS’ VIEW

During the journey toward automation in cytology, we can make out goals and achieved solutions. The primary goal is to help cytologists to reduce turnaround time with lesser “false” cases. Though none of the false cases is desired either, false-negative cases should be delimited from the diagnostic point of view. Because one false-negative case would leave the patient untreated leading to a fatal situation. False-positive cases, however, puts the patient in a traumatic condition. Thus, the machine-driven output should be at least comparable to human evaluation. The second goal is the cost-effectiveness, so that it can be run in low resource clinical settings.

It is observed from Figure 9 that a lot of experiments have been done on cervical and breast cancer detection whereas methods explored for other types of cytology are few. Some of the major segmentation techniques found extensively in literature are jotted down in Table 7. This survey points out the progress of automation in terms of research to reach a viable commercial output. During the period of 1988 to 1998, though meagre amount of cytology-based research works are registered, most of the researches tend to concentrate on contour and deformable models as depicted in Figure 10. During the period of 1998 to 2008, texture-based features using wavelet analysis were in vogue. During the past decade, deformable contour models and clustering-based segmentation approach showed encouraging results in diagnosing carcinoma, thus jointly bagging nearly
Table 6. Brief Overview of Different Works in Miscellaneous Domain of Cytology Where Accuracy (Acc), Jaccard Index (JI), Specificity (Sp), Sensitivity (Se), and Dice Similarity Coefficient (DSC) Are Represented in Their Corresponding Bucket

| SL# | Author’s name | Name/Source and number of samples (Size) in dataset | Method in Image Segmentation | Features and Classifiers used | Quantitative results and Findings |
|-----|---------------|-----------------------------------------------------|------------------------------|-------------------------------|----------------------------------|
| 1.  | Wu et al. (2009) [205] | Thyroid FNAC smears | Morphological self dual reconstruction using Watershed algorithm. | Bhattacharyya distance | Acc.- 95.19% FP- 2.12% Findings: Only segmentation of the images. No standart dataset |
| 2.  | Gopinath et al. (2013) [51] | Thyroid FNAB collected from on-line image atlas of Papanicolaou Society of Cytopathology | Morphological operations Watershed transformation | Features: Two-level discrete wavelet decomposition GLCM, Gabor filters Classifiers: k-NN, ENN (Elman neural network), SVM | Max. Acc: 93.33 % using ENN classifier Findings: CAD-based system for multi-stain images. |
| 3.  | Sanyal et al. (2018) [151] | Thyroid FNAC smears collected from tertiary care centers of North India | Not applicable | ANN | Combined result with 10x and 40x magnification: Se: 90.48% Sp: 83.33% negative predictive value: 96.49% Acc: 85.06% Findings: No standart dataset |
| 4.  | Hossain et al. (2019) [61] | Renal cytology collected from National Cancer Institute, USA | K-means clustering to classify background and nucleus region. SVM to differentiate between normal and abnormal nuclei regions. Selective search algorithm is used to detect irregular shaped structures in abnormal regions. | Features: patches of normal and abnormal cells. Classifier: RCNN used for normal and abnormal cell detection | Precision: 99.01% Recall: 98.7% F-measure: 98.8% Findings: proliferation rate estimation for successful prognosis of the disease. |

45% of the research works. As neural network-based deep learning techniques have become the cutting edge technology from the past few years due to its automatic feature learning mechanism, it is slowly replacing previously used classification techniques using KNN, Random Forest, SVM, and so on. In Figure 11, we have plotted various tasks over the past five years. It is observed that segmentation using traditional algorithms has decreased over the years, whereas deep-learning-based segmentation approaches had its first spike in 2017. Classification using deep-learning is gaining popularity over the past four years. However, addressing both segmentation and classification problems using the traditional approach is still finding more applications except for the past two years where numbers are almost comparable. The overall works using deep-learning have witnessed an incremental pattern, which signifies the potential of deep-learning approaches to classify images accurately. Its performance is largely dictated by an exhaustive dataset that has paucity in the cytology domain.

Though there are diverse and plenty of works in the cervical and breast cytology domains, whereas a few are realized in respiratory tract cytology. At the same time, research works on
Table 7. Major Segmentation Algorithms Used in Cytology Images

| SL# | Segmentation techniques | Methodologies |
|-----|------------------------|---------------|
| 1   | Thresholding-based     | a. Otsu thresholding [177, 196], b. Modified Otsu with class prior probability [91] | |
|     |                        | c. Seed-based region growing [79], b. MSBRG [108], [98], c. MSER [101] d. Grow-Cut [63], e. Anisotropic diffusion and anisotropic kernel mean shift [177] |
| 2   | Region-based           | a. Seed-based region growing [79], b. MSBRG [108], [98], c. MSER [101] d. Grow-Cut [63], e. Anisotropic diffusion and anisotropic kernel mean shift [177] |
| 3   | Edge/Contour-based     | a. CNC detector [181], [129], b. EENCC detector [209], c. Sobel edge detector [159] d. Non-maximum suppression [98], e. Hough transform [62, 165], f. Compact Hough transform [114] g. Two-group edge enhance method [98], h. Laplacian, Prewitt, Roberts, Robinson [108], i. Superpixel Partitioning and Cell-Wise Contour Refinement [94] j. Parametric elliptical waveform fitting [55], k. Riemannian dilation [104] |
| 4   | Deformable contour models | a. Active Contour Model [177, 207], b. Hough transform with deformable models [95], c. Adaptive active contour modelling [213], d. Snake model [6, 123, 199], e. GVF snakes [103], f. Radiating GVF Snake [97], [149], g. Dynamic sparse contour and GVF Snake [53], h. Viterbi search-based dual ACM [142], i. Level set method [98, 116], j. Joint level set [101], k. Multiple level set [102], l. Level set method active contour model [41] m. Multi-step level set method [72] |
| 5   | Watershed-based        | a. Watershed-based [63], [115, 126, 134, 158], b. Multi-pass fast watershed [175], c. Multi scale Watershed [78], d. color-based watershed [96], e. Hierarchical watershed [127] |
| 6   | Texture-based          | a. Coarseness [211], b. Texture Filter Bank [15], c. Scale space features [81] d. Wavelet transforms [3, 5, 73, 191] e. GLCM [166, 186] f. Gabor [123], [43, 115, 141], g. Rotation invariant LBP [125] h. Conventional LBP [43], rotation invariant patterns, local patterns with anisotropic structure, completed local binary pattern (CLBP) and local ternary pattern (LTP) [54] |
| 7   | Graph-based            | a. Global and local graph cuts [217], b. Graph search [218], c. Graph-search based MRF [221] |
| 8   | Clustering             | a. Spatial k-means [97], b. k-means algorithm [108], [44], c. Fuzzy C means [70, 86, 146, 157], d. Spatial patch-based Fuzzy C-means [22], e. Multi-scale Fuzzy C-means [65, 188], f. GMM-based [88, 139, 164], g. GMM with HMRF [32], h. Superpixel-based MRF [221], i. SLIC [136], j. Superpixel Partitioning and Cell-Wise Contour Refinement [94], k. Entropy-based superpixel method [112], l. Superpixel with Voronoi [183] |
| 9   | Hierarchical segmentation | a. Hierarchical shape approximation, and shape regularization [177] b. Non-parametric hierarchical segmentation algorithm [139] |
| 10  | Deep learning-based    | a. Multiscale CNNs [161], b. Multiscale CNN and graph-based partitioning [162] c. Deep learning and dynamic shape modeling [176] d. Patch-based CNN [4] e. Modified DeepLab V2 [187] f. Mask Regional Convolutional Neural Network (Mask R-CNN) [160], g. Instance relation network IRNet [223] h. Progressive Growing of U-net (PGU-Net) [220] i. Stacked sparse autoencoders [8] |
| 11  | Other Segmentation techniques | a. Color-based [216], [39], [16], [70] b. Phansalkar’s local search [183] c. Intersecting Cortical Model (ICM) [174], [3] d. Spatially adaptive active physical model [135] e. Minimax optimization of an energy functional [1] f. Mean Shift [13] g. QPSO with Fuzzy KNN [218], h. Star shape prior and Voronoi energy term, [124] i. Viola-Jones object detection approach [11] j. Multifractal [93] k. Neuro-Fuzzy [144], l. CRF [204] |

Fig. 10. Major segmentation techniques and their usages in past three decades on cytology images.
Fig. 11. Number of papers in different tasks for automation in cytology in past two decades.

histopathology images of lung are much more pronounced and diverse compared to cytology. It is observed that methodologies in lung cytology tend to concentrate on particular techniques. So, techniques also lack diversity at the end. Also, the majority of techniques produce only segmented images. Thus, the end-user interference is mandatory. Again, major works are based on sputum cytology whereas, other modalities of cytology are almost unattended creating, a large void and exploring opportunities to work on those modalities.

Developing certified systems for an automated application requires a deep insight into the functionality of the design, including power, cost and time to market to help ensure market success. Although some automated devices have become de-facto medical standards in a few parts of the world, third world countries like, India, are still to exploit the advantage of screening devices on a full scale. So, efforts are now streamlined toward producing a screening unit that can be operated in a semi-automated fashion. Researches are presently heading toward producing intelligent and remote web-based diagnostic systems that can handle a large group of patients with reduced false cases. It is observed that false cases are almost part and parcel of automated systems. Despite several disadvantages, many software and hardware automated systems are available in laboratories that act as a human companion, assisting in various decision-making processes by generating a reference from a machine-generated output. Existing systems though require end-user interaction; nevertheless, they reduced the workload to a greater extent. Thus, cytologists apart from reviewing slides can also join in research activities.

6 CONCLUSION

This article reviews and captures the diversity of state-of-the-art methodologies in the analysis of cytology images. It is realized that most of the existing works revolve around the cervical, breast and respiratory tract cytology whereas other domains do not receive much of the researcher’s attention. It is also realized that some segmentation frameworks surpass all contemporary results following recent trends of deep learning module, while other traditional/conventional algorithms are used to segment cells using handcrafted features. To be precise, we can now encapsulate a few outcomes of this article as possible future directions to bring cytology-based commercial systems into the mainstream.

- Appropriate segmentation algorithm is yet to be designed for the segmentation of clusters of overlapped nuclei. In addition, majority of segmentation algorithms are developed focusing on supervised data, whereas robust unsupervised algorithms need to be designed for automatic labelling of data.
• Processing high-resolution images with conventional segmentation techniques increases time complexity.
• Lack of freely available standard datasets for cytology images, except for pap smear cytology. Also, the size of available datasets is not sufficient enough to train a deep learning module efficiently.
• Absence of standardized staining technique hinders segmentation techniques to be appropriately implemented.
• Non-availability of viable commercial systems apart from cervical cytology.
• There exists no single system for handling cytology images in different domains.

Possible future directions to successfully run the existing systems can be enumerated below:

• The limitation of the available standard dataset can be overcome by synthetically generating realistic data using contemporary deep learning techniques, such as Adversarial Neural Network, Variational Autoencoder, and so on.
• Appropriate image resizing techniques and parallelization of algorithms to make computational time faster.
• Implementation of techniques for handheld devices using embedded modules like Raspberry Pi, Arduino, and so on.
• A typical procedure (see Figure 12) may be followed to prepare standard datasets and can be uploaded in different competition websites, like www.grand-challenge.org or www.kaggle.com, to motivate researchers to work on it.

Automatic understanding of the nature of cytology images is a challenge to researchers because of their diverse nature and presence of unusual artifacts. Systems that are available for cervical, breast, and respiratory tract cytology are not robust enough to deal with all kinds of data existing globally, to be marketed on a large scale. Also, the production of high precision screening machines is not very cost-effective. All these factors coagulated to hinder current screening systems reaching third-world countries like India. So, efforts are now streamlined toward finding a feasible solution or a screening unit that can at least be run in a semi-automatic fashion. The success in this regard, to date, is limited; still, there is a lot of scope to work in this domain due to improvements in data collection and staining techniques. We, finally, hope that this survey will help researchers to comprehend the latest state-of-the-art methods and progress of cytology-based research for the systematic and selective design of algorithms and test ideas to develop a conceptual framework suitable for analysis in the relevant problem domain.

ACKNOWLEDGMENTS

Authors are thankful to the Theism Medical Diagnostics Centre, Kolkata, India, and Saroj Gupta Cancer Centre & Research Institute, Kolkata, India.
REFERENCES

[1] P. Agarwal, A. Sao, and A. Bhavsar. 2015. Mean-shift-based segmentation of cell nuclei in cervical PAP-smear images. In Proceedings of the 5th National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG’15). 1–4. DOI: https://doi.org/10.1109/NCVPRIPG.2015.7490039

[2] I. Al and J. S. Ploem. 1979. Detection of suspicious cells and rejection of artefacts in cervical cytology using the leyden television analysis system. J. Histochem. Cytochem. 27, 1 (1 1979), 629–634.

[3] G. Anna Lakshmi, S. Ravi, G. Anna Lakshmi, and S. Ravi. 2017. Automated segmentation algorithm for cervical cell images by employing cuckoo search-based ICM. J. Ambient Intell. Human. Comput. (11 2017), 1–11. DOI: https://doi.org/10.1007/s12652-017-0640-z

[4] Flavio H. D. Araujo, Romuere R. V. Silva, Daniela M. Ushizima, Claudia M. Carneiro, Andrea G. Campos Bianchi, and Fatima N. S. Medeiros. 2019. Deep learning for cell image segmentation and ranking. Comput. Med. Imag. Graph. 72 (2019), 13–21.

[5] Mithlesh Arya. 2018. Texture-based feature extraction of smear images for the detection of cervical cancer. IET Comput. Vision 12, 8 (Dec. 2018), 1049–1059. Retrieved from https://digital-library.theiet.org/content/journals/10.1049/iet-cvi.2018.5349.

[6] Pascal Bamford and Brian Lovell. 1998. Unsupervised cell nucleus segmentation with active contours. Signal Process. 71, 2 (1998), 203–213.

[7] Eleanor P. Bankes. 2007. Cervical Cancer Research Trends (Nova Biomedical). Nova Science Pub.

[8] Elif Baykal, Hulya Dogan, Mustafa Emre Ercin, Safak Ersoz, and Murat Ekinci. 2018. Automated nuclei detection in serous effusion cytology with stacked sparse autoencoders. In Proceedings of the 26th Signal Processing and Communications Applications Conference (SIU’18). IEEE, 1–4.

[9] Elif Baykal, Hulya Dogan, Mustafa Emre Ercin, Safak Ersoz, and Murat Ekinci. 2020. Modern convolutional object detectors for nuclei detection on pleural effusion cytology images. Multimedia Tools Appl. 79, 21 (June 2020), 15417–15436.

[10] Elif Baykal, Hulya Dogan, Mustafa Emre Ercin, Safak Ersoz, and Murat Ekinci. 2020. Transfer learning with pre-trained deep convolutional neural networks for serous cell classification. Multimedia Tools Appl. 79, 21 (June 2020), 15593–15611.

[11] E. Baykal, H. Dogan, M. Ekinici, M. E. Ercin, and S. Ersoz. 2017. Automated nuclei detection in serous effusion cytology based on machine learning. In Proceedings of the 25th Signal Processing and Communications Applications Conference (SIU’17). 1–4.

[12] Christoph Bergmeir, Miguel García Silvente, and Jose Manuel Benitez. 2012. Segmentation of cervical cell nuclei in high-resolution microscopic images: A new algorithm and a web-based software framework. Comput. Methods Prog. Biomed. 107, 3 (2012), 497–512.

[13] C. Bergmeir, M. Garcia Silvente, J. Esquivias Lopez-Cuervo, and J. M. Benitez. 2010. Segmentation of cervical cell images using mean-shift filtering and morphological operators. In Medical Imaging 2010: Image Processing, B. M. Dawant and D. R. Haynor (Eds.), Vol. 7623. SPIE, 1337–1345.

[14] Kangkana Bora, Manish Chowdhury, Lipi B. Mahanta, Malay K. Kundu, and Anup K. Das. 2016. Pap smear image classification using convolutional neural network. Proceedings of the 10th Indian Conference on Computer Vision, Graphics and Image Processing (ICVGIP’16). 1–8.

[15] Kangkana Bora, Manish Chowdhury, Lipi B. Mahanta, Malay Kumar Kundu, Anup Kumar Das, and Anup Kumar Das. 2017. Automated classification of pap smear image to detect cervical dysplasia. Comput. Methods Prog. Biomed. 138, 1 (1 2017), 31–47.

[16] Omelkhir Boughzala. 2016. Automatic segmentation of cervical cells in Pap smear images. In Proceedings of the 24th Conference on Computer Graphics, Visualization and Computer Vision 2016. 343–349.

[17] Cathy J. Bradley, Charles W. Given, and Caralee Roberts. 2001. Disparities in cancer diagnosis and survival. Cancer 91, 1 (1 2001), 178–188.

[18] Alex brollo. 2006. Retrieved from https://upload.wikimedia.org/wikipedia/commons/9/95/Urine_citology_urothelial_carcinoma_2.jpg.

[19] Anne E. Carpenter, Thouis R. Jones, Michael R. Lamprecht, Colin Clarke, In Han Kang, Ola Friman, David Guertin et al. 2006. CellProfiler: Image analysis software for identifying and quantifying cell phenotypes. Genome Biol. 7, 10 (10 2006), 1–11.

[20] Turgay Celik and Tardi Tjahjadi. 2010. Texture classification and retrieval based on complex wavelet subbands. In Computer and Information Sciences, E. Gelenbe, R. Lent, G. Sakellari, A. Sacan, H. Toroslu, and A. Yazici (Eds.). Springer Netherlands, Dordrecht, 259–264.

[21] Thanatip Chankong, Nipon Theera-Umporn, and Sansanee Auephanwiriyakul. 2009. Cervical cell classification using fourier transform BT. In Proceedings of the 13th International Conference on Biomedical Engineering, C. T Lim and J. C H Goh (Eds.). Springer, Berlin, 476–480.
[22] Thanatip Chankong, Nipon Theera-Umpon, and Sansanee Auephanwiriyakul. 2014. Automatic cervical cell classification using patch-based fuzzy clustering and minimum average correlation energy filter. In Proceedings of the International Federation of Medical and Biological Engineering (IFMBE’14), Vol. 43, 164–167.

[23] Thanatip Chankong, Nipon Theera-Umpon, and Sansanee Auephanwiriyakul. 2014. Automatic cervical cell segmentation and classification in PAP smears. Comput. Methods Prog. Biomed. 113, 2 (2014), 539–556. DOI: https://doi.org/10.1016/j.cmpb.2013.12.012

[24] Saunak Chatterjee, Debaleena Nawn, Mousumi Mandal, Jyotirmoy Chatterjee, Somdutta Mitra, Mousumi Pal, and Ranjan Rashmi Paul. 2018. Augmentation of statistical features in cytopathology towards computer aided diagnosis of oral precancer/cancer. In Proceedings of the 4th International Conference on Biosignals, Images and Instrumentation (ICBSII’18). IEEE, 206–212.

[25] Y. Chen, P. Huang, K. Lin, H. Lin, L. Wang, C. Cheng, T. Chen, Y. Chan, and J. Y. Chiang. 2014. Semi-automatic segmentation and classification of pap smear cells. IEEE J. Biomed. Health Info. 18, 1 (2014), 94–108.

[26] Gaurav Chhabra. 2018. Automated hematometry analyzers: Recent trends and applications. J. Lab. Phys. 10, 01 (2018), 015–016.

[27] Edmund S. Cibas and Barbara S. Ducatman. 2013. Cytology E-Book: Diagnostic Principles and Clinical Correlates. Elsevier Health Sciences.

[28] Teresa Conceição, Cristiana Braga, Luís Rosado, and Maria João M. Vasconcelos. 2019. A review of computational methods for cervical cells segmentation and abnormality classification. Int. J. Mol. Sci. 20, 20 (Oct. 2019), 5114.

[29] Goodarz Danaei, Stephen Vander Hoorn, Alan D. Lopez, Christopher J. L. Murray, and Majid Ezzati. 2005. Causes of cancer in the world: Comparative risk assessment of nine behavioural and environmental risk factors. Lancet 366, 9499 (2005), 1784–1793.

[30] P. Das, T. Chatterjee, S. Chakraborty, D. Mondal, and N. Das. 2015. A texture-based approach for automatic identification of benign and malignant tumor from FNAC images. In Proceedings of the IEEE 2nd International Conference on Recent Trends in Information Systems (ReTIS’15). 249–254.

[31] Antonis Daskalakis, Spiros Kostopoulos, Panagioti Spyridonos, Dimitris Glotsos, Panagioti Ravazoula, Maria Kardari, Ioannis Kalatzi, Dionisia Cavouros, and George Nikiforidis. 2008. Design of a multi-classifier system for discriminating benign from malignant thyroid nodules using routinely H&E-stained cytological images. Comput. Biol. Med. 38, 2 (2008), 196–203.

[32] Moumita Dholey, Maitreya Maity, Atasi Sarkar, Amita Giri, Ampu Sadhu, Koel Chaudhury, Soumen Das, and Jyotirmoy Chatterjee. 2018. Combining GMM-based hidden markov random field and bag-of-words trained classifier for lung cancer detection using pap-stained microscopic images. In Advanced Computational and Communication Paradigms, S. Bhattacharyya et al. (Ed.). Springer, 695–705.

[33] Moumita Dholey, Atasi Sarkar, Maitreya Maity, Amita Giri, Ampu Sadhu, Koel Chaudhury, Soumen Das, and Jyotirmoy Chatterjee. 2018. A computer vision approach for lung cancer classification using FNAC-based cytological images. In Proceedings of the 2nd International Conference on Computer Vision and Image Processing, B. B Chaudhuri, M. S Kankanahalli, and H. (Eds.). Springer, Singapore, 181–195.

[34] E. G. Diaconakos, Emerson Day, and M. J. Kopac. 1962. Exfoliated cell studies and the cytoanalyzer. Ann. N.Y. Acad. Sci. 97, 2 (9 1962), 498–513. DOI: https://doi.org/10.1111/j.1749-6632.1962.tb43660.x

[35] Giovanni Dimaruo, Giorgio Ciprandi, Francesca Deperte, Francesca Girardi, Enrico Ladisa, Sergio Latrofa, and Matteo Gelardi. 2019. Nasal cytology with deep learning techniques. Int. J. Med. Info. 122 (2019), 13–19.

[36] Giovanni Dimaruo, Francesca Girardi, Matteo Gelardi, Vitoantonio Bevilacqua, and Danilo Caivano. 2018. Rhinocyty: A system for supporting the rhinologist in the analysis of nasal cytology. In Intelligent Computing Theories and Paradigms, H. Huang, K. Jo, and X. Zhang (Eds.). Springer.

[37] Giovanni Dimaruo, Davide Di Pierro, Francesca Deperte, Lorenzo Simone, and Pio Raffaele Fina. 2020. A smartphone-based cell segmentation to support nasal cytology. Appl. Sci. 10, 13 (June 2020), 4567. DOI: https://doi.org/10.3390/app10134567

[38] Na Dong, Meng die Zhai, Li Zhao, and Chun Ho Wu. 2020. Cervical cell classification based on the CART feature selection algorithm. J. Ambient Intell. Human. Comput. 12, 2 (June 2020), 1837–1849. DOI: https://doi.org/10.1007/s12652-020-02256-9

[39] C. Donner, N. Werghi, F. Taher, and H. Al-Ahmad. 2012. Cell extraction from sputum images for early lung cancer detection. In Proceedings of the 16th IEEE Mediterranean Electrotechnical Conference. 485–488. DOI: https://doi.org/10.1109/MELCON.2012.6196478

[40] R. Erhardt, E. R. Reinhardt, W. Schlipf, and W. H. Bloss. 1980. FAZYTAN: A system for fast automated cell segmentation, cell image analysis and feature extraction based on TV-image pickup and parallel processing. Anal. Quant. Cytol. 2, 1 (1980), 25–40.

[41] J. Fan, R. Wang, S. Li, and C. Zhang. 2012. Automated cervical cell image segmentation using level set-based active contour model. In Proceedings of the 12th International Conference on Control Automation Robotics Vision (ICARCV). 877–882. DOI: https://doi.org/10.1109/ICARCV.2012.6485273
Cytology Image Analysis Techniques Toward Automation: Systematically Revisited

[52] Pawel Filipczuk, Thomas Fevens, Adam Krzyzak, Roman Monczak, Adam Krzy, and Roman Monczak. 2013. Computer-aided breast cancer diagnosis based on the analysis of cytological images of fine needle biopsies. *IEEE Trans. Med. Imag.* 32, 12 (2013), 2169–2178.

[53] Pawel Filipczuk, Bartosz Krawczyk, and Michal Woźniak. 2013. Classifier ensemble for an effective cytological image analysis. *Pattern Recogn. Lett.* 34, 14 (2013), 1748–1757. DOI: https://doi.org/10.1016/j.patrec.2013.04.030

[54] D. Forsberg and N. Monsef. 2014. Evaluating cell nuclei segmentation for use on whole-slide images in lung cytology. In *Proceedings of the 22nd International Conference on Pattern Recognition*. 3380–3385. DOI: https://doi.org/10.1109/ICPR.2014.582

[55] Hrushikesh Garud, S. P. K. Karri, Debdoott Sheet, Jyotirmoy Chatterjee, Manjunatha Mahadevappa, Ajoy K. Ray, Arindam Ghosh, and Ashok K. Maity. 2017. High-magnification multi-views-based classification of breast fine needle aspiration cytology cell samples using fusion of decisions from deep convolutional networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW’17)*. 828–833.

[56] Srishti Gautam, Harinarayan K. K., Nirmal Jith, Anil K. Sao, Arnav Bhavasr, and Adarsh Natarajan. 2018. Considerations for a PAP Smear Image Analysis System with CNN Features. Retrieved from http://arxiv.org/abs/1806.09025

[57] Asli Genctav, Selim Aksoy, and Sevgen Onder. 2012. Unsupervised segmentation and classification of cervical cell images. *Pattern Recogn.* 45, 12 (2012), 4151–4168. DOI: https://doi.org/10.1016/j.patcog.2012.05.006

[58] Negar M. Harandi, Saeed Sadri, Noushin A. Moghaddam, and Rassul Amirfattahi. 2010. An automated method for segmentation of epithelial cervical cells in images of ThinPrep. *J. Med. Syst.* 34, 6 (Dec. 2010), 1043–1058. DOI: https://doi.org/10.1007/s10916-009-9323-4

[59] J. Holmquist, E. Bengtsson, O. Eriksson, B. Nordin, and B. Stenkivist. 1978. Computer analysis of cervical cells. *J. Histochem. Cytochem.* 26, 11 (Nov. 1978), 1000–1017. DOI: https://doi.org/10.1177/26.11.569164

[60] Hologic. 2018. ThinPrep imaging system operation summary and clinical information. Retrieved from https://www.thinprep.com.au/.

[61] Md Shamim Hossain, Hamid A. Jalab, Fariha Zulfiquar, and Mahfuza Pervin. 2019. Renal cancer cell nuclei detection from cytological images using convolutional neural network for estimating proliferation rate. *J. Telecommun., Electr. Comput. Eng.* 11, 1 (2019), 63–71.

[62] Maciej Hrebień, Jozef Korbicz, and Andrzej Obuchowicz. 2007. Hough transform, (1+1) search strategy and watershed algorithm in segmentation of cytological images. In *Advances in Soft Computing*, M. Kurzynski, E. Puchala, M. Wozniak, and A. Zolnierzek (Eds.), Vol. 45. Springer, 550–557.

ACM Computing Surveys, Vol. 54, No. 3, Article 52. Publication date: April 2021.
[63] Maciej Hrebiel, Piotr Stec, Tomasz Nieczkowski, and Andrzej Obuchowicz. 2008. Segmentation of breast cancer fine needle biopsy cystological images. Appl. Math. Comput. Sci. 18 (2008), 159–170.

[64] Zilong Hu, Jinshan Tang, Ziming Wang, Kai Zhang, Lin Zhang, and Qingling Sun. 2018. Deep learning for image-based cancer detection and diagnosis—A survey. Pattern Recogn. 83, 11 (2018), 134–149. DOI: https://doi.org/10.1016/j.patcog.2018.05.014

[65] Jinjie Huang, Tao Wang, Dequan Zheng, and Yongjun He. 2020. Nucleus segmentation of cervical cytology images based on multi-scale fuzzy clustering algorithm. Bioengineering 11, 1 (2020), 484–501. DOI: https://doi.org/10.3390/bioengineering1101023

[66] Elima Hussain, Lipi B. Mahanta, Chandana Ray Das, and Ratna Kanta Talukdar. 2020. A comprehensive study on the multi-class cervical cancer diagnostic prediction on pap smear images using a fusion-based decision from ensemble deep convolutional neural network. Tissue Cell 65 (2020), 101347. DOI: https://doi.org/10.1016/j.tice.2020.101347

[67] J. Hyeon, H. J. Choi, B. D. Lee, and K. N. Lee. 2017. Diagnosing cervical cell images using pre-trained convolutional neural network as feature extractor. In Proceedings of the IEEE International Conference on Big Data and Smart Computing (BigComp’17). 390–393. DOI: https://doi.org/10.1109/BIGCOMP.2017.7881741

[68] Nanette Icho. 2000. The automation trend in cytology. Lab. Med. 31, 4 (04 2000), 218–221. Retrieved from arXiv:https://academic.oup.com/labmed/article-pdf/31/4/218/24957807/labmed31-0218.pdf.

[69] M. Iliyasu and Chastine Fatichah. 2017. A quantum hybrid PSO combined with fuzzy k-NN approach to feature selection and cell classification in cervical cancer detection. Sensors (Switzerland) 17, 12 (Dec. 2017), 1–17. DOI: https://doi.org/10.3390/s17122935

[70] Indrabayu, A. R. Fatmasari, and Ingrid Nurtanio. 2017. A colour space-based detection for cervical cancer using fuzzy C-means clustering. In ACM International Conference Proceeding Series, Vol. Part F1309. 137–141. DOI: https://doi.org/10.1145/3121138.3121196

[71] Nor Ashidi Mat Isa, Esugasini Subramaniam, Mohd Yusoff Mashor, and Nor Hayati Othman. 2007. Fine needle aspiration cytology evaluation for classifying breast cancer using artificial neural network. Amer. J. Appl. Sci. 4, 12 (2007), 999–1008.

[72] Z. Islam and M. A. Haque. 2015. Multi-step level set method for segmentation of overlapping cervical cells. In Proceedings of the IEEE International Conference on Telecommunications and Photonics (ICTP’15). 1–5. DOI: https://doi.org/10.1109/ICTP.2015.7427948

[73] S. Issac Niwas, P. Palanisamy, K. Sujathan, E. Bengtsson, S. Issac Niwas, P. Palanisamy, K. Sujathan, and E. Bengtsson. 2013. Analysis of nuclei textures of fine needle aspirated cytology images for breast cancer diagnosis using Complex Daubechies wavelets. Signal Process. 93, 10 (2013), 2828–2837. DOI: https://doi.org/10.1016/j.sigpro.2012.06.029

[74] Kourosh Jafari-Khouzani and Hamid Soltanian-Zadeh. 2005. Rotation-invariant multisresolution texture analysis using radon and wavelet transforms. IEEE Trans. Image Process. 14, 6 (2005), 783–795. DOI: https://doi.org/10.1109/TIP.2005.847302

[75] Lukasz Jeleń, Thomas Fevens, and Adam Krzyżak. 2008. Classification of breast cancer malignancy using cytological images of fine needle aspiration biopsies. Int. J. Appl. Math. Comput. Sci. 18, 1 (Mar. 2008), 75–83. DOI: https://doi.org/10.2478/v10006-008-0007-x

[76] Yang Jian and Zhou Yikai. 2004. Detection of DNA aneuploidy in exfoliated airway epithelia cells of sputum specimens by the automatic image cytomtery and its clinical value in the identification of lung cancer. J. Huazhong Univ. Sci. Technol. [Med. Sci.] 24, 4 (2004), 407–410.

[77] Xin Jing, Momin T. Siddiqui, and Qing Kay Li (Eds.). 2018. Atlas of Non-Gynecologic Cytology. Springer International Publishing.

[78] Asli Kale and Selim Aksoy. 2010. Segmentation of cervical cell images. In Proceedings of the International Conference on Pattern Recognition. 2399–2402.

[79] K. Kancherla and S. Mukkamala. 2013. Early lung cancer detection using nucleus segmentation-based features. In Proceedings of the IEEE Symposium on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB). 91–95. DOI: https://doi.org/10.1109/CIBCB.2013.6595393

[80] Michael Kass, Andrew Witkin, and Demetri Terzopoulos. 1988. Snakes: Active contour models. Int. J. Comput. Vision 1, 4 (Jan. 1988), 321–331. DOI: https://doi.org/10.1007/BF00135570

[81] S. Sajith Kecheril, D. Venkataraman, J. Suganthi, and K. Sujathan. 2015. Automated lung cancer detection by the Daubechies wavelets. Signal Processing (BigComp’2017). 390–393.

[82] Roger A. Kemp, Daniel M. Reinders, and Bojana Turic. 2007. Detection of lung cancer by automated sputum cytology evaluation for classifying breast cancer using artificial neural network. Amer. J. Appl. Sci. 4, 12 (2007), 999–1008.

[83] Aditya Khamparia, Deepak Gupta, Victor Hugo C. de Albuquerque, Arun Kumar Sangaiah, and Rutvij H. Jhaveri. 2020. Internet of health things-driven deep learning system for detection and classification of cervical cells using transfer learning. J. Supercomput. 76, 11 (Jan. 2020), 8590–8608.
[84] SanaUllah Khan, Naveed Islam, Zahoor Jan, Ikrum Ud Din, and Joel J. P. C. Rodrigues. 2019. A novel deep learning-based framework for the detection and classification of breast cancer using transfer learning. Pattern Recogn. Lett. 125 (2019), 1–6.

[85] S. U. Khan, N. Islam, Z. Jan, H. U. Shah, and A. ud Din. 2018. Automated counting of cells in breast cytology images using level set method. In Proceedings of the IEEE 4th International Conference on Data Science and Systems (HCPC/SmartCity/DSS’18). 1578–1584.

[86] Kwang-Baek Kim, Sungshin Kim, and Gwang-Ha Kim. 2007. Nucleus classification and recognition of uterine cervical pap-smears using FCM clustering algorithm. In Proceedings of the International Conference on Adaptive and Natural Computing Algorithms (ICANNGA’07), Bartlomiej Helieczynski, Andrzej Dzielinski, Marcin Iwanowski, and Bernadette Ribeiro (Eds.), Vol. 4432. Springer, Berlin, 290–299.

[87] Leopold G. Koss, Eunice Lin, Klaus Schreiber, Paul Elgert, and Laurie Mango. 1994. Evaluation of the PAPNET™ cytoplastic screening system for quality control of cervical smears. Amer. J. Clin. Pathol. 101, 2 (1994), 220–229.

[88] Marek Kowal, Pawel Filipczuk, Andrzej Obuchowicz, and Jozef Korbicz. 2011. Computer-aided diagnosis of breast cancer using Gaussian mixture cytoplastic image segmentation. J. Med. Info. Technol. 17 (2011), 257–262.

[89] Marek Kowal, Michal Zemko, Marcin Skobel, Jozef Korbicz, and Roman Monczak. 2020. Cell nuclei segmentation in cytological images using convolutional neural network and seeded watershed algorithm. J. Dig. Imag. 33, 1 (Feb. 2020), 231–242.

[90] Mohammed Kuko and Mohammad Pourhomayoun. 2020. Single and clustered cervical cell classification with ensemble and deep learning methods. Info. Syst. Front. 22, 5 (2020), 1039–1051. DOI: https://doi.org/10.1007/s10796-020-10028-1

[91] P. Kumar, S. L. Happy, S. Chatterjee, D. Sheet, and A. Routray. 2016. An unsupervised approach for overlapping cervical cell cytoplasm segmentation. In Proceedings of the IEEE EMBS Conference on Biomedical Engineering and Sciences (ICBES’16). 106–109. DOI: https://doi.org/10.1109/ICBES.2016.7843424

[92] Vineet Kumar and Hitesh Sharma. 2013. Extraction of nuclear region from sputum images through pixel classification for early lung cancer detection. Int. J. Comput. Appl. 72, 22 (6 2013), 35–39.

[93] N. Lassouaoui and L. Hamami. 2003. Genetic algorithms and multifractal segmentation of cervical cell images. In Proceedings of the 7th International Symposium on Signal Processing and Its Applications, Vol. 2. 1–4.

[94] Hansang Lee and Jummo Kim. 2016. Segmentation of overlapping cervical cells in microscopic images with superpixel partitioning and cell-wise contour refinement. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops. 1367–1373.

[95] Kyoung-mi Lee and Nick Street. 2000. Generalized hough transforms with flexible templates. In Proceedings of the International Conference on Artificial Intelligence (IC-AI’00). 1133–1139.

[96] O. Lezoray and H. Cardot. 2002. Histogram and watershed-based segmentation of color images. In Proceedings of the European Conference on Colour in Graphics, Imaging and Vision (CGIV’02). 358–362.

[97] Kuan Li, Zhi Lu, Wenyin Liu, and Jianping Yin. 2012. Cytoplasm and nucleus segmentation in cervical smear images using Radiating GVF Snake. Pattern Recogn. 45, 4 (2012), 2563–2571. DOI: https://doi.org/10.1016/j.patcog.2011.09.018

[98] Chen Horng Lin, Yung Kuan Chan, and Chun Chieh Chen. 2009. Detection and segmentation of cervical cell cytoplasm and nucleus. Int. J. Imag. Syst. Technol. 19, 3 (2009), 260–270. DOI: https://doi.org/10.1002/ima.20198

[99] Geert Litjens, Thijs Kooi, Babak Ehteshami Bejnordi, Arnaud Arindra Adiyoso Setio, Francesco Ciompi, Mohsen Ghafoorian, Jeroen A. W. M. van der Laak, Bram van Ginneken, and Clara L. Sánchez. 2017. A survey on deep learning in medical image analysis. Med. Image Anal. 42, 12 (2017), 60–88. DOI: https://doi.org/10.1016/j.media.2017.07.005

[100] Yiming Liu, Pengcheng Zhang, Qingche Song, Andi Li, Peng Zhang, and Zhiguo Gui. 2018. Automatic segmentation of cervical nuclei based on deep learning and a conditional random field. IEEE Access 6 (2018), 53709–53721.

[101] Zhi Lu, Gustavo Carneiro, and Andrew P. Bradley. 2013. Automated nucleus and cytoplasm segmentation of overlapping cervical cells. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI’13). K. Mori et al. (Ed.), Springer, Berlin, 452–460.

[102] K. Lu, G. Carneiro, and A. P. Bradley. 2015. An improved joint optimization of multiple level set functions for the segmentation of overlapping cervical cells. IEEE Trans. Image Process. 24, 4 (2015), 1261–1272. DOI: https://doi.org/10.1109/TIP.2015.2389619

[103] Jiheen Malak, Abderrahim Sebri, Souhir Mabrouk, Khoudoun Torki, and Rachid Torki. 2009. Automated breast cancer diagnosis based on GVFS-snake segmentation, wavelet features extraction and fuzzy classification. J. Signal Process. Syst. 55, 1–3 (4 2009), 49–66.

[104] Patrik Malm and Anders Brun. 2009. Closing curves with riemannian dilation: Application to segmentation in automated cervical cancer screening. In Advances in Visual Computing, Springer, Berlin, 337–346. DOI: https://doi.org/10.1007/978-3-642-10331-5_32

[105] Edwin Jayasingh Mariarupathum and Allwin Stephen. 2015. Nominated texture-based cervical cancer classification. Comput. Math. Methods Med. 2015 (Jan. 2015), 1–10. DOI: https://doi.org/10.1155/2015/586928
[106] Yannis Marinakis, Georgios Dounias, and Jan Jantzen. 2009. Pap smear diagnosis using a hybrid intelligent scheme focusing on genetic algorithm-based feature selection and nearest neighbor classification. *Comput. Biol. Med.* 39, 1 (2009), 69–78.

[107] Shahla Masood, Dian Feng, Osman Tutuncuoglu, and Gabor Fischer et al. 2011. Diagnostic value of imprint cytology during image-guided core biopsy in improving breast health care. *Ann. Clin. Lab. Sci.* 41, 1 (2011), 8–13.

[108] N. A. Mat-Isa and Nor Ashidi Mat Isa. 2005. Automated edge detection technique for Pap smear images using moving K-means clustering and modified seed-based region growing algorithm. *Int. J. Comput. Internet Manage.* 13, 3 (Jan. 2005), 45–59.

[109] Nor Ashidi Mat-Isa, Mohd Yusoff Mashor, and Nor Hayati Othman. 2008. An automated cervical pre-cancerous diagnostic system. *Artif. Intell. Med.* 42, 1 (2008), 1–11. DOI: https://doi.org/10.1016/j.artmed.2007.09.002

[110] B. Mehdi, D. Stacey, and G. Harauz. 1994. A hierarchical neural-network assembly for classification of cervical cells in automated screening. *Anal. Cell. Pathol.* 7, 3 (1994), 171–180.

[111] Xu Meiquan, Zeng Weixiu, Sun Yanhua, Wu Junhui, Wu Tingting, Yang Yajie, Zhang Meng, Zhu Zeji, and Chen Longsen. 2018. Cervical cytology intelligent diagnosis based on object detection technology. In *MIDL 2018 1st Conference*.

[112] Shyamali Mitra, Soumyajyoti Dey, Nilbaran Das, Sukanta Chakrabarty, Mita Nasipuri, and Mrinal Kanti Naskar. 2018. Identification of benign and malignant cells from cytological images using superpixel-based segmentation approach. In *Proceedings of the 52nd Annual Convention of CSI: Social Transformation and Digital Way*, Jyotsna Kumar Mandal and Devadatta Sinha (Eds.). Springer, 257–269.

[113] Theodoros Mouroutis, Stephen J. Roberts, and Anil A. Bharath. 1997. Robust cell nuclei segmentation using statistical modelling. *Bioimaging* 6, 2 (1998), 79–91. DOI: https://doi.org/10.1002/1361-6374(199806)6:2<79::AID-BIO3>3.0.CO;2-#.

[114] T. Mouroutis, S. J. Roberts, A. A. Bharath, and G. Alusi. 1997. Compact {Hough} {Transform} and a maximum likelihood approach to cell nuclei detection. *Proceedings of the 13th International Conference on Digital Signal Processing* 2 (1997), 869–872. DOI: https://doi.org/10.1109/ICDSP.1997.628493

[115] Izzati Muhimmah, Rahadian Kurniawan, and Indrayanti. 2012. Automated cervical cell nuclei segmentation using morphological operation and watershed transformation. In *Proceeding of the IEEE International Conference on Computational Intelligence and Cybernetics (CyberneticsCom’12)*, 163–167. DOI: https://doi.org/10.1109/CyberneticsCom.2012.6381639

[116] N. Mustafa, N. A. M. Isa, M. Y. Mashor, and N. H. Othman. 2007. Colour contrast enhancement on preselected cervical cell for ThinPrep images. In *Proceedings of the 3rd International Conference on Intelligent Hiding and Multimedia Signal Processing (IIH-MSP’07)*, Vol. 2. 209–212.

[117] Loris Nanni, Stefano Ghidoni, Sheryl Brahmam, Shaoxiong Liu, and Ling Zhang. 2020. Ensemble of handcrafted and deep learned features for cervical cell classification. In *Intelligent Systems Reference Library*. Springer International Publishing, 117–135.

[118] Nephron. 2010. Retrieved from https://upload.wikimedia.org/wikipedia/commons/8/88/Mesothelioma_cytology_1.jpg.

[119] Nephron. 2010. Retrieved from https://upload.wikimedia.org/wikipedia/commons/f/f0/Mucoepidermoid_carcinoma_2_-_high_mag.jpg.

[120] Nephron. 2010. Retrieved from https://upload.wikimedia.org/wikipedia/commons/0/02/Serous_carcinoma_2_-_high_mag.jpg.

[121] Nephron. 2015. Retrieved from https://upload.wikimedia.org/wikipedia/commons/1/19/Benign_mesoethelial_cells_-_pleural_fluid_-_high_mag.jpg.

[122] Nephron. 2016. Retrieved from https://upload.wikimedia.org/wikipedia/commons/2/2b/Benign_urine_cytology_-_3b_-_very_high_mag.jpg.

[123] S. I. Niwas, P. Palanisamy, and K. Sujathan. 2010. Wavelet-based feature extraction method for breast cancer cytology images. In *Proceedings of the IEEE Symposium on Industrial Electronics and Applications (ISIEA’10)*, 686–690. DOI: https://doi.org/10.1109/ISIEA.2010.5679377

[124] M. S. Nosrati and G. Hamarneh. 2015. Segmentation of overlapping cervical cells: A variational method with star-shape prior. In *Proceedings of the IEEE 12th International Symposium on Biomedical Imaging (ISBI’15)*, 186–189. DOI: https://doi.org/10.1109/ISBI.2015.7163846

[125] T. Ojala, M. Pietikainen, and T. Maenpaa. 2002. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Trans. Pattern Anal. Mach. Intell.* 24, 7 (2002), 971–987. DOI: https://doi.org/10.1109/TPAMI.2002.1017623

[126] M. Orozco-Monteagudo, T. Geerinck, C. Mihai, H. Sahli, and A. Taboada-Crispi. 2013. Pap-smear cell nucleus extraction by using morphological watershed. In *Proceedings of the International Federation of Medical and Biological Engineering (IFMBE’13)*, Vol. 33, 1070–1073. DOI: https://doi.org/10.1007/978-3-642-21198-0_272
Cytology Image Analysis Techniques Toward Automation: Systematically Revisited

\[\text{Maykel Orozco-Monteagudo, Cosmin Mihai, Hichem Sahli, and Alberto Taboada-Crispi. 2012. Combined hierarchical watershed segmentation and SVM classification for pap smear cell nucleus extraction. In Computación y Sistemas, Vol. 16. 133–145.}\]

\[\text{P. S. Oud, D. J. Zahnisre, R. Harbers-Hendriks, M. C. van Boekel, M. C. Raaijmakers, and P. G. Vooijs. 1981. The development of a cervical smear preparation procedure from the BioPEPR image analysis system. Anal. Quant. Cytol. 3, 1 (3 1981), 73–79.}\]

\[\text{Pei Yan Fai, Chin Chen Chang, and Yung Kuan Chan. 2012. Nucleus and cytoplasm contour detector from a cervical smear image. Expert Syst. Appl. 39, 1 (2012), 154–161. DOI: https://doi.org/10.1016/j.eswa.2011.06.034}\]

\[\text{Jürgen Pannek, Franziska Rademacher, and Jens Wöllner. 2017. Clinical usefulness of urine cytology in the detection of bladder tumors in patients with neurogenic lower urinary tract dysfunction. Res. Rep. Urol. 9 (2017), 219.}\]

\[\text{Liron Pantanowitz and Anil V. Parwani. 2014. Practical Informatics for Cytopathology. Springer.}\]

\[\text{Hady Ahmadly Phoulady and Peter R. Mouton. 2018. A new cervical cytology dataset for nucleus detection and image classification (cervix93) and methods for cervical nucleus detection. Retrieved from https://arXiv:1811.09651.}\]

\[\text{Marina E. Plissiti, C. Nikou, and A. Charchanti. 2011. Automated detection of cell nuclei in pap smear images using morphological reconstruction and clustering. Trans. Info. Tech. Biomed. 15, 2 (Mar. 2011), 233–241. DOI: https://doi.org/10.1109/TITB.2010.2087030}\]

\[\text{Marina E. Plissiti, Christophoros Nikou, and Antonia Charchanti. 2011. Combining shape, texture and intensity features for cell nuclei extraction in Pap smear images. Pattern Recogn. Lett. 32, 6 (2011), 838–853. DOI: https://doi.org/10.1016/j.patrec.2011.01.008}\]

\[\text{Marina E. Plissiti, Christophoros Nikou, and Senior Member. 2012. Overlapping cell nuclei segmentation using a spatially adaptive active physical model. IEEE Trans. Image Process. 21, 11 (2012), 4568–4580.}\]

\[\text{Marina E. Plissiti, Michalis Vrigkas, and Christophoros Nikou. 2015. Segmentation of cell clusters in Pap smear images using intensity variation between superpixels. In Proceedings of the 22nd International Conference on Systems, Signals and Image Processing (IWSSIP’15). 184–187.}\]

\[\text{Abraham Pouliakis, Efrossyni Karakitsou, Niki Margari, Panagiotis Bountris, Maria Haritou, John Panayiotides, Dimitrios Koutsouris, and Petros Karakitsos. 2016. Artificial neural networks as decision support tools in cytopathology: Past, present, and future. Biomed. Eng. Comput. Biol. 7 (Feb. 2016), 1–18. DOI: https://doi.org/10.4137/BECB.S31601.26917984[pmid].}\]

\[\text{S. Rachid, N. Niki, H. Nishitani, S. Nakamura, and S. Mori. 1997. Segmentation of sputum color image for lung cancer diagnosis. In Proceedings of International Conference on Image Processing, Vol. 1. 243–246. DOI: https://doi.org/10.1109/ICIP.1997.647750}\]

\[\text{S. Ragothaman, S. Narasimhan, M. G. Basavaraj, and R. Dewar. 2016. Unsupervised segmentation of cervical cell images using gaussian mixture model. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW’16). 1374–1379.}\]

\[\text{M. M. Rahaman, C. Li, X. Wu, Y. Yao, Z. Hu, T. Jiang, X. Li, and S. Qi. 2020. A survey for cervical cytopathology image analysis using deep learning. IEEE Access 8 (2020), 61687–61710.}\]

\[\text{Rahmadwati, Golshah Naghdy, Montserrat Ros, Catherine Todd, and Eviana Norahmawati. 2016. Cervical cancer classification using gabor filters. In Proceedings of the IEEE 1st International Conference on Healthcare Informatics, Imaging and Systems Biology. IEEE, 48–52. DOI: https://doi.org/10.1109/IIHSB.2011.15}\]

\[\text{D. A. Ramli, A. F. Kadmin, M. Y. Mashor, and Nor Ashidi Mat Isa. 2004. Diagnosis of cervical cancer using hybrid multilayered perceptron (HMLP) network. In Lecture Notes in Computer Science. Vol. 3213. 591–598.}\]

\[\text{Catherine S. Abendroth. 1998. Cytology: Diagnostic principles and clinical correlates. Authors: E. Cibas and B. Ducatman WB Saunders, Philadelphia, 1996. Diagn. Cytopathol. 16, 2 (Dec. 1998), 193. DOI: https://doi.org/10.1002/(SICI)1097-0339(19970216)193:AID-DC243.0.CO;2-G}\]

\[\text{S. Kar and D. D. Majumder. 2019. A novel approach of mathematical theory of shape and neuro-fuzzy-based diagnostic analysis of cervical cancer. Pathol. Oncol. Res. 25, 2 (2019) 777–790.}\]

\[\text{Monjoy Saha, Rashmi Mukherjee, and Chandan Chakraborty. 2016. Computer-aided diagnosis of breast cancer using cytological images: A systematic review. Tissue Cell 48, 5 (2016), 461–474. DOI: https://doi.org/10.1016/j.tice.2016.07.006}\]

\[\text{Ratna Saha, Mariusz Bajger, and Gobert Lee. 2016. Spatial shape constrained fuzzy C-Means (FCM) clustering for nucleus segmentation in pap smear images. In Proceedings of the International Conference on Digital Image Computing: Techniques and Applications (DICTA’16). 1–8.}\]

\[\text{R. Saha, M. Bajger, and G. Lee. 2019. SRM superpixel merging framework for precise segmentation of cervical nucleus. In Proceedings of the Conference on Digital Image Computing: Techniques and Applications (DICTA’19). 1–8.}\]

\[\text{Amartya Ranjan Saikia, Kangkana Bora, Lipi B. Mahanta, and Anup Kumar Das. 2019. Comparative assessment of CNN architectures for classification of breast FNAC images. Tissue Cell 57 (2019), 8–14.}\]

\[\text{T. A. Sajeena and A. S. Jereesh. 2015. Cervical cancer detection through automatic segmentation and classification of Pap smear cells. Int. J. Appl. Eng. Res. 10, 18 (9 2015), 39078–39084.}\]

ACM Computing Surveys, Vol. 54, No. 3, Article 52. Publication date: April 2021.
[150] Rachid Sammouda, Noboru Niki, Hiromu Nishitani, and Emi Kyokage. 1998. Segmentation of sputum color image for lung cancer diagnosis-based on neural network. *IEICE Trans. Inf. Syst.* E81-D, 8 (1998), 862–871.

[151] Parikshit Sanyal, Tanushri Mukherjee, Sanghita Barui, Avinash Das, and Prabaha Gangopadhyay. 2018. Artificial intelligence in cytopathology: A neural network to identify papillary carcinoma on thyroid fine-needle aspiration cytology smears. *J. Pathol. Info.* 9 (Dec. 2018), 43.

[152] Dharmesh A. Sarvaiya and Mehul Barot. 2014. Detection of lung cancer using sputum image segmentation. *Int. J. Adv. Eng. Res. Dev.* 1, 6 (2014), 1–9.

[153] Abid Sarwar, Abrar Ali Sheikh, Jatinder Manhas, and Vinod Sharma. 2020. Segmentation of cervical cells for automated screening of cervical cancer: A review. *Artific. Intell. Rev.* 53, 4 (Apr. 2020), 2341–2379. DOI: https://doi.org/10.1007/s10462-019-09735-2

[154] Marin Scalbert, Florent Couzinie-Devy, and Riadh Fezzani. 2019. Generic isolated cell image generator. *Cytoometry Part A* 95, 11 (Oct. 2019), 1198–1206. DOI: https://doi.org/10.1002/cyto.a.23899

[155] L. Shajy, Varghese Paul, P. Marichamy, and K. Sujathan. 2015. Analysis of textures of sputum cytology images for lung cancer diagnosis using discrete wavelets transform. *Int. J. Appl. Eng. Res.* 10, 4 (1 2015), 9217–9229.

[156] Howard M. Shapiro. 2004. The evolution of cytomters. *Cytoometry Part A* 58, 1 (2004), 13–20.

[157] B. Sharma and K. K. Mangat. 2016. An improved nucleus segmentation for cervical cell images using FCM clustering and BPNN. In *Proceedings of the International Conference on Advances in Computers, Communications and Informatics (ICACCI’16)*, 1924–1929. DOI: https://doi.org/10.1109/ICACCI.2016.7732332

[158] Sanjay Kumar Singh and Anjali Goyal. 2017. A novel approach to segment nucleus of uterine cervix pap smear cells using watershed segmentation. In *Advanced Informatics for Computing Research*, D. Singh, R. Balasubramanian, A. K. Luhach, and P. Lingras (Eds.). Springer, 164–174.

[159] S. Anantha Sivaprasakam and E. R. Naganathan. 2017. Segmentation and classification of cervical cytology images using morphological and statistical operations. *ICTACT J. Image Video Process.* 7 (2017), 1445–1455.

[160] N. Sompawong, J. Mopan, P. Pooprasert, W. Himakhun, K. Suwannarurk, J. Ngaanvirojcharoen, T. Vachiramon, and C. Tantibundhit. 2019. Automated pap smear cervical cancer screening using deep learning. In *Proceedings of the 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC’19)*, 7044–7048.

[161] Y. Song, E. L. Tan, X. Jiang, J. Z. Cheng, D. Ni, S. Chen, B. Lei, and T. Wang. 2017. Accurate cervical cell segmentation from overlapping clumps in pap smear images. *IEEE Trans. Med. Imag.* 36, 1 (1 2017), 288–300. DOI: https://doi.org/10.1109/TMI.2016.2606380

[162] Youyi Song, Ling Zhang, Siping Chen, Dong Ni, Baiying Lei, and Tianfu Wang. 2015. Accurate segmentation of cervical cytoplasm and nuclei based on multiscale convolutional network and graph partitioning. *IEEE Trans. Biomed. Eng.* 62, 10 (2015), 2421–2433.

[163] Youyi Song, Ling Zhang, Siping Chen, Dong Ni, Baopu Li, Yongjing Zhou, Baiying Lei, and Tianfu Wang. 2014. A deep learning-based framework for accurate segmentation of cervical cytoplasm and nuclei. In *Proceedings of the 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, 2903–2906.

[164] Y. Srinivasan, B. Nutter, S. Mitra, B. Phillips, and E. Sinzinger. 2006. Classification of cervix lesions using filter bank-based texture mode. In *Proceedings of the 19th IEEE Symposium on Computer-Based Medical Systems (CBMS’06)*, 832–840. DOI: https://doi.org/10.1109/CBMS.2006.66

[165] W. N. Street. 2000. Xcyt: A System for remote cytological diagnosis and prognosis of breast cancer. In *Artificial Intelligence Techniques in Breast Cancer Diagnosis and Prognosis*. Series in Machine Perception and Artificial Intelligence, Vol. 39. World Scientific Publishing Co., 297–326.

[166] P. Sukumar and R. K. K. Gnanamurthy. 2016. Computer aided detection of cervical cancer using pap smear images based on adaptive neuro fuzzy inference system classifier. *J. Med. Imag. Health Info.* 6, 2 (2016), 312–319. DOI: https://doi.org/10.1166/jmihi.2016.1690

[167] Bilal Taha, Jorge Dias, and Naoufel Werghi. 2017. Classification of cervical-cancer using pap-smear images: A convolutional neural network approach. In *Medical Image Understanding and Analysis*, M. V. Hernández and V. González-Castro (Eds.). Springer, 261–272.

[168] F. Taher and R. Sammouda. 2007. Identification of lung cancer based on shape and color. In *Proceedings of the Innovations in Information Technologies (IIT’07)*, 481–485. DOI: https://doi.org/10.1109/IIT.2007.4430454

[169] Fatma Taher, Rachid Sammouda, Fatma Taher, and Rachid Sammouda. 2010. Artificial neural network and fuzzy clustering methods in segmenting sputum color images for lung cancer diagnosis. In *Image and Signal Processing, Abderrahim Elmoataz, Olivier Lezoray, Fathallah Nouboud, Driss Mammass, and Jean Meunier (Eds.). Vol. 3*. Springer, Berlin, 513–520.

[170] F. Taher, N. Werghi, and H. Al-Ahmad. 2012. Bayesian classification and artificial neural network methods for lung cancer early diagnosis. In *Proceedings of the 19th IEEE International Conference on Electronics, Circuits, and Systems (ICECS’12)*, 773–776. DOI: https://doi.org/10.1109/ICECS.2012.6463545

ACM Computing Surveys, Vol. 54, No. 3, Article 52. Publication date: April 2021.
[171] F. Taher, N. Werghi, and H. Al-Ahmad. 2015. Computer aided diagnosis system for early lung cancer detection. In Proceedings of the International Conference on Systems, Signals and Image Processing (IWSSIP’15). 5–8. DOI: https://doi.org/10.1109/IWSSIP.2015.7313923

[172] Noboru Tanaka, Hideo Ikeda, Tetsuo Ueno, Akio Mukawa, Sadakazu Watanabe, Kozo Okamoto, Sugao Hosoi, and Sho Tsunekawa. 1987. Automated cytologic screening system (CYBEST model 4): An integrated image cytometry system. Appl. Optics 26, 16 (1987), 3301–3307.

[173] Arti Taneja, Priya Ranjan, and Amit Ujlayan. 2018. Multi-cell nuclei segmentation in cervical cancer images by integrated feature vectors. Multimedia Tools Appl. 77, 8 (Apr. 2018), 9271–9290. DOI: https://doi.org/10.1007/s11042-017-4864-x

[174] J. R. Tang, N. A. M. Isa, and E. S. Ch’ng. 2015. Segmentation of cervical cell nucleus using intersecting cortical model optimized by particle swarm optimization. In Proceedings of the IEEE International Conference on Control System, Computing and Engineering (ICCSCE’15). 111–116.

[175] Afaf Tareef, Yang Song, Heng Huang, Dagan Feng, Mei Chen, Yue Wang, and Weidong Cai. 2018. Multi-pass fast watershed for accurate segmentation of overlapping cervical cells. IEEE Trans. Med. Imag. 37, 9 (September 2018), 2044–2059.

[176] Afaf Tareef, Yang Song, Heng Huang, Yue Wang, Dagan Feng, Mei Chen, and Weidong Cai. 2017. Optimizing the cervix cytological examination based on deep learning and dynamic shape modeling. Neurocomputing 248 (2017), 28–40.

[177] Walter E. Tolles. 1955. The cytoanalyzer—an example of physics in medical research. Trans. N.Y. Acad. Sci. 173 (1955), 250–6.

[181] Meng-Husiun Tsai, Yung-Kuan Chan, Zhe-Zheng Lin, Shys-Fan Yang-Mao, and Po-Chi Huang. 2008. Nucleus and cytoplasm contour detector of cervical smear image. Pattern Recogn. Lett. 29, 9 (7 2008), 1441–1453. DOI: https://doi.org/10.1016/j.patrec.2008.02.024

[182] J. H. Tucker. 1976. Cerviscan: An image analysis system for experiments in automatic cervical smear prescreening. Comput. Biomed. Res. 9, 2 (1976), 93–107. DOI: https://doi.org/10.1016/0010-4809(76)90033-1

[183] Daniela M. Ushizima, Andrea Gomes Campos Bianchi, and Claudia M. Carneiro. 2014. Segmentation of subcellular compartments combining superpixel representation with Voronoi diagrams. In Proceedings of the Overlapping Cervical Image Segmentation Challenge (ISBI’14). 1–2.

[184] Ed Uthman. 2010. Retrieved from https://live.staticflickr.com/4055/4706650000_c62c6c9b4e_b.jpg

[185] Philip T. Valente and H. Daniel Schantz. 2001. Cytology automation: An overview. Lab. Med. 32, 11 (2001), 686–690.

[186] Tao Wan, Shusong Xu, Chen Sang, Yulan Jin, and Zengchang Qin. 2019. Accurate segmentation of overlapping cells in cervical cytology with deep convolutional neural networks. Neurocomputing 365 (2019), 157–170. DOI: https://doi.org/10.1016/j.neucom.2019.06.086

[188] T. Wang, J. Huang, D. Zheng, and Y. He. 2020. Nucleus segmentation of cervical cytology images based on depth information. IEEE Access 8 (2020), 75846–75859.

[189] Sadakazu Watanabe. 1974. An automated apparatus for cancer prescreening: CYBEST. Comput. Graph. Image Process. 3, 4 (1974), 350–358.

[190] N. Werghi, C. Donner, F. Taher, and H. Al-Ahmad. 2012. Detection and segmentation of sputum cell for early lung cancer detection. In Proceedings of the 19th IEEE International Conference on Image Processing. 2813–2816. DOI: https://doi.org/10.1109/ICIP.2012.6467484

[191] B. Weyn, G. van de Wouwer, A. van Daele, P. Scheunders, D. van Dyck, E. van Marck, and W. Jacob. 1998. Automated breast tumor diagnosis and grading based on wavelet chromatin texture description. Cytometry 33, 1 (9 1998), 32–40.

[192] WHO. 2018. WHO (World Health Organization) Report. Retrieved from https://www.who.int/cancer/prevention/diagnosis-screening/cervical-cancer/en/.
[193] Wasswa William, Andrew Ware, Annabella Habinka Basaza-Ejiri, and Johnes Obungoloch. 2019. A review of image analysis and machine learning techniques for automated cervical cancer screening from pap-smear images. *Acta Cytol.* 55, 3 (2011), 227–238.

[194] Wasswa William, Andrew Ware, Annabella Habinka Basaza-Ejiri, and Johnes Obungoloch. 2017. Automated segmentation of cell nuclei in cytology pleural fluid images using OTSU thresholding. In *Proceedings of the International Conference on Digital Arts, Media and Technology (ICDAMT’17)*. 14–18.

[195] Khin Yadanar Win, Somsak Choomchuay, Kazuhiko Hamamoto, and Manasanan Raveesunthornkiat. 2017. Automated segmentation of cell nuclei in cytology pleural fluid images using OTSU thresholding. In *Proceedings of the International Conference on Digital Arts, Media and Technology (ICDAMT’17)*. 14–18.

[196] Chenyang Xu and Jerry L. Prince. 1998. Snakes, shapes, and gradient vector flow. *IEEE Trans. Image Process.* 7, 3 (1998), 359–369.

[197] Lin Yang, Peter Meer, and David J. Foran. 2005. Unsupervised segmentation based on robust estimation and color active contour models. *IEEE Trans. Info. Technol. Biomed.* 9, 3 (2005), 475–486. DOI: https://doi.org/10.1109/TITB.2005.847515

[198] Ziming Zeng, Siping Chen, Sheng Tang, and Lidong Yin. 2016. Unsupervised segmentation of cell nuclei in cervical smear images using active contour with adaptive local region fitting energy modelling. In *Proceedings of the 8th ACM Computing Surveys, Vol. 54, No. 3, Article 52. Publication date: April 2021.*
[214] J. Zhang, J. He, T. Chen, Z. Liu, and D. Chen. 2019. Abnormal region detection in cervical smear images based on fully convolutional network. *IET Image Process.* 13, 4 (2019), 583–590.

[215] J. Y. Zhang and Y. X. Liu. 2004. Cervical cancer detection using SVM-based feature screening. In *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention 2004*, C. Barillot, D. R. Haynor, and P. Hellier (Eds.), Vol. 3217. Springer, Berlin, 873–880.

[216] Ling Zhang, Siping Chen, Tianfu Wang, Yan Chen, Shaoxiong Liu, and Minghua Li. 2011. A practical segmentation method for automated screening of cervical cytology. In *Proceedings of the International Conference on Intelligent Computation and Bio-Medical Instrumentation*. IEEE, 140–143.

[217] Ling Zhang, Hui Kong, Chien Ting Chin, Shaoxiong Liu, Zhi Chen, Tianfu Wang, and Siping Chen. 2014. Segmentation of cytoplasm and nuclei of abnormal cells in cervical cytology using global and local graph cuts. *Comput. Med. Imag. Graphics* 38, 5 (2014), 369–380.

[218] Ling Zhang, Hui Kong, Shaoxiong Liu, Tianfu Wang, Siping Chen, and Milan Sonka. 2017. Graph-based segmentation of abnormal nuclei in cervical cytology. *Comput. Med. Imag. Graphics* 56 (3 2017), 38–48.

[219] Ling Zhang, Le Lu, Isabella Nogues, Ronald M. Summers, Shaoxiong Liu, and Jianhua Yao. 2017. DeepPap: Deep convolutional networks for cervical cell classification. *IEEE J. Biomed. Health Info.* 21, 6 (2017), 1633–1643. DOI: https://doi.org/10.1109/JBHI.2017.2705583

[220] Jie Zhao, Lei Dai, Mo Zhang, Fei Yu, Meng Li, Hongfeng Li, Wenjia Wang, and Li Zhang. 2019. PGU-net: Progressive Growing of U-net for Automated Cervical Nuclei Segmentation. In *Multiscale Multimodal Medical Imaging*. Springer International Publishing, 51–58.

[221] Lili Zhao, Kuan Li, Mao Wang, Jianping Yin, En Zhu, Chengkun Wu, Siqi Wang, and Chengzhang Zhu. 2016. Automatic cytoplasm and nuclei segmentation for color cervical smear image using an efficient gap-search MRF. *Comput. Biol. Med.* 71 (2016), 46–56.

[222] L. Zhao, K. Li, J. Yin, Q. Liu, and S. Wang. 2017. Complete three-phase detection framework for identifying abnormal cervical cells. *IET Image Process.* 11, 4 (2017), 258–265.

[223] Yanning Zhou, Hao Chen, Jiaqi Xu, Qi Dou, and Pheng-Ann Heng. 2019. IRNet: Instance relation network for overlapping cervical cell segmentation. In *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 640–648.

Received November 2019; revised November 2020; accepted January 2021