Check for
updates

A systematic review of deep learning-based cervical cytology screening: from cell identification to whole slide image analysis

Peng Jiang[1] · Xuekong Li[2] · Hui Shen[2] · Yuqi Chen[1] · Lang Wang[1] · Hua Chen[1] · Jing Feng[1] · Juan Liu[1]

Abstract
Cervical cancer is one of the most common cancers in daily life. Early detection and diagnosis can effectively help facilitate subsequent clinical treatment and management. With the growing advancement of artificial intelligence (AI) and deep learning (DL) techniques, an increasing number of computer-aided diagnosis (CAD) methods based on deep learning have been applied in cervical cytology screening. In this paper, we survey more than 80 publications since 2016 to provide a systematic and comprehensive review of DL-based cervical cytology screening. First, we provide a concise summary of the medical and biological knowledge pertaining to cervical cytology, since we hold a firm belief that a comprehensive biomedical understanding can significantly contribute to the development of CAD systems. Then, we collect a wide range of public cervical cytology datasets. Besides, image analysis approaches and applications including cervical cell identification, abnormal cell or area detection, cell region segmentation and cervical whole slide image diagnosis are summarized. Finally, we discuss the present obstacles and promising directions for future research in automated cervical cytology screening.

Keywords Cervical cytology · Deep learning · Cancer screening · Artificial intelligence · Cytopathology

1 Introduction

1.1 Introduction of cervical cancer

Cervical cancer is a common malignancy that poses a serious threat to women's health. It is the fourth most common cancer in terms of both incidence and mortality. In 2020,

✉ Juan Liu
liujuan@whu.edu.cn

[1] Institute of Artificial Intelligence, School of Computer Science, Wuhan University, Wuhan 430072, China

[2] Wei County Maternal and Child Health Hospital, Handan 056000, China

approximately 600,000 new cases of cervical cancer were diagnosed and more than 340,000 people died from this disease globally (Sung et al. 2021). The incidence and mortality of cervical cancer may vary among countries and regions, which is related to the level of health services, the implementation of screening and prevention measures, lifestyle, and environmental factors in that region (Cramer 1974).

Cervical cancer is a kind of malignant tumor arising from the cervix and threatens the life and health of women. There are two main types of cervical cancer: (1) squamous cell carcinoma (SCC); and (2) adenocarcinoma. About 90% of cervical cancer cases are SCC, most of which begin in the transformation zone and develop from cells in the outer part of the cervix (Waggoner 2003). Cervical cancer is by far the most common human papillomavirus (HPV) related disease, and almost all cervical cancers (more than 95%) are caused by persistent infection with some types of HPV. There are at least 13 known types of HPV that can persist and progress to cancer, called high-risk HPV, the most common being HPV 16 and 18 strains. Cervical cancer has a long period of the precancerous stage, and its development is continuous, as shown in Fig. 1. The main characteristics of precancerous cells focus on changes in the nucleus. For example, nuclear enlargement results in an increased nuclear-to-cytoplasmic ratio (N/C ratio). It's common to see Binucleation and multinucleation. Besides, nucleoli are generally absent or inconspicuous if present. The contour of the nuclear membrane is quite irregular. Early forms of cervical cancer may have no symptoms or signs, but as the disease progresses, symptoms such as abnormal vaginal bleeding, vaginal discharge, and pelvic pain may appear. Thus, early diagnosis is crucial for the treatment and prognosis of cervical cancer (Cohen et al. 2019). There is compelling evidence that cervical cancer is one of the most preventable and treatable cancers if detected early and managed effectively through regular screening programs. There are currently three World Health Organization (WHO) recommended screening tests for cervical cancer: (1) HPV testing for high-risk HPV types; (2) cervical cytology screening; and (3) visual inspection with acetic acid (VIA). Nowadays, cervical cytology screening has been the basic method worldwide since the cytological features are significant indications of cervical cancer.
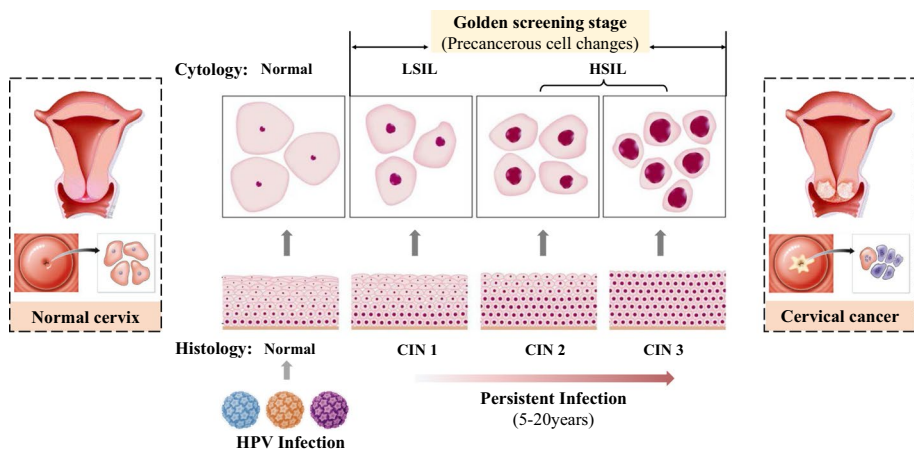

Fig. 1 The natural evolution of HPV-infected cervical cancer

1.2 Motivation of this review

Traditionally, cervical cytology screening program requires the manual identification of abnormal cells under a microscope, which is time-consuming, tedious, and error-prone (Elsheikh et al. 2013). In this context, an increasing number of automatic screening systems have been proposed to reduce the burden on cytopathologists and improve diagnosis efficiency (Koss et al. 1994; Biscotti et al. 2005; Kardos 2004). With the advancement of artificial intelligence (AI) and digital image processing, machine learning (ML) technology has been widely applied in cervical cytology screening to analyze cytological images due to its high-performance results (Marinakis et al. 2009; Chen et al. 2013; William et al. 2018). Nevertheless, traditional machine learning approaches have complex image pre-processing and feature selection steps that limit the further progress of human–machine collaboration.

In the past few years, deep learning (DL), a branch of machine learning, has exploded in the field of computer vision (LeCun et al. 2015; Krizhevsky et al. 2017; Simonyan and Zisserman 2015; He et al. 2016; Ren et al. 2015). The end-to-end automatic feature extraction and learning process of DL eliminates the need for manual feature design and selection. DL has made a breakthrough in various fields of image processing, medical image analysis is no exception (Litjens et al. 2017; Liu et al. 2019; Rajpurkar et al. 2022). DL solutions have been successfully applied in many medical imaging tasks, such as thoracic Imaging, neuroimaging, cardiovascular imaging, abdominal imaging, and microscopy imaging (Zhou et al. 2021). The development of DL has also greatly accelerated automatic image analysis in cervical cytology screening. To understand the popularity and development trend of deep learning in cervical cytology, multiple literature databases (PubMed, Scopus, IEEE Xplore, ACM Digital Library, and Web of Science) are searched using the keywords related to cervical cytology screening (cervical cytology, cervical cancer diagnosis, deep learning, Pap smear, etc.). Fig. 2 illustrates the
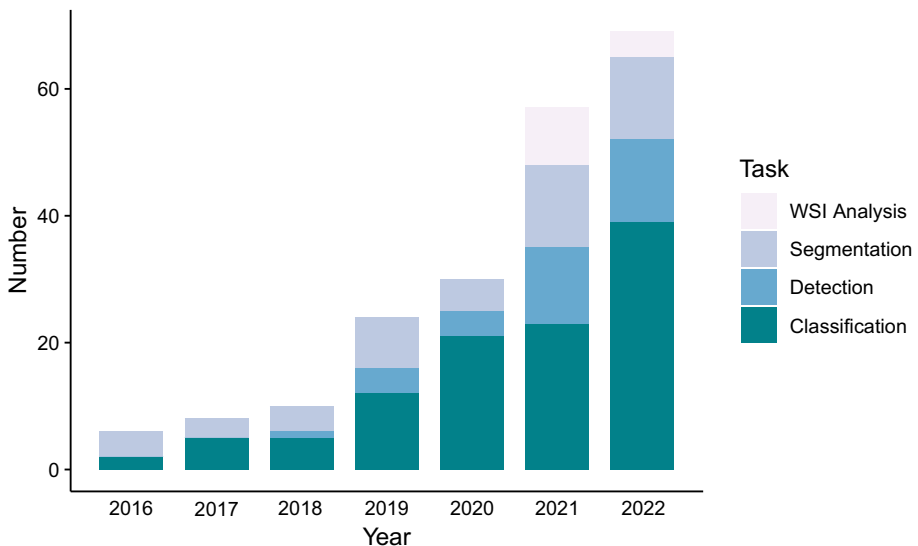
Fig. 2 Number of publications in DL-based classification, detection, segmentation and WSI analysis for automated cervical cytology

number of related publications from 2016 to 2022. Since 2016, there has been a notable surge in the use of DL for cervical cytology screening. Moreover, the object detection task has experienced significant growth since 2018, while the task of whole slide image (WSI) analysis has emerged in 2021 and shown impressive expansion recently.

There exist several surveys in the field of automated cervical cytology (William et al. 2018; Rahaman et al. 2020; Conceição et al. 2019; Chitra and Kumar 2022; Hou et al. 2022; Shanthi et al. 2022). Although these reviews provide valuable insights into automated cervical cytology, they are not exhaustive and some areas remain unexplored, calling for a further comprehensive investigation. First of all, the above reviews focus on classification and segmentation tasks at the cell level, and none of them investigate the application of object detection algorithms in automated cervical cytology screening. Secondly, the majority of these reviews primarily concentrate on conventional machine learning approaches, with comparatively limited coverage of DL-based methods. Moreover, few reviews provide biomedical context pertaining to cervical cytology, which is relevant for understanding the applicability of DL-based methods in this field. Last but not least, there is currently no review specialized for automatic WSI analysis of cervical cytology as the related works have only recently started to emerge in 2021. Automatic WSI analysis of cervical cytology holds great promise for improving the efficiency and accuracy of cervical cancer screening. Staying abreast of the latest developments and advancements in this field will be important for researchers and practitioners.

1.3 Contribution and organization of paper

To address the above issues, a comprehensive overview of relevant works for automated cervical cytology is presented in this survey including over 80 publications since 2016. For researchers just entering this field, this survey provides background knowledge on cervical cytology such as a brief introduction to cervical cancer, popular cervical cytology screening procedures, and definite cell categories in the Bethesda system (TBS). It is worth noting that the comparison of different reporting terms is elaborated as well. This can often cause confusion and impact the construction of a correct and reasonable DL model. In addition, the historical development of automated screening systems and the specific tasks in cervical cytology screening have been introduced in detail. Besides, this survey has also compiled the most extensive collection of publicly available cervical cytology image datasets. Moreover, this survey summarizes the latest DL-based classification, detection, segmentation, and WSI analysis methods in automated cervical cytology screening. Towards the end of this paper, several challenges and opportunities (stain normalization, image super-resolution, incorporating medical domain knowledge, annotation-efficient learning, internet of medical things, etc.) are presented that may provide promising research directions in cervical cytology screening.

The paper is organized as follows: Sect. 1 introduces the background and objective of this survey. In Sect. 2, an overview of biomedical knowledge related to cervical cytology is provided. Section 3 elaborates on the research methodology to construct this systematic review. Section 4 lists the public datasets in cervical cytology screening and summarizes the detailed progress in the DL-based automated cervical cytology from cell identification to WSI analysis. In Sect. 5, existing challenges and potential opportunities in automated cervical cytology screening are discussed. Finally, Sect. 6 concludes this review paper.

2 Overview of cervical cytology

Before the review of deep learning-based methods for cervical cytology screening, a preliminary overview of cervical cytology is presented in this section. We believe that medical and biological domain knowledge has a critical impact on the construction of computational models and the design of Computer-aided diagnosis (CAD) systems. In Sect. 2.1, the detailed procedure of cervical cytology screening is first described. After that, we introduce the history of reporting terminology for cervical cytology and explain the corresponding relations and differences between the four reporting systems in Sect. 2.2. Next, we elaborate the cell categories in TBS in Sect. 2.3. Finally, the historical development of automated screening systems are briefly introduced in Sect. 2.4.

2.1 Procedure of cervical cytology screening

Cervical cytology screening is the most effective and widely used screening program for discovering cancerous or precancerous lesions. The primary goal of screening is to identify abnormal cervical cells with severe cell changes so that they can be monitored or treated in time to prevent the development of invasive cancer (Sankaranarayanan et al. 2001). A large number of medical organizations suggest conducting routine cervical cytology screening every few years. Currently, conventional Papanicolaou smear (CPS) test and liquid-based cytology (LBC) are performed for cervical cytology screening worldwide (Siebers et al. 2009). CPS is a procedure in which cervical cells are scraped and observed under a microscope. Figure 3a illustrates the whole process of CPS. Under the guidance of a vaginal speculum, a soft brush will insert into the vagina to collect cells from the cervix. Then a pap smear can be acquired by evenly spreading the cells from the brush onto the glass slide. After staining, cytologists can observe the sample under a microscope and make a diagnosis.

Due to the influence of blood, mucus, inflammation, and other factors, CPS often acquires blurred samples, resulting in poor imaging results and detection errors. In recent years, with the improvement of sample preparation level, LBC can significantly improve the imaging quality of cervical cell samples, and thus has gradually become the mainstream implement for cervical cytology screening. As shown in Fig. 3b, the collected cells will be placed in a preservation solution for further process. After oscillation and centrifugation, a liquid-based glass slide can be obtained by natural sedimentation. Then, the liquid-based sample preparation is completed via staining and air drying. Nowadays, with the development of imaging equipment and digital processing techniques, CPS and LBC samples are usually transformed into digital slides via pathological scanners to facilitate retrospective examination. Digital pathology brings a positive and profound impact on traditional pathological diagnosis, which digitizes glass slides into whole slide images (WSIs) to greatly reduce the workload of pathologists and improve the diagnosis efficiency compared to microscope-based visual observation (Al-Janabi et al. 2012; Niazi et al. 2019). Liquid-based preparation together with digital slides is a satisfactory alternative to conventional smear and has a great application prospect for nowadays' large-scale cervical cancer screening programs.
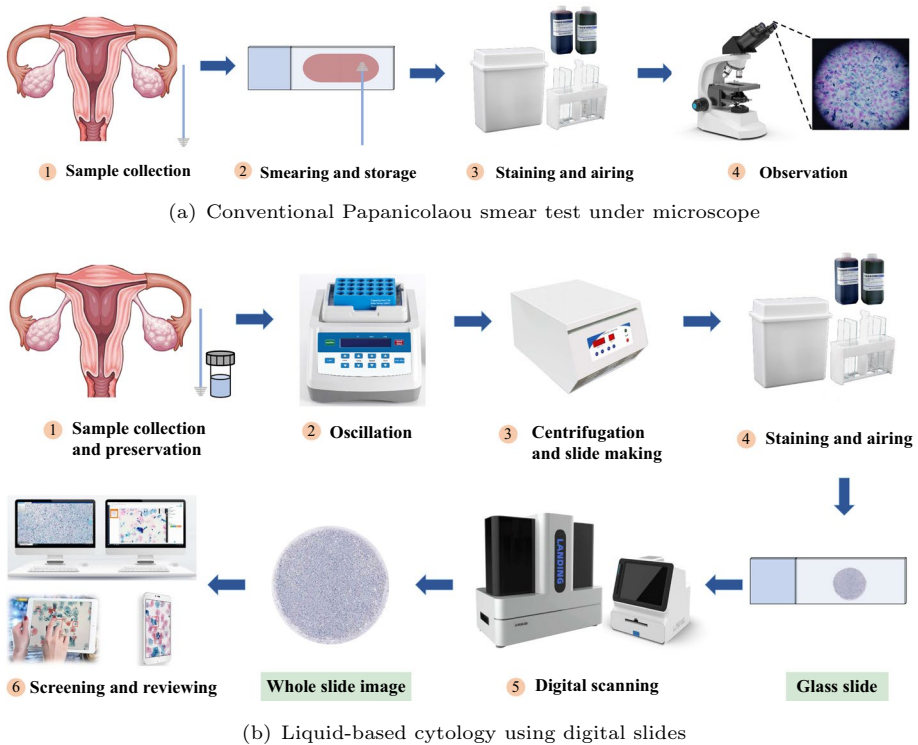
(a) Conventional Papanicolaou smear test under microscope

(b) Liquid-based cytology using digital slides

Fig. 3 Two prevalent procedures of cervical cytology screening

2.2 History of reporting terminology

The establishment of a standard cervical cytology report system plays a vital role in the universality of diagnosis methods and the acceptance of diagnosis results. In practice, the standard report system can cross the gap between different regions and different countries, strengthen the exchange of relevant scientific research results, and greatly improve the efficiency of cervical cancer diagnosis (St Clair and Wright 2009). The earliest report system for cervical cytological diagnosis was the Papanicolaou classification system, which developed a numeric classification terminology to grade cervical cells for 5 levels (Traut and Papanicolaou 1943). Class I to Class V respectively indicated the absence of abnormal or atypical cells; atypical cells, but no evidence of malignancy; cytology suggestive of but not conclusive for malignancy; cytology strongly suggestive of malignancy; and cytology conclusive for malignancy. However, many pointed out that the Papanicolaou classification system was strongly subjective and there was no strict objective standard for the difference between Class II, III, and IV. In addition, the Papanicolaou classification system did not have a clear definition of precancerous lesions and was not able to correspond to histopathological diagnosis terms.

With the development and refinement of both cytological and histological diagnoses of cervical cancer, an understanding of the natural history of cervical intraepithelial neoplasms (CIN) has developed progressively. The term dysplasia was introduced to refer to precancerous abnormalities of squamous cells and the 3-tiers dysplasia system

(mild/moderate/severe dysplasia, or carcinoma in situ) was proposed (Reagan et al. 1953). Recognizing the difficulty in differentiating severe dysplasia and carcinoma in situ (CIS), in 1966 (RICHART 1967), the CIN classification system was developed to describe CIN as a continuum of neoplastic change with progressively increasing risk of invasion, which was subdivided into grades I, II, and III. The advantage of both the 3-tiers dysplasia system and CIN classification was the ability to use it for cytological as well as histological samples.

In the 1970 s and 1980 s, as HPV testing became more available, vast epidemiological and biochemical evidence manifested the link between HPV and cervical dysplasia, which supported the role of high-risk HPV as a necessary factor in the development of cervical cancer (Hausen 1977; Crum et al. 1985). As a result, the first edition of the Bethesda system (TBS) for reporting cervical cytology was promulgated in 1988 (Workshop 1989). TBS aims to provide a uniform interpretation of cervical cytology, thereby facilitating communication between the clinician and the laboratory. With the change in practice to increased utilization of new technologies and findings in the last few decades, such as further insights into HPV biology and the development of liquid-based preparations, TBS has been updated three times to meet the evolving cervical cytology. The newest TBS 2014 guideline (Nayar and Wilbur 2015) offers comprehensive terminology for the reporting of cervical cytology.

Since TBS was established to better unify the reporting system, nowadays TBS has been the most standard reporting term for cervical cytology. 3-tier dysplasia or CIN systems are no longer used to report cervical cytology. However, when abnormalities are detected by cervical cytology, a further histological biopsy is performed. CIN system remains the standard reporting terminology for cervical histopathology and is used for the final diagnosis of cervical cancer because histopathology is the 'gold standard' for determining cancer. Referring to Herbert et al. (2007), Kedra et al. (2012), the specific classification criteria and corresponding relations of these four systems are shown in Table 1.

Table 1 Different cervical pathology reporting systems

Histopathology		Cytopathology	
3-tier dysplasia	CIN	The Pap classification	TBS
Normal	Normal	I (Normal)	NIML
Atypical cells	Squamous atypia	II (Atypical)	ASCUS
			ASCH
			AGC
Mild dysplasia	CIN1	III (Suspicious)	LSIL
Moderate dysplasia	CIN2		HSIL
Severe dysplasia carcinoma in situ	CIN3	IV (Suggestive)	
Invasive cancer	Invasive cancer	V (Indicative)	SCC
			Adenocarcinoma

2.3 Cell categories in TBS

TBS lays the foundation for our further comprehension of HPV biology and provides the necessary framework for the development of systematic evidence-based guidelines for cervical cancer screening and management. Since TBS is the widely recognized standard for cervical cytology reporting, in this section, cell categories in the latest version of TBS (TBS 2014) (Nayar and Wilbur 2015) are introduced for a better understanding of cervical cytology.

The specimen is reported as negative for intraepithelial lesion or malignancy (NIML) when there is no cellular evidence of neoplasia or epithelial abnormalities. Normal cellular elements include normal squamous cells and glandular cells. Squamous cells located in different positions of cervical epithelium have different characteristics, from shallow to deep can be divided into the superficial cell, intermediate cell, parabasal cell and basal cell. In a stained sample, the cytoplasm of superficial cells is pink or orange while the cytoplasm of all of the less mature cells is light green or cyan. Superficial cells and intermediate cells are large polygonal with a very low nuclear-to-cytoplasmic ratio (N/C ratio) while parabasal cells and basal cells are generally round or oval with a relatively high N/C ratio. Basal cells are small, and undifferentiated cells which are rarely seen in a Pap smear unless there is severe atrophy. Glandular cells consist of endocervical cells and endometrial cells. Viewed from above, sheets of endocervical cells have a honeycomb appearance, whereas when viewed from the side line up like "picket-fence" palisades. Endocervical glandular cells exhibit polarity with nuclei at one end of the cytoplasm and mucus present at the other. Endometrial cells which are spontaneously shed are derived from epithelial or stromal and often in a 3-dimensional cluster referred to as an "exodus" ball, which is generally present at the end of menstrual flow. Figure 4 exhibits various normal cervical cells.

Abnormal squamous cells or glandular cells can be discovered during cervical cytology screening which can be categorized as following types according to TBS reporting terminology:
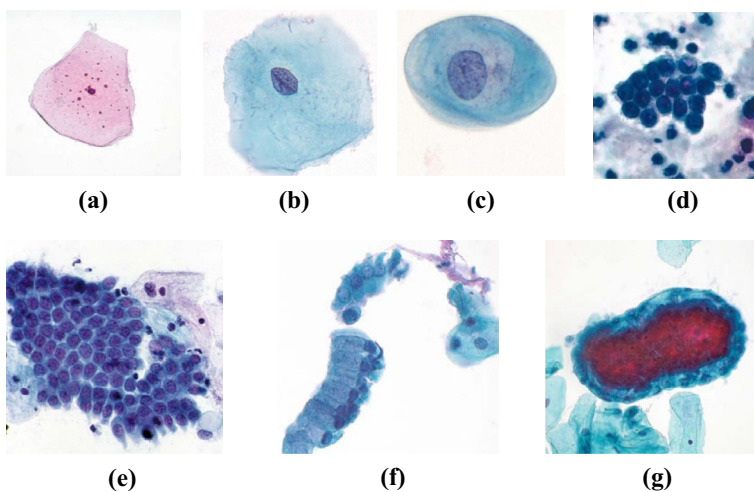

Fig. 4 Different normal cervical cells: **a** superficial cell, **b** intermediate cell, **c** parabasal cell, **d** basal cell, **e** endocervical cell (viewed from above), **f** endocervical cell (viewed from the side), **g** endometrial cell

- *Atypical squamous cells - undetermined significance (ASC-US)* This type refers to changes that are suggestive of the low-grade squamous intraepithelial lesion (LSIL). The nuclei of ASC-US are about 2.5 to 3 times the area of a normal intermediate squamous cell nucleus (approximately 35 mm2) and the N/C ratio is slightly increased.

- *Atypical squamous cells - cannot exclude a high-grade squamous intraepithelial lesion (ASC-H)* ASC-H primarily affects the squamous metaplastic cells and the nuclei are usually approximately 1.5–2.5 times larger than normal metaplastic cells' nuclei. The cytological changes of ASC-H are suggestive of the high-grade squamous intraepithelial lesion (HSIL) but are insufficient for a definitive diagnosis of HSIL.

- *Low-grade squamous intraepithelial lesion (LSIL)* To render an LSIL diagnosis, explicit abnormal changes must be found in the squamous cells. Cytological changes of LSIL usually occur in mature intermediate or superficial squamous cells and the nuclear enlargement is more than three times the area of normal intermediate nuclei. Additional characteristics of LSILs include hyperchromatic nuclei, absent or inconspicuous nucleoli, binucleation or multinucleation, and increased koilocytosis.

- *How-grade squamous intraepithelial lesion (HSIL)* In general, the cells affected by HSIL are immature parabasal or basal cells. HSIL cells can appear in sheets, singly, or in syncytial clusters which may result in hyperchromatic crowded groups (HCG). The nuclear enlargement and small size of HSIL cells lead to a marked increase in the N/C ratio. The nucleoli are generally absent and the contour of the nuclear membrane is quite irregular.

- *Squamous cell carcinoma (SCC)* SCC is defined as "an invasive epithelial tumor composed of squamous cells of varying degrees of differentiation" according to 2014 WHO terminology (Young 2014), which is the most common malignant tumor of cervical cancer. Cytological features of SCC usually include pleomorphic hyperchromatic nuclei, irregularly dispersed chromatin with nuclear clearing, prominent irregular often multiple nucleoli, keratinization of cells, and keratinous debris.

- *Atypical glandular cells (AGC)* AGC is a generic terminology for atypical endocervical cells or atypical endometrial cells when there is difficulty in locating the origin of the cells. Atypical endocervical cells may be further qualified as "NOS" or "favor neoplasia", while atypical endometrial cells don't need it. The cytological features of AGC may include nuclear enlargement, crowding, variation in size, hyperchromasia, chromatin heterogeneity, and evidence of proliferation.

- *Endocervical adenocarcinoma in situ (AIS)* AIS is considered to be the glandular counterpart of HSIL and the precursor to invasive endocervical adenocarcinoma. The criteria of AIS comprise of the following aspects: The cells present as sheets, pseudostratified strips or clusters - with loss of well-defined honeycomb patterns; The nuclei tend to be enlarged, variably sized, oval or elongated, and the loosely superficial cells of the cell groups incline to be tapered and spread out, referred to as "feathering"; Nucleoli are usually small or inconspicuous and may not be present; The quantity of cytoplasm is diminished and N/C ratio is increased; Nuclear hyperchromasia with evenly dispersed, coarsely granular chromatin; The chromatin pattern is coarsely granular with even distribution and mitoses are common.

- *Adenocarcinoma* The Cytological criteria for adenocarcinoma may overlap those outlined for AIS. There are abundant abnormal cells, typically with columnar configuration. Nuclei tend to be enlarged, pleomorphic with nuclear membrane irregularities, and may be hypochromatic with irregularly distributed chromatin or chromatin clearing. Multinucleation and Macronucleoli are common features. Adenocarcinoma may coexist with squamous lesions.

Figure 5 shows an illustration of various abnormal cervical cells. In a large-scale cervical cell screening program for the general population, the number of abnormal squamous cases is far more than abnormal glandular cases and ASC-US, LSIL, ASC-H, and HSIL are the four most common types. ASC-US and LSIL lesions usually occur in superficial cells or intermediate cells while ASC-H and HSIL lesions usually occur in parabasal cells and basal cells.
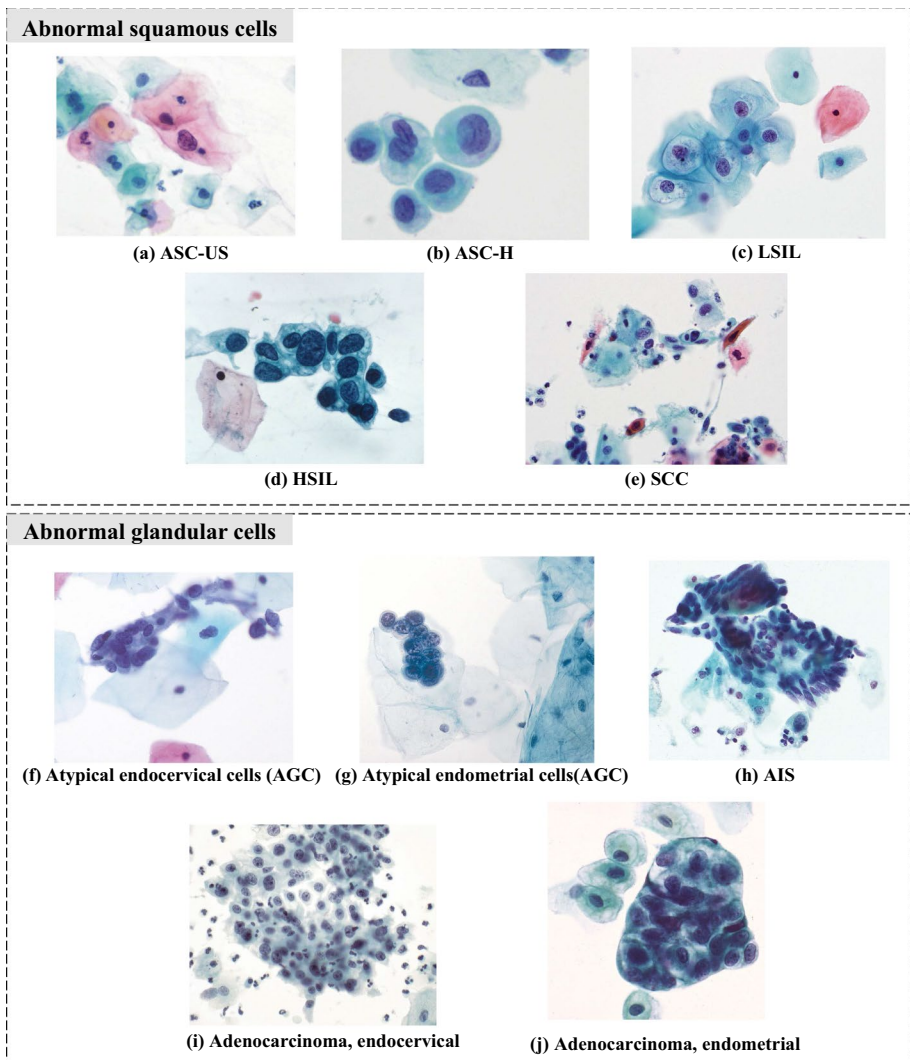

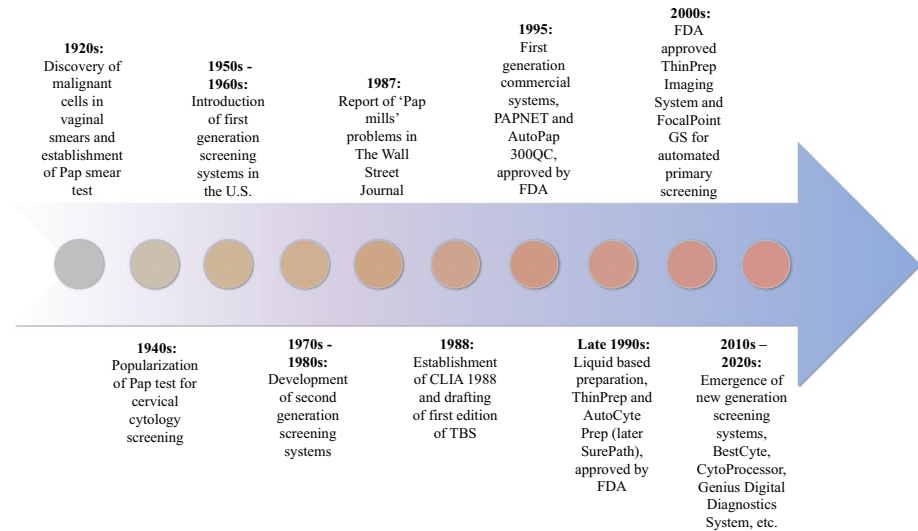**Fig. 5** Illustration of various abnormal cervical cells

Fig. 6 Timeline of major events in the technologic development of automated screening for cervical cytology

2.4 Automation in cervical cytology screening

This subsection reviews the development and evolutionary history of automated screening systems for cervical cytology. The timeline of some major events is shown in Fig. 6.

2.4.1 Origin of Pap smear

The earliest cervical cytology screening originated in the late 1920s. At that time, Papanicolaou first described malignant cells in vaginal smears and suggested the Pap smear test which is poorly received by Papanicolaou's contemporaries (Papanicolaou 1928). With the widespread adoption and practice of Pap smear tests in the 1940s, the incidence of cervical cancer was remarkably reduced (Papanicolaou and Traut 1943). Depending on the difficulty of the pap smear and the expertise of the cytopathologist, it takes about 5–10 mins on average to screen a sample (Traut and Papanicolaou 1943). Due to the lack of cytotechnologists and the increasing demand for Pap tests, manual screening of Pap smear is obviously a tedious and error-prone task. In this context, the development of automated screening systems is in full swing.

2.4.2 Early-stage screening Systems

The first attempt at automated Pap smear analyzers was the Cytoanalyzer (Tolles and Bostrom 1956), which utilized nuclear size and optical density to distinguish cancer cells from normal cells. Since then, the first-generation systems were all developed based on a primary concept that cancer cells could be differentiated from normal cells by morphometric features, such as the TICAS device (Wied et al. 1975) and CYBEST (Watanabe and Group 1974). These systems generated 2-dimensional histograms showing morphometric

differences between normal and abnormal cells through hard-wired analogue video processing circuits. However, such systems lacked interactive computers or display units that could show available digital images and often produced too many false results. Thus, a series of second-generation screening systems came into being by the 1980s, such as Bio-PEPR, FAZYTAN, Cerviscan, LEYTAS and Discanner (Bengtsson and Malm 2014). This generation of systems realized simple interaction for users and was used to explore new image segmentation, feature extraction, and classification methods for cervical cytology. But they still encountered several problems: first of all, computers were slow and unable to process the whole Pap smear at once which contained up to 300,000 cells. Besides, it was difficult to deal with three-dimensional (3D) clumps of cells and detect cell boundaries.

2.4.3 First-generation commercial systems

In the late 1980s, the "Pap mills" problem was reported that laboratories engaged in screening the greatest number of slides, often at the expense of quality. The public became aware of the reality of "false-negative" slides and the potential risk of Pap tests (Watanabe and Group 1974; Boronow 1998). Therefore, relevant laws and guidelines were established, such as the Clinical Laboratory Improvement Amendments of 1988 (CLIA 1988) and the first edition of TBS criteria (Nayar and Wilbur 2017). At the same peroid, computers had significantly advanced in terms of processing speed, memory capacity, and image display. These improvements had revitalized the hopes for automation in cervical cytology screening and many new projects were started.

AutoPap 300QC (Neopath, USA) and PAPNET (Neuromedical Systems, USA) were the first-generation commercially available screening systems that received the Food and Drug Administration (FDA) approval for rescreening of manually screened conventional cervical smears in 1995 (Lew et al. 2021). PAPNET (Koss et al. 1994) developed an artificial neural network to select up to 128 images of potentially abnormal cells for display and further diagnosis by cytopathologists. AutoPap 300QC (Patten et al. 1996) was a computerized image processor that developed algorithms to give a cumulative slide score to determine the abnormality of potential overall slide. Both PAPNET and AutoPap were considered capable of performing primary screening in their initial configurations, but the data to evaluate the practicality of these new technologies were not adequately prepared. Thus, they initially focused on quality control (QC) applications and successfully obtained approval for the rescreening of conventional Pap smear preparations to reduce false negative rates.

2.4.4 Currently available screening systems

Simultaneous to these cytology screening systems, sample preparation techniques were also constantly evolving and being commercialized. A new way of sample preparation for liquid-based cytology (LBC), ThinPrep, was developed by the Cytyc corporation (Hutchinson et al. 1991). Later, TriPath Imaging developed another similar preparation method, called AutoCyte Prep (ultimately SurePath) (Howell et al. 1998). The FDA approved ThinPrep and AutoCyte Prep for manual screening in 1996 and 1999, respectively. The advent of LBC enabled computers to locate and visualize individual cells easily by creating a uniform spread of cells, and further facilitated the development of automated cervical cytology screening tools. Afterward, the FDA successively approved two commercial products, the ThinPrep Imaging System (Biscotti et al. 2005) and the FocalPoint GS (Kardos 2004), that used these two products as their main specimen types for automated primary

screening. These approvals were granted in 2004 and 2008, respectively. Both systems were semi-automated slide scanning systems composed of a highly automated microscope and a processer that interpreted images of the FoV. The ThinPrep Imaging System detects and displays 22 FoVs containing the most suspicious cells on the slide, using a motorized microscope stage. Similarly, the FocalPoint GS operates in a similar way, but it also categorizes slides into quantiles based on the likelihood that they contain abnormalities.

2.4.5 Emerging cervical cytology screening systems

Recently, the new-generation analysis system for automated cervical cytology screening is in development, such as BestCyte (CellSolutions, USA) (Delga et al. 2014; Chantziantoniou 2022), CytoProcessor (DATEXIM, France) (Crowell et al. 2019), and Genius Digital Diagnostics System (Hologic, USA) (Ikenberg et al. 2023). The BestCyte consists of a digital scanner, networked storage and WSI analysis algorithm and enables remote access through web-based software. The CytoProcessor is a full web application that empowers users with virtual microscopy-like natural working environment. The CytoProcessor utilizes machine learning methods to select all suspicious abnormal cells to display in the gallery for cytopathologists' further review. The Genius Digital Diagnostics System is a digital cytology cloud platform that enables seamless and dynamic collaboration across laboratories within a network. Consisting of a digital imager, an image management server (IMS), and a review station, the system utilizes a new artificial intelligence (AI) algorithm and advanced volumetric imaging technology to detect (pre-)cancerous cells.

All these three screening systems enable web connection and exploit AI algorithms for primary diagnosis. The future emerging cervical cytology screening systems will be the combination of high-quality imaging devices, convenient viewing software/website, and powerful AI (ML/DL-based) analysis algorithms.

3 Research methodology

3.1 Research questions

In this systematic literature review, referred to the PRISMA (Preferred Reporting Items for Systematic reviews and Meta-Analyses) guidelines (Liberati et al. 2009), we searched for studies that applied or validated a DL-based method for automatic screening of cervical cytology. Overall, we assessed the tendencies and main problems present in this research field, highlighting the biomedical background knowledge of cervical cytology, the availability of the used datasets, DL-based approaches and result evaluation metrics. Besides, we aimed to identify the research gaps based on the existing findings and to discuss the feasibility analysis and future research directions. To obtain a more detailed and comprehensive view of the subject, the overall objective is motivated by the following research questions (RQs).

RQ 1   What are the DL-based techniques in automated cervical cytology screening?
RQ 2   What tasks are involved in automated cervical cytology screening?
RQ 3   How are DL-based approaches helping doctors in screening lesions?
RQ 4   Which data sources can be reached?
RQ 5   What is the best achieved performance in each study?

RQ 6   What are the challenges faced by the researchers while using DL models in cervi-
cal cytology?

3.2 Data source and search strategy

In this study, we searched PubMed, Scopus, IEEE Xplore, ACM Digital Library, and Web
of Science from 2016 to 2022. Because this review focuses on DL-based methods applied
to cervical cytology, the search was not limited to medical databases such as Medline or
PubMed, which were more inclined to biomedical topics and health informatics. Instead,
several databases in the field of computer sciences (CS) were also accessed. The articles
were searched using several keywords related to cervical cytology (e.g. 'Cervical cell' or
'Liquid-based cytology'), deep learning (e.g. 'Deep Learning' or 'Convolutional Neural
Networks') and screening tasks (e.g. 'Classification' or 'WSI Analysis'). The full search
strategy for each database can be found in Table S1 of the supplementary material.

3.3 Inclusion and exclusion criteria

The inclusion criteria for selecting an article were as follows: (a) It is written in English;
(b) It is published in 2016 or later; (c) It is published as an original journal article or con-
ference proceeding; and (d) It adopts, or proposes DL methods, systems, or applications to
realize related tasks in automated cervical cytology screening. Studies were excluded from
this review if: (a) the study is published as a book, book chapter, book review, executive
summary, scientific report, conference abstract, newspaper article, social media content
or workshop report, and research protocol. (b) the study is a duplicate that can be found
through several scholarly databases; (c) articles research on animals or non-human sam-
ples; and (d) articles do not focus on one of the research questions.

3.4 Study selection

The Covidence software (www.covidence.org) was used for screening and study selec-
tion. Eligibility assessment was conducted by two investigators, who screened titles and
abstracts of retrieved articles independently, and selected all relevant citations for full-text
review. Disagreements were resolved through discussion with a third reviewer.

   The search initially identified 1364 records from the stated databases, of which 799 were
screened after removing 565 duplicated articles. Then, according to the predetermined
inclusion criteria, 661 were excluded and a total of 138 articles were included through the
second round of the selection process. After assessing the full-text articles, 89 papers in
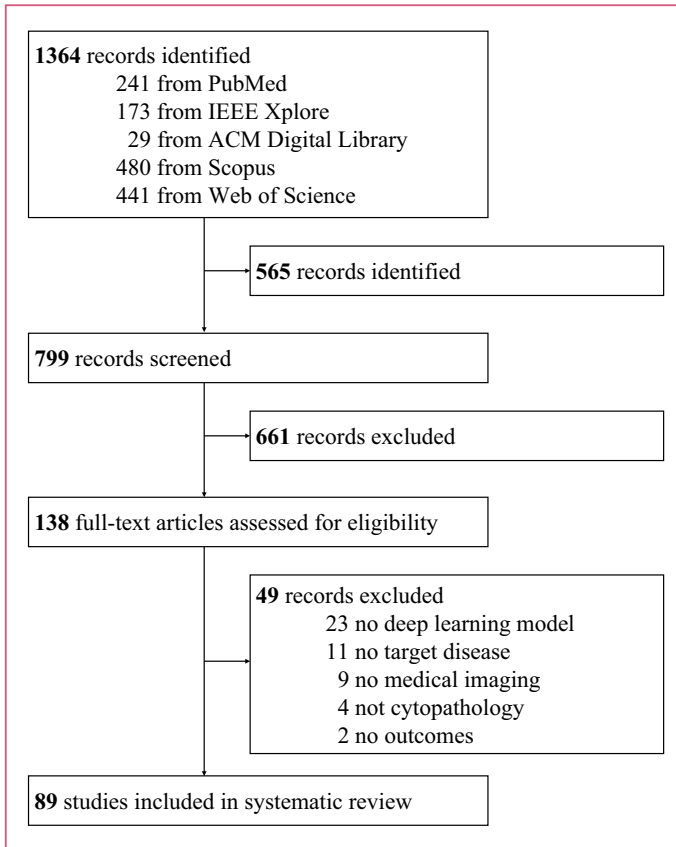total were included in this systematic review (see Fig. 7).

Fig. 7 PRISMA flowchart of the study selection process

3.5 Data extraction

Data extraction was conducted to explore different DL-based methods proposed or applied in cervical cytology screening programs. To ensure the reliability and quality of this review article, two reviewers with expertise in deep learning and biomedical engineering extracted study characteristics and diagnostic performance data. For each selected article, this study extracted the following data: authors, publication time, research objectives, study context, specific task, methodology used, dataset used, and study outcomes/findings. These items were extracted to enable researchers to find and compare current DL-based studies in their research fields or tasks. The extracted data were also synthesized and analyzed to summarize the existing research and identify the potential scopes for future research.

4 Deep learning in cervical cytology

In this section, we first give an overall introduction and structurized analysis of the automation of cervical cytology screening (Sect. 4.1). Next, we survey publicly available cervical cytology datasets and describe them in detail (Sect. 4.2). Then we comprehensively summarize the literature on various deep learning methods applied in cervical cytology, including several representative clinical tasks: cell-level identification (Sect. 4.3), detection (Sect. 4.4), segmentation (Sect. 4.5), and slide-level diagnosis (Sect. 4.6).

4.1 Overall introduction and problem analysis

The ultimate goal of automated cervical cytology screening is to improve the overall effectiveness of cervical cancer screening programs, reduce the incidence and mortality of cervical cancer, and ultimately improve women's health outcomes. Additionally, automated systems can potentially reduce healthcare costs associated with cervical cancer screening and treatment, making it more accessible to women in resource-limited settings. From the evolutionary history of automated screening systems in the above Sect. 2.4, we know that the developing systems are all going to use a combination of digital scanners and AI technology to assist doctors in making cytological diagnoses. Future automated screening systems will fully enable autonomous diagnosis and reporting of results via the use of such techniques as medical imaging, computer vision, and machine learning. This analysis procedure of the whole specimen involves searching the region of interest (RoI), segmenting cells, and classifying precancerous or cancerous cells. The specific process is as follows (Fig. 8):

- *Image acquisition* The first step in automating cervical cytology screening is to acquire high-quality images of the cervical cells. This can be done using various imaging techniques such as optical microscopy, digital imaging, or automated slide scanning.
- *Image preprocessing* The acquired images are then preprocessed to enhance image quality and remove any noise or artifacts that may interfere with the analysis. This step involves image filtering, noise reduction, stain normalization, and contrast enhancement.
- *RoI detection* The next step is to detect and classify different regions of interest (suspicious cells or areas) in the images for further analysis. Object detection algorithms such as Faster R-CNN, or YOLO can be employed in this step.
- *Cell segmentation* Once RoIs are detected, they can be segmented into individual cells and different parts of cells can be identified. This step is not necessary for recent diagnosis methods of whole slide images since DL-based models can directly extract features from RoIs without precise segmentation results. However this step is critical to obtaining fine-grained characteristics of (pre-)cancerous cells for quantitative cytology. Quantitative computation of the segmentation results can provide morphological features that have clear medical diagnostic significance. This can further improve the classification accuracy and ensure the reliability of the classification results.
- *Feature extraction* After the individual cells are segmented, features such as size, shape, texture, and color can be extracted from each cell. For example, the brightness, elongation, roundness, perimeter, and area (of nucleus and cytoplasm), N/C
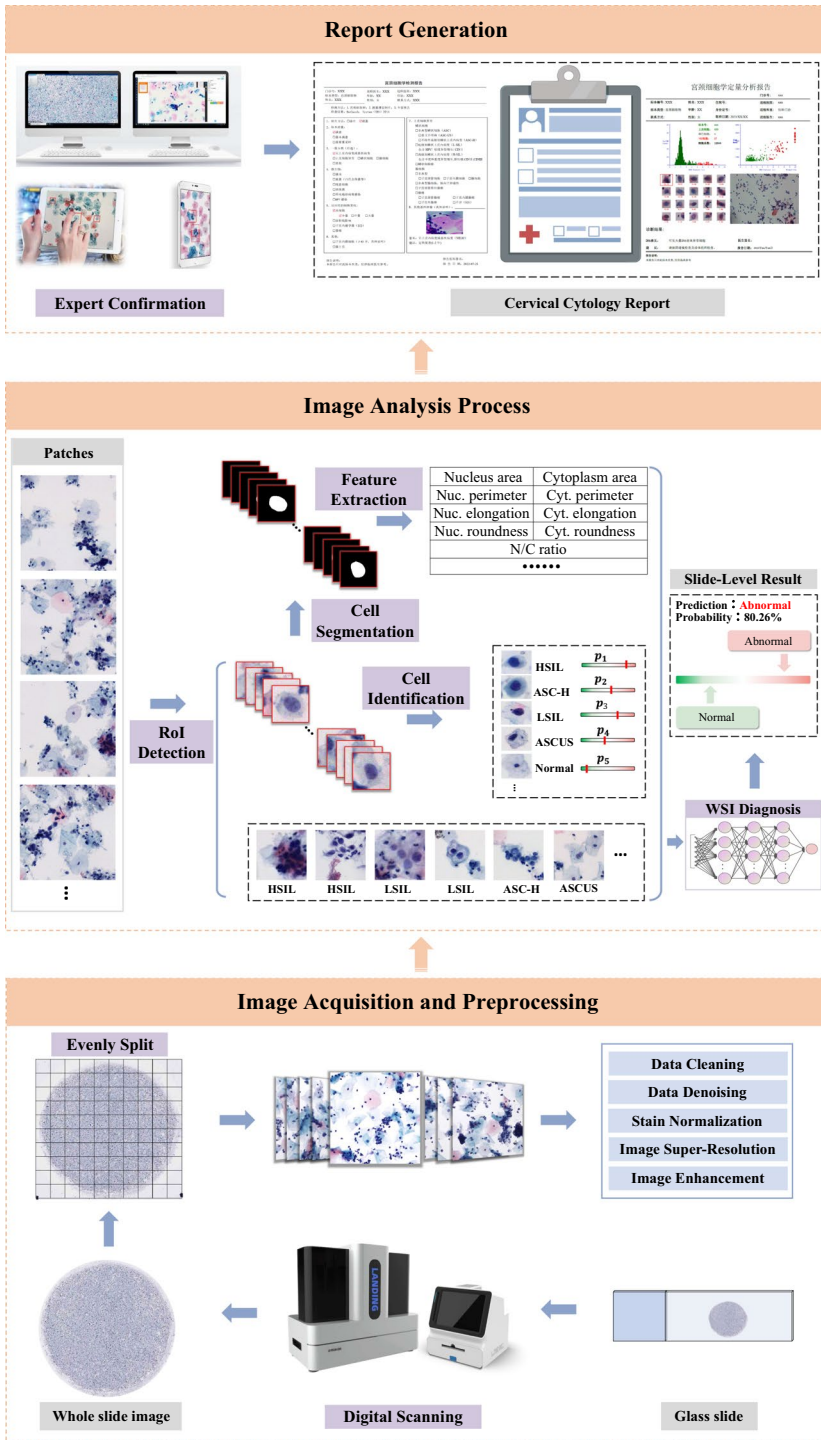
Fig. 8 The overall process of automated cervical cytology screening

ratio, and nucleus relative position, which are discriminative features of biomedical significance. These features are then used for cell identification.

- *Cell identification* Subsequently, the detected RoIs or the extracted features are exploited to classify the cells into different categories according to the TBS criteria, such as normal, LSIL, or HSIL. This can be done using various machine learning algorithms, such as decision trees, support vector machines, and deep learning models.
- *WSI diagnosis* Then, all cell-level or patch-level classification results will be integrated to perform the slide-level prediction. There are two prevailing ways to integrate the results: the first one is to fuse the cell-level or patch-level features to generate a slide-level feature for WSI diagnosis and the other one is to directly combine all prediction probabilities and output the final slide-level classification probability.
- *Result reporting* The final step is to generate a report that summarizes the results of the whole WSI analysis process. The report can cover information such as the location and characteristics of the identified cells or areas, the slide-level diagnosis result, the level of suspicion, etc. Besides, a recommended prognosis scheme may also be included.

The goal of image acquisition and preprocessing is to provide high-quality data support for subsequent image analysis steps. Deep learning can empower these two steps. Firstly, pathological staining is a time-consuming and labor-intensive process that requires specialized laboratory infrastructure, chemical reagents, and trained technicians. Developing virtual staining models based on deep learning can omit conventional staining steps and free up medical workers (Bai et al. 2023). Additionally, physical principles guided deep learning methods can help improve the quality of optical imaging (Li et al. 2022). Besides, because of differences in staining procedures, staining materials, imaging settings, and scanning devices, there are often variations in the style of cytological images collected. The utilization of generative adversarial networks (GANs) (Creswell et al. 2018), such as CycleGAN, for stain normalization can unify image styles. For example, Kang et al. realized stain normalization on cervical cytology images via StainNet (Kang et al. 2021). Furthermore, deep generative models based on GANs and diffusion models (Croitoru et al. 2023) can be used for out-of-focus and low-resolution images to enhance the resolution of acquired images, allowing for more detailed visualization of small structures.

During the process of automatic image analysis, achieving better feature representation has always been a pursuit. Initially, simple neural networks with only a few layers were used for recognizing cervical cytology images. Later, deeper and wider networks have become a consensus for learning better representation and achieving better performance (Simonyan and Zisserman 2015; He et al. 2016; Szegedy et al. 2016; Xie et al. 2017). Various deep CNN models are used for cervical cell identification (Rahaman et al. 2020). However, too deep a network is prone to gradient disappearance and requires more data for fitting. In recent years, the attention mechanism has been proposed as a way to mimic the operation of the human visual system (Niu et al. 2021; Guo et al. 2022). The human visual system has varying levels of perception in different parts, with the highest sensitivity and strongest information processing ability located at the center of the retina. During the process of receiving external information, people first quickly scan the global information and then focus their gaze on a specific field of view for local information. The brain analyzes and processes this local information more carefully, while other information is filtered and ignored. The introduction of the attention mechanism to deep learning models allows models to selectively focus on the most important features of the input image while ignoring irrelevant or noisy information. Most recently, self-attention and vision transformer (ViT) architecture further expands visual attention capabilities and has been used for a wide

range of computer vision tasks (Dosovitskiy et al. 2021; Liu et al. 2021; Touvron et al. 2021; Khan et al. 2022). In the field of automated cervical cytology, CVM-Cervix firstly employed ViT to identify abnormal cervical cells, which demonstrated its strong ability of feature extraction (Liu et al. 2022).

To further enhance the performance for practical production and broad application, introducing more additional information into deep learning models beyond existing cervical cytology datasets is a promising approach. Pretraining deep learning models on natural image datasets such as ImageNet (Russakovsky et al. 2015), and then fine-tuning on the target cervical cytology dataset, implicitly introduces the information learned from natural images. Other medical datasets with similar tasks to cervical cytology screening can also be used to introduce more relevant medical features. In addition to the above transfer learning approaches, incorporating medical domain knowledge is also helpful (Xie et al. 2021). Experienced cytopathologists can provide relatively accurate diagnoses, so their knowledge can better assist deep learning models with tasks related to cervical cytology screening. Medical knowledge related to cervical cytology screening includes the structure and function of cervical cells, the process of cell division and differentiation, how pathologists view cytology images, the specific areas they typically focus on, and the features they are particularly concerned with. This knowledge has been accumulated, summarized, and validated by numerous cytopathologists over many years based on a large number of cases. By incorporating this specialized knowledge into the model structuring or training process, DL-based models can better understand the underlying cytopathology of the images they are analyzing, resulting in more accurate and reliable diagnoses. Handcrafted features related to cell morphology give guidance to cervical cell identification (Dong et al. 2020) and shape priors are invaluable in assisting deep learning models in overlapping cell segmentation (Xu et al. 2018).

Compared to natural image datasets, there is currently a lack of large-scale high-quality image datasets for cervical cytology. The lack of cervical cytology datasets manifests in three aspects. Firstly, the number of images in the dataset is usually limited due to the high cost of data collection. The acquisition of cervical cytology images requires processes such as slide preparation, staining, and scanning stitching, which are expensive in terms of necessary equipment and labor. Secondly, only a small portion of the samples are annotated. These annotations include overall classification labels for the samples, as well as the location and category of abnormal cells. The annotation process requires a significant amount of experience from professional cytologists. Thirdly, cervical cytology screening is a preliminary screening for cervical cancer aimed at serving a broad population of women. Therefore, it is difficult to collect enough positive cases for rare and severe conditions to achieve a balanced dataset. To alleviate the above issues, annotation-efficient learning has been proposed to make full use of limited annotations and excavate potential discriminative information in unlabeled samples, including semi-supervised learning (Van Engelen and Hoos 2020), multiple instance learning (MIL) (Carbonneau et al. 2018), etc.

In addition to the above issues, the speed and convenience of automated screening systems also require consideration. Therefore, some researchers have studied lightweight modules (Wang et al. 2019; Zhao et al. 2022) and IoMT architectures (Jiang et al. 2022) to improve the efficiency in processing cervical cytology images and system operation, thereby reducing the overall costs of the system. Besides, DL-based models lack clinical interpretability and transparency, which limits their practicability and generality. Thus, the development of explainable methods in cervical cytology can foster trust between AI technologies and cytopathologists (Li et al. 2022). The use of visualization techniques is often considered a primary method for interpreting DL-based models. In the analysis of cervical

cytology images, a number of studies have utilized class activation mapping (CAM) (Zhou et al. 2016) based methods (Grad-CAM Selvaraju et al. 2017, Score-CAM Wang et al. 2020, Relevance-CAM Lee et al. 2021, etc.), to generate heatmaps or attention scores for further investigation of the decision-making process. More visualization techniques are under exploration for improving interpretability.

4.2 Public datasets of cervical cytology

At the beginning of developing automatic methods for cervical cytology screening, many human and material resources were devoted to the collection of cervical cytological images because automatic analysis methods rely on large amounts of labeled data and there are few public datasets available. We summarize publicly available datasets for cervical cytology screening, as listed in Table 2. These public cervical cytology datasets can be utilized to develop automatic analysis algorithms for multiple tasks, including image classification, object detection, semantic segmentation, etc.

Herlev (Jantzen et al. 2005). Herlev is the most widely used dataset for the analysis of cervical cytology, which consists of 917 Papanicolaou (Pap) smear cervical images in 7 classes (3 normal classes and 4 abnormal classes) based on the classification rule of the 3-tiers dysplasia system. All images are collected by using a microscope connected to a frame grabber with a resolution of 0.201 µm/pixel at Herlev University Hospital(Denmark). Each cell image is segmented manually into the background, cytoplasm, and nucleus for further feature extraction.

ISBI 2014 (Lu et al. 2015). This dataset is released for the first Overlapping Cervical Cytology Image Segmentation Challenge under the auspices of the IEEE International Symposium on Biomedical Imaging (ISBI 2014). The main target of this challenge is to extract the boundaries of individual cytoplasm and nucleus from overlapping cervical cytology images. The dataset consists of 16 Extended Depth Field (EDF) cervical cytology images and 945 synthetic images. Each image consists of 20 to 60 Papanicolaou-stained cervical cells with different degrees of overlap. This dataset is built by the University of Adelaide in Australia and the images are captured by an Olympus BX40 microscope with a 40× objective and a four mega-pixel SPOT Insight camera. The resolution of the image is about 0.185 µm/pixel.

ISBI 2015 (Lu et al. 2016). This dataset is used for the second cervical cell segmentation challenge in ISBI 2015, consisting of a collection of 17 multi-layer cervical cell volumes, from which 8 will be used for training and 9 for testing. The main difference between ISBI 2015 and ISBI 2014 is that the input data will consist of a multi-layer cytology volume, which means that the input data is now a volume consisting of a set of multi-focal images acquired from the same specimen. This richer input dataset may provide more information on the task of detecting and segmenting cervical cells, thus enabling more accurate cytoplasmic and nuclear detection and segmentation of cervical cells.

SIPaKMeD (Plissiti et al. 2018). This database consists of 4049 images of isolated cervical cells which are acquired from the University of Ioannina, Greece, through a CCD camera adapted to an optical microscope (OLYMPUS BX53F). The cells are annotated by experienced cytopathologists into five different classes (superficial-intermediate, parabasal, koilocytotic, dyskeratotic, and meta-plastic cells), depending on their cytological appearance and morphology. Among these five classes, superficial-intermediate and parabasal are normal cells. Koilocytes and dyskeratotic are abnormal but not malignant cells while

Table 2 Summary of publicly available datasets for cervical cytology screening

Dataset	year	Preparation	Terminology	Task	Description	Link
Herlev (Jantzen et al. 2005)	2005	CPS	three-tiers dysplasia	Classification; Segmentation	917 cells in 7 classes	https://mde-lab.aegean.gr/index.php/downloads/
ISBI 2014 (Lu et al. 2015)	2014	LBC	–	Segmentation	16 EDF cell images and 945 sythetic images	https://cs.adelaide.edu.au/~carneiro/isbi14_challenge/index.html
ISBI 2015 (Lu et al. 2016)	2015	LBC	–	Segmentation	17 multi-layer cervical cell volumes	https://cs.adelaide.edu.au/~carneiro/isbi15_challenge/index.html
SIPaKMeD (Plissiti et al. 2018)	2018	CPS	cytological morphology	Classification	4,049 cervical cells in 5 catogories	https://www.cs.uoi.gr/~marina/sipakmed.html
CERVIX93 (Phoulady and Mouton 2018)	2018	LBC	TBS	Classification; Detection	93 stacks of images with 2,705 annotated nuclei	https://github.com/parham-ap/cytology_dataset
BHS (Araújo et al. 2019)	2019	CPS	–	Segmentation	194 images	https://sites.google.com/view/centercric
BTTFA (Zhang et al. 2019)	2019	LBC	–	Segmentation	104 cervical LBC images	https://data.mendeley.com/datasets/jks43dkjj7/1
Mendeley LBC (Hussain et al. 2020)	2020	LBC	TBS	Classification	963 LBC images in 4 classes	https://data.mendeley.com/datasets/zddtpgzv63/4
Cric (Rezende et al. 2021)	2021	CPS	TBS	Classification	400 images with 11,534 cells	https://database.cric.com.br
Comparison Detector (Liang et al. 2021)	2021	LBC	TBS	Detection	7,410 images with 48,587 objects in 11 categories	https://github.com/kuku-sichuan/ComparisonDetector
RepoMedUNM (Riana et al. 2021)	2021	LBC + CPS	TBS	Classification	6,168 Pap smear cell images	http://repomed.nusamandiri.ac.id/
CCEDD (Liu et al. 2022)	2022	LBC	–	Edge Detection; Segmentation	686 raw cervical images and 33,614 cut images	https://github.com/nachifur/LLPC
Cx22 (Liu et al. 2022)	2022	LBC	–	Segmentation	1,320 images of 14,946 cellular instances	https://github.com/LGQ330/Cx22

metaplastic belongs to benign. In each image of the SIPaKMeD database, the areas of the cytoplasm and the nucleus are manually defined.

CERVIX93 (Phoulady and Mouton 2018). This dataset consists of 93 stacks (frames) of images provided by Moffitt Cancer Center (Tampa, FL). All images are acquired by an integrated microscope system (Stereologer, SRC Biosciences, Tampa, FL) at 40× magnification. Each of the stacks has 10-20 images and all images are size 1280 × 960 pixels. Based on TBS, all frames are examined by cytologists and graded with three categories (Negative, LSIL, HSIL). A total of 2705 nuclei are manually annotated with bounding boxes according to all grade categories.

BHS (Araújo et al. 2019). This database collects 194 conventional pap smears from the Brazilian Health System (BHS). The collected glass slides are digitized by a Zeiss Axio-Cam MRc camera with a magnification of 40× to construct the training dataset (26 images) and test dataset (168 images). Each image has a resolution of 0.255 mm/pixel with a size of 1392 × 1040. The images are labeled into two classes (normal/abnormal) and abnormal images contain 5 different types of (pre-)cancerous cells (Carcinoma, HSIL, LSIL, ASCUS, and ASCH).

BTTFA (Zhang et al. 2019). South China University of Technology releases this real-world clinical dataset with well-annotated nuclei. This dataset contains 104 cervical LBC images with a size of 1024 × 768. All images are scanned via the Olympus microscope B x 51 with a magnification of 200× and the resolution of the image is 0.32 μm/pixel. All collected images are manually segmented by a professional pathologist to get the pixel-level segmentation label.

Mendeley LBC (Hussain et al. 2020). This dataset collects a total of 460 specimens from three medical diagnostic centers in India, including Babina Diagnostic Pvt. Ltd in Imphal, Gauhati Medical College and Hospital in Guwahati and Dr. B. Barooah Cancer Institute in Guwahati. The dataset consists of a total of 963 liquid-based cytology (LBC) images captured by a Leica ICC50 HD microscope collected in 400× (40× objective lens and 10× eyepiece) magnification. The size of each image is 2048 × 1536. Images have been subdivided into four categories: NIML (613), LSIL (163), HSIL (113), and SCC (74).

CRIC (Rezende et al. 2021). The collection of the CRIC dataset has 400 images of conventional cervical pap smears and 11,534 classified cells. The Pap smears are collected from 118 female patients in the Southeast region of Brazil and prepared and analyzed in the Cytology Laboratory of the Pharmacy School, Federal University of Ouro Preto, Minas Gerais, Brazil. All images are captured by conventional bright-field microscopy with a 40× objective and a 10× eyepiece, using a Zeiss AxionCam MRc digital camera coupled to the Zeiss AxioImager. The image size is 1376 × 1020 and the resolution is 0.228 μm/pixel. CRIC collection covers six types based on TBS nomenclature: NILM (6779), ASC-US (606), LSIL (1360), ASC-H(925), HSIL (1703), and SCC (161).

Comparison Detector (Liang et al. 2021). This database is collected by Central South University in China with samples scanned by Pannoramic MIDI II digital slide scanner. It consists of 7410 cervical images cropped from the WSIs. There is a total of 48,587 object instance bounding boxes labeled by experienced cytopathologists. According to TBS categories, the annotated objects belong to 11 categories: ASC-US, ASC-H, LSIL, HSIL, SCC, AGC, trichomonas (TRICH), candida (CAND), flora, herps and actinomyces (ACTIN).

RepoMedUNM (Riana et al. 2021). This database is released by Universitas Nusa Mandiri in Indonesia which is comprised of 6168 Pap smear cell images collected from 24 slides including both non-ThinPrep Pap test images and ThinPrep Pap test images. Images are obtained by an OLYMPUS CX33RTFS2 optical microscope and an X52-107BN microscope with a Logitech camera. For non-ThinPrep images, there are 3083 images in

total containing two categories, normal and LSIL. ThinPrep images are divided into three categories: normal cells (1513), koilocyt cells (434), and HSIL (410).

CCEDD (Liu et al. 2022). This dataset collects 686 cervical images with a size of 2048 × 1536 from Liaoning Cancer Hospital & Institute. All samples are scanned with a Nikon ELIPSE Ci slide scanner, SmartV350D lens and a 3-megapixel digital camera. The magnification is 100× for negative patients and 400× for positive patients. The captured images contain overlapping cervical cell masses in various complex backgrounds and are labeled by 6 experienced cytologists to outline the contours of the cytoplasm and nucleus. The original images are divided into training, validation, and test sets using a ratio of 6:1:3. All raw images are cut into 512 × 384 pixels and 33,614 cut images are obtained.

Cx22 (Liu et al. 2022). This dataset is an extension of the CCEDD dataset released by Key Laboratory of Opto-Electronic Information Processing, Chinese Academy of Sciences, Shenyang in which more precise instances (cytoplasm and nucleus) are annotated. A total of 14,946 cell instances in 1320 images with the size of 512 × 512 are collected and divided into two sub-sets, Cx22-Multi (containing multiple instances) and Cx22-Pair (only containing a pair of instances).

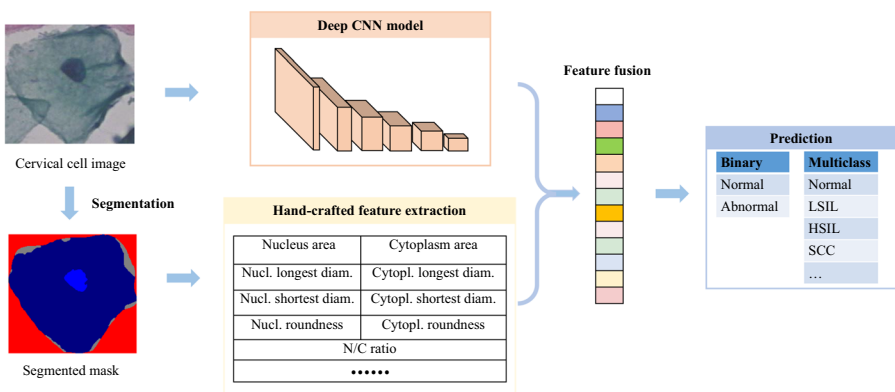## 4.3 Cervical cell identification

Cell-level identification is one of the most successful tasks applied by deep learning in cervical cytology screening. Traditional machine learning methods need to accurately segment the cell outline and even the nucleus, and then manually design the features (nucleus area, cytoplasm area, nucleus perimeter, cytoplasm perimeter, N/C ratio, etc.). The extracted hand-crafted features are fused and utilized for final classification, to realize the identification of cervical cells. Most of the traditional machine learning-based methods rely on the accuracy of cell segmentation, which is the key to feature extraction. However, in actual clinical practice, the complex background and fuzzy overlapping cells bring serious difficulties to the accurate segmentation of cervical cells. Conversely, a DL-based identification scheme in the form of a convolutional neural network (CNN) avoids complex image preprocessing steps such as pixel-level cell segmentation, feature selection, and extraction. Owing to the learning of abundant training data, DL-based approaches have gradually become a promising research direction use can realize end-to-end and high-performance identification of cervical cells. The most straightforward approach is to feed the cell image directly into a deep CNN model to extract the feature maps, then use the output layer and a classifier to obtain the predicted category. Bora et al. (2016) utilized AlexNet (Krizhevsky et al. 2017) and an unsupervised feature selection technique (Mitra et al. 2002) together with two classifiers, namely Least Square Support Vector Machine (LSSVM) and Softmax Regression, to classification of cervical dysplasia in Pap smear images. Shanthi et al. (2019) designed a CNN architecture composed of three convolutional layers, three max-pooling layers and one fully connected layer. They evaluated the proposed network on four different datasets using different settings (2 class, 3 class, 4 class, and 5 class), showing its ability for cervical cell identification. Chen et al. (2021) proposed a novel network CompactVGG, which is adapted from VGGNet to realize the high-performance classification of cervical cells. On public datasets Herlev and SIPaKMeD, and their collected private dataset, CompactVGG achieved the best performance compared to some classical CNN models. Similarly, DCAVN is proposed to identify cervical cells as normal or abnormal by using deep convolutional and variational autoencoder network (Khamparia et al. 2021).
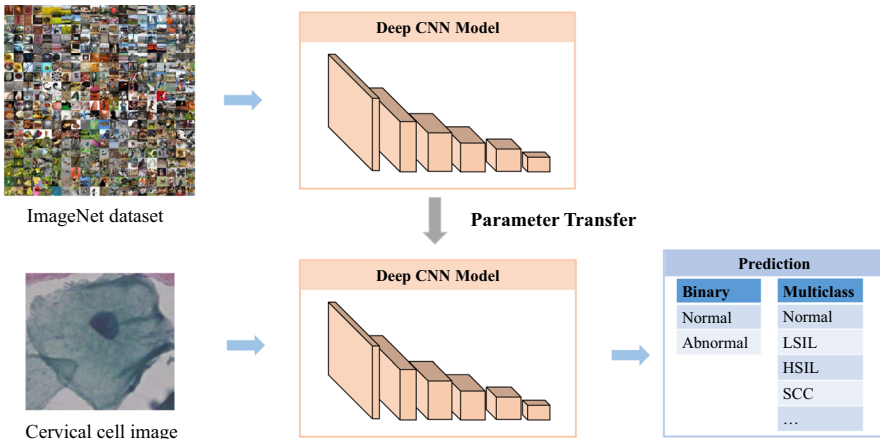
(a) Transfer learning based identification
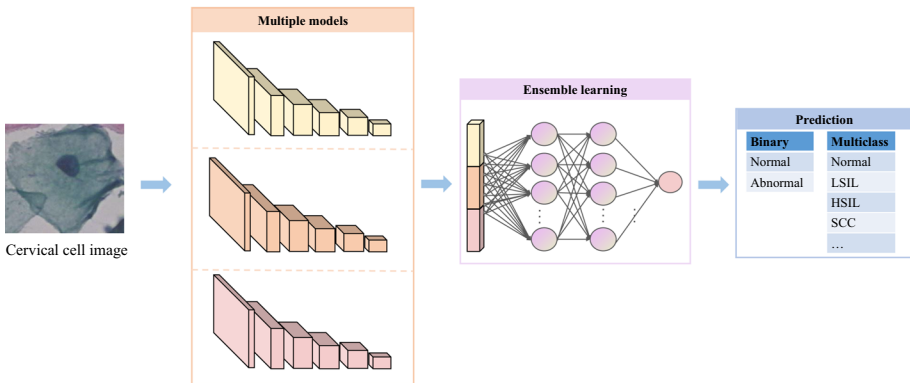
(b) Multi-model ensemble based identification

(c) Multi-type feature fusion based identification

Fig. 9 Three prevalent deep learning based approaches for cervical cell identification

Table 3 Summary of deep learning-based studies for cervical cell classification. Accuracy (Acc), Precision (Pre), Recall (Rec), Specificity (Spec), Sensitivity (Sens), F1-score (F1)

Reference	Method	Dataset	Classes	Result
Classical or Self-designed Model				
Bora et al. (2016)	AlexNet + Unsupervised feature selection + Classifiers (LSSVM and Softmax)	Herlev and generated database (1001 NILM cells, 400 LSIL cells, and 210 HSIL cells)	2-class, 3-class	Herlev (2-class): Acc = 94.61%(LSSVM) Herlev (3-class): Acc = 89.97%(LSSVM), Acc = 88.88%(Softmax) Generated (2-class): Acc = 98.88%(LSSVM) Generated (3-class): Acc = 92.24%(LSSVM), Acc = 93.1%(Softmax)
Shanthi et al. (2019)	Simple CNN (3 Conv + 1 pooling + 1 FC)	Herlev	2-class, 3-class, 4-class, 5-class	(2-class): Acc = 96.11%, (3-class): Acc = 94.80%, (4-class): Acc = 94.62%, (5-class): Acc = 95.31%
Chen et al. (2021)	CompactVGG (Adapted from VGGNet)	Herlev, SIPaKMed and private dataset (60,238 positive cells, 25,001 negative cells, and 113,713 junk images)	2-class, 3-class	Herlev (2-class): Acc = 94.81%, Sens = 95.52%, Spec = 92.76%, F1 = 96.46%; SIPaKMed (3-class): Acc = 98.94%, Sens = 97.80%, Spec = 99.17%, F1 = 98.28%; Private (3-class): Acc = 88.30%, Sens = 92.83%, Spec = 91.03%, F1 = 87.04%
Khamparia et al. (2021)	Deep convolutional and variational autoencoder network (DCAVN)	Herlev	2-class	Acc = 99.4%, Pre = 99.4%, Rec = 99.1%, F1 = 99%
Transfer Learning				
Zhang et al. (2017)	Simple ConvNet + Transfer learning (Fine-tune)	Herlev and private dataset HEMLBC (989 abnormal cells and 989 normal cells)	2-class	Herlev: Sens=98.2%, Spec=98.3%, Acc=98.3%, F1=98.8%, AUC=99.8%; HEMLBC: Sens = 98.3%, Spec = 99.0%, Acc = 98.6%
Hyeon et al. (2017)	VGG-16 + Transfer learning (Feature extraction)	Private dataset (8,373 abnormal cells and 8,373 normal cells)	2-class	Pre = 78.17%, Rec = 78.17%, F1 = 78.17%

Table 3 (continued)

Reference	Method	Dataset	Classes	Result
Nguyen et al. (2018)	Pretrained CNN (Inception-v3, ResNet152, and Inception-ResNet-v2) + Transfer learning + Feature concatenation	Herlev dataset	2-class	Acc = 92.63%
Ghoneim et al. (2020)	CNN (shallow CNN, VGG-16 and CaffeNet) + ELM-based classifier + Transfer learning (Fine-tune)	Herlev	2-class, 7-class	(2-class): Acc = 99.5%, (7-class): Acc = 91.2%
Khamparia et al. (2020)	CNN (InceptionV3, VGG-19, SqueezeNet, ResNet-50) + Transfer learning (Feature extraction)	Herlev	2-class	InceptionV3: Acc = 94.88%; VGG-19: Acc = 96.76%; SqueezeNet: Acc = 95.17%; ResNet-50: Acc = 97.89%
Wang et al. (2020)	Adaptive pruning deep transfer learning model (PsiNet-TAP)	Private dataset (120 normal cell images, 206 uninvolved cell images and 63 abnormal cell images)	2-class	Normal vs abnormal: Acc = 98.41%, Sens = 97.83%, Spec = 98.75%; Normal vs uninvolved: Acc = 98.18%, Sens = 98.57%, Spec = 97.50%; Normal vs abnormal + uninvolved: Acc = 98.49%, Sens = 99.64%, Spec = 95.83%
Bhatt et al. (2021)	Progressive resizing technique + CNN (EfficientNet, VGGNet, ResNet) + Transfer learning (Fine-tune)	Herlev and SIPaKMed	2-class, 5-class, 7-class	SIPaKMed (2-class): Acc = 99.01%, Pre = 99.15%, Rec = 98.89%; SIPaKMed (5-class): Acc = 99.70%, Pre = 99.70%, Rec = 99.72%; Herlev (7-class): Acc = 93.14%, Pre = 94.56%, Rec = 93.98%
Multi-Model Ensemble Rahaman et al. (2021)	CNN (VGG-16, VGG-19, XceptionNet, and ResNet-50) + Hybrid deep feature fusion	Herlev and SIPaKMed	2-class, 3-class, 5-class, 7-class	SIPaKMed: (2-class) Acc = 99.85%, (3-class) Acc = 99.38%, (5-class) Acc = 99.14%; Herlev: (2-class) Acc = 98.32%, (7-class) Acc = 90.32%

Table 3 (continued)

Reference	Method	Dataset	Classes	Result
Manna et al. (2021)	CNN (Inception v3, XceptionNet and DenseNet-169) + Fuzzy rank-based ensemble	Mendeley LBC and SIPaKMed	2-class, 4-class, 5-class	SIPaKMed (2-class): Acc = 98.55%, Pre = 98.57%, Rec = 98.52%, F1 = 98.54%; SIPaKMed (5-class): Acc = 95.43%, Pre = 95.34%, Rec = 95.38%, F1 = 95.36%; Mendeley LBC (4-class): Acc = 99.23%, Pre = 99.13%, Rec = 99.23%, F1 = 99.18%
Diniz et al. (2021)	CNN (EfficientNet, MobileNet, XceptionNet and Inception v3) + Vote from three best-trained models	CRIC	2-class, 3-class, 6-class	(2-class): Acc = 96%, Pre = 96%, Rec = 96%; (3-class): Acc = 96%, Pre = 94%, Rec = 94%; (6-class): Acc = 95%, Pre = 85%, Rec = 85%
Madhukar et al. (2022)	CNN (VGG16 and ResNet50) + Feature Fusion	CRIC	2-class, 3-class, 6-class	(2-class): Acc = 96.07%, Pre = 94%, Rec = 92%; (3-class): Acc = 93.30%, Pre = 91%, Rec = 92%; (6-class): Acc = 85.07%, Pre = 81%, Rec = 82%
Liu et al. (2022)	CNN (Xception) + Vision transformer (tiny DeiT) + MLP module	Herlev, CRIC and SIPaKMeD combined dataset	2-class, 11-class	Herlev (2-class): Acc = 92.35%, Pre = 88.90%, Rec = 93.50%, F1 = 90.70%; Combined dataset (11-class): Acc = 91.72%, Pre = 91.80%, Rec = 91.60%, F1 = 91.70%
Kundu and Chattopadhyay (2022)	CNN (GoogLeNet and ResNet-18) + genetic algorithm for feature selection	Mendeley LBC and SIPaKMed	2-class, 4-class, 5-class	SIPaKMed (2-class): Acc = 99.65%, Pre = 99.60%, Rec = 99.58%, F1 = 99.59%; SIPaKMed (5-class): Acc = 98.94%, Pre = 98.79%, Rec = 98.80%, F1 = 98.79%; Mendeley LBC (4-class): Acc = 99.07%, Pre = 98.39%, Rec = 98.18%, F1 = 98.31%

Table 3 (continued)

Reference	Method	Dataset	Classes	Result
Hybrid Feature Fusion				
Jia et al. (2020)	Strong features (extracted by GLCM and Gabor) + Abstract features (Extracted from LeNet-5) + Multi feature serial fusion	Herlev and private dataset (1000 positive cell images and 1000 negative cell images)	2-class, 7-class	Herlev (2-class): Acc = 99.3%, Sens = 98.9%, Spec = 99.4%; Herlev (7-class): Acc = 93.8%, Sens = 93.7%, Spec = 93.7%; Private dataset (2-class): Acc = 94.9%, Sens = 93.3%, Spec = 93.3%
Dong et al. (2020)	Hand-crafted features (Color, texture and morphological featurs) + Inception v3	Herlev	2-class, 7-class	Herlev (2-class): Acc = 98.23%, Sens = 99.44%, Spec = 96.73%; Herlev (7-class): Acc = 94.68%
Zhang et al. (2022)	MDHDN (spectrum images + VGG-19 + Hand-crafted features)	Herlev, SIPaKMed and private BJTU dataset (735 normal cells and 1756 abnormal cells)	2-class, 5-class, 7-class	Herlev (2-class): Acc = 98.7%, Sens = 98.2%, Spec = 98.9%; Herlev (7-class): Acc = 94.8%, Sens = 93.7%, Spec = 91.1%
Yaman and Tuncer (2022)	Multi-resolution images + DarkNet + NCA + SVM	Mendeley LBC and SIPaKMed	4-class, 5-class	SIPaKMed (5-class): Acc = 98.26%, Pre = 98.27%, Rec = 98.28%, F1 = 98.73%; Mendeley LBC (4-class): Acc = 99.47%, Pre = 99.26%, Rec = 98.21%, F1 = 98.73%
Qin et al. (2022)	Multi-task feature fusion model (manual features fitting brance + multi-task classification branch)	SIPaKMeD and private HUSTC dataset (70,197 single-cell images in 5 categories)	2-class, 5-class	HUSTC (2-class): Acc = 99.52%, F1 = 99.71%; HUSTC (5-class): Acc = 81.88%, F1 = 79.10%; SIPaKMeD (2-class): Acc = 98.96%, F1 = 98.72%; SIPaKMeD (5-class): Acc = 98.67%, F1 = 98.67%

In addition to using the classical CNN architectures or self-designed models, there are three commonly used approaches for cervical cell identification: transfer learning, multi-model ensemble, and hybrid feature fusion, as shown in Fig. 9. A summary of the deep learning-based methods for cervical cell classification is exhibited in Table 3.

4.3.1 Transfer learning based identification

The success of deep learning is closely related to large amounts of data, which means that insufficient training data can seriously affect the performance of deep learning models. However, one problem with applying deep learning to medical image analysis is the lack of effective annotation. Limited labels result in limited available data, which makes deep learning models difficult to train well and brings overfitting problems. Therefore, transfer learning is an effective alternative in this case (Pan and Yang 2009). In contrast to general deep learning algorithms that solve isolated tasks, transfer learning attempts to transfer learned knowledge in the source task and apply it to improve learning in the target task, such as transferring knowledge from a large public dataset (e.g. ImageNet) to a dome-specific task (e.g. Cervical cell identification), as shown in Fig. 9a. The application of transfer learning in the field of cervical cell identification can save a significant amount of labeling effort, reduce overfitting problems and improve the generalization ability of deep learning models. To transfer deep learning models, fine-tuning and feature extraction are two common strategies (Yu et al. 2022). Fine-tuning needs to train the pre-trained model which is obtained from the source dataset on the target dataset to fine-tune all parameters in the learnable layers of the networks. Feature extraction remains the same parameters in all layers except the top layer. The top layer connects with the classifier and is related to the specific classification task.

Zhang et al. (2017) first introduced a transfer learning approach to cervical cytology screening for both conventional Pap smear and liquid-based cytology datasets. They proposed a simple ConvNet, DeepPap, to classify the cervical cells into healthy and abnormal. The proposed ConvNet was firstly pre-trained on a natural image dataset, ImageNet and then fine-tuned on cervical cytological datasets. On both the CPS dataset, Herlev, and the LBC dataset, HEMLBC, the proposed ConvNet presented high-performance classification results.

Hyeon et al. (2017) utilized VGGNet-16 which was pre-trained on the ImageNet dataset to extract features of cervical cells and then trained an SVM classifier to perform the prediction. They collected 71,344 Pap smear microscopic images classified into six categories according to TBS criteria. To mitigate the imbalanced distribution they downsampled and regrouped all images into two classes: normal (8373) and abnormal (8373). Using 80% of the images for training and the rest for testing, the SVM classifier achieved the best performance with a 0.7817 F1 score when compared to logistic regression, random forest, and AdaBoost.

Nguyen et al. (2018) proposed a DL-based approach for microscopic image classification based on transfer learning and feature concatenation. They leveraged three different deep CNN models, namely Inception-v3, Resnet152, and Inception-Resnet-v2 which were pre-trained on ImageNet, to extract the initial features of cervical cells. Then, they concatenated the extracted features and used two extra fully connected layers to fuse and compress the feature for final classification. The proposed method achieved an average accuracy of 92.63% on the Herlev dataset showing its nice performance for cervical cell classification.

Ghoneim et al. (2020) introduced CNNs and extreme learning machine (ELM)-based classifier in cervical cell classification. They compared the shallow CNN model with two deep CNN models, VGG-16 and CaffeNet. Three deep learning models were fine-tuned on the Herlev dataset, and the proposed CNN-ELM-based system achieved 99.5% accuracy in the 2-class classification and 91.2% in the 7-class classification.

Khamparia et al. (2020) proposed a novel Internet of health Things (IoHT)-driven diagnostic system for cervical cancer. To classify abnormal cervical cells, they leveraged several classical CNN models (InceptionV3, VGG19, SqueezeNet, and ResNet50) as the feature extractor in conjunction with multiple machine learning classifiers (K nearest neighbor, naive Bayes, logistic regression, random forest, and support vector machines.) for final prediction. ResNet50 together with the random forest classifier achieved the highest classification accuracy of 97.89%. They also developed a web application for the prediction of uploaded test images and the proposed IoHT system can greatly improve the diagnosis efficiency of cytologists.

Wang et al. (2020) presented an adaptive pruning deep transfer learning model (PsiNet-TAP) to classify Pap smear images. PsiNet-TAP consists of 10 convolution layers and is firstly pre-trained on the ImageNet dataset. After that, transfer learning is applied by using the pre-trained weights as the initialized weights to fine-tune the model on Pap smear images. Furthermore, to discard all unimportant convolution kernels, they designed an adaptive pruning method based on the product of l1-norm and output excitation mean. Using their collected 389 cervical Pap smear images, PsiNet-TAP achieved a remarkable performance of more than 98% accuracy.

Bhatt et al. (2021) utilized progressive resizing together with a transfer learning technique to train several generic CNN models for the identification of cervical cells. They performed binary and multiclass experiments on Herlev and SIPaKMed datasets. The experimental results demonstrated the high performance of the proposed method and the activation results of GradCam highlights the pre-malignant or malignant lesions located by the proposed model.

4.3.2 Multi-model ensemble based identification

Ensemble learning is a machine learning technology that exploits multiple base learners to produce predictive results and fuse results with various voting mechanisms to achieve better performances of the learning systems (Yang et al. 2022). The basic guiding principle of ensemble learning is 'many heads are better than one'. In recent years, with the rapid development of deep learning, ensemble deep learning has been widely applied in biomedical and bioinformatic fields (Cao et al. 2020; Ganaie et al. 2022). The multi-model ensemble is the most straightforward way to realize ensemble deep learning. The diversity of individual networks is the essential characteristic of multi-model ensemble learning and various integration strategies can assist the basic model for better performance. The ensemble across multiple models has been a promising direction to improve accuracy for cervical cell identification, as illustrated in Fig. 9b.

Rahaman et al. (2021) proposed a hybrid deep feature fusion (HDFF) approach, Deep-Cervix for the multiclass classification task of cervical cells. Four deep learning networks, VGG16, VGG19, XceptionNet, and ResNet50 were used to extract the features and the subsequent feature fusion network was utilized to concatenate the extracted features to perform the final prediction. The HDFF Network achieved an accuracy of 99.85% for 2-class,

99.38% for 3-class, and 99.14% for 5-class classification on the SIPaKMeD dataset. For the Herlev dataset, the proposed method achieved 98.32% and 90.32% for 2-class and 7-class classification respectively.

Manna et al. (2021) developed an ensemble-based model for cervical cell classification using three general CNN models, Inception v3, Xception, and DenseNet-169. They presented a novel ensemble technique that the prediction scores of three CNN models were taken into account to make the final decision. The proposed ensemble method leveraged a fuzzy ranking-based approach, where two non-linear functions were applied to the probability scores of each base learner to determine the fuzzy ranks of the classes. The ranks assigned by the two non-linear functions were multiplied and the ranks of the three base learners were added and the lowest rank was assigned as the predicted class. Extensive experiments on two public datasets, SIPaKMeD, and Mendeley LBC demonstrated the high performance of the proposed method in terms of classification accuracy and sensitivity.

Diniz et al. (2021) proposed a simple but effective ensemble method to classify cervical cells. After the selection of the three best-trained models from all models, the final prediction was generated by the vote of these three models' predictions. Using the public CRIC dataset, the proposed ensemble method outperformed EfficientNet, MobileNet, Inception-NetV3, and XceptionNet, showing its effectiveness in cervical cell classification.

Madhukar et al. (2022) integrated two classical DL networks, VGG16 and ResNet50, to realize the classification of cervical cells. By fusing the 2048 dimensional feature vector from each network, the concatenated features were finally used for classification. On the publicly available CRIC dataset, the proposed method achieved accuracies of 96.07%, 93.30%, and 85.07% on the test set for 2-class, 3-class, and 6-class classification, respectively.

Liu et al. (2022) proposed a DL-based framework, CVM-Cervix for cervical cell classification. CVM-Cervix first combined a CNN module with a visual transformer module to extract local and global features from cervical cell images. The Xception model was used as the CNN module to generate 2048-dimensional local features and the tiny DeiT model was used as a vision transformer module to generate 192-dimensional global features. Then a multilayer perceptron module fused the local and global features to perform the final identification. CVM-Cervix was evaluated on the combination of CRIC and SIPaKMeD datasets, which included 11 categories in total. The experimental results demonstrated the effectiveness of the proposed CVM-Cervix to classify cervical Pap smear images. To meet the practical needs of clinical work, they also introduced a lightweight post-processing to compress the model by using a quantization technique to reduce the storage space of each weight from 32 to 16 bits. The model parameter size was greatly reduced while the classification accuracy remained almost unchanged.

Kundu and Chattopadhyay (2022) employed an evolutionary metaheuristic algorithm, named Genetic Algorithm to select the features which were extracted from GoogLeNet and ResNet-18 models. After feature selection, an SVM served as the classifier to perform the final prediction. The proposed method achieved 99.07% accuracy and 98.31% F1-score on the Mendeley LBC dataset. For the SIPaKMeD dataset, the proposed method achieved 99.65% and 98.94% for 2-class and 5-class classification.

4.3.3 Hybrid feature fusion based identification

Although the DL-based model has achieved good results in the task of cervical cell classification, there is still a lot of room for improvement. Hand-crafted features, especially some features related to cell morphology, contain rich domain knowledge in the medical field. Incorporating medical domain knowledge with the deep learning network can promote the effective attention of the network and further improve the network performance. Figure 9c shows a general example of combining DL-based features with manual cytological characteristics.

Jia et al. (2020) proposed a novel deep learning-based framework called strong feature CNN-SVM. Gray-Level Co-occurrence Matrix (GLCM) and Gabor were used to calculate the strong features. The strong features were fused with abstract features extracted by CNN and then they were sent into the SVM for final prediction. The experimental results on two independent datasets indicated the effectiveness of the strong feature CNN-SVM model in cervical cytology screening.

Dong et al. (2020) proposed an innovative cell recognition algorithm that combines hand-crafted features with automatically extracted features via the Inception v3 network. To address the low universality of artificial feature extraction while maintaining the cervical cell domain knowledge, they extracted both deep features and hand-crafted features and leveraged a fully connected layer to fuse these features. Furthermore, this paper also utilized an image enhancement algorithm to reduce noise generated during image acquisition and conversion and improve the overall performance. Based on the public Herlev dataset, the proposed method achieved an accuracy of 98.23% for 2-class classification and an accuracy of 94.68% for 7-class classification.

Zhang et al. (2022) proposed a novel multi-domain hybrid deep learning framework (MDHDN) to classify cervical cells. It was the first time to apply cell spectrum for cervical cell classification. MDHDN was a three-path cooperative framework, in which two subpaths were used to extract deep features from the time and frequency domains respectively using the VGG-19 network, and the other subpath was used to extract and select hand-crafted features. The final classification results were obtained through the correlation analysis of the prediction of the three paths. On the Herlev dataset, MDHDN acquired an
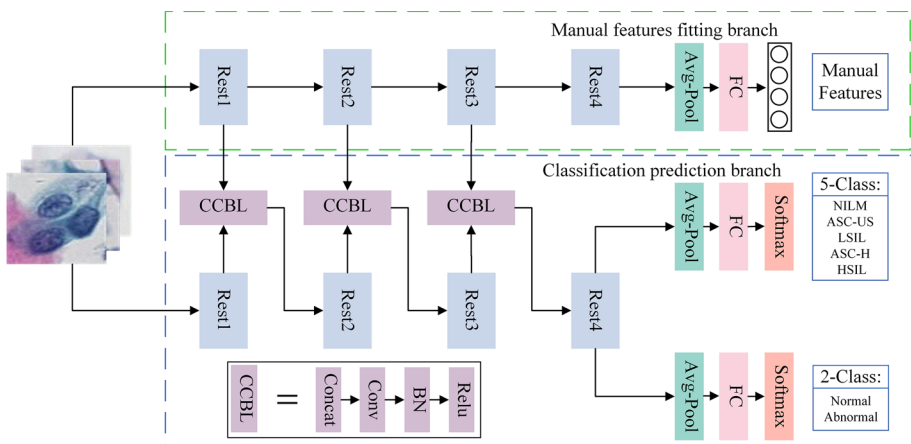

Fig. 10 Multi-task feature fusion model for cervical cell classification (Qin et al. 2022)

accuracy of 98.7% for 2-class classification and 94.8% for 7-class classification. The proposed framework also presented an excellent performance on the public SIPaKMeD dataset and their collected in-house dataset BJTU.

In Yaman and Tuncer (2022), Yaman et al. designed an exemplar pyramid deep feature extraction model for the classification of cervical cells. They fed pap-smear images of different resolutions into the DarkNet19/DarkNet53 to get the pyramid features. Then, a Neighborhood Component Analysis (NCA) algorithm was deployed to select the most discriminative features and an SVM classifier was utilized to execute the final classification. SIPaKMeD and Mendeley LBC datasets were used for method validation. Experimental results demonstrated that the proposed method outperformed some mainstream classification models such as ResNet, DenseNet, InceptionV3, Xception, etc.

Qin et al. (2022) presented a multi-task feature fusion model that performed binary classification and 5-class classification for cervical cells (see Fig. 10). The whole model consisted of a manual features fitting branch and a multi-task classification branch. They utilized CE-Net (Gu et al. 2019) to segment cervical cells for further manual feature acquirement. Multiple discriminatively hand-crafted features including morphological features, integral optical Dens, and texture features were obtained and utilized in the manual features fitting branch to supply prior knowledge for more precise classification. They also utilized smoothing noisy label regularization and supervised contrastive learning strategy for model training. On the SIPaKMeD dataset, the proposed method achieved accuracy of 98.96% and 98.67% for 2-class and 5-class classification, which surpassed other SOTA methods. On the self-built dataset, the proposed method also achieved the best performance.

4.3.4 Method analysis and summary

In this section, we have surveyed in detail the application of deep learning in the task of cervical cell identification. As deep learning models continue to make breakthroughs in computer vision, many researchers have attempted to leverage them in cervical cytology identification. In the early period (2016–2017), simple models with a few convolutional layers, or classical networks such as AlexNet, VGG, ResNet, etc. were used. Due to the small number of images in cervical cytology datasets compared to natural image datasets such as ImageNet, transfer learning brings a good parameter initialization method to the cell identification task, which can help in the model training process. Tranfer learning reduced overfitting problems and improved the generalization ability of DL-based models for cervical cell identification. DeepPap (Zhang et al. 2017) is a typically successful example that has achieved excellent performance on images of both sample preparation schemes, CPS and LBC. After the release of the public SIPaKMeD dataset (Plissiti et al. 2018) in 2018, more options are available for practice. At present, Herlev and SIPaKMeD datasets are still the two most used publicly available datasets in the automated analysis of cervical cytology. Later, attempts have been made to adapt the network to make the model more suitable for cervical cell identification rather than simply applying the classical models (Chen et al. 2021; Khamparia et al. 2021). For instance, ensemble learning has been verified to be effective in improving the models, and multi-model ensemble based identification for cervical cells has been widely used (Rahaman et al. 2021; Manna et al. 2021; Diniz et al. 2021; Madhukar et al. 2022; Liu et al. 2022; Kundu and Chattopadhyay 2022). With the popularity of self-attention mechanism and Vision Transformer (ViT), CVM-Cervix (Liu et al. 2022) effectively integrates CNN and ViT modules to construct a

more powerful classification system. CVM-Cervix has been extensively validated on multiple cervical cell datasets and even the peripheral blood cell dataset for similar tasks. In addition to model improvement, data augmentation, and pre-processing methods can also enhance the model's accuracy and generalization performance (Martinez-Mas et al. 2020; Yu et al. 2021).

To further improve performance for better application in practical production, introducing more new information into deep learning models beyond existing medical datasets is a more promising approach (Xie et al. 2021). Experienced cytopathologists can usually provide fairly accurate diagnostic results, so their knowledge can better assist deep learning models in classifying cervical cells. A most straightforward solution is to concatenate hand-crafted features with the features extracted from deep learning models, as hand-crafted features contain rich biomedical knowledge, especially related to cell morphology (Jia et al. 2020; Dong et al. 2020; Zhang et al. 2022; Yaman and Tuncer 2022; Qin et al. 2022). Visual attention can also be used to simulate the focus areas of doctors' diagnoses (Yu et al. 2022; Su et al. 2021; Jiang et al. 2022). When deploying the model in actual production to conduct large-scale screening programs, the execution speed of the model should be also given priority, thereby some researchers have used knowledge distillation methods (Gao et al. 2022; Chen et al. 2022) or designed lightweight modules (Wang et al. 2019) to improve the real-time performance of the model. In the future, DL-based methods for cervical cell identification should be high-performance and efficient. More deep learning methods considering stronger feature extractors together with lightweight design and incorporation with biomedical domain knowledge should be explored.
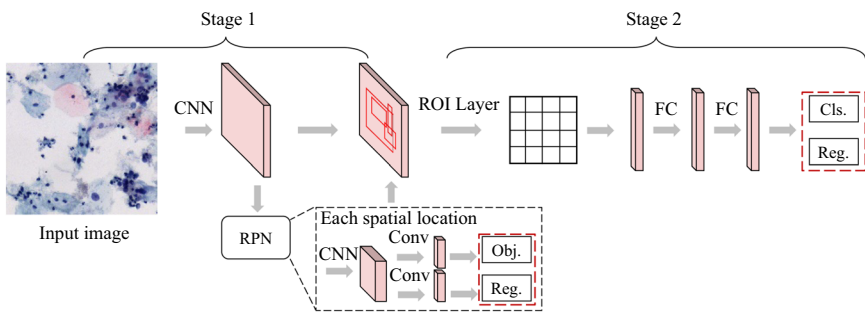
4.4 Abnormal cell detection

Identifying thousands of cells in a specimen using a classification network alone is time-consuming and inefficient. Thus, a fast search and localization of suspicious abnormal cervical cells are essential for cervical image analysis which further affects the slide-level diagnosis in cervical cytology screening. Object detection models from the computer vision field which simultaneously locate the objects and predict the categories have been well studied and applied in abnormal cervical cell detection.

After the first CNN-based object detection framework R-CNN (Girshick et al. 2014) was put forward, a series of improved algorithms have been proposed which greatly promote the development of generic object detection (Zou et al. 2023; Wu et al. 2020). There are mainly two types of generic object detection methods: two-stage object detection which involves two stages of region proposal and object detection, and one-stage object detection which directly predicts object bounding boxes and class labels in a single pass (Zhao et al. 2019). The two-stage object detection is preferred in scenarios where high detection accuracy is required, and the object instances are small or densely packed. In the region proposal stage, the algorithm first generates a set of candidate regions of interest in the image. These regions are proposed as potential locations of objects, and the goal is to reduce the number of regions to be processed in the second stage. This is usually achieved by using algorithms like Selective Search (Uijlings et al. 2013), EdgeBoxes (Zitnick et al. 2014), or Region Proposal Networks (RPN) (Ren et al. 2015). In the object detection stage, the algorithm processes the candidate regions generated in the previous stage and assigns object class labels and bounding boxes to each region. Object detection is usually performed using deep learning models, such as the popular Faster R-CNN (Ren et al. 2015)(see Fig. 11a), R-FCN (Dai et al. 2016), FPN (Lin et al. 2017) or Cascade R-CNN (Cai and Vasconcelos

2019), which use convolutional neural networks (CNNs) for feature extraction and classification. When considering the problem of detection speed, one-stage methods are better choices. One-stage object detection algorithms are typically faster and more efficient than two-stage approaches, as they don't require an initial region proposal step. However, they are generally less accurate, particularly for smaller objects or objects with high levels of occlusion. Some popular examples of one-stage object detection algorithms include YOLO (Redmon et al. 2016) (see Fig. 11b), SSD (see Fig. 11c), RetinaNet (Lin et al. 2017) and RefineDet (Zhang et al. 2018).

In clinical practice, it's hard to build a high-quality dataset for cervical cell detection since the annotation of cervical cells depends heavily on professional medical knowledge. Thus, some semi-supervised methods have also been explored to detect abnormal cervical cells. In this section, we not only review supervised learning-based methods (Sect. 4.4.1

(a) Two-stage detection model: Faster R-CNN [16]
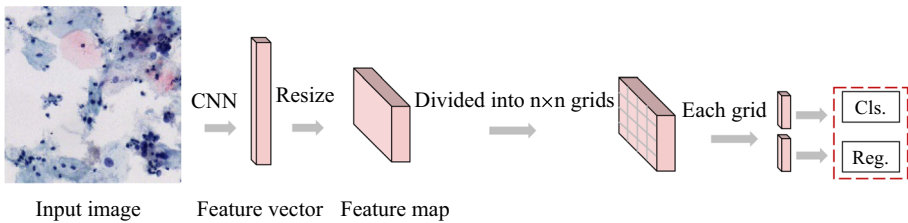
(b) One-stage detection model: YOLO [143]
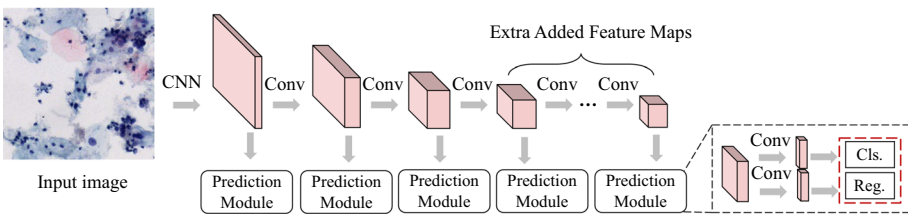
(c) One-stage detection model: SSD [146]

Fig. 11 Three commonly used detection models

and Sect. 4.4.2) for cervical cell detection but survey the latest semi-supervised learning-based methods as well (Sect. 4.4.3) (Table 4).

4.4.1 One-stage supervised learning based detection

Zhang et al. (2019) constructed the Deep Cervical Cytological Lesions (DCCL) dataset, which contained 14,432 image patches with 27,972 annotated lesion cells collected from 1,167 WSIs. They leveraged a two-stage algorithm, Faster R-CNN (Ren et al. 2015) and a one-stage algorithm, RetinaNet (Lin et al. 2017) to evaluate the collected DCCL dataset. This new benchmark dataset provided large-scale samples with full annotations on six types of abnormal cells, that could benefit future research and clinical studies on cervical cytology analysis.

Xiang et al. (2020) utilized CNN-based object detection to achieve the recognition of cervical cells. They exploited YOLOv3 as the baseline model and cascaded a further task-specifical classifier to improve the classification performance of hard examples. Furthermore, to relieve the problem of unreliable annotations, they smoothed the distribution of noisy labels. To evaluate the proposed method, they built a dataset composed of 12,909 cervical images with 58,995 ground truth boxes. All labels corresponded to 10 categories. The proposed method eventually achieved an mAP of 63.4% and improved the detection precision of hard samples.

In Ma et al. (2020), the authors designed a specialized booster, CCDB, for cervical cancer detection according to the medical knowledge of cervical cytology and the characteristics of cancerous cervical cells. CCDB consisted of two components: a refinement module (RM) to make better use of the detail features in Pap smear images and a spatial-aware module (SM) to consider the spatial context information of the cell. The whole detection model included a ResNet50 as backbone, an FPN for feature fusion, a CCDB module, and a dense detection head used in RetinaNet. On the Tian-chi competition dataset, with the installation of the CCDB module, several mainstream detection models, Faster R-CNN (Ren et al. 2015), Cascade R-CNN (Cai and Vasconcelos 2019), FreeAnchor (Zhang et al. 2019), RetinaNet (Lin et al. 2017), et al. all improved its performance for the detection of abnormal cervical cells. Experimental results showed the CCDB module can be used in general detectors to achieve better performance in cervical cancer detection tasks.

Nambu et al. (2022) proposed a two-step screening assistance system for detecting atypical cervical cells. The first step was a quick detection based on YOLOv4 and the second one was a further classification of the localized cells using a ResNeSt model. Experimental results showed that the developed system enabled high sensitivity with fast detection speed.

To relieve the problem that general CNN-based detectors might yield too many false positive predictions, Liang et al. (2021) proposed a global context-aware framework based on YOLOv3 using an image-level classification branch (ILCB) and a weighted loss to filter false positive predictions. Besides, they presented a soft scale anchor matching (SSAM) method to assign objects to anchors more softly. This paper carried out substantial experiments to evaluate the proposed method and the experimental results validated the effectiveness of the proposed method, which achieved an mAP of 65.44% and gained a 5.7% increase in mAP together with an 18.5% increase in specificity.

Jia et al. studied one-stage detection methods for cervical cancer cells carefully (Jia et al. 2022, 2022). They improved the SSD model by fusing feature maps between different layers in the first work. For the second work, they improved the YOLOv3 model by using dense blocks and S3Pool algorithm. To further enhance the performance of cervical cell

Table 4 Summary of deep learning-based studies for abnormal cell detection. Accuracy (Acc), Precision (Pre), Recall (Rec), Specificity (Spec), Sensitivity (Sens), Average precision (AP), Mean average precision (mAP)

Reference	Method	Dataset	Classes	Result
One-stage supervised learning				
Zhang et al. (2019)	Faster R-CNN, RetinaNet	Private DCCL dataset (14,432 cervical images with 27,972 annotations corresponded to 7 categories)	2-class, 7-class	Faster R-CNN (2-class): $mAP_1 = 0.2616$, $mAP_3 = 0.2218$, $AP_5 = 0.1935$; RetinaNet (2-class): $mAP_1 = 0.2401$, $mAP_3 = 0.2185$, $AP_5 = 0.1807$; Faster R-CNN (7-class): $mAP_1 = 0.2116$, $mAP_3 = 0.198$, $AP_5 = 0.171$; RetinaNet (7-class): $mAP_1 = 0.1961$, $mAP_3 = 0.1837$, $AP_5 = 0.1593$
Xiang et al. (2020)	YOLOv3 + Inception V3+ Label smoothing	Private (12,909 cervical images with 58,995 annotations corresponded to 10 categories.)	10-class	mAP = 0.634
Ma et al. (2020)	RetinaNet + CCDB module	Tian-chi competition dataset	2-class	mAP = 0.653
Nambu et al. (2022)	YOLOv4 + ResNeSt	Private (919 cervical cell images)	6-class	Detection: AP = 0.542; Classification: Acc = 0.905, Pre = 0.718, Rec = 0.708
Liang et al. (2021)	YOLOv3 + ILCB+ SSAM	Private (12,909 cervical images with 58,995 annotations corresponded to 10 categories.)	10-class	mAP = 0.6544
Jia et al. (2022)	Improved SSD	Private	4-class	mAP = 0.8153
Jia et al. (2022)	Improved YOLOv3	Private (2,000 cell images of 200 patients)	7-class	mAP = 0.7887
Two-stage supervised learning				
Liu et al. (2018)	Pre-trained VGG16 network + Task oriented anchor network + Fully convolutional network	Private (73 images with 98 annotated lesion cells)	2-class	AP = 0.27, Sens = 0.30
Sompawong et al. (2019)	Mask R-CNN	Private (2,734 normal cells, 494 atypical cells, 148 low-grade cells, and 84 high-grade cells)	4-class	mAP = 0.578, Acc = 0.917, Sens = 0.917, Spec = 0.917
Zhang et al. (2019)	R-FCN (Net-22 as feature extractor)	Private (62 cervical cell images including 180 abnormal regions)	2-class	AP = 0.932

Table 4 (continued)

Reference	Method	Dataset	Classes	Result
Yi et al. (2020)	Dense-Cascade R-CNN	Herlev dataset	2-class	mAP = 0.979
Li et al. (2021)	DGCA-RCNN	Tian-chi competition dataset	2-class	$AP_{.1} = 0.505$, $AP_{.3} = 0.486$, $AP_{.75} = 0.445$
Yan and Zhang (2021)	HSDet	Private (1,000 WSIs)	2-class	mAP = 0.571
Liang et al. (2021)	Comparison detector (Faster R-CNN + Prototype + Few-shot learning)	ComparisonDetector (7,410 cervical microscopical images with 48,587 object instance bounding boxes in 11 categories)	11-class	Small dataset: mAP = 0.263, Rec = 0.357; Medium-sized dataset: mAP = 0.488, Rec = 0.640
Wang et al. (2022)	3cDe-Net (DC-ResNet as backbone + FPN)	Tian-chi competition dataset and Herlev	2-class	mAP = 0.504
Xu et al. (2022)	Faster R-CNN + FPN + Transfer Learning + Multi-scale Learning	ComparisonDetector	11-class	mAP = 0.616, Rec = 0.877
Liu et al. (2022)	RepPoints + FPN + Grad-Libra Loss	Private dataset	10-class	mAP = 0.531
Chen et al. (2022)	TDCC-Net	Private (6,935 cervical cytologic images including 22,054 ground-truth boxes in 5 categories: ASCUS, LSIL, ASCH, HSIL and SCC)	5-class	mAP = 0.256, $AP_{.5} = 0.477$, $AP_{.75} = 0.255$
Semi-supervised learning				
Zhang et al. (2021)	CLCR-STNet	Private (6 categories: Normal, ASCUS, LSIL, ASCH, HSIL and SCC)	6-class	30% labeled: mAP = 0.3207; 50% labeled: mAP = 0.3532; 70% labeled: mAP = 0.3556
Du et al. (2021)	RetinaNet + MT + attention mechanism	Private (9,000 images and 2,050 testing images)	2-class	20% labeled: mAP = 0.579, Pre= 0.28, Rec = 0.78
Chai et al. (2022)	Faster R-CNN + Proposal alignment + Prototype alignment	Private (240,860 images containing lesions of ASCUS, LSIL, HSIL and AGC)	4-class	25% labeled: mAP=0.170; 50% labeled: mAP = 0.195; 75% labeled: mAP = 0.254

detection, they did anchor cluster analysis based on k-means++ to select the proper anchor size for cervical cells and adjusted the loss function for better training. Both of these two works achieved good detection accuracy for abnormal cervical cells.

4.4.2 Two-stage supervised learning based detection

Liu et al. (2018) presented a multi-task learning network to detect squamous intraepithelial lesions on cervical cytology images. They first proposed a task-oriented anchor (TOA) network based on a pre-trained VGG16 model to generate potential RoIs. Then, a multi-task learning network for both localization and classification tasks was designed to realize the detection of lesional cells. The experimental results demonstrated the proposed method achieved the best detection accuracy when compared to Faster R-CNN and YOLO.

Sompawong et al. (2019) applied the Mask Regional Convolutional Neural Network (Mask R-CNN) to detect cervical cancer. In detail, they leveraged ResNet-50 which was pre-trained from ImageNet as the backbone and used a feature pyramid network (FPN) as the detection neck to better select and fuse features. Based on their collected liquid-based dataset, the proposed method obtained an mAP of 57.8%, accuracy of 91.7%, sensitivity of 91.7%, and specificity of 91.7%.

Zhang et al. (2019) utilized a region-based, fully convolutional network (R-FCN) for abnormal region detection in cervical cytology screening. Inspired by ResNet, they designed a new feature extractor called Net-22, which consisted of 22 convolutional layers including the structure of the residual block. Experimental results showed that the R-FCN gained an average precision of 93.2%.

Yi et al. (2020) presented Dense-Cascade Region-based Convolutional Neural Networks (Dense-Cascade R-CNN) to automatically detect cervical cells. Dense-Cascade R-CNN was adapted from Cascade R-CNN and replaced the backbone of it from 101-layer ResNet to 121-layer DenseNet. They also carefully selected the specific combinations of data augmentation operations and used a training set balancing (TSB) algorithm to balance the training set. Based on the public Herlev dataset, the proposed Dense-Cascade R-CNN achieved high detection accuracy with an mAP of 97.9% and mAR of 98.8%.

Li et al. (2021) proposed a novel detection model, deformable and global context aware Faster R-CNN (DGCA-RCNN), to detect abnormal cervical cells in cytology images. DGCA-RCNN improved the original FPN-based Faster R-CNN by introducing deformable convolutional layers and a global context aware (GCA) module. The proposed DGCA-RCNN was evaluated on the public Tian-chi competition dataset and achieved the best performance compared with other SOTA detectors.

In Yan and Zhang (2021), the authors proposed a novel cervical cell detector, HSDet to make better use of negative samples. They adopted HRNet (Sun et al. 2019) as a feature extractor to cooperate with the cascade R-CNN (Cai and Vasconcelos 2018). Besides, they proposed a pair sampling method to generate the sample pair images and a hybrid sampling strategy to balance hard samples with simple samples. Combining the above methods with HSDet, false detections were effectively decreased. On the in-house dataset consisting of 1000 WSIs, HSDet achieved an mAP of 57.1%, surpassing the Faster R-CNN and Cascade R-CNN models.
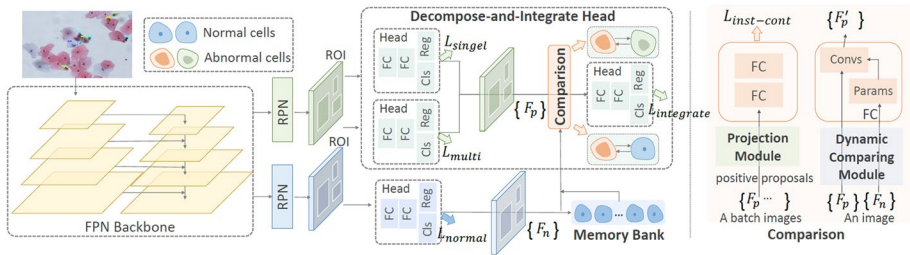
Fig. 12 The architecture of TDCC-Net (Chen et al. 2022)

Liang et al. (2021) proposed an end-to-end cervical cells/clumps detection method called Comparison detector. The Comparison detector utilized Faster R-CNN with FPN as the basic network and adapted the classifier to compare each proposal with the prototype representations of each category. They also investigated the generation method of prototype representations for the background category and considered different designs of the head model. For experiments, the Comparison detector obtained an mAP of 48.8% on its collected dataset. It's worth noting that on the constructed small dataset, the Comparison detector improved by about 20% accuracy than the baseline model.

Wang et al. (2022) presented a cervical cancer cell detection algorithm called 3cDe-Net, to address the issue of cell overlap with blurred cytoplasmic boundaries in clinical practice. 3cDe-Net consisted of an improved backbone network named DC-ResNet by introducing dilated convolution and group convolution and a multiscale feature fusion based detection head. Based on the Faster R-CNN algorithm, this paper also generated adaptive anchors and defined a new balanced loss function. The proposed method was evaluated on two publicly available datasets, the Tian-chi competition dataset (Data-T) and the Herlev dataset. Extensive experiments demonstrated the effectiveness of a novel backbone network, DC-ResNet. Besides, the proposed detection algorithm 3cDe-Net achieved an mAP of 50.4%, which significantly improved the performance of the original Faster R-CNN for cervical cancer cell detection.

Xu et al. (2022) studied a transfer learning-based method for the detection of cervical cells or clumps. Specifically, Faster R-CNN together with FPN was pre-trained on the COCO dataset and then fine-tuned on cervical cytological images for abnormal cell detection. The authors also utilized a multi-scale training strategy that randomly selected input scales to further improve the performance. The proposed method ultimately obtained an mAP of 0.616 and an average recall of 0.877.

Liu et al. (2022) proposed a Grad-Libra Loss to address the long-tailed data distribution problem in cervical cytology screening that normal or inflammatory cells were much more than cancerous or precancerous cells. Grad-Libra Loss considered the "hardness" of each sample and helped the detection model focus on hard samples in all categories. Various mainstream detectors were utilized to verify the performance of Grad-Libra Loss against the conventional cross-entropy loss. On the collected long-tailed CCA-LT dataset, Grad-Libra Loss presented excellent detection performance superior to other loss functions.

Chen et al. (2022) proposed a novel task decomposing and cell comparing network, TDCC-Net for cervical lesion cell detection (Fig. 12). To cope with the large appearance

variances between single-cell and multi-cell lesion regions, they decomposed the original detection task into two subtasks detecting single-cell and multi-cell regions, respectively. In addition, to better obtain lesion features and conform with clinical practice, they designed a dynamic comparing module to perform normal-and-abnormal cells comparison adaptively and present an instance contrastive loss to perform abnormal-and-abnormal cells comparison. Extensive experiments on a large cervical cytology dataset demonstrated that TDCC-Net achieved state-of-the-art performance in cervical lesion detection.

4.4.3 Semi-supervised learning based detection

In general, cervical cell detection has been done using supervised learning, where a model is trained on a set of labeled images to learn the patterns that indicate the presence of abnormal cells. However, obtaining a large number of labeled images can be difficult and time-consuming, especially in areas where access to healthcare is limited. Semi-supervised learning-based methods for cervical cell detection have been proposed in recent years to alleviate the above problem, which combines the use of labeled and unlabeled data to improve the accuracy of the model (Van Engelen and Hoos 2020). The model uses the labeled data to learn the patterns that indicate the presence of abnormal cells and then applies this knowledge to the unlabeled data to identify additional cases of abnormality. Semi-supervised learning for cervical cell detection has the potential to improve the accuracy of automated systems for detecting abnormal cells, especially in areas where labeled data is scarce.

Zhang et al. (2021) proposed a novel semi-supervised cervical cell detection method, called Classification and Localization Consistency Regularized Student-Teacher Network (CLCR-STNet). Since it was difficult to acquire large amounts of labeled data in the field of medical image analysis, this paper introduced a novel semi-supervised method that utilized both labeled and unlabeled data with online pseudo-label mining. Faster R-CNN was employed as the backbone network and Jensen-Shannon (JS) divergence was used to compute the consistency loss between student and teacher models. The experimental results demonstrated that the proposed CLCR-STNet effectively exerted the potential of unlabeled data and outperformed the supervised methods counterpart.

In Du et al. (2021), the authors devised a semi-supervised detection network to reduce the false positive rate in cervical cytology screening. To be specific, a RetinaNet was first employed to find the suspicious abnormalities and then a false positive suppression network based on the Mean Teacher (MT) model was utilized to execute the further fine-grained classification and decrease the false positive samples. MT model utilized both labeled and unlabeled data for training via the enforced consistency between the teacher network and student network. Moreover, the authors used the generated mask as an attention map to further improve the MT model. Using 20% labeled data and 80% unlabeled data for training, the proposed method achieved 88.6% accuracy which was comparable with the fully supervised method. Besides, the proposed method successfully reduced the false positive rate after using false positive suppressing.

Chai et al. (2022) delved into the semi-supervised method for cervical cancer cell detection. To learn more discriminative features, they proposed a deep semi-supervised metric learning network that performed a dual alignment of semantic features on both the proposal level and the prototype levels. Concretely, the pseudo labels were generated for the

unlabeled data to align the proposal features with the class proxy derived from the labeled data. Besides, to reduce the influence of possibly noisy pseudo labels, they further aligned the labeled and unlabeled prototypes. They also utilized a memory bank to store the labeled prototypes. The proposed method achieved an average mAP of 27.0% and surpassed another two state-of-the-art semi-supervised object detection methods, the consistency-based semi-supervised detection (CSD) model and the Mean Teacher model. Extensive experiments showed that the proposed method could improve the fully-supervised baseline through the use of metric learning.

4.4.4 Method analysis and summary

In this section, we have investigated DL-based methods in abnormal cervical cell detection. It was not until 2018 that researchers began to gradually apply deep learning methods to this task. Initially, detection algorithms such as two-stage methods (Faster R-CNN, R-FCN, Mask R-CNN, etc.) or single-stage methods (YOLO, SSD, etc.) were simply applied, and anchor sizes were reset based on the characteristics of cervical cell datasets (Zhang et al. 2019; Liu et al. 2018; Sompawong et al. 2019; Zhang et al. 2019). With the proposal of FPN (Lin et al. 2017), it was widely verified that multi-scale feature fusion is effective, and an increasing number of scholars studied various kinds of feature fusion networks. The design of multi-scale feature fusion networks has become an indispensable part of mainstream detectors. In the past few years, improvements to cervical cell detection models have mainly focused on the following aspects: 1) using stronger backbone networks to extract efficient basic representations (Yan and Zhang 2021; Jia et al. 2022; Yi et al. 2020; Wang et al. 2022), 2) using different feature fusion networks to aggregate and fuse multi-level features (Ma et al. 2020; Jia et al. 2022; Li et al. 2021), and 3) designing reasonable and efficient detection heads to better complete classification and regression tasks (Liang et al. 2021). Besides, some works explored two-step detection frameworks that first leveraged a detection algorithm to preliminary detect the suspicious cells and then used a classification network to further identify the detection results (Xiang et al. 2020; Nambu et al. 2022). In addition to model design, the device of loss function is also important to acquire high-performance detectors (Liang et al. 2021; Liu et al. 2022; Chen et al. 2022). Liang et al. (2021) introduced an extra image-level classification branch to predict whether there exists abnormal cells in images from a global perspective. This synergistic classification loss utilizes image global information to mimic the diagnostic approach of cytopathologists, who first determine whether suspicious abormal cells exist in a broad view then conduct fine-grained search and identify the categories. Similarly, the construction of TDCC-Net also uses the diagnostic habits of cytopathologists that the identification of abnormal cells should refer to the normal cells in the same image (Chen et al. 2022). TDCC-Net utilized a contrastive learning method to realize the comparison of different kinds of cells. The authors exploited a memory bank to store the normal cells and used a dynamic comparing module to complete the comparison of normal cells and abnormal cells. Besides, an instance contrastive loss was proposed to further compare the different abnormal cells. This is the first work to introduce a contrastive learning approach in cervical cell detection. By far there is still plenty of room for improvement in detection accuracy and stronger

detectors need to be explored via the development of high-performance network and the incorporation with medical knowledge.

Large-scale high-quality medical datasets are often difficult to construct, and cervical cytology datasets are no exception. Thus, some researchers have introduced semi-supervised learning methods to the detection of abnormal cervical cells since 2021 (Zhang et al. 2021; Du et al. 2021; Chai et al. 2022). They utilized a small number of labeled samples and a large number of unlabeled samples to deeply mine potentially diagnostic information from the unlabeled samples and achieved performance similar to those of fully supervised methods. For instance, CLCR-STNet (Zhang et al. 2021), the first work utilizing a semi-supervised deep learning method for cervical cell detection, employed a teacher model to generate pseudo labels for the student model and used a consistency loss to enforce a prediction consistency between these two models. Nevertheless, semi-supervised learning is a relatively blank and yet-to-be-explored research direction for abnormal cell detection in automated cervical cytology screening.

4.5 Cell region segmentation

Cervical cell segmentation is a process to identify and separate individual cervical cells from a digital image in cervical cytology screening. Even though the DL-based classification method has been widely applied to cervical cell identification, which does not need to accurately segment the contours of cervical cells, the segmentation of cervical cell regions is still the fundamental link to carrying out quantitative cell analysis (shape, size, texture, etc.). Besides, a precise segmentation of cell regions can provide detailed cytological features of clinical significance which can further support fine-grained cervical cell identification.

Cervical cell segmentation can be expressed as the problem of classifying pixels with semantic labels (semantic segmentation), especially the differentiation of the nucleus and cytoplasm, or the division of individual cells (instance segmentation). Traditional segmentation methods are generally based on thresholding, edge detection, region growing, k-means, clustering, or watershed methods (Minaee et al. 2021). With the development of deep learning and CNN, a new generation of deep learning-based segmentation models has yielded remarkable performance improvements, and gradually presented their potential for medical image segmentation (Hesamian et al. 2019). The Fully Convolutional Network (FCN) is a milestone in DL-based segmentation models which first introduces CNN into the task of semantic segmentation (Long et al. 2015). Inspired by FCN, U-Net (Ronneberger et al. 2015) has been proposed for biomedical image segmentation and received a good reputation and promotion. In addition to the above models, SegNet (Badrinarayanan et al. 2017), Mask R-CNN (He et al. 2017), DeepLab (Chen et al. 2017), and a series of improved methods have been developed to further enhance the performance of image segmentation. Figure 13 presents three commonly used models for cervical cell segmentation. In this section, the reviewed works encompass segmentation of both cell components (See Sect. 4.5.1) and overlapping cells (See Sect. 4.5.2), with the most relevant DL-based approaches being summarized in Table 5.
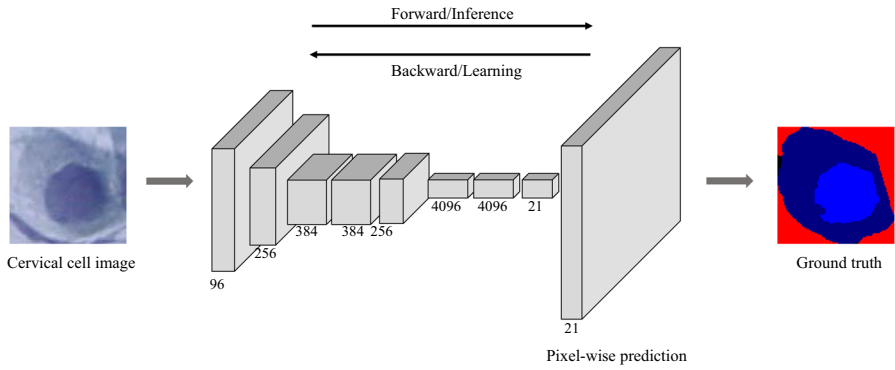
(a) The FCN model [172]
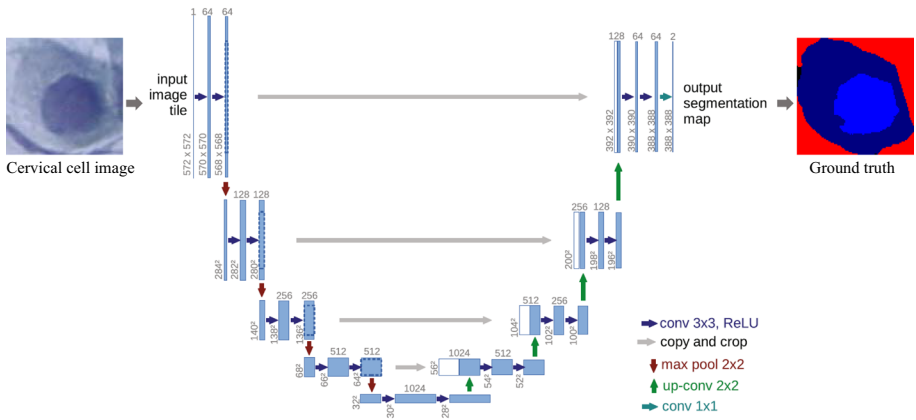
(b) The U-Net model [173]
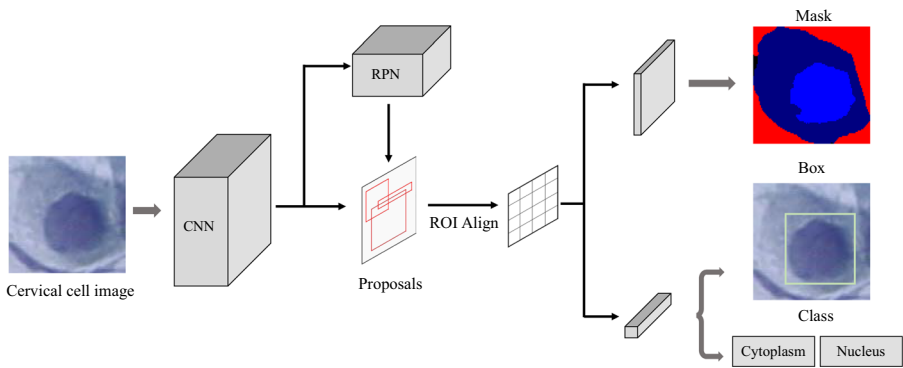
(c) The Mask R-CNN model [175]

Fig. 13 Three commonly used commonly used models for cervical cell segmentation

Table 5 Summary of deep learning-based studies for cervical cell segmentation. Nucleus (Nuc), Cytoplasm (Cyt), Accuracy (Acc), Precision (Pre), Recall (Rec), Specificity (Spec), Sensitivity (Sens), False negative rate (FNR), True positive rate (TPR), Zijdenbos similarity index (ZSI), Dice similarity coefficient (DSC), Average Jaccard Index (AJI), Mean intersection over union (mIoU)

Reference	Method	Dataset	Segmented region	Result
Segmentation of nucleus and cytoplasm				
Zhang et al. (2017)	FCN + graph-based approach	Herlev dataset	Nucleus	ZSI = 0.92
Gautam et al. (2018)	VGGNet-like network + selective pre-processing	Herlev dataset	Nucleus	ZSI = 0.90, Pre = 0.89, Rec = 0.91
Liu et al. (2018)	Mask R-CNN + LFCCRF	Herlev dataset	Nucleus	ZSI = 0.95, Pre = 0.96, Rec = 0.96
Zhang et al. (2019)	BTTFA (Binary-tree-like network topology + Two-path fusion attention)	ISBI 2014 and a new released dataset	Nucleus	ISBI 2014: DSC = 0.931; Released dataset: DSC = 0.91
Zhao et al. (2019)	D-MEM (Unet + Dense block + Deformable convolution + Multi-path ensemble model)	Herlev dataset	Nucleus	ZSI = 0.933, Pre = 0.946, Rec = 0.984
Zhao et al. (2020)	PGU-net+ (Unet + Residual module + Progressive growing method)	Herlev dataset	Nucleus	ZSI = 0.925, Pre = 0.901, Rec = 0.968
Hussain et al. (2020)	Shape context FCN	Combination of public Herlev, private LBC dataset (1670 raw images) and private conventional dataset (1320 raw images)	Nucleus	ZSI = 0.97
Yang et al. (2020)	ICPN (ICM + IPRCM)	Herlev dataset	Nucleus	ZSI = 0.972
Zhao et al. (2022)	LFANet (UNet + LFA module)	Herlev dataset	Nucleus and cytoplasm	Cyt DSC = 0.9454, AJI = 0.8979, Pre = 0.9453, and Rec = 0.9762; Nuc DSC = 0.9743, AJI = 0.9401, Pre = 0.9592, and Rec = 0.9748
Luo et al. (2022)	DSSNet (Dual-Supervised Sampling Network)	Mendeley-LBC, ISBI2014, and private TJ_sparse (3,600 images), TJ_dense (1,400 images)	Nucleus	TJ_sparse: mIoU = 0.7923; TJ_dense: mIoU = 0.6927; Mendeley-LBC: mIoU = 0.7458; ISBI 2014: mIoU = 0.8445

Table 5 (continued)

Reference	Method	Dataset	Segmented region	Result
Li et al. (2022)	U-Net + GDLA module	Herlev dataset	Nucleus and cytoplasm	Cyt ZSI = 0.796, Pre = 0.805, and Rec = 0.787; Nuc ZSI = 0.913, Pre = 0.888, and Rec = 0.936
Segmentation of overlapping cells				
Song et al. (2016)	Multi-scale CNN and deformation model	ISBI 2015 and private dataset (21 cervical cytology images and each image has 30-80 H &E stained cervical cells distributed in 7 clumps)	Overlapping cell	ISBI 2015: Cyt DSC=0.91, Nuc DSC = 0.93; Private SZU: Cyt DSC = 0.90, Nuc DSC = 0.92
Tareef et al. (2017)	Super pixel-wise CNN + dynamic shape modeling	ISBI 2014	Overlapping cell	Cyt ZSI = 0.90, Nuc ZSI = 0.94
Xu et al. (2018)	Light CNN model + SLIC + multi-cell labeling	ISBI 2014, ISBI 2015 and private dataset (14 scanned cervical cytology images)	Overlapping cell	ISBI 2014: DSC = 0.91, FNRo = 0.13, TPRp = 0.96; ISBI 2015: DSC = 0.90, FNRo = 0.21, TPRp = 0.89; Private: DSC = 0.90, FNRo = 0.24, TPRp = 0.90
Wan et al. (2019)	TernausNet model + double-window based cell localization + DeepLab V2 model + CRFs + DRLSE	ISBI 2014, ISBI 2015 and private dataset (14 cervical cytology images and each image has 2–20 cells)	Overlapping cell	ISBI 2014: DSC = 0.93; ISBI 2015: DSC = 0.92; Private dataset: DSC = 0.92
Zhou et al. (2019)	IRNet (Mask R-CNN + DRM + IRM)	Private dataset (413 images with 4,439 cytoplasm and 4,789 nuclei annotated)	Overlapping cell	Cyt AJI = 0.7185, F1 = 0.7497; Nuc AJI = 0.5496, F1 = 0.7554
Zhang et al. (2020)	Attention U-Net + graph-based Random Walk	ISBI 2014	Overlapping cell	Cyt DSC = 0.917, TP = 0.937, FP = 0.003; Nuc DSC = 0.93, Pre = 0.94, Rec = 0.95

Table 5 (continued)

Reference	Method	Dataset	Segmented region	Result
Zhou et al. (2020)	MMT-PSM (Mean Teacher framework + Perturbation-Sensitive Samples Distillation + Mask-Guided Feature distillation)	Private dataset (413 labeled and 4,371 unlabeled images)	Overlapping cell	100% labeled: AJI = 0.6643, mAP = 0.4052; 80% labeled: AJI = 0.6692, mAP = 0.4013; 40% labeled: AJI = 0.6449, mAP = 0.3726.
Huang et al. (2021)	Mask R-CNN	ISBI 2014 and ISBI 2015	Overlapping cell	ISBI 2014: DSC = 0.92, FNR = 0.02, TPR = 0.90; ISBI 2015: DSC = 0.89, FNR = 0.27, TPR = 0.92
Mahyari et al. (2022) Mahyari and Dansereau (2022)	Residual CNN model + multi-layer random walker image segmentation + Hungarian algorithm	ISBI 2014 and extended ISBI 2014 dataset (100,000 overlapped cervical cell images are newly created)	Overlapping cell	ISBI 2014: DSC = 0.89, FNR = 0.10,TPR = 0.90; Extended: DSC = 0.97, FNR = 0.03, TPR = 0.97

4.5.1 Segmentation of nucleus and cytoplasm

According to TBS (Nayar and Wilbur 2015), the morphological features, especially variations in the nucleus, are decisive factors supporting the precancerous lesions. There are a number of important specific cytological features that need a precise segmentation of nucleus and cytoplasm, such as nucleus area, cytoplasm area, nucleus/cytoplasm ratio, nucleus roundness, cytoplasm roundness, distribution of nucleus, etc. Thus, the accuracy and reliability of the segmentation algorithm can greatly affect the accuracy of subsequent cell feature extraction. Besides, the segmentation of the nucleus and cytoplasm plays a crucial role in the quantitative analysis of abnormal cells and the accurate diagnosis of cervical cancer. Numerous DL-based studies for the segmentation of cervical cell components have been investigated below.

Zhang et al. (2017) combined fully convolutional networks (FCN) and a graph-based approach for the automatic segmentation of cervical nuclei. The overall framework included two steps. FCN was first employed to coarsely split the background, cytoplasm, and nuclei in cervical cell images. Later, the graph-based approach was applied and incorporated with the FCN-learned nucleus probability map to yield fine-grained cell nucleus segmentation results. The proposed method finally obtained a ZSI of 0.92 on the Herlev dataset, superior to several traditional machine learning-based segmentation methods.

Gautam et al. (2018) put forward a novel approach for nuclei segmentation in Pap smear images based on deep CNN and selective pre-processing. They emphasize the importance of selective pre-processing since there were significant differences in the image characteristics (e.g. object sizes, chromatin pattern variability) for normal and abnormal cells. Using a VGGNet-like network, the proposed approach excellently accomplished nucleus segmentation on the Herlev dataset with a ZSI of 0.90.

Liu et al. (2018) provided pixel-level prior information to train a Mask R-CNN for cervical nucleus segmentation. ResNet together with the feature pyramid network (FPN) was utilized as the backbone of the Mask R-CNN to extract multi-scale features of the nuclei. To refine the segmentation result from Mask R-CNN's output, the authors leveraged a local fully connected conditional random field (LFCCRF). The experimental results on the Herlev dataset showed that the proposed method outperformed other prevailing methods with a precision of 0.96 and an average ZSI of 0.95.

Zhang et al. (2019) proposed a binary tree-like network with two-path fusion attention feature (BTTFA) for segmenting cervical cell nuclei. Due to the lack of real-world data for the cervical nucleus segmentation task, at the beginning of the work, they constructed a real-world clinical dataset including 104 LBC-based images with pixel-wise labels manually annotated by professional pathologists. BTTFA model selected ResNeXt as the backbone and utilized a binary tree-like network together with a two-path fusion attention to incorporate multi-level features which compensated for the information loss caused by the pooling layers. The proposed BTTFA was evaluated on the collected real-world dataset and the ISBI 2014 dataset. BTTFA obtained a DSC score of 0.91 on the released dataset, and a DSC score of 0.931 on the ISBI 2014 dataset which outperformed three classical segmentation networks, U-Net, FCN, and DeepLabv3+. The experimental results demonstrated that BTTFA provided a feasible method for cervical cell nucleus segmentation.

Zhao et al. (2019) suggested a unique method to segment cervical nuclei using the Deformable Multipath Ensemble Model (D-MEM). To build the D-MEM, U-Net was adopted as the basic network and dense blocks were exploited to transfer feature

information more effectively. To capture the irregular shape of abnormal cervical nuclei and make the network sensitive to subtle changes in objects, deformable convolutions were employed. Moreover, this paper created the multi-path ensemble model by training several networks simultaneously and integrating all paths' predictions for final results.

In Zhao et al. (2020), a progressive growing U-net (PGU-net+) model was presented to segment nuclei of cervical cells. Residual modules were inserted into different stages of the U-net to enhance the extraction ability of multi-scale features. Furthermore, the authors adopted the progressive growing method as the network training strategy that could significantly reduce computational consumption and effectively improve the segmentation performance. PGU-net+ gained a ZSI of 0.925 on the Herlev dataset and outperformed the original U-net.

Hussain et al. (2020) proposed a shape context fully convolutional neural network (FCN) to extract cervical nuclei which accomplished instance segmentation and classification on Pap smear images simultaneously. Based on standard U-Net architecture, they added residual blocks, densely connected blocks, and a bottleneck layer to build the final segmentation network. Besides, a stacked auto-encoder based shape representation model (SRM) was introduced to enhance the strength and robustness of the proposed FCN. To evaluate the performance of the proposed method, extensive experiments were carried out on the combination of three datasets (two clinical datasets and one public dataset, Herlev). The proposed method realized an average Zijdenbos similarity index (ZSI) of 0.97 and surpassed another two deep learning-based models U-Net and Mask R-CNN.

Yang et al. (2020) proposed an interacting convolution with pyramid structure network (ICPN) for end-to-end segmentation of cervical nuclei. ICPN was comprised of a sufficient aggregating path and a selecting path by using Interacting Convolutional Modules (ICM) and Internal Pyramid Resolution Complementing Modules (IPRCM) respectively. The proposed ICPN was evaluated on the Herlev dataset and achieved state-of-the-art performance with an average ZSI of 0.972.

To meet the needs of clinical application in practice, Zhao et al. (2022) proposed a lightweight feature attention network (LFANet) for abnormal cervical cell segmentation. Two plug-and-play modules, the lightweight feature extraction (LFE) module, and the feature layer attention (FLA) module were introduced to improve the feature extraction ability and reduce the computational consumption. The proposed LFANet achieved the best segmentation results on the Herlev dataset with a low computational complexity showing that LFANet was effective for splitting the nucleus and cytoplasm regions of cervical cells. Besides, the authors also carried out comparative experiments on three other medical image segmentation datasets to further verify the robustness of LFANet.

Luo et al. (2022) proposed a dual-supervised sampling network (DSSNet) to accelerate the speed of cervical nucleus segmentation. Via the supervised-down sampling module using the compressed images rather than raw images, the amount of the convolution computation was dramatically reduced. Besides, a boundary detection network was exploited to supervise the up-sampling process of the decoding layer to ensure segmentation accuracy. The proposed DSSNet achieved the same level of accuracy as U-Net while speeding up 5 times.

Li et al. (2022) proposed a global dependency and local attention (GDLA) module to improve the capability of contextual information modeling and feature refinement for the classical U-Net network. GDLA consisted of three parallel components: a channel attention module based on Squeeze-and-Excitation operations (Hu et al. 2018), a spatial attention module accomplished by 1×1 convolutional layer and sigmoid function, and a global dependency module implemented by the simplified Non-local Block (Cao et al. 2019).

On the Herlev dataset, the proposed method achieved ZSI of 0.913 for nuclei and 0.796 for cytoplasm, which gained performance improvement of 0.028 and 0.057 from the original U-Net.

4.5.2 Segmentation of overlapping cells

Segmentation of overlapping cervical cells refers to the process of separating individual cells that are overlapped in a cervical cytological image. Early systems focus on segmenting the nucleus and cytoplasm of isolated cells, which is not entirely practical. In clinical practice, the overlap of cervical cells is a very common phenomenon. The large degree of overlap and poor cytoplasmic boundary contrast increase the complexity of the cell segmentation task, which may lead to incorrect diagnoses. In recent years, with the successful organization of the Overlapping Cervical Cytology Image Segmentation Challenge in ISBI 2014 and 2015, an increasing number of works have paid attention to this topic. To address this issue, researchers tend to adopt multi-stage approaches in which a coarse segmentation of cell elements is performed first and then the extraction and refinement of overlapping regions followed.

To segment individual cells from overlapping clumps in Pap smear images, Song et al. (2016) proposed a novel framework based on a multi-scale CNN and deformation model. The overall segmentation framework consisted of the following three parts: cell component segmentation part to classify the region of nuclei, cytoplasm or background; multiple cells labeling part for splitting of the detected overlapping cytoplasm; and cell boundary refinement and inference part to achieve accurate segmentation results. They evaluated the proposed method with two different datasets, the ISBI 2015 Challenge Dataset and the Shenzhen University (SZU) Dataset. The experimental results demonstrated that the proposed method outperformed state-of-the-art methods and achieved the highest dice coefficient (DSC) value on both two datasets.

Tareef et al. (2017) proposed a variational segmentation framework for cervical cells using super pixel-wise CNN and dynamic shape modeling. The cellular components were first classified into background, nuclei, and cytoplasm based on a CNN model. Then, individual cytoplasm was separated from overlapping cellular mass using Voronoi segmentation and learned shape prior-based evolution. On both versions of the ISBI 2014 datasets (preliminary version and final challenge version), the proposed framework achieved the highest segmentation performance.

Xu et al. (2018) presented a novel method for automated segmentation of overlapping cervical cells using a light CNN model and fast multi-cell labeling. They first leveraged a light CNN model which is composed of a convolutional layer a pooling layer and a fully connected layer, to discriminate nuclei part as accurate initialization. Then, for the segmentation of overlapping cytoplasm, they utilized the simple linear iterative clustering (SLIC) method to generate a superpixel map and devised a fast multi-cell labeling method to roughly split clumped cytoplasm. Finally, the cell boundary was refined by an improved distance regularized level set method. The proposed method was validated on three datasets including the ISBI 2014 dataset, the ISBI 2015 dataset, and an in-house dataset. The experimental results showed the effectiveness of the proposed method for the segmentation of overlapping cervical cells.

Wan et al. (2019) presented a unique DCNN-based framework to automatically segment overlapping cervical cells. The workflow of the proposed method included cell detection, cytoplasm segmentation, and boundary refinement. TernausNet model and the
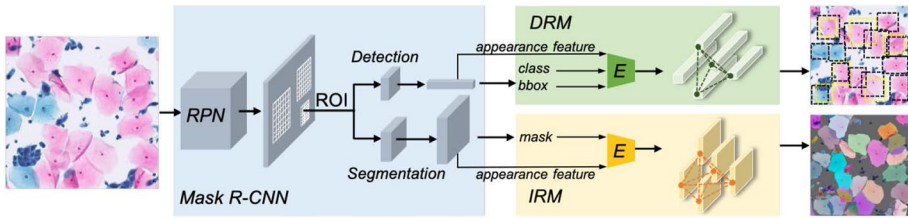
Fig. 14 Overview of IRNet (Zhou et al. 2019)

double-window based cell localization method were first utilized to extract the individual cells for cell detection. Then, a modified DeepLab V2 model was constructed to segment the cytoplasm. To refine the cell outer contours, fully connected conditional random fields (CRFs) and distance regularized level set evolution (DRLSE) served as post-processing methods. Three datasets including one in-house dataset and public datasets, ISBI 2014 and ISBI 2015, were served to evaluate the proposed method. The developed DCNN method achieved DSCs of 0.93, 0.92, and 0.92 on ISBI 2014, ISBI 2015, and the in-house dataset, respectively. The high-performance segmentation results showed the effectiveness and potential of the proposed method to be applied for automatic cervical cancer diagnosis.

Zhou et al. (2019) proposed an Instance Relation Network (IRNet) to segment overlapping cervical cells which explored instance relation interaction, as illustrated in Fig. 14. Based on Mask R-CNN, IRNet introduced Instance Relation Module (IRM) and Duplicate Removal Module (DRM) to improve the network's ability for cell-instance segmentation. IRM could make good use of contextual information and enhance semantic consistency. DRM benefited candidate selection which calibrated the misalignment between classification score and localization accuracy. A large cervical Pap smear (CPS) dataset was built to validate the performance of IRNet and the experimental results demonstrated the effectiveness of IRNet for overlapping cervical cell segmentation.

Zhang et al. (2020) proposed a polar coordinate sampling-based approach for overlapping cervical cell segmentation using Attention U-Net and graph-based Random Walk (RW). Attention U-Net was utilized to separate nuclei from the cellular clumps and graph-based RW was exploited to extract the cytoplasm. On the ISBI 2014 dataset, the proposed approach gained DSC scores of 0.93 and 0.917 for the nucleus and cytoplasm, respectively. The experimental results demonstrated that the proposed approach was effective and reliable for segmenting overlapping cervical cells.

To address the problem of limited data for cervical cell segmentation since the instance segmentation task required voluminous pixel-level annotations, Zhou et al. (2020) proposed a novel semi-supervised method, Mask-guided Mean Teacher framework with Perturbation-sensitive Sample Mining (MMT-PSM), which utilized both labeled and unlabeled data for cervical cell segmentation. MMT-PSM consisted of a teacher network and a student network using the same backbone. The teacher's self-ensemble predictions from augmented samples were used to generate reliable pseudo-labels to supervise the student network. Moreover, mask-guided feature distillation was leveraged to reduce the interference of the background noise. Experiments demonstrated the proposed MMT-PSM outperformed other semi-supervised methods and significantly improved the segmentation accuracy.

Huang et al. (2021) proposed a two-stage framework based on Mask R-CNN for automated segmentation of overlapping cells. The first stage involved presenting candidate

bounding boxes for the cytoplasm, while the second stage employed pixel-to-pixel alignment to refine the boundary and category classification. On ISBI 2014 and 2015 datasets, the proposed method achieved DSC of 0.92 and 0.89, respectively.

Mahyari and Dansereau (2022) designed a three-phase scheme for the segmentation of overlapping cells. In the first phase, a self-created residual CNN model was used to generate probabilistic image maps for cell components. In the second phase, high-probability nuclei nodes were used as seeds for a multi-layer random walker image segmentation for nuclei-seeded region growing. In addition, a cytoplasm approximation could be acquired by thresholding the cytoplasm probabilistic output maps. In the last phase, the Hungarian algorithm was applied to refine the individual pixel locations for the final cell segmentation. On the extended ISBI 2014 dataset, the proposed three-phase method achieved the highest segmentation performance with a DSC of 0.97 over nine different segmentation techniques.

4.5.3 Method analysis and summary

In this section, several DL-based methods for cervical cell segmentation have been surveyed. Cervical cancer cell segmentation can be divided into two parts: one is overlapping and the other is non-overlapping. The segmentation of cytoplasm and nucleus is often performed for individual cells. FCN (Long et al. 2015), U-Net (Ronneberger et al. 2015), and Mask R-CNN (He et al. 2017) are three frequently used basic models. U-Net is one of the most successful segmentation models that has been widely applied in various medical specialties. U-Net adopts encoder-decoder architecture for end-to-end semantic segmentation. The encoder uses a series of convolutional layers to extract high-level features from the input image, while the decoder uses upsampling and concatenation operations to generate a pixel-level segmentation map. The architecture also includes skip connections to combine low-level and high-level features. For the segmentation of cytoplasm and nucleus for individual cells, multi-level feature fusion and attention mechanism are the two most effective ways to improve accuracy (Zhang et al. 2019; Yang et al. 2020; Zhao et al. 2022; Li et al. 2022). In addition, deformable convolution (Zhao et al. 2019) and conditional random field (Liu et al. 2018) are also good choices to enhance the segmentation performance.

For the segmentation of overlapping cells, shape priors-based methods are preferable (Tareef et al. 2017; Xu et al. 2018; Wan et al. 2019). In Wan et al. (2019), fully connected conditional random fields (CRFs) including the prior knowledge of relationships between pixels were used to refine the cytoplasm segmentation results. IRNet (Zhou et al. 2019) proposed an Instance Relation Module (IRM) and Duplicate Removal Module (DRM) to take advantage of relation information between overlapping cells.

4.6 Whole slide image analysis

Automated WSI analysis has been widely studied in digital histopathological images for cancer diagnosis since the histopathological examination is the most reliable diagnostic basis and the gold standard for clinical diagnosis of cancer (Tellez et al. 2019; Dimitriou et al. 2019). In general, automated WSI analysis is realized by multiple instance learning (MIL) (Carbonneau et al. 2018), in which each tissue specimen is represented as a bag of instances and each instance is a small image patch extracted from the WSI. MIL belongs to weakly-supervised learning and there is only the slide-level label for all patches in the
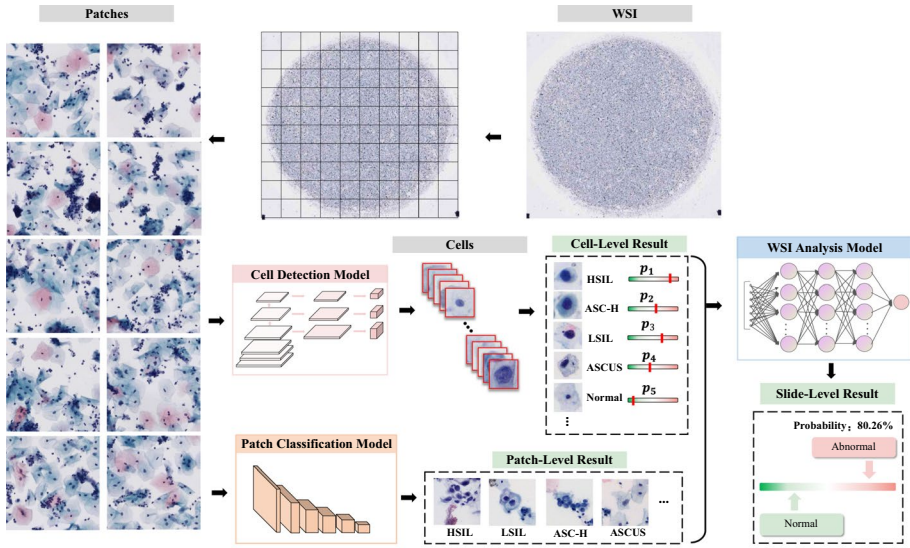
Fig. 15 The general process of cervical WSI analysis

same WSI. The core of MIL algorithms is to associate the slide-level label (e.g., normal specimen or cancerous specimen) with patch-level features. MIL-based WSI analysis has the potential to improve diagnostic accuracy and has been well studied in histopathology (Xiang et al. 2022; Zhang et al. 2022).

However, it is still arduous work to perform WSI analysis in cytopathology since the lesion area is continuous in histopathological WSI, and even the presence of a single isolated diseased cell may lead to an abnormal sample in cytopathological WSI. Thus, it is important to leverage an object detection algorithm to search abnormal cells and collect cell-level features. Both cell-level features and patch-level features are crucial for the final slide-level prediction, as shown in Fig. 15. It was not until 2021 that automated cervical cytology screening entered the thorough WIS analysis stage with the presence of the first DL-based WSI analysis methods in cervical cytology screening (Lin et al. 2021). In the past two years, several DL-based WSI analysis methods for cervical cytology successively emerged (Table 6).

4.6.1 Reference review

In Chen et al. (2021), an automatic WSI diagnosis was proposed using unit stochastic selection and attention fusion. Chen et al. first constructed a unit-level CNN based on VGG16b and ResNet50 to extract features of each unit (patch or cell). Next, they leveraged a UOI selection method to select the representative features of the WSI and employed an attention module to fuse all units' features for WSI diagnosis. The authors evaluated the proposed framework on three different types of pathological images. For the diagnosis of cervical cytological WSIs, the proposed method achieved good performance with a mean AUC of 0.851.

Lin et al. (2021) presented the first work for the specific analysis of cervical whole slide images. Firstly, an efficient deep learning-based dual-path network (DP-Net) was

Table 6 Summary of deep learning-based studies for cervical WSI analysis. Accuracy (Acc), Precision (Pre), Recall (Rec), Specificity (Spec), Sensitivity (Sens), Area Under Curve (AUC), F1-score (F1)

Reference	Method	Dataset	Result
Chen et al. (2021)	Unit-level CNN (VGG16, ResNet50) + attention module	Private dataset (264 positive slides and 108 negative slides in total)	AUC = 0.851
Lin et al. (2021)	DP-Net + SGL + RRS	Private dataset (19,303 WSIs in 6 classes from 4 centers)	Sens = 0.907, Spec = 0.80, AUC = 0.925
Zhou et al. (2021)	Hierarchical pathology screening (RetinaNet + Patch Encoder Module + SVM)	Private dataset (237 WSIs for traing and 361 WSIs for testing)	Acc = 0.905, F1 = 0.867, Sens = 0.891
Zhu et al. (2021)	AIATBS (YOLOv3 + Xception + Patch-based models + U-Net + XGBoost + logical decision tree)	Private dataset (81,727 WSIs retrospective samples for training and 34,403 prospective clinical samples for clinical validation)	Sens = 0.9474
Cao et al. (2021)	AttFPN + ResNet50	Private dataset (325 cases)	AUC =0.934, Sens = 0.913, Spec = 0.906, and Acc = 0.909
Cheng et al. (2021)	LR and HR model (ResNet50) + RNN	Private dataset (3,545 WSIs with 79,911 annotations for training and validation, 1,170 WSIs for testing)	Spec = 0.935, Sens = 0.951
Pirovano et al. (2021)	Tile classification and localizaiton + Slide level aggregation	Private dataset (40 slides)	Acc = 0.775, Spec = 0.83
Kanavati et al. (2022)	EfficientNet + RNN	Private dataset (1,503 WSIs for training, 150 for validation, and 1,468 WSIs of different settings for testing)	Full agreement set: AUC = 0.960, Acc = 0.907, Sens = 0.850, Spec = 0.911; Equal Balance-rev set: AUC = 0.915, Acc = 0.885, Sens = 0.839, Spec = 0.920; Clinical Balance-rev set: AUC = 0.890, Acc = 0.903, Sens = 0.886, Spec = 0.904
Geng et al. (2022)	FCOS + ResNet34	Private dataset (2,625 WSIs)	2-category task: Sens = 0.9784, Spec=0.8550, Pre = 82.55; 5-category task: Acc = 79.74
Zhang et al. (2022)	RetinaNet + SE-ResNeXt-50 + graph attention network + supervised contrastive learning	Private dataset (3,485 negative WSIs and 3,462 positive WSIs)	Acc = 0.8579, AUC = 0.9252, Rec = 0.8263, Pre = 0.8815, F1 = 0.8528

designed for lesion detection. Inspired by medical domain knowledge that different pre-cancerous cervical cells belonged to different groups (epidermal group and basal group), a synergistic grouping loss (SGL) was proposed for fine-grained cell classification. Then, a slide-level classifier called rule-based risk stratification (RRS) was introduced to perform the final WSI diagnosis, which simulated the clinical diagnostic criteria of cytopathologists. To evaluate the proposed method, a large number of samples were collected from multiple medical centers to construct the cervical WSI dataset (19,303 WSIs). The proposed method achieved a high sensitivity of 0.907 and a specificity of 0.80, showing strong robustness for practical cervical cytology screening (Fig. 16).

Zhou et al. (2021) proposed a hierarchical framework for case-level automatic diagnosis of cervical smears, which consisted of three stages. In the first stage, a large number of cytological images were extracted from the scanned WSI, and cell-level detection was performed for each image using RetinaNet. In the second stage, top-k regions with the highest confidence were selected and fed into the subsequent Patch Encoder Module (PEM) for image-level classification. In the last stage, the confidence scores of all images in each case were collected and used as the feature vectors to train an SVM classifier for final case-level diagnosis. Experiments showed that the proposed framework presented better accuracy than applying object detection and classification networks directly.
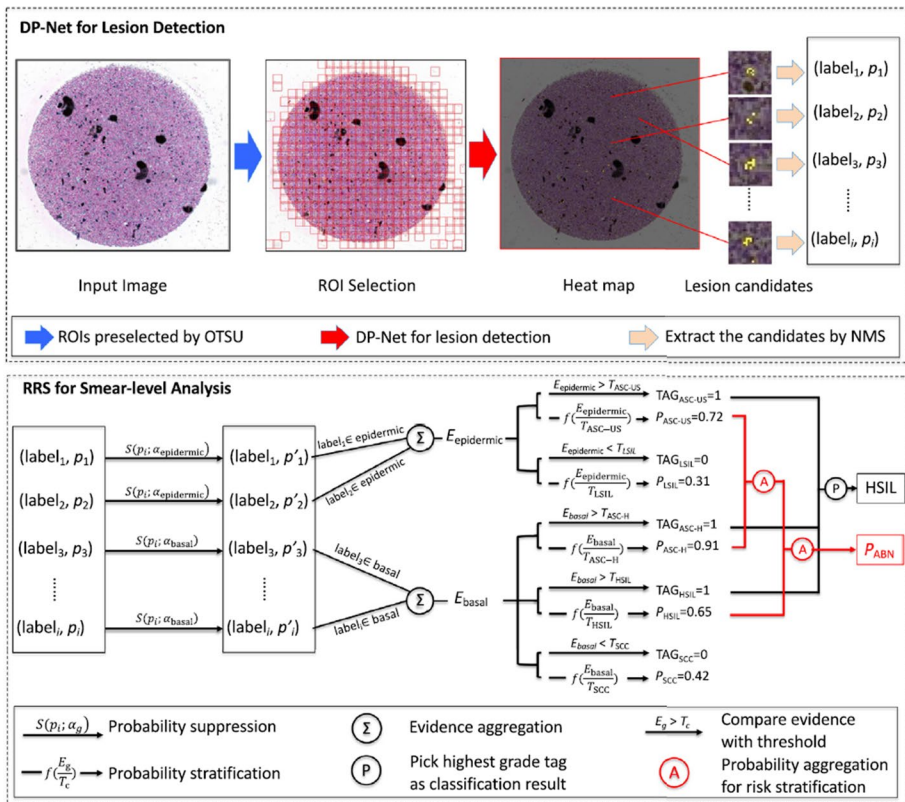

Fig. 16 First WSI analysis framework in cervical cytology screening, DP-Net with synergistic grouping loss and rule-based risk stratification (Lin et al. 2021)

Zhu et al. (2021) developed an AI-aided diagnostic system for automated cervical cytology screening, called AIATBS, which could help cytologists interpret in strict accordance with TBS standards. This system integrated five AI models including YOLOv3 for object detection, Xception for further fine-grained classification, DenseNet-50 for patch-based classification, U-Net for nucleus segmentation, and XGBoost model together with the logical decision tree for final slide-level diagnostic decisions. This paper also presented a digital pathology image quality control (DPIQC) system to ensure the quality of digitized images. AIATBS system was validated at 11 medical centers, and the outstanding performance demonstrated its adoption applicability and robustness for routine assistive diagnostic screening which could reduce the workload of cytologists, and improve the accuracy of cervical cancer screening.

Cao et al. (2021) devised a three-phase framework for automatic cervical cytology screening. Firstly, they proposed a novel attention feature pyramid network (AttFPN) to automatically detect abnormal cervical cells. AttFPN leveraged both channel and spatial attention for multi-scale feature fusion to improve the accuracy of abnormal cervical cells at different scales. Then, the image-level classification results were obtained by using the ResNet50 according to the corresponding probability prediction of detected abnormal cervical cells. At last, The classification results of all image patches in the same WSI were summarized to determine the ultimate case-level result. Extensive experiments demonstrated that AttFPN was effective for abnormal cell detection and the whole system had the potential for routine cervical cancer screening programs.

In Cheng et al. (2021), the authors proposed a robust WSI analysis method for cervical cancer screening by imitating the diagnosis process of cytopathologists, in which suspicious cells were found at low magnification and then scrutinized for confirmation at high magnification. They utilized a low-resolution model cascaded with a high-resolution model to recommend the 10 most suspicious lesion cells in each WSI. Then, an RNN-based WSI classification model was constructed by integrating the extracted feature representations of the top 10 lesion cells. The proposed system achieved 93.5% specificity and 95.1% sensitivity on multi-center WSI datasets with 1170 samples.

Pirovano et al. (2021) devised an explainable region classifier in cervical cytological WSIs. A created dataset and a novel loss were proposed to train an efficient region classifier to perform weakly supervised localization for malignancy regions in WSIs. Besides, they extended their approach to a more general detection task for cell abnormality and a real clinical slide dataset. The results demonstrated its effectiveness and potential to be applied in the current workflow of cytopathologists.

Kanavati et al. (2022) developed a DL-based method for WSI analysis of LBC specimens. They utilized a CNN model together with an RNN model to realize the slide-level classification. The EfficientNetB0 (Tan and Le 2019) model was employed to extract features of all tiles in one WSI. The output of the CNN model was adjusted as the input of the RNN model, which then gave a final WSI diagnosis. On 1468 collected test WSIs, the proposed method achieved AUCs in the range of 0.89–0.96, which fully demonstrated its effectiveness for cervical WSI diagnosis.

Geng et al. (2022) developed a two-stage learning framework for analyzing gigapixel cervical WSIs including a patch-level feature learning module and a WSI-level feature learning module. Patch-level model leveraged a one-stage object detector FCOS (Tian et al. 2019) and the WSI-level feature learning module utilized a modified ResNet34. The proposed approach achieved state-of-the-art classification performance on both 2-class and 5-class tasks.

Zhang et al. (2022) developed a deep learning-based framework for cervical cancer screening which explored the relationships between the suspicious cells and took advantage of other cells for comparison. This system was comprised of a ranking and feature extractor based on RetinaNet (Lin et al. 2017) and SE-ResNeXt-50 (Hu et al. 2018) model, and a graph attention network (GAT) to model the intrinsic relationships between different patches. They also proposed a supervised contrastive learning strategy to enhance the feature learning capacity for better classification. Extensive experiments validated the effectiveness of the proposed GAT and contrastive learning strategy, which outperformed other prevalent WSI classification approaches.

4.6.2 Method analysis and summary

In this section, we have reviewed automatic analysis methods for cervical Whole slide images based on deep learning. At present, the WSI analysis methods for cervical cytology can be divided into three main categories: (1) cell-level detection + WSI diagnosis (Chen et al. 2021; Pirovano et al. 2021; Zhang et al. 2022; 2) patch-level classification or feature + WSI diagnosis (Lin et al. 2021; Cheng et al. 2021; Kanavati et al. 2022; Geng et al. 2022; 3) cell-level detection + patch-level classification + WSI diagnosis (Zhou et al. 2021; Zhu et al. 2021; Cao et al. 2021). In the first scheme, an object detection algorithm is first employed to find the abnormal cells. Then, the slide-level diagnosis is performed via the information aggregation (feature or predicted probability) from the detected cells. With regard to the second type, since WSI is composed of many patches, these patches can be predicted as probability or encoded into features via a deep neural network and then fused for slide-level diagnosis. As for the last one, the cell-level information isn't directly used for the final diagnosis but is integrated to represent the image patch. The WSI diagnosis is then carried out by leveraging all patch information. In Lin et al. (2021), the authors proposed a Dual-path network with synergistic grouping loss for patch-level classification according to the origin area of cells. They further employed a rule-based risk stratification approach, inspired by clinical practice, to combine the prediction results of each patch and make a final diagnosis on the WSI.

Unlike general WSI classification methods for pathology (Li et al. 2022), most WSI analysis methods for cervical cytology tend to focus on identifying abnormal cells or areas first, rather than directly utilizing multiple instance learning (MIL) methods (Kemp et al. 1997; Hallinan 2005; Moshavegh et al. 2012). However, several studies have shown that normal cells in positive samples may also exhibit subtle changes, known as malignancy associated changes (MACs). Therefore, the presence of MACs enables the use of MIL methods for whole-slide analysis in cervical cytology, without necessarily identifying abnormal cells first. In the future, more MIL-based weakly supervised learning methods are worth exploring.

5 Challenges and opportunities

Despite significant progress in automated cervical cytology screening in recent years, there are still considerable challenges and opening issues that need to be resolved. Furthermore, the development of DL technology and computational cytology is accelerating

the advancement of this field. This section further discusses the prospects and potential research directions in automated cervical cytology screening.

Stain Normalization Due to the variations in staining procedures, staining durations, imaging environments, and scanning instruments, there always exists diverse image styles of the collected cytological images. Such image style inconsistency makes it difficult to build robust and generalized DL-based models of cervical cytology since training data and testing data may have different image styles causing the low performance of trained models in actual deployment. Stain normalization is an ideal way to eliminate the differences in image style. Traditional stain normalization methods such as color transfer, stain spectral matching, color deconvolution, etc. need one or several template images to estimate stain parameters, but a few template images cannot represent the color distribution of the entire reference dataset. Therefore, DL-based stain normalization methods using generative adversarial networks (GANs) are a better substitute because the whole dataset of the target style is leveraged as the template to execute color normalization by image-to-image translation. For example, Chen et al. (2021) proposed a two-stage domain adversarial style normalization framework for cervical cytopathological images and Kang et al. (2021) presented StainNet by using StainGAN (Shaban et al. 2019) and distillation learning to complete the stain normalization of cervical cell images.

Image Super-Resolution Image super-resolution is another promising research direction for cervical cytology images. Out-of-focus and low-resolution images will interfere with the precise diagnosis in cervical cytology screening. However, in real-world screening programs, blurred field of view (FoV) caused by scanning too fast without proper focusing is a frequent occurrence in scanned images. In addition, the acquisition of high-resolution digital slides needs advanced scanners which increases the financial burden in remote and underdeveloped regions. To address this problem, single-image super-resolution (SISR) brings an effective solution by converting low-resolution slides into high-resolution slides. Two DL-based SISR methods, PathSRGAN (Ma et al. 2020) and STSRNet (Ma et al. 2021) have been proposed for cervical cytopathological images. Both stain normalization and image super-resolution are urgently needed image preprocessing tools to assist DL-based diagnoses to improve inter-laboratory comparability and facilitate the development of CAD systems in cervical cytology screening.

Overlapping Cells For objective and quantitative analysis of cervical cytological images, precise segmentation of each individual is the most crucial step. However, earlier attempts focused on segmenting cellular nuclei of isolated cells, which is not completely realistic, as cell clumps along with the translucent overlapping cytoplasm are always present in clinical practice. Thus, the segmentation of overlapping cells became an urgent problem to be solved for cervical cytology. The first work concerning the segmentation of overlapping cell nuclei in cervical cytology images was proposed by Bengtsson et al. based on the smoothed difference code (Bengtsson et al. 1981). In recent years, with the organization of ISBI challenges and the release of public datasets in 2014 and 2015, more people are paying attention to this issue (Tareef et al. 2017, 2018). However, most conventional methods rely on the precise nuclei detection results for further cytoplasm division and refinement, which can be easily interfered with by blood, mucus, and other miscellaneous conditions in practice. More recently, deep learning-based methods can obtain better segmentation results with the support of big data (Zhang et al. 2020; Song et al. 2016; Tareef et al. 2017). Whereas the pixel-level prediction of each cytology image requires a longer computational time, especially when dealing with the WSI owning tens of thousands of cells. Thus, more efficient and lightweight DL-based methods are major research trends in the future (Xu et al. 2018; Zhao et al. 2022; Luo et al. 2022).

Effective Feature Extractor Feature extractors are used to learn discriminative features of cytology images in computational cytology (Jiang et al. 2022). The feature representation capability of the feature extractor will greatly affect the downstream tasks (cervical cell identification, abnormal cell detection, and cell region segmentation). During the initial period of rapid development of deep learning, researchers aimed to enhance the feature extraction ability of deep neural networks by either increasing the depth or width of the network (Simonyan and Zisserman 2015; He et al. 2016; Szegedy et al. 2016; Xie et al. 2017). In the past few years, attention mechanism has been introduced into the field of computer vision and various visual attention module has been proposed (Hu et al. 2018; Woo et al. 2018; Wang et al. 2020). visual attention modules that make DL-based models focus on lesion-related parts while inhibiting irrelevant information have been widely employed in automated cervical cytology screening (Zhang et al. 2020; Chen et al. 2021; Cao et al. 2021). Most recently, with the successful practice of the Transformer (Vaswani et al. 2017) in multiple computer vision tasks (Dosovitskiy et al. 2021; Liu et al. 2021; Touvron et al. 2021), iVision Transformer (ViT) quickly spreads in various research fields. CVM-Cervix (Liu et al. 2022) demonstrates the superior performance of the ViT to serve as an effective feature extractor for cervical cell classification. More ViT-based approaches in automated cervical cytology screening are expected in the future.

Incorporating Medical Domain Knowledge Since experienced cytopathologists can often give fairly accurate diagnoses, it's not surprising that their knowledge may guide DL-based models to do their assigned tasks better. The specialized knowledge of cytopathologists for cervical cytology refers to the cytological characteristics they learned, the way they browse the slides, the features they pay special attention to, and the training process they experienced (Xie et al. 2021). A simple way to realize the incorporation of medical domain knowledge is to combine the hand-crafted features with DL models since the manual features contain cytological characteristics related to diagnosis which are definitely pointed out in guidelines and criteria of cervical cytology, as mentioned in Sect. 4.3.3. Besides, Lin et al. presented a synergistic grouping loss and a rule-based risk stratification system using the cell grouping rules in the TBS criterion (Lin et al. 2021). Cheng et al. devised a DL-based model mimicking the cytopathologists' habits of viewing specimens (Cheng et al. 2021). Cao et al. attention-guided network, AttFPN, to pay special attention to lesion-related areas (Cao et al. 2021). Moreover, Chen et al. (2022) built TDCC-Net by leveraging the diagnosis experience of cytopathologists that normal cells in the same image should be used as a reference for better identification of abnormal ones. All the above studies make good use of medical domain knowledge to guide the construction of the DL model so as to achieve excellent results. There is a wealth of untapped medical knowledge that could be leveraged to develop high-performance and interpretable DL models.

Malignancy Associated Changes (MACs) In the approaches described so far to the cervical cytology screening, these methods have been the so-called "rare event approach" to mimic the way cytologists do the screening by looking for potentially rare (pre-)cancerous cells showing clear signs of atypia or malignancy (Bengtsson and Malm 2014). However, if there are no abnormal cells in the area where a specimen is taken and prepared, a patient who may have (pre-)cancerous lesions will be diagnosed as normal. Thus, the theory of malignancy associated changes (MACs) was first proposed in the late 1950s to describe subtle morphological and physiologic changes that are found in normal cells of patients harboring malignant disease. Lots of researchers have studied MAC for cervical cancer by quantitative cytology (Kemp et al. 1997; Hallinan 2005; Moshavegh et al. 2012; Mehnert et al. 2014; Tang et al. 2015). Kemp et al. utilized feed-forward neural networks with 53 designed nuclear features to detect MAC cells which were normal intermediate

cells selected from severe dysplasia slides. The proposed method achieved an accuracy of 76.2% for slide-level classification, respectively (Kemp et al. 1997). Mehnert et al. presented a structural approach to quantitatively characterizing nuclear chromatin texture in Pap smears based on mean-shift and watershed transform methods. The proposed method successfully verified the existence of MACs in conventional Pap smears (Mehnert et al. 2014). Since MAC cells belong to normal cells, it is hard to analyze MAC from a single cell image. However, when it comes to slide-level diagnosis, the presence of a large number of MAC cells may present an implicit pattern that, if detected, could help better predict malignant diseases. Thus, this approach is particularly interesting in relation to WSI analysis systems which hasn't been explored in recent methods for slide-level cytology image analysis and may be an alternative approach to the "rare event" approach.

Annotation-Efficient Learning Unlike natural images, the annotation of medical images requires specialized medical knowledge. The extensive annotation work can be a heavy burden for cytologists, making it difficult to obtain a large-scale dataset of high quality in cervical cytology screening. To address limited and noisy labels, annotation-efficient learning has emerged which is generally accomplished by transfer learning, domain adaptation, weakly supervised learning(multiple instance learning), semi-supervised learning, and self-supervised learning (Cheplygina et al. 2019). Wang et al. proposed a novel annotation-efficient learning method for medical image segmentation based on noisy pseudo labels and adversarial learning (Wang et al. 2020). Hu et al. utilized semi-supervised contrastive Learning to segment MRI and CT images (Hu et al. 2021). For cervical cytology, several semi-supervised learning-based methods have also been proposed to detect abnormal cell detection or segment overlapping cells (Zhang et al. 2021; Du et al. 2021; Chai et al. 2022; Zhou et al. 2020). These approaches have successfully improved the labeling efficiency and exhibit high accuracies which are comparable with full-supervised methods. More methods deserve to be explored and studied in cervical cytology screening by using annotation-efficient learning.

Multi-modal Data Fusion With the recent advancements in multi-modal deep learning technologies, significant progress has been made in the field of cancer diagnosis and prognosis analysis (Cui et al. 2023; Arya and Saha 2020). Multi-modal data fusion aided decision is also a good choice to realize slide-level diagnosis in cervical cytology screening. In addition to cervical cytopathological images, clinical data such as electronic medical records (EMRs) is also a critical reference in the final slide-level diagnosis. EMR contains a great deal of helpful personal information (Age, duration of menstrual period, medical history, cytology screening record, etc.) that can be utilized to guide more accurate diagnosis. At present, there is no related work of multi-modal data-based diagnosis in cervical cytology but this is a potential task in the future. Combining natural language processing (NLP) and computer vision (CV) technology to extract image features and clinical text features, and building a multi-modal classification model to realize the interactive fusion of multi-source data are meaningful to realize the precise diagnoses and personalized recommendations for cervical cytology.

Internet of Medical Things (IoMT) IoMT is an emerging challenge of the conventional Internet of Things (IoT) which enables the connection of medical devices, software applications, and health systems that collect and exchange healthcare data (Kakhi et al. 2022). IoMT can provide significant benefits by increasing access to care, improving the quality of medical service, and reducing healthcare costs, particularly for patients in remote or underdeveloped regions. IoMT also enables the integration of deep learning algorithms into healthcare, allowing for real-time disease diagnosis and personalized treatment plans based on individual patient data. Liu et al. proposed a Dental IoMT system based on

intelligent hardware, deep learning, and mobile terminals, aiming to explore the feasibility of in-home dental health (Liu et al. 2019). Guo et al. proposed a hybrid intelligence-driven IoMT system to diagnose pathological myopia for remote patients by combining conventional machine learning with deep learning (Guo et al. 2021). In automated cervical cytology screening, there are also some works that designing smart scanners and IoMT systems to promote digital pathology in rural areas and remote hospitals (Huang et al. 2018; Tang et al. 2021; Holmström et al. 2021; Jiang et al. 2022, 2023). The design of a universal and efficient cytopathological IoMT system is the ultimate pursuit for automated cervical cytology screening and there is still a long way to go.

Federated Learning In order to successfully apply the DL-based model to the actual clinical screening programs, strong generalization is a guarantee. Currently, most DL methods can achieve considerable performance on their internal datasets whereas the results are less than satisfactory when applied to clinical environments (Zhou et al. 2021). With the improvement of medical services and the promotion of IoMT, there are increasing concerns about the security and privacy of healthcare data. The lack of data privacy has restricted data sharing among medical institutions and further affected the construction and verification of the DL-based model with superior generalization (Rajpurkar et al. 2022). Recently, federated learning (FL) has been proposed to address the above issue which allows multiple parties to collaborate on the training of a shared model without sharing their data (Rieke et al. 2020). IoMT and FL can work together to improve the accuracy of DL-based models and guarantee generalization while maintaining patient data privacy and security. For example, a COVID-19 IoMT System is proposed by using FL and blockchain (Samuel et al. 2022) and a novel skin disease detection system is presented with the integration of federated machine learning (Hossen et al. 2022). Overall, IoMT together with FL has the potential to revolutionize cervical cytology screening for cancer prevention and timely treatment.

6 Conclusion

In this survey, an overview of cervical cytology and its current screening procedures is first introduced. Then, we offer a comprehensive collection of public image datasets for cervical cytology. Next, the most relevant DL-based image analysis methods in automated cervical cytology screening have been analyzed. From these summarized approaches, different learning paradigms (transfer learning, ensemble learning, semi-supervised and weakly supervised learning) have been applied to multiple tasks (cell identification, abnormal cell or suspicious area detection, cell component or overlapping cell segmentation, and WSI diagnosis) in cervical cytology screening. The primary objective of this review is to aid the advancement of automated tools that can effectively facilitate cervical screening procedures. The primary objective of this survey is to aid the advancement of CAD tools that can effectively facilitate automated cervical cytology screening programs. Additionally, this work provides insights into potential directions for future research including data preprocessing, feature representation, model design, clinical application, and privacy security.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10462-023-10588-z.

References

Al-Janabi S, Huisman A, Van Diest PJ (2012) Digital pathology: current status and future perspectives. Histopathology 61(1):1–9

Araújo FH, Silva RR, Ushizima DM, Rezende MT, Carneiro CM, Bianchi AGC, Medeiros FN (2019) Deep learning for cell image segmentation and ranking. Comput Med Imaging Gr 72:13–21

Arya N, Saha S (2020) Multi-modal classification for human breast cancer prognosis prediction: proposal of deep-learning based stacked ensemble model. IEEE/ACM Trans Comput Biol Bioinform 19(2):1032–1041

Badrinarayanan V, Kendall A, Cipolla R (2017) Segnet: a deep convolutional encoder-decoder architecture for image segmentation. IEEE Trans Pattern Anal Mach Intell 39(12):2481–2495

Bai B, Yang X, Li Y, Zhang Y, Pillar N, Ozcan A (2023) Deep learning-enabled virtual histological staining of biological samples. Light: Sci Appl 12(1):57

Bengtsson E, Eriksson O, Holmquist J, Jarkrans T, Nordin B, Stenkvist B (1981) Segmentation of cervical cells: detection of overlapping cell nuclei. Comput Gr Image Process 16(4):382–394

Bengtsson E, Malm P (2014) Screening for cervical cancer using automated analysis of pap-smears. Comput Math Methods Med. https://doi.org/10.1155/2014/842037

Bhatt AR, Ganatra A, Kotecha K (2021) Cervical cancer detection in pap smear whole slide images using convnet with transfer learning and progressive resizing. Peer J Comput Sci 7:348

Biscotti CV, Dawson AE, Dziura B, Galup L, Darragh T, Rahemtulla A, Wills-Frank L (2005) Assisted primary screening using the automated thinprep imaging system. Am J Clin Pathol 123(2):281–287

Bora K, Chowdhury M, Mahanta LB, Kundu MK, Das AK (2016) Pap smear image classification using convolutional neural network. In: Proceedings of the Tenth Indian Conference on Computer Vision, Graphics and Image Processing, pp. 1–8

Boronow RC (1998) Death of the papanicolaou smear? a tale of three reasons. Am J Obstet Gynecol 179(2):391–396

Cai Z, Vasconcelos N (2018) Cascade r-cnn: Delving into high quality object detection. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 6154–6162

Cai Z, Vasconcelos N (2019) Cascade r-cnn: high quality object detection and instance segmentation. IEEE Trans Pattern Anal Mach Intell 43(5):1483–1498

Cao L, Yang J, Rong Z, Li L, Xia B, You C, Lou G, Jiang L, Du C, Meng H et al (2021) A novel attention-guided convolutional network for the detection of abnormal cervical cells in cervical cancer screening. Med Image Anal 73:102197

Cao Y, Geddes TA, Yang JYH, Yang P (2020) Ensemble deep learning in bioinformatics. Nat Mach Intell 2(9):500–508

Cao Y, Xu J, Lin S, Wei F, Hu H (2019) Gcnet: Non-local networks meet squeeze-excitation networks and beyond. In: Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops, pp. 1971–1980

Carbonneau M-A, Cheplygina V, Granger E, Gagnon G (2018) Multiple instance learning: a survey of problem characteristics and applications. Pattern Recognit 77:329–353

Chai Z, Luo L, Lin H, Chen H, Han A, Heng P-A (2022) Deep semi-supervised metric learning with dual alignment for cervical cancer cell detection. In: 2022 IEEE 19th International Symposium on Biomedical Imaging (ISBI), pp. 1–5. IEEE

Chantziantoniou N (2022) Bestcyte® cell sorter imaging system: Primary and adjudicative whole slide image rescreening review times of 500 thinprep pap test thin-layers-an intra-observer, time-surrogate analysis of diagnostic confidence potentialities. J Pathol Inform 13:100095

Chen H, Liu J, Wen Q-M, Zuo Z-Q, Liu J-S, Feng J, Pang B-C, Xiao D (2021) Cytobrain: cervical cancer screening system based on deep learning technology. J Comput Sci Technol 36(2):347–360

Chen L-C, Papandreou G, Kokkinos I, Murphy K, Yuille AL (2017) Deeplab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. IEEE Trans Pattern Anal Mach Intell 40(4):834–848

Chen P, Liang Y, Shi X, Yang L, Gader P (2021) Automatic whole slide pathology image diagnosis framework via unit stochastic selection and attention fusion. Neurocomputing 453:312–325

Chen T, Zheng W, Ying H, Tan X, Li K, Li X, Chen DZ, Wu J (2022) A task decomposing and cell comparing method for cervical lesion cell detection. IEEE Trans Med Imaging 41:2432–2442

Chen W, Gao L, Li X, Shen W (2022) Lightweight convolutional neural network with knowledge distillation for cervical cells classification. Biomed Signal Process Control 71:103177

Chen X, Yu J, Cheng S, Geng X, Liu S, Han W, Hu J, Chen L, Liu X, Zeng S (2021) An unsupervised style normalization method for cytopathology images. Comput Struct Biotechnol J 19:3852–3863

Chen Y-F, Huang P-C, Lin K-C, Lin H-H, Wang L-E, Cheng C-C, Chen T-P, Chan Y-K, Chiang JY (2013) Semi-automatic segmentation and classification of pap smear cells. IEEE J Biomed Health Inform 18(1):94–108

Cheng S, Liu S, Yu J, Rao G, Xiao Y, Han W, Zhu W, Lv X, Li N, Cai J et al (2021) Robust whole slide image analysis for cervical cancer screening using deep learning. Nat Commun 12(1):5639

Cheplygina V, de Bruijne M, Pluim JP (2019) Not-so-supervised: a survey of semi-supervised, multi-instance, and transfer learning in medical image analysis. Med Image Anal 54:280–296

Chitra B, Kumar S (2022) Recent advancement in cervical cancer diagnosis for automated screening: a detailed review. J Ambient Intell Humaniz Comput. https://doi.org/10.1007/s12652-021-02899-2

Cohen PA, Jhingran A, Oaknin A, Denny L (2019) Cervical cancer. The Lancet 393(10167):169–182

Conceição T, Braga C, Rosado L, Vasconcelos MJM (2019) A review of computational methods for cervical cells segmentation and abnormality classification. Int J Mol Sci 20(20):5114

Cramer DW (1974) The role of cervical cytology in the declining morbidity and mortality of cervical cancer. Cancer 34(6):2018–2027

Creswell A, White T, Dumoulin V, Arulkumaran K, Sengupta B, Bharath AA (2018) Generative adversarial networks: an overview. IEEE Sign Process Mag 35(1):53–65

Croitoru F-A, Hondru V, Ionescu RT, Shah M (2023) Diffusion models in vision: a survey. IEEE Trans Pattern Anal Mach Intell. https://doi.org/10.1109/TPAMI.2023.3261988

Crowell EF, Bazin C, Saunier F, Brixtel R, Caillot Y, Lesner B, Toutain M, Ferreri C, Garcia I, Mathieu M-C et al (2019) Cytoprocessortm: a new cervical cancer screening system for remote diagnosis. Acta Cytol 63(3):215–223

Crum CP, Mitao M, Levine RU, Silverstein S (1985) Cervical papillomaviruses segregate within morphologically distinct precancerous lesions. J Virol 54(3):675–681

Cui C, Yang H, Wang Y, Zhao S, Asad Z, Coburn LA, Wilson KT, Landman B, Huo Y (2023) Deep multimodal fusion of image and non-image data in disease diagnosis and prognosis: a review. Prog Biomed Eng. https://doi.org/10.1088/2516-1091/acc2fe

Dai J, Li Y, He K, Sun J (2016) R-fcn: Object detection via region-based fully convolutional networks. Advances in neural information processing systems **29**

Delga A, Goffin F, Kridelka F, Marée R, Lambert C, Delvenne P (2014) Evaluation of cellsolutions bestprep® automated thin-layer liquid-based cytology papanicolaou slide preparation and bestcyte® cell sorter imaging system. Acta Cytol 58(5):469–477

Dimitriou N, Arandjelović O, Caie PD (2019) Deep learning for whole slide image analysis: an overview. Front Med 6:264

Diniz DN, Rezende MT, Bianchi AGC, Carneiro CM, Luz EJS, Moreira GJP, Ushizima DM, de Medeiros FNS, Souza MJF (2021) A deep learning ensemble method to assist cytopathologists in pap test image classification. J Imaging 7(7):111

Dong N, Zhao L, Wu C-H, Chang J-F (2020) Inception v3 based cervical cell classification combined with artificially extracted features. Appl Soft Comput 93:106311

Dosovitskiy A, Beyer L, Kolesnikov A, Weissenborn D, Zhai X, Unterthiner T, Dehghani M, Minderer M, Heigold G, Gelly S, *et al.* (2021) An image is worth 16x16 words: Transformers for image recognition at scale. In: International Conference on Learning Representations (ICLR), pp. 1–21

Du X, Huo J, Qiao Y, Wang Q, Zhang L (2021) False positive suppression in cervical cell screening via attention-guided semi-supervised learning. In: Predictive Intelligence in Medicine: 4th International Workshop, PRIME 2021, Held in Conjunction with MICCAI 2021, Strasbourg, France, October 1, 2021, Proceedings 4, pp. 93–103 . Springer

Elsheikh TM, Austin RM, Chhieng DF, Miller FS, Moriarty AT, Renshaw AA (2013) American society of cytopathology workload recommendations for automated pap test screening: developed by the productivity and quality assurance in the era of automated screening task force. Diagn Cytopathol 41(2):174–178

Ganaie MA, Hu M, Malik A, Tanveer M, Suganthan P (2022) Ensemble deep learning: a review. Eng Appl Artif Intell 115:105151

Gao W, Xu C, Li G, Zhang Y, Bai N, Li M (2022) Cervical cell image classification-based knowledge distillation. Biomimetics 7(4):195

Gautam S, Bhavsar A, Sao AK, Harinarayan K (2018) Cnn based segmentation of nuclei in pap-smear images with selective pre-processing. In: Medical Imaging 2018: Digital Pathology, vol. 10581, pp. 246–254. SPIE

Geng R, Liu Q, Feng S, Liang Y (2022) Learning deep pathological features for wsi-level cervical cancer grading. In: ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 1391–1395. IEEE

Ghoneim A, Muhammad G, Hossain MS (2020) Cervical cancer classification using convolutional neural networks and extreme learning machines. Future Gener Comput Syst 102:643–649

Girshick R, Donahue J, Darrell T, Malik J (2014) Rich feature hierarchies for accurate object detection and semantic segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 580–587

Gu Z, Cheng J, Fu H, Zhou K, Hao H, Zhao Y, Zhang T, Gao S, Liu J (2019) Ce-net: context encoder network for 2d medical image segmentation. IEEE Trans Med Imaging 38(10):2281–2292

Guo M-H, Xu T-X, Liu J-J, Liu Z-N, Jiang P-T, Mu T-J, Zhang S-H, Martin RR, Cheng M-M, Hu S-M (2022) Attention mechanisms in computer vision: a survey. Comput Vis Media 8(3):331–368

Guo Z, Shen Y, Wan S, Shang W-L, Yu K (2021) Hybrid intelligence-driven medical image recognition for remote patient diagnosis in internet of medical things. IEEE J Biomed Health Inform 26(12):5817–5828

Hallinan J (2005) Evolving neural networks for the classification of malignancy associated changes. In: IDEAL, pp. 382–389 . Springer

He K, Gkioxari G, Dollár P, Girshick R (2017) Mask r-cnn. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 2961–2969

He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778

Herbert A, Bergeron C, Wiener H, Schenck U, Klinkhamer P, Bulten J, Arbyn M (2007) European guidelines for quality assurance in cervical cancer screening: recommendations for cervical cytology terminology. Cytopathology 18(4):213–219

Hesamian MH, Jia W, He X, Kennedy P (2019) Deep learning techniques for medical image segmentation: achievements and challenges. J Digit Imaging 32:582–596

Holmström O, Linder N, Kaingu H, Mbuuko N, Mbete J, Kinyua F, Törnquist S, Muinde M, Krogerus L, Lundin M et al (2021) Point-of-care digital cytology with artificial intelligence for cervical cancer screening in a resource-limited setting. JAMA Netw Open 4(3):211740–211740

Hossen MN, Panneerselvam V, Koundal D, Ahmed K, Bui FM, Ibrahim SM (2022) Federated machine learning for detection of skin diseases and enhancement of internet of medical things (iomt) security. IEEE J Biomed Health Inform 27:835–841

Hou X, Shen G, Zhou L, Li Y, Wang T, Ma X (2022) Artificial intelligence in cervical cancer screening and diagnosis. Front Oncol 12:851367

Howell LP, Davis RL, Belk TI, Agdigos R, Lowe J (1998) The autocyte preparation system for gynecologic cytology. Acta Cytol 42(1):171–177

Huang J, Yang G, Li B, He Y, Liang Y (2021) Segmentation of cervical cell images based on generative adversarial networks. IEEE Access 9:115415–115428

Huang Y-N, Peng X-C, Ma S, Yu H, Jin Y-B, Zheng J, Fu G-H et al (2018) Development of whole slide imaging on smartphones and evaluation with thinprep cytology test samples: Follow-up study. JMIR mHealth uHealth 6(4):9518

Hu J, Shen L, Sun G (2018) Squeeze-and-excitation networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 7132–7141

Hussain E, Mahanta LB, Borah H, Das CR (2020) Liquid based-cytology pap smear dataset for automated multi-class diagnosis of pre-cancerous and cervical cancer lesions. Data Br 30:105589

Hussain E, Mahanta LB, Das CR, Choudhury M, Chowdhury M (2020) A shape context fully convolutional neural network for segmentation and classification of cervical nuclei in pap smear images. Artif Intell Med 107:101897

Hutchinson ML, Cassin CM, Ball HG III (1991) The efficacy of an automated preparation device for cervical cytology. Am J Clin Pathol 96(3):300–305

Hu X, Zeng D, Xu X, Shi Y (2021) Semi-supervised contrastive learning for label-efficient medical image segmentation. In: Medical Image Computing and Computer Assisted Intervention–MICCAI 2021: 24th International Conference, Strasbourg, France, September 27–October 1, 2021, Proceedings, Part II 24, pp. 481–490. Springer

Hyeon J, Choi H-J, Lee BD, Lee KN (2017) Diagnosing cervical cell images using pre-trained convolutional neural network as feature extractor. In: 2017 IEEE International Conference on Big Data and Smart Computing (BigComp), pp. 390–393. IEEE

Hz Hausen (1977) Human papillomaviruses and their possible role in squamous cell carcinomas. Curr Top Microbiol Immunol 78:1–30

Ikenberg H, Lieder S, Ahr A, Wilhelm M, Schön C, Xhaja A (2023) Comparison of the hologic genius digital diagnostics system with the thinprep imaging system-a retrospective assessment. Cancer Cytopathol. https://doi.org/10.1002/cncy.22695

Jantzen J, Norup J, Dounias G, Bjerregaard B (2005) Pap-smear benchmark data for pattern classification. Nature inspired Smart Information Systems (NiSIS 2005), 1–9

Jia AD, Li BZ, Zhang CC (2020) Detection of cervical cancer cells based on strong feature cnn-svm network. Neurocomputing 411:112–127

Jia D, He Z, Zhang C, Yin W, Wu N, Li Z (2022) Detection of cervical cancer cells in complex situation based on improved yolov3 network. Multimed Tools Appl 81(6):8939–8961

Jia D, Zhou J, Zhang C (2022) Detection of cervical cells based on improved ssd network. Multimed Tools Appl 81(10):13371–13387

Jiang H, Zhou Y, Lin Y, Chan RC, Liu J, Chen H (2022) Deep learning for computational cytology: a survey. Med Image Anal 84:102691

Jiang P, Liu J, Chen H, Li C, Pang B, Cao D (2022) Channel spatial collaborative attention network for fine-grained classification of cervical cells. In: International Conference on Neural Information Processing, pp. 540–551. Springer

Jiang P, Liu J, Luo Q, Pang B, Xiao D, Cao D (2023) Development of automatic portable pathology scanner and its evaluation for clinical practice. J Dig Imaging 36:1110–1122

Jiang P, Liu J, Xiao D, Pang B, Hao Z, Cao D (2022) A novel iomt system for pathological diagnosis based on intelligent mobile scanner and whole slide image stitching method. In: Intelligent Computing Methodologies: 18th International Conference, ICIC 2022, Xi'an, China, August 7–11, 2022, Proceedings, Part III, pp. 463–472. Springer

Kakhi K, Alizadehsani R, Kabir HD, Khosravi A, Nahavandi S, Acharya UR (2022) The internet of medical things and artificial intelligence: trends, challenges, and opportunities. Biocybern Biomed Eng 42:749–771

Kanavati F, Hirose N, Ishii T, Fukuda A, Ichihara S, Tsuneki M (2022) A deep learning model for cervical cancer screening on liquid-based cytology specimens in whole slide images. Cancers 14(5):1159

Kang H, Luo D, Feng W, Zeng S, Quan T, Hu J, Liu X (2021) Stainnet: a fast and robust stain normalization network. Front Med 8:746307

Kardos TF (2004) The focalpoint system: focalpoint slide profiler and focalpoint gs. Cancer Cytopathol: Interdiscip Int J Am Cancer Soc 102(6):334–339

Kedra B, Chomczyk M, Zlotkowski M, Stokowska W, Borsuk A, Bicz M, Pietruska M, Tokajuk G, Charkiewicz R, Czajka P et al (2012) Cytological picture of the oral mucosa in patients with gastric and colon cancer. Folia Histochem et Cytobiol 50(3):375–380

Kemp RA, MacAulay C, Garner D, Palcic B (1997) Detection of malignancy associated changes in cervical cell nuclei using feed-forward neural networks. Anal Cell Pathol 14(1):31–40

Khamparia A, Gupta D, de Albuquerque VHC, Sangaiah AK, Jhaveri RH (2020) Internet of health things-driven deep learning system for detection and classification of cervical cells using transfer learning. J Supercomput 76(11):8590–8608

Khamparia A, Gupta D, Rodrigues JJ, de Albuquerque VHC (2021) Dcavn: cervical cancer prediction and classification using deep convolutional and variational autoencoder network. Multimed Tools Appl 80(20):30399–30415

Khan S, Naseer M, Hayat M, Zamir SW, Khan FS, Shah M (2022) Transformers in vision: a survey. ACM Comput Surv (CSUR) 54(10s):1–41

Koss LG, Lin E, Schreiber K, Elgert P, Mango L (1994) Evaluation of the papnet™cytologic screening system for quality control of cervical smears. Am J Clini Pathol 101(2):220–229

Krizhevsky A, Sutskever I, Hinton GE (2017) Imagenet classification with deep convolutional neural networks. Commun ACM 60(6):84–90

Kundu R, Chattopadhyay S (2022) Deep features selection through genetic algorithm for cervical pre-cancerous cell classification. Multimed Tools Appl 82:13431–13452

LeCun Y, Bengio Y, Hinton G (2015) Deep learning. Nature 521(7553):436–444

Lee JR, Kim S, Park I, Eo T, Hwang D (2021) Relevance-cam: Your model already knows where to look. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 14944–14953

Lew M, Wilbur DC, Pantanowitz L (2021) Computational cytology: lessons learned from pap test computer-assisted screening. Acta Cytol 65(4):286–300

Li G, Sun C, Xu C, Zheng Y, Wang K (2022) Cervical cell segmentation method based on global dependency and local attention. Appl Sci 12(15):7742

Li X, Li C, Rahaman MM, Sun H, Li X, Wu J, Yao Y, Grzegorzek M (2022) A comprehensive review of computer-aided whole-slide image analysis: from datasets to feature extraction, segmentation, classification and detection approaches. Artif Intell Rev 55(6):4809–4878

Li X, Xiong H, Li X, Wu X, Zhang X, Liu J, Bian J, Dou D (2022) Interpretable deep learning: Interpretation, interpretability, trustworthiness, and beyond. Knowl Inform Syst 64(12):3197–3234

Li X, Xu Z, Shen X, Zhou Y, Xiao B, Li T-Q (2021) Detection of cervical cancer cells in whole slide images using deformable and global context aware faster rcnn-fpn. Curr Oncol 28(5):3585–3601

Li Y, Su Y, Guo M, Han X, Liu J, Vishwasrao HD, Li X, Christensen R, Sengupta T, Moyle MW et al (2022) Incorporating the image formation process into deep learning improves network performance. Nat Methods 19(11):1427–1437

Liang Y, Pan C, Sun W, Liu Q, Du Y (2021) Global context-aware cervical cell detection with soft scale anchor matching. Comput Methods Prog Biomed 204:106061

Liang Y, Tang Z, Yan M, Chen J, Liu Q, Xiang Y (2021) Comparison detector for cervical cell/clumps detection in the limited data scenario. Neurocomputing 437:195–205

Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JP, Clarke M, Devereaux PJ, Kleijnen J, Moher D (2009) The prisma statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. Ann Intern Med 151(4):65

Lin H, Chen H, Wang X, Wang Q, Wang L, Heng P-A (2021) Dual-path network with synergistic grouping loss and evidence driven risk stratification for whole slide cervical image analysis. Med Image Anal 69:101955

Lin T-Y, Dollár P, Girshick R, He K, Hariharan B, Belongie S (2017) Feature pyramid networks for object detection. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2117–2125

Lin T-Y, Goyal P, Girshick R, He K, Dollár P (2017) Focal loss for dense object detection. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 2980–2988

Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, Van Der Laak JA, Van Ginneken B, Sánchez CI (2017) A survey on deep learning in medical image analysis. Med Image Anal 42:60–88

Liu G, Ding Q, Luo H, Sha M, Li X, Ju M (2022) Cx22: a new publicly available dataset for deep learning-based segmentation of cervical cytology images. Comput Biol Med 150:106194

Liu J, Fan H, Wang Q, Li W, Tang Y, Wang D, Zhou M, Chen L (2022) Local label point correction for edge detection of overlapping cervical cells. Front Neuroinform 16:895290

Liu L, Wang Y, Wu D, Zhai Y, Tan L, Xiao J (2018) Multitask learning for pathomorphology recognition of squamous intraepithelial lesion in thinprep cytologic test. In: Proceedings of the 2nd International Symposium on Image Computing and Digital Medicine, pp. 73–77

Liu L, Xu J, Huan Y, Zou Z, Yeh S-C, Zheng L-R (2019) A smart dental health-iot platform based on intelligent hardware, deep learning, and mobile terminal. IEEE J Biomed Health Inform 24(3):898–906

Liu M, Li X, Gao X, Chen J, Shen L, Wu H (2022) Sample hardness based gradient loss for long-tailed cervical cell detection. In: Medical Image Computing and Computer Assisted Intervention–MICCAI 2022: 25th International Conference, Singapore, September 18–22, 2022, Proceedings, Part II, pp. 109–119. Springer

Liu S, Wang Y, Yang X, Lei B, Liu L, Li SX, Ni D, Wang T (2019) Deep learning in medical ultrasound analysis: a review. Engineering 5(2):261–275

Liu W, Anguelov D, Erhan D, Szegedy C, Reed S, Fu C-Y, Berg AC (2016) Ssd: Single shot multibox detector. In: Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I 14, pp. 21–37 . Springer

Liu W, Li C, Xu N, Jiang T, Rahaman MM, Sun H, Wu X, Hu W, Chen H, Sun C et al (2022) Cvm-cervix: a hybrid cervical pap-smear image classification framework using cnn, visual transformer and multilayer perceptron. Pattern Recognit 130:108829

Liu Y, Zhang P, Song Q, Li A, Zhang P, Gui Z (2018) Automatic segmentation of cervical nuclei based on deep learning and a conditional random field. IEEE Access 6:53709–53721

Liu Z, Lin Y, Cao Y, Hu H, Wei Y, Zhang Z, Lin S, Guo B (2021) Swin transformer: Hierarchical vision transformer using shifted windows. In: Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 10012–10022

Long J, Shelhamer E, Darrell T (2015) Fully convolutional networks for semantic segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 3431–3440

Lu Z, Carneiro G, Bradley AP (2015) An improved joint optimization of multiple level set functions for the segmentation of overlapping cervical cells. IEEE Trans Image Process 24(4):1261–1272

Lu Z, Carneiro G, Bradley AP, Ushizima D, Nosrati MS, Bianchi AG, Carneiro CM, Hamarneh G (2016) Evaluation of three algorithms for the segmentation of overlapping cervical cells. IEEE J Biomed Health Inform 21(2):441–450

Luo D, Kang H, Long J, Zhang J, Chen L, Quan T, Liu X (2022) Dual supervised sampling networks for real-time segmentation of cervical cell nucleus. Comput Struct Biotechnol J 20:4360–4368

Ma D, Liu J, Li J, Zhou Y (2020) Cervical cancer detection in cervical smear images using deep pyramid inference with refinement and spatial-aware booster. IET Image Process 14(17):4717–4725

Ma J, Liu S, Cheng S, Chen R, Liu X, Chen L, Zeng S (2021) Stsrnet: self-texture transfer super-resolution and refocusing network. IEEE Trans Med Imaging 41(2):383–393

Ma J, Yu J, Liu S, Chen L, Li X, Feng J, Chen Z, Zeng S, Liu X, Cheng S (2020) Pathsrgan: multi-supervised super-resolution for cytopathological images using generative adversarial network. IEEE Trans Med Imaging 39(9):2920–2930

Madhukar RK, Joshi RC, Dutta MK (2022) A robust deep learning and feature fusion-based multi-class classification of cervical cells. In: 2022 International Conference on Edge Computing and Applications (ICE-CAA), pp. 1269–1275. IEEE

Mahyari TL, Dansereau RM (2022) Multi-layer random walker image segmentation for overlapped cervical cells using probabilistic deep learning methods. IET Image Process 16(11):2959–2972

Manna A, Kundu R, Kaplun D, Sinitca A, Sarkar R (2021) A fuzzy rank-based ensemble of cnn models for classification of cervical cytology. Sci Rep 11(1):1–18

Marinakis Y, Dounias G, Jantzen J (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):69–78

Martinez-Mas J, Bueno-Crespo A, Martinez-Espana R, Remezal-Solano M, Ortiz-Gonzalez A, Ortiz-Reina S, Martinez-Cendan J-P (2020) Classifying papanicolaou cervical smears through a cell merger approach by deep learning technique. Expert Syst Appl 160:113707

Mehnert A, Moshavegh R, Sujathan K, Malm P, Bengtsson E (2014) A structural texture approach for characterising malignancy associated changes in pap smears based on mean-shift and the watershed transform. In: 2014 22nd International Conference on Pattern Recognition, pp. 1189–1193. IEEE

Minaee S, Boykov Y, Porikli F, Plaza A, Kehtarnavaz N, Terzopoulos D (2021) Image segmentation using deep learning: a survey. IEEE Trans Pattern Anal Mach Intell 44(7):3523–3542

Mitra P, Murthy C, Pal SK (2002) Unsupervised feature selection using feature similarity. IEEE Trans Pattern Anal Mach Intell 24(3):301–312

Moshavegh R, Bejnordi BE, Mehnert A, Sujathan K, Malm P, Bengtsson E (2012) Automated segmentation of free-lying cell nuclei in pap smears for malignancy-associated change analysis. In: 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 5372–5375 . IEEE

Nambu Y, Mariya T, Shinkai S, Umemoto M, Asanuma H, Sato I, Hirohashi Y, Torigoe T, Fujino Y, Saito T (2022) A screening assistance system for cervical cytology of squamous cell atypia based on a two-step combined cnn algorithm with label smoothing. Cancer Med 11(2):520–529

Nayar R, Wilbur DC (2015) The Bethesda System for Reporting Cervical Cytology: Definitions, Criteria, and Explanatory Notes. Springer, New York

Nayar R, Wilbur DC (2017) The bethesda system for reporting cervical cytology: a historical perspective. Acta Cytol 61(4–5):359–372

Nguyen LD, Lin D, Lin Z, Cao J (2018) Deep cnns for microscopic image classification by exploiting transfer learning and feature concatenation. In: 2018 IEEE International Symposium on Circuits and Systems (ISCAS), pp. 1–5. IEEE

Niazi MKK, Parwani AV, Gurcan MN (2019) Digital pathology and artificial intelligence. The Lancet Oncol 20(5):253–261

Niu Z, Zhong G, Yu H (2021) A review on the attention mechanism of deep learning. Neurocomputing 452:48–62

Pan SJ, Yang Q (2009) A survey on transfer learning. IEEE Trans knowl Data Eng 22(10):1345–1359

Papanicolaou GN (1928) New cancer diagnosis. In: Proceedings of the 3rd Race Betterment Conference, pp. 528–534

Papanicolaou GN, Traut HF (1943) vol 46. New York

Patten SF Jr, Lee JS, Nelson AC (1996) Neopath, inc. neopath autopap 300 automatic pap screener system. Acta Cytol 40(1):45–52

Phoulady HA, Mouton PR (2018) A new cervical cytology dataset for nucleus detection and image classification (cervix93) and methods for cervical nucleus detection. arXiv preprint arXiv:1811.09651

Pirovano A, Almeida LG, Ladjal S, Bloch I, Berlemont S (2021) Computer-aided diagnosis tool for cervical cancer screening with weakly supervised localization and detection of abnormalities using adaptable and explainable classifier. Med Image Anal 73:102167

Plissiti ME, Dimitrakopoulos P, Sfikas G, Nikou C, Krikoni O, Charchanti A (2018) Sipakmed: A new dataset for feature and image based classification of normal and pathological cervical cells in pap smear images. In: 2018 25th IEEE International Conference on Image Processing (ICIP), pp. 3144–3148. IEEE

Qin J, He Y, Ge J, Liang Y (2022) A multi-task feature fusion model for cervical cell classification. IEEE J Biomed Health Inform 26(9):4668–4678

Rahaman MM, Li C, Wu X, Yao Y, Hu Z, Jiang T, Li X, Qi S (2020) A survey for cervical cytopathology image analysis using deep learning. IEEE Access 8:61687–61710

Rahaman MM, Li C, Yao Y, Kulwa F, Wu X, Li X, Wang Q (2021) Deepcervix: a deep learning-based framework for the classification of cervical cells using hybrid deep feature fusion techniques. Comput Biol Med 136:104649

Rajpurkar P, Chen E, Banerjee O, Topol EJ (2022) Ai in health and medicine. Nat Med 28(1):31–38

Reagan JW, Seidemann IL, Saracusa Y (1953) The cellular morphology of carcinoma in situ and dysplasia or atypical hyperplasia of the uterine cervix. Cancer 6(2):224–235

Redmon J, Divvala S, Girshick R, Farhadi A (2016) You only look once: Unified, real-time object detection. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 779–788

Ren S, He K, Girshick R, Sun J (2015) Faster r-cnn: Towards real-time object detection with region proposal networks. Advances in neural information processing systems **28**

Rezende MT, Silva R, Bernardo FdO, Tobias AH, Oliveira PH, Machado TM, Costa CS, Medeiros FN, Ushizima DM, Carneiro CM et al (2021) Cric searchable image database as a public platform for conventional pap smear cytology data. Sci Data 8(1):151

Riana D, Hadianti S, Rahayu S, Hasan M, Karimah IN, Pratama R (2021) Repomedunm: A new dataset for feature extraction and training of deep learning network for classification of pap smear images. In: Neural Information Processing: 28th International Conference, ICONIP 2021, Sanur, Bali, Indonesia, December 8–12, 2021, Proceedings, Part V, pp. 317–325. Springer

Richart RM (1967) Natural history of cervical intraepithelial neoplasia. Clin Obstet Gynecol 10(4):748–784

Rieke N, Hancox J, Li W, Milletari F, Roth HR, Albarqouni S, Bakas S, Galtier MN, Landman BA, Maier-Hein K et al (2020) The future of digital health with federated learning. NPJ Digit Med 3(1):119

Ronneberger O, Fischer P, Brox T (2015) U-net: Convolutional networks for biomedical image segmentation. In: Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III pp. 234–241. Springer

Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma S, Huang Z, Karpathy A, Khosla A, Bernstein M et al (2015) Imagenet large scale visual recognition challenge. Int J Comput Vis 115:211–252

Samuel O, Omojo AB, Onuja AM, Sunday Y, Tiwari P, Gupta D, Hafeez G, Yahaya AS, Fatoba OJ, Shamshirband S (2022) Iomt: a covid-19 healthcare system driven by federated learning and blockchain. IEEE J Biomed Health Inform 27(2):823–834

Sankaranarayanan R, Budukh AM, Rajkumar R (2001) Effective screening programmes for cervical cancer in low-and middle-income developing countries. Bull World Health Org 79(10):954–962

Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D (2017) Grad-cam: Visual explanations from deep networks via gradient-based localization. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 618–626

Shaban MT, Baur C, Navab N, Albarqouni S (2019) Staingan: Stain style transfer for digital histological images. In: 2019 Ieee 16th International Symposium on Biomedical Imaging (Isbi 2019), pp. 953–956. IEEE

Shanthi P, Faruqi F, Hareesha K, Kudva R (2019) Deep convolution neural network for malignancy detection and classification in microscopic uterine cervix cell images. Asian Pac J Cancer Prev: APJCP 20(11):3447

Shanthi P, Hareesha K, Kudva R (2022) Automated detection and classification of cervical cancer using pap smear microscopic images: a comprehensive review and future perspectives. Eng Sci 19:20–41

Siebers AG, Klinkhamer PJ, Grefte JM, Massuger LF, Vedder JE, Beijers-Broos A, Bulten J, Arbyn M (2009) Comparison of liquid-based cytology with conventional cytology for detection of cervical cancer precursors: a randomized controlled trial. Jama 302(16):1757–1764

Simonyan K, Zisserman A (2015) Very deep convolutional networks for large-scale image recognition. In: International Conference on Learning Representations (ICLR), pp. 1–14

Sompawong N, Mopan J, Pooprasert P, Himakhun W, Suwannarurk K, Ngamvirojcharoen J, Vachiramon T, Tantibundhit C (2019) Automated pap smear cervical cancer screening using deep learning. In: 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 7044–7048. IEEE

Song Y, Tan E-L, Jiang X, Cheng J-Z, Ni D, Chen S, Lei B, Wang T (2016) Accurate cervical cell segmentation from overlapping clumps in pap smear images. IEEE Trans Med Imaging 36(1):288–300

St Clair C, Wright J (2009) Cervical intraepithelial neoplasia: history and detection. Glob Libr Women's Med 10:3843

Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, Bray F (2021) Global cancer statistics 2020: globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: Cancer J Clin 71(3):209–249

Sun K, Xiao B, Liu D, Wang J (2019) Deep high-resolution representation learning for human pose estimation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 5693–5703

Su X, Shi J, Peng Y, Zheng L (2021) Cervical cell image classification based on multiple attention fusion. In: 2021 14th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), pp. 1–5. IEEE

Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z (2016) Rethinking the inception architecture for computer vision. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2818–2826

Tang JR, Mat Isa NA, Ch'ng ES (2015) A fuzzy-c-means-clustering approach: Quantifying chromatin pattern of non-neoplastic cervical squamous cells. PloS one 10(11):0142830

Tang H-P, Cai D, Kong Y-Q, Ye H, Ma Z-X, Lv H-S, Tuo L-R, Pan Q-J, Liu Z-H, Han X (2021) Cervical cytology screening facilitated by an artificial intelligence microscope: a preliminary study. Cancer Cytopathol 129(9):693–700

Tan M, Le Q (2019) Efficientnet: Rethinking model scaling for convolutional neural networks. In: International Conference on Machine Learning, pp. 6105–6114. PMLR

Tareef A, Song Y, Huang H, Wang Y, Feng D, Chen M, Cai W (2017) Optimizing the cervix cytological examination based on deep learning and dynamic shape modeling. Neurocomputing 248:28–40

Tareef A, Song Y, Cai W, Huang H, Chang H, Wang Y, Fulham M, Feng D, Chen M (2017) Automatic segmentation of overlapping cervical smear cells based on local distinctive features and guided shape deformation. Neurocomputing 221:94–107

Tareef A, Song Y, Huang H, Feng D, Chen M, Wang Y, Cai W (2018) Multi-pass fast watershed for accurate segmentation of overlapping cervical cells. IEEE Trans Med Imaging 37(9):2044–2059

Tellez D, Litjens G, van der Laak J, Ciompi F (2019) Neural image compression for gigapixel histopathology image analysis. IEEE Trans Pattern Anal Mach Intell 43(2):567–578

Tian Z, Shen C, Chen H, He T (2019) Fcos: Fully convolutional one-stage object detection. In: Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 9627–9636

Tolles WE, Bostrom R (1956) Automatic screening of cytological smears for cancer: the instrumentation. Ann N.Y. Acad Sci 63(6):1211–1218

Touvron H, Cord M, Douze M, Massa F, Sablayrolles A, Jégou H (2021) Training data-efficient image transformers & distillation through attention. In: International Conference on Machine Learning, pp. 10347–10357. PMLR

Traut HF, Papanicolaou GN (1943) Cancer of the uterus: the vaginal smear in its diagnosis. Calif West Med 59(2):121

Uijlings JR, Van De Sande KE, Gevers T, Smeulders AW (2013) Selective search for object recognition. Int J Comput Vis 104:154–171

Van Engelen JE, Hoos HH (2020) A survey on semi-supervised learning. Mach Learn 109(2):373–440

Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, Kaiser Ł, Polosukhin I (2017) Attention is all you need. Advances in neural information processing systems **30**

Waggoner SE (2003) Cervical cancer. The Lancet 361(9376):2217–2225

Wan T, Xu S, Sang C, Jin Y, Qin Z (2019) Accurate segmentation of overlapping cells in cervical cytology with deep convolutional neural networks. Neurocomputing 365:157–170

Wang H, Jiang C, Bao K, Xu C (2019) Recognition and clinical diagnosis of cervical cancer cells based on our improved lightweight deep network for pathological image. J Med Syst 43:1–9

Wang H, Wang Z, Du M, Yang F, Zhang Z, Ding S, Mardziel P, Hu X (2020) Score-cam: Score-weighted visual explanations for convolutional neural networks. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, pp. 24–25

Wang L, Guo D, Wang G, Zhang S (2020) Annotation-efficient learning for medical image segmentation based on noisy pseudo labels and adversarial learning. IEEE Trans Med Imaging 40(10):2795–2807

Wang P, Wang J, Li Y, Li L, Zhang H (2020) Adaptive pruning of transfer learned deep convolutional neural network for classification of cervical pap smear images. IEEE Access 8:50674–50683

Wang Q, Wu B, Zhu P, Li P, Zuo W, Hu Q (2020) Eca-net: Efficient channel attention for deep convolutional neural networks. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)

Wang W, Tian Y, Xu Y, Zhang X-X, Li Y-S, Zhao S-F, Bai Y-H (2022) 3cde-net: a cervical cancer cell detection network based on an improved backbone network and multiscale feature fusion. BMC Med Imaging 22(1):1–13

Watanabe S, Group C et al (1974) An automated apparatus for cancer prescreening: cybest. Computer Graphics and Image Processing 3(4):350–358

Wied GL, Bahr G, Bibbo M, Puls JH, Taylor J Jr, Bartels PH (1975) The ticas-rtcip real time cell identification processor. Acta Cytol 19(3):286–288

William W, Ware A, Basaza-Ejiri AH, Obungoloch J (2018) A review of image analysis and machine learning techniques for automated cervical cancer screening from pap-smear images. Comput Methods Progr Biomed 164:15–22

Woo S, Park J, Lee J-Y, Kweon IS (2018) Cbam: Convolutional block attention module. In: Proceedings of the European Conference on Computer Vision (ECCV), pp. 3–19

Workshop NCI (1989) The 1988 bethesda system for reporting cervical/vaginal cytological diagnoses. JAMA 262(7):931–934

Wu X, Sahoo D, Hoi SC (2020) Recent advances in deep learning for object detection. Neurocomputing 396:39–64

Xiang T, Song Y, Zhang C, Liu D, Chen M, Zhang F, Huang H, O'Donnell L, Cai W (2022) Dsnet: a dual-stream framework for weakly-supervised gigapixel pathology image analysis. IEEE Trans Med Imaging 41(8):2180–2190

Xiang Y, Sun W, Pan C, Yan M, Yin Z, Liang Y (2020) A novel automation-assisted cervical cancer reading method based on convolutional neural network. Biocybern Biomed Eng 40(2):611–623

Xie S, Girshick R, Dollár P, Tu Z, He K (2017) Aggregated residual transformations for deep neural networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1492–1500

Xie X, Niu J, Liu X, Chen Z, Tang S, Yu S (2021) A survey on incorporating domain knowledge into deep learning for medical image analysis. Med Image Anal 69:101985

Xu C, Li M, Li G, Zhang Y, Sun C, Bai N (2022) Cervical cell/clumps detection in cytology images using transfer learning. Diagnostics 12(10):2477

Xu S, Sang C, Jin Y, Wan T (2018) Robust segmentation of overlapping cells in cervical cytology using light convolution neural network. In: Neural Information Processing: 25th International Conference, ICONIP 2018, Siem Reap, Cambodia, December 13–16, 2018, Proceedings, Part VII 25, pp. 387–397. Springer

Yaman O, Tuncer T (2022) Exemplar pyramid deep feature extraction based cervical cancer image classification model using pap-smear images. Biomed Signal Process Control 73:103428

Yang X, Wu J, Yin Y (2020) Interacting convolution with pyramid structure network for automated segmentation of cervical nuclei in pap smear images. In: 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI), pp. 499–502. IEEE

Yang Y, Lv H, Chen N (2022) A survey on ensemble learning under the era of deep learning. Artif Intell Rev 56:5545–5589

Yan X, Zhang Z (2021) Hsdet: A representative sampling based object detector in cervical cancer cell images. In: Bio-Inspired Computing: Theories and Applications: 15th International Conference, BIC-TA 2020, Qingdao, China, October 23-25, 2020, Revised Selected Papers 15, pp. 406–418. Springer

Yi L, Lei Y, Fan Z, Zhou Y, Chen D, Liu R (2020) Automatic detection of cervical cells using dense-cascade r-cnn. In: Pattern Recognition and Computer Vision: Third Chinese Conference, PRCV 2020, Nanjing, China, October 16–18, 2020, Proceedings, Part II 3, pp. 602–613. Springer

Young R (2014) Who classification of tumours of female reproductive organs. Kurman RJ Carcangiu ML Herrington CS Young RH Monodermal teratomas and somatic-type tumours arising from a dermoid cyst, 63–66

Yu J, Wang G, Cheng S (2022) Cervical cell classification based on attention mechanism and multi-resolution feature fusion. In: 2022 3rd International Conference on Information Science, Parallel and Distributed Systems (ISPDS), pp. 50–53. IEEE

Yu S, Zhang S, Wang B, Dun H, Xu L, Huang X, Shi E, Feng X (2021) Generative adversarial network based data augmentation to improve cervical cell classification model. Math Biosci Eng 18:1740–1752

Yu X, Wang J, Hong Q-Q, Teku R, Wang S-H, Zhang Y-D (2022) Transfer learning for medical images analyses: a survey. Neurocomputing 489:230–254

Zhang C, Jia D, Li Z, Wu N (2022) Auxiliary classification of cervical cells based on multi-domain hybrid deep learning framework. Biomed Signal Process Control 77:103739

Zhang C, Liu D, Wang L, Li Y, Chen X, Luo R, Che S, Liang H, Li Y, Liu S *et al.* (2019) Dccl: A benchmark for cervical cytology analysis. In: Machine Learning in Medical Imaging: 10th International Workshop, MLMI 2019, Held in Conjunction with MICCAI 2019, Shenzhen, China, October 13, 2019, Proceedings 10, pp. 63–72 . Springer

Zhang H, Meng Y, Zhao Y, Qiao Y, Yang X, Coupland SE, Zheng Y (2022) Dtfd-mil: Double-tier feature distillation multiple instance learning for histopathology whole slide image classification. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 18802–18812

Zhang H, Zhu H, Ling X (2020) Polar coordinate sampling-based segmentation of overlapping cervical cells using attention u-net and random walk. Neurocomputing 383:212–223

Zhang J, He J, Chen T, Liu Z, Chen D (2019) Abnormal region detection in cervical smear images based on fully convolutional network. IET Image Process 13(4):583–590

Zhang J, Liu Z, Du B, He J, Li G, Chen D (2019) Binary tree-like network with two-path fusion attention feature for cervical cell nucleus segmentation. Comput Biol Med 108:223–233

Zhang L, Lu L, Nogues I, Summers RM, Liu S, Yao J (2017) Deeppap: deep convolutional networks for cervical cell classification. IEEE J Biomed Health Inform 21(6):1633–1643

Zhang L, Sonka M, Lu L, Summers RM, Yao J (2017) Combining fully convolutional networks and graph-based approach for automated segmentation of cervical cell nuclei. In: 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017), pp. 406–409. IEEE

Zhang M, Li X, Shen L (2021) Classification and localization consistency regularized student-teacher network for semi-supervised cervical cell detection. In: 2021 IEEE 34th International Symposium on Computer-Based Medical Systems (CBMS), pp. 289–294 . IEEE

Zhang S, Wen L, Bian X, Lei Z, Li SZ (2018) Single-shot refinement neural network for object detection. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 4203–4212

Zhang X, Cao M, Wang S, Sun J, Fan X, Wang Q, Zhang L (2022) Whole slide cervical cancer screening using graph attention network and supervised contrastive learning. In: Medical Image Computing and Computer Assisted Intervention–MICCAI 2022: 25th International Conference, Singapore, September 18–22, 2022, Proceedings, Part II, pp. 202–211. Springer

Zhang X, Wan F, Liu C, Ji R, Ye Q (2019) Freeanchor: Learning to match anchors for visual object detection. Advances in neural information processing systems **32**

Zhao J, Dai L, Zhang M, Yu F, Li M, Li H, Wang W, Zhang L (2020) Pgu-net+: progressive growing of u-net+ for automated cervical nuclei segmentation. In: Multiscale Multimodal Medical Imaging: First International Workshop, MMMI 2019, Held in Conjunction with MICCAI 2019, Shenzhen, China, October 13, 2019, Proceedings 1, pp. 51–58. Springer

Zhao J, Li Q, Li X, Li H, Zhang L (2019) Automated segmentation of cervical nuclei in pap smear images using deformable multi-path ensemble model. In: 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019), pp. 1514–1518. IEEE

Zhao Y, Fu C, Xu S, Cao L, Ma H-f (2022) Lfanet: lightweight feature attention network for abnormal cell segmentation in cervical cytology images. Comput Biol Med 145:105500

Zhao Z-Q, Zheng P, Xu S-T, Wu X (2019) Object detection with deep learning: a review. IEEE Trans Neural Netw Learn Syst 30(11):3212–3232

Zhou B, Khosla A, Lapedriza A, Oliva A, Torralba A (2016) Learning deep features for discriminative localization. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2921–2929

Zhou M, Zhang L, Du X, Ouyang X, Zhang X, Shen Q, Luo D, Fan X, Wang Q (2021) Hierarchical pathology screening for cervical abnormality. Comput Med Imaging Gr 89:101892

Zhou SK, Greenspan H, Davatzikos C, Duncan JS, Van Ginneken B, Madabhushi A, Prince JL, Rueckert D, Summers RM (2021) A review of deep learning in medical imaging: imaging traits, technology trends, case studies with progress highlights, and future promises. Proc IEEE 109(5):820–838

Zhou Y, Chen H, Lin H, Heng P-A (2020) Deep semi-supervised knowledge distillation for overlapping cervical cell instance segmentation. In: Medical Image Computing and Computer Assisted Intervention–MICCAI 2020: 23rd International Conference, Lima, Peru, October 4–8, 2020, Proceedings, Part I 23, pp. 521–531. Springer

Zhou Y, Chen H, Xu J, Dou Q, Heng P-A (2019) Irnet: Instance relation network for overlapping cervical cell segmentation. In: Medical Image Computing and Computer Assisted Intervention–MICCAI 2019: 22nd

International Conference, Shenzhen, China, October 13–17, 2019, Proceedings, Part I 22, pp. 640–648. Springer

Zhu X, Li X, Ong K, Zhang W, Li W, Li L, Young D, Su Y, Shang B, Peng L et al (2021) Hybrid ai-assistive diagnostic model permits rapid tbs classification of cervical liquid-based thin-layer cell smears. Nat Commun 12(1):3541

Zitnick CL, Dollár P (2014) Edge boxes: Locating object proposals from edges. In: Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part V 13, pp. 391–405 . Springer

Zou Z, Chen K, Shi Z, Guo Y, Ye J (2023) Object detection in 20 years: A survey. Proceedings of the IEEE