Multimedia Tools and Applications
https://doi.org/10.1007/s11042-024-19799-0

Review of cervical cell segmentation

Qian Huang[1] · Wei Zhang[1,2] · Yulin Chen[1] · Junzhou Chen[1] · Zheng Yang[2]

Received: 12 January 2024 / Revised: 27 May 2024 / Accepted: 28 June 2024

Abstract
Cervical cell segmentation is a significant task in medical image analysis and can be used for screening various cervical diseases. In recent years, substantial progress has been made in cervical cell segmentation techniques, leading to notable improvements in the performance of cervical cancer auxiliary diagnostic systems. This review summarizes and analyzes the recent research on cervical cell segmentation. The main contents include an introduction to cervical cell segmentation datasets, commonly used evaluation metrics, and various segmentation methods. Currently, mainstream segmentation methods can be classified into two categories: traditional and deep learning-based. Building upon this foundation, we unfold according to the context of segmentation objectives, evaluating the performance of each method in achieving specific segmentation objectives and exploring the relationships among different methods. Through this review, other researchers can clearly understand the development of cervical cell segmentation technology and future trends, and explore new methods and technologies based on integrating and sorting out existing technologies, so as to help cervical cancer auxiliary diagnosis systems achieve more accurate cell image segmentation.

Keywords Cervical cell segmentation · Auxiliary diagnosis · Deep learning · Nucleus · Cytoplasm

Wei Zhang, Yulin Chen, Junzhou Chen, and Zheng Yang contributed equally to this work.

✉ Qian Huang
huangqian@hhu.edu.cn

Wei Zhang
2920810788@qq.com

Yulin Chen
chenyulina@hhu.edu.cn

Junzhou Chen
junzhouchen7@gmail.com

Zheng Yang
1528960648@qq.com

[1] College of Computer Science and Software Engineering, Hohai University, Nanjing, China

[2] Nanjing Huiying Electronic Technology Corporation, Nanjing, China

Published online: 02 August 2024

1 Introduction

Cervical cancer ranks as the fourth most malignant tumor threatening women's lives, following breast, colorectal, and lung cancer [1]. Clinical research data indicate that the incidence of cervical cancer in China has been on the rise annually, with high mortality rates. In recent years, the probability of cervical cancer occurring in young women has been increasing yearly due to changes in people's lifestyles and behaviors. Since the time interval from initial inflammation to the development of cervical cancer is usually long, typically ranging from 8 to 10 years, regular screening has become the most effective preventive measure against cervical cancer. If during the precancerous lesion period, patients can find cancer cells or precancerous cells through cervical cytology screening and get timely diagnosis and treatment, the lesion can be prevented from developing into life-threatening invasive cancer [2, 3].

Currently, cervical cancer screening techniques can be categorized as single or combined. Single screening methods include cytology detection, iodine staining, acetic acid staining, and HPV testing. Among them, commonly used cytology detection methods are the Papanicolaou (Pap) smear and ThinPrep cytologic test (TCT), both of which are time-consuming, laborious, prone to errors, and can result in diagnostic discrepancies between pathologists [4, 5]. Low-cost and high-efficiency Computer-Aided Diagnosis (CAD) systems have emerged to address these issues. The typical process of computer-aided diagnosis for cervical cell images consists of four steps: image acquisition, preprocessing, segmentation, and feature extraction and classification. Among these, cervical cell image segmentation is one of the most critical technologies in CAD system research. It is the foundation for content analysis and understanding of cervical cell images and provides the necessary groundwork for subsequent feature extraction and classification [6]. The quality of image segmentation fundamentally determines the accuracy of the final diagnosis. Furthermore, precise cell image segmentation assists healthcare professionals and researchers in gaining a deeper understanding of cellular structures and tissue morphology. Automated and accurate image segmentation techniques can improve medical workflow efficiency, reduce human errors, and provide more reliable data to support medical research, thus paving the way for novel possibilities in future medical diagnostics and treatment methods.

Cervical cell image segmentation techniques utilize information such as intensity, color, shape, boundaries, and regional textures to divide the images into cell clusters, nucleus, and cytoplasm regions. Therefore, research in the field of cervical cell image segmentation can be conducted based on different segmentation targets, such as the detection of cancer cells or precancerous cells, including cell clusters, nucleus, and cytoplasm, involving the segmentation of different targets, such as single cells and overlapping cells. During the segmentation process, various challenges are encountered, such as utilizing multi-modal information, distinguishing boundaries and quantities of overlapping cells, segmenting complex, diverse-shaped, and irregularly bounded cell clusters, and eliminating the influence of image noise, blurriness, uneven illumination, and other interfering factors. These challenges make the task of cervical cell segmentation highly complex and challenging. To address these challenges, researchers have proposed various cervical cell segmentation methods, which can be broadly categorized into traditional methods and deep learning methods. Traditional methods include threshold segmentation [7], region-based segmentation [8], clustering methods [9] and so on. Deep learning methods refer to the use of deep neural networks such as FCN [10], U-Net [11], Mask R-CNN [12], and other network models to learn and extract effective features from original images for the segmentation of different regions or objects within the images.

In recent years, cervical cell segmentation techniques have been advancing continuously. Researchers seeking to engage in this field and conduct more in-depth research and analysis therein must understand these technologies. Some scholars have undertaken comprehensive reviews of these techniques. For instance, Sarwar et al. [13] reviewed the literature on cervical cell segmentation spanning 40 years from 1977, organized chronologically. However, their review lacks clarity regarding recent technological advancements and in-depth analysis and comparison of different methods. Most existing reviews on cervical cell segmentation integrate detection and classification aspects, resulting in incomplete coverage of segmentation techniques. Additionally, no published reviews have focused explicitly on cervical cell segmentation techniques of the last decade, thereby lacking a comprehensive summary and analysis of both traditional and deep learning methods. A comprehensive review of these techniques is crucial for driving advancements in this field and is an important means to continuously pursue technological progress. Therefore, conducting a systematic, comprehensive, and critical review of cervical cell segmentation techniques in recent years is highly necessary.

The relevant features of the cervical nucleus or cytoplasm are important indicators for determining whether cells are diseased. Cervical cell nucleus segmentation can help locate and extract morphological features of the nucleus, while cytoplasm segmentation can facilitate automated cell counting and measurement of cell morphological features. Accurate segmentation of the cytoplasm region enables the calculation of quantitative evaluation indicators such as cell diameter and nucleus-to-cytoplasm ratio. Whether it is the segmentation of the cervical nucleus, cytoplasm, or both, it can help identify potential abnormal cells and assist computer-aided systems in subsequent analysis and diagnosis. Therefore, this review is organized according to the segmentation targets: nucleus, cytoplasm, and multiple targets. This helps readers understand the segmentation requirements, algorithm design, and evaluation metrics differences for different targets, thereby gaining a more comprehensive understanding of the entire field of cervical cell segmentation. In this review, a comparative analysis of various methods within the same target is conducted based on the two major categories of traditional and deep learning methods. This reveals the current research focus and outstanding issues and helps the academic community quickly grasp the development process, current status, and trends in this field. Furthermore, it provides references and guidance for further research and improvement by integrating and organizing existing technologies.

2 Datasets and evaluation metrics

Cervical cell datasets and evaluation indicators are important for studying segmentation techniques. Therefore, this chapter summarizes the commonly used data sets and evaluation indicators.

2.1 Public datasets

There are currently many cervical cell datasets available for study, but few datasets are suitable for segmentation. In this paper, four datasets of cervical cells published for multi-task segmentation were extracted from selected papers: HEMLBC [14], ISBI 2014 [15], ISBI 2015 [16], and Herlev [17]. Except for HEMLBC, which consists of liquid-based H & E slides, the rest of the datasets use pasteurized staining techniques. Most papers used the Herlev, ISBI 2014, and 2015 datasets. Table 1 shows the characteristics of the dataset.

Table 1 Introduction to datasets

Dataset name	Sample size	Data source	Dyeing techniques	Dataset characteristics
Herlev [14]	917	Herlev University Hospital	Pap staining	It is originally a classification dataset, and the provided masks can be used as segmentation masks for different segmentation targets.
HEMLBC [17]	917	National Cancer Institute (NCI)	H & E staining	The nucleus is prominent, and the shape is mostly like a solid circle, providing manual annotation.
ISBI 2014 [15]	16 real EDF 945synthesis	Overlapping Cervical Cytology Image Segmentation Challenge	Pap staining	The image contains only one cell clump and the overlapping areas are less contrasting.
ISBI 2015 [16]	17 real EDF	Overlapping Cervical Cytology Image Segmentation Challenge	Pap staining	There are multiple cell clumps in the image, and the number of cells in the cell clump and the overlap rate between adjacent cells vary.

Note: This table shows the characteristics of the dataset,including the name of the dataset, the sample size, the staining technique, and the description of the corresponding characteristics.Except for HEMLBC, which consists of liquid-based H & E slides, the rest of the datasets use pasteurized staining techniques

Herlev dataset The dataset consists of 917 single-cell color cervical cell images at various resolutions, each with corresponding pathologist-annotated hand-labeled images. In this dataset, there are seven types of cells present, with four corresponding to abnormal cells and the remaining three to normal cells. Additionally, each image in the dataset is accompanied by a corresponding mask. An example of this dataset section is shown (see Fig. 1).

ISBI 2014 dataset The dataset consists of a total of 16 Extended Depth of Field (EDF) images of real cervical cytology and 945 composite images. There were 135 synthetic images and 8 real cervical cytology EDF images in the training set and 810 synthetic and 8 real cervical cytology EDF images in the test set. The composite image is generated by making small transformations to the background and brightness of different isolated cells in a real EDF image that is published without annotating the individual cytoplasm. The partial plot of the dataset is shown (see Fig. 2).

ISBI 2015 dataset The dataset consists of 17 real EDF images, of which 8 will be used for training and 9 for testing. Each group of cervical cell images contains 20 stacked images and 1 EDF image, and the resolution of each image is 1024*1024. Each image in each set of overlapping images corresponds to a different focus plane for the same sample. Test images are artificially annotated to mark cytoplasm boundaries. The partial plot of the dataset is shown (see Fig. 3).

HEMLBC dataset This dataset contains 917 images of cervical cells, including 241 pictures of abnormal cells and 676 pictures of normal cells. This dataset is labeled by professional pathologists. The partial plot of the dataset is shown (see Fig. 4).

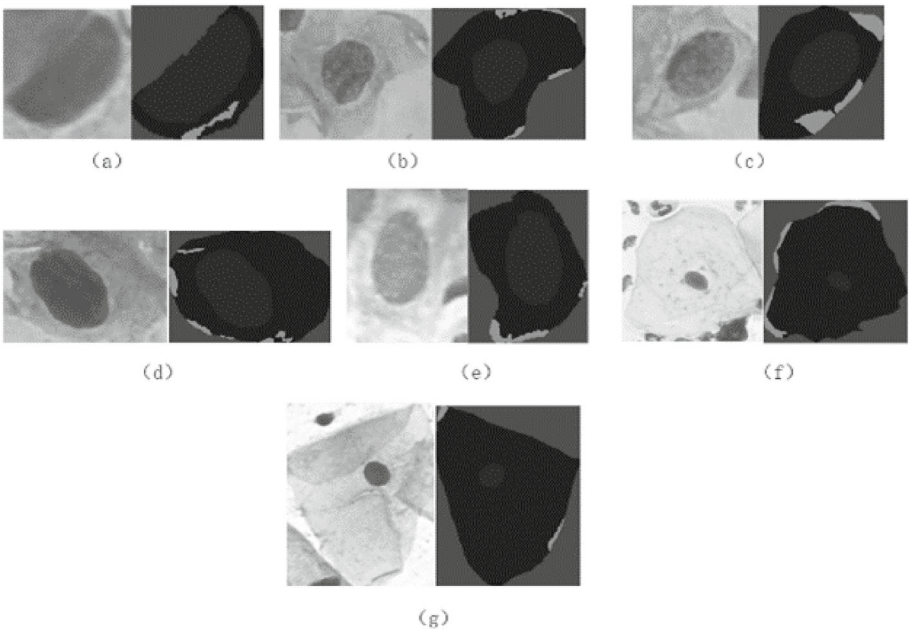


Fig. 1 Example images of cervical cells from the Herlev dataset. There are two images of each type, the original image on the left and the mask image of the original image on the right

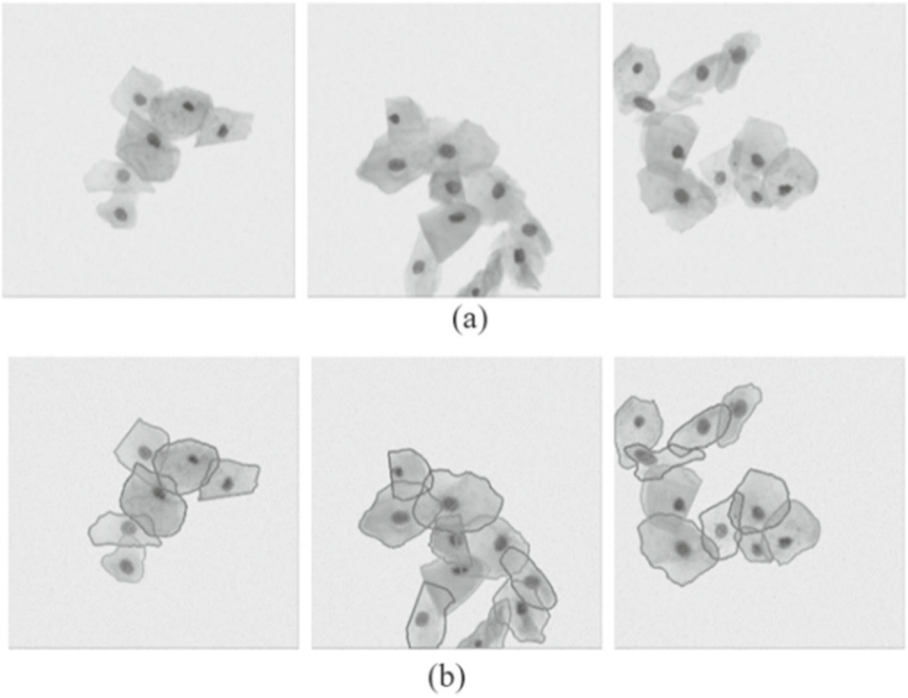


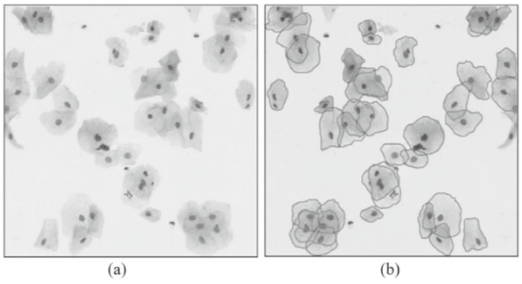
Fig. 2 (a) Example DEF and (b) Ground Truth (Real Label Image)

2.2 Evaluation metrics

Image evaluation metrics are crucial for assessing the performance and effectiveness of image processing algorithms. They provide researchers with an objective and quantitative method to measure the quality of algorithms and serve as important reference criteria for algorithm improvement and optimization. Therefore, image evaluation metrics hold significant importance in image processing. This paper summarizes commonly used evaluation metrics in multiple studies on cervical cell segmentation, specifically focusing on object-level metrics [18, 19].

Object-level metrics are based on a confusion matrix, and in order to describe each metric, you first need to describe the four basic parameters used to calculate these metrics. Among them, True Positive (TP) represents that the true category of the sample is a positive case,

Fig. 3 (a) Example DEF and (b) Ground Truth (Real Label Image)



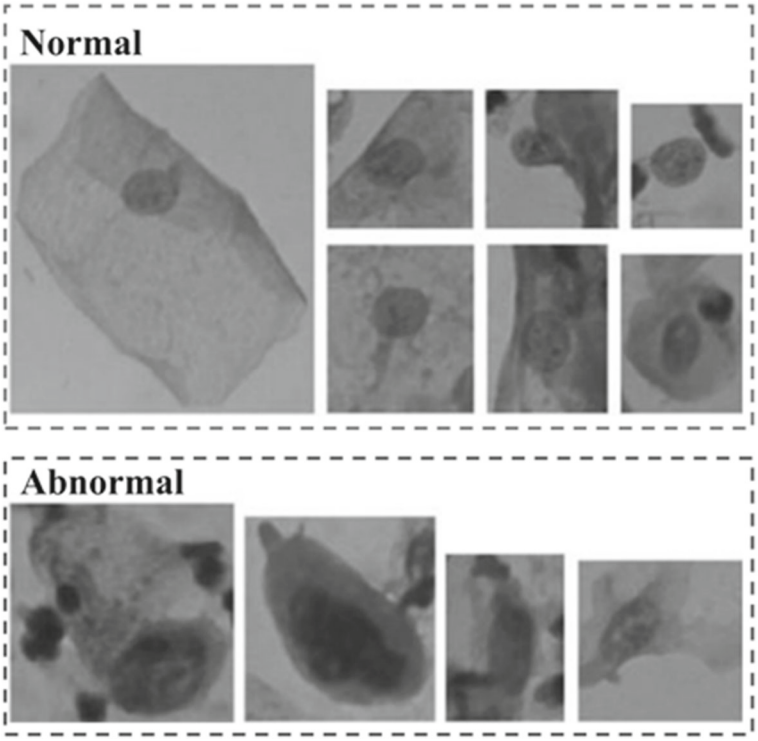


Fig. 4 Example images of normal and abnormal cervical cells from the HEMLBC dataset

and the result predicted by the model is also a positive case; True Negative (TN) indicates that the true class of the sample is a negative example, and the model predicts it as a negative example; False positive (FP) indicates that the true category of the sample is a negative case, but the model predicts it as a positive case; A false negative (FN) indicates that the true class of the sample is positive, but the model predicts it as a negative case.

1) Accuracy: The probability of predicting the negative and correct positive sample is an indicator of completeness to measure the performance of segmentation. For the specific formula, see (1):

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

2) Precision: refers to the probability of correct identification from all detected results (predict the probability of a certain class of correctness; the higher the accuracy, the better the performance of the segmentation method. The specific formula is shown in (2):

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

3) Recall: refers to the probability that it can be correctly identified from real data; the higher the recall rate, the better the segmentation performance of the method. For the specific formula, see (3):

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

4) Dice coefficient: measures the similarity between two sets (such as real and predicted labels). The value ranges from 0 to 1; the closer to 1 means, the more similar. The Dice coefficient is also known as the Dice similarity coefficient (DSC) and the Zijdenbos similarity index (ZSI), F1 Score. ZSI is calculated based on the degree to which the two segmentation regions overlap. If the ZSI is greater than 0.7, it indicates that the detected segmentation boundary closely matches the real image. The F1 score is the harmonic mean of precision and recall. For the specific formula, see (4):

$$Dice = \frac{2TP}{2TP + FP + FN} \tag{4}$$

5) IoU: measures the degree of overlap between prediction segmentation and labels. It ranges from 0 to 1, where 0 means no overlap and 1 means complete coincide. For the specific formula, see (5):

$$IoU = \frac{TP}{TP + FP + FN} \tag{5}$$

6) Jaccard index : A measure of similarity and difference between finite sample sets, focusing on the characteristics common to samples. For the specific formula, see (6):

$$Jaccard = \frac{TP}{TP + FP + FN} \tag{6}$$

Among them, TP represents the true yang class, the true class of the sample is the yang class, and the result of the model recognition is also the yang class. FN represents false negative, and the true class of the sample is positive, but the recognition result of the model is negative. FP represents false positive, and the true category of the sample is negative, but the result of model recognition is positive. TN represents the true negative class, the true class of the sample is the negative class, and the result of the model recognition is also the negative class.

3 Literature review

In medical image processing, image segmentation is a prerequisite for identification, and the segmented results directly affect subsequent recognition and extraction processes. Segmentation techniques in cytology aim to provide labels for each pixel in an image, classifying them between desired objects (such as cells or nucleus) and undesired objects (such as the background) to facilitate image interpretation.

This section presents a literature review and comprehensive analysis of recent techniques on cervical cell segmentation, primarily categorizing them into traditional and deep learning-based image segmentation methods. To systematically summarize and evaluate the strengths and weaknesses of different methods in achieving specific segmentation objectives, namely nucleus segmentation, cytoplasm segmentation, and cell segmentation segmentation of cervical cells, relevant papers on cell segmentation using both traditional and deep learning methods are elaborated upon.

3.1 Tradition methods

Traditional cervical cell image segmentation methods are early segmentation techniques that utilize various image processing and segmentation algorithms to automatically segment

cervical cell images. These methods are often simple yet effective. This section will elaborate on relevant papers that employ traditional methods for cell segmentation.

3.1.1 Nucleus segmentation

When cervical cells are infected, the nucleus's shape, color, and texture are the first to show abnormalities. The cervical cancer auxiliary diagnosis system first needs to accurately segment the nucleus and then determine whether the cell has undergone pathological changes after feature extraction and classification. As the first step in screening for cancerous cells, the segmentation result affects the accuracy of the diagnostic system in screening for pathological cells. Usually, one nucleus corresponds to one cervical cell, and the nucleus serves as an indicator for cytoplasm detection. Therefore, the accuracy of nuclear detection directly affects the subsequent recognition and segmentation of cytoplasm. At the same time, factors such as cervical cell overlap, uneven staining, poor contrast, and the presence of neutrophils pose a significant challenge to cervical cell nuclear segmentation.

Regarding cervical cell nucleus segmentation, related work can be divided into two categories: single-cell image nucleus segmentation and multi-cell image nucleus segmentation. For single-cell image nucleus segmentation, the focus is often on improving the accuracy of nucleus segmentation to more precisely assess the degree of cell abnormality [20, 21]. In the multi-cell image nucleus segmentation research, the emphasis is on accurately identifying the nucleus regions in the image, which serve as indicators for subsequent cytoplasm segmentation. Early studies primarily focused on single-cell image nucleus segmentation, and to date, satisfactory segmentation results have been achieved for single cervical cell images. However, cervical nucleus are commonly found within clusters with adhesion or overlap in multi-cell images.Therefore, in recent years, some researchers have tried to segment the nucleus in multicellular images. Segmentation of overlapping or adhesion nuclei is very challenging in multi-cervical cell images. These traditional methods can be categorized into several types, including threshold-based methods, level set methods, deformable model-based methods, superpixel methods, edge-based methods, cluster-based methods, and other traditional image segmentation methods.

Threshold segmentation This method is a region-based image segmentation technique. The basic idea of the threshold method is to calculate one or more grayscale thresholds based on the grayscale characteristics of the image, compare the grayscale value of each pixel in the image with the threshold, and finally classify the pixels into appropriate categories according to the comparison results. Therefore, the most critical step of this class of methods is to solve the optimal gray threshold according to a criterion function. Jaya et al. [22] proposed a multi-class channel threshold segmentation method based on mean and standard deviation to preprocess, segment, and feature extract cervical Pap images. First, the input Pap smear image is converted to a grayscale image, and the image is averaged filtered to obtain a clear and noiseless microscopic image and enhanced with limiting adaptive histogram equalization (CLAHE). After that, threshold segmentation is used to segment by taking the mean and standard deviation of the three channels. Finally, the area properties, shape, texture, and some statistical characteristics of the segmented image are extracted.The authors evaluate the performance of the method on a private dataset. Threshold segmentation is characterized by weak marginalization, less noise sensitivity, and weak target space structural information, so the segmentation effect of images with complex backgrounds is not good.

Level set segmentation This method is a digital method of tracking the motion of contours and surfaces, which does not operate directly on the profile. However, it sets the profile as a

zero-level set of a high-dimensional function called the level set function. Then it differentiates the level set function to obtain the contour of the motion by extracting the zero level set from the output. Braga [23] proposed a stratified segmentation algorithm for isolated and overlapping cervical nuclei based on a narrowband horizontal set. First, the cell clusters in the input image are segmented, each with a bounding box. After that, a new multi-scale analysis algorithm is applied to estimate the number of clusters in each image region containing cells and converted into input to the narrow band level set algorithm, and for each cell cluster segment, the narrow band level set algorithm is applied for recursive division. The new segment divided is again used as the input of the median narrow band level set. Finally, the new multi-scale cluster detection algorithm is applied to the initialization and guidance of the hierarchical division process of the median narrow band level set. The authors evaluate the method's performance on the Herlev dataset and ISBI 2014. This method adopts hierarchical clustering, which can quickly and accurately identify the position and shape of the nucleus, and the narrow band method used can converge quickly and reduce the complexity of the calculation. However, the method has high requirements for image preprocessing and needs to be denoised and enhanced. Otherwise, it will affect the accuracy of segmentation, and the bandwidth size of the narrow band is more sensitive, and the choice of bandwidth size needs to be adjusted according to the specific situation.

Edge-based segmentation This method is mainly based on the law that the edge gray value is expressed as step or roof type and is suitable for contour extraction in cellular and non-cellular regions. The most significant disadvantage of this type of method is that it is sensitive to noise. Therefore, cell contour extraction is often combined with filters such as the more common differential operators such as the Robert operator, Sobel operator, and Canny operator. Among them, the Canny edge detection operator [24] is a multi-level edge detection algorithm developed by John F. Canny in 1986. It was once considered the strongest edge contour extraction operator. Zhao et al. [25] propose a novel method for nuclear segmentation based on selective edge enhancement (SEENS). First, ROIs that may contain one or more nuclei are extracted from the background by selective search, and then subsequent steps are performed for each extracted ROI. Segmentation results of cell images. After that, on this basis, the original ROI is screened twice, removing the duplicate selection region and the non-nuclear region. Finally, on the basis of contrast grouping, the ROI of the low contrast group was enhanced and segmented by the Chan-Vese(CV) model [26], while the CV model directly segmented the ROI of the high contrast group, and the Canny operator was used to extract the edge information combined with mathematical morphology to enhance the nuclear ROI to be segmented, thereby improving the segmentation accuracy of the CV model. The authors evaluate the performance of the method on a private dataset. The experimental results show that SEENS has high accuracy in cervical nucleus segmentation and improves the segmentation results of cell images under low contrast.

Cluster segmentation The Fuzzy C-Means clustering method (FCM) [27] is widely popular in medical image segmentation and analysis algorithms. Saha et al. [28] first used the FCM algorithm for nuclear segmentation in overlapping cervical cell images, proposing a Pap smear nuclear segmentation method based on spatial shape-constrained fuzzy C-means clustering. The proposed technique introduces a circular shape function (CSF) to increase the robustness of Pap nuclear segmentation. CSF imposes shape constraints on the formed clusters while improving the boundary of the nucleus. The shape function helps to distinguish pixels with similar intensity values but located in different spatial regions. The authors evaluated the method on the ISBI 2014 dataset, and qualitative and quantitative experimental results showed that the new technique performed well, improving the boundary segmentation

performance of the clustering algorithm. Saha et al. [29] did further work based on 2016, using an adaptive MSER algorithm to adaptively calculate the spatial shape force threshold. First, in the background subtraction stage, the foreground image is obtained by removing the background from the Pap image through morphological reconstruction operations. Afterward, the proposed circular shape-constrained fuzzy clustering is performed using the supplementary foreground image and initialized fuzzy partition matrix, and the cluster with the highest average intensity is subsequently detected as a nuclear cluster from the clustering output. Finally, in the reduced false positive stage, outliers are removed to find the final candidate nucleus. The authors evaluated the performance of this method on the ISBI 2014 dataset. The algorithm obtains a suitable number of clusters for the FCM algorithm through the training set statistical methods, thereby solving the problem that there is a large difference in segmentation results due to cluster number selection and has a certain improvement compared to previous recall and accuracy rates. It can adaptively adjust cluster centers and fuzziness parameters to adapt to different image features. However, it is only suitable for circular shapes of nucleus and may not be able to adapt to other more complex shapes of nucleus. The selection of cluster numbers is often a key factor in the performance of FCM and its improved algorithms. Huang et al. [30] proposed a cervical cell nucleus segmentation method based on a multi-scale fuzzy clustering algorithm for the problem of selecting category numbers in clustering algorithms. First, identify cell clusters, separate foreground and background areas from cervical cytology images. Then, a multi-scale FCM algorithm is used to obtain multi-scale image fragments and construct hierarchical trees. Finally, determine interest nodes in hierarchical trees according to the region prior interest degree, use the DRLSE method for fine segmentation of cell nucleus boundary, apply concave point detection to obtain candidate cell nucleus, then use the feature threshold method to remove some non-nuclear regions with different shapes and intensities from fine segmentation and candidate cell nucleus, to obtain nuclear region. The authors evaluated the performance of this method on IBSI 2014 and IBSI 2015 datasets. Experimental results show that this algorithm achieves high nuclear segmentation accuracy.

Advanced methods of nucleus segmentation of cervical cells typically work from coarse to fine. In the coarse phase, general segmentation techniques (e.g., morphology, Hough transform, K-means, or other methods to generate coarse nucleus candidates) are generally used. The fine stage operates on each candidate and is designed to provide more accurate segmentation. Zeng et al. [31] proposed an unsupervised segmentation method for active contour cervical smear nuclei based on an adaptive locoregional fitting energy model. The basic idea is to make the contour adaptively deformed so that the energy function of a given region is minimized. First, the main distribution of nuclei was detected after the coarse segmentation of the cervical smear image by a morphological method. After that, the Gaussian kernel function is used to extract the local region and define its local region fitting energy function, and some parameters in the improved active profile model are estimated using the previous segmentation results. Finally, the Split Bregman [32] method is used to solve the numerical solution of the active contour model to generate the segmentation result. The authors evaluate the performance of the proposed method on the private dataset, and the segmentation results are relatively accurate and robust, and the proposed adaptive energy fitting function can reduce the influence of intensity inhomogeneity on the evolution process of active contours. Zhang et al. [33] propose a general graph-based method for improving the segmentation of abnormal nuclei in cervical cytology images. First, the initial segmentation of the nucleus, that is, the coarse segmentation, determines the rough center and size of the nucleus. Then, according to the annotation protocol that relies on the nucleus boundary of the initial segmentation, the rectangle around each nucleus candidate (sub-image) is cropped, the cropped

sub-image is image-expanded, and the ellipse-like boundary in the Cartesian coordinate system is mapped to a straight line in the polar coordinate system, thus constructing a graph that corrects the local cost function using a priori constraints on the nucleus-cytoplasm position. Finally, dynamic programming and iterative methods determine the globally optimal path in the constructed graph to ensure the optimal closed profile. The authors evaluated the performance of the method on the Herlev and HEMLBC datasets. Graph-based solutions offer several advantages in image segmentation. One of the key advantages is their ability to capture global information and long-range dependencies in the image [34]. However, this segmentation method primarily follows a coarse-to-fine approach, which is partly dependent on initial segmentation results. If the initially detected nucleus center is far from the correct nucleus region, or if the cytoplasm boundaries extracted initially are incorrect, the segmentation results may fail.

In cell images, the contrast between targets is an important feature. Cell images can be enhanced in the data preprocessing part due to factors such as staining and illumination, which will cause the contrast between the cells themselves and the background. Agarwal et al. [35] proposed a method for cervical nuclei segmentation based on the mean shift method. First, the cell images are correlated with stretched contrast enhanced by using the decorrelation stretching method. After that, the mean drift method is applied for segmentation. Finally, morphological manipulation is used to further optimize the segmentation results. The author evaluated the method's performance on the ISBI 2014 dataset, and the experimental results show that this simple and effective method produces a high validation rate on large image datasets. For well-made datasets, contrast can be used directly as a feature. Saha et al. [36] propose a superpixel pooling framework based on a statistical region merging (SRM) algorithm for precise segmentation of cervical nuclei. This method combines the superpixels generated by the SRM algorithm by using paired area contrast and gradient boundaries. First, the input image is converted to grayscale with Gaussian filtering and histogram equalization. After that, the image is segmented into multiple superpixels using an SRM algorithm [37] with adaptive parameter values, merging adjacent superpixels into larger regions based on paired area contrast and gradient boundaries. Finally, morphological manipulations and connected region markers are used to remove noise and small regions, and the final nucleus segmentation results are obtained. The authors evaluated the performance of the method on the ISBI 2014 dataset. The algorithm automatically selects the optimal superpixel size and merge threshold, reducing the effort of manual intervention and parameter adjustment.

3.1.2 Cytoplasm segmentation

Cytoplasm signatures have proven to be critical for the identification of abnormal cells. Accurate cytoplasm segmentation in cervical cell images is the core link of cervical cell segmentation. Once the cytoplasm boundary is located, quantitative evaluation index values such as cell diameter and nucleus-to-plasma ratio can be calculated. Some of the earlier methods addressed free cell and partially overlapping cell segmentation, which did not overlap with neighboring cells or only a tiny fraction of them, and when the cervical cell image intensity met the bimodal distribution, the automatic threshold segmentation method could give better segmentation results. The threshold technique was used to identify cytoplasm in single-cell images in [38], but due to poor contrast and uneven staining of cervical cell images, the segmentation results obtained by the threshold method were not satisfactory. A multi-resolution threshold cell mass and cytoplasm detection method is proposed in [39]. This method locates cell clumps in low-resolution images and detects cytoplasm and nucleus boundaries in high-resolution images.The geometric activity model was used to segment

the cytoplasm of single-cell cervical images and cell images with a low overlap rate. For cytoplasm segmentation, you can start from the nucleus, locate the nucleus first, and then segment the cytoplasm through other characteristics. Tareef et al. [40] proposed a segmentation framework to segment overlapping cervical cells. First, cell block estimation is performed by morphological filtering and maximum region calculation. After that, the nuclei within the cell clump are localized by the H-maxima transform and threshold method. Finally, discrete contour extraction, fractional shape approximation, and shape regularization perform partially overlapping cytoplasm segmentation. The authors test the proposed method on a private dataset. Experimental results show that this method can achieve accurate and stable cytoplasm segmentation, better nucleus segmentation, and lower time complexity. Given the computational efficiency of the superpixel method, some scholars use the simple linear iterative clustering superpixel method to complete single-cell and overlapping cervical cell segmentation. Xia et al. [41] proposed a new method for cervical cell image segmentation based on region merging and simple linear iterative clustering (SLIC) [42]. First, the mean drift algorithm is used to smooth the cell image. After that, the initial contour of cervical cells is extracted using an adaptive threshold algorithm. Then, a simple linear iterative clustering of the smoothed cell image is performed to obtain the superpixels of the entire cell image. Finally, based on the initial outline, the background and foreground labels are set for the cell image full of superpixels, and the superpixel merging is carried out according to the principle of maximum similarity. The authors evaluated the performance of this method on the Herlev dataset, demonstrating that it can accurately extract the contours of the cytoplasm from single-cell Pap smear images in a relatively short time. Lee and Kim [43] proposed an automated segmentation method for overlapping cervical cells using superpixel segmentation and cell contour refinement. Firstly, the cell clumps are detected by superpixel generation and adaptive triangle threshold, the search region is determined, and the triangular threshold method is used to extract cell clumps and candidate nuclei from the superpixel points obtained by SLIC. After that, nuclei are extracted by local threshold and outlier removal. Finally, cytoplasm segmentation is performed using superpixels, and cell contours are refined by map cutting. The authors evaluated the method's performance on the ISBI 2014 and ISBI 2015 datasets.

In addition, merging the shape a prior into the parametric segmentation process can significantly enhance the segmentation effect of overlapping cells. Phoulady et al. [44] first used elliptic shape a prior to segment cells, and Nosrati et al. [45] optimized the above algorithm by replacing the original elliptic shape a priori by using a star shape a priori based on the elliptic shape priori model and obtained more accurate cytoplasm segmentation results. However, elliptical and star shape priors are too simplistic for representing the true shape of cervical cells, and overlapping cell segmentation algorithms using fixed-shape priors cannot provide sufficient shape information about cell overlaps. Song et al. [46] propose an adaptive shape prior information method based on contour fragments to segment overlapping cytoplasm in cervical images to address this issue. First, their contours are cut into profile fragments, and the corresponding profile fragments of each cytoplasm are localized through a grouping process. Later, for each cytoplasm, based on the grouped fragments and a set of known shape references, its shape is constructed, the fragments are joined to form a closed profile as a result of segmentation, the strength and curvature information is further integrated, and the contour fragments extracted into the frame by supplementing the shape a prior are extracted into the frame to improve segmentation accuracy. In addition, the authors propose iterative fragment grouping, shape construction, and fragment joining to gradually refine the shape a prior and improve the segmentation results. The authors evaluate the method's performance on ISBI 2015 and private datasets. Song et al. [47] proposed a method for overlapping cell nucleus

segmentation based on constrained multi-shape evolution. This method jointly evolves the shape of each cytoplasm under the guidance of modeled shape priors while segmenting all overlapping cytoplasms within a clump. First, the lack of intensity of segmentation is compensated for by introducing a local shape prior (cytoplasmic level) and a global shape prior (clump–level), where modeling a local shape prior is created by creating an infinite set of shape assumptions with statistical shape information from the cytoplasm and modeling a global shape prior is created by considering the mutual shape constraints of the nuclei in the cluster. Second, a learning algorithm was developed to learn the importance of each shape instance in the shape statistics calculation, ensuring that the constructed shape hypothesis set encompasses all possible shapes of the cytoplasm. The authors evaluated the performance of this method on two typical cervical smear datasets, and extensive experimental results confirmed that the method effectively addresses the issue of insufficient intensity, which is the most challenging problem in the task of overlapping cytoplasm segmentation.

3.1.3 Cell segmentation

More advanced image segmentation techniques are used to identify and segment different cell areas. Reviewing cell segmentation methods can facilitate research and applications with integrated objectives. For example, performing joint segmentation of nucleus and cytoplasm can enhance the overall segmentation accuracy and consistency, aiding in a more accurate assessment of cellular morphological features and identification of abnormal cells.

Level set segmentation The level set algorithm can effectively describe the curve of curvature-dependent rate evolution, which is not based on gradient changes, and has a natural advantage when processing images with weak edge features. Islam and Haque [48] have developed a novel multistep level set (LVS) method for segmenting cytoplasm and nuclei in overlapping cells from a single EDF image. First, a new level set method is used to detect clumps from Gaussian-filtered EDF images using intensity and edge information. After that, the nuclei are detected by applying the intensity-based level set multiple times until no nuclei remain. Finally, the modified LVS method was used for each nucleus to detect their corresponding cellular boundaries. The authors evaluate the performance of the method on the ISBI 2015 dataset. Bhan et al. [49] propose a new method for supervised segmentation of overlapping cervical Pap smears by segmenting the image into many non-overlapping pixel blocks and then extracting texture features from the grayscale symbiotic matrix GLCM, thereby precisely segmenting overlapping cells. First, the image is divided into blocks of 16 $\times$ 16 size. After that, the texture features of the EDF image are extracted; Then, based on changes in texture characteristics, the area of interest, the overlapping area, is finally determined. Finally, independent level sets are used to segment the cytoplasm and nucleus. The authors evaluated the performance of the method on the ISBI 2014 dataset.

The Active Profile Model (ACM) is also commonly known as the Chan-Vese (CV) model based on the Level Set method, which is widely used in visual tracking and image segmentation tasks to minimize the characteristics of the desired segmentation of energy functions [50], in addition, the new Gradient Vector Flow (Gradient et al.- ACM, GVF-ACM) model is not only robust to the initial contour of the target. Moreover, it has the advantage of convergence to the concave boundary of the cell, so this method is widely used for the segmentation of cell images. According to the model's different forms of curve expression, active profile models can be divided into parametric active profile models and geometric active profile models. The parametric active contour model is based on the Lagrange framework, which directly expresses the curve in the parametric form of the curve, the most representative being

the Snake model proposed by Kasset a1 (1987). Guan et al. [51] proposed a method to precisely segment partially overlapping cervical cells based on a dynamic sparse profile search and GVF Snake model. Firstly, the K-means clustering algorithm based on morphological filtering was used to extract nuclei and background. After that, an edge enhancement method based on gradient decomposition is proposed, and the basic contour points can be extracted using the extracted background and enhanced edge map. Finally, the dynamic sparse contour search algorithm is used to automatically locate the cell contour points, arrange the localized contour points, construct the initial contour of GVF Snake, and accurately segment the cell image. The authors evaluated the method's performance on both the Herlev and private datasets, and the high-precision cell profile extraction results verified the method's effectiveness.

Cluster segmentation The K-means clustering algorithm is used in the proposed method by Guan et al., which is a region-based segmentation method. The region-based segmentation method is to divide it into different area blocks according to the similarity criterion of the image, which can be achieved by the classic method of finding clustering in the data or the modern method of extracting "superpixels" from the image with clustering, which is divided into cluster-based area segmentation and superpixel-based area segmentation. Among them, the area segmentation based on superpixels relies on the image's color and spatial relationship information. The image is divided into superpixel blocks that far exceed the number of targets and are much smaller than the number of pixels to retain the edge information of all targets in the image as much as possible. SLIC superpixel algorithm is a popular method in superpixel-based region segmentation, which efficiently generates regular and compact superpixel points to complete cervical nucleus segmentation by the K-means clustering method. Zhao et al. [52] proposed a superpixel-based Markov random field (MRF) segmentation framework for extracting the nucleus, cytoplasm, and image background of cervical cells. Firstly, the described technique considers the non-overlapping superpixel block of a single image and represents the image as an undirected probability map of the superpixel node and the connecting edge. The connecting edge represents the spatial information between adjacent superpixels through local probability description. After that, using a labeling mechanism, the three regions of the cell are identified by distinguishing the brightness of the region, such as the darkest nucleus, the brightest background, and the rest as the cytoplasm. Finally, the authors propose a gap search algorithm based on the superpixel model, and the gap search algorithm in MRF accelerates the solution by overlapping the set of superpixel elements with the searched Gap block and determines the search gap block through the Canny edge detector. The edge gap is calculated by drawing a square centered on the pixel point of the edge. The authors processed each image in 0.63s on the Herlev dataset and 1.55s on the real dataset. Thus, the proposed gap search algorithm saves a lot of time. Tareef et al. [53] propose a multichannel rapid watershed-based approach (MPFW) that splits nuclei and cytoplasm from large cell clumps of overlapping cervical cells into three watershed channels. The triangulation transformation algorithm was used to extract the cell masses, and the SLIC algorithm and the label-based watershed algorithm were combined to complete the extraction of candidate nuclei. Applying the SLIC algorithm produces more over-segmented nuclei, reducing the recognition accuracy. Due to the wide range of variations in cervical cell size, depending on the degree of abnormality, the enlargement of the nucleus can even be six times the size of the normal nucleus. Therefore, it is difficult to know the optimal average size of pixels before running the SLIC algorithm. And due to cell overlap and uneven staining, varying intensity and poor contrast of cervical cell images often occur. Based on the above two points, the possibility of nuclear hyper fractionation and under-segmentation is greatly increased using

fixed-size SLIC superpixel points. First, the barrier watershed on the gradient-based edge map locates the nucleus. Isolated and overlapping cells are segmented using morphology-based preprocessing and improved distance transformation. Finally, regional growth and boundary adjustment strategies precisely segment the cytoplasm. In MPFW, the linear contour of the watershed unit is deformed by ellipse fitting and contour adjustment to better represent the cell shape. The authors evaluated the performance of the proposed method in ISBI 2014 and ISBI 2015, and the performance evaluation in the paper is very detailed and can be viewed in the downloadable paper. Experimental results show that the proposed algorithm is superior to the existing equivalent algorithms regarding segmentation accuracy, detection rate, and time complexity.

Watershed segmentation The watershed algorithm is combined with other methods in the approaches mentioned above. Besides the traditional watershed algorithm, there is an algorithm inspired by the watershed algorithm called the Maximally Stable Extremal Region (MSER) algorithm, which can identify cell nucleus regions in cervical cell images. Lu et al. [54] applied the MSER algorithm to detect nucleus in cell clusters and proposed an improved joint optimization method using multiple-level set functions for segmenting overlapping cervical cells. The proposed method consists of two steps: initial cluster segmentation and detailed segmentation of each cell. The first step includes the following stages: 1) unsupervised classification-based cell block detection, 2) cell nucleus detection using the MSER algorithm, and 3) shape prior estimation of overlapping cell regions and initial segmentation of each cell. Subsequently, in the second step, the initial segmentation is optimized using the level set method, which utilizes multiple level set functions to minimize the energy function. The authors evaluated the performance of this method on their own publicly available dataset. Nisar et al. [55] developed a segmentation method for cell clusters, nuclei, and cytoplasm using image processing algorithms. Firstly, data preprocessing was performed using median filtering, contrast-limited adaptive histogram equalization, and H-maxima transformation. Then, cell clusters were segmented from the background using an Iterative Region-based Otsu (IRO) thresholding method [56]. Next, nucleus were segmented using the Maximally Stable Extremal Region (MSER) method. Finally, the cytoplasm was segmented using the distance regularized level set evolution method. The authors evaluated the performance of this method on the ISBI 2014 and ISBI 2015 datasets. The segmentation threshold for extracting cell regions was determined using the Otsu method, which adaptively determines the threshold to accommodate variations in the image better.

Deformable model Furthermore, there are segmentation methods based on deformable models and the application of shape features. Ragothaman et al. [57] proposed an unsupervised segmentation method for cervical cell images based on Gaussian Mixture Models (GMM). Firstly, the image was roughly segmented into a certain number of classes using GMM, and the number of classes was determined heuristically using the Akaike Information Criterion. After segmenting the image, further clustering of image regions based on user-defined parameters was performed to reduce the final number of classes to three: nucleus, cytoplasm, and background. Morphological operations were then applied to remove over-segmented nucleus based on their characteristics. Finally, the Hough transform was utilized for cell nucleus shape recognition based on edge information. The authors evaluated the performance of this method on a private dataset. The use of GMM in this study allowed the segmentation of images without requiring prior knowledge and could handle complex image backgrounds and multiple objects. Tareef et al. [58] proposed an automatic segmentation method for nucleus and cytoplasm in clusters of cervical cells based on unique local features and guided sparse shape deformation. Firstly, image superpixels were classified based on

shape and appearance using Support Vector Machines to segment the image into nucleus, clusters, and background. Then, based on the sparse coding theory, a robust shape deformation framework guided by typical shape features was proposed to construct cytoplasm shapes for each overlapping cell. Finally, the obtained shapes were refined using the distance regularized level set evolution model. The authors evaluated the performance of this method on the ISBI 2014 dataset.

Most of the above methods use two well-known public overlapping cervical cell datasets, ISBI 2014 and ISBI 2015, which were announced at the first and second cervical cell image segmentation challenges held by the IEEE International Biomedical Imaging Society in 2014 and 2015. In recent years, most overlapping cell segmentation methods have been proposed to solve these two datasets' overlapping cell segmentation tasks, and most multi-target segmentation tasks are performed on these two datasets. Ushizima's team won the first challenge. Ushizima et al. [59] used a segmentation method combining superpixel representation with Voronoi diagrams to segment nuclei. This method includes three main steps: cytoplasm mass estimation, nuclear detection through superpixel representation, and cytoplasm detection through nuclear narrowband inoculation, graph-based region growth, and Voronoi diagrams. Our algorithm relies on three main steps: segmenting cell clusters by defining superpixels, using Phansalkar local search to segment nuclei from low-contrast images, and detecting and segmenting cytoplasm by calculating the Voronoi diagram of previously detected nuclei as a constraint on cytoplasm boundaries. The authors evaluated the performance of this method on the ISBI 2014 dataset. The Phoulady team won the second challenge. Phoulady et al. [60] improved the previous method, using an iterative thresholding method to complete nuclear segmentation, and proposed a new multi-layer cervical cell image overlapping cell detection and segmentation method. First, locate the nucleus within the extended depth of field (EDF) image using an iterative method. Then use a learned Gaussian mixture model to segment clusters from pixel intensity through the expectation-maximization (EM) algorithm. Finally, use a new distance measure to approximate the overlapping boundaries of cytoplasm on multiple focal planes, then refine it in two post-processing steps. The authors evaluated the performance of this method on the ISBI 2015 dataset. Experimental results show that this method can effectively detect and segment nuclei and their corresponding cytoplasm in highly overlapping cytological multi-layer Pap smear volumes. However, this method has problems, such as many false nuclei in the segmentation results using the iterative thresholding method applied and sensitivity to changes in image intensity. Phoulady et al. [61] proposed an iterative threshold algorithm based on regional solidity features to segment overlapping nuclei and cytoplasm in response to the above iterative threshold problem. This algorithm has low sensitivity to changes in image intensity and improves the accuracy of nuclear detection. It uses EDF images for nuclear detection, cell cluster segmentation, and volume images to segment overlapping cytoplasm. First, an iterative algorithm designed using three features of size, average intensity, and solidity is used to detect and segment nuclei. Then use the expectation-maximization algorithm to learn a Gaussian mixture model with two components on pixel intensity to separate cell clusters from the background. One Gaussian model estimates the distribution of foreground (cell block) pixel intensity, while the other estimates background pixel intensity. Finally, these two criteria are used to segment overlapping cytoplasm. The authors evaluated the performance of this framework on ISBI 2014 and ISBI 2015 datasets. The author did not mention the efficiency and scalability of this method when processing large-scale data, so it is impossible to evaluate its feasibility in practical applications, whether there are certain limitations when processing different types of cell images is also unknown.3.2 deep methods

3.1.4 Traditional methods analysis

The purpose of segmentation is mainly to provide a region of interest for subsequent classification and analysis, almost all cell images contain four parts: nucleus, cytoplasm, background and impurities, first of all, background and impurities are the interference factors that affect the analysis, so segmentation is to eliminate the interference factors, leaving important areas that are the nucleus, cytoplasm or the whole cell region. The characteristics of the nucleus are important for the grading and prognosis of cancer. At the same time, the location of the nucleus can also help locate the cell. Segmentation of the cytoplasm region can help to calculate the quantitative evaluation index values such as cell diameter and nucleocytoplasm ratio, which is helpful for subsequent analysis and classification. Therefore, it is necessary to summarize different segmentation objectives.

Cellular images often suffer from reduced contrast between the cells and the background due to factors such as staining and illumination. This contrast deficiency is particularly pronounced in instances where cell nuclei are small, and the cytoplasm exhibits light colors. In addressing this issue, traditional methods have employed enhancement techniques during data preprocessing. For instance, the de-correlation stretch method, as applied by Agarwal et al. [35], is utilized to enhance the contrast of cell images through correlated stretching. Furthermore, the presence of numerous overlapping, proliferative, and closely adjacent cells poses challenges in distinguishing between the boundaries of cell nuclei and cytoplasm. To tackle this problem, conventional approaches often leverage shape constraints. Saha et al. [28] introduced a nuclear segmentation method for Pap smear images based on spatial shape constraints using fuzzy C-means clustering. The proposed technique incorporates a circular shape function (CSF) that imposes shape constraints on the formed clusters, simultaneously improving the boundaries of the nuclei. Additionally, Song et al. [46] presented an adaptive shape prior information method based on contour fragments for segmenting overlapping cytoplasm in cervical images.

Most traditional methods can achieve greater than 90% accuracy for nuclear segmentation targets for single-cell datasets. When you first come into contact with the field of cervical cell segmentation, you can start with nuclear segmentation. For the multi-cell dataset for nucleus segmentation, the superpixel method and clustering method based on the Statistical Region Merging (SRM) algorithm are better than the traditional methods, and most of the studies using the clustering method to achieve nucleus segmentation are the most common methods. There are few traditional methods for cytoplasm segmentation of cervical cells, and we can sometimes achieve the segmentation of the nucleus in this step by locating the nucleus first, so it is usually rare to study the method of cytoplasm task alone, but some of the methods used in it are also worth summarizing and learning, and the SLIC superpixel method is more commonly used in the traditional method. In the multi-objective segmentation of cervical cells, in terms of traditional methods, there are many research methods and rich materials, especially in the refinement of segmentation, that is, the post-processing part, the traditional methods are very effective, among which threshold, superpixel, level set and watershed algorithm are still commonly used algorithms. Table 2 summarizes the traditional methods used for segmentation.

3.2 Deep learning methods

Traditional cell segmentation methods rely on manually engineered features to segment images, mostly utilizing surface-level information. These methods are unsuitable for

Table 2 Summary of traditional methods

Segmentation objects	Methods classification	Year	Methods and Literature	Dataset	Performance
Nucleus	Threshold	2020	Multi-channel threshold [22]	private dataset	Accuarcy=0.946
	level set	2021	Narrow band level set [23]	ISBI 2014	ISBI 2014: DSC=0.92± 0.04,
			Herlev		Herlev: Mean DSC=0.93
	Edge-based	2021	Edge enhancement	private dataset	mean Dice=0.9375, Jaccard=0.8834,
			morphological algorithm [25]		Precision=0.9641, Recall=0.9177
	Cluster	2016	FCM algorithm circle shape function [28]	ISBI 2014	Precision=0.918, Recall=0.915
		2017	MSER algorithm (improved) [29]	ISBI 2014	Dice=0.938, Precision=0.968, Recall=0.939
		2020	Multi-scale FCM algorithm [30]	ISBI 2014	ISBI 2014 train-set :Precision=0.98, Recall=0.94
	Other	2015	Adaptive Local Area	private dataset	mean Recall=0.87,
			Fitting Energy Model [31]		mean DSC=0.85
		2017	Graph cut [33]	Herlev	HEMLBC: Precision=0.91±0.04, Recall=0.96±0.04,
				HEMLBC	F1score=0.930±0.03, IoU(overlap)=0.87±0.05
		2015	Mean-shift [35]	ISBI 2014	Precision=0.9357, Recall=0.8275, Dice=0.8737
		2019	SRM Super Pixel [36]	ISBI 2014	Dice=0.956, pixel-based Recall=0.962, Precision=0.930,
					object-based Recall=0.944, Precision=0.987
Cytoplasm	SLIC	2016	Region consolidation SLIC [41]	Herlev	IoU=0.9722±0.0216
		2016	Super Pixel SLIC [43]	ISBI 2014	ISBI2014: Dice=0.897±0.075,
				ISBI 2015	ISBI2015: Dice=0.879± 0.087
	Other	2015	Morphology thresholdmethod [40]	private dataset	Dice=0.92, Precision=0.94
		2019	Based on shape prior information [46]	ISBI 2015	Degree of overlap 0.3: TPR=0.95±0.07, DSC=0.84±0.06,
					Degree of overlap 0.3-0.6: TPR=0.92± 0.06, DSC=0.81± 0.05

Table 2 continued

Segmentation objects	Methods classification	Year	Methods and Literature	Dataset	Performance
		2024	Constrained multi-shape evolution [47]	pap stain dataset	DSC=0.83±0.05
Cell	Level Set	2015	Multi-step level set [48]	private dataset	Accuarcy=0.946
		2016	Independent level set [49]	ISBI 2014	Dice=0.76
		2015	MSER algorithm, level set [54]	private dataset	Precision=0.96±0.06, Dice=0.92±0.05
		2015	GVF Snake model [51]	Herlev	mean Recall=0.87, mean DSC=0.85
	Super Pixel	2015	Superpixel,Voronoi figure [59]	ISBI 2014	Train-45: Dice=0.880±0.074, Test-90: Dice=0.867±0.0832
		2016	Super Pixel Gap search MRF [52]	Herlev	nucleus: ZSI=0.917±0.07, cytoplasm: ZSI=0.817±0.12
		2018	SLIC	ISBI 2014	ISBI 2014: Dice=0.897±0.075,
			watershed algorithm [53]	ISBI 2015	ISBI2015: Dice=0.879±0.087
	Threshold	2017	MSER algorithm,thresholds [55]	ISBI 2014	nucleus: synthetic image Dice=0.92, real image Dice=0.79
				ISBI 2015	cytoplasm: synthetic image Dice=0.84, real image Dice=0.73
		2016	Iterative threshold [60]	ISBI 2015	Dice=0.861
		2017	Iteration threshold based on	ISBI 2014	ISBI2014: Precision=0.961, Recall=0.933
			the solidity feature of the region [61]	ISBI 2015	ISBI2015: DSC=0.873
	Deformable Model	2016	Gaussian mixture model [57]	private dataset	ISBI 2014train-set: Precision=0.9146, Recall=0.9490
	Other	2017	Shape deformation [58]	ISBI 2014	ISBI 2014train-set: Dice=0.89±0.072, Recall=0.91±0.09

Note: This table summarizes the traditional methods used for segmentation. On the basis of classification based on segmentation targets, We summarised the method classification for each section, the year of each article reviewed, and the datasets used and experimental performance

segmentation tasks requiring a significant amount of semantic information and cannot address real-world demands. They only perform well on specific tasks. With the advancement and introduction of deep learning, computer vision has made breakthrough progress. Deep learning-based methods can automatically extract deep features from large data and possess strong generalization capabilities. Deep learning segmentation algorithms can be categorized into semantic segmentation and instance segmentation. Semantic segmentation involves dividing an image into different semantic parts and assigning each pixel to a class (e.g., cell foreground or background) [62]. Instance segmentation aims to identify each instance of the same class by individually detecting and delineating the cells displayed in the image. This section will describe papers on segmentation methods based on neural networks.

Most deep learning papers apply classic neural networks for cell segmentation, such as FCN [13], Mask R-CNN [14], U-Net [15], and others. Some papers utilize variant models based on classic architectures, for example, TernausNet, which is a modified U-Net in the DeepLabv2 network [76]; PGUnet+ [75], which adds residual modules to the model and trains it with progressively growing input sizes.

Among these, FCN performs pixel-level classification of images. Unlike classical CNNs that use fully connected layers to obtain fixed-length feature vectors for classification in the convolutional layers, FCN can accept images of arbitrary sizes. It employs deconvolution layers to upsample the feature maps from the last convolutional layer, restoring them to the same size as the input image. This allows the generation of predictions for each pixel while preserving spatial information from the original input image. The FCN network architecture is illustrated (see Fig. 5)., where "conv" represents the convolution operation, and "pool" represents the pooling operation.

Mask R-CNN, a region-based convolutional neural network, is a method that effectively detects objects while outputting high-quality instance segmentation masks. It commonly uses ResNet networks with 50 or 101 layers as backbones. The authors also explore another efficient backbone structure called Feature Pyramid Network (FPN). The network architecture is shown (see Fig. 6). The first part consists of ResNet and FPN, which serve as the backbone network for feature extraction and generation of feature maps. The second part is the Regional Proposal Network (RPN), which performs object localization. It takes the feature maps obtained from CNN feature extraction as input. 9 anchor boxes are generated using a sliding window for each pixel in the feature map, and feature extraction is performed on them. On the one hand, softmax is used to classify the anchor boxes, distinguishing foreground and

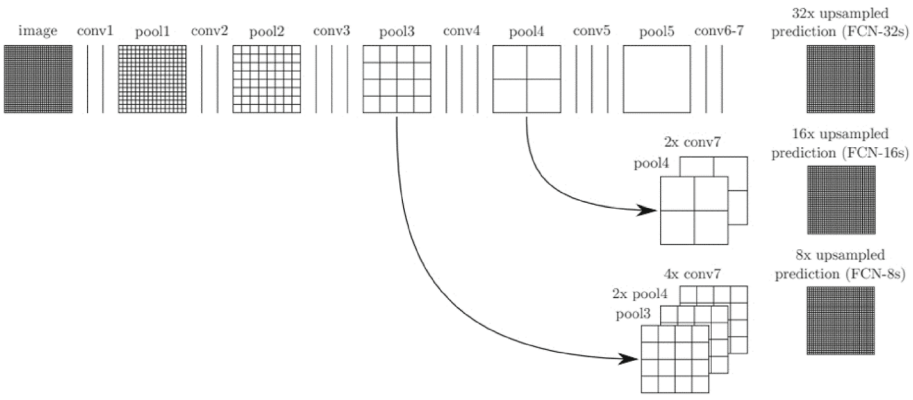


Fig. 5 FCN network structure

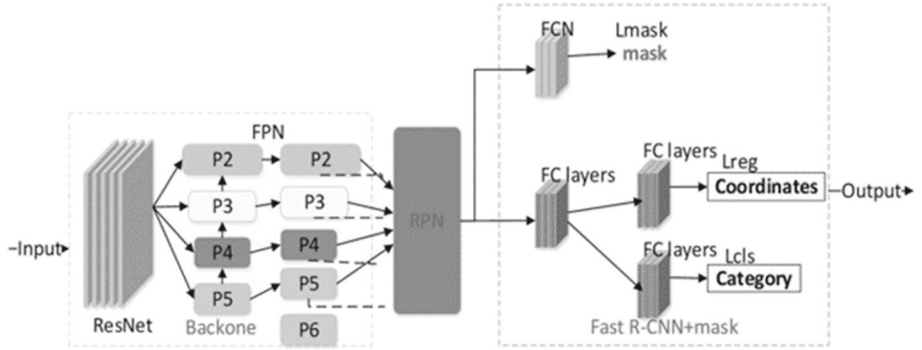


Fig. 6 Mask R-CNN network structure

background in the RPN. On the other hand, the bounding box regression offset of the anchor boxes is computed to achieve a more precise localization of the regions of interest (ROIs). The final anchor boxes are subjected to non-maximum suppression (NMS) based on their scores to remove overlapping and inaccurately positioned boxes. The third part consists of the Fast R-CNN and mask branches. It is responsible for classifying the ROIs and refining their positions using bounding box regression for improved accuracy. The mask branch performs pixel-level classification and regression for the ROIs, accomplishing segmentation and generating masks. In the Fast R-CNN part of the third section, Mask R-CNN predicts the segmentation masks. First, the Fast R-CNN classification results are obtained, and then these results are input to the mask prediction branch for semantic segmentation, resulting in instance segmentation. Mask R-CNN utilizes pixel-level prior information to obtain better semantic features. The backbone network can flexibly adopt different architectures as its backbone, eliminating the need for complex preprocessing steps. Compared to traditional neural networks, Mask R-CNN has a smaller computational overhead, allowing for the implementation of fast systems for training.

U-Net is a convolutional neural network used for biomedical image segmentation. Its superior performance stems from its ability to effectively utilize limited annotated data by employing data augmentation, unlike other neural networks. U-Net employs a novel feature fusion approach by concatenating features in the channel dimension to create thicker features, whereas Fully Convolutional Networks (FCNs) use element-wise summation during feature fusion without creating thicker features. Additionally, U-Net utilizes five pooling layers to achieve multi-scale feature recognition in the network. However, U-Net is known for its slow runtime efficiency. The network needs to run separately for each neighborhood, and for overlapping regions, the network performs redundant computations. There is a trade-off between accurate localization and acquiring contextual information. Larger patches require more max-pooling layers, which can reduce localization accuracy, while smaller neighborhoods provide less contextual information. The U-Net network structure is shown (see Fig. 7), features a symmetrical U-shaped architecture comprising a contracting path (encoder) and an expansive path (decoder). This structure effectively leverages contextual and localization information and enhances segmentation accuracy and robustness through data augmentation. The encoder, located on the left side of the diagram, consists of two 3x3 convolutional layers with a ReLU activation function followed by a 2x2 max-pooling layer for downsampling. The decoder on the right side of the diagram consists of an upsampling layer (deconvolutional layer), feature concatenation (concat), and two 3x3 convolutional layers (ReLU).

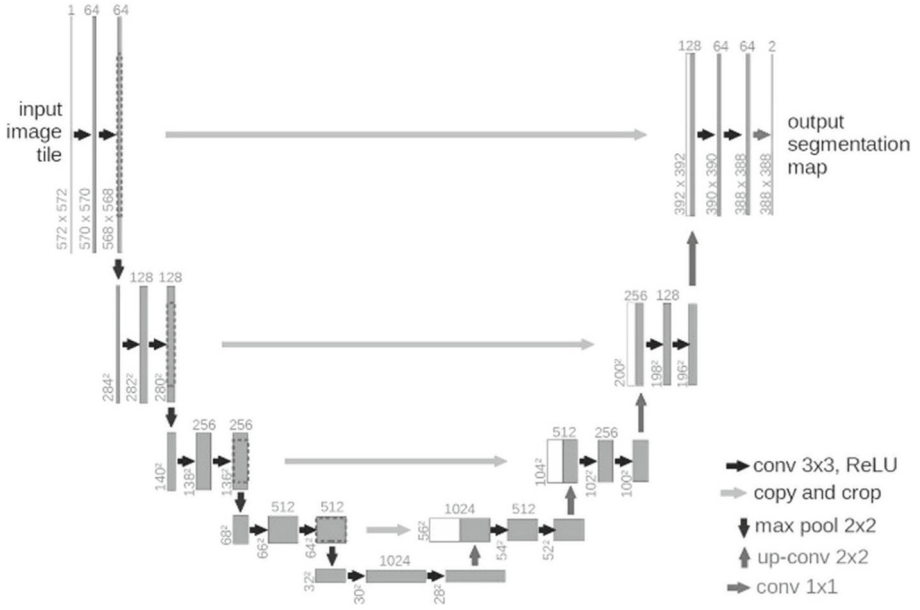


Fig. 7 U-Net network structure

3.2.1 Nucleus segmentation

Segmenting cell nuclei in histopathology images is the preliminary step in analyzing current imaging data for biological and biomedical purposes [63], which provides rich spatial and morphological information about the nucleus and is the basis for abnormal cell detection.

In recent years, neural network methods have been applied to segment the nuclei of cervical cells and have achieved good segmentation results. Some papers apply fully convolutional networks (FCN) to segment cells. FCN is the pioneering work of deep learning in image segmentation, which can achieve end-to-end segmentation. It is more efficient than traditional networks based on CNN for segmentation, but the details of the segmentation results are not good enough. To solve this problem, many researchers use traditional methods such as graph-based segmentation methods to improve the output of the segmentation results by FCN and further optimize the results. Zhang et al. [64] proposed a new cervical cell nucleus segmentation method combining FCN and graph-based methods (FCNG). First, a fully convolutional neural network is used for initial cell segmentation, dividing the entire cell image into background, cytoplasm, and nucleus while generating a probability map of the nucleus. Then, a graph-based fine segmentation is developed based on an annotation protocol that ensures that the entire nuclear region is involved in cropping rectangles around the nuclei segmented by FCN (sub-images), using a graph-based method combined with the nuclear probability map learned by the FCN mentioned above to detect improved nuclear boundaries. The authors evaluated the performance of the proposed method on the Herlev dataset. However, the aforementioned proposed method did not effectively use location information and local features in the image, and this method did not perform well in overlapping cervical cell image segmentation. Therefore, many researchers have studied this issue. The emergence of attention mechanisms has well-integrated multi-layer information and fully utilized important

image features. Zhang et al. [65] proposed a network called Binary Tree-shaped Network (BTTFA). The simplified diagram of BTTFA is similar to a binary tree structure. Integrating adjacent layer pairs each time combines the last four layers of information of ResNext to fully integrate multi-layer information and recover information lost by pooling layers. BTTFA applies dual-path fusion attention to selectively use information close to the root node, thereby fully utilizing low-level details and high-level semantic information, selectively emphasizing important features while suppressing less useful features. At some nodes of the binary tree-shaped network, focus loss is applied to calculate the loss between ground truth and feature maps during training. The authors evaluated the performance of this network on ISBI 2014 and private datasets. This network is dedicated to solving cluster cell nucleus segmentation. It has been effective, but if the color of cytoplasm and nucleus is close, BTTFA's segmentation result is not ideal either. The binary tree-shaped network is well applied in overlapping cervical cell image segmentation, and some scholars have researched single-cell images. Based on FCN, an attention mechanism was added to improve the segmentation effect. Cheng et al. [66] proposed a fully convolutional attention network FCANet suitable for medical image segmentation. Unlike previous work that performed image segmentation through multi-scale feature fusion, a Fully Convolutional Attention Network (FCANet) aggregates long-distance and short-distance contextual information. Specifically, the authors combine spatial and channel attention modules into a Res2Net network with a dilation strategy to better learn feature representation of biomedical images. By aggregating features at each location through spatial attention modules, similar features are mutually promoted in terms of spatial size. At the same time, channel attention modules treat each channel of feature maps as feature detectors and emphasize channel dependencies between any two-channel maps. Finally, the sum of output features from two attention modules is weighted to retain long-distance and short-distance information, further improving feature representation performance and making biomedical image segmentation more accurate. The authors evaluated the performance of this network on the Herlev dataset and other datasets. Experimental results show that FCANet can improve the biomedical image segmentation effect, but when the target area boundary is unclear and similar to the background, the FCANet segmentation result cannot completely match ground truth. Improving the decoder module by retaining more boundary information during upsampling process can help improve segmentation results. In addition, based on FCN, adding other different modules, such as residual and densely connected blocks, can also achieve performance optimization. Hussain et al. [67] proposed a shape context-based fully convolutional neural network (SC-FCN), for instance segmentation and classification of cervical nucleus in Pap smear images. First, pre-training is performed using a nucleus pixel mask combined with a shape representation model (SRM). In the next step, this model serves as a regularizer for the fully convolutional neural network. Then, the fully convolutional neural network is trained using the output images of SRM and combined with the pre-trained SRM model by including its network parameters, resulting in instance segmentation output images. Finally, the output of the SRM model is combined with the nucleus pixel mask along the nucleus type prediction mask to obtain instance segmentation and classification output images. The authors evaluated the performance of the proposed model on a private dataset, achieving a ZSI score of approximately 97%. Although this network is based on shape, a microscopic feature, and uses FCN to optimize performance by adding residual blocks, densely connected blocks, and bottleneck layers, it is limited by its reliance on limited datasets and their associated mask annotations.

The above methods achieve instance segmentation of nucleus based on FCN. Instance segmentation is a combination of object detection and semantic segmentation, which can not only segment the edges of objects but also label different individuals in the same type of

objects in the image. Existing instance segmentation methods often require accompanying object detection, such as using Mask-RCNN, leading to overly complex models [68]. For example, Liu et al. [69] proposed a cervical cell nucleus segmentation method based on pixel-level prior information. This method uses pixel-level prior information to provide supervision for the training of the Mask Region Convolutional Neural Network (Mask-RCNN). It uses a Feature Pyramid Network (FPN) based on ResNet and modified according to nucleus images as the backbone of Mask-RCNN. First, using the Mask-RCNN algorithm, pyramid feature mapping suitable for images is extracted through pixel-level prior information of nuclei. Then, nucleus recognition and coarse segmentation are performed, and the nucleus region of interest is obtained by expanding the bounding box of the nucleus provided by Mask-RCNN's recognition. Finally, the coarse segmentation in this region is fed into a local fully connected conditional random field (LFCCRF) for refinement, and the final segmentation is achieved by minimizing the energy of LFCCRF. The authors evaluated the performance of this method on the Herlev dataset. Although this method is more accurate and stable in cervical cell nucleus segmentation, the accuracy of abnormal nucleus segmentation still needs to be improved.

Among the neural network methods, there is a classical model - U-Net, which is a convolutional neural network for biomedical image segmentation, proposed by Olaf Ronneberger et al. in 2015. U-Net achieved good segmentation results at that time, which influenced the design of many subsequent segmentation networks. Some work enhances the segmentation performance of classical U-Net networks by designing complex aggregation feature modules between the backbone and segmentation semantic headers. Zhao et al. [70] proposed a cervical cell nucleus automatic segmentation method based on a deformable multi-path ensemble model (D-MEM). The method employs a U-shaped convolutional network as the backbone, utilizing dense blocks to convey feature information. Deformable convolutions are utilized to handle the irregular shapes and sizes of different nucleus to enhance the model's flexibility. In order to further improve the segmentation process, the model is organized in a multi-path manner, training multiple networks with different configurations and integrating the results using a majority voting strategy. The authors evaluated the performance of this method on the Herlev dataset, achieving optimal segmentation accuracy and demonstrating the potential for generalization to other medical image segmentation tasks. U-Net and its latest extensions, such as Transunet [71] have been leading medical image segmentation methods in recent years. U-Net combined with Transformer [72] achieved good performance in several medical imaging tasks; However, in the cervical cell segmentation task, Transformer is not well applied to this task due to the limitation of data volume [73], but some other models have achieved good segmentation results for U-Net variant models. Qin et al. [74] propose a novel multi-tasking region-enhanced nucleus segmentation network (REU-Net). It uses a combination of serial and parallel stacking of three U-shaped structures to build a multi-tasking architecture, the basic module adopts attention U-Net and adopts two auxiliary tasks, namely contour extraction and coarse segmentation. After the image is input into the model, the coarse split branch decoder and the contour extract the predicted region and contour of the output kernel of the branch decoder, and the kernel suggestion is obtained by integrating the contour of the predicted region and the kernel. The kernel suggestion is multiplied with the input image elements to enhance the significance of the nuclear region and the kernel profile. After that, the region-enhanced image is input into the subdivision branch together with the original image for further segmentation, and the coded features of the different branches are aggregated by hopping connections using attention gates. Finally, the prediction results of the auxiliary task are used to enhance the salient area, and the main task segments the enhanced image. The authors evaluated the method on the private cervical cell dataset HUST and other tissue datasets, conducted a series of ablation experiments in HUSTS and other

existing methods, the multi-task model focused on the nucleus profile, the proposed new feature aggregation strategy allows the network to learn better feature representations, and the authors propose that in future work, we can try to map distances to the network to obtain more accurate segmentation results. Yang et al. [75] proposed a gated context-aware pooling network(GCP-Net) for cervical nucleus segmentation. The proposed U-Net-based GCP-Net consists of a pre-trained ResNet-34 model as an encoder, a gated context-aware pool (GCP) module, and an improved decoder. The GCP module is the main building block of the network to improve the quality of feature learning, which allows GCP-Net to refine the details of the feature map by leveraging multi-scale context gating and global context attention of spatial and texture dependencies, including the decoder block of the global context concern (GCA) residual block to help establish long-range dependencies and global context interactions in the decoder to refine the prediction mask. The authors evaluated the network's performance on private cervical cell and other datasets. GCP-Net is useful for primary cytology screening or triage in practical applications, but it is not precise enough to segment challenging cases. Rasheed et al. [76] introduced the Cervical-UNet (C-UNet) framework for nucleus segmentation in cervical cell smear images characterized by overlaps and blurriness. The architecture begins with an encoder that incorporates the Cross-Scale Features Integration (CSFI) module, facilitating the fusion of extracted features through a cross-scale weighted ensemble methodology. Subsequently, these refined features undergo optimization and fusion via dual interconnected decoders to generate precise segmentation masks. The model's efficacy was evaluated on datasets from ISBI 2014 and ISBI 2015. Experimental findings demonstrate the superiority of the proposed approach over existing models such as CGAN (Conditional Generative Adversarial Network), achieving a target-level accuracy of 0.93 on intricate cervical image datasets.

Most neural networks have the characteristic of having a large number of parameters, which can lead to high complexity, long computation time, and other issues. To address this problem, researchers have proposed different solutions, such as processing from the network architecture and image aspects. Sabeena et al. [77] developed a deep convolutional framework using FCDensenet 56 to detect and analyze cell components in cervical smears for segmenting nuclei and cytoplasm from Pap-stained cell images. To achieve accurate results, this method combines two networks, FC-DenseNet 56 and ResNet 101. The authors evaluated the performance of this method on the Herlev dataset. The proposed segmentation method was evaluated using Precision and Dice coefficients, with performance parameters such as accuracy and Dice coefficient greater than 90%, recall rate, and IoU greater than 85%. This method fully uses deep convolutional networks to extract cell components, whereas the DenseNet used has dense block convolutional layers with fewer network parameters than traditional convolutional networks. Luo et al. [78] proposed a dual-supervised sampling network structure, using compressed images instead of original nucleus segmentation images in the supervised-downsampling module and introducing network supervision of up-sampling decoding layers for accurate segmentation through boundary detection. This strategy can significantly reduce the convolution calculation during image feature extraction and ensure the segmentation accuracy. The authors evaluated the performance of this network on ISBI 2014 and other datasets.

In addition, some other basic backbones are used for segmenting cervical nucleus. For example, Battula et al. [79] used EfficientNet based on deep learning for cervical smear image classification and segmentation. To enhance the image, contrast-limited adaptive histogram equalization (CLAHE) was used in this study. Then, features including wavelet, morphological features, and gray level co-occurrence matrix (GLCM) were extracted from the cervical image. These derived features were efficiently trained and tested using EfficientNet

to distinguish between normal and abnormal cervical images. On abnormal cervical images, the SegNet method was used to identify and segment cancerous areas. The authors evaluated the performance of the proposed method on the Herlev dataset. The EfficientNet model used by the authors can achieve better performance with limited computational resources, and by using data augmentation techniques to expand the dataset, the model's generalization ability is improved. Hu et al. [80] propose a Pixel Adaptive Transformer(PATrans) to improve the segmentation performance of nuclei edges on small datasets through adaptive pixel tuning. First, they designed the Continuous Pixel Patch (CPP) to embed rich multi-scale contexts into isolated image patches. This allows PATrans to quickly establish long-range dependencies and accelerate convergence. Second, CPPI is applied in front of the Transformer blocks in stage 1, while CPPII is applied in front of the Transformer blocks in stages 2, 3, and 4. Finally, to address the issue of imbalanced pixel distribution, a Pixel Adaptive Transformer Block (PATB) is designed to replace the Transformer blocks in stages 3 and 4. PATB effectively models the relationships between different pixels across the entire feature map in a data-dependent manner, adaptively reducing the interference of irrelevant pixels through collaborative learning of local features and global dependencies. The authors evaluated the performance of this network on the Herlev, ISBI, and private datasets, achieving finer segmentation edges.

3.2.2 Cytoplasm segmentation

Variants of the U-Net model can perform not only nucleus segmentation but also cytoplasm segmentation on both single-cell and multi-cell images with different modifications. Zhao et al. [81] proposed a progressive growing U-Net (PGU-Net+) model that incorporates two paradigms to extract image features at different scales more independently. Firstly, residual modules are introduced between different scales of the U-Net, enabling the model to learn the approximate shape annotations at coarse scales and the residual between the annotations and the approximate shapes at finer scales. Subsequently, the model is trained starting from the coarsest part and progressively adding more fine-grained parts until the full model is included. When training a finer part, the learning rate of the previous coarser parts is reduced, further ensuring that the model independently extracts information from different scales. The authors evaluated the performance of this model on the Herlev dataset, and adding residual modules to the U-Net architecture improved the performance, particularly in multi-scale information extraction. Wan et al. [82] proposed a novel framework based on deep convolutional neural networks (DCNNs) for the automatic segmentation of overlapping cells. Firstly, the image is scanned using a sliding window approach based on the dual-window method. The TernausNet (a modification of the U-Net architecture) is employed to classify image pixels into the nucleus, cytoplasm, and background, generating candidate cell regions. Subsequently, a synthesis method is utilized to generate cell clusters containing touching or overlapping cells, and an improved DeepLab V2 model is applied for cytoplasm segmentation. Finally, a fully connected conditional random field and distance regularized level set evolution method are employed as post-processing techniques to refine the outer contours of cells. The authors evaluated the proposed method's performance on two public datasets, ISBI 2014 and ISBI 2015, as well as a private dataset. This framework is more sensitive to cell pixels and boundaries, and combining the two post-processing methods achieves more accurate segmentation than using either method alone. The authors primarily focused on evaluating methods for normal cervical cell images by varying cell numbers and overlap levels, with less emphasis on abnormal cervical cells.

In addition, Mask R-CNN, as a backbone, also has applications in cytoplasm segmentation. Chen et al. [83] propose a two-stage framework based on Mask R-CNN to automatically segment overlapping cells. In the first stage, candidate cytoplasm bounding boxes are proposed. In the second stage, pixel alignment is used to refine the boundaries and classify categories. The authors evaluated the performance of the framework on the ISBI2014 and ISBI2015 datasets, and the experimental results showed that the Mask RCNN-based segmentation method is effective in cytology analysis, which requires only a small number of annotated images and any prior knowledge of cell shape without the need for it but is not fine enough for some cytoplasm embedded in cell clumps, especially in real EDF images.

3.2.3 Cell segmentation

The goal of cell segmentation is to predict segmentation masks in a given image so that each pixel is labeled according to its category.In deep learning methods, U-Net and its variants or combinations with traditional methods are commonly used for multi-object segmentation, mostly segmenting the nuclei and cytoplasm in cervical cell images. For example, Huang et al. [84] proposed a new fusion algorithm that combines the convolutional neural network U-Net with an improved level set method to segment overlapping cervical smear cells. First, the convolutional network structure U-Net is used to separate the nucleus, cytoplasm, and background from the cell image. Then, after obtaining the initial boundary of each cell, an improved level set energy to function and distance map are used along with a new shape before updating the boundary of each cell to make it more realistic. The authors evaluated the performance of the proposed fusion algorithm on the public dataset ISBI 2014.

The above semantic segmentation loses fine image features due to pooling operations and convolutional strides, resulting in poor segmentation results. To improve the accuracy of segmentation results, researchers have proposed many attention mechanism models for image semantic segmentation. Attention mechanisms are used to obtain target areas that need to be focused on to obtain more details and key information about the current task in image semantic segmentation. Zhang et al. [85] proposed a new framework based on an attention mechanism-based U-Net (ATTU-Net) network and a graph-based random walk (RW) for extracting individual nucleus and cytoplasm from overlapping cervical cell images. First, ATTU-Net is used to separate nuclei from overlapping cell clusters. Then, based on each detected nucleus, a transformed image is obtained by polar coordinate sampling. Then, ATTU-Net is applied to predict the cytoplasm boundary of the transformed image. Finally, based on graph-based RW, the contour of individual cells is updated. The authors evaluated the performance of this framework on ISBI 2014 dataset. Zhao et al. [86] proposed a lightweight feature attention network (LFANet), which can extract feature information with different richness at different resolutions of objects and can accurately segment nucleus and cytoplasm regions in cervical images. A lightweight feature extraction module is designed as an encoder, combining depthwise separable convolution, residual connection, and attention mechanism to extract rich features from input images. In addition, to accurately recover pixel positions, a feature layer attention module is added, which uses global high-level information as guidance for low-level features to capture channel feature dependencies. The authors evaluated the performance of this network on the Herlev dataset and other datasets. The proposed lightweight feature attention network can segment cervical cells with low computational complexity and solve the problem of parameter redundancy brought by complex network models. On this basis, further optimization or replacement of network structure may lead to further performance improvement. Zhao et al. [87] proposed a multi-task network for segmenting

cervical cells. Firstly, the images are fed into an encoder consisting of a pre-trained ResNet-34 and a context encoding layer. Then, context encoding is added after each merge to extract contextual features, and the obtained features are input to an attention learning module with skip connections. Finally, the outputs of all attention modules are upsampled to the size of the input image, auxiliary task losses are computed, and the net loss is calculated by weighting the losses of the two tasks using prior knowledge. The authors evaluated the performance of this model on the ISBI2014 dataset and other datasets. Furthermore, improvements in the encoder structure of the attention module or replacements and increasing the number of features in the shape loss function may lead to further performance improvements. Li et al. [88] proposed a context-aware information modeling and feature refinement network based on global dependencies and local attention (GDLA). Specifically, the network consists of three main components: global dependency module, spatial attention module, and channel attention module. The global dependency module first computes the weighted average of the feature map's spatial features as global contextual features and aggregates them into each location feature to enhance its representation. The spatial and channel attention modules perform dimension reduction on the channel and spatial dimensions, respectively, and activate them using the sigmoid function to obtain spatial attention maps and channel attention maps. These three modules are computed in parallel and applied to the original feature map to adaptively process the input information. The authors evaluated the performance of this network on the Herlev dataset and, through comparisons and ablation experiments, demonstrated the effectiveness of the proposed method. Regarding the cell nucleus and cytoplasm segmentation metrics, the method outperformed the state-of-the-art channel attention, hybrid attention, and contextual networks, achieving superior segmentation performance compared to most previous advanced methods.

Attention mechanisms assign weights to focus on important parts of input data, helping models better capture long-distance dependencies and directly focus on input parts related to current output. These characteristics enable attention mechanisms to improve model performance in image segmentation tasks. However, whether it is a classic neural network or a variant model that changes different modules,the biggest challenge in medical image processing is the scarcity of publicly available datasets and the small size of these datasets, which results in the inability of models to learn the features of medical images thoroughly [89]. Especially in cervical cell image segmentation, public datasets and picture numbers are too small, so many researchers have studied this issue. The main method is through generative adversarial networks (GANs). For example, Huang et al. [90] proposed a new method based on generative adversarial networks (Cell-GAN) for cervical cell image segmentation. First, Cell-GAN is trained by comparing the difference between generated and annotated single-cell images to learn the probability distribution of cell morphology. Then, the trained Cell-GAN generates a complete single-cell image for each cell located by a guiding factor. Finally, the contour defined by the generated cells defines the segmentation line and uses R-crop recursively on the image using cell size information until the area of the generated cells changes within a small range. The authors evaluated the performance of this method on a private dataset. They achieved good comprehensive segmentation capabilities in cases where cells were highly overlapping, and backgrounds were complex with low failure rates. Hao et al. [91] (2022) proposed an improved deep convolutional network-based method for cervical nucleus and cytoplasm segmentation. The method comprises a cell region proposal and pixel-level segmentation network (CRP-PSN). In the CRP network, online hard example mining (OHEM) and soft non-maximum suppression (Soft NMS) algorithms are combined to address background noise, impurities, and other interferences in trailing images. A generative adversarial network (GAN) algorithm and least squares loss function are employed for the

PSN network to generate cell region segmentation, enhancing cytoplasm segmentation performance. The authors evaluated the performance of this network on the ISBI 2014 dataset. The experimental results demonstrate that Soft NMS and OHEM improve the integrity and confidence of the model in cell region detection, and the improvement of the GAN method on PSN also enhances the accuracy of pixel segmentation.

In addition, embedding soft non-maximum suppression (Soft NMS) algorithms can improve cell region integrity. Long et al. [92] proposed a cervical image object detection and segmentation model based on multi-scale feature fusion based on Mask R-CNN. Using Mask R-CNN as the basic framework during feature extraction, through NAS-FPN neural network structure search, automatically finds the best feature pyramid structure and merges different levels of feature maps, finally using Soft NMS instead of NMS to reduce bounding box mis-deletion.

Instance segmentation aims to identify each instance of the same class by separately detecting and depicting cells displayed in an image. Zhou et al. [93] proposed a new instance relationship network (IRNet) for robust overlapping cell segmentation by studying the interaction of instance relationships. Specifically, an instance relationship module is proposed to construct a unit association matrix for passing information between single-cell instance features. Through collaboration between instances, enhanced features can fully utilize contextual information and improve semantic consistency. At the same time, a sparse constraint deduplication module is proposed to eliminate the deviation between candidate classification and positioning accuracy. The authors constructed the largest cervical Pap smear (CPS) dataset. They evaluated the performance of this network on this dataset, and experimental results proved the effectiveness of mining instance relationships in cervical cell segmentation.

3.2.4 Deep learning methods analysis

Traditional methods are sensitive to image quality and conditions, and the performance of these methods can be greatly affected if the image quality is poor or if there is noise. In addition, for different types of cells and cell tissues, parameters need to be adjusted to suit different conditions, which adds to the complexity of the method. In recent years, deep learning methods have made significant progress in the field of cell segmentation, which can better handle complex image structures, reduce the dependence on manually set features, and improve the accuracy and efficiency of segmentation. As can be seen from Table 3, U-Net and its variant models are recommended to be learned, and attention mechanism, feature fusion, and sampling are also good learning objects. Semantic segmentation reduces the size of the feature map through pooling operation and convolution step size, which loses the fine features of the image, resulting in poor segmentation effect. In order to obtain more details and key information of the current task in image semantic segmentation, Zhang et al. [85] proposed a U-Net (ATTU-Net) network based on the attention mechanism [101] and a graph-based random walk (RW) Li et al. [88] proposed a contextual information modeling and feature refinement network based on global dependence and local attention (GDLA). Although the performance of neural networks is very superior, researchers often ignore the cost of the model in order to improve the accuracy in the early stage, and as the accuracy improves more and more slowly, some researchers have designed models with high accuracy and low parameter quantity, which makes the research content show a development trend of equal emphasis on high precision and light weight, such as the Lightweight Feature Attention Network (LFANet) proposed by Zhao et al. [86] can segment cervical cells with low computational complexity, which solves the problem of parameter redundancy caused by complex network models. LFANet is also a variant model of U-Net, therefore, U-Net and its variant model are the key

Table 3 Research results of deep learning method based on basic network classification

Backbone	Methods Method Abbreviation	Segmentation task Nucleus	 Cytoplasm	Dataset Herlev	 ISBI 2014	 ISBI 2015	 Private	Performance Dice	 Precision	 Recall
FCN	FC-DenseNet [77]	√		√				0.9342	0.9678	0.8768
	SC-FCN [67]	√		√				–	0.9734±0.0154	0.9778±0.0191
	DCNN [91]	√	√		√			–	cytoplasm:0.92 nucleus: 0.986	–
	FCANet [66]	√		√				0.9321	0.9413	–
Mask R-CNN	MR-C+LFCCRF [69]	√		√				–	0.96	0.96
	MR-C+ Multi-scale feature fusion [92]	√			√			–	–	–
	MR-C [83]		√		√	√		0.92	–	–
	IRNet [93]	√					√	–	–	–
U-Net	PGU-Net+ [79]		√	√				–	0.925	–
	Ternausnet [82]		√		√	√		0.93	–	–
	AL-Net [87]	√	√		√			0.7890	–	–
	GCP-Net [75]	√					√	0.880	–	–
	REU-Net [74]	√					√	0.876	–	–
	C-UNet [76]	√				√	√	0.9312	0.93	0.9532
	U-Net + GDLA [88]	√		√				–	0.888±0.11	0.936±0.13
	LFANet [86]	√	√	√				0.9411	0.9301	0.9610
	U-Net+Level set [84]	√	√		√			0.90	–	–
	ATTU-Net+Cut [85]	√	√		√			0.93	–	0.95

Table 3 continued

Backbone	Methods Method Abbreviation	Segmentation task Nucleus	 Cytoplasm	Dataset Herlev	 ISBI 2014	 ISBI 2015	 Private	Performance Dice	 Precision	 Recall
other	BTTFA [65]	√			√			0.91	0.91	–
	EfficientNet [79]	√		√				–	0.9829	–
	Cell-GAN [90]	√	√				√	0.94	–	–
	Double-supervised sampling [78]	√			√			–	–	–
	FC-DenseNet+ResNet [77]	√			√			0.9342	–	0.8768
	PATrans [80]	√		√	√	√	√	Herlev:0.9400 ISBI:0.9406	–	–

Note: This table summarizes the deep learning methods used for segmentation. On the basis of classification based on network structure, We summarised the specific methods of each article reviewed, segmentation task, the datasets used and experimental performance

research networks, and there are some networks that are not used in the image segmentation of cervical cells can be further studied from the large field of image segmentation.

4 Comparative analysis and algorithm recommendations

This chapter mainly includes the comparative analysis of traditional and deep learning methods for different segmentation targets, and gives corresponding algorithm recommendations.

4.1 Analysis and recommendation of nucleus segmentation algorithms

For single-cell datasets, both traditional methods and deep learning-based methods are applicable to nuclear segmentation targets, and there is little difference in accuracy performance, and most of the methods can achieve an accuracy of more than 90. For multi-cell datasets, the performance of each method is different, there are algorithms and networks with good performance and weak performance in traditional methods and deep learning methods. Among the traditional methods, the superpixel, threshold segmentation and clustering method are better, and most of the studies using clustering method to achieve nucleus segmentation are the most popular. The variant models of U-Net and FCN in the deep learning method show well, among which SC-FCN is an instance segmentation method that optimizes performance by adding residual blocks, dense connection blocks, and bottleneck layers, but it is limited by a limited data set and its mask labels. In addition, in order to solve the problems of high complexity and long operation time caused by the large number of parameters used by most neural networks, Luo et al. [78] proposed a dual-supervised sampling network to significantly reduce the convolution calculation and ensure the segmentation accuracy during image feature extraction. Recently, Transformer has shown excellent performance in various visual tasks, making its application in medicine an inevitable trend. However, simply using Transformer for a small-scale cervical nucleus dataset will result in disastrous performance. Thus, Hu et al. [80] propose a Pixel Adaptive Transformer(PATrans) to achieve fine-grained segmentation of fuzzy kernel edges on small datasets through adaptive pixel tuning.

4.2 Analysis and recommendation of cytoplasm segmentation algorithms

There are few traditional methods and neural network methods for cervical cell cytoplasm segmentation tasks, and we can sometimes achieve nucleus segmentation in this step by locating the nucleus first, so it is usually rare to study the methods of cytoplasm tasks alone, but some of the methods used are also worth summarizing and learning, the SLIC superpixel method is more commonly used in traditional methods, and the deep learning methods with better performance such as PGU-Net+ and Ternausnet.

4.3 Analysis and recommendation of cell segmentation algorithms

When performing cell segmentation of cervical cells, it is recommended to use deep learning methods, which can automatically extract deep features from a large amount of data and have strong generalization ability. With the proposal of more and more image segmentation network models, the application prospect of deep learning methods in medical cervical cell image segmentation will be broader. However, in terms of traditional methods, there are

many research methods, materials are abundant, and both aspects need to be understood, and the combination of the two will also bring different effects, such as [64], especially in the refinement of segmentation, that is, the post-processing part, the traditional methods are very effective, among which threshold, superpixel, level set and watershed algorithm are still commonly used algorithms. In addition, the biggest challenge in medical image processing is the scarcity of publicly available datasets and the small size of these datasets, which prevents models from learning the features of medical images thoroughly. Especially in cervical cell image segmentation, the number of publicly available datasets and pictures is too small, so many researchers have studied this problem. A new method based on Generative Adversarial Network (Cell-GAN) was proposed to segment cervical cell images through Generative Adversarial Network (GAN), which has a good comprehensive segmentation ability and a low segmentation failure rate in the case of highly overlapping cells and complex nuclear backgrounds.

5 Conclusion

This article reviews the research done in the field of cervical cell segmentation in the past 9 years, and the literature review section provides a detailed description of various methods of cervical cell segmentation. The summary is as follows:

1) Introduce the main characteristics of public datasets and the segmentation indicators commonly used in the selected papers.
2) On the basis of elaborating the concept of cervical cell image segmentation, the traditional and deep learning-based segmentation methods are introduced. The representative algorithm in each type of method is selected to segment the target context for research and analysis. The performance evaluation is carried out on the commonly used image segmentation dataset.
3) The overall performance of the traditional method and the deep learning method is compared and analyzed by segmentation object classification.
4) Summary and outlook for the future of cervical cell segmentation.

Throughout the development of cervical cell image segmentation in recent years, in the actual segmentation task, researchers flexibly choose segmentation methods according to different application scenarios to obtain the best segmentation effect.

With the continuous development of segmentation technology, the accuracy and speed of cervical cell image segmentation have been significantly improved, but there are still some problems. It can be considered from three dimensions, namely, the cervical cell data set, the segmentation of each target, and the segmentation algorithm itself. Table 4 briefly summarizes the current problems encountered in cervical cell segmentation from these three dimensions. These problems will become a hot spot in future research and have extremely important research value and significance.

This review summarizes the relevant techniques for cervical cell segmentation. It can be observed that methods based on deep learning have been rapidly evolving. However, further research to address the existing challenges remains critically important. The primary challenge facing cervical cell segmentation is the scarcity of publicly available datasets, which hinders comprehensive learning of medical image features by models. Currently, the predominant approach is data augmentation through Generative Adversarial Networks (GAN), leading to derivative models such as CycleGAN and DCGAN. DCGAN has achieved significant success in image generation tasks, including generating realistic facial images and

Table 4 Analysis of current problems

Dimension	Existing problems	Popular directions for research
Dataset	1. Cervical cell segmentation dataset is small. 2.The number of images in the dataset is too small, which makes the model easy to overfitand the model is not robust. enough during the training process 3.The low image quality in the dataset affects the segmentation task.	1. Dig into the availability of datasets for other tasks 2. Data augmentation 3. Image reconstruction
The segmentation of each target	1. Nucleus: Poor nucleus segmentation in overlapping cells 2. Cytoplasm: The cytoplasm boundary is not clear due to the low contrast between the cytoplasm and the complex background 3. Cells: Cell clusters have the characteristics of complex composition and diverse shapes, which bring some difficulties to the segmentation of cell clusters.	1. Instance segmentation study of the nucleus 2. The nucleus is localized first, and then the cytoplasm is divided 3. Incorporate traditional methods for refinement and prior knowledge of shapes
The segmentation algorithm itself	1. The algorithm is complex and takes a long time to run. 2. The performance of the cervical cell instance segmentation algorithm is insufficient.	1. At present, deep learning algorithms are the most popular directions, focusing on solving the parameter redundancy of the model and reducing the convolution calculation. 2. Research on Phase 1 and Phase 2 networks

Note: This table briefly summarizes the current problems encountered in cervical cell segmentation, including three dimensions, existing problems and popular directions for research. Each point in the third column of the popular directions for research corresponds to each point of the existing problems in the second column in each dimension

bedroom scenes, but has not yet been applied to cell images. On the other hand, Cycle-GAN has demonstrated the capability to generate diverse cytological images and annotations. Furthermore, given the inherent blurriness in existing cervical cell image datasets, image reconstruction techniques, such as super-resolution reconstruction and blind deconvolution methods, can be employed to enhance image details and clarity, or to remove image blurriness and noise, contributing to more precise segmentation tasks. It is anticipated that future researchers will devise improved methods, enhance existing technologies, and overcome current model limitations, thus achieving more precise computer-aided image analysis.

Funding This work was supported in part by the Key Research and Development Program of Yunnan Province under grant No. 202203AA080009, the 14th Five-Year Plan for Educational Science of Jiangsu Province under grant No. D/2021/01/39, the Jiangsu Higher Education Reform Research Project under grant No. 2021JSJG143, the Fundamental Research Funds of China for the Central Universities under grant B230205048, and the Postgraduate Research & Practice Innovation Program of Jiangsu Province under grant KYCX23_0685.

Data Availability Data sharing not applicable to this article as no datasets were generated or analyzed during the current study.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Ferlay J, Soerjomataram I, Dikshit R et al (2015) Cancer incidence and mortality worldwide: sources, methods and major patterns in globocan 2012. Int J Cancer 136(5):359–386
2. Chaturvedi AK (2012) Epidemiology and clinical aspects of hpv in head and neck cancers. Head Neck Pathol 6:16–24
3. Saslow D, Solomon D, Lawson HW et al (2012) American cancer society, american society for colposcopy and cervical pathology, and american society for clinical pathology screening guidelines for the prevention and early detection of cervical cancer. Am J Clin Pathol 137(4):516–542
4. Malm P, Balakrishnan BN, Sujathan VK et al (2013) Debris removal in pap-smear images. Comput Meth Programs Biomed 111(1):128–138
5. 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
6. Bengtsson E (2010) Recognizing signs of malignancy—the quest for computer assisted cancer screening and diagnosis systems. In: 2010 IEEE international conference on computational intelligence and computing research, pp 1–6. IEEE, Coimbatore
7. Shen X, Liu X, Chen H (2017) Fast computation of threshold based on multi-threshold otsu criterion. J Electron Inf Technol 39(1):144–149
8. Stutz D, Hermans A, Leibe B (2018) Superpixels: An evaluation of the state-of-the-art. Comput Vis Image Underst 166:1–27
9. Bai X, Sun C, Sun C (2018) Cell segmentation based on fopso combined with shape information improved intuitionistic fcm. IEEE J Biomed Health Inform 23(1):449–459
10. 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
11. Ronneberger O, Fischer P, Brox T (2015) U-net: Convolutional networks for biomedical image segmentation. In: Medical image computing and computer-assisted intervention, pp 234–241
12. 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
13. Sarwar A, Sheikh AA, Manhas J, Sharma V (2020) Segmentation of cervical cells for automated screening of cervical cancer: a review. Artif Intell Rev 53:2341–2379
14. 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

15. Lu Z, Carneiro G, Bradley AP et al (2016) Evaluation of three algorithms for the segmentation of overlapping cervical cells. IEEE J Biomed Health Inform 21(2):441–450
16. Jantzen J, Norup J, Dounias G, Bjerregaard B (2005) Pap-smear benchmark data for pattern classification. Nature inspired smart information systems, pp 1–9
17. Shi F, Xia L, Shan F et al (2021) Large-scale screening to distinguish between covid-19 and community-acquired pneumonia using infection size-aware classification. Phys Med Biol 66(6):065031
18. Dice LR (1945) Measures of the amount of ecologic association between species. Ecology 26(3):297–302
19. Oliveira PHC, Moreira G, Ushizima DM et al (2017) A multi-objective approach for calibration and detection of cervical cells nuclei. In: 2017 IEEE congress on evolutionary computation, pp 2321–2327. IEEE, Donostia
20. Plissiti ME, Nikou C (2012) Overlapping cell nuclei segmentation using a spatially adaptive active physical model. IEEE Trans Image Process 21(11):4568–4580
21. Ngo TA, Carneiro G (2014) Fully automated non-rigid segmentation with distance regularized level set evolution initialized and constrained by deep-structured inference. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 3118–3125. IEEE, Columbus
22. Jaya S, Latha M (2020) Channel based threshold segmentation of multi-class cervical cancer using mean and standard deviation on pap smear images. In: 2020 International conference on electronics and sustainable communication systems, pp 721–726. IEEE, Coimbatore
23. Braga AM, Marques RCP, Medeiros FNS et al (2021) Hierarchical median narrow band for level set segmentation of cervical cell nuclei. Measurement 176:109232
24. Canny J (1986) A computational approach to edge detection. IEEE Trans Pattern Anal Mach Intell 6:679–698
25. Zhao M, Wang H, Han Y et al (2021) Seens: Nuclei segmentation in pap smear images with selective edge enhancement. Futur Gener Comp Syst 114:185–194
26. Chan TF, Vese LA (2001) Active contours without edges. IEEE Trans Image Process 10(2):266–277
27. Bezdek JC (2013) Pattern Recognition with Fuzzy Objective Function Algorithms. Springer, New York
28. Saha R, Bajger M, Lee G (2016) Spatial shape constrained fuzzy c-means (fcm) clustering for nucleus segmentation in pap smear images. In: 2016 International conference on digital image computing: techniques and applications, pp 1–8. IEEE, Gold Coast
29. Saha R, Bajger M, Lee G (2017) Circular shape constrained fuzzy clustering (ciscfc) for nucleus segmentation in pap smear images. Comput Biol Med 85:13–23
30. Huang J, Wang T, Zheng D, He Y (2020) Nucleus segmentation of cervical cytology images based on multi-scale fuzzy clustering algorithm. Bioengineered 11(1):484–501
31. Zeng Z, Chen S, Tang S, Yin L (2015) Unsupervised segmentation of cell nuclei in cervical smear images using active contour with adaptive local region fitting energy modelling. In: 2015 8th International conference on biomedical engineering and informatics, pp 250–254. IEEE, Shenyang
32. Goldstein T, Bresson X, Osher S (2010) Geometric applications of the split bregman method: segmentation and surface reconstruction. J Sci Comput 45:272–293
33. Zhang L, Kong H, Liu S et al (2017) Graph-based segmentation of abnormal nuclei in cervical cytology. Comput Med Imaging Graph 56:38–48
34. Zaki N, Qin WJ, Krishnan A (2023) Graph-based methods for cervical cancer segmentation: Advancements, limitations, and future directions. AI Open
35. Agarwal P, Sao A, Bhavsar A (2015) Mean-shift based segmentation of cell nuclei in cervical pap-smear images. 2015 Fifth national conference on computer vision. pattern recognition, image processing and graphics, pp 1–4. IEEE, Patna
36. Saha R, Bajger M, Lee G (2019) Srm superpixel merging framework for precise segmentation of cervical nucleus. In: 2019 Digital image computing: techniques and applications, pp 1–8. IEEE, Perth
37. Nock R, Nielsen F (2004) Statistical region merging. IEEE Trans Pattern Anal Mach Intell 26(11):1452–1458
38. Wu H-S, Barba J, Gil J (1998) A parametric fitting algorithm for segmentation of cell images. IEEE Trans Biomed Eng 45(3):400–407
39. Harandi NM, Sadri S, Moghaddam NA, Amirfattahi R (2010) An automated method for segmentation of epithelial cervical cells in images of thinprep. J Med Syst 34:1043–1058
40. Tareef A, Song Y, Lee M-Z et al (2015) Morphological filtering and hierarchical deformation for partially overlapping cell segmentation. In: 2015 International conference on digital image computing: techniques and applications, pp 1–7. IEEE, Adelaide
41. Xia H, Jin K, Jiang F, Tran QA (2016) A novel method of cervical cell image segmentation via region merging and slic. In: Proceedings of the 7th symposium on information and communication technology, pp 153–158

42. Achanta R, Shaji A, Smith K et al (2012) Slic superpixels compared to state-of-the-art superpixel methods. IEEE Trans Pattern Anal Mach Intell 34(11):2274–2282
43. Lee H, Kim J (2016) Segmentation of overlapping cervical cells in microscopic images with superpixel partitioning and cell-wise contour refinement. In: Proceedings of the IEEE conference on computer vision and pattern recognition workshops, pp 63–69. IEEE, Las Vegas
44. Phoulady HA, Goldgof DB, Hall LO, Mouton PR (2015) An approach for overlapping cell segmentation in multi-layer cervical cell volumes. The second overlapping cervical cytology image segmentation challenge-IEEE ISBI
45. Nosrati MS, Hamarneh G (2015) Segmentation of overlapping cervical cells: A variational method with star-shape prior. In: 2015 IEEE 12th International symposium on biomedical imaging, pp 186–189. IEEE, Brooklyn
46. Song Y, Zhu L, Qin J et al (2019) Segmentation of overlapping cytoplasm in cervical smear images via adaptive shape priors extracted from contour fragments. IEEE Trans Med Imaging 38(12):2849–2862
47. Song YY, Zhang A, Zhou JL, Luo Y, Lin ZZ, Zhou T (2024) Overlapping cytoplasms segmentation via constrained multi-shape evolution for cervical cancer screening. Artif Intell Med 148:102756
48. Islam Z, Haque MA (2015) Multi-step level set method for segmentation of overlapping cervical cells. In: 2015 IEEE international conference on telecommunications and photonics, pp 1–5. IEEE, Dhaka
49. Bhan A, Vyas G, Mishra S (2016) Supervised segmentation of overlapping cervical pap smear images. In: 2016 International conference on signal processing and communication, pp 225–228. IEEE, Noida
50. Soille P (1999) Morphological Image Analysis: Principles and Applications, vol 2. Springer, Berlin
51. Guan T, Zhou D, Liu Y (2014) Accurate segmentation of partially overlapping cervical cells based on dynamic sparse contour searching and gvf snake model. IEEE J Biomed Health Inform 19(4):1494–1504
52. Zhao L, Li K, Wang M et al (2016) Automatic cytoplasm and nuclei segmentation for color cervical smear image using an efficient gap-search mrf. Comput Biol Med 71:46–56
53. Tareef A, Song Y, Huang H et al (2018) Multi-pass fast watershed for accurate segmentation of overlapping cervical cells. IEEE Trans Med Imaging 37(9):2044–2059
54. 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
55. Nisar H, Wai LY, Hong LS (2017) Segmentation of overlapping cells obtained from pap smear test. In: 2017 IEEE life sciences conference, pp 254–257. IEEE, Sydney
56. Khan MB, Nisar H, Aun NC, Lo PK (2016) Iterative region based otsu thresholding of bright-field microscopic images of activated sludge. In: 2016 IEEE EMBS conference on biomedical engineering and sciences, pp 533–538. IEEE, Kuala Lumpur
57. Ragothaman S, Narasimhan S, Basavaraj MG, Dewar R (2016) Unsupervised segmentation of cervical cell images using gaussian mixture model. In: Proceedings of the IEEE conference on computer vision and pattern recognition workshops, pp 70–75. IEEE, Las Vegas
58. Tareef A, Song Y, Cai W et al (2017) Automatic segmentation of overlapping cervical smear cells based on local distinctive features and guided shape deformation. Neurocomputing 221:94–107
59. Ushizima DM, Bianchi AGC, Carneiro CM (2015) Segmentation of subcellular compartments combining superpixel representation with voronoi diagrams. Technical report, Lawrence Berkeley National Lab.(LBNL), Berkeley, CA (United States)
60. Phoulady HA, Goldgof DB, Hall LO, Mouton PR (2016) A new approach to detect and segment overlapping cells in multi-layer cervical cell volume images. In: 2016 IEEE 13th international symposium on biomedical imaging, pp 201–204. IEEE, Prague
61. Phoulady HA, Goldgof D, Hall LO, Mouton PR (2017) A framework for nucleus and overlapping cytoplasm segmentation in cervical cytology extended depth of field and volume images. Comput Med Imaging Graph 59:38–49
62. Wang R, Lei T, Cui R et al (2022) Medical image segmentation using deep learning: A survey. IET Image Process 16(5):1243–1267
63. Basu A, Senapati P, Deb M, Rai R, Dhal KG (2024) A survey on recent trends in deep learning for nucleus segmentation from histopathology images. Evol Syst 15(1):203–248
64. Zhang L, Sonka M, Lu L et al (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, pp 406–409. IEEE, Melbourne
65. Zhang J, Liu Z, Du B et al (2019) Binary tree-like network with two-path fusion attention feature for cervical cell nucleus segmentation. Comput Biol Med 108:223–233
66. Cheng J, Tian S, Yu L et al (2020) Fully convolutional attention network for biomedical image segmentation. Artif Intell Med 107:101899
67. Hussain E, Mahanta LB, Das CR et al (2020) A shape context fully convolutional neural network for segmentation and classification of cervical nuclei in pap smear images. Artif Intell Med 107:101897

68. Chen JZ, Huang Q, Chen YL, Qian LY, Yu CY (2023) Enhancing nucleus segmentation with haru-net: A hybrid attention based residual u-blocks network. arXiv:2308.03382 [eess.IV]
69. Liu Y, Zhang P, Song Q et al (2018) Automatic segmentation of cervical nuclei based on deep learning and a conditional random field. IEEE Access 6:53709–53721
70. Zhao J, Li Q, Li X et al (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, pp 1514–1518. IEEE, Venice
71. Chen J, Lu Y, Yu Q et al (2021) Transunet: Transformers make strong encoders for medical image segmentation. arXiv:2102.04306
72. Dosovitskiy A, Beyer L, Kolesnikov A et al (2020) An image is worth 16x16 words: Transformers for image recognition at scale. arXiv:2010.11929
73. Liu Z, Lin Y, Cao Y et al (2021) Swin transformer: Hierarchical vision transformer using shifted windows. In: Proceedings of the IEEE/CVF international conference on computer vision, pp 10012–10022
74. Qin J, He Y, Zhou Y et al (2022) Reu-net: Region-enhanced nuclei segmentation network. Comput Biol Med 146:105546
75. Yang G, Huang J, He Y et al (2022) Gcp-net: A gating context-aware pooling network for cervical cell nuclei segmentation. Mob Inf Syst 2022
76. Rasheed A, Shirazi SH, Umar AI, Shahzad M, Yousaf W, Khan Z (2023) Cervical cell's nucleus segmentation through an improved unet architecture. PloS One 18(10)
77. Sabeena K, Gopakumar C, Thampi R (2020) An improved deep convolutional model for segmentation of nucleus and cytoplasm from pap stained cell images. In: 2020 6th International conference on advanced computing and communication systems, pp 1274–1278. IEEE, Coimbatore
78. Luo D, Kang H, Long J et al (2022) Dual supervised sampling networks for real-time segmentation of cervical cell nucleus. Comp Struct Biotechnol J 20:4360–4368
79. Battula KP, Chandana BS (2022) Deep learning based cervical cancer classification and segmentation from pap smears images using an efficientnet. Int J Adv Comput Sci Appl 13(9)
80. Hu HX, Zhang JY, Yang TJ, Hu Q, Yu Y, Huang Q (2024) Patrans: Pixel-adaptive transformer for edge segmentation of cervical nuclei on small-scale datasets. Comput Biol Med 168:107823
81. Zhao J, Dai L, Zhang M et al (2020) Pgu-net+: progressive growing of u-net+ for automated cervical nuclei segmentation. In: Multiscale multimodal medical imaging: first international workshop, pp 51–58. Springer, Cham
82. Wan T, Xu S, Sang C et al (2019) Accurate segmentation of overlapping cells in cervical cytology with deep convolutional neural networks. Neurocomputing 365:157–170
83. Chen J, Zhang B (2021) Segmentation of overlapping cervical cells with mask region convolutional neural network. Comput Math Method Med 2021
84. Huang Y, Zhu H, Wang P, Dong D (2019) Segmentation of overlapping cervical smear cells based on u-net and improved level set. 2019 IEEE international conference on systems, man and cybernetics, pp 3031–3035. IEEE, Bari
85. 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
86. Zhao Y, Fu C, Xu S et al (2022) Lfanet: Lightweight feature attention network for abnormal cell segmentation in cervical cytology images. Comput Biol Med 145:105500
87. Zhao J, He Y-J, Zhao S-Q et al (2021) Al-net: Attention learning network based on multi-task learning for cervical nucleus segmentation. IEEE J Biomed Health Inform 26(6):2693–2702
88. Li G, Sun C, Xu C et al (2022) Cervical cell segmentation method based on global dependency and local attention. Appl Sci 12(15):7742
89. Qian LY, Huang Q, Chen YL, Chen JZ (2023) A voting-stacking ensemble of inception networks for cervical cytology classification. arXiv:2308.02781 [cs.CV]
90. Huang J, Yang G, Li B et al (2021) Segmentation of cervical cell images based on generative adversarial networks. IEEE Access 9:115415–115428
91. Hao X, Pei L, Li W et al (2022) An improved cervical cell segmentation method based on deep convolutional network. Math Probl Eng 2022
92. Long M, Liang G, Zheng Y et al (2021) Cervical cell tct image detection and segmentation based on multi-scale feature fusion. 2021 IEEE 5th advanced information technology, vol 5. electronic and automation control conference, pp 192–196. IEEE, Chongqing
93. Zhou Y, Chen H, Xu J et al (2019) Irnet: Instance relation network for overlapping cervical cell segmentation. In: Medical image computing and computer assisted intervention, pp 640–648. Springer, Cham