



## Article

# Automated Cervical Cancer Screening Framework: Leveraging Object Detection and Multi-Objective Optimization for Interpretable Diagnostic Rules

Weijian Ye  and Binghao Dai \* 

School of Computer Science and Engineering, South China University of Technology, Guangzhou 510006, China; yeweijian137@gmail.com

\* Correspondence: dbinghao37@gmail.com

**Abstract:** Cervical cancer is one of the most common malignant tumors, with high incidence and mortality rates. Recent studies mainly adopt Artificial Intelligence (AI) models to detect cervical cells. Yet, due to the imperceptible symptoms of cervical cells, there are three problems that may hinder the performance of the existing approaches: (a) poor quality of the whole-slide image (WSI) performed on cervical cells may lead to undesirable performance; (b) several types of abnormal cervical cells are involved in the progression of cervical cells from normal to cancer, which requires extensive clinical data for training; and (c) the diagnosis of the WSI is medical-rule-driven and requires the AI model to provide interpretability. To address these issues, we propose an integrated automatic cervical cancer screening (IACCS) framework. First, the IACCS framework incorporates a quality assessment module utilizing binarization-based cell counting and a Support Vector Machine (SVM) approach to identify fuzzy regions, ensuring WSI suitability for analysis. Second, to overcome the data limitations, the framework employs data enhancement techniques alongside incremental learning (IL) and active learning (AL) mechanisms, allowing the model to adapt progressively and learn efficiently from new data and expert feedback. Third, recognizing the need for interpretability, the diagnostic decision process is modeled as a multi-objective optimization problem. A multi-objective optimization algorithm is used to generate a set of interpretable diagnostic rules that offer explicit trade-offs between sensitivity and specificity. Extensive experiments demonstrate the effectiveness of the proposed IACCS framework. Applying our comprehensive framework yielded significant improvements in detection accuracy, achieving, for example, a 6.34% increase in  $mAP_{50:95}$  compared to the baseline YOLOv8 model. Furthermore, the generated Pareto-optimal diagnostic rules provide superior and more flexible diagnostic options compared to traditional manually defined rules. This research presents a validated pathway towards more robust, adaptable, and interpretable AI-assisted cervical cancer screening.

**Keywords:** cervical lesion cell detection; active learning; incremental learning; multi-objective optimization; automatic assisted diagnosis



Academic Editor: Heung-Il Suk

Received: 8 April 2025

Revised: 8 May 2025

Accepted: 12 May 2025

Published: 15 May 2025

**Citation:** Ye, W.; Dai, B. Automated Cervical Cancer Screening Framework: Leveraging Object Detection and Multi-Objective Optimization for Interpretable Diagnostic Rules.

*Electronics* **2025**, *14*, 2014.

<https://doi.org/10.3390/electronics14102014>

electronics14102014

**Copyright:** © 2025 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license

(<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Cervical cancer is a common and fatal cancer that affects women's health and ranks fourth in the world in terms of incidence and mortality. It was estimated that, in 2020, there would be approximately 600,000 newly diagnosed cases and more than 340,000 deaths worldwide [1,2]. Its incidence and mortality rates vary across countries and regions and are related to factors such as the quality of local healthcare services, the extensiveness of

screening and prevention programs, and the living habits and environmental conditions of the population [3].

Initially, cervical cancer may not present obvious symptoms, but, when the disease progresses, it may cause problems such as abnormal vaginal bleeding, increased vaginal discharge, and pelvic pain. Early detection is critical for effective treatment and improved prognosis [4]. Strong evidence shows that regular cervical cancer screening is effective in preventing and treating this disease, making it one of the most manageable and treatable types of cancer.

With the development of AI [5,6], more and more AI models are being applied to cervical cytology screening [7]. AI can effectively reduce the burden on cytopathologists and improve diagnostic efficiency, providing automated screening for early assisted diagnosis. However, the transition to AI-driven diagnosis using digitized whole-slide images (WSIs) faces several significant challenges:

(1) WSI Quality: Variability in slide preparation, staining, and scanning can result in poor-quality WSIs (e.g., with uneven staining, artifacts, or insufficient cell density). This variability can seriously affect the reliability of AI analysis, potentially leading to missed detections or misclassifications.

(2) Data Scarcity and Imbalance: Currently, high-quality large-scale annotated datasets for cervical cytology are limited. This scarcity, coupled with the challenge of collecting sufficient examples of diverse abnormal cell types (especially rare but critical high-grade lesions), makes building balanced datasets difficult and hinders the training of robust and generalizable AI models. For example, the significant class imbalance observed in the instance distribution detected by the baseline model (YOLOv8) in our study vividly illustrates this challenge. This initial distribution, where common findings or lower-grade lesions vastly outnumber rarer high-grade abnormalities, highlights the data imbalance problem addressed by our subsequent strategies.

(3) Lack of Interpretability: Clinical diagnosis in cytopathology often relies on established rule-based reasoning. The “black-box” nature of many deep learning models makes it difficult for clinicians to understand and trust their predictions, posing a barrier to adoption.

To address these challenges, we propose an integrated automatic cervical cancer screening (IACCS) framework. This framework aims to provide a robust, adaptable, and interpretable solution for AI-assisted screening. The main contributions of this study include the following:

(1) Enhanced Quality Control: IACCS uses a binarization-based method for cell counting and an SVM approach to identify fuzzy image areas, enabling automated assessment of WSI quality and filtering out unreliable data inputs.

(2) Adaptive Learning Strategy: Leveraging data enhancement, an incremental learning mechanism and an active learning mechanism are proposed to progressively adjust the model parameters. This allows the model to continuously improve by learning from new data and expert feedback, addressing data scarcity and imbalance issues.

(3) Interpretable Rule Generation: A multi-objective evolutionary algorithm is utilized to generate interpretable diagnostic rules that offer explicit trade-offs between specificity and sensitivity, addressing the need for transparency in AI-based diagnostic support.

The remainder of the paper is organized as follows. Section 2 provides an overview of the background, while Section 3 outlines the related work. Section 4 provides a detailed account of the implementation of IACCS. Section 5 outlines the experimental design, which validates the effectiveness of IACCS. The final section of this study, Section 6, provides a conclusion.

## 2. Background

Cervical cancer is a malignant tumor primarily caused by HPV (Human Papillomavirus), mainly occurring in the cervix. It is divided into two main types: squamous cell carcinoma and adenocarcinoma, with squamous cell carcinoma being the most common. This cancer has a prolonged precancerous stage characterized by significant changes in cell nuclei and irregular nuclear membrane contours.

Cervical cytology screening is a widely used procedure effective in detecting early precancerous lesions, which is crucial for preventing cervical cancer [8]. It is recommended for women to undergo this screening regularly, using methods such as Pap smears or liquid-based cytology testing [9].

With the advancement in imaging and digital processing technologies, cervical cancer screening samples can now be converted into digital slides through pathology scanners, greatly transforming the method of cervical cancer screening. Digital pathology reduces the workload of pathologists, improves diagnostic efficiency, and enables rapid and accurate examination of cervical cells [10,11].

The Bethesda System (TBS), introduced in 1988, marked a significant advancement in cervical cytology reporting [12]. Abnormal cells can be classified into several types according to latest TBS [13], including the following:

Atypical Squamous Cells of Undetermined Significance (ASC-US) in cervical cytology screening indicate possible low-grade squamous intraepithelial lesions (LSILs), characterized by unusual cytological features with an unspecified cause.

Atypical Squamous Cells-Cannot Exclude High-grade Squamous Intraepithelial Lesion (ASC-H) indicates a possible high-grade squamous intraepithelial lesion (HSIL) but remains undiagnosed.

Low-grade squamous intraepithelial lesion (LSIL) involves mild cytological abnormalities, mainly appearing in mature intermediate or superficial squamous cells. The features include nuclear enlargement, hyperchromasia, inconspicuous nucleoli, binucleation or multinucleation, and increased vacuolated cells, typically associated with HPV infection.

High-grade squamous intraepithelial lesion (HSIL) mainly occurs in immature basal parabasal cells or basal cells, presenting as sheets, single cells, or clustered dark crowded groups. The features include nuclear enlargement with decreased cytoplasmic volume, increased nuclear-to-cytoplasmic ratio, common nucleolar absence, and irregular nuclear membranes.

Squamous cell carcinoma (SCC) is the primary type of cervical cancer, accounting for approximately 90% of cases, often associated with HPV infection. Its cytological features include pleomorphic deeply staining nuclei, irregular nuclear clearing, multiple irregular nucleoli, cell keratinization, and keratin fragments.

In large-scale cervical cytology screening projects targeting the general population, the number of abnormal squamous cell cases far exceeds that of abnormal glandular cell cases, with ASC-US, LSIL, ASC-H, and HSIL being the four most common types. ASC-US and LSIL lesions typically occur in superficial or intermediate cells, while ASC-H and HSIL lesions typically occur in parabasal and basal cells. Our proposed IACCS is primarily based on these five types of lesions for auxiliary diagnosis.

## 3. Related Work

In cytopathology, conducting an analysis of the entire WSI remains a significant and complex task. For cervical cytology screening, rapid searching and accurate localization of suspicious abnormal cervical cells are crucial to ensuring the accuracy of screening. In the field of computer vision, object detection models have been widely studied and applied [14]. These models not only accurately locate objects but also simultaneously

predict their categories. In the detection of abnormal cervical cells, these object detection models have made significant progress [15], providing reliable tools for the rapid and effective detection of abnormal cervical cells [16–18].

The application of these methods is closely associated with a large amount of data, which means that insufficient training data may severely affect the performance of deep learning models. One challenge in applying deep learning to cervical cytology screening is the lack of effectively annotated data. Due to limited labeled data, the available datasets are consequently restricted, making it difficult for deep learning models to obtain sufficient training and leading to overfitting issues. Although transfer learning [19] can compensate for certain data problems, it is still crucial to construct high-quality cervical cytometry datasets in complex clinical settings. Extensive cervical cancer screening involves a vast female population, making it difficult to collect a sufficient number of rare and severe cases, resulting in dataset imbalances [20]. To address this issue, we introduce an incremental learning [21,22] mechanism in IACCS, which progressively learns based on new data without retraining the entire model. This method allows the model to continuously learn and quickly adapt to new contexts and data distributions.

An ensemble of multiple models is the most direct way to achieve integrated deep learning [20,23]. The key feature of ensemble learning is the diversity of individual networks, and various integration strategies can assist basic models in achieving better performance. Inspired by ensemble learning, we introduce an active learning [24] mechanism into IACCS. In complex clinical scenarios, once the object detection model locates and predicts an object, we introduce a classification model for that predicted object. This classification model, through active learning based on expert-reviewed results, is used to distinguish abnormal cervical cells from normal cells. The active learning mechanism allows the model to selectively pick the most informative samples, focusing on those the model deems most challenging or uncertain.

Despite many studies applying deep learning methods to cervical cell studies, research on the interpretation and diagnostic application of model results is relatively limited. With the emergence of the first deep-learning-based WSI analysis method in cervical cell screening, automated cervical cell screening has entered a comprehensive WSI analysis stage. Unlike the above methods, we propose for the first time the use of multi-objective optimization [25] to find interpretable diagnostic rules.

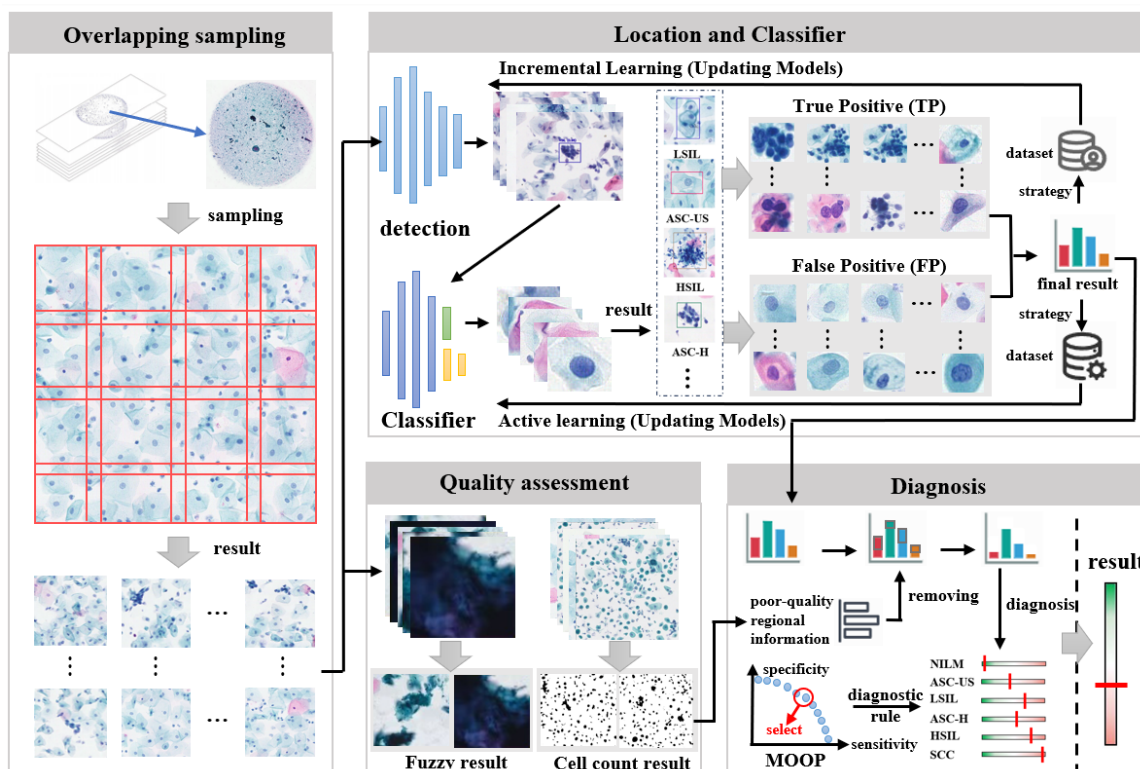
## 4. Proposed Method

In this section, we first outline the integrated automated cervical cancer screening framework we propose. Subsequently, we delve into the discussion of the evaluation module used for automatically assessing the quality of WSI. Following that, we will introduce the dataset established for abnormal squamous or glandular cells and provide a detailed description of the data processing strategies we have employed. Finally, we introduce a multi-objective optimization algorithm for generating diagnostic rules that have different trade-offs among sensitivity and specificity.

### 4.1. Framework Overview

The IACCS framework is illustrated in Figure 1. When there are new WSIs awaiting screening, the WSI is first subjected to preprocessing. This preprocessing involves removing edge portions to enhance detection efficiency and partitioning the WSI for subsequent detection. Subsequently, the quality of the WSI is assessed, including counting the number of cells and determining fuzzy areas. Meanwhile, the partitions are sequentially fed into a pre-trained object detection model to detect abnormal cervical cells. Next, the detected results of abnormal cervical cells are input into a classification model to further determine

if they are abnormal. An expert will review these results, and data from rare and severe cases will be added to the object detection dataset for incremental learning, while cases of model misjudgment will be added to the classification dataset for active learning. Finally, a diagnostic rule for generating preliminary diagnosis results is formulated based on multi-objective optimization, considering various types and quantities of abnormal cervical cells.



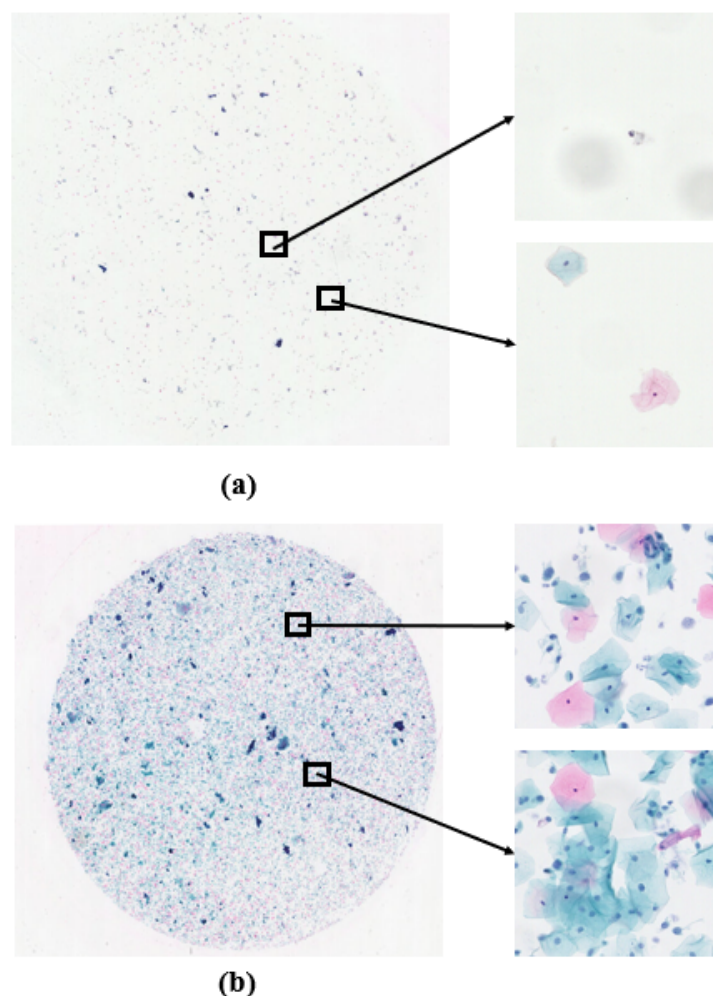
**Figure 1.** Integrated automated cervical cancer screening framework.

#### 4.2. Quality Assessment

In the quality assessment module, by statistically analyzing the cell count and determining the presence of fuzzy regions in the WSI, potential issues in the preparation process can be identified. This process contributes to enhancing the quality of WSIs, thereby bolstering the accuracy and reliability of subsequent analyses.

##### 4.2.1. WSI Cell Count Statistics

Figure 2a shows the case of low cell density, while Figure 2b represents the normal situation. Conducting cell counting on the WSIs not only helps to evaluate potential flaws in the sample preparation process, such as uneven cell dispersion or low cell density, but also serves as a reminder for operators to strengthen technical controls and quality management during preparation. Leveraging the characteristic of cell staining, where cell nuclei appear darker compared to other parts, we utilize traditional image analysis methods to perform cell counting. The main steps include, first, converting the color image to a grayscale image and applying erosion and dilation operations to remove noise; then identifying connected regions in the image through thresholding and Gaussian filtering, representing potential cell nucleus regions; next, filtering out potential cells based on the area of connected regions, excluding small-area noise; and, finally, counting the remaining connected regions to determine the number of cells in the image.



**Figure 2.** Cell count: (a) the case of low cell density; (b) the normal situation.

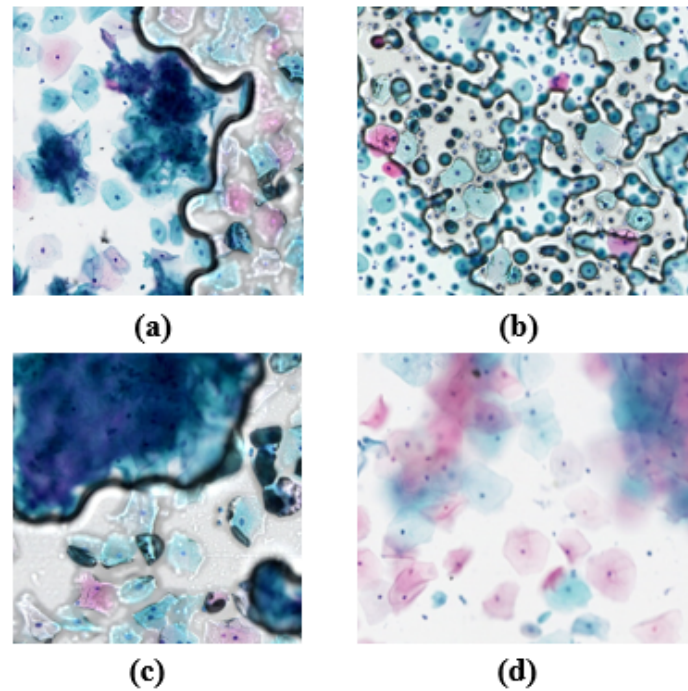
#### 4.2.2. Fuzzy Region Determination

The presence of fuzzy regions may indicate poor image quality, possibly due to issues with imaging equipment (such as poor focus, vibration, or movement) or improper sample preparation (such as uneven tissue slicing or uneven staining). By determining fuzzy regions in WSIs, it is possible to effectively improve the accuracy and reliability of subsequent analyses. The fuzzy regions, as shown in Figure 3, prompted the development of a method for determining image blurriness based on gradient features and Support Vector Machine (SVM). This method first converts the image into a grayscale image. To evaluate image clarity, four gradient features are employed as follows:

The equation for calculating the Brenner gradient is as follows:

$$B = \sum_{x,y} [f(x+2,y) - f(x,y)]^2 \quad (1)$$

where  $B$  represents the calculation result of the Brenner gradient, serving as an indicator of the overall sharpness of the image.  $x$  represents the horizontal coordinate, and  $y$  represents the vertical coordinate.  $f(x,y)$  denotes the intensity or grayscale value of the pixel at coordinates  $(x,y)$ . For each pixel in the image, we find the intensity of the pixel two positions to its right and calculate the square of the intensity difference between them. Then, summing up all these squared differences yields a total  $B$ . A larger total  $B$  indicates higher image contrast, typically implying greater image sharpness.



**Figure 3.** Fuzzy regions in WSIs. (a) Boundary effect causing significant one-sided blur/defocus; (b) Generalized blur with artifacts obscuring cellular details; (c) Large, severely out-of-focus/artifactual area obscuring cells; (d) Subtle, generalized soft focus reducing overall sharpness and detail.

The equation for calculating the Laplacian gradient is as follows:

$$L(x, y) = \sum_{i=-1}^1 \sum_{j=-1}^1 k(i, j) \cdot f(x + i, y + j) \quad (2)$$

where  $L(x, y)$  represents the result after processing with the Laplacian operator, typically used in image processing to indicate the edge intensity at position  $(x, y)$ .  $k(i, j)$  is an element in the kernel matrix, where the kernel is a  $3 \times 3$  matrix used to multiply with the corresponding region of the image to compute the edge intensity. By applying the kernel matrix to each pixel and its neighborhood, new pixel values are calculated for detecting edges and textures in the image.

The equation for calculating the SMD is as follows:

$$SMD = \sum_{x,y} (|2f(x, y) - f(x + 1, y) - f(x - 1, y)| + |2f(x, y) - f(x, y + 1) - f(x, y - 1)|) \quad (3)$$

where  $SMD$  stands for Sum of Modified Laplacian, which is a measure of image sharpness.  $x, y$  represent pixel positions on the image.  $f(x, y)$  is the pixel value of the image at coordinates  $x, y$ . This equation evaluates image sharpness by calculating the squared sum of changes (gradients) in pixel values in both horizontal and vertical directions, indicating that higher  $SMD$  values typically imply higher image sharpness or contrast as this indicates larger variations between pixel values.

The equation for calculating the SMD2 is as follows:

$$SMD2 = \sum_{x,y} \left( \sum_{i=-1}^1 \sum_{j=-1}^1 |f(x + i, y + j) - f(x, y)| \right) \quad (4)$$

SMD2 is an enhanced version of the Sum of Modified Laplacian method used to measure image sharpness. Image sharpness is evaluated by calculating the differences between each pixel and its surrounding pixels and summing the absolute values of these differences. Higher SMD2 values typically indicate more edge or texture information in the image, thus higher image sharpness.

The extracted gradient features are normalized to meet the input requirements of the SVM model. The SVM model, trained on fuzzy data, is used to classify whether images are clear or fuzzy. By inputting standardized features, the model can effectively determine whether images are clear or fuzzy.

#### 4.3. Detection of Abnormal Cervical Cells

As mentioned above, given the significantly higher number of abnormal squamous cell cases compared to abnormal glandular cell cases in large-scale cervical cell screening projects, we focus on ASC-US, LSIL, ASC-H, HSIL, and SCC and have accordingly constructed a dataset. We selected initial data, consisting of 500 positive WSIs and 500 negative WSIs. To establish the dataset, we partitioned the WSIs with a 50% overlap ( $640 \times 640$  pixels) and selected the partitions annotated with abnormal cervical cells. Regarding the processing of new WSIs, it follows Algorithm 1, where, in lines 1–10, the WSIs are initially partitioned, followed by detection on each partition. Subsequently, in lines 11–18, the detected results with the same object are merged. Finally, in lines 19–25, each detection result undergoes further classification to obtain the ultimate detection outcome.

---

#### Algorithm 1 Whole-slide images (WSIs) to be inspected

---

**Input:** WSI  $X$ , the height of WSI  $X_h$ , the width of WSI  $X_w$ , the size of the blocks:  $\sigma$ , Overlap:  $\delta$

**Output:** detection results: boxes

```

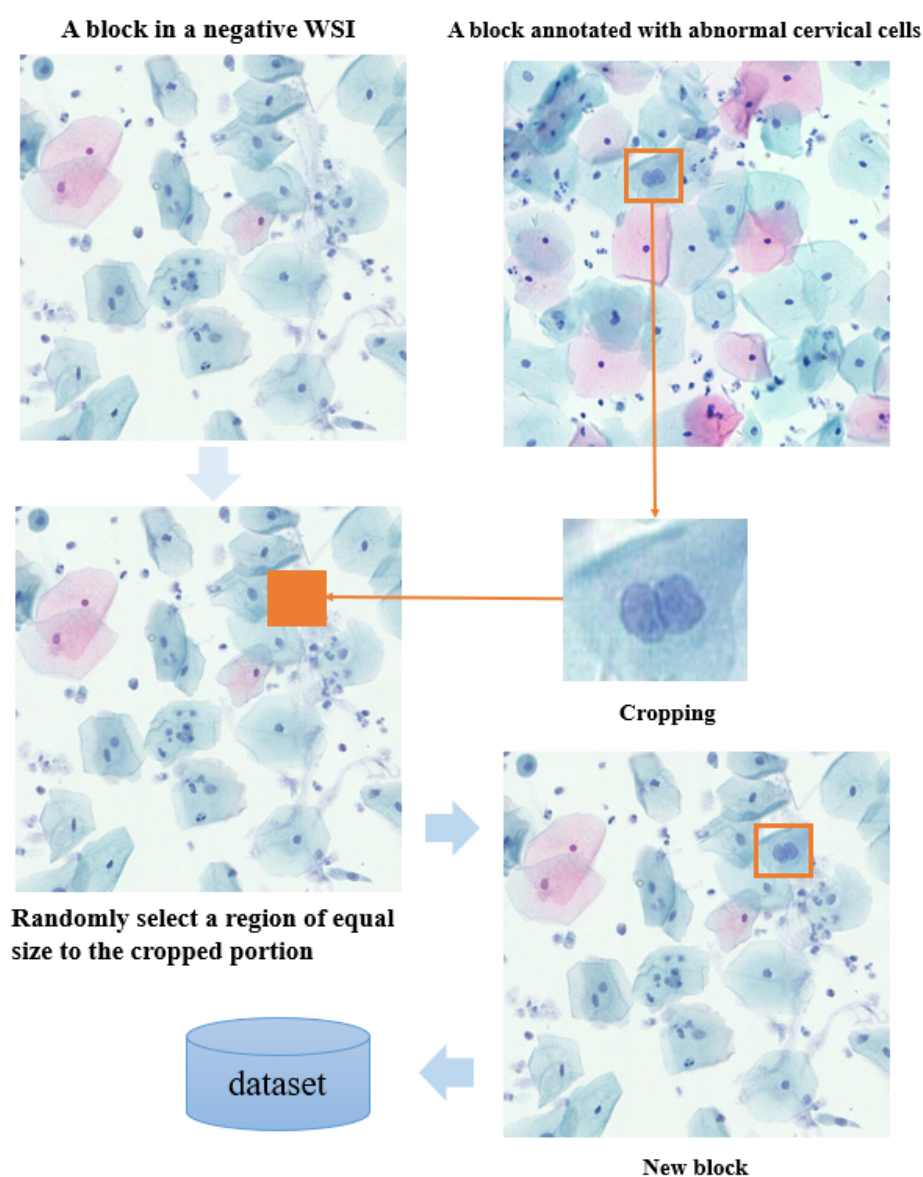
1:  $boxes \leftarrow \{\}, stride \leftarrow \sigma \times \delta, i \leftarrow 0$ 
2: While  $i < X_h$ 
3:    $j \leftarrow 0$ 
4:   While  $j < X_w$ 
5:      $box \leftarrow detect(j, i, \sigma)$ 
6:      $boxes \leftarrow boxes \cup box$ 
7:      $j \leftarrow j + stride$ 
8:   EndWhile
9:    $i \leftarrow i + stride$ 
10: EndWhile
11: For  $m$  in  $boxes$ 
12:   For  $n$  in  $boxes[i + 1 :]$ 
13:     If  $IOU(boxes_m, boxes_n) > \xi$  and  $Label(m, n)$ 
14:        $boxes_n \leftarrow [xmin, ymin, xmax, ymax](m, n)$ 
15:        $delete(boxes_m)$ 
16:     EndIf
17:   EndFor
18: EndFor
19: For  $k$  in  $boxes$ 
20:   If  $classification(boxes_k)$  is abnormal
21:      $save(boxes_k)$ 
22:   Else
23:      $delete(boxes_k)$ 
24:   EndIf
25: EndFor
26: End

```

---

#### 4.3.1. Data Enhancement

In cervical cell detection, positive annotation refers to images containing annotations of abnormal cervical cells, while negative blocks refer to images without abnormal cells. CutMix technology allows the replacement of positive annotations into a random region of negative blocks, as shown in Figure 4. By incorporating positive annotations into the background of negative blocks, training sample diversity is increased, and the balance between positive and negative samples is improved. This helps the model to better learn the features of abnormal cells and reduces training biases caused by class imbalances. Additionally, embedding positive annotations into negative blocks enables the model to learn features of both normal and abnormal cells simultaneously, enhancing model robustness to correctly identify abnormal cells even in the absence of anomalies. By introducing CutMix, the generated mixed images retain the label information of the original images but are not exact duplicates. This helps to reduce the risk of overfitting and improves the model's generalization ability to unseen data.



**Figure 4.** The process of generating new data using CutMix. The orange frames highlight: (top right) the abnormal region for cropping; (middle left) the random region in a negative WSI for replacement; and (bottom right) the pasted abnormal region in the new block.

#### 4.3.2. Incremental Learning

Due to the limited amount of annotated data and the relatively small number of severe cases, the distribution of the dataset is uneven. To address this issue, we introduced an incremental learning mechanism into IACCS. This mechanism enables progressive learning based on new data, gradually integrating data detected as rare or severe cases and reviewed by experts into the original dataset. This allows the model to continuously learn and quickly adapt to new situations and data distributions. Let  $D_n$  represent the current dataset of the model and  $D_p$  denote the set of samples reviewed by experts. Let  $N$  represent the number of categories in the dataset,  $n_i$  denote the number of samples in the  $i$ th category, and  $r_i$  represent the proportion of the  $i$ th category.

When relatively small proportions of categories are detected, we aim to incorporate these new data into the original dataset for learning. For each category  $i$ , we can calculate its proportion in the new dataset,  $r_{i,new}$ . If  $r_{i,new}$  is less than a certain threshold  $\alpha$ , we add the samples of that category to the set of samples reviewed by experts  $D_p$ , i.e.,

$$D_p = \bigcup_{i=1}^N \{x_j | r_{i,new} < \alpha, x_j \in D_{new}^i\} \quad (5)$$

Here, the union symbol  $\bigcup_{i=1}^N$  indicates that  $D_p$  is formed by collecting all samples  $x_j$  that satisfy the condition ( $r_{i,new} < \alpha$ ) across all relevant categories  $i$  (from 1 to  $N$ ) in the new data  $D_{new}^i$ .  $D_{new}^i$  represents the set of samples belonging to category  $i$  in the new data.

#### 4.3.3. Active Learning

After conducting cervical cell detection on the WSI, we introduced the ResNet50 model to classify cells and determine if they are abnormal. If the ResNet50 model classifies a cell as abnormal, the sample is retained; otherwise, it is removed. To improve the performance of the classification model, we employed active learning mechanism. It selects the most informative samples based on the current model's prediction uncertainty or other metrics and actively requests expert annotations to minimize the number of labeled samples. We chose abnormal cervical cells detected in negative blocks and those confirmed by experts as errors in positive blocks as normal cervical cells. Let  $D_n$  represent the current dataset of the model,  $D_p$  denote the set of positive block samples after expert review,  $D_n^+$  represent the set of abnormal cervical cells detected in negative blocks, and  $D_p^-$  represent the set of abnormal cervical cells confirmed as errors in positive blocks. Then, the new sample set  $D_{new}$  can be expressed as

$$D_{new} = D_n^+ \cup D_p^- \quad (6)$$

where samples in  $D_n^+$  are selected under the condition that the confidence  $O$  is greater than or equal to the threshold  $\beta$ , i.e.,

$$D_n^+ = \{x_i | O(x_i) \geq \beta, x_i \in D_n\} \quad (7)$$

The active learning process utilizes the selection strategy of Equations (6) and (7) to actively learn problematic cases, further classifying and filtering the detection results to improve accuracy.

#### 4.4. Diagnostic Rules

After completing the detection of abnormal cervical cells in WSI, automatic diagnosis based on the detection results is required. This diagnosis is based on the counts of five types of abnormal cervical cells (i.e., five variables) and rules corresponding to different

lesion grades. The evaluation of diagnosis results mainly relies on two objective functions: sensitivity and specificity. The calculation of sensitivity and specificity is as follows:

$$\text{Sensitivity} = \frac{TP}{TP + FN} = \frac{1}{f_1(x)} \quad (8)$$

and

$$\text{Specificity} = \frac{TN}{TN + FP} = \frac{1}{f_2(x)} \quad (9)$$

We take another thousand WSIs and aim for the diagnostic rules to maximize the sensitivity and specificity of the diagnostic results for this batch of WSIs; i.e., the optimization objective is to minimize  $f_1$  and  $f_2$ . Therefore, this problem can be transformed into a multi-objective optimization problem, which satisfies the following:

$$\begin{aligned} \min \quad & y = F(x) = (f_1(x), f_2(x)) \\ \text{subject to:} \quad & TN(x) + FP(x) \neq 0 \\ & TN(x) \neq 0 \\ & TP(x) + FN(x) \neq 0 \\ & TP(x) \neq 0 \\ & x \in [0, N]^5 \end{aligned} \quad (10)$$

We utilize the Pareto dominance relationship to obtain a set of mutually non-dominating solutions. Pareto optimality ensures that specificity (sensitivity) is improved without decreasing sensitivity (specificity). Our problem can be addressed through the following steps to obtain the Pareto-optimal solution set. Firstly, compute the objective function values  $F(x) = (f_1(x), f_2(x))$  for each candidate solution  $x$ , where  $f_1(x)$  and  $f_2(x)$  represent the reciprocals of sensitivity and specificity, respectively. Based on the computed objective function values, categorize all candidate solutions into different non-dominated solution sets. A solution  $x^*$  is considered non-dominated over another solution  $x$  if  $f_1(x^*) \leq f_1(x)$  and  $f_2(x^*) \leq f_2(x)$ . For each non-dominated solution set, select the Pareto-optimal solution set, comprising solutions that cannot be dominated by other solutions. Specifically, from each non-dominated solution set, choose solutions that satisfy the following conditions: there is no other solution  $x$  such that  $f_1(x) < f_1(x^*)$  and  $f_2(x) \leq f_2(x^*)$ ; there is no other solution  $x$  such that  $f_1(x) \leq f_1(x^*)$  and  $f_2(x) < f_2(x^*)$ .

The final set of solutions obtained through the multi-objective optimization algorithm GDE3 [26] is the set of Pareto-optimal solutions, where the solutions show the best sensitivity and specificity. GDE3 was selected due to its demonstrated effectiveness in handling multi-objective problems through its differential evolution mechanism, which can efficiently explore the complex parameter space of diagnostic rules [26]. It is important to note that our proposed framework is designed to be general, and other multi-objective optimization algorithms, such as NSGA-II or SPEA2, could also be readily integrated and applied. In this study, we utilized GDE3 primarily as a means to validate the framework's capability to generate effective and interpretable diagnostic rules. Each solution  $x$  corresponds to a diagnostic rule, where  $x = [x_{SCC}, x_{HSIL}, x_{LSIL}, x_{ASC-H}, x_{ASC-US}]$  includes the quantity of each type of abnormal squamous cell. When diagnosing a new WSI, diagnosis proceeds according to Algorithm 2 and the selected rule  $x$ . In Algorithm 2, the first step is to perform statistical analysis of the detection results of various abnormal cervical cells (as seen in lines 2 to 4). Subsequently, diagnoses are conducted based on the quantity of each type of abnormal squamous cell, ranked by severity from high to low (as seen in lines 6 to 26), ultimately resulting in the graded diagnostic outcome of the WSI.

**Algorithm 2** Automated diagnosis**Input:** detection results *boxes*, diagnostic rule *x***Output:** Positive Grading  $\zeta$ 

```

1:  $\theta_{ASC-US}, \theta_{ASC-H}, \theta_{LSIL}, \theta_{HSIL}, \theta_{SCC} \leftarrow 0$ 
2: For box in boxes
3:    $\theta_{box-class} \leftarrow \theta_{box-class} + 1$ 
4: EndFor
5:  $\zeta \leftarrow NULL$ 
6: If  $\theta_{SCC} > x_{SCC}$ 
7:    $\zeta \leftarrow SCC$ 
8: Else
9:   If  $\theta_{HSIL} > x_{HSIL}$ 
10:     $\zeta \leftarrow HSIL$ 
11:   Else
12:     If  $\theta_{LSIL} > x_{LSIL}$ 
13:        $\zeta \leftarrow LSIL$ 
14:     Else
15:       If  $\theta_{ASC-H} > x_{ASC-H}$ 
16:          $\zeta \leftarrow ASC-H$ 
17:       Else
18:         If  $\theta_{ASC-US} > x_{ASC-US}$ 
19:            $\zeta \leftarrow ASC-US$ 
20:         Else
21:            $\zeta \leftarrow NSIL$ 
22:         EndIf
23:       EndIf
24:     EndIf
25:   EndIf
26: EndIf
27: End

```

**5. Experiments and Analysis**

To validate the effectiveness of our integrated automatic cervical cancer screening framework, we introduce the achieved results from the following three aspects. Firstly, quality assessment is implemented to automatically judge the quality of WSIs. Secondly, we present the situation of the established datasets for abnormal cervical cells, along with the corresponding improvements we made and their effects. Finally, we demonstrate the specificity and sensitivity of the diagnostic rules found through multi-objective optimization.

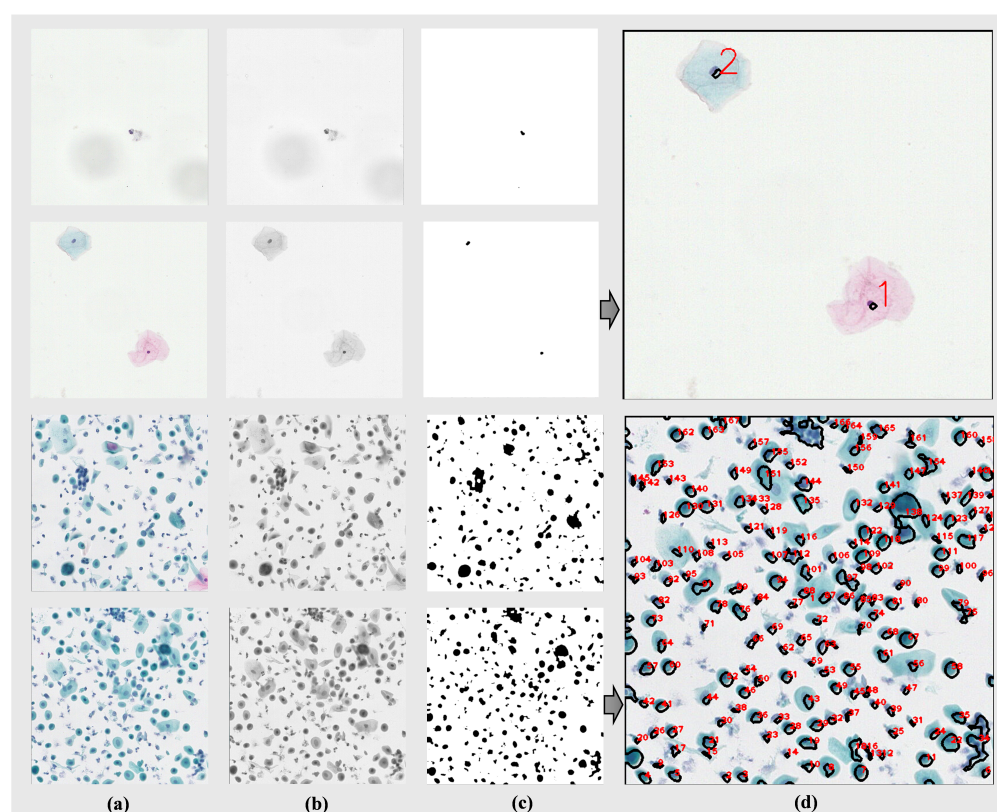
**5.1. WSI Cell Count Statistics**

As shown in Figure 5a, four segmented images from two types of WSIs (those with fewer instances and those with normal cell quantities) are displayed. These segmented images are then converted into grayscale images, as depicted in Figure 5b. Subsequently, we applied erosion and dilation operations to the images using a  $2 \times 2$  matrix of all ones as the kernel, iterating each operation five times to eliminate noise and unnecessary details from the images. Next, the images were thresholded at 155/255 to convert them into binary images, further highlighting the contours of the cells, as shown in Figure 5c. Following this, a  $3 \times 3$  Gaussian filter was applied to the binary images for smoothing, reducing the influence of noise. Finally, by identifying connected components in the images and excluding components with an area less than 25 pixels, we removed smaller noise or non-cellular areas. We then plotted and quantified the connected components, the results of which are illustrated in Figure 5d. Based on the darker color of the cell nuclei compared to the background and cytoplasm, our method can achieve cell count statistics with minimal

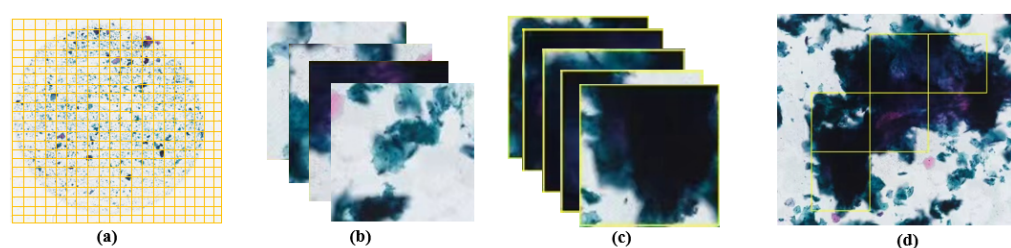
computational resources. This approach not only eliminates the cost of manual annotation but also operates independently of algorithmic models.

### 5.2. Fuzzy Region Determination

The process of determining fuzzy regions is illustrated in Figure 6. Firstly, the WSI is segmented, as shown in Figure 6a. Then, each segment is processed sequentially using four different gradient functions (Brenner, Laplacian, SMD, and SMD2) to extract gradient features from the images. Subsequently, the extracted gradient feature data are normalized, and a pre-trained SVM model is utilized for fuzzy region determination. Figure 6b displays the non-fuzzy segments, while Figure 6c illustrates the fuzzy segments. Following the assessment of all the segments, the fuzzy regions are annotated, as depicted in Figure 6d. When the proportion of fuzzy to WSI was small, we removed the abnormal cervical cells whose detection was in the fuzzy region, and, when the proportion was too large, the WSI was considered to be of poor quality. The determination of the fuzzy region was thus realized.



**Figure 5.** Cell count statistics display: (a) Segmented Whole Slide Image (WSI) patches; (b) Grayscale converted images; (c) Binary images after thresholding, highlighting cell contours; (d) Final cell count results after connected component analysis and noise removal.

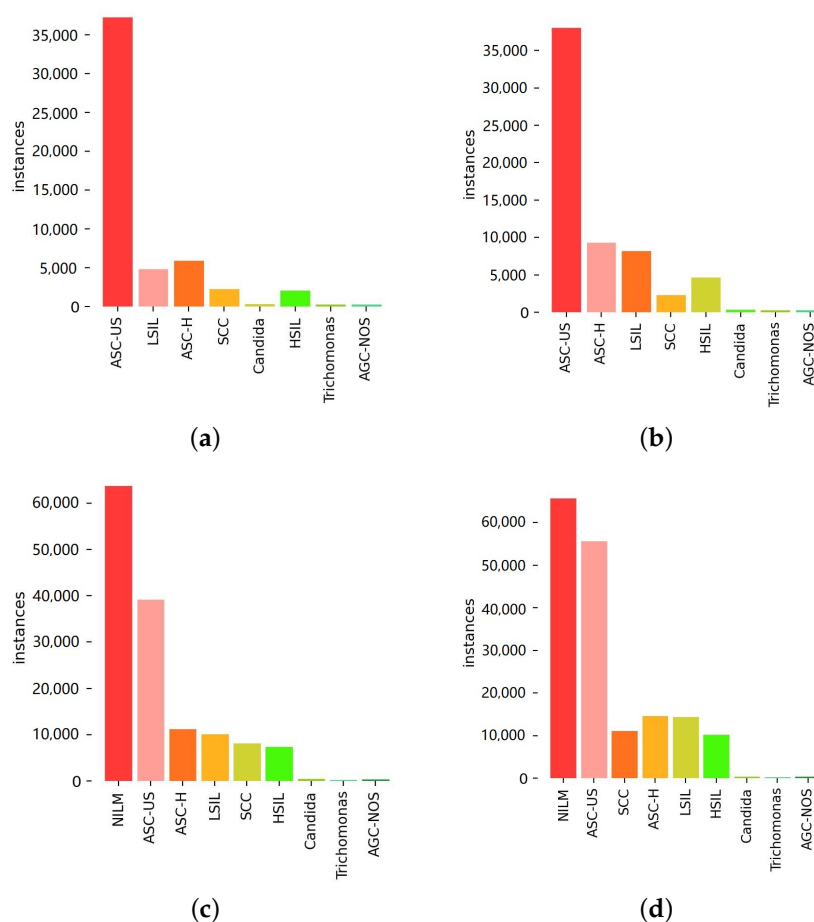


**Figure 6.** Fuzzy region determination display: (a) Segmented Whole Slide Image (WSI); (b) Identified non-fuzzy segments; (c) Identified fuzzy segments; (d) Annotated fuzzy regions on the WSI.

### 5.3. Object Detection

We chose YOLOv8 as the object detection model and created a dataset of 500 positive WSIs with annotations and 500 negative WSIs. We divided the dataset into training, validation, and test sets according to the ratio of 7:2:1. The dataset included abnormal squamous cells (ASC-US, ASC-H, LSIL, HSIL, and SCC), glandular cells (AGC-NOS), microbial infections (Candida and Trichomonas), and normal squamous cells (NILM).

To address the inherent class imbalance and illustrate the impact of our proposed data handling strategies, Figure 7 presents the per-class sample distributions at various stages. Specifically, it shows the distribution for (a) the baseline YOLOv8 dataset, (b) after applying data enhancement (DE), (c) after subsequently incorporating incremental learning (IL), and (d) the final distribution after including active learning (AL). These distributions highlight how our framework progressively adjusts the dataset composition to potentially improve model training on under-represented classes.



**Figure 7.** Per-class sample distributions (y-axis: number of instances) at different stages of the proposed framework. (a) Baseline (YOLOv8); (b) after data enhancement (DE); (c) after DE + incremental learning (IL); (d) after DE + IL + active learning (AL).

We evaluated the model performance using COCO-style average precision (AP) and average recall (AR). Additionally, we employed  $mAP_{50}$  (mean average precision at an IoU threshold of 0.5) and  $mAP_{50:95}$  (mean average precision within the IoU range of 0.5 to 0.95). In our proposed IACCS, we utilized CutMix data augmentation (DE) and introduced incremental learning (IL) mechanisms and active learning (AL) mechanisms.

Table 1 presents the performance comparison of different components of our proposed method. From Table 1, it is evident that adopting DE significantly improves all the metrics compared to the baseline model. Furthermore, further enhancements in  $mAP_{50}$  and

$mAP_{50:95}$  are achieved with the incorporation of *IL* and *AL*, demonstrating their effectiveness. Clearly, leveraging *DE* for negative WSI data utilization and continuously updating new data with *IL* and *AL* can effectively enhance the model performance.

We further validate the proposed strategy on the current state-of-the-art object detection algorithms RT-DETR [27] and YOLOV9 [28]. The results are shown in Table 2, covering the metrics before and after applying the strategy and focusing on the calculation of the performance improvement, for which the improved parts of the performance metrics are labeled. The results show that the proposed strategy can effectively improve the performance of the model.

**Table 1.** Performance of ablation study.

Comparison	AP	AR	$mAP_{50}$	$mAP_{50:95}$
YOLOv8	94.52	87.16	94.3	88.4
YOLOv8+ <i>DE</i>	<b>96.54</b>	90.88	96.1	92.48
YOLOv8+ <i>DE</i> + <i>IL</i>	93.32	<b>95.72</b>	97.24	92.46
YOLOv8+ <i>DE</i> + <i>IL</i> + <i>AL</i>	95.46	95.4	<b>98.06</b>	<b>94.74</b>

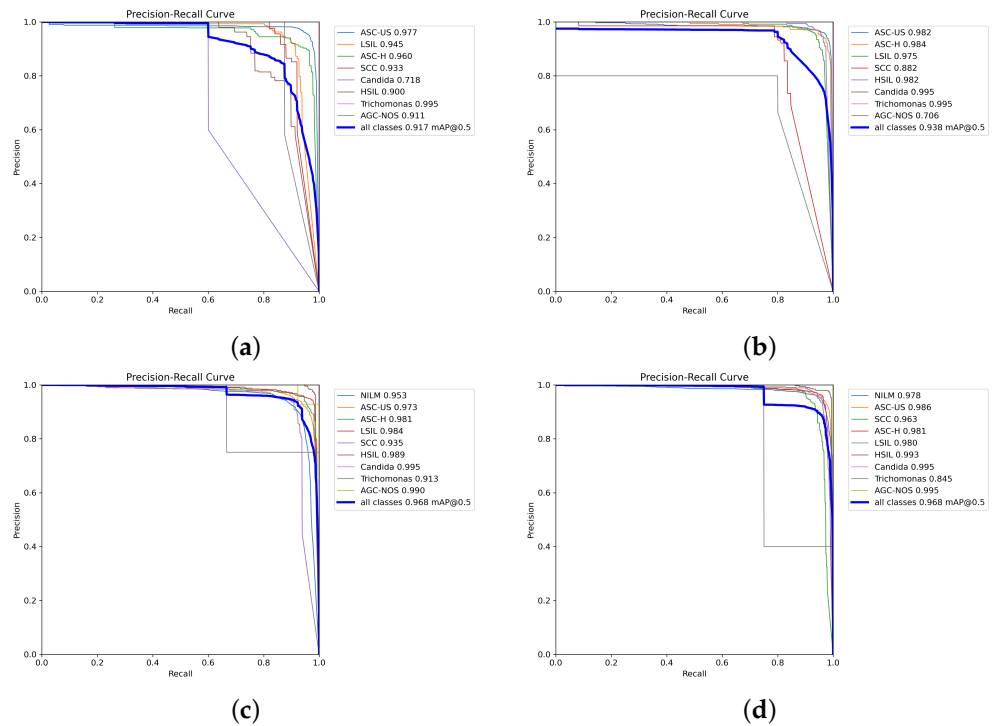
Note: Bold values indicate the best performance achieved for that metric across the different ablation settings.

**Table 2.** Performance of the proposed strategy on state-of-the-art methods. The five rightmost columns are for each lesion category AP.

Method	AP	AR	$mAP_{50}$	$mAP_{50:95}$	ASC-US	LSIL	ASC-H	HSIL	SCC
YOLOv8	94.52	87.16	94.3	88.4	96.2	96.2	90.2	94.5	95.5
YOLOv8 + Our Strategy Improvement	95.46	95.4	98.06	94.74	93.8	96.4	94.4	98.0	94.7
	↑ <b>0.94</b>	↑ <b>8.24</b>	↑ <b>3.76</b>	↑ <b>6.34</b>	↓ 2.4	↑ <b>0.2</b>	↑ <b>4.2</b>	↑ <b>3.5</b>	↓ 0.8
RT-DETR	94.46	92.34	93.88	85.44	95.4	94.9	91.8	95.2	95.0
RT-DETR + Our Strategy Improvement	97.08	95.04	97.18	93.18	96.3	96.6	96.8	98.2	97.5
	↑ <b>2.62</b>	↑ <b>2.7</b>	↑ <b>3.3</b>	↑ <b>7.74</b>	↑ <b>0.9</b>	↑ <b>1.7</b>	↑ <b>5.0</b>	↑ <b>3.0</b>	↑ <b>2.5</b>
YOLOv9	95.76	86.5	93.52	90.02	94.3	96.7	93.7	95.6	98.5
YOLOv9 + Our Strategy Improvement	96.38	90.26	96.68	93.58	96.2	97.4	98.3	92.1	97.9
	↑ <b>0.62</b>	↑ <b>3.76</b>	↑ <b>3.16</b>	↑ <b>3.56</b>	↑ <b>1.9</b>	↑ <b>0.7</b>	↑ <b>4.6</b>	↓ 3.5	↓ 0.6

Note: Arrows ↑ and ↓ indicate performance improvement and degradation, respectively, after applying our strategy. Bold values in the "Improvement" rows highlight these changes.

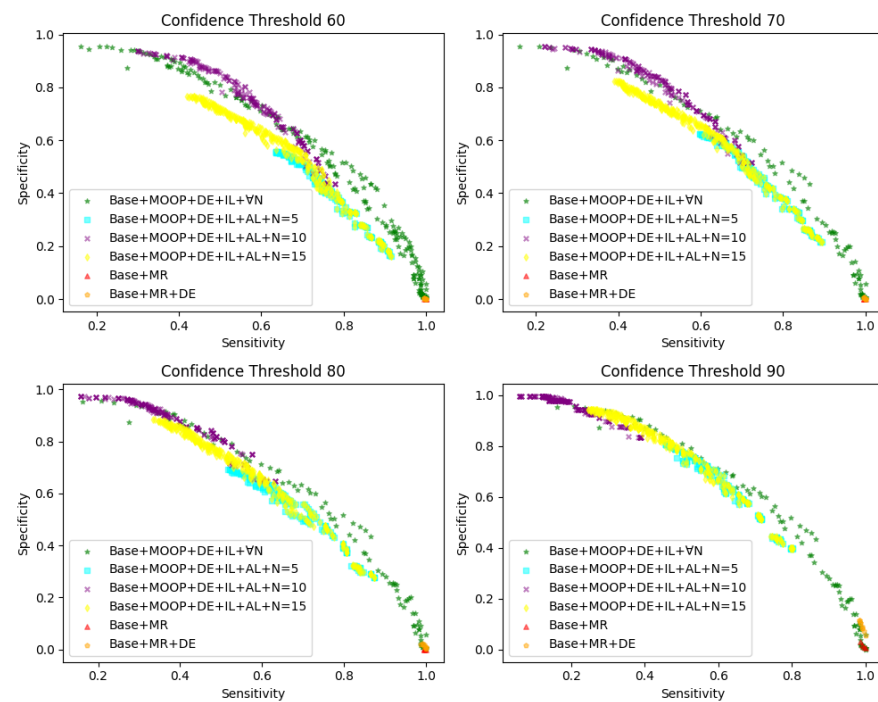
To further assess the model's performance across different confidence thresholds and evaluate its probabilistic outputs, which are crucial in medical diagnosis for risk stratification, we present the precision–recall curves (PRCs) in Figure 8. These curves illustrate the trade-off between precision and recall (sensitivity) at various decision thresholds for each class and the overall model performance. The area under the PRC (PRC-AUC), often summarized by metrics like mean average precision (mAP) at a specific IoU (e.g.,  $mAP@0.5$ , which we report), provides an indication of the model's ability to distinguish between classes across the full range of operating points. Figure 8 shows the PRCs for (a) the baseline YOLOv8 model, (b) after applying data enhancement (DE), (c) after subsequently incorporating incremental learning (IL), and (d) the final PRCs after including active learning (AL). These visualizations demonstrate how our proposed strategies impact the model's diagnostic confidence and its ability to correctly classify instances at varying levels of certainty.



**Figure 8.** Precision–recall curves (PRCs) at different stages of the proposed framework. (a) PRC: baseline (YOLOv8); (b) PRC after data enhancement (DE); (c) PRC after DE + incremental learning (IL); (d) PRC after DE + IL + active learning (AL).

#### 5.4. Diagnostics

We collected 990 WSIs diagnosed by experts under a microscope to establish the diagnostic rules. Our algorithm first detected abnormal cervical cells in the WSI and then identified the Pareto-optimal solution set based on Equation (10) after obtaining the detection results. The Pareto-optimal solution sets under various conditions are shown in Figure 9.



**Figure 9.** Multi-objective optimization results.

For each condition, we diagnosed using different thresholds, namely confidence thresholds of 60, 70, 80, and 90. In Figure 9, *Base* represents the baseline model, *MR* denotes manual rules, and *N* represents the range of parameter  $x$  in multi-objective optimization (MOOP) from  $[0, N]$ , with sensitivity on the horizontal axis and specificity on the vertical axis. For example, as shown in Figure 9, *Base + MR* represents the diagnosis of YOLOv8 test results using manually formulated rules for evaluating its sensitivity and specificity, while *Base + MOOP + DE + IL + AL + (N = 5)* represents the diagnosis of YOLOv8 using *DE*, *IL*, and *AL* detection results by a multi-objective optimization method under  $N = 5$  to assess its sensitivity and specificity.

Comparing the scenarios of *Base* and *Base + DE*, it can be observed that *Base + DE* can yield better results than *Base*. However, unlike the results of MOOP, the sensitivity and specificity corresponding to *MR* have significant limitations. The sensitivity of *MR* tends towards 1, while the specificity tends towards 0, indicating that almost all the WSIs are diagnosed as positive in different confidence intervals, which clearly does not serve as effective auxiliary diagnosis.

After applying MOOP and combining *DE*, *IL*, and *AL*, the effectiveness of our approach surpasses manually devised rules in various scenarios. Additionally, each point on the graph represents a diagnostic rule, indicating that they are optimal or nearly optimal on both objectives, and no solution can further improve one objective without sacrificing the other.

Furthermore, we explored the impact of different values of  $N$  on the results, where  $\forall N$  represents no restriction on the scale of  $N$ , yielding the best results, albeit far beyond what experts can accept. We aim for  $N$  to be as small as possible for diagnosis. Comparing the scenarios of  $\forall N$  and  $N = 5/10/15$ , it is evident that our method can approach the performance of  $\forall N$  and even outperform it in some cases. Therefore, by selecting a certain Pareto solution, effective auxiliary diagnosis can be achieved to enhance the diagnostic efficiency of experts under this rule.

## 6. Conclusions

In this work, we proposed an IACCS framework. Firstly, regarding the quality assessment, cell count statistics and the identification of fuzzy regions are conducted on the WSI to ensure high-quality WSI detection and diagnosis. Then, by fully utilizing annotated information and negative WSIs, an incremental learning mechanism and active learning mechanism are introduced to effectively utilize and continuously update the data. Finally, by employing multi-objective optimization, interpretable diagnostic rules are identified, providing more reliable auxiliary diagnosis.

**Future Perspectives, Limitations, and Methodological Enhancements:** Looking ahead, the proposed IACCS framework holds considerable promise for clinical deployment. It could serve as an efficient assistive tool for cytopathologists, pre-screening WSIs to prioritize suspicious cases, thereby potentially reducing workload and diagnostic turnaround times. In regions with limited access to pathology expertise, IACCS could offer valuable preliminary screening support. Furthermore, the interpretability offered by its MOO-generated rules may facilitate clinician trust and adoption and aid in standardizing diagnostic approaches.

However, several limitations and avenues for future work should be acknowledged. While our current dataset encompassed common lesion types, its diversity concerning rarer conditions like adenocarcinoma (which we plan to focus on in future research), varied patient demographics, and different WSI acquisition protocols may be limited. This highlights the need for future validation on larger multi-center datasets to assess broader applicability. Consequently, the generalization of the current models to entirely new clinical settings or scanner types without further fine-tuning or domain adaptation techniques

remains a challenge. To more rigorously assess model robustness and generalization, particularly as we expand our research to include glandular abnormalities and more diverse datasets, future work will incorporate stratified k-fold cross-validation. The use of a fixed train/validation/test split in this study, while facilitating consistent comparisons, may not fully capture generalization across all data partitions, and k-fold cross-validation will provide a more comprehensive evaluation. Practical integration with the existing laboratory information systems and the need for continuous model monitoring also represent important considerations for real-world deployment. Ultimately, rigorous prospective clinical trials will be essential to fully validate the efficacy and safety of IACCS before its widespread clinical adoption.

**Author Contributions:** Conceptualization, W.Y. and B.D.; methodology, W.Y.; software, W.Y.; validation, W.Y. and B.D.; formal analysis, W.Y.; investigation, W.Y.; resources, B.D.; data curation, W.Y.; writing—original draft preparation, W.Y.; writing—review and editing, B.D.; visualization, W.Y.; supervision, B.D.; project administration, B.D.; funding acquisition, B.D. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Data Availability Statement:** The anonymized WSI dataset used in this study was obtained through a research collaboration. Due to patient privacy and the terms of the original collaboration agreement, these data are not publicly available.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

1. Li, Y.; Chen, J.; Xue, P.; Tang, C.; Chang, J.; Chu, C.; Ma, K.; Li, Q.; Zheng, Y.; Qiao, Y. Computer-aided cervical cancer diagnosis using time-lapsed colposcopic images. *IEEE Trans. Med. Imaging* **2020**, *39*, 3403–3415. [[CrossRef](#)] [[PubMed](#)]
2. Sung, H.; Ferlay, J.; Siegel, R.L.; Laversanne, M.; Soerjomataram, I.; Jemal, A.; Bray, F. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* **2021**, *71*, 209–249. [[CrossRef](#)]
3. Cramer, D.W. The role of cervical cytology in the declining morbidity and mortality of cervical cancer. *Cancer* **1974**, *34*, 2018–2027. [[CrossRef](#)] [[PubMed](#)]
4. Cohen, P.A.; Jhingran, A.; Oaknin, A.; Denny, L. Cervical cancer. *Lancet* **2019**, *393*, 169–182. [[CrossRef](#)] [[PubMed](#)]
5. Kuo, P.H.; Huang, C.T.; Yao, T.C. Optimized transfer learning based dementia prediction system for rehabilitation therapy planning. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2023**, *31*, 2047–2059. [[CrossRef](#)]
6. Lu, C.; Ge, R.; Tang, Z.; Fu, X.; Zhang, L.; Yang, K.; Xu, X. Multi-channel FES gait rehabilitation assistance system based on adaptive sEMG modulation. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2023**, *31*, 3652–3663. [[CrossRef](#)]
7. William, W.; Ware, A.; Basaza-Ejiri, A.H.; Obungoloch, J. A review of image analysis and machine learning techniques for automated cervical cancer screening from pap-smear images. *Comput. Methods Programs Biomed.* **2018**, *164*, 15–22. [[CrossRef](#)]
8. Sankaranarayanan, R.; Budukh, A.M.; Rajkumar, R. Effective screening programmes for cervical cancer in low-and middle-income developing countries. *Bull. World Health Organ.* **2001**, *79*, 954–962.
9. Siebers, A.G.; Klinkhamer, P.J.; Grefte, J.M.; Massuger, L.F.; Vedder, J.E.; Beijers-Broos, A.; Bulten, J.; Arbyn, M. Comparison of liquid-based cytology with conventional cytology for detection of cervical cancer precursors: A randomized controlled trial. *JAMA* **2009**, *302*, 1757–1764. [[CrossRef](#)]
10. Al-Janabi, S.; Huisman, A.; Van Diest, P.J. Digital pathology: Current status and future perspectives. *Histopathology* **2012**, *61*, 1–9. [[CrossRef](#)]
11. Niazi, M.K.K.; Parwani, A.V.; Gurcan, M.N. Digital pathology and artificial intelligence. *Lancet Oncol.* **2019**, *20*, e253–e261. [[CrossRef](#)] [[PubMed](#)]
12. Solomon, D. The 1988 Bethesda system for reporting cervical/vaginal cytological diagnosis. *JAMA* **1989**, *262*, 931–934.
13. Nayar, R.; Wilbur, D.C. *The Bethesda System for Reporting Cervical Cytology: Definitions, Criteria, and Explanatory Notes*; Springer: Berlin/Heidelberg, Germany, 2015.
14. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster r-cnn: Towards real-time object detection with region proposal networks. In Proceedings of the Advances in Neural Information Processing Systems, Montreal, QC, Canada, 7–12 December 2015; Volume 28.
15. Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You only look once: Unified, real-time object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 779–788.

16. Cao, L.; Yang, J.; Rong, Z.; Li, L.; Xia, B.; You, C.; Lou, G.; Jiang, L.; Du, C.; Meng, H.; et al. A novel attention-guided convolutional network for the detection of abnormal cervical cells in cervical cancer screening. *Med. Image Anal.* **2021**, *73*, 102197. [[CrossRef](#)] [[PubMed](#)]
17. Chen, T.; Zheng, W.; Ying, H.; Tan, X.; Li, K.; Li, X.; Chen, D.Z.; Wu, J. A task decomposing and cell comparing method for cervical lesion cell detection. *IEEE Trans. Med. Imaging* **2022**, *41*, 2432–2442. [[CrossRef](#)]
18. Cheng, S.; Liu, S.; Yu, J.; Rao, G.; Xiao, Y.; Han, W.; Zhu, W.; Lv, X.; Li, N.; Cai, J.; et al. Robust whole slide image analysis for cervical cancer screening using deep learning. *Nat. Commun.* **2021**, *12*, 5639. [[CrossRef](#)]
19. Khamparia, A.; Gupta, D.; de Albuquerque, V.H.C.; Sangaiah, A.K.; Jhaveri, R.H. Internet of health things-driven deep learning system for detection and classification of cervical cells using transfer learning. *J. Supercomput.* **2020**, *76*, 8590–8608. [[CrossRef](#)]
20. Jiang, P.; Li, X.; Shen, H.; Chen, Y.; Wang, L.; Chen, H.; Feng, J.; Liu, J. A systematic review of deep learning-based cervical cytology screening: From cell identification to whole slide image analysis. *Artif. Intell. Rev.* **2023**, *56*, 2687–2758. [[CrossRef](#)]
21. Wu, Y.; Chen, Y.; Wang, L.; Ye, Y.; Liu, Z.; Guo, Y.; Fu, Y. Large scale incremental learning. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 374–382.
22. Van de Ven, G.M.; Tuytelaars, T.; Tolias, A.S. Three types of incremental learning. *Nat. Mach. Intell.* **2022**, *4*, 1185–1197. [[CrossRef](#)]
23. Kundu, R.; Chattopadhyay, S. Deep features selection through genetic algorithm for cervical pre-cancerous cell classification. *Multimed. Tools Appl.* **2023**, *82*, 13431–13452. [[CrossRef](#)]
24. Sener, O.; Savarese, S. Active learning for convolutional neural networks: A core-set approach. *arXiv* **2017**, arXiv:1708.00489.
25. Deb, K.; Pratap, A.; Agarwal, S.; Meyarivan, T. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Trans. Evol. Comput.* **2002**, *6*, 182–197. [[CrossRef](#)]
26. Kukkonen, S.; Lampinen, J. GDE3: The third evolution step of generalized differential evolution. In Proceedings of the 2005 IEEE Congress on Evolutionary Computation, Edinburgh, UK, 2–5 September 2005; Volume 1, pp. 443–450.
27. Zhao, Y.; Lv, W.; Xu, S.; Wei, J.; Wang, G.; Dang, Q.; Liu, Y.; Chen, J. Detrs beat yolos on real-time object detection. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 16–22 June 2024; pp. 16965–16974.
28. Wang, C.Y.; Yeh, I.H.; Liao, H.Y.M. Yolov9: Learning what you want to learn using programmable gradient information. *arXiv* **2024**, arXiv:2402.13616.

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.