

Article

Mask Guidance Pyramid Network for Overlapping Cervical Cell Edge Detection

Wei Zhang ^{1,2,3} , Huijie Fan ^{1,2,*} , Xuanhua Xie ^{1,2,4}, Qiang Wang ⁵ and Yandong Tang ^{1,2}

¹ State Key Laboratory of Robotics, Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang 110016, China; zhangwei5@sia.cn

² Institutes for Robotics and Intelligent Manufacturing, Chinese Academy of Sciences, Shenyang 110016, China

³ University of Chinese Academy of Sciences, Beijing 101408, China

⁴ College of Information Engineering, Shenyang University of Chemical Technology, Shenyang 110142, China

⁵ Key Laboratory of Manufacturing Industrial Integrated, Shenyang University, Shenyang 110044, China

* Correspondence: fanhuijie@sia.cn

Abstract: An important indicator of cervical cancer diagnosis is to calculate the proportion of diseased cells and cancer cells, so it is necessary to segment cells and judge the cell status. The existing methods are difficult to deal with the segmentation of overlapping cells. In order to solve this problem, we put forward such a hypothesis by reading a large number of literature, that is, image segmentation and edge measurement tasks have unity in high-level features. To prove this hypothesis, in this paper, we focus on the complementary between overlapping cervical cell edge information and cell object information to get higher accuracy cell edge detection results. Specifically, we present a joint multi-task learning framework for overlapping cell edge detection by the mask guidance pyramid network. The main component of the framework is the Mask Guidance Module (MGM), which integrates two tasks and stores the shared latent semantics to interact in the two tasks. For semantic edge detection, we propose the novel Refinement Aggregated Module (RAM) fusion to promote semantic edges. Finally, to improve the edge pixel accuracy, the edge consistency constraint loss function is introduced to our model training. Our extensive experiments have proved that our method outperforms other edge detection efforts.

Keywords: edge detection; medical image segmentation; overlapping cervical cell; mask guidance; pyramid network



Citation: Zhang, W.; Fan, H.; Xie, X.; Wang, Q.; Tang, Y. Mask Guidance Pyramid Network for Overlapping Cervical Cell Edge Detection. *Appl. Sci.* **2023**, *13*, 7526. <https://doi.org/10.3390/app13137526>

Academic Editors: Sami Bourouis, Hammam Alshazly and Ali Javed

Received: 28 April 2023

Revised: 14 June 2023

Accepted: 14 June 2023

Published: 26 June 2023



Copyright: © 2023 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 one of the most common malignant tumors. The incidence rate of malignant tumors in the female reproductive system ranks first, and it also ranks fourth among all female malignant tumors [1]. Early diagnosis and active treatment of cervical cancer can effectively inhibit the development of cancer cells and increase the life of patients. Therefore, the detection of cervical cancer cells by computer vision technology has become a research hotspot in recent years [2–4].

From a pathological point of view, normal cervical cells first become diseased cells, and then become cancerous cells. Cervical cancer is diagnosed through cancerous cells [5]. When the diseased cells become cancerous, they will change in shape and size, and the arrangement is irregular, and may have double nucleus and other aberrations. Therefore, the diagnosis of cervical cancer needs to segment the cervical cells in the picture of cervical cells first, and analyze the shape and size of the nucleus and the ratio of nucleus to cytoplasm. This information is an important basis for experts to diagnose cervical cancer [6].

With the development of computer vision technology, more and more researches have applied image segmentation related algorithms to cervical cell segmentation, hoping

to develop methods that can automatically diagnose cervical cancer. According to the classification of image segmentation algorithms, the existing cervical cell segmentation methods can be divided into traditional core depth segmentation methods.

Traditional image segmentation algorithms include two methods, namely intensity change method and energy functional method [7]. The intensity change method mainly realizes cervical cell edge detection through object color, image strength, scene texture and other low-level information features [8]. However, in medicine, the edge of lesion cells is usually blurry, and the intensity change method has some limitations. The representative energy functional model is the active contour model and the derived algorithms. The advantage of active contour model is that it is simple and can effectively segment and detect the edges of non overlapping cervical cells [9]. However, the similarity between overlapping nuclear malignant cells is easy to lead to the classification error of overlapping areas, which brings difficulties to the diagnosis of cervical cancer. Therefore, the focus of overlapping cell edge detection is to solve the problem of overlapping cell detection.

With the rapid development of deep learning theory, the edge detection algorithm of cervical cells based on deep neural network has been widely studied [10]. For example, Wan et al. used the modified DeepLab V2 model for cytoplasmic edge detection and proposed a cell detection method based on double window localization [11]. This method divides image pixels into nucleus and cytoplasm through TernauesNet, and then realizes overlapping cell segmentation based on attention model. In order to obtain useful features, some deep models attempt to use multi-scale and multi-level networks to improve the results of cell edge detection. However, according to the edge results with high score, the existing methods have limitations in improving the edge detection accuracy. These methods usually blur and deviate from the image boundary, which makes the quality of the obtained cell edge score low. Some recent works use edge detection results to assist network models to improve the expressiveness of semantic segmentation tasks [12–14]. It is a common sense that the edge feature map is a part of the segmentation feature map, that is, the segmentation feature map contains all the object edge information. One of the works proposed a two-stream CNN architecture for semantic segmentation. In this architecture, shape information and boundary information are processed separately, and the two complement each other to improve the perception ability of edge features. However, the edge semantic information in instance segmentation is not processed. In fact, instance edge detection and instance segmentation are more coupled than binary edge maps.

To overcome the above problems, we propose a multi-task collaboration framework that combines instance segmentation and semantic edge detection. By combining these two tasks, the instance segmentation task and edge detection task advantages complement each other. We use pyramid context feature information learning to achieve the the process of one task to another. Specifically, for edge detection, we propose a novel framework by the Mask Guidance Module (MGM) and Refinement Aggregated Module (RAM) fusion to promote the cell edge detection of overlapping cervix. Since the distinction between semantic edge and non-semantic edge is crucial, and how to suppress non-semantic edge information is a primary problem. In this regard, we use instance segmentation mask as the guiding task of the edge detection, cascade the instance mask feature map from the pyramid network features, and use its feature to guide the semantic edge probability map. To obtain an accurate boundary map, we perform multi-feature fusion on the edge detection results at different levels. Because there are duality constraints between semantic segmentation and edge detection, to eliminate the constraints, we propose a novel loss function to enhance the edge consistency of semantic segmentation. For the predicted mask, edge are exported as the outer contour that can be used to constrain masks. Thus, the differences between the prediction result and the ground-truth are expressed as a loss term to impose edge consistency on the semantic mask during model training. Due to the pixel-wise operation, the duality loss term is differentiable, and the whole network can be trained in an end-to-end strategy.

Specifically, the main contributions of this paper can be summarized as follows.

- To our best knowledge, we are the first study to integrate cell segmentation task into the learning framework of edge detection task to guide the research of edge detection of cervical cells.
- In the edge detection module, we introduce a novel one-to-one mask guidance module (MGM) strategy to suppress non-semantic edges detection by fusing the derived edge from mask probability map with edge detection probability map. Then we proposed Refinement Aggregated Module (RAM) to integrating multi-level coarse edge maps for generating final refined edge detection prediction.
- Experimentally proved our proposed method simultaneously optimizes the two complementary tasks to help each other, which can improve the accuracy of edge detection.

The main contents of this paper are as follows. Section 2 introduces the related work of cervical cell edge detection. In Section 3, we introduce the proposed method. Section 4 introduces the commonly experimental datasets, evaluation metrics and experimental results. We conduct extensive to analyze the effectiveness of the representative UIC algorithms. We conclude this paper in the last section.

2. Related Work

The segmentation and edge detection of overlapping cancer cells based on computer vision technology has become a research hotspot [15]. Nosrati proposed a new segmentation method based on incorporating shape prior knowledge, using a star shape prior to segment the overlapping cervical cells in Pap smear images [16]. There are also methods based on shape coding, they segment the nucleus and cytoplasm separately through a two-stage strategy, segment the image into nuclei, cell clusters and background, and then based on the sparse coding (SC) theory and guided by representative shape features [17]. The level set evolution model is used to refine the segmentation. Among traditional multi-stage segmentation methods, the watershed-based method (MPFW) is able to segment nuclei and cytoplasm from a large number of overlapping cervical cell clusters [18]. And in subsequent MPFW, for a better representation of cell shapes, the line-shaped contours are deformed with cell contour adjustment. An efficient deep learning MIU-net is proposed for the nuclei segmentation of histopathology images, benefited from two blocks of modified inception module and attention module [19].

Song et al. propose a learning based overlapping cell segmentation method [20]. This method decomposes the overlapping cell segmentation problem into discrete cell labeling tasks with multiple cost functions through shape prior. By inputting the marking results into a dynamic multi template deformation model, accurate edge segmentation is achieved. Flavjo et al. propose a layered overlapping cell segmentation method [21]. By segmenting the cell block and nucleus respectively, cytoplasm is recognized with active contour, and the precise segmentation of overlapping cells is realized. Lu et al. propose an overlapping cell segmentation method based on joint optimization of multiple level set functions [22]. Through the restriction between cells and within cells, the contour length, edge strength and cell shape are used within cells, and the area is used between cells. Finally, the accurate segmentation of overlapping cells is achieved. For improving the signal-to-noise ratio of the image and also retaining edge detail information, another denoise study proposed with wavelet transform before edge detection [23]. Huang et al. propose a segmentation method based on confrontation generation network to simultaneously solve the problems of poor contrast, irregularity and overlap of cell object [24]. This method learns the probability distribution image of cell morphology and annotated single cell image by comparing the differences between the generated single cells. The trained cell GAN generates a complete single cell image for each cell to avoid generating multi cell images in the case of overlapping. The contour of the generated cell defines the segmentation line, and uses the cell size information to obtain the edge of the input cell image.

3. Materials and Methods

Some recent researches show that the pixel-wise based medical cell edge detection methods has shown superior over the region based methods [25]. However, pixel-wise based methods ignore the spatial coherence in the cell images. This may lead to the unsatisfactory results of cell boundary detection [26]. Most of the improved methods want to solve this problem by using multiple scale features. Other methods use post-processing techniques (such as CRF) to highlight the boundary of the object [27]. In [28], the authors propose Non-local deep features (NLDF) for salient object detection. They put forward an IOU loss, which can highlight the boundary by influencing the gradient around the edge. Although these methods improve the effect of edge segmentation at a certain level, but they do not realize that edge detection and semantic segmentation are complementary. An object segmentation result can help to detect the edge information in object edge detection, and vice versa. Based on this idea, we propose a mask guided network to overlapping cell edge detection, which called MGP-Net. The proposed network uses a single network to simultaneously model the overlapping cervical cell edge information and semantic mask information. For the network training, we use an end-to-end strategy.

The overall architecture of the proposed network is shown in Figure 1. Specifically, the network includes two parts, which are GlobalNet and LocalNet. In the following subsections, we will introduce the main modules of the propose network in details. Section 3.1 introduces the architecture of GlobalNet, including the backbone network and the Cascade Edge Feature Presentation module (CEFP). In Section 3.2, we introduce the architecture of LocalNet, including the one-to-one Mask Guidance module (MGM) and the Refinement Aggregated Module (RAM). In the last part of this section, we introduce the proposed Duality Consistency Loss (DCL), and demonstrate the training process of the entire network.

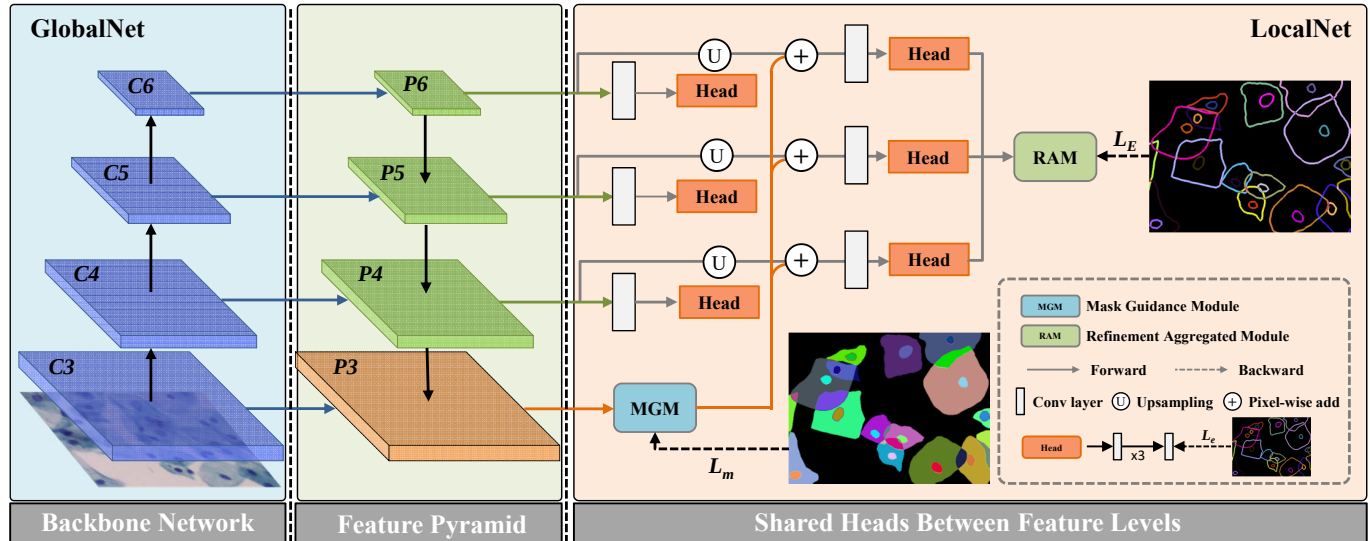


Figure 1. The overall architecture of the proposed network, a Mask Guided Pyramid Network to overlapping cervical cell edge detection, which called MGP-Net. Our framework consists of two stages: GlobalNet is constructed by backbone and feature pyramid to obtain more robust features for the cell objects. LocalNet is designed for context refinement to generate more refined features of cell edges.

3.1. GlobalNet

As shown in Figure 1, GlobalNet is constructed by backbone and feature pyramid. In this paper, we use a pre-training residual network with dilated operation as the backbone. Different from the previous ResNet, we remove the down-sampling layers, and use the dilation convolution layer in the last blocks. Thus, the size of the final feature map is 1/16 of the size of the input image. This operation can preserve the details of the input

images without adding additional parameters. The feature maps with different scales from ResNet-101 backbone are first fed into a 3×3 convolution, and then they followed by ReLU and batch normalization (BN) layer. This architecture can reduce the number of feature map to 256, and the output is regarded as the input of the feature pyramid module. We perform task-interaction by mask context guidance refining on multiple levels. Besides, we leverage high-level feature maps, including the same level feature map, to refine the low-level feature maps.

To perform mask context guidance refining on multiple levels, we propose CEFPM. As shown in Figure 1, the proposed network uses the widely used pyramid network architecture to generate the multi-resolution image features. Therefore, we obtain an useful context features. Different from the traditional pyramid networks, to obtain more robust features for the cell objects, we add three layers on each stream. These layers are convolutional operations, and we also add one ReLU layer after each convolution layer to ensure nonlinearity of the network. In addition, we conduct deep supervision on each stream. For dimensional needs, we adopt a convolutional layer to convert the feature maps to the single-channel prediction edge map. At the head of each stream, we adopt edge supervision to constrict edge feature. This operation can preserve the edge feature at each level.

3.2. LocalNet

In this paper, we assume that the edge detection and the semantic segmentation can achieve the same global feature. To realize these two tasks correlation, we adopt a global pyramid to capture the global context. Thus, the global context from instance segmentation task is to purify the feature maps from edge detection task. In Figure 2, we design a network to realize this assumption. Specifically, the feature map of P4 is refined by the pyramid context representations. This means that different level feature maps with different scales and different context collection from different patch partition. On the one hand, the context refinement can help to generate more refined features. On the other hand, this promotes the unification between semantic segmentation and edge detection. Therefore, we use pyramid mask feature presentation to get cell instance segmentation mask. As shown in Figure 2, since the pyramid context structure is implemented on the two tasks, their feature maps are closely related and complementary to each other to improve performance.

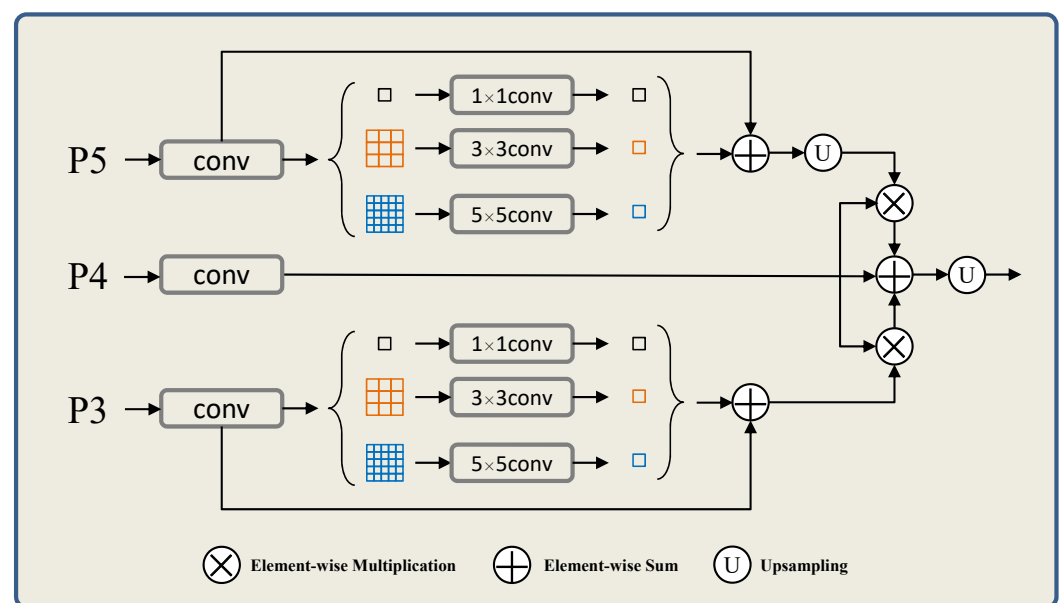


Figure 2. The Pyramid Context Structure (PCS) for LocalNet to promote the unification between two tasks of cell semantic segmentation and edge detection.

3.2.1. One-to-One Mask Guidance Module (MGM)

After obtaining cervical cell margin features and instance mask features with overlapping complementary information. In this part, our goal is to use cell instance mask features to guide cell edge features so that boundary detection results can be better segmented and localized. Based on the previously obtained complementary information, i.e., overlapping cervical cell edge features and instance mask features, integrating F_E and the $\hat{F}^{(3)}$ is the simplest and most straightforward way to take advantage of the characteristics of multi-resolution feature objects. However, in the process of gradually fusing cell edge features and multi-resolution cell target features, edge features are gradually weakened when fusing cell target features. Moreover, our goal is to fuse cell object features and cell edge features, and gradually use complementary information to achieve better prediction results. Therefore, we propose a one-to-one mask guidance module.

The specific method is that we add sub branches after network feature learning. After each sub-branch, by correcting the cell object features to enhanced cell edge features, the high-level localization prediction effect in the network structure is more accurate, and the detection effect at the final edge is improved even more. Cell mask guidance features (i -features) can be expressed as follows.

$$E^{(i)} = Up(\hat{F}^{(i)}; \theta, F_M) + F_M, i \in [3, 6] \quad (1)$$

where $\hat{F}^{(i)}$ denotes the enhanced features of side path $P^{(i)}$, F_M is the final mask features, $Up(*; F_M)$ is bilinear interpolation operation which aims to get the same size of up-sample * operation as F_M , θ is a parameter of the convolutional layer. In Equation (1), we can get the enhanced i -th feature $E^{(i)}$.

At the same time, we add deep supervision at each enhancement. For each branch output, the predicted edge graph is supervised one-to-one, and its loss function can be calculated by the cross-entropy loss L_e between the prediction edge and the ground truth value of the edge.

3.2.2. Refinement Aggregated Module (RAM)

In this part, we propose RAM to obtain an edge detection prediction. Specifically, we integrate multi-level relatively coarse edge prediction maps at different scales to generate refined edge detection results. The edge information of different levels can be captured through the multi-scale input model structure. This guarantee that the network can adaptively learn and integrate edge features of different scales, and refine the edge detection results. The architecture of the RAM is shown in Figure 3.

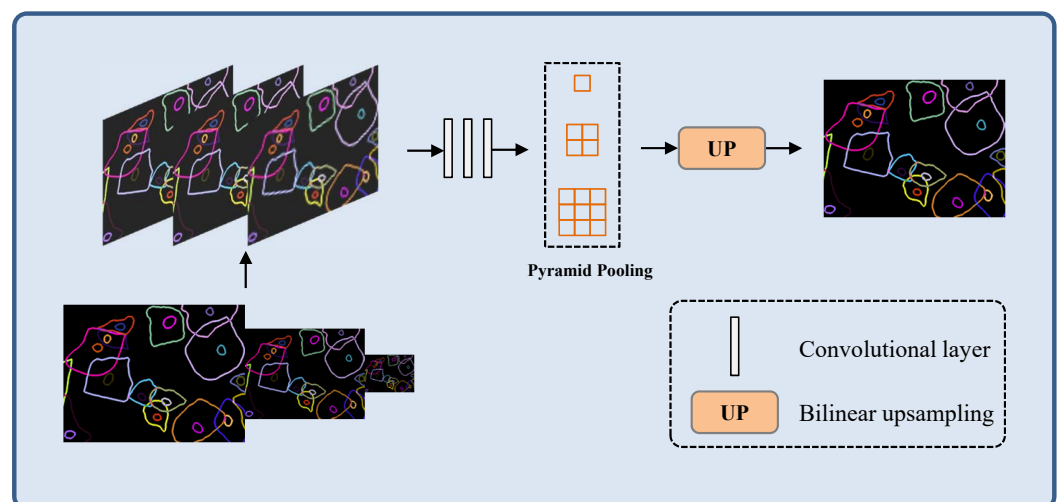


Figure 3. Refinement Aggregated Module (RAM). Integrating multi-level coarse edge prediction maps at different scales to generate final refined edge detection prediction.

In Figure 3, bilinear up-sampling operator is used to make the input edge feature maps of low resolution reach the same size to the mask. To capture global context information, we use the dimensions of the pyramid pool size with [1, 2, 3]. To fuse the edge detection results, we use several convolutional layers to unify the edge fusion operations, which generated by convolutional layers. In addition to the final boundary generation prediction graph, our model is more focused on the fixed input overall structure of the model boundary mapping. In terms of specific function implementation, we skip step 2 to correct the local error edge, making this step more flexible.

To capture pixel level image detail lost during feature extraction, hop-style joins are employed and feature fusion occurs after up-sampling blocks. In the future, the bilinear up-sampling features of autonomous branches and the features after jump connections are fused, that is, these two parts of the features are processed by convolutional layer integration. Finally, 2-layer 1×1 convolution and sigmoid activation are used to generate the edge result.

3.3. Duality Consistency Loss

As shown in Figure 1, the proposed model mainly includes two sub module structures. For each module, we calculate the loss of the probabilistic feature mapping of the two responses, separately.

Mask Context Loss. Cross-entropy loss is a commonly used calculation of the loss of each pixel when the mask instance is segmented, called L_M . The function considers every pixel to be equal, but in our task, there is a gradual blurring of pixels around the edges, which does not correspond to the actual groundtruth. Therefore, we introduce duality loss for instance segmentation loss, assuming that the two can maintain consistency between the boundary of the split object and the groundtruth of the object edge.

When we use the inconsistency between the semantic edge ∇E and the semantic edge groundtruth derived from the predicted cell instance mask, the specific formula is expressed as,

$$L_D = \sum_i \left| \nabla E_i - \nabla E_i^{gt} \right| \quad (2)$$

where E^{gt} is the semantic edge groundtruth obtained from the cell instance segmentation mask.

We introduce two consistency constraints to improve the performance of cell instance masks. where the loss function of the consistency calculation is the cross-entropy loss term L_M , which calculates the consistency between the mask and its groundtruth. The consistency loss between the derived edge of the cell instance mask and the semantic edge groundtruth is measured using the loss term L_D . Thus, the total error loss function L_{MD} measured by the cell instance segmentation task is,

$$L_{MD} = L_M + \lambda_1 L_D \quad (3)$$

where λ_1 is a constant to balance two losses. Edge Context Loss Compared with the semantic segmentation task, the pixels near the object edge are sparse. This leads to the problem of high loss rate in edge detection tasks. To alleviate this impact, we define the following cross entropy loss function L_E :

$$L_E = \sum_k \sum_i \left(\beta y_i^k \log Y_i^k + (1 - \beta)(1 - y_i^k) \log(1 - Y_i^k) \right) \quad (4)$$

where β is the percentage of non-edge pixels in the edge ground-truth and y_i^k is the ground-truth edge labels and binary indicating whether pixel i belongs to class k . Thus, the total loss function L_{total} is defined as follows:

$$L_{total} = L_{MD} + \lambda_2 L_E \quad (5)$$

where λ_2 is a weight for balancing the edge loss.

4. Results

In this part, we conduct experiments on the common dataset, named Cervical cell edge detection datasets (CCEDD) [29], to prove the effectiveness of our algorithm. Firstly, we describe the information of the datasets, the evaluation metrics and the implementation details. Then, we report the experiment results, and conduct comparison experiments to evaluate the performance of the proposed algorithm. Finally, we conduct ablation study on the propose algorithm to prove the key function of each modules.

4.1. Datasets Description

CCEDD datasets is a high resolution large-scale open source datasets. The samples in the datasets are collected by a digital camera with 3,000,000 pixel, smart v50D lens, and Nikon Ellipse CI slide scanner. All the samples in the database are collected in Liaoning Provincial Cancer Hospital. The datasets includes 686 images of 2048×1536 thin layer cytology test (TCT) cervical cells. In addition, we obtain the mask by annotating the overlapping cell instance edges. For fairness, we selected a label correction method LLPC of [29] to the CCEDD datasets for generating higher-quality edge labels.

To a joint multi-task learning framework for overlapping cell edge detection and cell mask semantic segmentation, we make a pre-processing of image data according to the cervical cell picture (a) and label file (b) provided by the CCEDD database. Since the CCEDD database is acquired by manually labeling individual cells one by one, we operate the label file separately for a single instance target, as shown in Figure 4c. Each color represents a different instance (cell or nucleus). Among them, the cell instance edge label (d) is also obtained from the cell instance mask. The final two parts of obtained cell labeled data (c) and (d) can be used for our model training.

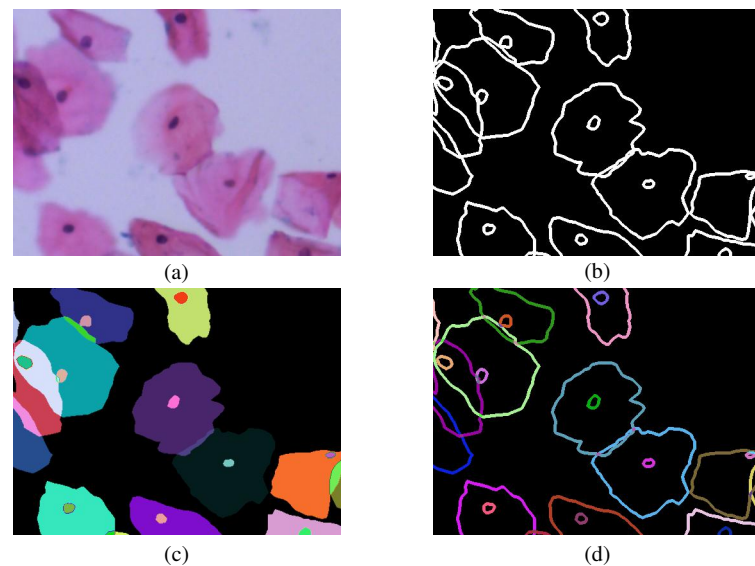


Figure 4. Pre-processing CCEDD database. (a) CCEDD cervical cell image (b) CCEDD cell edge label image (c) Obtained cell instance mask label (d) Obtained cell instance edge label.

4.2. Evaluation Metrics

In edge detection, optimal data set scale (ODS), optimal image scale (OIS) and average precision (AP) are three common evaluation metrics, which analyze the algorithm performance from different perspectives [30]. ODS uses a fixed threshold method to provide the best edge detection by global calibration in the entire dataset. OIS evaluates the edge detection performance of the entire image by selecting the best threshold. AP is the area of the precision recall curve.

4.3. Implementation Details

To expand the data volume of the datasets and ensure the network convergence, we adopt augmentation techniques for the samples in each dataset, such as image flipping, image rotation, image scaling, image random clipping and image affine transformation. Specifically, we randomly select 10% of the training data set as the verification data set, and the rotation angle ranges from -10° to 10° . We use a pre-trained network for the initialization of network parameters, and adopt the Adam optimization method to update the weights. According to the experience, we choose the step of random gradient descent with momentum is 0.99, and train the network with 3000 epochs. In each dataset, the initial learning rate is selected as e^{-2} , and the decay rate set as 0.995.

4.4. Results and Comparison

4.4.1. Ablation Study

In this part, we conduct ablation experiments to assess the importance of the modules of the proposed algorithm. For this reason, we conducted experiments on CCEDD datasets, and analyzed the impact of different modules on the proposed MGP by reducing different components in turn. In the ablation study, we use a U-Net architecture as our baseline model. Different from general U-Net, the outputs of each encoder layer are directly added, rather than connected to the corresponding decoder layer. We use this operation to improve the inference speed of U-Net. To prove the effectiveness of the MGM module, we replace the convolutional layer of U-Net with the MGM module. By further adding the RAM module to the baseline model, we obtained another model, named Baseline + RAM. We also integrated these two modules into one baseline model, and named FRCNet. The experimental results under different conditions are shown in Table 1.

Table 1. Evaluate the effectiveness of MGM and RAM.

Method	AP	ODS	OIS
Baseline	0.729	0.679	0.689
Baseline+MGM	0.751	0.6875	0.702
Baseline+RAM	0.736	0.680	0.693
MGP-Net	0.763	0.691	0.714

As shown in Table 1, the baseline method can not obtain satisfactory segmentation results, especially under the harsh conditions of irregular shapes or low contrast regions with different sizes. In comparison, by adding MGM to baseline, it can collect more instructive contextual information for each object location, the Baseline+MGM has obtained better results than the Baseline. However, we found that the background region at the edge of the cell is weakened, which can lead to the loss of boundary information during learning. In addition, in order to overcome the problem of loss of output gap information at different levels, Baseline+RAM can learn multi-range contextual information by gradually integrating local and global features at target locations to cope with edge detection challenges with different cell sizes and shapes. After this operation, the RAM module is added after the baseline, and combined with the effectiveness of the pyramid pooling mechanism, the final model can adaptively integrate multi-level output features, so that the final edge detection results after refinement can achieve satisfactory results. As can be seen from the last row of Table 1, the MGP-Net proposed in this paper gets the best performance on the AP compared to other methods. In addition, the quantitative scores of ODS and OIS for different methods are also given, as shown in Table 1. Both modules, Baseline + MGM + RAM, received higher scores in each evaluation indicator. When only the MGM module is added to the baseline, the AP metric is improved by nearly 2% compared with the baseline, as shown in Table 1. The segmentation accuracy of the baseline+RAM model is also higher than that of the baseline model, which shows that multi-level feature fusion is beneficial to improve the performance of edge detection. Our final web framework, MGP-Net, combines

the above two modules, which were trained and tested on the CCEDD dataset nearly 3.5% ahead of baseline.

4.4.2. Comparison with State-of-the-Art Algorithms

In this part, we conduct comparative experiments on several representative algorithms, including RCF [31], ENDE [32], UNet++ [33], DexiNed [34], FINED [35], PiDiNet [36], and MSU-Net [37]. To comprehensively analyze the performance of these algorithms, we conducted quantitative and qualitative experiments. To make the comparative experiment fair under the same conditions, we implemented all of the comparison methods and evaluated them on the CCEDD corrected labelled datasets by MSU-Net with their higher accuracy than no corrected. The tested data use the same experimental settings, such as data augmentation methods, operation system and hardware environments. We present our model performance and results for other methods in Table 2. From the experimental results we can find that the proposed algorithm is superior to other algorithms on different datasets, and all the three evaluation metrics have considerable margin. In terms of test data, MGP-Net outperformed the previous best method by 0.8%, 0.4% and 1.3 respectively.

Table 2. Quantitative results comparisons of our method.

Method	AP	ODS	OIS
RCF	0.667	0.638	0.645
ENDE	0.733	0.682	0.691
UNet++	0.755	0.691	0.701
DexiNed	0.723	0.671	0.680
FINED	0.703	0.660	0.621
PiDiNet	0.648	0.624	0.628
MSU-Net	0.749	0.689	0.699
MGP-Net	0.763	0.695	0.714

To observe the edge detection performance of the proposed algorithm more intuitively, we report the qualitative test results in Figure 5. The qualitative results show the edge detection results of different algorithms in CCEDD datasets. The characteristic of CCEDD datasets is that it includes many challenging cases of irregular overlapping cervical cells, which brings difficulties to the general edge detection algorithm. Besides, the extremely low contrast between foreground and background organizations may increase the probability of inaccurate edge detection.

For easy comparison, we convert our edge visualization results to binary images. From the experimental results, we can find that the classical RCF cannot deal with the complex situation of cell superposition due to the inherent limitations of its architecture. U-Net++ is superior to RCF because it uses residual technology to effectively combine image features and can use more image information. But as shown in the third column in Figure 5, UNet++ is also not applicable to edge detection in the case of overlapping cells. The main reason why we analyze this phenomenon is that it does not have sufficient global receptive field and context information.

Our MGP-Net detects more edge pixels through the proposed module and achieves satisfactory performance compared to other existing methods. By gradually fusing context information in multiple ranges and guiding according to the mask guidance of individual cells, features such as different shapes of cells at multiple scales can be effectively extracted.

The interference of non-marginal pixels can be effectively suppressed, so that the network can learn more useful and discriminating features. In addition, the MGM-based mask guidance mechanism and RAM multi-level cascading feature aggregation strategy also help to obtain accurate edge prediction and effectively obtain the final edge detection.

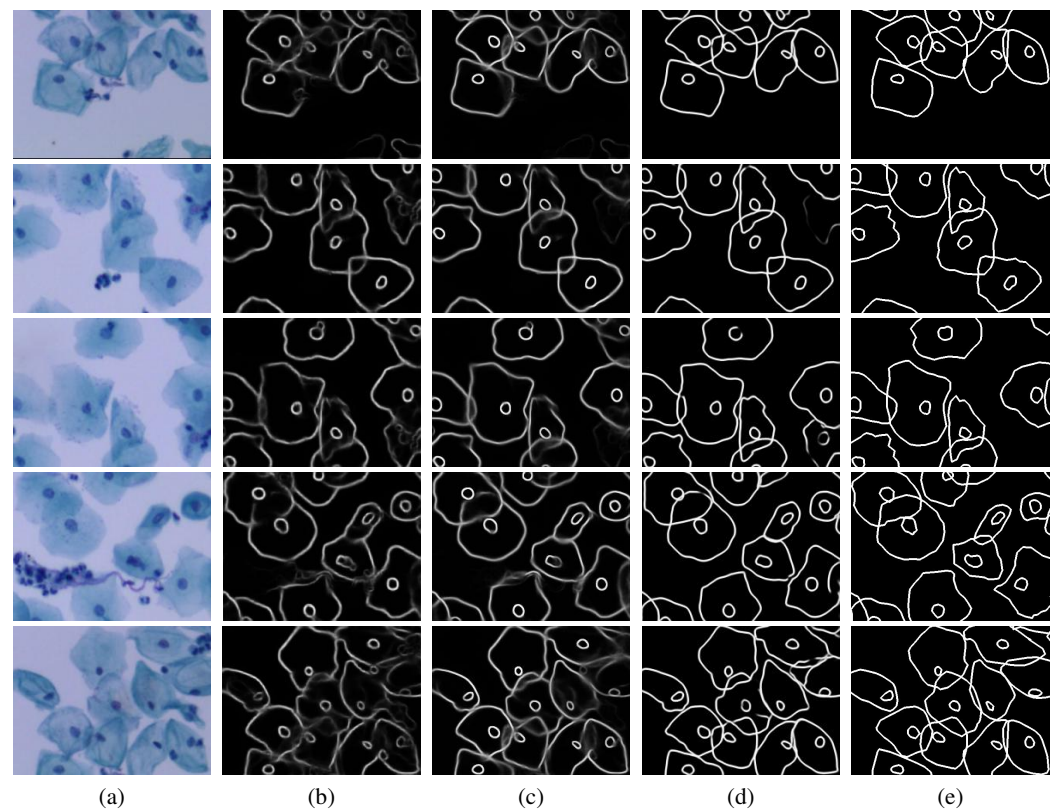


Figure 5. The edge detection performance compared with other algorithms. (a) Input cervical cell images. (b) RCF method results. (c) UNet++ method results. (d) Our proposed MGP-Net results. (e) Ground Truth label images.

5. Conclusions

The means of preventing and diagnosing cervical cancer in medicine has always been to screen cervical TCT cell images. Automatic identification and accurate cell edge detection of cervical cell images are key technologies for this medical diagnosis. In this paper, we propose a segmented mask guided joint task framework for edge detection. We use a one-to-one mask guidance module for one task on multiple cell scales to refine the feature mapping of another task, and the interaction between the two tasks contributes to the final edge detection result. In order to solve the edge sparse problem, we fuse the segmentation mask guide features into the semantic edge maps to suppress the interference of non-semantic edge pixels. In the model training stage, we design a loss function based on double constraints, which further improves the model performance and ensures the consistency between the mask-guided prediction edge and the edge ground-truth. The experimental part of this paper verifies the effectiveness of the framework under different evaluation indicators on the CCEDD datasets, and it can be seen that the proposed MGP-Net is superior to other existing methods in cell edge detection tasks. However, our method has a low detection rate in cervical cell images with a high overlap rate, which is related to the standardization of pathological cell image acquisition and the effect of model mask feature capture, which will be improved in subsequent studies.

Author Contributions: In this paper, H.F. contribute to the conception of the study. W.Z. contribute significantly to analysis and manuscript preparation. X.X. perform the experiment. W.Z. also perform the data analyses and wrote the manuscript. Q.W. and Y.T. help to perform the analysis with constructive discussions. All authors have read and agreed to the published version of the manuscript.

Funding: This work is supported by the National Natural Science Foundation of China (61873259), Natural Science Foundation of Liaoning Province (2022-YGJC-66), Youth Innovation Promotion Association, Chinese Academy of Sciences (2019203).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Jin, J.; Liu, Z.; Chen, Y. Effect of nano-tracer on identification of sentinel lymph nodes in pelvic cavity and postoperative complications in patients with cervical cancer. *J. Nanosci. Nanotechnol.* **2021**, *21*, 971–976. [[CrossRef](#)] [[PubMed](#)]
- Xia, L.I.; Xu, Z.; Shen, X.; Zhou, Y.; Li, T.Q. Detection of cervical cancer cells in whole slide images using deformable and global context aware faster RCNN-FPN. *Curr. Oncol.* **2021**, *23*, 3585–3601.
- Jia, D.; He, Z.; Zhang, C.; Yin, W.; Wu, N.; Li, Z. Detection of cervical cancer cells in complex situation based on improved YOLOv3 network. *Multimed. Tools Appl.* **2022**, *81*, 8939–8961. [[CrossRef](#)]
- Jia, F.; Li, S.; Zhang, T. Detection of cervical cancer cells based on strong feature CNN-SVM network. *Neurocomputing* **2020**, *441*, 112–127.
- Elakkiya, R.; Teja, K.; Deborah, L.J.; Bisogni, C.; Medaglia, C. Imaging based cervical cancer diagnostics using small object detection—Generative adversarial networks. *Multimed. Tools Appl.* **2022**, *81*, 191–207. [[CrossRef](#)]
- Xue, J.H.; Liu, T.; Le, B.O.; Radiology, D.O.; Xian, F.H. Value of MRI imaging in the diagnosis of cervical cancer and the judgment of surrounding invasion. *Chin. J. MRI* **2019**, *17*, 68–71.
- Bo, L.; Fei, W.; Sikai, L.; Jinhong, T.; Guang, H.L.; Meiling, Z.; Xiaohui, G. CA-Unet++: An improved structure for medical CT scanning based on the Unet++ Architecture. *Int. J. Intell. Syst.* **2022**, *37*, 8814–8832.
- Lorenzo-Ginori, J.V.; Curbelo-Jardines, W.; López-Cabrera J.D.; Huergo-Suárez, S.B. Cervical cell classification using features related to morphometry and texture of nuclei. *Iberoam. Congr. Pattern Recognit.* **2013**, *21*, 222–229.
- Krishnan, N.; Sujatha, S. Segmentation of cervical cancer images using active contour models. In Proceedings of the 2010 IEEE International Conference on Computational Intelligence and Computing Research, Coimbatore, India, 28–29 December 2010.
- Jia, D.; Zhou, J.; Zhang, C. Detection of cervical cells based on improved SSD network. *Multimed. Tools Appl.* **2022**, *81*, 13371–13387. [[CrossRef](#)]
- Wan, T. and Xu, S. and Sang, C. and Jin, Y. and Qin, Z. Accurate segmentation of overlapping cells in cervical cytology with deep convolutional neural networks. *Neurocomputing* **2019**, *365*, 157–170. [[CrossRef](#)]
- Liu, W.; Lu, Z.; Xu, H. Auxiliary edge detection for semantic image segmentation. In Proceedings of the 2020 6th International Conference on Computing and Artificial Intelligence, Tianjin, China, 23–26 April 2020; pp. 182–187.
- Wang, L. and Liu, F. and Xu, J. A semantic segmentation and edge detection model based on edge information constraint training. *J. Phys. Conf. Ser. IOP Publ.* **2020**, *1518*, 012046. [[CrossRef](#)]
- Jiang, Z.; Chen, Z.; Ji, K.; Yang, J. Semantic segmentation network combined with edge detection for building extraction in remote sensing images. In Proceedings of the MIPPR 2019: Pattern Recognition and Computer Vision, Wuhan, China, 2–3 November 2019.
- Jia, W.; Yuxuan, G.; Fangfang, G.; Zhehao, D. A medical assistant segmentation method for MRI images of osteosarcoma based on DecoupleSegNet. *Int. J. Intell. Syst.* **2022**, *37*, 8436–8461.
- Nosrati, M.S.; Hamarneh, G. Segmentation of overlapping cervical cells: A variational method with star-shape prior. In Proceedings of the 2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI), Brooklyn, NY, USA, 16–19 April 2015.
- Tareef, A.; Song, Y.; Cai, W.; Huang, H.; Chang, H.; Wang, Y.; Chen, M. Automatic segmentation of overlapping cervical smear cells based on local distinctive features and guided shape deformation. *Neurocomputing* **2017**, *221*, 97–107. [[CrossRef](#)]
- Tareef, A.; Song, Y.; Huang, H.; Feng, D.; Chen, M.; Wang, Y.; Cai, W. Multi-pass fast watershed for accurate segmentation of overlapping cervical cells. *IEEE Trans. Med. Imaging* **2018**, *37*, 2044–2059. [[CrossRef](#)] [[PubMed](#)]
- Li, J.; Li, X. MIX-Attention and Inception U-Net for Histopathology Image Nuclei Segmentation. *Appl. Sci.* **2023**, *13*, 4842. [[CrossRef](#)]
- Song, Y.; Tan, E.L.; Jiang, X.; Cheng, J. Z.; Dong, N.; Chen, S.; Lei, B.; Wang, T. Accurate cervical cell segmentation from overlapping clumps in pap smear images. *IEEE Trans. Med. Imaging* **2016**, *36*, 288–300. [[CrossRef](#)] [[PubMed](#)]
- Flavio, H.D.A.; Romuere, R.V. Silva and Fatima N.S. Medeiros and Jeova Farias Rocha Neto and Paulo Henrique Calaes Oliveira and Andrea G. Campos Bianchi and Daniela Ushizima. Active contours for overlapping cervical cell segmentation. *Int. J. Biomed. Eng. Technol.* **2021**, *35*, 70–926.
- Zhi, L.; Gustavo, C.; Andrew, P.B. An improved joint optimization of multiple level set functions for the segmentation of overlapping cervical cells. *IEEE Trans. Image Process.* **2015**, *24*, 1261–1272. [[CrossRef](#)]
- You, N.; Han, L.; Zhu, D.; Song, W. Research on Image Denoising in Edge Detection Based on Wavelet Transform. *Appl. Sci.* **2023**, *13*, 1837. [[CrossRef](#)]
- Jinjie, H.; Guihua, Y.; Biao, L.; Yongjun, H.; Yani, L. Segmentation of cervical cell images based on generative adversarial networks. *IEEE Access* **2021**, *9*, 115415–115428.

25. Guo, X.U.; Zhang, X.; Jin-Xian, H.U.; Xu-Xiang, N.I. Medical cell edge detection algorithm based on threshold segmentation and contour extraction. *Comput. Technol. Dev.* **2015**, *25*, 64–67.
26. Xiang, Y.; Qinghua, Z.; Shuihua, W.; YuDong, Z. A systematic survey of deep learning in breast cancer. *Int. J. Intell. Syst.* **2022**, *37*, 152–215.
27. Jiaxiang, J.; Po-Yu, K.; Samuel, B.A.; Daniel B.S.; Manjunath, B.S. Accurate 3D cell segmentation using deep features and CRF refinement. In Proceedings of the 2019 IEEE International Conference on Image Processing (ICIP), Taipei, Taiwan, 22–25 September 2019.
28. Luo, Z.; Mishra, A.; Achkar, A.; Eichel, J.; Jodoin, P.M. Non-local deep features for salient object detection. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017.
29. Jiawei, L.; Qiang, W.; Huijie, F.; Shuai, W.; Wentao, L.; Yandong, T.; Danbo, W.; Mingyi, Z.; Li, C. Automatic label correction for the accurate edge detection of overlapping cervical cells. *Front. Neuroinform.* **2022**, *16*, 1662–5196.
30. Arbelaez, P.; Maire, M.; Fowlkes, C.; Malik, J. Contour detection and hierarchical image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2011**, *5*, 33–53. [[CrossRef](#)] [[PubMed](#)]
31. Yun, L.; Ming-Ming, C.; Xiaowei, H.; Jia-Wang, B.; Le, Z. Richer convolutional features for edge detection. *IEEE Trans. Pattern Anal. Mach. Intell.* **2018**, *4*, 1939–1946.
32. Nazeri, K.; Ng, E.; Joseph, T.; Qureshi, F.; Ebrahimi, M. EdgeConnect: Structure guided image inpainting using edge prediction. In Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW), Seoul, South Korea, 27–28 October 2019.
33. Zhou, Z.; Siddiquee, M.; Tajbakhsh, N.; Liang, J. UNet++: Redesigning skip connections to exploit multiscale features in image segmentation. *IEEE Trans. Med. Imaging* **2020**, *39*, 1856–1867. [[CrossRef](#)]
34. Soria, X.; Riba, E.; Sappa, A. Dense extreme inception network: Towards a robust CNN model for edge detection. In Proceedings of the 2020 IEEE Winter Conference on Applications of Computer Vision (WACV), Snowmass Village, CO, USA, 1–5 March 2020.
35. Wibisono, J.K.; Hang, H.M. Fined: Fast inference network for edge detection. *arXiv* **2021**, arXiv:2012.08392.
36. Zhuo, S.; Wenzhe, L.; Zitong, Y.; Dewen, H.; Qing, L.; Qi, T.; Matti, P.; Li, L. Pixel difference networks for efficient edge detection. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Montreal, BC, Canada, 11–17 October 2021; pp. 5117–5127.
37. Su, R.; Zhang, D.; Liu, J.; Cheng, C. MSU-Net: Multi-scale U-Net for 2D medical image segmentation. *Front. Genet.* **2021**, *12*, 639930. [[CrossRef](#)]

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.