This MICCAI paper is the Open Access version, provided by the MICCAI Society. It is identical to the accepted version, except for the format and this watermark; the final published version is available on SpringerLink.

Controllable Image Synthesis Workflow for Enhancing Cervical Cell Detection

Yihuang Hu[1], Qi Chen[1], Linbo Liao[1], Weiping Lin[1], Huisi Wu[2], and Liansheng Wang[1](✉)

[1] Department of Computer Science at School of Informatics, Xiamen University, Xiamen, China
{huyihuang, qchen, wplin}@stu.xmu.edu.cn, linboliao123@163.com, lswang@xmu.edu.cn

[2] College of Computer Science and Software Engineering, Shenzhen University, Shenzhen, China
hswu@szu.edu.cn

Abstract. Cervical cancer is the only cancer that can be eliminated, yet it causes over 300,000 deaths annually. Early detection of its precancerous lesions can significantly reduce both incidence and mortality rates, while the process is labor-intensive and demands highly trained professionals. The application of artificial intelligence for cervical cell detection shows great promise but frequently encounters challenges such as limited data scale and class imbalance, stemming from the difficulties associated with expert annotation and the diverse types of cervical cells. To address this, current studies tend to design advanced detection models, while little attention is given to the potential improvements of data augmentation. In this work, we innovatively present the first controllable image synthesis workflow with adaptive cell segmentation and style transfer to synthesize realistic cervical cell images with bounding box annotations. Specifically, an adaptive cell segmentation method was introduced to cut target cells of varying sizes and morphologies from real images. These cells are then controllably pasted onto blank backgrounds to synthesize coarse images, which were further refined to realistic ones through the style transfer approach. The extensive experiment on a private long-tailed dataset demonstrated that our proposed workflow can generate realistic cervical cell images, thereby enhancing model training and improving the performance of cervical cell detection, generally and categorically. The code is available at https://github.com/huyihuang/ImageSynthesisForCCD.

Keywords: Cervical Cell Detection · Controllable Image Synthesis · Adaptive Cell Segmentation · Style Transfer.

1 Introduction

Cervical cancer is the fourth most common cancer in women worldwide, with more than 300,000 deaths annually [22]. Meanwhile, it is the only cancer that can

Y. Hu and Q. Chen——Contributed equally.

be eliminated worldwide [24], the precancerous lesions of which can be detected in the early stages. Screening methods like the Pap test and liquid-based cytology (LBC) are effective and widely used for detection while requiring highly trained expertise to accurately interpret abnormal cervical cells [2]. This dependence not only increases the workload of expertise but also leads to delayed diagnoses and reduced efficiency in large-scale screening efforts, particularly in resource-limited settings. Therefore, there is a growing interest in incorporating artificial intelligence into cervical cell detection.

Cervical cell detection is a challenging task due to its inherent complexities, such as the diversity of abnormal cells with different sizes, their morphological similarity to normal cells or each other, and the heterogeneity within each cell class [14]. Current research tends to design advanced models to achieve more accurate automated detection [4,9,21,12,25,6]. In addition to advanced models, high-quality and sufficiently large-scale data are also widely considered crucial for the detection task. However, the challenges associated with annotation and the wide types of abnormal cells make it difficult to construct large-scale datasets. Hence, the cervical cell data are typically small-scale and exhibit a long-tailed distribution, which exacerbates the training of detection models [13]. To improve data diversity, some studies have explored the use of GAN-based methods to synthesize cervical cell images [26,19,29] or individual cells [28,18] under given labels, demonstrating effectiveness in supporting image classification tasks. However, most of these methods primarily focus on the overall visual appearance, while the fidelity of fine cellular details and the morphological diversity of the synthesized images or individual cells remain insufficiently validated. In addition, these approaches often lack effective control over the morphology and spatial arrangement of individual cells, making it challenging to meet the requirements of more complex detection or segmentation tasks.

Unlike GAN-based methods, CutPaste synthesizes new images by leveraging individual real instances, which was proposed [3] and showed success in natural image domains [3,23,11]. Specifically, CutPaste involves cutting a region of interest from an image and pasting it onto another image in a different context to synthesize data for augmenting model training. In recent years, researchers have successfully adapted CutPaste to various medical imaging domains. Yap et al. [27] proposed a simple semi-supervised learning method for lesion segmentation using CutPaste augmentation and consistency regularization, demonstrating superior performance on eye fundus and brain CT scan datasets. Athalye et al. [1] demonstrated the effectiveness of a context-preserving CutPaste data augmentation strategy for view classification on fetal ultrasound FETAL-125 and OB-125 datasets. Sato et al. [17] introduced a self-supervised learning model utilizing an anatomy-aware pasting (AnatPaste) augmentation tool for unsupervised anomaly detection in chest radiographs.

However, to our knowledge, few studies have applied CutPaste to enhance cervical cell detection. Given that cells in screening images of the Pap test or LBC are generally isolated and independent, CutPaste could also potentially improve cervical cell detection. In this work, we proposed the first controllable

image synthesis workflow with adaptive cell segmentation and style transfer for cervical cell detection, which mainly consists of three stages. **Stage 1:** Adaptive cell segmentation method, based on a well-pretrained cell segmentation model, is designed to cut target cervical cells of varying sizes and morphologies from real images. **Stage 2:** The cut cells are controllably placed onto blank backgrounds to generate coarse images, the process of which allows for precise manipulation of the cell placement and orientation. **Stage 3:** The style transfer model is trained with coarse images as the source domain and real images as the target domain. Then, it is applied to transfer coarse images into refined ones. In summary, the main contributions of this work are listed as follows:

(1) We proposed an innovative, controllable image synthesis workflow with adaptive cell segmentation and style transfer for cervical cell detection. To our knowledge, this is the first approach to enhance cervical cell detection from a data synthesis perspective.

(2) We proposed an adaptive segmentation method to effectively cut cervical cells of varying sizes and morphologies. Besides, we leveraged the style transfer approach to eliminate stitching artifacts and generate more realistic images.

(3) Through extensive experiments, we demonstrated that our proposed workflow can controllably generate realistic cervical cell images, which are effective for augmenting model training and improving cervical cell detection.

2 Method

2.1 Adaptive Cell Segmentation

The precise segmentation of cervical cells with varying sizes and morphologies from real images is a critical prerequisite for coarse image synthesis. Even though there is no dedicated pretrained segmentation model for cervical cells, we noticed the existence of well-pretrained segmentation models [20,5,7] for other types of cells. Among them, we identified CellPose [20] as a promising solution, requiring a suitable cell diameter as input. On this basis, we first cut sub-images slightly larger than the labeled bounding box to ensure including the whole cells in real images, as shown in Fig. 1(a). Then, the sub-image is inputted to CellPose with an adaptive cell diameter (denoted as d), as shown in Fig. 1(a). Furthermore, cervical cell images inevitably present overlapping cells at times, which may lead to additional segmentation of cells within the sub-image, as shown in Fig. 1(a). Therefore, we applied an intersection over union (IoU) check to ensure accurate segmentation, as shown in Fig. 1(a). After that, we cut the target cells from real images and refresh the labeled bounding box by cell size. Additionally, due to limitations in computational resources, the whole slide image is typically split into patches, resulting in several cervical cells being split across different patches. Hence, we further classified cut cells into internal and boundary cells based on their position in patches, as shown in Fig. 1(b).

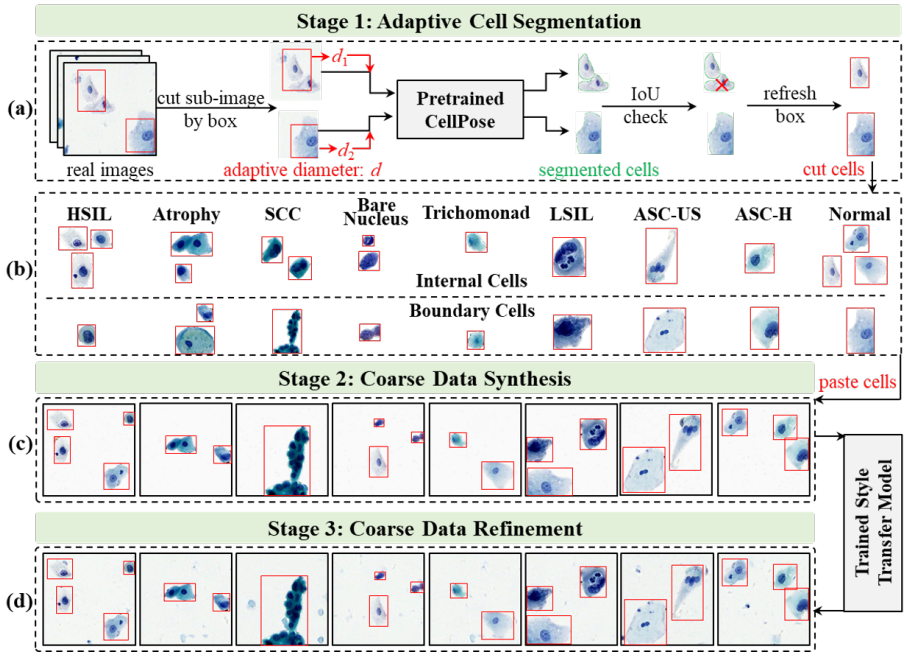

Fig. 1. Overview of our proposed controllable image synthesis workflow.

2.2 Coarse Image Synthesis

The cut cells are then pasted onto blank backgrounds of the same size as the real image to generate coarse synthetic images. It is worth noting that internal cells can be arbitrarily rotated and pasted within the coarse image, while the placement of boundary cells is restricted to simulate real scenarios, as shown in Fig. 1(c). In addition, the normal cell is also cut and distributed around the abnormal cell in coarse images to simulate the real scenario.

2.3 Refined Image Synthesis

When cells are directly pasted onto blank backgrounds, the edges of the cells often exhibit stitching artifacts, and the backgrounds of coarse images differ significantly from those of real images. These issues can potentially degrade the training of detection models [3]. We identified style transfer learning as a promising solution due to its impressive performance [30,15]. Specifically, the style transfer model is trained with coarse images as the source domain and real images as the target domain. After that, it is applied to refine coarse images with its learned transformation, smoothing out cell boundaries and replacing the unrealistic background with a natural one. Through the process above, we have established a pipeline for controllable image synthesis to generate realistic cervical cell images and corresponding annotations based on real detection data.

3 Experiments

3.1 Dataset and Evaluation Metrics

Public cervical cancer datasets [8,16] are limited in size and balanced in categories, not accurately representing real-world clinical distributions. Therefore, we obtained a private cervical cell dataset closer to the real long-tailed data, as shown in Table 1. Specifically, it contains 11748 images with 512*512 resolution and annotates 9 types of cells, including normal cells and 8 kinds of abnormal cells. As shown in Table 1, the dataset was divided into train (10,331 images), val (612 images), and test (805 images) sets. It should be noted that the train and val sets are created through overlap-cropping, while the test set remains independent for evaluation. The dataset is available on request for non-commercial and academic purposes from the author (hswu@szu.edu.cn).

Table 1. Details of the private cervical cell dataset.

Class		HSIL	Atrophy	SCC	Bare Nucleus	Tricho-monad	LSIL	ASC-US	ASC-H	Normal
train	images	3339	1349	1658	1572	842	819	770	151	4521
	instances	6019	2696	2237	2095	1742	1089	851	187	10157
val	images	216	144	37	122	73	29	26	4	348
	instances	300	258	43	162	84	29	26	4	1066
test	images	252	150	52	112	82	93	64	14	454
	instances	331	299	60	138	123	95	69	14	1350

For detection assessment, we adopted mean average precision (mAP) with an IoU threshold of 0.5 and mAP across multiple IoU thresholds (from 0.5 to 0.95), commonly referred to as mAP50-95, as the evaluation metrics, along with their values for each detected class.

3.2 Implementation Details

Adaptive Cell Segmentation. The sub-image is cropped to twice the dimensions of the corresponding bounding box while keeping the center of the box unchanged. The adaptive diameter is given by the minimum of the bounding box's width and height, which is simple yet proven effective. The IoU threshold for checking additional cells is set to 0.5. The cut cells are only from the train and val sets of the private dataset.

Coarse Image Synthesis. In each coarse image, only one type of abnormal cell (1-2 cells) is included, along with several normal cells (0-5 cells), to simulate real scenarios. During the placement, an IoU check with a threshold of 0.2 is also applied to ensure different pasted cells do not overlap with each other excessively.

Refined Image Synthesis. We synthesized 1,000 images for each type of abnormal cell as the source domain. 10,331 images from the train set were utilized as the target domain to train the style transfer models, including CycleGAN [30] and CUT/FastCUT [15]. The total number of training epochs was set to 100, with the learning rate remaining constant during the first half of the epochs and gradually decreasing to zero in the second half.
Detection Models. Since the code of cervical cell detection research is not publicly available [4,9,21,12,25,6], we conducted experiments by YOLOv11x and YOLOv11n, the largest and smallest models in the YOLOv11 detection series [10]. All models were initialized with the officially provided pretrained weights and then, trained for 200 epochs with a batch size of 64. All experiments were carried out on NVIDIA GeForce RTX 4090.

3.3 Results of Cell Segmentation and Style Transfer

Cell Segmentation Results. The cell segmentation results of different methods are shown in Fig. 2, with the segmentation boundaries indicated by green contours. The red boxes highlight the bounding boxes of the target cells. It can be observed that Hover-net [5] primarily segments cell nuclei, but fails to capture the full structure. Cellvit [7] is almost unable to segment relevant features. CellPose with a default cell diameter struggles with cells that vary in size. Our adaptive method overcomes these limitations by using a dynamic diameter, providing effective segmentation across various cells. Quantitatively, our approach achieves an average IoU of 0.758 between approximately 25,000 segmented cells and their annotated bounding boxes, outperforming other methods (IoU $<$ 0.617). When using an IoU threshold of 0.5, our method achieves a cell hit rate of 93 %, much higher than other methods (hit rate $<$ 72 %).

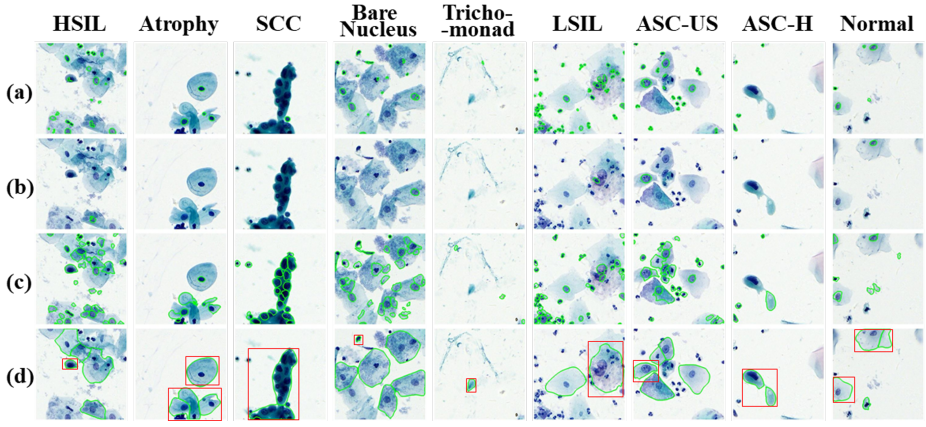

Fig. 2. cell segmentation results of different methods (green contours). Red boxes represent target cells' bounding boxes. (a) Hover-net, (b) Cellvit, (c) CellPose, (d) Ours.

Style Transfer Results. CycleGAN and CUT/FastCUT are widely recognized models for style transfer, each with its unique approach. We chose CUT as an example because its contrastive loss-based approach potentially results in better preservation of fine details, the style transfer results of which are shown in Fig. 3. As observed, trained CUT effectively eliminates cell stitching artifacts and improves the background of coarse images, generating more realistic images.

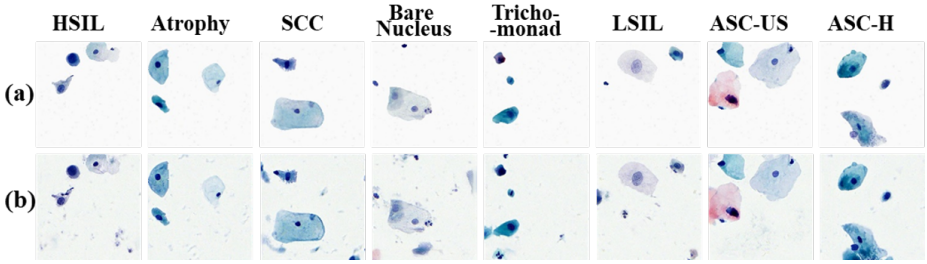

Fig. 3. The style transfer results of CUT. (a) Coarse images, (b) Refined images.

3.4 Results of Detection

To distinguish the training data for the style transfer model, we regenerated 1,000 coarse images for each type of abnormal cell and input them into the trained style transfer model to synthesize refined images for the detection experiment.

Overall Performance Analysis. We assigned the training of different models by the train set as the baseline. On this basis, we adopted a simple mixing strategy by adding the same number of synthetic images for each class of abnormal cell to the train set for data augmentation. Table 2 shows the results of YOLOv11n and YOLOv11x with varying numbers of images per class (denoted as num_per_class) and style transfer models. As observed in Table 2, incorporating refined images generated by our proposed workflow achieved better mAP50 and mAP50-95 compared to the baseline. Furthermore, incorporating coarse images typically resulted in smaller improvements than those of refined images in most cases. This further underscores the effectiveness of our proposed workflow and highlights the importance of synthetic data and realistic image generation for cervical cell detection.

Class Performance Analysis. We conducted a class performance analysis to clarify the impact of synthetic data on different categories of abnormal cells. We presented examples with YOLOv11n and CUT, illustrating two scenarios with num_per_class set at 250 and 1000, as shown in Table 3. With num_per_class equal to 250, the enhancement in head categories was minimal. However, for tail categories, such as ASC-H, there was a noticeable improvement. This is because even 250 images could provide significant augmentation for the limited training

Table 2. Results of overall performance analysis based on YOLOv11n and YOLOv11x. The **best** values are highlighted.

	mAP50				mAP50-95			
num_per_class	250	500	750	1000	250	500	750	1000
YOLOv11n								
baseline	0.433	0.433	0.433	0.433	0.326	0.326	0.326	0.326
w/ coarse	0.428	0.452	0.451	0.463	0.320	0.338	0.335	0.344
w/ refined (CycleGAN)	0.450	0.455	0.452	0.439	0.333	**0.342**	0.337	0.328
w/ refined (FastCUT)	**0.453**	0.448	**0.471**	0.455	**0.337**	0.333	**0.352**	0.333
w/ refined (CUT)	0.429	**0.458**	0.448	**0.481**	0.321	0.340	0.338	**0.356**
YOLOv11x								
baseline	0.462	0.462	0.462	0.462	0.350	0.350	0.350	0.350
w/ coarse	0.470	0.464	0.461	0.481	0.354	0.348	0.349	0.363
w/ refined (CycleGAN)	0.480	0.448	0.477	0.472	0.353	0.339	0.363	**0.368**
w/ refined (FastCUT)	0.460	**0.487**	**0.487**	**0.488**	0.346	**0.367**	**0.364**	0.364
w/ refined (CUT)	**0.489**	0.468	0.471	0.483	**0.370**	0.354	0.357	**0.368**

Table 3. Results of class performance analysis based on YOLOv11n and CUT. The **best** values are highlighted.

Metrics	Train Data	HSIL	Atrophy	SCC	Bare Nucleus	Tricho--monad	LSIL	ASC-US	ASC-H	all
numer of images per class = 250										
mAP50	baseline	**0.462**	0.939	**0.147**	**0.729**	0.226	**0.695**	0.203	0.063	**0.433**
	w/ coarse	**0.462**	**0.942**	0.087	0.691	**0.281**	0.639	**0.232**	0.091	0.428
	w/ refined	0.430	0.935	0.144	0.706	0.165	0.683	0.227	**0.139**	0.429
mAP50-95	baseline	**0.329**	0.794	**0.117**	**0.512**	0.150	0.518	0.141	0.049	**0.326**
	w/ coarse	0.327	**0.803**	0.073	0.484	**0.193**	0.459	**0.147**	0.072	0.320
	w/ refined	0.302	0.798	0.110	0.493	0.124	0.490	0.146	**0.109**	0.321
numer of images per class = 1000										
mAP50	baseline	0.462	0.939	0.147	0.729	0.226	0.695	0.203	0.063	0.433
	w/ coarse	0.458	**0.940**	0.096	0.703	**0.464**	0.689	0.227	**0.129**	0.463
	w/ refined	**0.499**	0.934	**0.215**	**0.743**	0.374	**0.710**	**0.257**	0.118	**0.481**
mAP50-95	baseline	0.329	0.794	0.117	0.512	0.150	0.518	0.141	0.049	0.326
	w/ coarse	0.324	**0.800**	0.068	0.493	**0.293**	0.519	0.155	**0.101**	0.344
	w/ refined	**0.352**	0.795	**0.156**	**0.520**	0.258	**0.522**	**0.162**	0.082	**0.356**

data of tail categories. When the volume of added data is sufficiently high, as in the case of num_per_class equal to 1000, synthetic data improved detection metrics for nearly all categories, with refined images typically yielding greater enhancements than coarse images, leading to significant improvements in mAP50 and mAP50-95 over the baseline. This further demonstrates the effectiveness of our proposed workflow, highlighting its potential to address challenges related to limited data and class imbalance.

4 Conclusion

In conclusion, this work introduces an innovative and controllable image synthesis workflow for enhancing cervical cell detection. It specifically combines an adaptive cell segmentation method to cut various cervical cells of differing sizes and morphologies, and the style transfer approach to eliminate stitching artifacts, thereby generating realistic cervical cell images with bounding box annotations. The results of overall and class performance demonstrate its effectiveness in improving cervical cell detection through data augmentation, highlighting its potential to address challenges such as limited data scale and class imbalance.

Acknowledgments. This work was supported by National Natural Science Foundation of China (Grant No. 62371409) and Fujian Provincial Natural Science Foundation of China (Grant No. 2023J01005).

Disclosure of Interests. The authors have no competing interests to declare that they are relevant to the content of this article.

References

1. Athalye, C., Arnaout, R.: Domain-guided data augmentation for deep learning on medical imaging. PloS one **18**(3), e0282532 (2023)
2. Dasgupta, S.: The efficiency of cervical pap and comparison of conventional pap smear and liquid-based cytology: a review. Cureus **15**(11) (2023)
3. Dwibedi, D., Misra, I., Hebert, M.: Cut, paste and learn: Surprisingly easy synthesis for instance detection. In: Proceedings of the IEEE international conference on computer vision. pp. 1301–1310 (2017)
4. Fei, M., Shen, Z., Song, Z., Wang, X., Cao, M., Yao, L., Zhao, X., Wang, Q., Zhang, L.: Distillation of multi-class cervical lesion cell detection via synthesis-aided pretraining and patch-level feature alignment. Neural Networks **178**, 106405 (2024)
5. Graham, S., Vu, Q.D., Raza, S.E.A., Azam, A., Tsang, Y.W., Kwak, J.T., Rajpoot, N.: Hover-net: Simultaneous segmentation and classification of nuclei in multi-tissue histology images. Medical image analysis **58**, 101563 (2019)
6. Gupta, M., Das, C., Roy, A., Gupta, P., Pillai, G.R., Patole, K.: Region of interest identification for cervical cancer images. In: 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI). pp. 1293–1296. IEEE (2020)
7. Hörst, F., Rempe, M., Heine, L., Seibold, C., Keyl, J., Baldini, G., Ugurel, S., Siveke, J., Grünwald, B., Egger, J., et al.: Cellvit: Vision transformers for precise cell segmentation and classification. Medical Image Analysis **94**, 103143 (2024)

8. Jantzen, J., Norup, J., Dounias, G., Bjerregaard, B.: Pap-smear benchmark data for pattern classification. Nature inspired smart information systems (NiSIS 2005) pp. 1–9 (2005)
9. Jia, D., Zhou, J., Zhang, C.: Detection of cervical cells based on improved ssd network. Multimedia Tools and Applications **81**(10), 13371–13387 (2022)
10. Khanam, R., Hussain, M.: Yolov11: An overview of the key architectural enhancements. arXiv preprint arXiv:2410.17725 (2024)
11. Li, C.L., Sohn, K., Yoon, J., Pfister, T.: Cutpaste: Self-supervised learning for anomaly detection and localization. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 9664–9674 (2021)
12. Liang, Y., Pan, C., Sun, W., Liu, Q., Du, Y.: Global context-aware cervical cell detection with soft scale anchor matching. Computer Methods and Programs in Biomedicine **204**, 106061 (2021)
13. Liu, M., Li, X., Gao, X., Chen, J., Shen, L., Wu, H.: Sample hardness based gradient loss for long-tailed cervical cell detection. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 109–119. Springer (2022)
14. Pangarkar, M.A.: The bethesda system for reporting cervical cytology. Cytojournal **19**, 28 (2022)
15. Park, T., Efros, A.A., Zhang, R., Zhu, J.Y.: Contrastive learning for unpaired image-to-image translation. In: Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part IX 16. pp. 319–345. Springer (2020)
16. Plissiti, M.E., Dimitrakopoulos, P., Sfikas, G., Nikou, C., Krikoni, O., Charchanti, A.: Sipakmed: A new dataset for feature and image based classification of normal and pathological cervical cells in pap smear images. In: 2018 25th IEEE international conference on image processing (ICIP). pp. 3144–3148. IEEE (2018)
17. Sato, J., Suzuki, Y., Wataya, T., Nishigaki, D., Kita, K., Yamagata, K., Tomiyama, N., Kido, S.: Anatomy-aware self-supervised learning for anomaly detection in chest radiographs. Iscience **26**(7) (2023)
18. Scalbert, M., Couzinie-Devy, F., Fezzani, R.: Generic isolated cell image generator. Cytometry Part A **95**(11), 1198–1206 (2019)
19. Shen, Z., Cao, M., Wang, S., Zhang, L., Wang, Q.: Cellgan: Conditional cervical cell synthesis for augmenting cytopathological image classification. In: International conference on medical image computing and computer-assisted intervention. pp. 487–496. Springer (2023)
20. Stringer, C., Wang, T., Michaelos, M., Pachitariu, M.: Cellpose: a generalist algorithm for cellular segmentation. Nature methods **18**(1), 100–106 (2021)
21. Tan, X., Li, K., Zhang, J., Wang, W., Wu, B., Wu, J., Li, X., Huang, X.: Automatic model for cervical cancer screening based on convolutional neural network: a retrospective, multicohort, multicenter study. Cancer cell international **21**, 1–10 (2021)
22. Tewari, K.S.: Cervical cancer. New England Journal of Medicine **392**(1), 56–71 (2025)
23. Wang, H., Wang, Q., Zhang, H., Yang, J., Zuo, W.: Constrained online cut-paste for object detection. IEEE Transactions on Circuits and Systems for Video Technology **31**(10), 4071–4083 (2020)
24. Wu, J., Jin, Q., Zhang, Y., Ji, Y., Li, J., Liu, X., Duan, H., Feng, Z., Liu, Y., Zhang, Y., et al.: Global burden of cervical cancer: current estimates, temporal trend and future projections based on the globocan 2022. Journal of the National Cancer Center (2025)

25. Xia, M., Zhang, G., Mu, C., Guan, B., Wang, M.: Cervical cancer cell detection based on deep convolutional neural network. In: 2020 39th Chinese Control Conference (CCC). pp. 6527–6532. IEEE (2020)
26. Xu, L., Cai, F., Fu, Y., Liu, Q.: Cervical cell classification with deep-learning algorithms. Medical & Biological Engineering & Computing **61**(3), 821–833 (2023)
27. Yap, B.P., Ng, B.K.: Cut-paste consistency learning for semi-supervised lesion segmentation. In: Proceedings of the IEEE/CVF winter conference on applications of computer vision. pp. 6160–6169 (2023)
28. Yu, S., Zhang, S., Wang, B., Dun, H., Xu, L., Huang, X., Shi, E., Feng, X.: Generative adversarial network based data augmentation to improve cervical cell classification model. Math. Biosci. Eng **18**(2), 1740–1752 (2021)
29. Zhao, C., Shuai, R., Ma, L., Liu, W., Wu, M.: Improving cervical cancer classification with imbalanced datasets combining taming transformers with t2t-vit. Multimedia tools and applications **81**(17), 24265–24300 (2022)
30. Zhu, J.Y., Park, T., Isola, P., Efros, A.A.: Unpaired image-to-image translation using cycle-consistent adversarial networks. In: Proceedings of the IEEE international conference on computer vision. pp. 2223–2232 (2017)