Whole Slide Cervical Cancer Screening Using Graph Attention Network and Supervised Contrastive Learning

Xin Zhang[1], Maosong Cao[2], Sheng Wang[1], Jiayin Sun[1], Xiangshan Fan[3], Qian Wang[2], and Lichi Zhang[1(✉)]

[1] School of Biomedical Engineering, Shanghai Jiao Tong University, Shanghai, China
lichizhang@sjtu.edu.cn
[2] School of Biomedical Engineering, ShanghaiTech University, Shanghai, China
[3] Department of Pathology, The Affiliated Drum Tower Hospital, Nanjing University Medical School, Nanjing, China

Abstract. Cervical cancer is one of the primary factors that endanger women's health, and Thin-prep cytologic test (TCT) has been widely applied for early screening. Automatic whole slide image (WSI) classification is highly demanded, as it can significantly reduce the workload of pathologists. Current methods are mainly based on suspicious lesion patch extraction and classification, which ignore the intrinsic relationships between suspicious patches and neglect the other patches apart from the suspicious patches, and therefore limit their robustness and generalizability. Here we propose a novel method to solve the problem, which is based on graph attention network (GAT) and supervised contrastive learning. First, for each WSI, we extract and rank a large number of representative patches based on suspicious cell detection. Then, we select the top-K and bottom-K suspicious patches to construct two graphs seperately. Next, we introduce GAT to aggregate the features from each node, and use supervised contrastive learning to obtain valuable representations of graphs. Specifically, we design a novel contrastive loss so that the latent distances between two graphs are enlarged for positive WSIs and reduced for negative WSIs. Experimental results show that the proposed GAT method outperforms conventional methods, and also demonstrate the effectiveness of supervised contrastive learning.

Keywords: Cervical cancer · Whole slide image classification · Graph attention network · Supervised contrastive learning

1 Introduction

Cervical cancer is one of the most common malignant cancers in women [14]. Many studies show that cytology screening can effectively reduce the incidence and mortality of cervical cancer [16,19]. At present, the most advanced screening method is Thin-prep cytologic test (TCT) [10]. During the diagnosis process, the cytopathology doctors need to spend a lot of time traversing all the cells, diagnosing the suspicious lesion cells among them, and grading the whole slide [13].

L. Wang et al. (Eds.): MICCAI 2022, LNCS 13432, pp. 202–211, 2022.
https://doi.org/10.1007/978-3-031-16434-7_20

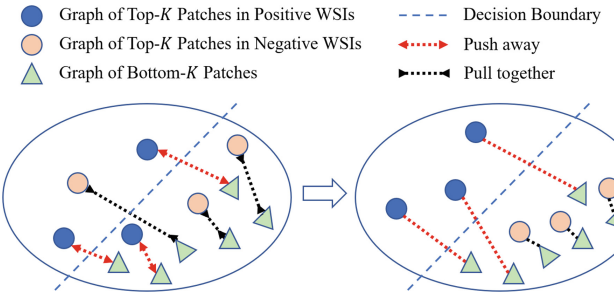

Fig. 1. The effect of supervised contrastive learning. **Left:** Some WSIs may be misclassified, which only depends on the graph of top-K patches. **Right:** In positive WSIs, the distances between graphs of top-K and bottom-K patches are expanded, while in negative WSIs, the distances are reduced, so that the WSIs can be properly classified.

Therefore, manual screening is labor-intensive and inevitably subjective. With the progress of digital whole-slide image-scanning instruments [18], accurate and efficient computer-aided cervical cancer screening becomes feasible.

Traditional methods, mainly based on morphological and textural characteristics, generally consist of cell segmentation [1,6], feature extraction [7], and cell classification [12]. With the development of deep learning [8], many attempts have been applied to the identification of cervical lesion cells based on convolutional neural networks (CNNs). For example, Yi et al. [20] proposed an automatic cervical cell detection method based on Dense-Cascade R-CNN. Du et al. [3] used CNN with attention-guided semi-supervised learning for the classification of cervical cell images. Zhou et al. [22] proposed a two-stage method based on CNN for cervical screening of pathology images. Shi et al. [15] used graph convolutional network (GCN) for the classification of cervical cell images. However, only a few studies tackled the whole slide image (WSI) diagnostic problem. Zhou et al. [21] proposed a three-stage method including cell-level detection, image-level classification, and case-level diagnosis obtained by an SVM classifier. Cheng et al. [2] designed a pipeline to find the most suspicious lesion cells in each slide and feed the features extracted by CNN to a recurrent neural network (RNN) to grade the whole slide.

The keynote of the methods mentioned above is finding the most suspicious lesion cells and integrating them to grade the WSIs. However, most methods ignore the intrinsic relationships between the suspicious lesion cells. Furthermore, in clinical practice, doctors not only focus on the suspicious lesion cells but also take the other cells into consideration. It is necessary due to the potential individual sample differences such as staining color variations. But the existing methods neglected information in the other areas which are not suspicious enough.

To solve the problems in WSI-level analysis, we propose a robust computer-aided diagnostic system for cervical cancer screening based on graph attention network (GAT) [17] and supervised contrastive learning. Our system focuses on

distinguishing positive and negative classes of WSIs. Inspired by the standard screening routine of pathologists, our screening framework is developed in a two-stage manner. In the first stage, a large number of representative patches are extracted by RetinaNet [9] and then ranked by a pre-trained SE-ResNeXt-50 [5]. In the second stage, to model the intrinsic relationships between the suspicious lesion patches, we choose top-K and bottom-K suspicious lesion patches to construct graphs separately and use GAT to aggregate their features. As illustrated in Fig. 1, we develop a novel supervised contrastive learning strategy that in positive WSIs, the distances between two graphs are forced to expand, while in negative WSIs the distances are reduced. Therefore, the model can learn from the different distances between two graphs in negative and positive WSIs, which further improves the accuracy and robustness of our screening system.

2 Method

The proposed framework is shown in Fig. 2 which consists of two major stages. The first stage aims to generate a large number of suspicious lesion patches. These representative patches are further ranked by a pre-trained model. In the second stage, two different groups of patches are selected: one group contains top-K suspicious lesion patches in the WSI and the other one contains bottom-K suspicious lesion patches, which can be seen as the false positive patches generated in the detection process. The two groups of patches are aggregated into two graphs separately. The representations of graphs are computed by GAT and max-pooling layers, then in latent space, their feature vectors are optimized based on our designed contrastive loss. In the end, the screening system utilizes the graph representation of top-K patches to make the final prediction of the WSI.

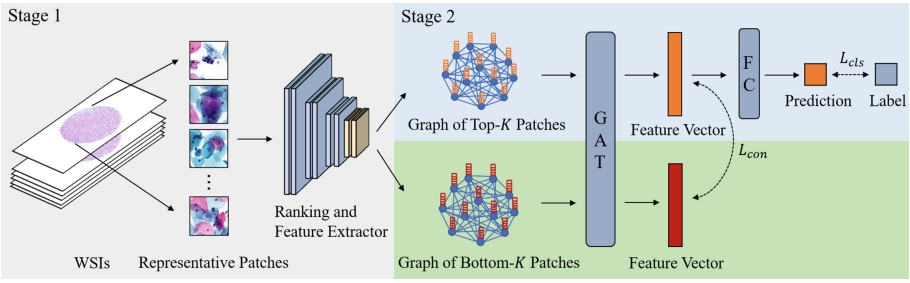

Fig. 2. The overview of our proposed framework. Stage 1 is to detect and extract all representative patches with their suspicious ranking information. Stage 2 is the proposed Graph Attention Network based on supervised contrastive learning for WSI-level classification.

2.1 Representative Patches Extraction and Ranking

Stage 1 aims to extract the representative patches based on the suspicious lesion cells detected in the WSI. Note that as it is impractical to directly implement detection on the WSI due to its enormous image size, we firstly crop every WSI into image tiles which are 1024×1024 in pixel. Usually, there are around 600 image tiles in one WSI. Then we adopt RetinaNet [9] as the detection model, which has demonstrated its effectiveness in this field. We train a RetinaNet which can automatically locate the suspicious lesion cervical cells by providing their bounding boxes, with the confidence scores as their initial suspicious ranking. For each bounding box of the detected cell, we extract a corresponding patch by starting from the center of the bounding box and expanding outward until reaching the patch size of 224×224 in pixel. We choose the top 200 patches as the representative patches according to confidence scores. The confidence scores output by the RetinaNet may not be accurate, so we need another classification model to regrade the representative patches. A large number of patches are collected in advance and manually divided by experienced pathologists into positive and negative classes, then we use SE-ResNext-50 [5] as the backbone and train the classification model. In this way, the suspicious ranking of the selected patches is updated by applying the classification model to them, experiments show that the results are more reliable than the initial ranking from the detection model.

2.2 WSI-Level Classification with GAT

The pipeline of the Graph Attention Network with supervised contrastive learning is shown in Stage 2 in Fig. 2. The inputs of the network are two graphs constructed by two groups of patches and the output is the prediction of the WSI.

Graph Construction. Let $S = \{(X_i, Y_i)\}, i = 1, 2, \ldots, N$ denote the dataset of WSIs. Here X_i represents the WSI, and Y_i represents the label of the WSI. After ranking the patches in Stage 1, we choose the top-K and bottom-K patches for further classification ($K = 20$). For negative WSIs, both groups of patches are negative and should share similar representations in latent space. For positive WSIs, the distances between two groups of patches in latent space should be large, because the confidence scores of the bottom-K patches are much smaller than 0.5 and can be seen as negative patches.

Through the pre-trained SE-ResNext-50, we can obtain the representations of the patches, which are all 2048-dim feature vectors. We construct two fully connected graphs based on two groups of patches for each WSI. Every node represents a patch and connects to the other nodes in the graph. The node features are the feature vectors of the corresponding patches. Here we define the graph of top-K patches in X_i as $G_i^+ = (V_i^+, E_i^+)$, and the graph of bottom-K patches as $G_i^- = (V_i^-, E_i^-)$ respectively. V_i includes a set of node features, $\boldsymbol{h}_i = \left\{\boldsymbol{h}_i^1, \boldsymbol{h}_i^2, \ldots, \boldsymbol{h}_i^K\right\}, \boldsymbol{h}_i^k \in \mathbb{R}^F$, where $\boldsymbol{h}_i^k$ is the output feature vector of the k-th patch in X_i, and F is the number of features in each node, which is 2048.

E_i represents the weights of edges. Each adjacency matrix of graphs is 20×20 where every number is initialized to 1.

Graph Attention Network. We propose a Graph Attention Network, which mainly includes two graph attention layers to solve the WSI analysis problem. For each X_i, the input to the network is G_i^+ and G_i^-. Both graphs are computed by the same graph attention layers.

Given $G_i = (V_i, E_i)$, the first graph attention layer produces a new set of node features, $\widehat{\boldsymbol{h}_i} = \left\{\widehat{\boldsymbol{h}_i^1}, \widehat{\boldsymbol{h}_i^2}, \ldots, \widehat{\boldsymbol{h}_i^K}\right\}, \widehat{\boldsymbol{h}_i^k} \in \mathbb{R}^{F'}$, as its output ($F' = 512$). We firstly perform self-attention mechanism on the nodes to compute the attention coefficients e_{ijk} , which represents the importance of node j's features to k's in G_i:

$$e_{ijk} = \text{LeakyReLU}(\boldsymbol{a}^T[\text{W}\boldsymbol{h}_i^j \parallel \text{W}\boldsymbol{h}_i^k]), \tag{1}$$

where $\text{W} \in \mathbb{R}^{F' \times F}$ is a learnable linear transformation, $\boldsymbol{a} \in \mathbb{R}^{2F'}$ is a single-layer feed-forward neural network, and $\parallel$ represents concatenation operation. To make coefficients comparable across different nodes, we normalize them across all of the nodes using the softmax function:

$$\alpha_{ijk} = \text{softmax}(e_{ijk}) = \frac{\exp(e_{ijk})}{\sum_{k=1}^{K} \exp(e_{ijk})}. \tag{2}$$

The normalized attention coefficients are used to compute a linear combination of the features corresponding to them, to serve as the output features for every node:

$$\widehat{\boldsymbol{h}_i^j} = \sigma(\sum_{k=1}^{K} \alpha_{ijk} \text{W}\boldsymbol{h}_i^k), \tag{3}$$

where σ denotes the logistic sigmoid operation.

We perform multi-head attention on both layers with 8 heads. The aggregated features from each head are concatenated in the first layer and averaged in the last layer. Through graph attention layers, each node in graphs aggregates the information of the other nodes, and the weights of edges and node features are updated. Then we obtain the graph representations, $\boldsymbol{H}_i \in \mathbb{R}^{F'}$, by a simple max-pooling layer.

Let $\boldsymbol{H}_i^+$ denote the representation of G_i^+, and $\boldsymbol{H}_i^-$ denote the representation of G_i^-. The inference of our GAT model is based on $\boldsymbol{H}_i^+$, so we connect a fully-connected layer after $\boldsymbol{H}_i^+$ and obtain the classification results of the network, $f(\boldsymbol{H}_i^+)$, where f represents the transform function. Finally, we use cross-entropy loss to minimize the predicted errors for the GAT model, and the classification loss is therefore written as:

$$L_{cls} = \frac{1}{N} \sum_{i=1}^{N} CE(Y_i, f(\boldsymbol{H}_i^+)), \tag{4}$$

where CE denotes cross-entropy loss.

Supervised Contrastive Learning. $\boldsymbol{H}_i^+$ is the representation of the most suspicious lesion cells in X_i and $\boldsymbol{H}_i^-$ is the representation of the least suspicious lesion cells in X_i. Here we define the cosine similarity of $\boldsymbol{H}_i^+$ and $\boldsymbol{H}_i^-$ as:

$$sim(\boldsymbol{H}_i^+, \boldsymbol{H}_i^-) = \left\langle \boldsymbol{H}_i^+, \boldsymbol{H}_i^- \right\rangle = \frac{\boldsymbol{H}_i^+ \cdot \boldsymbol{H}_i^-}{\left|\boldsymbol{H}_i^+\right| \cdot \left|\boldsymbol{H}_i^-\right|}, \tag{5}$$

where $sim(\boldsymbol{H}_i^+, \boldsymbol{H}_i^-) \in [-1, 1]$ computes the cosine of the angle between $\boldsymbol{H}_i^+$ and $\boldsymbol{H}_i^-$. Let D_i denotes the distance between $\boldsymbol{H}_i^+$ and $\boldsymbol{H}_i^-$, and we obtain

$$D_i = 1 - sim(\boldsymbol{H}_i^+, \boldsymbol{H}_i^-). \tag{6}$$

If X_i is negative ($Y_i = 0$), both G_i^+ and G_i^- are constructed by negative patches, which indicates their latent distance D_i should be small. If X_i is positive ($Y_i = 1$), D_i should be large. So we propose a novel training strategy based on contrastive learning, which takes advantage of the prior knowledge that the distances between top-K and bottom-K patches are different in positive and negative WSIs. In this way, here we design a novel contrastive loss function L_i to decrease D_i in the negative WSI and increase D_i in the positive WSI:

$$L_i = \begin{cases} \alpha D_i, & Y_i = 0 \\ \beta(2 - D_i), & Y_i = 1 \end{cases} \tag{7}$$

where α and β denote the different weights of loss. The final contrastive loss is:

$$L_{con} = \frac{1}{N} \sum_{i=1}^{N} L_i. \tag{8}$$

Total Loss. The total loss for our framework is written as follows:

$$L_{total} = L_{cls} + L_{con}. \tag{9}$$

3 Experimental Results

Dataset. For the suspicious cervical cell detection, our training dataset includes 9000 images with the size of 1024 × 1024 pixels from WSIs. All abnormal cervical cells are manually annotated in the form of bounding boxes. Then we extract the suspicious lesion patches from both positive and negative WSIs using the detection model, and collect 5000 positive cell patches and 5000 negative cell patches. The patches are used to train the SE-ResNext-50 for updating the ranking information and feature extracting, as previously mentioned in Sect. 2.1. For the WSI-level classification model, we collect 3485 negative WSIs and 3462 positive WSIs. Note that there is no data overlap when training these models.

Implementation Details. The backbone of the suspicious cell detection network is RetinaNet [9] with ResNet-50 [4]. The backbone of the pre-trained patch classification network is SE-ResNeXt-50 [5]. All parameters are optimized by Adam with the initial learning rate of 4e-5. The network is trained for 100 epochs. The batch size is 128. The model is implemented by PyTorch on 2 Nvidia Tesla P100 GPUs.

Evaluation of the Proposed Method. Our framework includes two stages. The first stage aims to extract suspicious lesion patches and classify them, which is a common method for WSI-level classification [2,23]. The second stage is to aggregate the patches extracted before and make the final prediction of WSIs. In this paper, we focus on the second stage and propose to use GAT for aggregation. So we choose some common classification methods including SVM, MLP and RNN to compare with GAT. For the SVM method, we use the confidence scores of top-K patches output by the first stage as features and train an SVM classifier. For the MLP method, we aggregate the output features of top-K patches by max-pooling and feed the features to a fully connected layer to train the model. For the RNN method, the features of top-K patches are integrated and trained by the RNN model according to [2]. For the GAT method, we perform an ablation study to further evaluate the contributions of contrastive learning by adjusting α and β. α is the weight of contrastive loss for negative WSIs and β is for positive WSIs. If they are both set to 0, the graphs of bottom-K patches will not be involved during training. If $\alpha = 0$ and $\beta = 1$, the contrastive loss will only have effects on positive WSIs and vice versa.

We conduct a 5-fold cross-validation and testing experiment to evaluate the performance of the proposed method. All of the WSIs are equally and randomly divided into 5 groups. The ratio of training, validation and testing WSIs is 3:1:1. In each fold, one group is selected as the testing group and the other four are training and validation groups.

Table 1. The comparison of different methods and the ablation study for contrastive learning, α and β are the weights of contrastive loss for negative and positive WSIs, respectively. (%)

Method		ACC	AUC	REC	PREC	F1
SVM		76.82 ± 0.75	84.32 ± 1.07	73.49 ± 0.99	78.62 ± 1.20	75.96 ± 0.84
MLP		78.70 ± 1.94	85.60 ± 0.63	60.78 ± 4.11	84.52 ± 1.08	73.90 ± 3.11
RNN		80.89 ± 1.29	$87.17 \pm$v1.42	73.71 ± 2.52	84.08 ± 2.77	79.82 ± 0.96
GAT	$\alpha = 0, \beta = 0$	82.31 ± 1.62	90.77 ± 0.94	75.93 ± 5.75	87.21 ± 2.79	80.95 ± 2.50
	$\alpha = 0, \beta = 1$	84.93 ± 1.51	92.24 ± 1.08	82.09 ± 1.68	87.03 ± 2.42	84.46 ± 1.36
	$\alpha = 1, \beta = 0$	85.12 ± 0.99	$\mathbf{92.57 \pm 0.92}$	79.67 ± 1.37	$\mathbf{89.35 \pm 1.51}$	84.22 ± 0.94
	$\alpha = 1, \beta = 1$	$\mathbf{85.79 \pm 1.21}$	92.52 ± 0.91	$\mathbf{82.63 \pm 2.04}$	88.15 ± 1.39	$\mathbf{85.28 \pm 1.27}$

Table 1 shows the experimental results of other methods and the proposed method. It is observed that the basic GAT model without contrastive learning ($\alpha = 0, \beta = 0$) already outperforms other classifiers, which reaches 82.31% accuracy. Only performing contrastive learning on positive WSIs ($\alpha = 0, \beta = 1$) leads to higher sensitivity of the model, which raises the recall from 75.93% to 82.09%. Respectively, if we perform contrastive learning on negative WSIs, the precision increases from 87.21% to 89.35%, indicating the specificity of the model is enhanced. The GAT model with contrastive learning ($\alpha = 1, \beta = 1$) reaches the highest accuracy 85.79% with more balanced recall and precision.

We visualize the feature distribution of graphs in testing groups of WSIs by t-SNE [11]. The dimensional-reduced features are shown in Fig. 3. The graphs of bottom-K patches represent negative information of WSIs, so their features are aligned in Fig. 3(a), and the graphs of top-K patches in negative WSIs are separated from those in positive WSIs. It is noted that in Fig. 3(b), the distances between graphs of top-K and bottom-K patches in positive WSIs are expanded. And in Fig. 3(c), the distances between graphs of top-K and bottom-K patches in negative WSIs are reduced. Figure 3(d) achieves the goal that graphs of top-K patches of positive WSIs and negative WSIs are well separated, which demonstrates the effectiveness of contrastive learning.

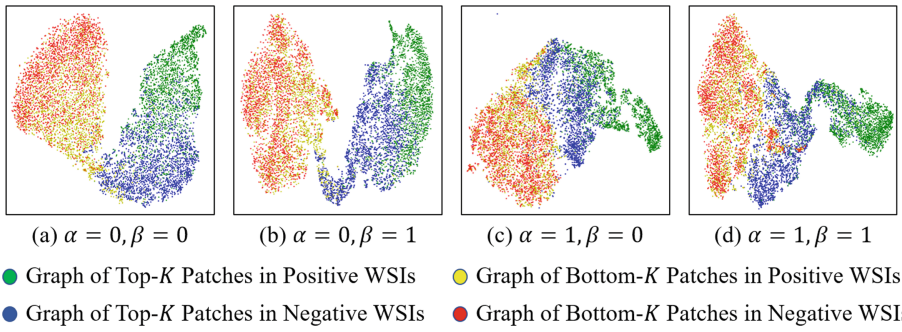

Fig. 3. The t-SNE visualizations of graphs in testing groups. (a) shows the feature distribution without contrastive learning. (b) shows the feature distribution with only applying contrastive learning in positive WSIs. (c) shows the feature distribution with only applying contrastive learning in negative WSIs. (d) shows the feature distribution with contrastive learning.

4 Conclusion

In this paper, a novel WSI classification method for cervical cancer with GAT and supervised contrastive learning is developed. Our method constructs graphs of the top-K and bottom-K suspicious lesion patches and aggregates node features into graph representations for WSI classification. Besides, the distances in latent space between top-K and bottom-K patches in positive and negative

WSIs are used for contrastive learning, which effectively improves the performance of GAT. Our work has great value in clinical applications, and can also be further applied to other WSI classification tasks in the computer-aided diagnosis of pathology images. Our source code and example data are available at https://github.com/ZhangXin1997/MICCAI-2022.

References

1. Chang, C.W., et al.: Automatic segmentation of abnormal cell nuclei from microscopic image analysis for cervical cancer screening. In: 2009 IEEE 3rd International Conference on Nano/Molecular Medicine and Engineering, pp. 77–80 (2009)
2. Cheng, S., et al.: Robust whole slide image analysis for cervical cancer screening using deep learning. Nature Commun. **12**, 5639 (2021)
3. Du, X., Huo, J., Qiao, Y., Wang, Q., Zhang, L.: False positive suppression in cervical cell screening via attention-guided semi-supervised learning. In: Rekik, Islem, Adeli, Ehsan, Park, Sang Hyun, Schnabel, Julia (eds.) PRIME 2021. LNCS, vol. 12928, pp. 93–103. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-87602-9_9
4. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778 (2016)
5. Hu, J., Shen, L., Sun, G.: Squeeze-and-excitation networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 7132–7141 (2018)
6. Kale, A., Aksoy, S.: Segmentation of cervical cell images. In: 2010 20th International Conference on Pattern Recognition, pp. 2399–2402 (2010)
7. Kim, K.B., Song, D.H., Woo, Y.W.: Nucleus segmentation and recognition of uterine cervical pap-smears. In: International Workshop on Rough Sets, Fuzzy Sets, Data Mining, and Granular-Soft Computing, pp. 153–160 (2007)
8. LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. Nature **521**(7553), 436–444 (2015)
9. Lin, T.Y., Goyal, P., Girshick, R., He, K., Dollár, P.: Focal loss for dense object detection. In: Proceedings of the IEEE international conference on computer vision, pp. 2980–2988 (2017)
10. Liu, Y., Zhang, L., Zhao, G., Che, L., Zhang, H., Fang, J.: The clinical research of Thinprep Cytology Test (TCT) combined with HPV-DNA detection in screening cervical cancer. Cell Mol. Biol. (Noisy-le-grand) **63**(2), 92–95 (2017)
11. Van der Maaten, L., Hinton, G.: Visualizing data using t-SNE. J. Mach. Learn. Res. **9**(86), 2579–2605 (2008)
12. Mariarputham, E.J., Stephen, A.: Nominated texture based cervical cancer classification. Comput. Math. Methods Med. **2015**, 1–10 (2015)
13. Nayar, R., Wilbur, D.C. (eds.): The Bethesda System for Reporting Cervical Cytology. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-11074-5
14. Schiffman, M., Castle, P.E., Jeronimo, J., Rodriguez, A.C., Wacholder, S.: Human papillomavirus and cervical cancer. The Lancet **370**(9590), 890–907 (2007)
15. Shi, J., Wang, R., Zheng, Y., Jiang, Z., Zhang, H., Yu, L.: Cervical cell classification with graph convolutional network. Comput. Methods Programs Biomed. **198**, 105807 (2021)

16. Solomon, D., Breen, N., McNeel, T.: Cervical cancer screening rates in the united states and the potential impact of implementation of screening guidelines. CA Cancer J. clin. **57**(2), 105–111 (2007)
17. Veličković, P., Cucurull, G., Casanova, A., Romero, A., Liò, P., Bengio, Y.: Graph attention networks. In: International Conference on Learning Representations (ICLR)), pp. 1–12 (2017)
18. Wright, A.M., et al.: Digital slide imaging in cervicovaginal cytology: a pilot study. Arch. Pathol. Lab. Med. **137**(5), 618–624 (2013)
19. Yang, D.X., Soulos, P.R., Davis, B., Gross, C.P., Yu, J.B.: Impact of widespread cervical cancer screening: number of cancers prevented and changes in race-specific incidence. Am. J. Clin. Oncol. **41**(3), 289 (2018)
20. Yi, L., Lei, Y., Fan, Z., Zhou, Y., Chen, D., Liu, R.: Automatic detection of cervical cells using dense-cascade R-CNN. In: Chinese Conference on Pattern Recognition and Computer Vision (PRCV), pp. 602–613 (2020)
21. Zhou, M., et al.: Hierarchical pathology screening for cervical abnormality. Comput. Med. Imaging Graph. **89**, 101892 (2021)
22. Zhou, M., et al.: Hierarchical and robust pathology image reading for high-throughput cervical abnormality screening. In: International Workshop on Machine Learning in Medical Imaging, pp. 414–422 (2020)
23. Zhu, X., et al.: Hybrid AI-assistive diagnostic model permits rapid TBS classification of cervical liquid-based thin-layer cell smears. Nat. Commun. **12**(1), 1–12 (2021)