2023 IEEE 4th International Conference on Pattern Recognition and Machine Learning (PRML) | 979-8-3503-2430-3/23/$31.00 ©2023 IEEE | DOI: 10.1109/PRML59573.2023.10348294

Deep Learning for Diagnosis of Endometrial Cancer and Atypical Endometrial Hyperplasia

Aihua Zhao, Xin Zhu
Graduate School of Computer Science and Engineering
The University of Aizu
Fukushima, Japan
zhuxin@u-aizu.ac.jp

Xin Du
Department of Gynecology
Maternal and Child Hospital of Hubei Province
Tongji Medical College
Huazhong University of Science and Technology
Wuhan, China

Wenwen Wang, Wenqing Ma, Shixuan Wang
Department of Obstetrics and Gynecology
Tongji Hospital, Tongji Medical College
Huazhong University of Science and Technology
Wuhan, China

***Abstract*—Endometrial cancer(EC) is the most common and rapidly increasing female cancer globally. Atypical endometrial hyperplasia (AEH) is a precancerous condition of EC. Although hysteroscopy serves as the primary modality for diagnosing lesions, it relies on the subjective judgment of hysteroscopists. Therefore, this study proposed a computer-aided diagnostic system utilizing the EfficientNet network as a baseline, incorporating ParNet attention mechanism and class weighting to accurately classify EC/AEH from benign lesions. This study included 49,556 hysteroscopy images from 1,237 cases as a training set and 3,412 hysteroscopy images from 85 cases as a testing set. AUC, accuracy, sensitivity, specificity, PPV, Kappa, and F_1-Score of the proposed method are 0.941, 89.4%, 93.7%, 87.1%, 73.3%, 0.755, and 0.8225, respectively. The proposed model may be used as a computer-aided tool for the diagnosis of EC/AEH.**

***Index Terms*—Hysteroscopy, deep learning, endometrial cancer, atypical endometrial hyperplasia, EfficientNet**

I. Introduction

Endometrial cancer (EC) is a common gynecological cancer that occurs in the endometrium [4]. EC is a rapidly increasing women's cancer globally and is the most common female genital tract cancer in middle- and high-income countries [6].EC has become the fourth prevalent female cancer in high-income countries [10]. In 2020, the world witnessed 417,367 new cases and 97,370 fatalities attributed to the condition [11]. Although EC is generally considered to have a favorable outcome, high-grade cancers tend to recur and a more advanced stage or in metastatic lesions, posing a significant therapeutic challenge [16]. Atypical endometrial hyperplasia (AEH) is a precancerous lesion of the endometrium. After hysterectomy, EC was reportedly diagnosed in 27% to 52% of patients who had preoperative AEH [8]. The risk of concurrent endometrial cancer in patients with AEH during hysterectomy is 40% [15]. Accurate diagnosis of EC and AEH is crucial for early treatment, given the rapid progression of these lesions.

Experienced surgeons often fail to curette more than half of the uterine cavity in 60% of D&C procedures, which can complicate the diagnosis, particularly in instances of focal uterine lesions [3] [20]. For the direct evaluation and detection of structural abnormalities in the endometrium and uterine cavity, hysteroscopy has been generally accepted as a safe and minimally invasive method [1]. With the rapid development in endoscopic technology, endometrial cancer surgery can be diagnosed through hysteroscopy, which is an effective tool.

However, hysteroscopy, relying solely on the subjective judgment of the hysteroscopists, can introduce inherent variability and potential errors in the diagnostic process [14]. Inexperienced hysteroscopists may face challenges in accurately identifying and interpreting abnormalities, leading to decreased diagnostic accuracies. This, in turn, can result in delays in initiating timely and appropriate treatment for patients, causing suboptimal outcomes and potential economic repercussions. Therefore, there is a critical demand for interventions focused on augmenting the objectivity and precision of hysteroscopy, notably through the incorporation of computer-aided diagnostic systems. These interventions aim to alleviate the challenges posed by limited clinical expertise, thereby reducing the risks associated with misdiagnosis, treatment delays, and financial implications.

II. Related Work

A. Deep Learning for Diagnosis of EC

Deep learning has been widely implemented in endometrial cancer image analysis. Zhang et al. [18] utilized MRIs from 158 patients with endometrial cancer to train and evaluate the performance of the LeNet-5 neural network. The model achieved an impressive AUC value of 0.897. Similarly, Dong et al. [2] developed and evaluated a U-Net neural network approach to determine the depth of endometrial cancer invasion using MRI scans. The model achieved a high accuracy of 79.2%. Some researchers [19] employed VGGNet-16 for endometrial lesion classification in hysteroscopic images. An-

other study [12] proposed a deep learning-based system for automatic localization of endometrial cancer.

B. Attention Mechanisms

Attention mechanisms are essential computational techniques in computer vision models, allowing them to focus on specific parts of the input data based on varying levels of importance assigned to different elements. These mechanisms play a crucial role in improving performance and enhancing understanding of visual content. In recent years, Woo et al. [17] proposed the Convolutional Block Attention Module (CBAM) to enhance the attention capabilities of Convolutional Neural Networks in image classification. In contrast, another approach known as the Bottleneck Attention Module (BAM) [9] introduces a different design. The BAM module incorporates a bottleneck structure that reduces computational complexity and memory requirements, making it more efficient while preserving attention capabilities. This module aims to strike a balance between performance and resource utilization, offering advantages over CBAM in terms of efficiency. Additionally, the Squeeze-and-Excitation (SE) module, proposed in [7], introduces a distinct attention mechanism. This SE module effectively enhances CNNs by selectively amplifying important channels and suppressing less relevant ones.

To summarize, attention mechanisms, such as the CBAM, BAM, and SE modules, have been proposed to enhance the attention capabilities of CNNs in computer vision tasks. These modules enable models to selectively focus on important regions and features in the input data, leading to improved performance, better understanding of visual content, and efficient resource utilization.

III. Methods

A. Dataset

The hysteroscopy images were collected from the Maternal and Child Hospital of Hubei Province (MCH) in 2008-2019. Table I displays the data distribution for the different categories in both the training and test sets. The categories are divided into "EC/AEH" and "Control" groups. In the training set, there are 131 cases belonging to the EC/AEH category, with a total of 3,122 images. For the control group, there are 1,106 cases with 46,434 images. The test set includes 23 cases from the EC/AEH category, comprising a total of 698 images. The control group in the test set consists of 62 cases and 2,714 images. The control group consisted of normal uterine cavities, uterine leiomyoma, endometrial polyps, and endometrial hyperplasia without atypia. Table II lists the detailed information of the control group. Pathologists diagnosed all lesions depicted in the pathological images. This study was approved by the Medical Ethics Committee of the Maternal and Child Hospital of Hubei Province. The training and analysis were conducted anonymously to comply with the privacy policy.

TABLE I
THE DATASETS FOR EC/AEH CLASSIFICATION

Category	Training set		Test set	
	Cases	Images	Cases	Images
EC/AEH	131	3,122	23	698
Control	1,106	46,434	62	2,714

TABLE II
DETAIL OF CONTROL GROUP.

Category	Training set		Test set	
	Cases	Images	Cases	Images
NE	499	10,977	41	2,151
P	260	9,166	21	563
UL	194	19,866	-	-
EH	153	6,425	-	-

B. Data Preprocessing

All hysteroscopy images were cropped to isolate the region of interest and eliminate any extraneous non-lesion areas. The hysteroscopy image augmentation process was performed using the following data transforms. The input size was first resized to 224×224 pixels, followed by center cropping to ensure consistency in image dimension. To introduce variability and increase robustness, the images underwent random affine transformations with a translation of up to 0.05 in each direction and a random horizontal flip. Finally, the augmented images were converted to tensor format and normalized with mean and standard deviation values of [0.485, 0.456, 0.406] and [0.229, 0.224, 0.225], respectively.

C. EfficientNet Neural Network

In previous convolutional neural networks, the accuracy of the network was often improved by increasing the image resolution, network width, or network depth individually. However, as the model size increased, its accuracy would often decrease. To address this challenge, Tan et al. [13] introduced a standardized approach for scaling and balancing different dimensions of the network. They used the Neural Architecture Search (NAS) technique to design the EfficientNet network, which introduced a novel model scaling method using a simple and efficient compound coefficient. The new method provides a new paradigm for model scaling that achieves enhanced performance while maintaining efficiency.

In this study, we propose an EC/AEH computer-aided diagnosis method based on EfficientNet-B0. EfficientNet-B0 is a baseline model in the EfficientNet architecture. The network structure is shown in Figure 1. Stage 1 is a convolutional layer with a convolution kernel size of 3×3. Stage 2-8 are repeatedly stacked the Mobile Inverted Bottleneck Convolution (MBConv) modules, and Stage 9 consists of a 1×1 convolutional layer, an average pooling layer, and a fully connected layer.

Stage i	Operator $\hat{\mathcal{F}}_i$	Resolution $\hat{H}_i \times \hat{W}_i$	#Channels $\hat{C}_i$	#Layers $\hat{L}_i$
1	Conv3x3	224×224	32	1
2	MBConv1, k3x3	112×112	16	1
3	MBConv6, k3x3	112×112	24	2
4	MBConv6, k5x5	56×56	40	2
5	MBConv6, k3x3	28×28	80	3
6	MBConv6, k5x5	28×28	112	3
7	MBConv6, k5x5	14×14	192	4
8	MBConv6, k3x3	7×7	320	1
9	Conv1x1 & Pooling & FC	7×7	1280	1

Fig. 1. EfficientNet-B0 network structure [13].

D. ParNet Attention

The ParNet Attention [5] is an attention mechanism, integrating elements of spatial and channel-wise feature learning with a modified Squeeze-and-Excitation (SE) block, termed as Skip-Squeeze-Excitation (SSE). The fundamental concept in this structure is the ability to capture both channel-wise and spatial information simultaneously, which is enabled by two separate convolutional layers. The ParNet structure is shown in Figure 2. A 1×1 convolutional layer is utilized for channel-wise feature extraction, while a 3×3 convolutional layer is employed for spatial context acquisition. The SSE block is a critical aspect of the ParNetAttention module, designed to overcome the limited receptive field challenge of non-deep networks with only 3×3 convolutions. The SSE design is based on the traditional SE network but features certain key modifications to better suit our objectives. Unlike the vanilla SE design, which increases the depth of the network, the SSE mechanism is applied alongside the skip connection, ensuring the network depth remains manageable. The SSE block uses Adaptive Average Pooling and a 1×1 convolution to capture global spatial information and channel-wise dependencies, respectively. The outputs from the convolutional layers and the SSE block are aggregated and passed through a Sigmoid Linear Unit (SiLU) activation function. This function introduces the desired level of non-linearity into the model, enabling it to model complex patterns effectively. The ParNet Attention module thus provides a comprehensive and nuanced attention mechanism, simultaneously capturing spatial and channel-wise features while also embedding global information through the SSE block. Its design ensures a balance between complexity and computational efficiency.

The integration of the ParNet Attention module within the EfficientNet-B0 architecture offers notable benefits over the traditional SE attention module. The ParNet Attention module, incorporating the SSE block, achieves this without increasing the network's complexity, resulting in a more resource-efficient model. By replacing the SE module with the ParNet Attention module, our model acquires a more nuanced feature representation, capturing spatial and channel-wise information simultaneously. This enhancement improves the model's comprehension of hysteroscopy images, leading to comparable performance while preserving the network depth.

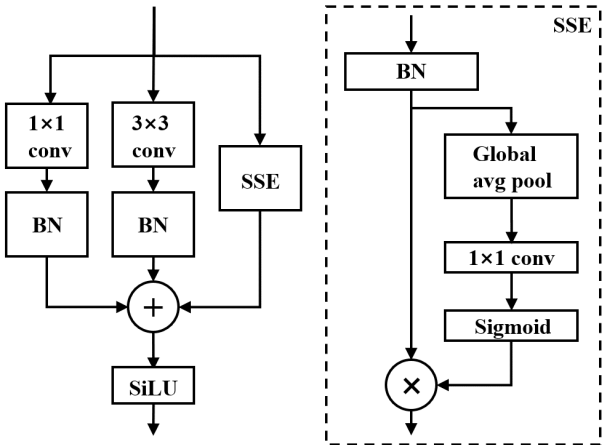

Fig. 2. ParNet attention mechanism.

The MBConv module in our model is illustrated in Figure 3.

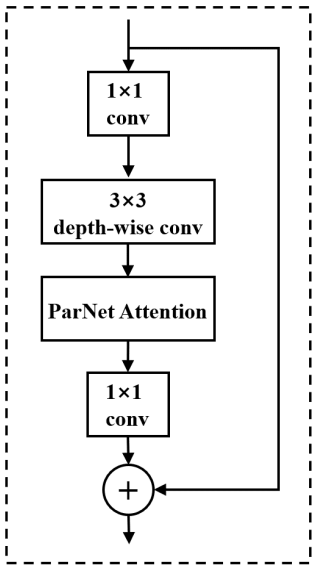

Fig. 3. MBConv module in EfficientNet-B0 based on ParNet attention mechanism.

E. Class Weighting

In our training dataset, we have 3,122 images belonging to the EC/AEH class and 46,434 images belonging to the Control class. There is a significant imbalance in the number of images between these classes. Therefore, during the training of the EfficientNet model, we employed a class weighting technique to address this issue. The class weighting involved assigning weights to the classes based on their respective sample sizes within the dataset. By applying these weights to the cross-entropy loss function, we aimed to increase the sensitivity of the loss function to prediction errors in the EC class, which has fewer samples. This approach allowed the model to focus more

TABLE III
EVALUATION RESULTS OF ABLATION EXPERIMENTS USING THE TEST SET.

	EfficientNet-B0	+ParNet attention	+Class weighting	+ParNet attention +Class weighting (Our model)
AUC(95% CI)	0.886(0.797-0.974)	0.923(0.858-0.988)	0.930(0.860-1.0)	0.941(0.891-0.990)
Accuracy (95% CI)	89.4%(89.2–89.6%)	88.2%(88.0-88.5%)	83.5%(83.2-83.8%)	89.4%(89.2-89.6%)
Sensitivity	82.6%(67.1-98.1%)	89.9%(79.8-1.0%)	89.9%(79.8-1.0%)	93.7%(87.3-1.0%)
Specificity	91.9% (85.2-98.7%)	87.1%(78.8-95.4%)	90.3%(83.0-97.7%)	87.1%(78.8-95.4%)
PPV (95% CI)	79.2%(62.9-95.4%)	72.4%(56.1-88.7%)	77.8%(62.1-93.5%)	73.3%(57.5-89.2%)
Kappa	0.735(0.573-0.898)	0.725(0.567-0.882)	0.774(0.626-0.922)	0.755(0.607-0.903)
F_1-Score	0.8086	0.8313	0.8341	0.8225

on the EC/AEH class during training, thereby addressing the class imbalance and improving its performance in the EC/AEH class.

The specific method for calculating the weights is as follows:

- For each class, calculate the reciprocal of the sample count to obtain a weight factor. For class EC/AEH, the weight factor is 1/3211; for class control, the weight factor is 1/46434.
- Normalize the weight factors so that their sum equals 1. Specifically, divide each weight factor by the sum of all weight factors to obtain normalized weights. For class EC/AEH, the normalized weight is 0.935; for class control, the normalized weight is 0.065.
- Utilize the normalized weights as class weights in the calculation of the cross-entropy loss function.

F. Evaluation Metrics

We utilize the classification performance of the model at the case level as a benchmark for comparing and evaluating the performance of different models. The predicted malignant score for each case was calculated following the approach of Zhou et al. [21]. The formula is shown in equation 1:

$$\theta = -\left[w_1 \times \log_{10}(1-p_1) + [w_2 \times \log_{10}(1-p_2) + \cdots + [w_n \times \log_{10}(1-p_n)]/n \right. \tag{1}$$

where n is the total number of images from the case, $P_{malignancy} = [p_1, p_2, \ldots, p_n]$ is the predicted probabilities of these n images classified as malignancy. w_1 is calculated as $w_i = p_i/(p_1 + p_2 + \cdots + p_n)$.

The classification performance of the models was evaluated using AUC, accuracy, sensitivity, specificity, PPV, F_1-Score Kappa, and related 95% confidence interval (CI). All values for the evaluation metrics were computed using the reportROC package (version 3.5) in the R programming language. The evaluation metrics are defined by the following equations.

$$Accuracy = \frac{TP + TN}{TP + TN + FN + FP}, \tag{2}$$

$$Sensitivity = \frac{TP}{TP + FN}, \tag{3}$$

$$Specificity = \frac{TN}{TN + FP}, \tag{4}$$

$$PPV = \frac{TP}{TP + FP}, \tag{5}$$

$$F_1 - Score = \frac{TP}{(FP + FN)/2 + TP} \tag{6}$$

where TP (True Positive), FP (False Positive), TN (True Negative), and FN (False Negative) are the values derived from the confusion matrix.

IV. RESULTS AND DISCUSSION

The proposed model was developed with Pytorch framework. A Stochastic Gradient Descent optimizer was used with a learning rate of 0.01, a momentum of 0.9, and a decay rate of 4e-4. The training epoch is set to 50 and batch size is set to 20. The model was configured with a scaling factor of 1.0 for both the channel dimension and the depth dimension. Within the MBConv module, the dropout layer was set with a dropout rate of 0.2 to introduce random dropout. The dropout rate of 0.2 was applied before the last fully connected layer of the model.

To assess the efficacy of the improvements to the Efficient-B0 model, we applied it in ablation experiments. Table III shows a detailed comparison of the classification performance achieved by different models on the test set. Our proposed model, incorporating ParNet attention and class weighting modifications, achieved the highest AUC of 0.941 compared to other models with AUCs of 0.886, 0.923, and 0.930. This indicates the superior discriminative ability of our model in accurately classifying EC/AEH from benign lesions. Specifically, our proposed model reached accuracy of 89.4%, sensitivity of 93.7%, specificity of 87.1%, PPV of 73.3%, Kappa of 0.755, and F_1-Score of 0.82. These results demonstrate the efficiency of combining ParNet attention and class weighting techniques to enhance EfficientNet-B0 performance. The ROC curves of different models are shown in Figure 4. Figure 5 shows the loss curve of our proposed model. Some examples of the classification results of our proposed model on the test set are displayed in Figure 6.

V. CONCLUSION

In this study, we improved the EfficientNet-B0 model with ParNet attention and class weighting to achieve higher AUC of 0.941 in the classification EC/AEH lesions from cases. Our

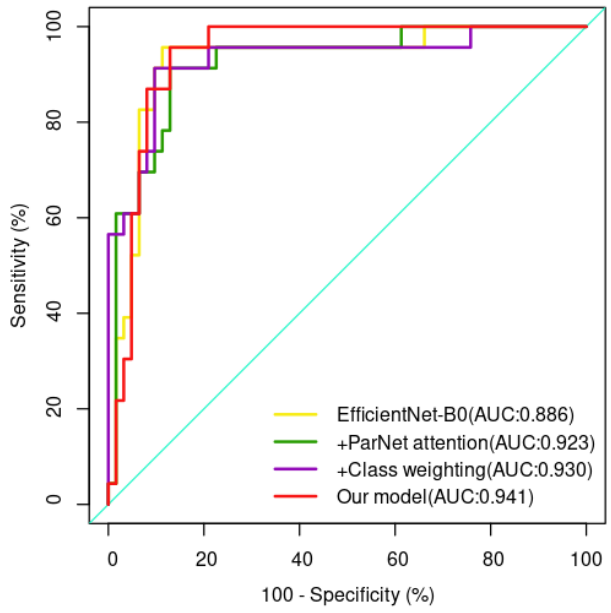

Fig. 4. ROC curves of the models.

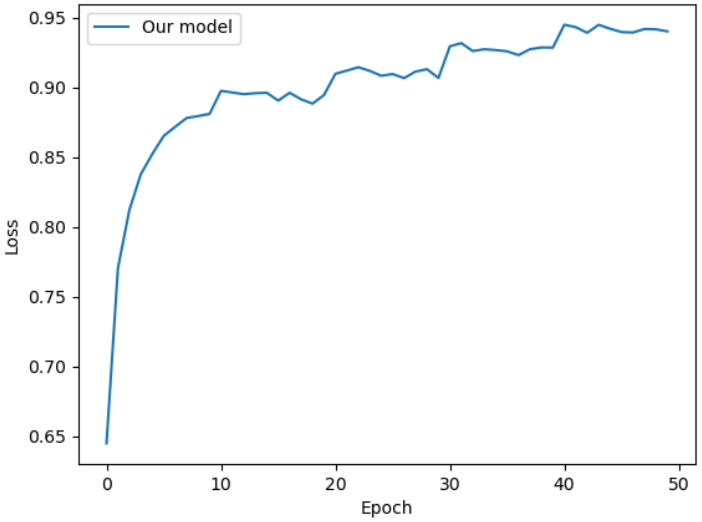

Fig. 5. The Loss value curve of our model during training.

proposed model may be useful in diagnosing of endometrial cancer and non-cancerous. In the future, we will evaluate the efficacy of this study in practical clinical applications.

REFERENCES

[1] Dusan Djokovic and Amal Drizi. Diagnostic hysteroscopy: patient assessment and preparation.

[2] Hsiang-Chun Dong, Hsiang-Kai Dong, Mu-Hsien Yu, Yi-Hsin Lin, and Cheng-Chang Chang. Using deep learning with convolutional neural network approach to identify the invasion depth of endometrial cancer in myometrium using mr images: a pilot study. *International journal of environmental research and public health*, 17(16):5993, 2020.

[3] Nishat Fatima, Gayatree Bharti, and Nishi Mishra. Role of hysteroscopy in the diagnosis of endometrial cancer.

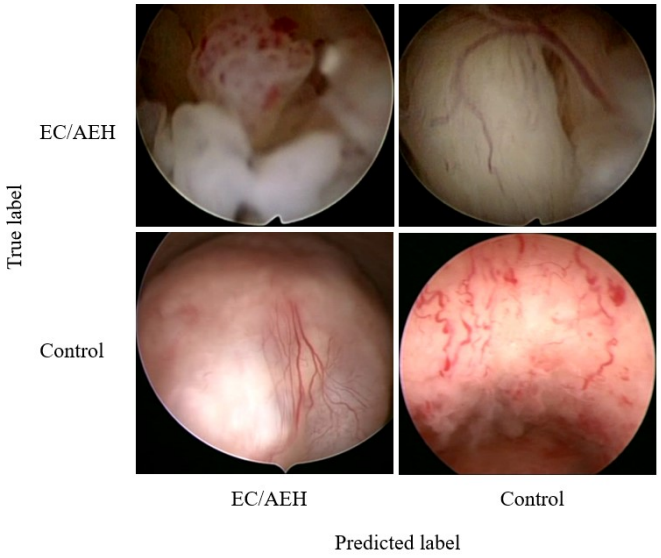

Fig. 6. Example classification results output by our model.

[4] Emily Goebel, August Vidal, Xavier Matias-Guiu, and C. Gilks. The evolution of endometrial carcinoma classification through application of immunohistochemistry and molecular diagnostics: past, present and future. *Virchows Archiv*, 472, 06 2018.

[5] Ankit Goyal, Alexey Bochkovskiy, Jia Deng, and Vladlen Koltun. Non-deep networks. *Advances in Neural Information Processing Systems*, 35:6789–6801, 2022.

[6] Jian-Zeng Guo, Qi-Jun Wu, Fang-Hua Liu, Chang Gao, Ting-Ting Gong, and Gang Li. Review of mendelian randomization studies on endometrial cancer. *Frontiers in Endocrinology*, 13, 2022.

[7] Jie Hu, Li Shen, and Gang Sun. Squeeze-and-excitation networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2018.

[8] Misato Kamii, Yoko Nagayoshi, Kazu Ueda, Motoaki Saito, Hirokuni Takano, and Aikou Okamoto. Laparoscopic surgery for atypical endometrial hyperplasia with awareness regarding the possibility of endometrial cancer. *Gynecology and Minimally Invasive Therapy*, 12(1):32, 2023.

[9] Jongchan Park, Sanghyun Woo, Joon-Young Lee, and In So Kweon. Bam: Bottleneck attention module. *arXiv preprint arXiv:1807.06514*, 2018.

[10] Kelly Passarello, Shiney Kurian, and Valerie Villanueva. Endometrial cancer: an overview of pathophysiology, management, and care. In *Seminars in oncology nursing*, volume 35, pages 157–165. Elsevier, 2019.

[11] Hyuna Sung, Jacques Ferlay, Rebecca L Siegel, Mathieu Laversanne, Isabelle Soerjomataram, Ahmedin Jemal, and Freddie Bray. Global cancer statistics 2020: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: a cancer journal for clinicians*, 71(3):209–249, 2021.

[12] Yu Takahashi, Kenbun Sone, Katsuhiko Noda, Kaname Yoshida, Yusuke Toyohara, Kosuke Kato, Futaba Inoue, Asako Kukita, Ayumi Taguchi, Haruka Nishida, et al. Automated system for diagnosing endometrial cancer by adopting deep-learning technology in hysteroscopy. *PLoS One*, 16(3):e0248526, 2021.

[13] Mingxing Tan and Quoc Le. Efficientnet: Rethinking model scaling for convolutional neural networks. In *International conference on machine learning*, pages 6105–6114. PMLR, 2019.

[14] Steffi van Wessel, Tjalina Hamerlynck, Benedictus Schoot, and Steven Weyers. Hysteroscopy in the netherlands and flanders: a survey amongst practicing gynaecologists. *European journal of obstetrics & gynecology and reproductive biology*, 223:85–92, 2018.

[15] Monica Hagan Vetter, Blair Smith, Jason Benedict, Erinn M Hade, Kristin Bixel, Larry J Copeland, David E Cohn, Jeffrey M Fowler, David O'Malley, Ritu Salani, et al. Preoperative predictors of endometrial cancer at time of hysterectomy for endometrial intraepithelial neoplasia

or complex atypical hyperplasia. *American journal of obstetrics and gynecology*, 222(1):60–e1, 2020.

[16] Peng-Hui Wang, Szu-Ting Yang, Chia-Hao Liu, Wen-Hsun Chang, Fa-Kung Lee, and Wen-Ling Lee. Endometrial cancer: part i. basic concept. *Taiwanese Journal of Obstetrics and Gynecology*, 61(6):951–959, 2022.

[17] Sanghyun Woo, Jongchan Park, Joon-Young Lee, and In So Kweon. Cbam: Convolutional block attention module. In *Proceedings of the European conference on computer vision (ECCV)*, pages 3–19, 2018.

[18] Yan Zhang, Cuilan Gong, Ling Zheng, Xiaoyan Li, and Xiaomei Yang. Deep learning for intelligent recognition and prediction of endometrial cancer. *Journal of Healthcare Engineering*, 2021, 2021.

[19] YunZheng Zhang, ZiHao Wang, Jin Zhang, CuiCui Wang, YuShan Wang, Hao Chen, LuHe Shan, JiaNing Huo, JiaHui Gu, and Xiaoxin Ma. Deep learning model for classifying endometrial lesions. *Journal of Translational Medicine*, 19:1–13, 2021.

[20] Aihua Zhao, Xin Du, Suzhen Yuan, Wenfeng Shen, Xin Zhu, and Wenwen Wang. Automated detection of endometrial polyps from hysteroscopic videos using deep learning. *Diagnostics*, 13(8), 2023.

[21] Dejun Zhou, Fei Tian, Xiangdong Tian, Lin Sun, Xianghui Huang, Feng Zhao, Nan Zhou, Zuoyu Chen, Qiang Zhang, Meng Yang, Yichen Yang, Xuexi Guo, Zhibin Li, Jia Liu, Jiefu Wang, Junfeng Wang, Bangmao Wang, Guoliang Zhang, Baocun Sun, and Xiangchun Li. Diagnostic evaluation of a deep learning model for optical diagnosis of colorectal cancer. *Nature Communications*, 11, 06 2020.