Journal on Oncology

www.journalononcology.org

Research Article

Open Access, Volume 3

Computer-Aided Diagnosis System Based on Deep Learning for Diagnosis of Atypical Endometrial Hyperplasia and Endometrial Cancer on Hysteroscopy

Wenwen Wang[1]†; Zhe Guo[2,5]†; Aihua Zhao[2]; Wenqing Ma[1]; Yan Wang[3]; Xin Du[3]; Wuliang Wang[4]; Xin Zhu[2]*; Shixuan Wang[1]*

[1]Department of Obstetrics and Gynecology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China.

[2]Graduate School of Computer Science and Engineering, The University of Aizu, Aizu-Wakamatsu, Japan.

[3]Department of Gynecology, Maternal and Child Hospital of Hubei Province, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China.

[4]Department of Obstetrics and Gynecology, The Second Affiliated Hospital of Zhengzhou University, Zhengzhou, Henan, China.

[5]Department of Laboratory Medicine, The Second Xiangya Hospital, Central South University, Changsha, China.

†W Wang and Z Guo have equivalent contributions.

Abstract

Hysteroscopy is an effective tool to diagnose and treat abnormal uterine bleeding or uterine cavity abnormalities, especially precancerous or cancerous disorders. In this study, we proposed an Endometrial Cancer (EC) Computer-Aided Diagnosis System (ECCADx) based on deep learning to boost the diagnostic accuracy for recognizing Atypical Endometrial Hyperplasia (AEH) and EC. ECCADx was developed using a training dataset with 49,646 images from 1,237 patients in Maternal and Child Hospital of Hubei Province (MCH) and two test datasets with 7,243 images from 209 patients in MCH, Tongji Hospital (TJH), and The Second Affiliated Hospital of Zhengzhou University (ZZSH). We compared the diagnostic efficiency between the proposed system and eight gynecological endoscopists from two different hospitals (MCH and TJH) using a hospital cross-testing method. The sensitivity, specificity and AUC of ECCADx were 92.8% (95% CI 85.7-100%), 92.5% (95% CI 86.7-98.3%), and 0.965 (95% CI 0. 931-1) on MCH test dataset (internal data), respectively, superior to two gynecological endoscopists and having no significant difference compared with the other two endoscopists at TJH. For TJH/ZZSH test dataset (external data), the sensitivity, specificity and AUC were 75.2% (95% CI 59.5-90.8%), 95.2% (95% CI 91.5-99.0%), and 0.881 (95% CI 0.789-0.967), respectively, superior to three gynecological endoscopists and having no significant difference compared with the other endoscopist at MCH. ECCADx demonstrated excellent performance in identifying AEH and EC in test datasets from different medical centers. The effectiveness of ECCADx was comparable or even better than those of experienced gynecological endoscopists.

Keywords: Atypical endometrial hyperplasia; Endometrial cancer; Hysteroscopy; Deep learning.

Manuscript Information: *Received: Mar 12, 2024; Accepted: Apr 12, 2024; Published: Apr 19, 2024*

Correspondance: *Xin Zhu[1] & Shixuan Wang[2]*

[1]Graduate School of Computer Science and Engineering, The University of Aizu, Aizu-Wakamatsu, Japan. Email: zhuxin@u-aizu.ac.jp

[2]Department of Obstetrics and Gynecology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China. Email: shixuanwang@tjh.tjmu.edu.cn

Citation: *Wang W, Guo Z, Zhao A, Zhu X, Wang S, et al. Computer-Aided Diagnosis System Based on Deep Learning for Diagnosis of Atypical Endometrial Hyperplasia and Endometrial Cancer on Hysteroscopy. J Oncology. 2024; 4(1): 1130.*

Introduction

Endometrial Cancer (EC) is a common gynecological malignant condition with a rising incidence worldwide [1]. High 5-year survival of EC patients relies on early diagnosis and treatment [2]. Atypical Endometrial Hyperplasia (AEH) is a precancerous condition of EC, and up to 40% of AEH would become EC without timely hysterectomy [3]. Furthermore, approximately 30% of AEH would develop into cancer within one year [4]. Considering the rapid progress of EC and AEH lesions, the accurate detection of EC and AEH is of utmost importance for early and effective diagnosis and treatment.

Hysteroscopic-guided curettage has been widely considered to be a useful tool to tailor treatment in patients with uterine malignancies. Hysteroscopy showed higher diagnosis performance than that of Dilation and Curettage (D&C) alone [5]. A meta-analysis by Gkrozou et al. involving over 9,000 patients assessed the accuracy of hysteroscopy in the diagnosis of polyps, submucosal myomas, hyperplasia and endometrial cancer, demonstrating a high diagnostic accuracy for endometrial cancer with a sensitivity of 82.6% and a specificity of 99.7% [6].

However, misdiagnosis could underestimate the risk of uterine conditions, leading to a treatment delay. A recent meta-analysis on 1,106 patients, with a preoperative diagnosis of atypical endometrial hyperplasia, showed an underestimation of endometrial cancer up to 32.7-45.3% following uterine curettage and hysteroscope guided biopsy [7]. Similarly, another systematic review and meta-analysis evaluating D&C and hysteroscopy in diagnosing cancer from women with postmenopausal bleeding, demonstrated that a high failure rate [11% (range 1-53%)] and infeasible endometrial samples [31% (range 7-76%)] would lead to a missing diagnosis in average 7% (0-18%) of cases [8]. Considering the high risk caused by missing diagnosis, a better diagnosis assist tool is urgently needed for enhancing the accuracy of this evaluation.

Recently, deep learning has been widely applied in endoscopy, especially for the detection of polyp, adenoma or gastrointestinal cancer using colonoscopy, gastroscopy, hysteroscopy, and capsule endoscopy [9-11]. In a single center study on hysteroscopy, a method was proposed for the classification of endometrial lesions and developed using 6,728 hysteroscopic images from 454 patients, and showed a 90.8% of accuracy, 83% of sensitivity and 96% of specificity for identifying lesions of benign or premalignant/malignant [10].

In this study, we performed a multicohort retrospective study involving 1,446 cases from three tertiary hospitals for the development and validation of an Endometrial Cancer Computer-Aided Diagnosis (ECCADx) system based on deep learning for identifying AEH and EC from benign lesions.

Materials and methods

Study design and participants

This multicohort retrospective study was conducted in three tertiary hospitals. A total of 1,446 cases with 55,874 images in png format were enrolled consecutively. The numbers of cases and images in training and test datasets were listed in Table 1. Images were captured by one of three high resolution devices (Olympus OTV-S190, Japan; Karl storz 26105FA or 26120BA, Germany). Pathological images of all lesions had been diagnosed by pathologists. All images have been confirmed by two experts WW.W. and W.M. The control category (benign lesions) included cases with endometrial polyps, submucosal uterine leiomyoma, endometrial hyperplasia without atypia and normal uterine cavities (details can be found in Table A1).

The training set was retrieved and collected from January 2008 to December 2017 at Maternal and Child Hospital of Hubei Province (MCH) by Olympus OTV-S190, Japan and Karl storz 26105FA or 26120BA, Germany. The internal test dataset was made up of images collected from January 2018 to June 2019 at MCH by the same devices. The external test dataset contained data obtained from January 2019 to December 2019 at Tongji Hospital of Huazhong University of science and technology (TJH) and the second affiliated hospital of Zhengzhou University (ZZSH). AEH/EC categories included cases with endometrial atypia hyperplasia and endometrial cancer. The external test datasets were mainly obtained by Olympus OTV-S190, Japan. The training and test datasets have no case overlap.

We recruited four gynecological endoscopists from either hospital of MCH and TJH to assess the counterpart's test dataset (TJH/ZZSH or MCH). The four endoscopists from either of the two hospitals included two senior endoscopists with at least 15 years of clinical experiences and two intermediate endoscopists with more than 6 years of clinical experiences. This study is the first attempt for multi-level evaluation of endometrial lesions. All eight endoscopists evaluated all images of each patient as "Must be benign", "Most likely to be benign", "May be benign", "May be to be malignant", "Must be malignant", according to their clinical experiences. Figure 1 illustrates the flowchart on the development and estimate of ECCADx.

This study was approved by Medical Ethics Committee of Tongji Hospital Affiliated to Tongji Medical College of Huazhong University of Science and Technology, Maternal and Child Hospital of Hubei Province and the Second Affiliated Hospital of Zhengzhou University. To comply with the privacy policy, the training and analysis were conducted anonymously.

Training and test datasets

Detailed information of training and test datasets is illustrated and listed in Figure 1 and Table 1. Additional two test datasets include 3,419 images from 23 AEH/EC and 62 control cases diagnosed between January 2018 and June 2019 at MCH, and 2,809 images from 24 AEH/EC and 100 control cases diagnosed between January 2019 to December 2019 at TJH/ZZSH, respectively. The former test dataset serves as an internal test dataset, and the latter one an external test dataset. Information of non-cancerous disorders in the training and test datasets can be found in Table A1. In the test datasets, all extracted images were put to use to estimate the efficiency of ECCADx and endoscopists. The test datasets from MCH and TJH/ZZSH were evaluated by endoscopists from TJH and MCH, respectively.

Model development

In this study, we proposed a convolutional neural network with a backbone of ResNet-50 [12] for the analysis of hysteroscopic images. ResNet-50 is a 50-layer convolutional neural network pretrained with over 100 million images in the ImageNet database

[13]. Skip shortcuts used in ResNet50 [12] mimicking pyramidal cells in cerebral cortex are employed to improve the performance of convolutional neural networks. Image crops and resizing were performed for all images in advance because images obtained by different hysteroscopes have different image sizes and excessive black background.

To overcome data unbalance caused by less cases of malignancies, i.e., AEH/EC, a focal loss in (1) multiplying the cross-entropy function in (2)(3) with a modulating factor is used in the proposed model to increase the sensitivity of misclassified AEH/EC observations [14]. Besides, an «over-sampling» technology was used to compensate the impact of data imbalance in the training dataset [15]. We also employed image augmentation [16] by generating additional training data to prevent overfitting and improve performance. Data augmentation was performed automatically including image scaling, translation, rotation, and reflection. Furthermore, all training data was resized to 224*224 pixels to be analyzed by ResNet50.

$$\text{focal loss } (p_t)=(1-p_t)^{0.1}\text{CE}(p,y) \quad (1)$$

$$\text{CE}(p,y)=-\log(p_t) \text{ and} \quad (2)$$

$$p_t=\begin{cases}-\log(p) & \text{if } y=1\\ -\log(1-p) & \text{otherwise}\end{cases} \quad (3)$$

where p is the model's estimated probability [14] for AEH/EC, and y is ground truth (1: AEH/EC; 0: control).

For endoscopists, "Must be benign", "Most likely to be benign", "May be benign", "May be malignant", "Most likely to be malignant", "Must be malignant" were set with AEH/EC probabilities of 0, 0.2, 0.4, 0.6, 0.8, and 1, respectively. These probabilities were used to calculate Receiver Operating Characteristic (ROC) curves and Area Under Curves of (AUC) of endoscopists.

The proposed method was developed with MATLAB R2020a (The MathWorks, Inc. US), and Deep Learning Toolbox™ and Parallel Computing Toolbox™. We «freeze» the initial 10 layers in the network by setting the learning rate to zero to prevent overfitting and also speed up network training. A stochastic gradient descent optimizer was used with a learning rate of 0.01, a momentum of 0.9, a decay rate of 0.1 every 10 epoch, and training epochs of 30. The hyperparameters were set by trials and errors.

Statistical analysis

The classification efficiency of ECCADx was evaluated using ROC curves, AUC, accuracy, sensitivity, specificity, Positive Predictive Value (PPV), Negative Predictive Value (NPV), F1 score, Kappa, Brier, and related 95% confidence interval (CI). All these statistical analyses were performed by R (version 4.0.2) programing language (R Development Core Team). Report ROC package (version 3.5) was used to calculate AUC, accuracy, sensitivity, specificity, PPV, NPV; irr package (version 0.84.1) to calculate Fleiss' Kappa and two-sided z-test; measures package (version 0.3) to calculate F1 score and Brier score; pROC package (version 1.17.0.1) to compute AUC, and confirm whether there is significant difference in the AUCs between ECCADx and endoscopists using DeLong's test. Accuracy, sensitivity, specificity, PPV, NPV are defined in the following equations.

$$Accuracy=\frac{TP+TN}{TP+TN+FP+FN} \quad (4)$$

$$Sensitivity=\frac{TP}{TP+FN} \quad (5)$$

$$Specificity=\frac{TN}{TN+FP} \quad (6)$$

$$PPV=\frac{TP}{TP+FP}$$

$$NPV=\frac{TN}{TN+FN} \text{ and}$$

$$F1Score=\frac{TP}{TP+(FN+FP)/2} \quad (9)$$

Where TP, TN, FP, FN indicates true positive, true negative, false positive, and false negative, respectively.

Finally, a predictive score of each lesion [16] was calculated from predicted probabilities of images classified as malignancies, i.e., AEH/EC. The predictive score is then used for the classification of AEH/EC and control.

Results

Performance of models on test datasets

ECCADx was trained and used to estimate the performance of the proposed model on two test datasets. Figure 2 illustrates the ROC curves of ECCADx and endoscopists in identifying AEH/EC. As listed in Table 2, AUC value, accuracy, sensitivity, specificity, and F1 of ECCADx was 0.965 (95%CI 0.931-1), 94.4% (95%CI 90.1-98.8%), 92.8% (95% CI 85.7-100%), 92.5% (95% CI 86.7-98.3%) and 0.939, respectively on the MCH test dataset. This indicated a nearly perfect discriminative ability. No significant difference was observed between the AUCs of ECCADx and endoscopists. For TJH/ZZSH test dataset, the AUC, accuracy, sensitivity, specificity, and F1 were 0.881 (95% CI 0.789-0.972), 92.2% (95% CI 87.8-96.7%), 75.2% (95% CI 59.5-90.8%), 95.2% (95% CI 91.5-99.0%), and 0.826, respectively. No significant difference was observed between the AUCs of ECCADx and the endoscopist (MCH-Exp2) with the best performance. Other evaluation metrics such as PPV, NPV, kappa coefficient and Brier were listed in Tables 2 and 3.

Six false negative cases of ECCADx included two cases with polyp cystic degeneration, and one with myomatoid. There is no abundant blood vessel distribution among them. The other two false negative cases had typical lesions but poor image quality due to necrotic tissues attached to the surfaces of lesions. The lesion was missed in the images of the final case.

Performance of deep learning versus endoscopists

For MCH test dataset, we compared ECCADx with the TJH senior endoscopist with the largest AUC value in Table 2, AUC of 0.965 (95% CI 0.931-1) vs 0.974 (95% CI 0.947-1), accuracy of 94.4% (95%CI 90.1-98.8%) vs 95.6% (95%CI 91.8-99.4%), and F1 metric of 0.939 vs 0.958 as listed in Table 2. For TJH/ZZSH test dataset, ECCADx and one senior endoscopist from MCH reached an AUC of 0.894 (95% CI 0.807-0.981) vs 0.881 (95% CI 0.789-0.972), accuracy of 84.4% (95% CI 78.2-90.6%) vs 92.2% (95% CI 87.8-96.7%) and F1 metric of 0.719 vs 0.826 as listed in Table 3. Other evaluation metrics such as sensitivity, specificity, negative predictive value, and Kappa coefficient for ECCADx and endoscopists were also listed in Tables 2 and 3. The interrater agreement rate for the four experienced endoscopists from TJH was 62.4% (Fleiss'

Kappa 0.58; two-sided z-test, p<0.001) in MCH test dataset and for four experienced endoscopists from MCH was 37.1% (Fless's Kappa 0.322; two sided z-test, p<0.001) in TJH/ZZSH test dataset.

Grad-CAM algorithm [17] was used to confirm important regions for predicting AEH/EC by ECCADx. These regions highlighted in Figure 3 may contain important morphological and vascular features such as a gross distortion of endometrial cavity, focal necrosis, friable consistency, and atypical vessels related to different pathological patterns of AEH and EC [18]. These features may play a crucial role in ECCADx for recognizing AEH and EC.

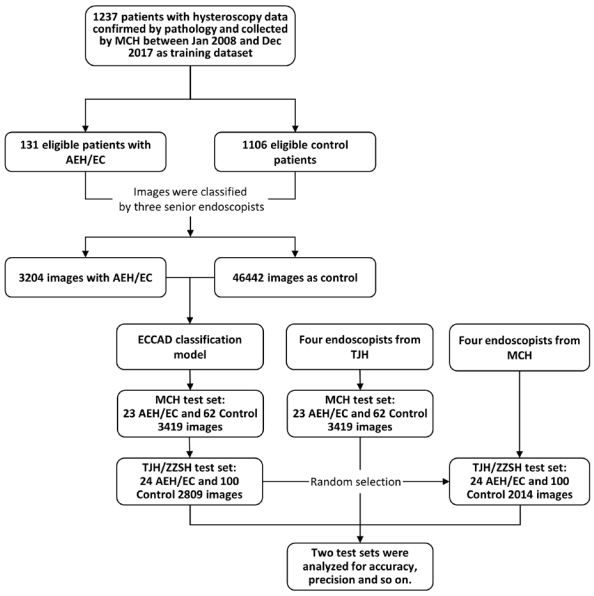

Figure 1: Flowchart on development and validation of ECCADx.

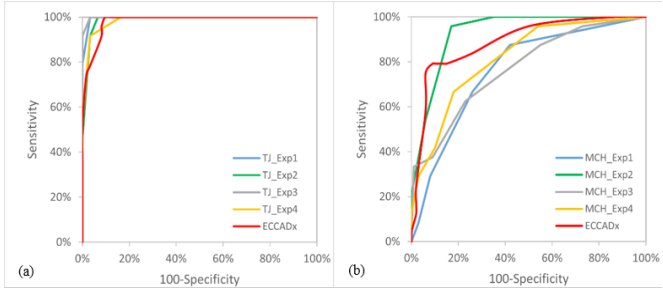

Figure 2: This is a figure. Schemes follow the same formatting.

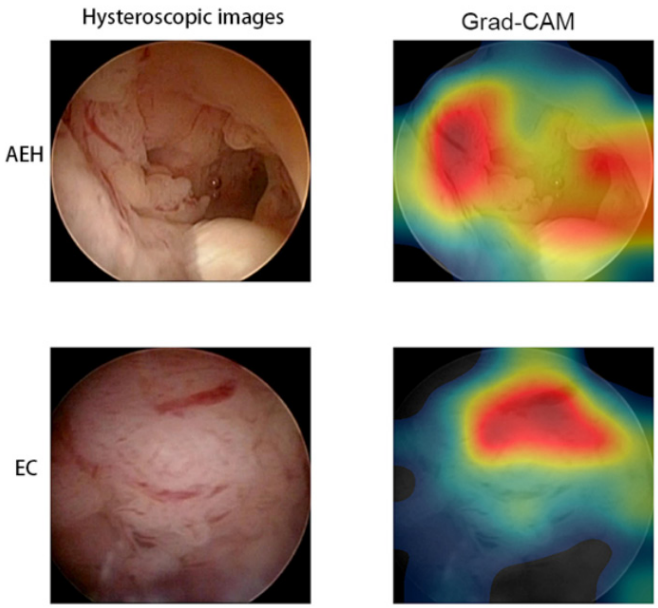

Figure 3: This is a figure. Schemes follow the same formatting.

Table A1: Information of non-cancerous disorders.

	Training dataset	MCH test dataset	TJH/ZZSH test dataset
P	260	21	53
NE	499	41	36
UL	194	-	9
EH	153	-	2

P: Polyp; NE: Normal Endometrium; UL: Uterine Leiomyomata; EH: Endometrial Hyperplasia.

Table 1: Baseline characteristics of training and test datasets.

	MCH training dataset		MCH test dataset		TJH/ZZSH test dataset	
	AEH/EC[1]	Control	AEH/EC	Control	AEH/EC	Control
Cases	131	1,106	23	62	24	100
Images	3,204	46,442	698	2,721	760	2,049

[1]AEH/EC: endometrial atypia hyperplasia and endometrial cancer

Table 2: Per patient diagnostic performance of endoscopists versus ECCADx in the MCH test dataset.

	Gynecological endoscopist				ECCADx
	TJ-Exp1	TJ-Exp2	TJ-Exp3	TJ-Exp4	
AUC (95% CI)	0.965(0.931-1)	0.951(0.902-1)	0.974(0.947-1)	0.911(0.828-0.995)	0.965(0.931-1)
P value (Exp vs ECCADx)*	0.41	0.62	0.25	0.48	-
Accuracy (95% CI)	94.4% (90.1-98.8%)	92.2% (87.0-97.5%)	95.6% (91.8-99.4%)	85.4% (78.3-92.6%)	94.4% (90.1-98.8%)
Sensitivity	92.8% (85.7-100%)	92.8% (85.7-100.0%)	92.8% (85.7-100.0%)	92.8% (85.7-100.0%)	92.8% (85.7-100%)
Specificity	92.5% (86.7-98.3%)	89.5% (82.5-96.5%)	94.0% (89.0-99.1%)	80.4% (71.0-89.8%)	92.5% (86.7-98.3%)
PPV (95% CI)	83.5% (71.0-96.0%)	78.2% (64.4-92.1%)	86.4% (75.0-97.8%)	65.9% (50.8-80.9%)	83.5% (71.0-96.0%)
NPV (95% CI)	97.0% (93.9-100%)	96.8% (93.7-100.0%)	97.0% (94.0-100.0%)	96.5% (93.0-100.0%)	97.0% (93.9-100%)
F1	0.939	0.902	0.958	0.807	0.939
Kappa	0.914	0.861	0.942	0.715	0.914
Brier	0.058	0.069	0.027	0.112	0.108

*DeLong's test

Table 3: Per patient diagnostic performance of endoscopists versus ECCADx in the TJH/ZZSH test datasets.

	Gynecological endoscopist				ECCADx
	MCH-Exp1	MCH-Exp2	MCH-Exp3	MCH-Exp4	
AUC (95% CI)	0.728 (0.605-0.85)	0.894 (0.807-0.981)	0.698 (0.571-0.824)	0.709 (0.584-0.834)	0.881 (0.789-0.972)
P value (Exp vs ECCADx)	0.03	0.66	0.008	0.04	-
Accuracy (95% CI)	63.3% (55.0-71.6%)	84.4% (78.2-90.6%)	73.4% (65.8-81.1%)	55.5% (46.9-64.1%)	92.2% (87.8-96.7%)
Sensitivity	82.30% (69.0-95.7%)	89.60% (79.8-99.3%)	60.8% (42.7-78.8%)	89.6% (79.8-99.3%)	75.2% (59.5-90.8%)
Specificity	57.70% (48.2-67.2%)	81.8% (74.5-89.1%)	76.0% (67.8-84.2%)	46.2% (36.6-55.7%)	95.2% (91.5-99.0%)
PPV (95% CI)	34.2% (22.9-45.6%)	56.8% (42.2-71.5%)	40.4% (25.6-55.3%)	30.8% (20.8-40.8%)	81.0% (66.7-95.3%)
NPV (95% CI)	92.4% (86.5-98.3%)	96.7% (93.6-99.8%)	87.8% (81.3-94.4%)	94.2% (88.9-99.6%)	93.4% (89.0-97.9%)
F1	0.483	0.719	0.484	0.455	0.826
Kappa	0.281	0.629	0.323	0.227	0.787
Brier	0.166	0.116	0.181	0.125	0.098

Discussion

The results demonstrated that ECCADx has a sensitivity and specificity nearly equivalent to experienced endoscopists in identifying AEH/EC patients of 2 test datasets from different medical centers.

The advantage of our ECCADx lies in recognizing AEH/EC from non-cancerous lesions including polyps, submucosal uterine leiomyoma, endometrial hyperplasia without atypia, and normal uterine cavity. Moreover, the proposed system maintains the stability of diagnostic capabilities in datasets from different medical centers. The combination of ECCADx and hysteroscopy systems could balance the diagnostic efficiency of endoscopists with diverse working experiences and speed up the diagnosis process. Meanwhile, the proposed system may serve as a second observer to enhance the ability of endoscopists to deal with patients at high risk of AEH/EC and reduce the misdiagnosis and unnecessary biopsy due to the perceptual bias and visual fatigue by endoscopists.

The establishment of ECCADx was based on the dataset over 9 years from a single hospital. Restricted sample size, population distribution, discrepancy devices and uneven image quality would lead to model instability in the analysis of other datasets. To comfirm this shortcoming, geographical and temporal test datasets from two other hospitals as external test data were used here to verify the classification ability of this model. The validation result reflects a true diagnostic ability of ECCADx in processing images from different devices with diverse quality and subject distributions. As we have introduced before, the training and internal test dataset were obtained by Olympus OTV-S190, Japan or Karl storz 26105FA or 26120BA, Germany, and the external data by Olympus OTV-S190, Japan. This may explain why ECCADx has less performance in validation using external dataset that that using internal data. Nevertheless, EXCADx demonstrated nearly equivalent performance to experienced endoscopists.

Hysteroscopic-guided curettage can accurately remove benign lesions and therefore reduce the probability of endometrial injury. However, this may bring a risk of missed diagnosis for precancerous/malignant lesions, which are recommended to be removed by hysterectomy and Bilateral Salpingo-Oophorectomy (TH/BSO) [19]. Therefore, an underestimated diagnosis could lead to treatment delay. A computer-aided diagnosis system such as ECCADx can play an important role in helping endoscopists identifying various precancerous and malignant lesions from benign ones.

Machine learning has been widely applied in gastrointestinal endoscopy system for the detection and classification of disorders [20,21]. However, only very few studies were conducted in hysteroscopy using computer-aided diagnosis. Neofytou et al. presented a computer-aided diagnosis system for the early detection of endometrial cancer [22]. The CADx system was validated using 516 Regions of Interest (ROIs) extracted from 52 subjects. In terms of ROI classification, the best results were achieved by using Statistical Features (SFs) and Gray-Level Difference Statistics (GLDS) features with an SVM classifier. For this combination, the proposed CAD system achieved an 81% correct classification rate [21]. Recently, Ma et al.'s team used VGGNet-16 model to classify endometrial lesions, and got a sensitivity of 84.0%, 68.0%, 78.0%, 94.0%, and 80.0% as endometrial hyperplasia without atypia, atypical hyperplasia, endometrial cancer, endometrial polyp, and submucous myoma [10]. Compared with these two models, ECCADx demonstrated superior performance for identifying endometrial cancer in a larger number of cases from multiple medical centers.

In general, according to the morphological and vascular patterns of AEH and EC, gynecological endoscopists could recognize AEH/EC from benign images. However, lower inter-rater agreement among gynecological endoscopists were observed. Especially for the testing results by the TJH/ZZSH test dataset, the agreement rates were from 48.3% to 71.8%. It might attribute to the lower ratio of AEH/EC patients in TJH/ZZSH dataset (24/124) than that in MCH dataset (23/85). This caused difficulty in identifying cases with malignancy tumors. In addition, the proposed model was trained by data only from MCH, and demonstrated lower performance in the analysis of the data from TJH/ZZSH possible because of inter-hospital difference mentioned above.

The limitations of this study are as follows

(1) Binary class model. ECCADx can only identify AEH/EC and non-cancerous disorders. The next step is to distinguish atypical hyperplasia and various pathological types of endometrial cancer, which can better guide the treatment strategies.

(2) Retrospective study. The application and evaluation of ECCADx should be taken out in a multicentral prospective study in the future.

Conclusion

The proposed ECCADx demonstrated satisfying performance in identifying AEH/EC lesions from cases in different medical centers. The effectiveness of ECCADx was comparable or even better than those of experienced gynecological endoscopists. In the future, this model should be validated in a prospective randomized study in multicenter for the evaluation of its clinical usefulness.

Declarations

Funding: This study and APC was funded by National Natural Science Foundation of China (No. 81701420), and the Competitive Research Fund from The University of Aizu (P-12-2022). The funding sources are not involved in the performance of the research nor in the preparation of this manuscript.

Institutional review board statement: The study was conducted in accordance with the Declaration of Helsinki, and approved by Medical Ethics Committee of Tongji Hospital Affiliated to Tongji Medical College of Huazhong University of Science and Technology (Approval No. TJ-IRB20190604; Date: June 10th, 2019), and Medical Ethics Committee of Maternal and Child Hospital of Hubei Province (Approval No. [2022] IEC (007); Date: Feb. 10th, 2022); was recorded at Institutional Review Board of the second affiliated hospital of Zhengzhou University (Approval No. 2022336; Date: May 31th, 2022).

Informed consent statement: Patient consent was waived because the study was performed complying with the privacy policy, and the training and analysis were conducted anonymously.

Data availability statement: The data used in this study is unavailable due to the rules of ethical approvals.

Acknowledgments: In this section, you can acknowledge any support given which is not covered by the author contribution or funding sections. This may include administrative and technical support, or donations in kind (e.g., materials used for experiments). We also thank 8 endoscopists for diagnosing cases in two test datasets.

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

References

1. Chen X, Xiang YB, Long JR, Cai H, Cai Q, Cheng J, et al. Genetic polymorphisms in obesity-related genes and endometrial cancer risk. Cancer. 2012, 118, 3356-3364. DOI: 10.1002/cncr.26552

2. Clarke MA, Long BJ, Del Mar Morillo A, Arbyn M, Bakkum-Gamez JN, Wentzensen N. Association of Endometrial Cancer Risk With Postmenopausal Bleeding in Women: A Systematic Review and Meta-analysis. JAMA Intern Med. 2018, 178, 1210-1222. DOI: 10.1001/jamainternmed.2018.2820

3. Giannella L, Delli Carpini G, Sopracordevole F, Papiccio M, Serri M, Giorda G, et al. Atypical Endometrial Hyperplasia and Unexpected Cancers at Final Histology: A Study on Endometrial Sampling Methods and Risk Factors. Diagnostics. 2020, 10. DOI: 10.3390/diagnostics10070474

4. Gucer F, Reich O, Tamussino K, Bader AA, Pieber D, Scholl W, et al. Concomitant endometrial hyperplasia in patients with endometrial carcinoma. Gynecologic Oncology. 1998, 69, 64-68. DOI: 10.1006/gyno.1997.4911

5. Nagele F, O'Connor H, Baskett TF, Davies A, Mohammed H, Magos AL. Hysteroscopy in women with abnormal uterine bleeding on hormone replacement therapy: a comparison with postmenopausal bleeding. Fertility and sterility. 1996, 65, 1145-1150. DOI: 10.1016/s0015-0282(16)58329-0

6. Gkrozou F, Dimakopoulos G, Vrekoussis T, Lavasidis L, Koutlas A, Navrozoglou I, et al. Hysteroscopy in women with abnormal uterine bleeding: a meta-analysis on four major endometrial pathologies. Archives of gynecology and obstetrics. 2015, 291, 1347-1354. DOI: 10.1007/s00404-014-3585-x

7. Bourdel N, Chauvet P, Tognazza E, Pereira B, Botchorishvili R, Canis M. Sampling in Atypical Endometrial Hyperplasia: Which Method Results in the Lowest Underestimation of Endometrial Cancer? A Systematic Review and Meta-analysis. Journal of minimally invasive gynecology. 2016, 23, 692-701. DOI: 10.1016/j.jmig.2016.03.017

8. van Hanegem N, Prins MM, Bongers MY, Opmeer BC, Sahota DS, Mol BW, et al. The accuracy of endometrial sampling in women with postmenopausal bleeding: a systematic review and meta-analysis. European journal of obstetrics, gynecology, and reproductive biology. 2016, 197, 147-155. DOI: 10.1016/j.ejogrb.2015.12.008

9. Min JK, Kwak MS, Cha JM. Overview of Deep Learning in Gastrointestinal Endoscopy. Gut Liver. 2019, 13, 388-393. DOI: 10.5009/gnl18384

10. Zhang Y, Wang Z, Zhang J, Wang C, Wang Y, Chen H, et al. Deep learning model for classifying endometrial lesions. Journal of translational medicine. 2021, 19, 10. DOI: 10.1186/s12967-020-02660-x

11. Soffer S, Klang E, Shimon O, Nachmias N, Eliakim R, Ben-Horin S, et al. Deep learning for wireless capsule endoscopy: a systematic review and meta-analysis. Gastrointestinal endoscopy. 2020, 92, 831-839. DOI: 10.1016/j.gie.2020.04.039

12. He K, Zhang X, Ren S and Sun J. Deep Residual Learning for Image Recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, 770-778. DOI: 10.1109/CVPR.2016.90

13. Deng J, Dong W, Socher R, Li L, Li K, and Li F. ImageNet: A large-scale hierarchical image database. 2009 IEEE Conference on Computer Vision and Pattern Recognition 2009, 248-255. DOI: 10.1109/CVPR.2009.5206848

14. Lin T, Goyal P, Girshick R, He K, and Dollar P. Focal Loss for Dense Object Detection. 2017 IEEE® International Conference on Computer Vision (ICCV). Venice 2017, 2999–3007. DOI: 10.1109/ICCV.2017.324

15. Ando S and Huang CY. Deep Over-sampling Framework for Classifying Imbalanced Data. In: Ceci M. HJ, Todorovski L., Vens C., Džeroski S. (eds), editor. European Conference, ECML PKDD. Skopje, Macedonia: Springer; 2017.

16. Zhou D, Tian F, Tian X, Sun L, Huang X, Zhao F, et al. Diagnostic evaluation of a deep learning model for optical diagnosis of colorectal cancer. Nature communications. 2020, 11, 2961. DOI: 10.1038/s41467-020-16777-6

17. Selvaraju RR, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra. Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization. Proc. IEEE International Conference on Computer Vision (ICCV), 2017, 618–626. DOI: 10.1007/s11263-

019-01228-7.

18. Garuti G, Mirra M, Luerti M. Hysteroscopic view in atypical endometrial hyperplasias: a correlation with pathologic findings on hysterectomy specimens. J Minim Invasive Gynecol 2006, 13, 325–330. 10.1016/j.jmig.2006.03.010.

19. Magrina JF, Mutone NF, Weaver AL, Magtibay PM, Fowler RS, Cornella JL. Laparoscopic lymphadenectomy and vaginal or laparoscopic hysterectomy with bilateral salpingo-oophorectomy for endometrial cancer: morbidity and survival. American journal of obstetrics and gynecology. 1999, 181, 376-381. DOI: 10.1016/S0002-9378(99)70565-X

20. Yahagi N. Is artificial intelligence ready to replace expert endoscopists? Endoscopy. 2021, 53, 478-479. DOI: 10.1055/a-1308-2121

21. Gottlieb K, Requa J, Karnes W, Chandra Gudivada R, Shen J, Rael E, et al. Central Reading of Ulcerative Colitis Clinical Trial Videos Using Neural Networks. Gastroenterology. 2021, 160, 710-719. DOI: 10.1053/j.gastro.2020.10.024

22. Neofytou MS, Tanos V, Constantinou I, Kyriacou EC, Pattichis MS, Pattichis CS. Computer-aided diagnosis in hysteroscopic imaging. IEEE Journal of Biomedical and Health Informatics. 2015, 19, 1129-1136. DOI: 10.1109/JBHI.2014.2332760