Computer Methods and Programs in Biomedicine 221 (2022) 106906

Contents lists available at ScienceDirect

Computer Methods and Programs in Biomedicine

journal homepage: www.elsevier.com/locate/cmpb

Diagnosis of endometrium hyperplasia and screening of endometrial intraepithelial neoplasia in histopathological images using a global-to-local multi-scale convolutional neural network

Fengjun Zhao[a], Didi Dong[a], Hongyan Du[b,*], Yinan Guo[b], Xue Su[b], Zhiwei Wang[a], Xiaoyang Xie[a], Mingjuan Wang[b], Haiyan Zhang[b], Xin Cao[a], Xiaowei He[a,*]

[a] Xi'an Key Lab of Radiomics and Intelligent Perception, School of Information Science and Technology, Northwest University, Xi'an, Shaanxi 710069, China
[b] Department of Pathology, Northwest Women and Children's Hospital, Xi'an, Shaanxi 710061, China

ARTICLE INFO

Article history:
Received 23 February 2022
Revised 10 May 2022
Accepted 23 May 2022

Keywords:
Endometrial hyperplasia
Endometrial intraepithelial neoplasia
Computer-aided-diagnosis
Convolutional neural network
Histopathological image

ABSTRACT

Background and objective: Endometrial hyperplasia (EH), a uterine pathology characterized by an increased gland-to-stroma ratio compared to normal endometrium (NE), may precede the development of endometrial cancer (EC). Particularly, atypical EH also known as endometrial intraepithelial neoplasia (EIN), has been proven to be a precursor of EC. Thus, diagnosing different EH (EIN, hyperplasia without atypia (HwA) and NE) and screening EIN from non-EIN are crucial for the health of female reproductive system. Computer-aided-diagnosis (CAD) was used to diagnose endometrial histological images based on machine learning and deep learning. However, these studies perform single-scale image analysis and thus can only characterize partial endometrial features. Empirically, both global (cytological changes relative to background) and local features (gland-to-stromal ratio and lesion dimension) are helpful in identifying endometrial lesions.

Methods: We proposed a global-to-local multi-scale convolutional neural network (G2LNet) to diagnose different EH and to screen EIN in endometrial histological images stained by hematoxylin and eosin (H&E). The G2LNet first used a supervised model in the global part to extract contextual features of endometrial lesions, and simultaneously deployed multi-instance learning in the local part to obtain textural features from multiple image patches. The contextual and textural features were used together to diagnose different endometrial lesions after fusion by a convolutional block attention module. In addition, we visualized the salient regions on both the global image and local images to investigate the interpretability of the model in endometrial diagnosis.

Results: In the five-fold cross validation on 7812 H&E images from 467 endometrial specimens, G2LNet achieved an accuracy of 97.01% for EH diagnosis and an area-under-the-curve (AUC) of 0.9902 for EIN screening, significantly higher than state-of-the-arts. In external validation on 1631 H&E images from 135 specimens, G2LNet achieved an accuracy of 95.34% for EH diagnosis, which was comparable to that of a mid-level pathologist (95.71%). Specifically, G2LNet had advantages in diagnosing EIN, while humans performed better in identifying NE and HwA.

Conclusions: The developed G2LNet that integrated both the global (contextual) and local (textural) features may help pathologists diagnose endometrial lesions in clinical practices, especially to improve the accuracy and efficiency of screening for precancerous lesions.

1. Introduction

Endometrial cancer (EC), also called uterine cancer or corpus uteri cancer, is the most common female reproductive system cancer in developed countries and the second most common gynecologic malignancy worldwide [1]. EC often occurs in perimenopausal or menopausal women at an average age of 55 years, but the age of onset has gradually become younger and the incidence continues to increase due to the unprecedented rate of obesity [2–5]. Recent studies suggest that endometrial hyperplasia (EH), a uterine pathology characterized by an increased gland-to-stroma ratio compared to normal proliferative endometrium, may precede the development of EC and that the two share common predisposing risk factors [6,7]. The World Health Organization (WHO) further di-

* Corresponding authors.
E-mail addresses: hongyandudu@163.com (H. Du), hexw@nwu.edu.cn (X. He).

https://doi.org/10.1016/j.cmpb.2022.106906
0169-2607/

vided EH into hyperplasia without atypia (HwA) and atypical endometrial hyperplasia (AEH) [8–10]. HwA is a group of benign lesions that prefer progestin conservative treatment, whereas AEH, also known as endometrial intraepithelial neoplasia (EIN), has been proven to be a precursor (premalignancy) of endometrioid (type I) EC [10,11]. Previous findings showed that without treatment, 20 – 52% of EIN would develop into endometriosis, compared to 2% of HwA [12]. Since EIN may progress to or coexist with EC, the most appropriate treatment is hysterectomy [13,14]. Conspicuous differences in severity and treatment options prompt us to diagnose different EH and to screen EIN from HwA and normal endometrium (NE) as early as possible.

Histopathological examination after endometrial biopsy is currently the gold standard for the diagnosis of different endometrial lesions [15]. Specifically, EIN shows a minimum dimension of 1 mm, gland area exceeding stromal area, cytological changes relative to background, and exclusion of carcinomas and mimics, while EH that does not meet these criteria is identified as HwA [12,16]. Computer-assisted morphometry (D-score) was also developed to quantify the architectural (volume percentage of stroma and outer surface density of glands) and cytological (standard deviation of the shortest nuclear axis) features of endometrium, and has been used to the assisted diagnosis of hematoxylin and eosin (H&E) slides [17,18]. Nevertheless, confirmation of these diagnoses still relies on specialized visual assessment, which leads to an excessive demand for pathologists [19]. According to the China Diagnostic Pathology Industry Analysis Report, we needed 84,000-168,000 pathologists based on 1-2 pathologists per 100 beds, but there were only 18,000 pathologists on record in 2018, leaving a gap of at least 66,000 pathologists. As a result, pathologists need to perform histological analysis of H&E slides as quickly as possible, which sometimes results in conflicting or even erroneous diagnostic results due to prolonged overwork. That is, the scarcity of pathologists as a bottleneck has largely hindered efficient and accurate endometrial medical care.

Computer-aided diagnosis (CAD), as a promising tool to assist clinicians in decision making, provides automated alternatives for effective diagnosis of various cancers such as rectal cancer [20], ocular adnexal lymphoma [21], renal cancer [22], lung cancer [23] and breast cancer [24] in high-throughput medical images. In particular, several studies designed CAD systems to the differentiation of endometrial lesions in histological [25–29], hysteroscopic [30,31], ultrasound [32,33] and magnetic resonance images [34]. For example, researchers at the University of Athens employed machine learning (ML) models including logistic regression [29], classification and regression trees [25] and artificial neural networks [26] to differentiate benign from malignant endometrial liquid-based cytological smears, achieving an overall classification accuracy of 81.33–85% and 90.87–95% for endometrial nuclei and lesions, respectively. Downing et al. [27] segmented endometrial histological slides into epithelium, cells and nuclei and then extracted 1413 image features, from which 75 representative features were selected to distinguish different endometrial tissues using a random forest model, resulting in a classification accuracy of 94.2% between the benign, EIN and malignant groups.

Despite encouraging results, these ML-based CAD requires labor-intensive and time-consuming cytological preprocessing, such as manual segmentation of the endometrial nuclei, which takes an average of 10–20 min per specimen [26]. In addition, the performance of ML models depends on to how well the handcrafted features match the classification task, which often shows suboptimal performance if domain knowledge is not available. Deep learning (DL), one of the most important branches of ML, can automatically extract task-relevant features without much preprocessing and can continuously improve the performance through end-to-end learning [35]. Notably, single convolutional neural network (CNN) [36] and the ensemble of multiple CNNs [37] have been used to classify hysteroscopic images of endometrial lesions into benign, premalignant and malignant lesions with an accuracy of 80.8% and 90.29%, respectively. Sun et al. [24] developed a CNN model with an attention mechanism (called HIENet) for diagnosing different endometrial tissues from H&E images, obtaining 76.91% and 95.94% accuracy in four-class classification (NE, EH, endometrial polyp and EC) and binary classification (benign and malignant), respectively.

These studies have shown the potential of DL in the classification of different endometrial tissues. However, they mainly focus on the diagnosis of malignant and benign endometrial lesions without much attention to the identification of precancerous lesions such as EIN. In addition, current CAD performs image analysis at one fixed single-scale and thus can only characterize local or global features of endometrial histopathological images. It is known that both global (cytological changes relative to background) and local features (gland-to-stromal ratio and lesion dimension) are helpful in identifying different endometrial lesions. Therefore, we proposed a global-to-local multi-scale convolutional neural network (G2LNet) to diagnose different EH and to screen EIN in endometrial histological images stained by H&E. The G2LNet first used a supervised CNN model in the global part to extract contextual features of endometrial lesions, and simultaneously deployed multi-instance learning in the local part to obtain textural features from multiple image patches (Fig. 1). The contextual and textural features were used together to diagnose different endometrial lesions after fusion by the convolutional block attention module (CBAM) [38]. We anticipate our work may provide an efficient and accurate diagnostic tool for premalignant endometrial diseases.

2. Materials and dataset

2.1. Study population

This retrospective study was approved by the Medical Ethics Committee of Northwestern Women's and Children's Hospital and the informed consent was waved. From January 2016 to December 2020, 602 endometrial specimens were collected by diagnostic curettage, hysteroscopic surgery or hysterectomy from 452 female patients who did not receive radiotherapy, chemotherapy or hormone therapy before surgery and had no other gynecologic tumor complications. Subsequently, these patients were pathologically labeled (Section II.C) as NE (140 patients with 185 specimens), HwA (132 patients with 177 specimens) and EIN (180 patients with 240 specimens) with a mean age of 38 $\pm$ 7.2 (23–54), 47 $\pm$ 5.4 (26–68) and 54.7 $\pm$ 6.5 (27–71) years, respectively.

2.2. Specimen preparation

Following the standard protocol for histological reporting of EC published by the Royal College of Pathologists [39], the endometrial specimens were prepared by the following 5 steps: (1) the endometrial specimens were fixed in 10% neutral buffered formalin at room temperature for 12–48 h; (2) the fixed specimens were dehydrated by immersion in ethanol solutions in a series of increasing concentration; (3) the dehydrated samples were cleared using xylene; (4) the cleared samples were immersed in melted paraffin wax for embedding and serially sectioned at 4 μm thickness; (5) routine staining of sections with H&E was started after paraffin dissolution, and slides were made and sealed.

2.3. Pathological labeling

According to the WHO diagnostic criteria (4th edition, 2014) [8], two experienced pathologists (M.J.W and H.Y.Z, over 20 years of

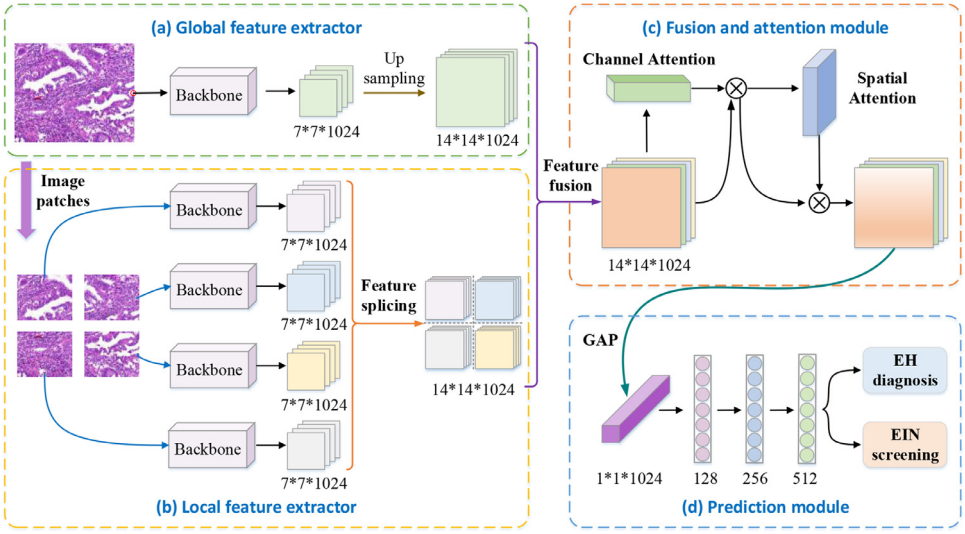

Fig. 1. Framework of the global-to-local multi-scale convolutional neural network (G2LNet). (a) Global feature extractor. (b) Local feature extractor. (c) Fusion and attention module. (d) Prediction module.

Table 1
Number of histological images in different datasets.

Class	# Images in training dataset	# Images in validation dataset	# Images in external validation	# Total images per class
NE (PE, SE)	2153 (1540, 614)	539 (385, 153)	525 (372, 153)	3217 (2297, 920)
HwA (SH, CH)	1244 (314, 931)	312 (78, 233)	465 (164, 301)	2021 (556, 1465)
EIN	2851	713	641	4205
Total:	6248	1564	1631	9443

PE: proliferative endometrium, SE: secretory endometrium, SH: simple hyperplasia, CH: complex hyperplasia.

endometrial pathology experience) labeled all H&E slides and divided them into three classes: i.e. NE ($n = 185$, including 120 proliferative endometria and 65 secretory endometria), HwA ($n = 177$, including 54 simple hyperplasia and 123 complex hyperplasia) and EIN ($n = 240$). Specifically, the pathologists first outlined all tissues on the H&E slide with a marker to prevent missing the field of view during microscopic observation and to eliminate areas with bubbles and dry pieces. Then, the slide was fixed on the carrier table and each field of view was observed with a low magnification lens (4X objective). If suspicious or difficult-to-identify areas were found, they switched to medium magnification lens (10X) and high magnification lens (40X) for closer observation to confirm the diagnosis. Notably, labeling of each slide was performed only with the agreement of the two pathologists, and controversial cases required further discussion and labeling after consensus was reached.

2.4. Digital imaging

We scanned the marked area of each H&E slide with a MoticEasyScan 60 scanner (I. Miller Microscopes, PA) at 400× magnification into high-resolution digital images and saved them in Joint Photographic Experts Group (JPEG) format, which facilitated the subsequent analysis using DL-based CAD. All marked areas of H&E slides were imaged sequentially, generating 10–20 non-overlapping digital images, with an average of approximately 16 per slide. The labels of these digital images were copied from the corresponding H&E slide, and the digital images that did not belong to this category were removed to ensure a match between images and labels. Considering the possible bias in tissue specimen preparation at different times and by different physicians, we did not preprocess the digital images before performing DL-based diagnosis. In this way, all endometrial H&E slides were converted into digital images of the same size of 1816 × 1519 pixels, obtaining 3217 images with NE (2297 proliferative endometria and 920 secretory endometria), 2021 images with HwA (556 simple hyperplasia and 1465 complex hyperplasia) and 4205 images with EIN (Table 1), respectively.

2.5. Dataset partitioning

This study used 467 specimens from 350 patients collected from January 2016 to December 2019 as the primary cohort, including 7812 digital images (2692 NE, 1556 HwA and 3564 EIN). Remaining 135 specimens from 102 patients collected from January 2020 to December 2020 were used as the secondary cohort. As shown in Table 1, we randomly selected 80 and 20% of the data in the primary cohort to train and validate the proposed model under five-fold cross-validation, obtaining 6248 (2153 NE, 1244 HwA and 2851 EIN) and 1564 images (539 NE, 312 HwA and 713 EIN) on each fold as the training and validation datasets, respectively. The secondary cohort was used as an external validation dataset consisting of 1631 images (525 NE, 465 HwA and 641 EIN) to further validate the robustness and generality of our model.

3. Methods

3.1. Global-to-local multi-scale convolutional neural network (G2LNet)

We propose a global-to-local multi-scale convolutional neural network (G2LNet) to perform the diagnosis of different endometrial lesions in H&E stained histological images. As shown in Fig. 1, the G2LNet consists of four modules, namely, a global feature extractor, a local feature extractor, a fusion and attention module and a prediction module. The backbone of the global and local feature extractors is DenseNet [40], a CNN-based network that mitigates the gradient vanishing issue in training by creating dense connections between all the preceding and following layers, thus enabling

the construction of deeper networks and improving classification performance.

Specifically, we partition each histological image of size 1816 × 1519 pixels into four image patches of the same size of 908 × 759 pixels. The large image is fed into the global feature extractor to learn the contextual features of endometrium tissues via supervised learning (Fig. 1(a)). Corresponding four image patches are fed into the local feature extractor to learn textural features via multi-instance learning (Fig. 1(b)), with each patch as one instance and four patches as one bag. Each bag is assigned the label of the large image, even though there may be several instances in this bag fall into other classes. For example, the large image labeled as EIN may contain regions of NE, leading to several of four patches belonging to NE rather than EIN, but the bag composed of these four patches is still labeled as EIN in multi-instance learning.

To fuse the contextual and textural features, we upsample the features of the global extractor and splice the four sets of features of the local extractor according to the relative position of the four patches, both of which generate the feature maps of the same size of 14 × 14 × 1024. Feature fusion is performed by element-wise max-pooling, which compares each pair of contextual (global) and textural (local) features in the feature space and selects the larger feature into the fused feature map. Then, a lightweight convolutional block attention module (CBAM) [38] is used to reweight the fused feature map from both the channel and spatial perspectives (Fig. 1(c), which has proven useful in enhancing important features and eliminating irrelevant ones elsewhere [41]. The channel attention uses max- and average-pooling to aggregate the spatial information and employs a multi-layer perceptron (MLP) to perform nonlinear feature mapping; element-wise summation and sigmoid activation are then performed to generate the channel attention that highlights the important channels of the fused feature map. The spatial attention module conducts channel-wise max- and average-pooling on the channel-weighted features and generate two feature maps; then we perform convolutional operations on them after dimensional concatenation, and use the sigmoid activation function to obtain spatial attention that pays more attention to the spatially important features for subsequent diagnosis. Of note, the CBAM does not change the feature dimension and the attention-weighted feature maps is still of size 14 × 14 × 1024.

The prediction module performs global average-pooling (GAP) on the weighted feature map and generates a 1 × 1 × 1024 feature vector (Fig. 1(d)), which is fed into three fully connected (FC) layers with 128, 256 and 512 neurons, respectively. The number of FC layers and the neurons per layer are determined empirically and experimentally. For the diagnosis of EH, the prediction module uses a softmax function that produces three outputs corresponding to the probabilities of NE, HwA and EIN, while for the screening of EIN, a sigmoid function is used as the activation function and the outputs are the binary probabilities of EIN and non-EIN (NE and HwA).

3.2. Manual diagnosis by pathologists

To compare the similarities and differences between the DL model (G2LNet) and humans in assessing endometrial morphology, three pathologists with different experience from the Department of Pathology of Northwest Women's and Children's Hospital participated in the manual pathological diagnosis, including a senior pathologist (Pathol_1, H.Y.D.), a mid-level pathologist (Pathol_2, X.S.) and a junior pathologist (Pathol_3, Y.N.G.) with fifteen, six and two years of pathological experience, respectively. Their diagnostic criteria were given as follows: (1) EIN was dominated by cytologic changes, including enlarged, rounded and irregular nuclei, distinct nucleoli, large amounts of eosinophilic cytoplasm, and lesions > 1 mm in size and usually consisting of at least 5–10 glands; (2) HwA exhibited excessive hyperplasia of glands with irregular size and morphology and a gland-to-stroma ratio exceeding that of normal proliferating endometrium without cytologic atypia; (3) Endometrium without the above-mentioned EIN and HwA characteristics was NE. All three pathologists were blinded to the pathological labels of digital images during the manually diagnostic process.

3.3. Performance evaluation

Under five-fold cross-validation, we evaluated the diagnostic performance of the G2LNet and state-of-the-arts for different EH in the validation dataset in terms of accuracy, precision, recall and F1-score, and visualized the results using confusion matrix, where each column represents the predicted class (NE, HwA or EIN) and each row represents the ground truth label, so that detailed misclassification can be easily analyzed. For the screening of EIN, the receiver operating characteristic (ROC) curve and the area under the curve (AUC), as well as the sensitivity and specificity with the threshold determined by the maximum Yuden index were used to validate the performance of different models in the binary classification of EIN from non-EIN.

We further compared our model with three pathologists for the EH diagnosis on the external validation dataset using accuracy, precision, recall, F1-score and confusion matrix. In addition, we calculated Kappa values between our model and each pathologist and between different pathologists to investigate the consistency of the prediction results between two subjects.

3.4. Image feature visualization

To provide more reliable and intuitive explanation to pathologists and help them understand the mechanism of the G2LNet, we used the gradient-weighted class activation mapping (Grad-CAM) proposed in [42] to visualize the regions noticed by the model during the diagnosis. The Grad-CAM can highlight the regions that associated with a given class by generating corresponding heap maps. In this study, we showed the salient regions of the G2LNet on both the global image and four local image patches, and compared these regions with those of human interest to see how the model performs in the case of correct and incorrect predictions.

4. Experiments and results

4.1. Experiment settings

We compared the proposed G2LNet to two state-of-the-art methods that were previously used for the diagnosis of endometrial or other lesions in histological images. The first one is the HIENet [24], a VGG-16-based CNN model with both position attention and channel attention, and the second one used a SVM classifier after extracting bottleneck features by a DenseNet-based CNN model (denoted as SVM_{CNN}) [43]. Moreover, to demonstrate the feasibility of the double pathways of the G2LNet (Fig. 1), we also constructed a GlobalNet and a LocalNet that extracted features by the global and local feature extractors, respectively, and then employed the same prediction module as the G2LNet to perform the diagnosis. In this way, the GlobalNet was a typical supervised learning method, while the LocalNet belonged to the weak-supervised multi-instance learning. The hyper- parameters of our G2LNet and other CNN models in comparative methods are given in Table 2.

Each H&E image was resized to 224 × 224 pixels to meet the size requirement of our G2LNet and other methods (HIENet, SVM_{CNN}, GlobalNet and LocalNet). The standard score, also known as z-score, was used to normalize the input image. The learning

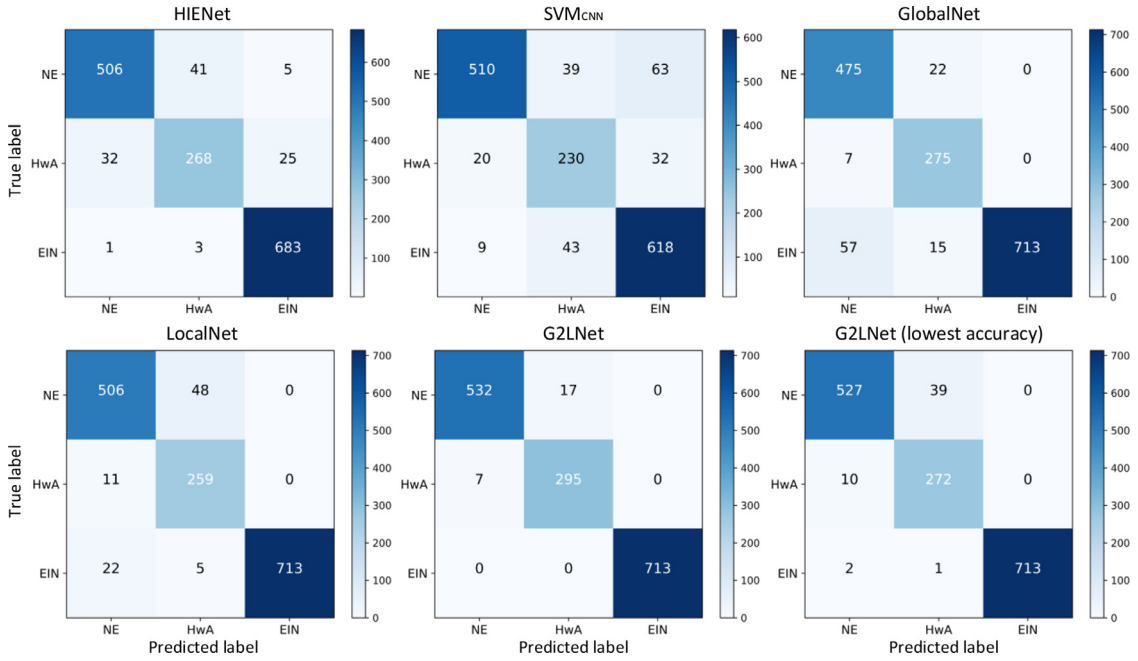

Fig. 2. Confusion matrices of diagnosis EH using different methods.

Table 2
Hyper-parameter settings for all models.

Hyper-parameter	CNN models in all methods
Learning rate	0.005
Loss function	Categorical/binary cross-entropy (patience =30)
Optimizer	Adam ($\beta_1 = 0.9$, $\beta_2 = 0.999$)
Batch size	8

rate was initially set to 0.005 and was annealed to 0.2 times when the performance stopped improving after 2 consecutive epochs. We pre-trained the backbone CNN models of all methods where appropriate using ImageNet dataset and fine-tuned their parameters (weights and biases) on our endometrial images until convergence. For the SVM_{CNN}, the extracted features were reduced to 50 using principal component analysis (PCA), and a SVM classifier with the radial basis function (RBF) kernel was used to perform the classification, where the parameters γ and C were set to 0.0078 and 2, respectively.

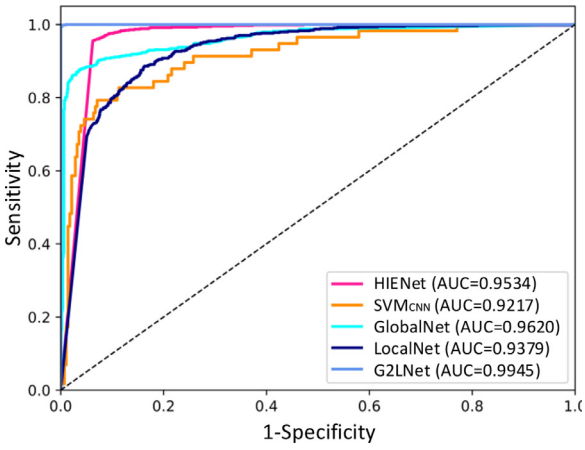

Fig. 3. ROC curves of EIN screening using different methods.

4.2. Results and comparison to state-of-the-arts

The results of diagnosing different EH using the five methods under five-fold cross-validation are given in Table 3. Our G2LNet achieved the best performance with a diagnostic accuracy of 97.01 ± 0.91%, significantly higher than the other four methods with the *p*-values being 0.0160, 0.0019, 0.0011 and 0.0067 comparing to HIENet, $SVMC_{NN}$, GlobalNet and LocalNet, respectively. The superiority of our G2LNet over the other four methods can also be demonstrated by its higher precision, recall and F1-score, which are 97.07 ± 0.87%, 97.01 ± 0.91% and 96.98 ± 0.93%, respectively. Fig. 2 shows the confusion matrix corresponding to the fold with median accuracy of each method. It can be seen that our G2LNet among the five methods achieved the best performance in the classification of all endometrial lesions including NE, HwA and EIN. In particular, when identifying EIN, the G2LNet did not misclassify any images of NE or HwA as EIN, nor did it misclassify any images of EIN as NE or HwA. Even on the fold with the lowest accuracy, the G2LNet was still better than or comparable to other four methods. Although the GlobalNet and LocalNet did not misclassify images of EIN as other types of EH, they did not performed well in identifying NE, which resulted in the overall performance being inferior to our G2LNet that combined the feature extractors of the GlobalNet and LocalNet. The SVM_{CNN} performed worst among the five methods, which may be due to its two-stage learning process rather than end-to-end learning like other methods, and thus extracted features may less relevant to subsequent EH diagnosis.

For the task of EIN screening, binary classification of EIN and non-EIN (including NE and HwA) was performed and the results under five-fold cross-validation are given in Table 4. Apparently, the G2LNet has the best performance with an AUC of 0.9902 ± 0.0057, significantly higher than the HIENet, SVMCNN, GlobalNet and LocalNet with the *p*-values of 0.0332, 0.0001, 0.0276 and 0.0013, respectively. The accuracy, sensitivity and specificity of our G2LNet with the threshold at the maximum Yuden index also surpass the four comparative methods, exception the LocalNet, which achieves the highest sensitivity of 99.71 ± 0.31%, although other metrics (such as accuracy and specificity) are not the case. The ROC curve corresponding to the median AUC (Fig. 3)

Table 3
Results of EH diagnosis using different methods (mean ± s.d.).

Method	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)	*P*-value (accuracy)
HIENet	93.46 ± 1.13	93.67 ± 0.94	93.46 ± 1.12	93.43 ± 1.09	0.0160
SVM_{CNN}	89.06 ± 3.38	89.20 ± 3.37	89.20 ± 3.37	89.20 ± 3.37	0.0019
GlobalNet	92.52 ± 1.59	93.22 ± 1.42	92.52 ± 1.58	92.41 ± 1.63	0.0011
LocalNet	92.81 ± 2.13	93.04 ± 1.90	92.81 ± 2.13	92.77 ± 2.03	0.0067
G2LNet	**97.01 ± 0.91**	**97.07 ± 0.87**	**97.01 ± 0.91**	**96.98 ± 0.93**	–

Table 4
Results of EIN screening using different methods (mean ± s.d.).

Methods	AUC	Accuracy (%)	Sensitivity (%)	Specificity (%)	*P*-value (AUC)
HIENet	0.9682 ± 0.0162	96.41 ± 1.09	94.61 ± 1.37	98.37 ± 1.80	0.0332
SVM_{CNN}	0.9120 ± 0.0110	94.87 ± 0.62	97.52 ± 0.32	84.88 ± 1.96	0.0001
GlobalNet	0.9687 ± 0.0149	94.77 ± 2.24	94.60 ± 4.16	95.44 ± 1.71	0.0276
LocalNet	0.9307 ± 0.0240	95.94 ± 1.51	91.38 ± 3.20	**99.71 ± 0.31**	0.0013
G2LNet	**0.9902 ± 0.0057**	**99.07 ± 0.55**	**98.54 ± 0.92**	99.51 ± 0.61	–

Table 5
Comparison between G2LNet and pathologists in external validation dataset (mean and 95% confidence interval).

	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Pathol_1	85.04 (83.31, 86.77)	85.70 (84.00, 87.40)	85.04 (83.31, 86.77)	85.01 (83.28, 86.74)
Pathol_2	95.71 (94.73, 96.69)	9577 (94.79, 96.75)	95.48 (94.47, 96.49)	95.60 (94.60, 96.60)
Pathol_3	**98.65** (98.09, 99.21)	**98.66** (98.10, 99.22)	**98.52** (97.93, 99.11)	**98.59** (98.02, 99.16)
G2LNet	95.34 (94.32, 96.36)	95.67 (94.32, 96.36)	95.34 (94.32, 96.36)	95.31 (94.28, 96.34)

shows that the classification performance ranks from the best to worst is the G2LNet, GlobalNet, HIENet, LocalNet and SVM_{CNN}, which is slightly different from the ranks in the task of EH diagnosis.

4.3. Comparison between G2LNet and human diagnosis

The comparison between the DL model and human diagnosis was performed on the external validation dataset and the results are given in Table 5. The G2LNet achieved an accuracy of 95.34% (95% CI: 94.32-96.36%), which exceeded that of the junior pathologist (Pathol_1, 85.04%, 95% CI: 83.31-86.77%) and was comparable to that of the mid-level pathologist (Pathol_2, 95.71%, 95% CI: 94.73-96.69%), but not as good as that of the senior pathologist (Pathol_3, 98.71%, 95% CI: 98.09-99.21%). Furthermore, we note from the confusion matrix corresponding to the median accuracy (Fig. 4) that the G2LNet performs best in identifying EIN, with only one image of NE being misclassified as EIN and no images of EIN being misclassified as other types of EH. This demonstrates the superiority of the DL model over human experts in diagnosing premalignant lesions. In contrast, human experts performed better than the G2LNet in identifying images with NE and HwA. For example, the mid-level pathologist (Pathol_2) misclassifies only 14 NE images as HwA, compared with 66 misclassifications for the G2LNet.

The kappa values between the DL model and human experts as well those between different human experts are given in Table 6. It is observed that the diagnosis among different pathologists achieved good repeatability with the kappa values ranging from 0.7775 (95% CI: 0.7573-0.7977) to 0.9732 (95% CI: 0.9654-0.9810). The agreement between the G2LNet and pathologists increases with the practicing experience of these pathologists, with the highest Kappa value of 0.9312 (95% CI: 0.9189-0.9435) achieved between the G2LNet and the senior pathologist (Pathol_3). This demonstrates that the consistencies between the DL model and

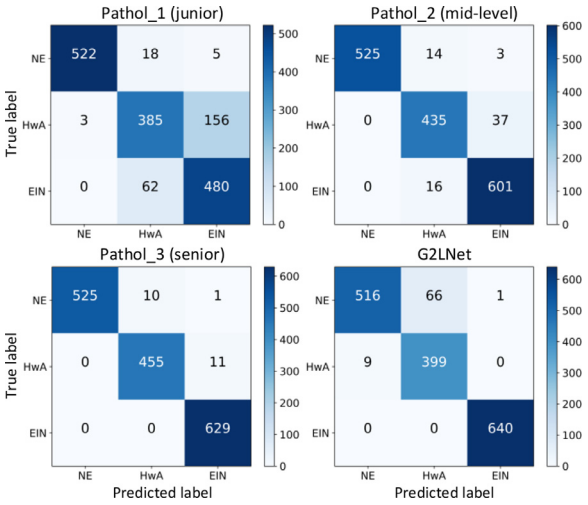

Fig. 4. Confusion matrices of the G2LNet and three pathologists for EH diagnosis in external validation dataset.

Table 6
Kappa values between G2LNet and pathologists in external validation dataset (mean and 95% confidence interval).

	G2LNet	Pathol_1	Pathol_2
Pathol_1	0.7168 (0.6949, 0.7387)	–	
Pathol_2	0.8939 (0.8790, 0.9088)	0.7968 (0.7773, 0.8163)	–
Pathol_3	0.9312 (0.9189, 0.9435)	0.7775 (0.7573, 0.7977)	0.9732 (0.9654, 0.9810)

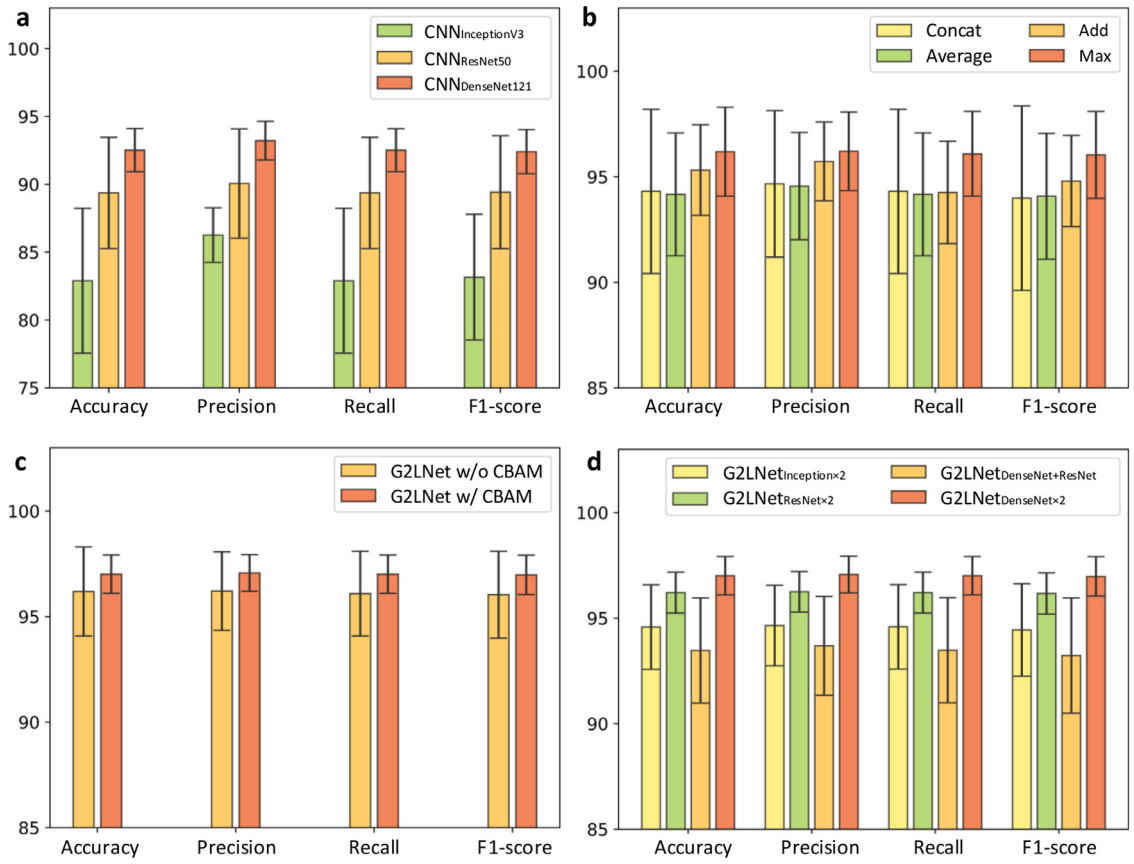

Fig. 5. Results of ablation study. (a) Results of three single-pathway CNN classifiers. (b) Results of different feature fusion strategies. (c) Results of G2LNet with and without CBAM. (d) Results of G2LNet with different combinations of backbone CNNs.

human experts are basically on par with those between different human experts.

4.4. Ablation study

To verify the effectiveness of each component of our G2LNet, we performed the following ablation studies: (a) three CNN classifiers including InceptionV3 [44], ResNet50 [45] and DenseNet121 [40] were compared to select the optimal backbone CNN used for the construction of the G2LNet; (b) four fusion strategies including concatenation, adding, element-wise average- and max-pooling were used to fuse the contextual and textural features extracted by the global and local feature extractors, respectively; (c) the G2LNet with and without CBAM were compared to investigate the effect of the attention mechanism; (d) different combinations of backbone CNNs in double pathways were compared to validate the usefulness of the global-to-local structure of the G2LNet, including $\text{G2LNet}_{\text{Inception}\times 2}$, $\text{G2LNet}_{\text{ResNet}\times 2}$, $\text{G2LNet}_{\text{DenseNet+ResNet}}$ and $\text{G2LNet}_{\text{DenseNet}\times 2}$, respectively.

As shown in Fig. 5(a), among three single-pathway CNNs that are used to diagnose between different EH (NE, HwA and EIN), the DenseNet121 achieved an accuracy of 92.52 $\pm$ 1.59%, outperforming the InceptionV3 ($p = 0.0086$) and ResNet50 ($p = 0.1901$), which is why we used the DenseNet121 as the backbone of our G2LNet. As for the fusion of contextual and textural features (Fig. 5(b)), the element-wise max-pooling obtained a higher accuracy of 96.19 $\pm$ 2.11% compared to other fusion strategies, although the difference was not significant ($p > 0.05$). The addition of CBAM in the fusion and attention module (Fig. 1(c)) further elevated the diagnostic accuracy to 97.01 $\pm$ 0.91% compared against the G2LNet without the attention mechanism (Fig. 5(c)).

Fig. 5(d) shows that the $\text{G2LNet}_{\text{DenseNet}\times 2}$ using DenseNet121 as the backbone in both the global and local feature extractors obtains an accuracy of 97.07 $\pm$ 0.91%, better than the $\text{G2LNet}_{\text{Inception}\times 2}$, $\text{G2LNet}_{\text{ResNet}\times 2}$ and $\text{G2LNet}_{\text{DenseNet+ResNet}}$ with the p-values of 0.0587, 0.2655, and 0.0286, respectively. Beyond our expectations, the lowest accuracy was obtained for $\text{G2LNet}_{\text{DenseNet+ResNet}}$ employing the DenseNet121 and ResNet50 as the backbone in the global and local feature extractors, respectively, which may be caused by the semantic mismatch of these two CNNs and the sub-optimal hyper-parameters. Nevertheless, each G2LNet with the global-to-local architecture with double pathways surpasses the single-pathway CNNs. For example, the diagnostic performance of the $\text{G2LNet}_{\text{DenseNet+ResNet}}$ is still higher than that of the DenseNet121 and ResNet50, with an accuracy improvement of 0.95% and 4.10%, respectively.

5. Discussion and conclusion

To relieve the burden of pathologists, we developed a DL-based endometrial diagnostic model (G2LNet) that fully exploits the contextual and textual features of histological images by the unique global-to-local double-pathway architecture. The G2LNet performs better on 7812 H&E images for both the EH diagnosis and EIN screening, compared to state-of-the-arts (HIENet and SVMCNN) and single-pathway networks (GlobalNet and LocalNet). In external validation of 1631 H&E images, the G2LNet is superior or comparable to the junior and mid-level pathologists with two and six years of pathological experience, respectively. In particular, the G2LNet outperforms comparative methods and pathologists in identifying EIN lesions in both the cross-validation and external validation, indicating that the proposed DL-based CAD has great potential for the automatic diagnosis of precancerous endometrial lesions.

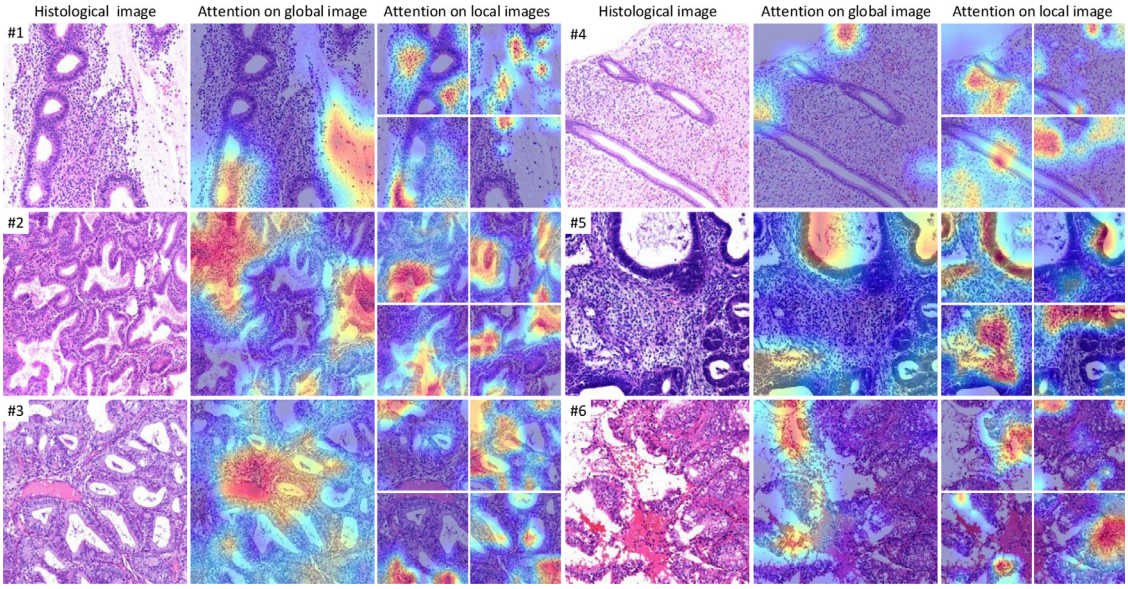

Fig. 6. Three histological images with correct predictions and three images with misclassifications from the external validation dataset. Cases #1–3 are correctly predicted as NE, HwA and EIN, respectively. Cases #4–6 are misclassified as HwA, EIN, and NE, respectively, while the ground truth labels for these cases are EN, HwA, and EIN, respectively. In the heat map of the Grad-CAM, the red, orange and blue colors indicate the model (or G2LNet) pays high, moderate and low attention to corresponding regions during endometrial diagnosis, respectively.

Table 7
Comparison between G2LNet's attention and human assessment on correctly and incorrectly predicted cases.

	Ground truth label	Predicted result	Characteristics of visual assessment by human	Attention of G2LNet on global and local images	Analysis of the attentional regions
#1	NE	NE	Less than 50% gland-to-stroma ratio, no atypia in glandular epithelial cells (long oval, sunken, regularly arranged and well polarized nuclei)	Global: continuous "bead-like" structure in glandular cross-section and part of stromal regions; Local: multiple stromal regions and normal glandular epithelial cells	Capturing critical features of NE in both the global and local images that are noticed by humans
#2	HwA	HwA	Slightly dilated, irregular & densely arranged glands, over 50% gland-to-stroma ratio but without atypia in glandular cells	Global: part of the stromal region, "back-to-back" densely arranged glands; Local: irregular dilatation of glandular lumens and details on multiple typical glandular structures	Identified more details on glandular architecture in four local images as a complement to the global image
#3	EIN	EIN	Closely arranged glands, disappearing stromata for most glands and atypical glandular epithelial cells (disorderly arranged and unpolarized cells, and enlarged, rounded and faded nuclei)	Global: central glands in the global image and disappeared stroma and a blood vessel region; Local: multiple glandular and stromal regions covering both the structural and cellular atypia	Small disturbance by the blood vessel in the global image was eliminated by the atypical features in the local images
#4	NE	HwA	Small and consistent glandular lumens, two uniformly distributed straight glands, far less than 50% gland-to-stroma ratio and no atypical glandular epithelial cells	Global: few stromal region in the upper left of the global image; Local: more stromal regions and a few incomplete glandular regions	Failure cause by secretory response in stroma and different view of glandular lumens from typical cross-sectional view
#5	HwA	EIN	Marked cystic dilatation in some glandular lumens, no atypical appearance in glandular epithelial cells	Global: part of glands with cystic dilation; Local: multiple glandular and stromal regions including the glands with cystic expansion	Failure caused by unclear cellular hierarchy and disorderly arrangement due to the thick specimen
#6	EIN	NE	Disappearing stromata, broken, incomplete but densely arranged glands, atypical appearance of some glandular epithelial cells and marked hemorrhagic background	Global: disappearing stromal regions and fragmented glandular regions of the left part of the global image; Local: more detailed fragmented glandular regions over all four local images	Failure to identify key glandular features due to the interference from fragmented glands on both the global and local images

To investigate the interpretability of the G2LNet model in EH diagnosis, we visualized the salient regions on both the global image and four local images (patches) by performing Grad-CAM to the model. Specifically, we randomly selected three histological images with correct predictions and three images with misclassifications from the external validation dataset as examples (Fig. 6). Detailed compassion between the G2LNet's attention and human expert visual assessment is given in Table 7. Overall, Grad-CAM focuses on both the stromal and glandular areas in the diagnosis of NE, HwA and EIN (e.g., cases #1–5). In addition to the increased gland-to-stroma ratio in case #2, the model captures atypical glandular epithelial cells in case #3, which is crucial for identifying EIN. For cases #1 and #4, Grad-CAM focus more on the stromal area, which may be caused by these NE cases having large stromal areas or small gland-to-stromal ratios.

For correctly predicted cases (#1–3), the G2LNet either captured key features of different EH in both the global and local images that are consistent with human visual assessment (case #1), or

identified more important structural details in local images as a complement to the global image (cases #2 and #3). The G2LNet can even eliminate small vessel disturbance in the global image by focusing more on atypical features in the local images (case #3). Similar to the ablation study in Fig. 5(d), these corrected predicted cases demonstrate that image features extracted from the global and local images benefit from each in our G2LNet, thus improving the accuracy of diagnosing endometrial lesions.

For those misclassified cases (#4–6), the G2LNet noticed only a part of the salient regions in the global image but captured most critical regions in the local images. However, the model is still unable to accurately identify the type of these cases due to their abnormal appearance compared to the common histological images. Specifically, case #4 has a secretory response in stroma and the glandular lumen in the longitudinal view differs greatly from that in the typical cross-sectional view. Case #5 exhibits a poorly layered and disordered cellular arrangement caused by an overly thick specimen. The glands in case #6 are fragmented, which disrupts the identification of atypical features of glandular epithelial cells. Obviously, failures on these cases are more or less related to inadequate preparation of histological specimens. However, this is unavoidable due to the excessive demand for histological examination in routine clinical diagnosis. Therefore, we will collect more abnormal histological images in the future and improve the G2LNet's awareness for these images in order to correctly diagnose different endometrial tissues in cases of inadequate specimen preparation.

Clinically, there are two commonly used histological classifications, WHO and EIN systems, where the former assesses glandular/stromal architecture and the cytological atypia, while the latter performs the combined morphometric analysis of nuclear and architectural features [6]. Based on the EIN system, an objective likelihood metric called D-score was developed by integrating volume percentage of stroma, outer surface density of the glands, and the standard deviation of the shortest glandular nuclear axis [17,18]. Although both focus on the gland-stroma ratio in endometrial tissues, studies suggest that the two systems are complementary to some extent, and the EIN system with the D-score may be more reliable than the WHO system in predicting the risk of EH progression to EC [46,47]. Therefore, the assessment experience in the WHO and EIN systems or direct combination of the D-score into our DL-based CAD system might further improve the performance of the model as well as the interpretability of the diagnostic results.

Furthermore, molecular classification based on The Cancer Genome ATLAS (TCGA) has been used to stratify EC as four subtypes with distinct clinical course and progression-free survival [48,49]. In particular, high expression of tumor protein p53 [50,51] and β-catenin [52,53], and low expression of phosphatase and tensin homologue (PTEN) [54,55], AT-rich interactive domain-containing protein 1 A (ARID1A) [56,57], DNA mismatch repair (MMR) (including MLH1, MSH2 and MSH6) [58,59] and B-cell lymphoma 2 (Bcl-2) [6,60] may predict the progression from NE/HwA, though EIN, to EC. Despite sometimes encountering inconsistent or even conflicting results, future integration of these molecular biomarkers into the DL-based histological analysis may benefit the overall diagnostic performance.

There are also limitations in our study. First, we focus on the classification between NE, HwA and EIN in EH diagnosis, or the distinction between EIN and non-EIN in EIN screening, thus neglecting the differentiation between benign lesions and EC. How to perform the classification between NE, HwA, EIN and EC or the distinction between EIN and EC needs further investigation in the future. Second, we performed the diagnosis on digital images rather than on whole slide images (WSIs) because (1) localized regions can better capture the morphological features of EIN such as the atypical glandular cells, and (2) the WSI consisting of ~16 digital images (each 1816 × 1519 pixels in size) would be too large, making the subsequent diagnosis challenging due to expensive storage and time consumption. In the following study, we may extend the model to WSIs by first performing the diagnosis on the digital images using the G2LNet, and then conducting the decision-level fusion on multiple digital images to obtain the diagnostic result on corresponding WSI.

In conclusion, we have proposed a DL model called G2LNet to diagnose different types of endometrial tissues in H&E stained histological images. The model extracted contextual and texture features from global and local feature extractors, respectively, which were then fused together by a CBAM model and used for predicting different EH and screening EIN. For EH diagnosis, the accuracy of the G2LNet was significantly better than the HIENet, SVM_{CNN}, GlobalNet and LocalNet ($p < 0.05$). In particular, the G2LNet did not misclassify NE or HwA as EIN, nor did it misclassify any EIN as NE or HwA. For EIN screening, the AUC of the G2LNet also outperformed the four comparative methods ($p < 0.05$). In external validation, the G2LNet exhibited complementarity with pathologists. The G2LNet performed better in identifying EIN, while human experts were skilled in identifying NE and HwA. The developed G2LNet may help pathologists diagnose endometrial lesions in clinical practices, especially to improve the accuracy and efficiency of screening for precancerous lesions.

Declaration of Competing Interest

The authors declare that no conflict of interest.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China (61971350), Shaanxi International Science and Technology Cooperation Program (2021KW-55), Shaanxi Key R&D Plan (2020SF-036), China Postdoctoral Science Foundation (2019M653717), and Xi'an Science and Technology Plan (201805060ZD11CG44).

References

[1] J. Ferlay, I. Soerjomataram, R. Dikshit, S. Eser, C. Mathers, M. Rebelo, D.M. Parkin, D. Forman, F. Bray, Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012, Int. J. Cancer 136 (2015) E359–E386.
[2] T. Evans, O. Sany, P. Pearmain, R. Ganesan, A. Blann, S. Sundar, Differential trends in the rising incidence of endometrial cancer by type: data from a UK population-based registry from 1994 to 2006, Br. J. Cancer 104 (2011) 1505–1510.
[3] C. Gonthier, A. Trefoux-Bourdet, M. Koskas, Impact of conservative managements in young women with grade 2 or 3 endometrial adenocarcinoma confined to the endometrium, Int. J. Gynecol. Cancer 27 (2017) 493–499.
[4] A.G. Renehan, I. Soerjomataram, M. Tyson, M. Egger, M. Zwahlen, J.W. Coebergh, I. Buchan, Incident cancer burden attributable to excess body mass index in 30 European countries, Int. J. Cancer 126 (2010) 692–702.
[5] M.R. Wise, V. Jordan, A. Lagas, M. Showell, N. Wong, S. Lensen, C.M. Farquhar, Obesity and endometrial hyperplasia and cancer in premenopausal women: a systematic review, Am. J. Obstet. Gynecol. 214 (2016) 689–697.
[6] P.A. Sanderson, H.O.D. Critchley, A.R.W. Williams, M.J. Arends, P.T.K. Saunders, New concepts for an old problem: the diagnosis of endometrial hyperplasia, Hum. Reprod. Update 23 (2017) 232–254.
[7] S.F. Lax, Precursor lesions of endometrial carcinoma, Pathologe 40 (2019) 13–20.
[8] R. Zaino, S. Carinelli, L. Ellenson, R. Kurman, M. Carcangiu, C. Herrington, R. Young, Uterine corpus: epithelial tumors and precursors, in: World Health Organization Classification of Tumors of Female Reproductive Organs, 4th ed., International Agency for Research on Cancer (IARC) Press, Lyon, France, 2014, pp. 125–135.
[9] G. Emons, M.W. Beckmann, D. Schmidt, P. Mallmann, G. Gynecological Oncology Working, New WHO classification of endometrial hyperplasias, Geburtshilfe und Frauenheilkd. 75 (2015) 135–136.
[10] R.A. Owings, C.M. Quick, Endometrial intraepithelial neoplasia, Arch. Pathol. Lab. Med. 138 (2014) 484–491.
[11] E.A. Jarboe, G.L. Mutter, Endometrial intraepithelial neoplasia, Semin. Diagn. Pathol. 27 (2010) 215–225.

[12] F. Edris, G.A. Vilos, A. Al-Mubarak, H.C. Ettler, J. Hollett-Caines, B. Abu-Rafea, Resectoscopic surgery may be an alternative to hysterectomy in high-risk women with atypical endometrial hyperplasia, J. Minim. Invasive Gynecol. 14 (2007) 68–73.
[13] C.L. Trimble, M. Method, M. Leitao, K. Lu, O. Ioffe, M. Hampton, R. Higgins, R. Zaino, G.L. Mutter, Management of endometrial precancers, Obstet. Gynecol. 120 (2012) 1160–1175.
[14] V. Chandra, J.J. Kim, D.M. Benbrook, A. Dwivedi, R. Rai, Therapeutic options for management of endometrial hyperplasia, J. Gynecol. Oncol. 27 (2016) e8.
[15] A. Revel, A. Shushan, Investigation of the infertile couple - Hysteroscopy with endometrial biopsy is the gold standard investigation for abnormal uterine bleeding, Hum. Reprod. 17 (2002) 1947–1949.
[16] J.L. Hecht, T.A. Ince, J.P.A. Baak, H.E. Baker, M.W. Ogden, G.L. Mutter, Prediction of endometrial carcinoma by subjective endometrial intraepithelial neoplasia diagnosis, Mod. Pathol. 18 (2005) 324–330.
[17] J.P.A. Baak, G.L. Mutter, EIN and WHO94, J. Clin. Pathol. 58 (2005) 1–6.
[18] G.L. Mutter, G. Endometrial Collaborative, Endometrial intraepithelial neoplasia (EIN): will it bring order to chaos? Gynecol. Oncol. 76 (2000) 287–290.
[19] S. Jha, E.J. Topol, Adapting to artificial intelligence radiologists and pathologists as information specialists, JAMA J. Am. Med. Assoc. 316 (2016) 2353–2354.
[20] Z. Liu, X.Y. Zhang, Y.J. Shi, L. Wang, H.T. Zhu, Z. Tang, S. Wang, X.T. Li, J. Tian, Y.S. Sun, Radiomics analysis for evaluation of pathological complete response to neoadjuvant chemoradiotherapy in locally advanced rectal cancer, Clin. Cancer Res. 23 (2017) 7253–7262.
[21] Y. Hou, X. Xie, J. Chen, P. Lv, S. Jiang, X. He, L. Yang, F. Zhao, Bag-of-features-based radiomics for differentiation of ocular adnexal lymphoma and idiopathic orbital inflammation from contrast-enhanced MRI, Eur. Radiol. 31 (2021) 24–33.
[22] L. Zhou, Z. Zhang, Y.C. Chen, Z.Y. Zhao, X.D. Yin, H.B. Jiang, A deep learning-based radiomics model for differentiating benign and malignant renal tumors, Transl. Oncol. 12 (2019) 292–300.
[23] A. Hosny, C. Parmar, T.P. Coroller, P. Grossmann, R. Zeleznik, A. Kumar, J. Bussink, R.J. Gillies, R.H. Mak, H.J.W.L. Aerts, Deep learning for lung cancer prognostication: a retrospective multi-cohort radiomics study, PLoS Med. 15 (2018) e1002711.
[24] H. Sun, X. Zeng, T. Xu, G. Peng, Y. Ma, Computer-aided diagnosis in histopathological images of the endometrium using a convolutional neural network and attention mechanisms, IEEE J. Biomed. Health Inform. 24 (2020) 1664–1676.
[25] A. Pouliakis, C. Margari, N. Margari, C. Chrelias, D. Zygouris, C. Meristoudis, I. Panayiotides, P. Karakitsos, Using classification and regression trees, liquid-based cytology and nuclear morphometry for the discrimination of endometrial lesions, Diagn. Cytopathol. 42 (2014) 582–591.
[26] G.M. Makris, A. Pouliakis, C. Siristatidis, N. Margari, E. Terzakis, N. Koureas, V. Pergialiotis, N. Papantoniou, P. Karakitsos, Image analysis and multi-layer perceptron artificial neural networks for the discrimination between benign and malignant endometrial lesions, Diagn. Cytopathol. 45 (2017) 202–211.
[27] M.J. Downing, D.J. Papke, S. Tyekucheva, G.L. Mutter, A new classification of benign, premalignant, and malignant endometrial tissues using machine learning applied to 1413 candidate variables, Int. J. Gynecol. Pathol. 39 (2020) 333–343.
[28] A.A. Vlachokosta, P.A. Asvestas, G.K. Matsopoulos, A. Kondi-Pafiti, N. Vlachos, Classification of histological images of the endometrium using texture features, Anal. Quant. Cytol. Histol. 35 (2013) 105–113.
[29] D. Zygouris, A. Pouliakis, N. Margari, C. Chrelias, E. Terzakis, N. Koureas, I. Panayiotides, P. Karakitsos, Classification of endometrial lesions by nuclear morphometry features extracted from liquid-based cytology samples a system based on logistic regression model, Anal. Quant. Cytopathol. Histpathol. 36 (2014) 189–198.
[30] A.A. Vlachokosta, P.A. Asvestas, F. Gkrozou, L. Lavasidis, G.K. Matsopoulos, M. Paschopoulos, Classification of hysteroscopical images using texture and vessel descriptors, Med. Biol. Eng. Comput. 51 (2013) 859–867.
[31] M.S. Neofytou, V. Tanos, I. Constantinou, E.C. Kyriacou, M.S. Pattichis, C.S. Pattichis, Computer-aided diagnosis in hysteroscopic imaging, IEEE J Biomed. Health Inform. 19 (2015) 1129–1136.
[32] J.Y. Wu, A. Tuomi, M.D. Beland, J. Konrad, D. Glidden, D. Grand, D. Merck, Quantitative analysis of ultrasound images for computer-aided diagnosis, J. Med. Imaging 3 (2016) 014501.
[33] J. Konrad, D. Merck, J.Y. Wu, A. Tuomi, M. Beland, Improving ultrasound detection of uterine adenomyosis through computational texture analysis, Ultrasound Q. 34 (2018) 29–31.
[34] Y. Ueno, B. Forghani, R. Forghani, A. Dohan, X.Z. Zeng, F. Chamming's, J. Arseneau, L. Fu, L. Gilbert, B. Gallix, C. Reinhold, Endometrial carcinoma: MR imaging-based texture model for preoperative risk stratification-a preliminary analysis, Radiology 284 (2017) 748–757.
[35] P. Afshar, A. Mohammadi, K.N. Plataniotis, A. Oikonomou, H. Benali, From handcrafted to deep-learning-based cancer radiomics: challenges and opportunities, IEEE Signal Process. Mag. 36 (2019) 132–160.
[36] Y. Zhang, Z. Wang, J. Zhang, C. Wang, Y. Wang, H. Chen, L. Shan, J. Huo, J. Gu, X. Ma, Deep learning model for classifying endometrial lesions, J. Transl. Med. 19 (2021) 10.
[37] Y. Takahashi, K. Sone, K. Noda, K. Yoshida, Y. Toyohara, K. Kato, F. Inoue, A. Kukita, A. Taguchi, H. Nishida, Y. Miyamoto, M. Tanikawa, T. Tsuruga, T. Iriyama, K. Nagasaka, Y. Matsumoto, Y. Hirota, O. Hiraike-Wada, K. Oda, M. Maruyama, Y. Osuga, T. Fujii, Automated system for diagnosing endometrial cancer by adopting deep-learning technology in hysteroscopy, PLoS One 16 (2021) e0248526.
[38] S. Woo, J. Park, J.Y. Lee, I.S. Kweon, CBAM: convolutional block attention module, Comput. Vis. 11211 (2018) 3–19 Eccv 2018Pt Vii.
[39] R. Ganesan, N. Singh, W.G. McCluggage, Standards and Datasets for Reporting Cancers Dataset for Histological Reporting of Endometrial Cancer, The Royal Colledge of Pathologists, 2014.
[40] G. Huang, Z. Liu, L. van der Maaten, K.Q. WeinbergerIEEE, Densely connected convolutional networks, in: Proceedings of the 30th IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, 2017, pp. 2261–2269.
[41] A. Khan, A. Sohail, U. Zahoora, A.S. Qureshi, A survey of the recent architectures of deep convolutional neural networks, Artif. Intell. Rev. 53 (2020) 5455–5516.
[42] R.R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, D. Batra, Grad–CAM: visual explanations from deep networks via gradient-based localization, Int. J. Comput. Vis. 128 (2020) 336–359.
[43] T. Araujo, G. Aresta, E. Castro, J. Rouco, P. Aguiar, C. Eloy, A. Polonia, A. Campilho, Classification of breast cancer histology images using convolutional neural networks, PLoS One 12 (2017) e0177544.
[44] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, Z. Wojna, Rethinking the inception architecture for computer vision, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 2818–2826.
[45] K. He, X. Zhang, S. Ren, J. SunIEEE, Deep residual learning for image recognition, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, 2016, pp. 770–778.
[46] A. Travaglino, A. Raffone, G. Saccone, A. Mollo, G. De Placido, L. Insabato, F. Zullo, Endometrial hyperplasia and the risk of coexistent cancer: WHO versus EIN criteria, Histopathology 74 (2019) 676–687.
[47] A. Raffone, A. Travaglino, G. Saccone, L. Insabato, A. Mollo, G. De Placido, F. Zullo, Endometrial hyperplasia and progression to cancer: which classification system stratifies the risk better? A systematic review and meta-analysis, Arch. Gynecol. Obstet. 299 (2019) 1233–1242.
[48] A. Raffone, A. Travaglino, D. Raimondo, D. Neola, F. Renzulli, A. Santoro, L. Insabato, P. Casadio, G.F. Zannoni, F. Zullo, A. Mollo, R. Seracchioli, Prognostic value of myometrial invasion and TCGA groups of endometrial carcinoma, Gynecol. Oncol. 162 (2021) 401–406.
[49] G. Getz, S.B. Gabriel, K. Cibulskis, E. Lander, A. Sivachenko, C. Sougnez, et al., Integrated genomic characterization of endometrial carcinoma, Nature 497 (2013) 67–73.
[50] N. Horree, P.J. van Diest, D.M.D.S. Sie-Go, A.P.M. Heintz, The invasive front in endometrial carcinoma: higher proliferation and associated derailment of cell cycle regulators, Hum. Pathol. 38 (2007) 1232–1238.
[51] L. Cinel, A. Polat, O. Aydin, D. Dusmez, R. Egilmez, Bcl-2, iNOS, p53 and PCNA expression in normal, disordered proliferative, hyperplastic and malignant endometrium, Pathol. Int. 52 (2002) 384–389.
[52] Y. Norimatsu, T. Moriya, T.K. Kobayashi, T. Sakurai, K. Shimizu, C. Tsukayama, E. Ohno, Immunohistochemical expression of PTEN and beta-catenin for endometrial intraepithelial neoplasia in Japanese women, Ann. Diagn. Pathol. 11 (2007) 103–108.
[53] Y. Xiong, Y.Y. Xiong, Y.F. Zhou, Expression and significance of beta-catenin, Glut-1 and PTEN in proliferative endometrium, endometrial intraepithelial neoplasia and endometrioid adenocarcinoma, Eur. J. Gynaecol. Oncol. 31 (2010) 160–164.
[54] G.L. Mutter, Histopathology of genetically defined endometrial precancers, Int. J. Gynecol. Pathol. 19 (2000) 301–309.
[55] H. Lee, H.J. Choi, C.S. Kang, H.J. Lee, W.S. Lee, C.S. Park, Expression of miRNAs and PTEN in endometrial specimens ranging from histologically normal to hyperplasia and endometrial adenocarcinoma, Mod. Pathol. 25 (2012) 1508–1515.
[56] T.L. Mao, I.M. Shih, The roles of ARID1A in gynecologic cancer, J. Gynecol. Oncol. 24 (2013) 376–381.
[57] H.M.J. Werner, A. Berg, E. Wik, E. Birkeland, C. Krakstad, K. Kusonmano, K. Petersen, K.H. Kalland, A.M. Oyan, L.A. Akslen, J. Trovik, H.B. Salvesen, ARID1A loss is prevalent in endometrial hyperplasia with atypia and low-grade endometrioid carcinomas, Mod. Pathol. 26 (2013) 428–434.
[58] M.J.W. Berends, H. Hollema, Y. Wu, T. van der Sluis, R.G.J. Mensink, K.A. ten Hoor, R.H. Sijmons, E.G.E. de Vries, E. Pras, M.J.E. Mourits, R.M.W. Hofstra, C. Buys, J.H. Kleibeuker, A.G.J. van der Zee, MLH1 and MSH2 protein expression as a pre-screening marker in hereditary and non-hereditary endometrial hyperplasia and cancer, Int. J. Cancer 92 (2001) 398–403.
[59] A. Orbo, M.N. Nilsen, M.S. Arnes, I. Pettersen, K. Larsen, Loss of expression of MLH1, MSH2, MSH6, and PTEN related to endometrial cancer in 68 patients with endometrial hyperplasia, Int. J. Gynecol. Pathol. 22 (2003) 141–148.
[60] G. Peiro, J. Diebold, G.B. Baretton, R. Kimmig, U. Lohrs, Cellular apoptosis susceptibility gene expression in endometrial carcinoma: correlation with Bcl-2, Bax, and caspase-3 expression and outcome, Int. J. Gynecol. Pathol. 20 (2001) 359–367.