Computational and Structural Biotechnology Journal 24 (2024) 205–212

Contents lists available at ScienceDirect

Computational and Structural Biotechnology Journal

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

Review Article

Role of artificial intelligence in digital pathology for gynecological cancers

Ya-Li Wang [a,b], Song Gao [c], Qian Xiao [a,c], Chen Li [d], Marcin Grzegorzek [e], Ying-Ying Zhang [a,f,g], Xiao-Han Li [h], Ye Kang [h], Fang-Hua Liu [a,f,g], Dong-Hui Huang [a,f,g], Ting-Ting Gong [c,*], Qi-Jun Wu [a,c,f,g,i,**]

[a] Department of Clinical Epidemiology, Shengjing Hospital of China Medical University, Shenyang, China
[b] Department of Information Center, The Fourth Affiliated Hospital of China Medical University, Shenyang, China
[c] Department of Obstetrics and Gynecology, Shengjing Hospital of China Medical University, Shenyang, China
[d] Microscopic Image and Medical Image Analysis Group, College of Medicine and Biological Information Engineering, Northeastern University, Shenyang, China
[e] Institute for Medical Informatics, University of Luebeck, Luebeck, Germany
[f] Clinical Research Center, Shengjing Hospital of China Medical University, Shenyang, China
[g] Liaoning Key Laboratory of Precision Medical Research on Major Chronic Disease, Shengjing Hospital of China Medical University, Shenyang, China
[h] Department of Pathology, Shengjing Hospital of China Medical University, Shenyang, China
[i] NHC Key Laboratory of Advanced Reproductive Medicine and Fertility (China Medical University), National Health Commission, Shenyang, China

ARTICLE INFO

Keywords:
Artificial intelligence
Digital pathology
Deep learning
Gynecological cancers
Machine learning

ABSTRACT

The diagnosis of cancer is typically based on histopathological sections or biopsies on glass slides. Artificial intelligence (AI) approaches have greatly enhanced our ability to extract quantitative information from digital histopathology images as a rapid growth in oncology data. Gynecological cancers are major diseases affecting women's health worldwide. They are characterized by high mortality and poor prognosis, underscoring the critical importance of early detection, treatment, and identification of prognostic factors. This review highlights the various clinical applications of AI in gynecological cancers using digitized histopathology slides. Particularly, deep learning models have shown promise in accurately diagnosing, classifying histopathological subtypes, and predicting treatment response and prognosis. Furthermore, the integration with transcriptomics, proteomics, and other multi-omics techniques can provide valuable insights into the molecular features of diseases. Despite the considerable potential of AI, substantial challenges remain. Further improvements in data acquisition and model optimization are required, and the exploration of broader clinical applications, such as the biomarker discovery, need to be explored.

1. Introduction

Gynecological cancers (GCs), primarily comprising ovarian cancer (OC), endometrial cancer (EC, also known as uterine cancer), and cervical cancer (CC), present a significant global public health concern with profound implications for women's health and quality of life, leading to substantial disease burden. The high mortality rate and poor prognosis associated with these cancers have imposed significant pressure for effective prevention and management strategies [1]. According to the 2020 global cancer statistics, the new cases of OC, CC and EC were 313,959, 604,127, and 417,367, the new deaths were 207,252, 341,831, and 97,370, respectively [2]. The high mortality rate of OC is related to late-stage diagnoses and a high rate of recurrence [3–5], with 5-year survival rates < 50% in most countries [6]. Premenopausal women account for 14% of EC cases, and 5% of them were younger than 40 years [7]. The standard treatment modalities for GCs entail surgical cytoreduction and systemic chemotherapy; nevertheless, a large proportion of patients experiences disease recurrence after completing chemotherapy [5]. Despite advancements in medical imaging techniques enhancing cancer detection rates, histopathological evaluation remains

* Correspondence to: Department of Obstetrics and Gynecology, Shengjing Hospital of China Medical University, No. 36, San Hao Street, Shenyang, Liaoning 110004, China.
** Correspondence to: Department of Clinical Epidemiology, Clinical Research Center, Liaoning Key Laboratory of Precision Medical Research on Major Chronic Disease, Department of Obstetrics and Gynecology, Shengjing Hospital of China Medical University, Key Laboratory of Reproductive and Genetic Medicine (China Medical University), No. 36, San Hao Street, Shenyang, Liaoning 110004, China.
E-mail addresses: gongtt@sj-hospital.org (T.-T. Gong), wuqj@sj-hospital.org (Q.-J. Wu).

https://doi.org/10.1016/j.csbj.2024.03.007
Received 28 December 2023; Received in revised form 8 March 2024; Accepted 9 March 2024
Available online 11 March 2024

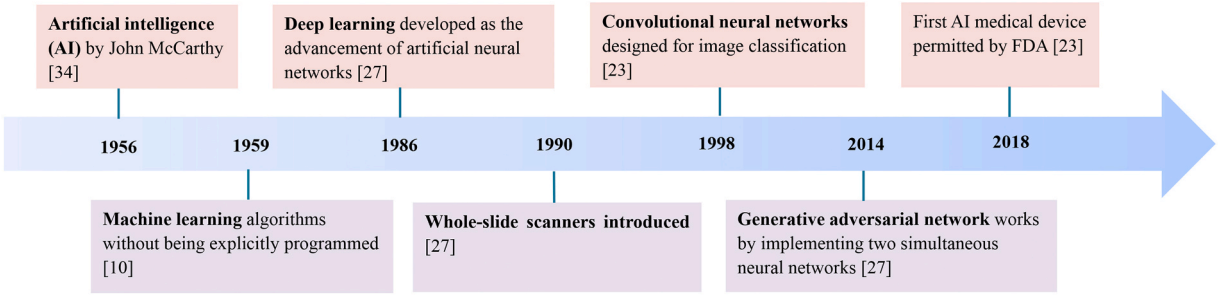

Fig. 1. A timeline of important historical events in artificial intelligence and digital pathology. FDA: Food and Drug Administration.

the gold standard for accurate cancer diagnosis and subsequent management. However, the escalating global incidence of GCs poses a growing challenge due to the expanding volume of pathological data and the shortage of pathologists [8].

Driven by high-dimensionality datasets, advances in computing hardware, and the utilization of deep learning (DL) models, the oncology is increasingly benefiting from artificial intelligence (AI) [9,10]. Various clinical applications of AI in oncology range from cancer detection [11] and classification [12], to predicting patient responses to therapy [13], lymph node metastasis [14,15], and prognosis [16,17]. The cancer types cover breast cancer [11,18], colorectal cancer [12,16], gastric cancer [17], lung cancer [19,20], prostate cancer [21], and lymphoma [22]. As the volume of healthcare data in cancer management continues to grow, the integration of AI holds promises for comprehensive utilization throughout the entire spectrum of cancer prevention and treatment, ultimately guiding clinical decision-making processes [10].

This review will describe the basic concepts and principles of digital pathology and AI in histopathology images analysis. Then, we will present the clinical applications achieved by AI in the digital pathology of GCs. Finally, the challenges of applying AI to the clinic will also be discussed.

2. Digital pathology

Histopathology is the examination and analysis of glass slides under a microscope, and it serves as the cornerstone of cancer diagnosis. Manual annotation of histological slides by pathologists is time-consuming, subjective, and susceptible to intra- and inter-observer variability [23,24]. The demand for diagnostic accuracy in cancer histopathology is increasing because accurate biomarker evaluation is required for personalized cancer therapy [25]. Digital pathology, originating in the 1960 s, is the process of digitizing histopathology slides into whole-slide images (WSIs) that can be reviewed by pathologists on computer monitor [26,27]. Digital slides are easier to preserve, share and annotate, and facilitate remote diagnosis or educational purposes [23]. The automation and efficiency afforded by digital pathology can enhance productivity and cost-efficiency [26], with its performance having demonstrated superiority over conventional microscopy [28,29]. Recently, the Food and Drug Administration (FDA) has approved the use of digital pathology for primary diagnosis [30]. Nevertheless, the abundance and intricate information among different cell types, along with the spatial context provided by digital pathology, has underscored the necessity for precise analysis of large datasets [31–33]. Thus, implementing robust and reproducible AI-based methods might potentially resolve the challenges faced by oncologists and pathologists.

3. AI in digital pathology

AI, which originated in the 1950 s, refers to a broad field of computer science that involves utilizes machine-based techniques to model the human decision-making process and generate predictions [34]. Machine learning (ML) refers to computer programs that process data for intelligent analysis, serving as a fundamental research method in the field of AI [35]. The main steps of ML involve annotation, feature extraction, and model prediction, empowering machines to automatically train and optimize models through statistical methods [25]. Supervised, unsupervised, and reinforcement learning represent the key learning types in ML, addressing tasks such as classification, regression, clustering and dimensionality reduction [23]. Weakly supervised learning represents an intermediate learning paradigm that lies between supervised and unsupervised learning. Common ML algorithms encompass linear or logistic regression, decision tree-based methods, and support vector machine (SVM) [23]. DL is a subset of ML which based on neural network structures, comprising interconnected input, hidden, and output layers that automatically extracts data features, overcomes the limitations and challenges of handcrafted features in ML [35]. Convolutional neural networks (CNNs) have gained widespread deployment in pathology image analysis since 2012 when AlexNet secured the first place in the ImageNet Large Scale Visual Recognition Challenge [35]. Subsequently, other deep CNN models have been developed and applied in medical domains. Fig. 1 provides a succinct summary of important historical events in AI and digital pathology.

4. AI in digital pathology for GCs

The application of AI in GCs predominantly commenced after 2017, with the classification of histopathological subtype accounting for the largest proportion (Table 1). DL methods are the most commonly used AI algorithms (Fig. 2), with CNNs being the most extensively employed model (Table 2). The model's performance, as assessed by the area under the curve (AUC), ranges from 0.71 to 0.99 across all tumor types (Table 3).

4.1. Classification of histopathological subtype

The treatment strategies and clinical prognosis for distinct histopathological types of GCs exhibit variability. ML and DL have been tested as methods for the classification of GCs subtypes.

4.1.1. Ovarian cancer

In 2017, BenTaieb et al.[36] utilized weakly-supervised ML approaches based on SVM for the classification of OC subtypes. The model achieved an average multi-class classification accuracy of 90%, obtaining substantial agreement with clinicians (Kappa=0.89). In discerning between the two epithelial OC types, ML models achieved 91–95% accuracy [37]. Meanwhile, in the cell-level classification of both tumor and stroma cells, the models demonstrated accuracy exceeding 90%. Wu et al. [38] employed deep CNNs (DCNN) based on AlexNet for OC subtype classification, achieving an accuracy of 78.20% on augmented data. In Farahani's investigation, DCNNs achieved a diagnostic concordance of 81.38% in the training set and 80.97% in the external set for classifying OC subtypes [39].

Table 1
Basic features for the included studies.

Author [ref], year	Cancer type	WSI type	Objectives	Sample size	Participants
BenTaieb et al. [36], 2017	Ovarian cancer	H&E	Subtype classification	133	Ovarian carcinoma patients with different subtypes (HGSC, EN, MC, LGSC, CC)
Jiang et al.[37], 2021	Ovarian cancer	H&E	Subtype classification	30	SBOT and HGSOC patients were retrieved from the institutional pathology system database
Wu et al.[38], 2018	Ovarian cancer	H&E	Subtype classification	85	Ovarian cancer patients with different subtypes (serous carcinoma, MC, endometrioid, and CC) were obtained from the First Affiliated Hospital of Xinjiang Medical University
Farahani et al.[39], 2022	Ovarian cancer	H&E IHC	Subtype classification	485	Patients from the OVCARE archives and the University of Calgary
Hong et al.[40], 2021	Endometrial cancer	H&E	Subtype classification	456	Train, validate, and test data from the TCGA and the Clinical Proteomic Tumor Analysis Consortium (CPTAC). Independent dataset from New York University (NYU) hospitals
Song et al.[41], 2022	Endometrial cancer and Cervical cancer	H&E	Subtype classification	230 (70 for CC, 160 for EC)	Data from The Cancer Genome Atlas (TCGA) program and The National Cancer Institute's Clinical Proteomic Tumor Analysis Consortium (CPTAC) endometrial cancer dataset
Li et al.[42], 2023	Cervical cancer	H&E	Subtype classification	229	Cervical specimens from January 2018 and December 2020 were acquired from the Department of Pathology, Xinhua hospital Chongming branch affiliated with Shanghai Jiaotong University
Habtemariam et al. [43], 2022	Cervical cancer	H&E	Subtype classification	915 WSIs	Four cervical cancer classes (normal, precancer, adenocarcinoma, and squamous cell carcinoma) were gathered from Jimma University Medical Center (JUMC) and St. Paul Hospital
Shin et al.[44], 2021	Ovarian cancer	H&E	Diagnosis (tumor *vs* non-tumor)	142	Ovarian serous cystadenocarcinoma (HGSC) data from the Cancer Image Archive and the Ajou University Medical Center
Sengupta et al.[45], 2022	Ovarian cancer	IHC	Diagnosis (tumor *vs* non-tumor)	NR	Ovarian cancer patients were obtained during frontline surgery at Tata Medical Center (TMC), Kolkata
Mohammadi et al. [46], 2022	Endometrial cancer	H&E	Diagnosis (benign *vs* malignant)	2910	The tissue blocks originate from Glasgow Royal Infirmary (NG), Southern General Hospital (SG), Royal Alexandria Hospital (RAH), and Queen Elizabeth University Hospital (QEUH) (all in Glasgow, Scotland)
Zhang et al.[47], 2022	Endometrial cancer	H&E	Diagnosis (tumor *vs* non-tumor)	1190 WSIs	Endometrial specimens collected from PUPH, including all main pathological subtypes of the endometrium and the Chinese PLA General Hospital (PLAGH)
Sun et al.[48], 2020	Endometrial cancer	H&E	Diagnosis (benign *vs* malignant)	498	Patients from the Third Affiliated Hospital of Zhengzhou University from October 2017 to August 2018
H.J. Fick et al.[49], 2021	Cervical cancer	H&E	Diagnosis (benign *vs* malignant)	1015 WSIs	NR
Du et al.[50], 2018	Ovarian cancer	H&E	Prognosis	154 WSIs	Breast cancer and ovarian cancer tissue, the former came from the Stanford Tissue Microarray Database (TMAD) and the latter came from the OUHSC
Laury et al.[51], 2021	Ovarian cancer	H&E	Prognosis	30	Stage III-IV high-grade extrauterine serous carcinoma who underwent primary cytoreductive surgery, and at least 6 cycles of adjuvant platinum-based chemotherapy from HUS Helsinki University Hospital between 1982 and 2013
Zeng et al.[52], 2021	Ovarian cancer	H&E	Prognosis	229	HGSOC patients from The Cancer Genome Atlas (TCGA)
Nero et al.[53], 2022	Ovarian cancer	H&E	Prognosis	664	EOC patients from the Fondazione Policlinico Universitario "Agostino Gemelli" IRCCS of Rome, Italy, from November 2016 to November 2020
Fremond et al.[54], 2022	Endometrial cancer	H&E	Prognosis	2028	Endometrial cancer patients from three randomised trials and four clinical cohorts: the randomised PORTEC-1 trial (recruited in the Netherlands); the randomised PORTEC-2 trial (the Netherlands); the randomised PORTEC-3 trial (the Netherlands, UK, France, Italy, Canada, Australia, and New Zealand); the retrospective TransPORTEC pilot Study (the Netherlands, UK, and France); the prospective Medisch Spectrum Twente (MST) cohort (the Netherlands); patients with POLEmut endometrial cancer from the Leiden Endometrial Cancer Repository (the Netherlands); and TCGA-Uterine Corpus Endometrial Carcinoma cohort (TCGA-UCEC), extracted from the cBioPortal for Cancer Genomics
Chen et al.[55], 2023	Cervical cancer	H&E	Prognosis	251	Patients with the International Federation of Gynecology and Obstetrics (FIGO) Stage IA1–IIA2 cervical cancer were collected from Nanfang Hospital of Southern Medical University (Guangzhou,China) from January 2009 to December 2016 and other hospitals
Wang et al.[56], 2022	Ovarian cancer	H&E	Therapeutic response	288 WSIs	HGSOC patients are collected from the tissue bank of the TriService General Hospital and the National Defense Medical Center, Taipei, Taiwan
Heindl et al.[57], 2018	Ovarian cancer	H&E IHC	Cancer microenvironment	514	Patients with International Federation of Gynecology and Obstetrics (FIGO) stage II-IV HGSOC from TCGA
Desbois et al.[58], 2020	Ovarian cancer	IHC	Cancer microenvironment	370	Epithelial ovarian cancer from mixed histology were collected from the Phase III ICON7 clinical trial. Independent validation collection was procured from Cureline, Inc (Brisbane, CA, USA)

Abbreviation: CC: clear cell carcinoma, EN: endometrioid carcinoma, EOC: epithelial ovarian cancer, H&E: hematoxylin and eosin, HGSC: high grade serous carcinoma, HGSOC: high grade serous ovarian cancer, IHC: immunohistochemistry, LGSC: low grade serous carcinoma, MC: mucinous carcinoma, NR: not reported, PUPH: Peking University People's Hospital, PLAGH: Chinese PLA General Hospital, SBOT: serous borderline ovarian tumor, WSIs: whole slide images.

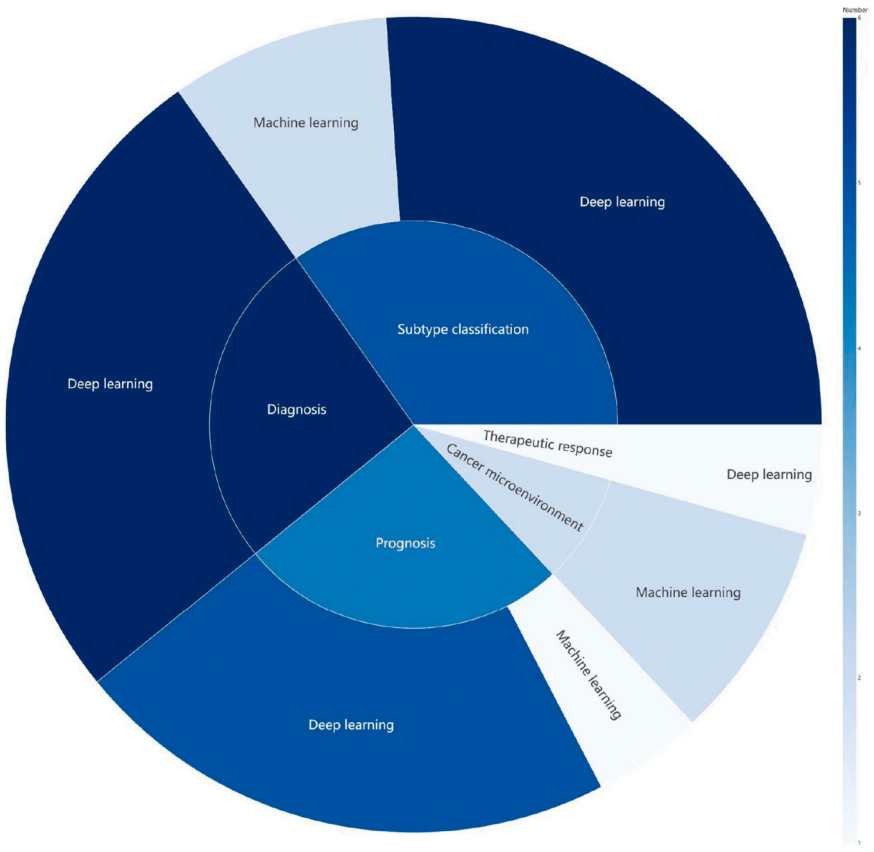

Fig. 2. The clinical applications of AI in digital pathology of gynecological cancer.

4.1.2. Endometrial cancer

The DCNN model achieved a per-patient level AUC of 0.969 (0.905–1) for differentiating samples into endometrioid or serous histological subtypes. Additionally, this model offers insights into molecular subtypes and mutation status [40]. On the other hand, the Inception-v3 model attained an AUC value of 0.944 for classifying the EC subtype [41].

4.1.3. Cervical cancer

In 2023, Li et al. [42] employed AlexNet, VGG-19, Xception, and ResNet-50 with five-fold cross-validation to identify cervical malignancies and provide diagnostic interpretability. The AUC for internal validation varied from 0.73 to 0.98. Habtemariam et al. used the EfficientNetB0 pre-trained model for CC classification, and the results were validated using histogram-matched histopathological images. The model achieved a test accuracy of 94.5% for classifying CC [43].

4.2. Cancer diagnosis

The diagnostic task involved the differentiation between tumors and non-tumors, as well as between benign and malignant lesions. All diagnostic tasks were based on DL methods (Table 2).

4.2.1. Ovarian cancer

In 2021, Shin et al. [44] utilized the Inception V3 model for the detection of malignancy on tissue slides. They examined a public set of tissue slide images comprising 142 patients diagnosed with ovarian serous cystadenocarcinoma from The Cancer Genome Atlas Ovarian. Notably, the researchers evaluated the classifier's performance stability using style transfer techniques on a limited institutional dataset. After the style transfer, the AUC and area under the precision recall curve improved from 0.737 (0.708–0.764) and 0.710 (0.672–0.748) to 0.916 (0.899–0.930) and 0.898 (0.872–0.922), respectively. The Inception V3 model was also used by Sengupta et al. [45] for OC diagnosis, with lamin-induced morphological changes of the nuclei as the input parameter. The model showed a higher performance in distinguishing between normal and OC tissues, achieving an AUC of 0.99 in both the training and validation sets.

4.2.2. Endometrial cancer

Mohammadi et al. [46] employed the CLAM (attention multiple instance learning) model to differentiate between malignant and benign tumors, achieving a validation accuracy of 85% and a test accuracy exceeding 87%. The attention heatmapping, feature visualization, and end-to-end saliency-mapping improved the interpretability of the model. Zhang et al. [47] utilized DeepLab v3 and ResNet-50 for the diagnosis of EC and non-EC in multiple datasets, demonstrating good performance (AUC, sensitivity, and specificity all >0.8). Sun and collaborators [48] proposed the HIENet framework, based on VGG-16 and incorporates two crucial blocks that utilize the visual attention mechanism. HIENet exhibited an AUC of 0.96 ± 0.01, with a sensitivity of $81.04 \pm 3.87\%$ and specificity of $94.78 \pm 0.87\%$ in the detection of endometrioid adenocarcinoma.

4.2.3. Cervical cancer

In 2021, models utilizing CNN and SVM algorithms achieved an 85% classification accuracy in 1015 annotated WSIs [49]. The CNN was employed to predict the probability of each of the four lesion classes at the patch level, and the SVM was utilized to predict the final slide-level lesion status.

4.3. Cancer prognosis

The prognosis of cancer is affected by various complex factors, and

Table 2
Model, algorithm, and model validation for the included studies.

Author [ref], year	Model	Algorithm	CV	External validation	AI *vs* clinicians
BenTaieb et al.[36], 2017	ML	SVM	leave-one-out	No	Yes
Jiang et al. [37], 2021	ML	SVM	NR	No	No
Wu et al.[38], 2018	DL	CNN (AlexNet)	10-fold	No	No
Farahani et al. [39], 2022	DL	CNN	3-fold	Yes	No
Hong et al. [40], 2021	DL	CNN (InceptionResnet)	NR	Yes	No
Song et al. [41], 2022	DL	CNN (Inception-v3)	5-fold	Yes	No
Li et al.[42], 2023	DL	CNN (AlexNet, VGG-19, Xception, ResNet-50)	5-fold	Yes	Yes
Habtemariam et al.[43], 2022	DL	EffecientNetB0	10-fold	Yes	No
Shin et al. [44], 2021	DL	CNN (Inception V3)	NR	Yes	No
Sengupta et al. [45], 2022	DL	CNN (Inception V3)	5-fold	No	No
Mohammadi et al.[46], 2022	DL	CLAM	NR	Yes	No
Zhang et al. [47], 2022	DL	DeepLab v3, ResNet-50	NR	Yes	No
Sun et al.[48], 2020	DL	CNN (VGG-16)	10-fold	Yes	Yes
H.J. Fick et al. [49], 2021	DL	SVM, CNN (DenseNet)	10-fold	No	No
Du et al.[50], 2018	DL	CNN (AlexNet, Places365-AlexNet, GoogLeNet)	NR	Yes	No
Laury et al. [51], 2021	DL	CNN	NR	No	No
Zeng et al. [52], 2021	ML	RF, GBDT, AdaBoost, LR, DT, SVM, NB, KNN	5-fold	Yes	No
Nero et al. [53], 2022	DL	CLAM	NR	No	No
Fremond et al. [54], 2022	DL	HoVer-Net, SVM	4-fold	Yes	No
Chen et al. [55], 2023	DL	CNN (ResNet-50)	NR	Yes	No
Wang et al. [56], 2022	DL	Inception V3	5-fold	Yes	No
Heindl et al. [57], 2018	ML	SVM	NR	Yes	No
Desbois et al. [58], 2020	ML	RF, k-means clustering	NR	Yes	No

Abbreviation: AI: artificial intelligence, AdaBoost: adaptive boosting, CV: cross-validation, CNN: convolutional neural network, CLAM: clustering-constrained attention multiple instance learning, DT: decision tree, DL: deep learning, GBDT: gradient boosting decision tree, KNN: K-nearest neighbor, SVM: support vector machine, LR: logistic regression, ML: machine learning, NB: naive Bayesian, NR: not reported, RF: random forest.

AI-based methods hold the potential to enhance prognosis prediction in GCs.

4.3.1. Ovarian cancer

Du et al. [50] explored diverse transfer learning strategies to effectively differentiate between epithelial and stromal regions in hematoxylin and eosin (H&E) stained histological images. Utilizing DCNNs (AlexNet, Places365-AlexNet, GoogLeNet, and two modified AlexNet models), they extracted natural-image features without fine-tuning, subsequently conducting end-to-end fine-tuning by training the classifiers at certain layers. An accuracy of 90.2 was achieved with the implementation of GoogLeNet. In study by Laury et al. [51], 205 WSIs from 30 patients with high-grade serous ovarian cancer exhibiting distinct treatment responses (platinum-free intervals of $\leq$6 months or $\geq$18 months) were analyzed for outcome prediction. CNN-based models effectively differentiated extreme patient responses to primary platinum-based chemotherapy, achieving a sensitivity of 73% and a specificity of 91%. Furthermore, besides prognostication, ML models combined with other data were utilized to infer molecular features. In Zeng's [52] study, the model's AUC for predicting 5-year overall survival (OS) was 0.825 and 0.703 in the test and validation sets, respectively. Notably, for the prediction of molecular features (BRCA mutation, microsatellite instability, and molecular subtypes), all AUC values were $>$ 0.9. The CLAM method demonstrated an AUC of 0.71 for predicting progression-free survival in 664 epithelial OC patients, while the AUC for predicting BRCA mutation was 0.55 [53]. Despite the relatively modest predictive performance, it also suggests that the AI model has the potential to provide information on molecular features of the disease.

4.3.2. Endometrial cancer

A comprehensive analysis integrating data from the PORTEC randomized trials and clinical cohorts was conducted to develop an interpretable DL pipeline aimed at predicting progression (5-year recurrence-free survival) [54]. This model could be clinically applied for pre-screening EC to identify occurrences of p53abn for further confirmatory immunohistochemistry or molecular testing.

4.3.3. Cervical cancer

Chen et al. [55] developed a CNN-based pathological risk score (RS) to predict patient prognosis. The performance of the RS in predicting OS and disease-free survival (DFS) was validated through Kaplan–Meier survival analysis in both the training and testing datasets. In the testing cohort, the RS exhibited an AUC of up to 0.80 for predicting both OS and DFS.

4.4. Cancer therapeutic response and microenvironment

In addition to the tasks mentioned above, DL models based on H&E staining have also been used to predict the therapeutic efficacy (invalid *vs.* effective) of bevacizumab in OC. The method achieved an accuracy of 0.882 $\pm$ 0.06, a precision of 0.921 $\pm$ 0.04, a recall of 0.912 $\pm$ 0.03; and an F-measure of 0.917 $\pm$ 0.07. The findings suggest that the utilization of DL techniques holds promise in providing valuable guidance for treatment decisions [56]. Furthermore, ML combined with omics and clinical data has elucidated the microenvironment of OC, including dysregulation of DNA repair, loss of nuclear integrity [57], and tumour-immune phenotypes [58].

5. Limitations and challenges

The studies illustrates that the utilization of AI has facilitated technological advances in GC's digital pathology, showing promising potential in various applications. However, the translation of these techniques into clinical practice may take several years attributable to existing limitations and challenges.

5.1. 'Black box' problem and interpretability

Despite the robust capabilities of AI-based models, the development of explainable AI models is imperative for clinical practice. Understanding and explaining how AI-based algorithms work and how the model arrives at its decisions is an important hurdle for the adoption of AI-based methods in clinical practice. This 'black box' nature limits its clinical application. In handcrafted approaches, relevant features from the data are manually selected, which requires close collaboration

Table 3
The performance of the model for the included studies.

Author [ref], year	Performance							
	Accuracy (%)	AUC	SE (%)	SP (%)	Precision	Recall	Kappa	F1 score
BenTaieb et al.[36], 2017	90.00	/	/	/	/	/	0.89	0.66
Jiang et al.[37], 2021	> 90.00	/	/	/	/	/	/	/
Wu et al.[38], 2018	78.20	/	/	/	/	/	/	/
Farahani et al.[39], 2022	/	0.95	/	/	/	/	0.74	0.79
Hong et al.[40], 2021	/	0.97 (0.91-1.00)	/	/	1	0.60	/	0.75
Song et al.[41], 2022	89.90	0.94 (0.92-0.97)	84.60	93.90	/	/	/	0.88
Li et al.[42], 2023	92.50 ± 1.90	0.95 ± 0.01	/	/	0.94 ± 0.02	0.95 ± 0.03	/	/
Habtemariam et al.[43], 2022	94.50	/	/	/	0.96	1	0.92	0.98
Shin et al.[44], 2021	80.80	0.92 (0.90-0.93)	95.80	65.80	0.74	0.96	/	0.83
Sengupta et al.[45], 2022	/	0.99	/	/	/	/	/	/
Mohammadi et al.[46], 2022	85.57	0.95	/	/	/	/	/	/
Zhang et al.[47], 2022	/	0.93	92.40	80.10	/	/	/	/
Sun et al.[48], 2020	93.53 ± 0.81	0.96 ± 0.01	81.04 ± 3.87	94.78 ± 0.87	/	/	/	/
H.J. Fick et al.[49], 2021	85.00	/	/	/	/	/	/	/
Du et al.[50], 2018	90.20	/	/	/	/	/	/	/
Laury et al.[51], 2021	82.00	/	73.00	91.00	/	/	/	/
Zeng et al.[52], 2021	/	0.83	/	/	/	/	/	/
Nero et al.[53], 2022	/	0.71	/	/	/	/	/	/
Fremond et al.[54], 2022	/	0.87 (0.86-0.89)	/	/	/	/	/	/
Chen et al.[55], 2023	/	0.87 (0.77-0.96)	/	/	/	/	/	/
Wang et al.[56], 2022	88.20 ± 6	/	/	/	0.92 ± 0.04	0.91 ± 0.03	/	0.92 ± 0.07
Heindl et al.[57], 2018	85.00	/	/	/	/	/	/	/
Desbois et al.[58], 2020	91.00	/	/	/	/	/	/	/

Abbreviation: AUC: area under the curve, SE: sensitivity, SP: specificity, /: The results were not reported.

between clinicians and experts. The integration of DL with handcrafted strategies leverages domain expertise to ensure the biological interpretability of the results generated [27]. Furthermore, some visualization methods have been developed to improve interpretability [46]. On the other hand, the integration of algorithms with other patient data, such as follow-up and clinicopathological information, has also contributed to improving interpretability to a certain extent [54].

5.2. Quality of data

The performance of an algorithm is influenced by the nature of the task, including the required level of accuracy and the quality of the samples being assessed [24]. To achieve optimal predictive performance and utility, AI algorithms need to be trained on clean and accurate data with a high signal-to-noise ratio [24,27]. This can be challenging when dealing with histological data obtained from various laboratories. Various color normalization methods, such as spectral sensing [59], stain color adaptive normalization [60], adaptive color deconvolution [61], and transfer learning approaches [23,50], can be employed to address this issue. Moreover, loss of data fidelity may occur when the scanner exceeds its maximum scanning capacity. Super-resolution microscopy techniques offer a solution by enabling higher resolution focusing on specific biological elements [62].

5.3. The generalization of AI models

The model's performance is affected by intrinsic variations in datasets [39]. Therefore, improving the model's generalization ability as a pressing issue that necessitates resolution. Cross-validation is a common method employed to improve the generalization ability of models. However, due to data constraints, several studies omit the implementation of cross-validation (Table 2). Additionally, refining the model through fine-tuning has the potential to reduce the generalization error [50].

5.4. AI algorithms and validation

The dataset utilized for training AI models is typically divided into training and validation sets. The training set is typically balanced, while the validation set is derived either from the original dataset or sourced from another institution [27]. Validation sets are necessary to prevent overfitting. However, external validation is often lacking due to the difficulty in acquiring data (Table 2). In addition, AI algorithms may exhibit limitations in specific domains, such as mutation detection, where their performance is typically lower [25,53].

6. Future directions

The advent of digital pathology has ushered in new prospects for generating extensive, high-resolution digital data. The integration of AI in image analysis has catalyzed advancements across various areas of medical imaging. Anticipated future developments in AI for GCs include the identification of biomarkers. However, this new technology is confronted with several challenges that necessitate resolution before its integration into clinical practice. For instance, it is still uncertain whether AI applications can replace some expensive molecular tests for cancer screening and molecular phenotype stratification. However, given the increasing global incidence of cancer and the auspicious potential of AI-based methods in accurately classifying pathological images at a lower cost, it is likely that these challenges will be overcome.

7. Conclusions

Digital pathology facilitates the digitized acquisition, management, and interpretation of information. AI in digital pathology provides opportunities for computational analysis. This review summarizes the application of AI-based digital pathology in the diagnosis, histopathological classification, prognosis, and assessment of therapeutic responses in GCs. DL algorithms, particularly CNNs, have demonstrated superior performance. Despite the advantages offered by AI, significant challenges such as interpretability, data quality, model generalization, and validation need to be addressed. Moreover, certain areas such as biomarker detection and multi-center studies remain underexplored. However, with the ongoing advancements in AI algorithms, it is envisaged that their application will improve and become more widespread in clinical practice. Furthermore, the anticipation is that large, prospective studies and clinical trials testing AI methods will become universal in the future.

CRediT authorship contribution statement

Y-LW, SG, QX, T-TG, and Q-JW conceived the study. Y-LW, SG, T-TG, and Q-JW contributed to the design. Y-LW, QX, Y-YZ, X-HL, YK, D-HH and F-HL collected the data. Y-LW and Y-YZ cleaned the data and checked the discrepancy. Y-LW, SG, QX, CL, MG, Y-YZ, T-TG, and Q-JW drafted the article and revised it critically for important intellectual content. T-TG and Q-JW agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. All authors interpreted the data, read the manuscript, and approved the final vision. Y-LW and SG contributed equally to this work.

Declaration of Competing Interest

The authors report no conflict of interest.

Acknowledgements

This work was supported by the National Key R&D Program of China (No. 2022YFC2704200 to Wu QJ), the Natural Science Foundation of China (No. 82073647 and No. 82373674 to Wu QJ and No. 82103914 to Gong TT), Liaoning Province Science and Technology Plan (No. 2023JH2/20200019 to Wu QJ), LiaoNing Revitalization Talents Program (No. XLYC1907102 to Wu QJ), Outstanding Scientific Fund of Shengjing Hospital (No. M1150 to Wu QJ), and 345 Talent Project of Shengjing Hospital of China Medical University (No. M0268 to Wu QJ and No. M0952 to Gong TT).

References

[1] Shrestha Pallabi, Poudyal Bhavya, Yadollahi Sepideh, Wright Darryl E, Gregory Adriana V, et al. A systematic review on the use of artificial intelligence in gynecologic imaging – background, state of the art, and future directions. Gynecol Oncol 2022;166(3):596–605.
[2] Sung Hyuna, Ferlay Jacques, Siegel Rebecca L, Laversanne Mathieu, Soerjomataram Isabelle, et al. Global Cancer Statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2021;71(3):209–49.
[3] Jayson Gordon C, Kohn Elise C, Kitchener Henry C, Ledermann Jonathan A. Ovarian cancer. Lancet 2014;384(9951):1376–88.
[4] Momenimovahed Zohre, Tiznobaik Azita, Taheri Safoura, Salehiniya Hamid. Ovarian cancer in the world: epidemiology and risk factors. Int J Women's Health 2019;11:287–99.
[5] González-Martín Antonio, Pothuri Bhavana, Vergote Ignace, DePont Christensen René, Graybill Whitney, et al. Niraparib in patients with newly diagnosed advanced ovarian cancer. N Engl J Med 2019;381(25):2391–402.
[6] Gong Ting-Ting, Liu Fang-Hua, Liu Ya-Shu, Yan Shi, Xu He-Li, et al. A follow-up study of ovarian cancer (OOPS): a study protocol. Front Nutr 2022;9:872773.
[7] Morice Philippe, Leary Alexandra, Creutzberg Carien, Abu-Rustum Nadeem, Darai Emile. Endometrial cancer. Lancet 2016;387(10023):1094–108.
[8] Robertson Stephanie, Azizpour Hossein, Smith Kevin, Hartman Johan. Digital image analysis in breast pathology-from image processing techniques to artificial intelligence. Transl Res 2018;194:19–35.
[9] Bhinder Bhavneet, Gilvary Coryandar, Madhukar Neel S, Elemento Olivier. Artificial intelligence in cancer research and precision medicine. Cancer Discov 2021;11(4):900–15.
[10] Kann Benjamin H, Hosny Ahmed, Aerts Hugo JWL. Artificial intelligence for clinical oncology. Cancer Cell 2021;39(7):916–27.
[11] Hu Qiyuan, Whitney Heather M, Li Hui, Ji Yu, Liu Peifang, et al. Improved classification of benign and malignant breast lesions using deep feature maximum intensity projection MRI in breast cancer diagnosis using dynamic contrast-enhanced MRI. Radio Artif Intell 2021;3(3):e200159.
[12] Rathore Saima, Iftikhar Muhammad Aksam, Chaddad Ahmad, Niazi Tamim, Karasic Thomas, et al. Segmentation and grade prediction of colon cancer digital pathology images across multiple institutions. Cancers 2019;11(11):1700.
[13] Ha Richard, Chin Christine, Karcich Jenika, Liu Michael Z, Chang Peter, et al. Prior to initiation of chemotherapy, can we predict breast tumor response? Deep learning convolutional neural networks approach using a breast MRI tumor dataset. J Digit Imaging 2019;32(5):693–701.
[14] David F. Steiner, Robert MacDonald, Yun Liu, Peter Truszkowski, Jason D. Hipp, et al. Impact of Deep Learning Assistance on the Histopathologic Review of Lymph Nodes for Metastatic Breast Cancer. Am J Surg Pathol 42(12): 1636–1646.
[15] Liu Yun, Kohlberger Timo, Norouzi Mohammad, Dahl George E, Smith Jenny L, et al. Artificial intelligence-based breast cancer nodal metastasis detection: insights into the black box for pathologists. Arch Pathol Lab Med 2019;143(7):859–68.
[16] Kather Jakob Nikolas, Krisam Johannes, Charoentong Pornpimol, Luedde Tom, Herpel Esther, et al. Predicting survival from colorectal cancer histology slides using deep learning: a retrospective multicenter study. PLoS Med 2019;16(1):e1002730.
[17] Chen Dexin, Fu Meiting, Chi Liangjie, Lin Liyan, Cheng Jiaxin, et al. Prognostic and predictive value of a pathomics signature in gastric cancer. Nat Commun 2022;13(1):6903.
[18] Jon Whitney German Corredor, Andrew Janowczyk Shridar Ganesan, Doyle Scott, et al. Quantitative nuclear histomorphometry predicts oncotype DX risk categories for early stage ER+ breast cancer. BMC Cancer 2018;18(1):610.
[19] Corredor Germán, Wang Xiangxue, Zhou Yu, Lu Cheng, Fu Pingfu, et al. Spatial architecture and arrangement of tumor-infiltrating lymphocytes for predicting likelihood of recurrence in early-stage non-small cell lung cancer. Clin Cancer Res 2019;25(5):1526–34.
[20] Yu Kun-Hsing, Zhang Ce, Berry Gerald J, Altman Russ B, Ré Christopher, et al. Predicting non-small cell lung cancer prognosis by fully automated microscopic pathology image features. Nat Commun 2016;7:12474.
[21] Nir Guy, Hor Soheil, Karimi Davood, Fazli Ladan, Skinnider Brian F, et al. Automatic grading of prostate cancer in digitized histopathology images: learning from multiple experts. Med Image Anal 2018;50:167–80.
[22] Achi Hanadi El, Belousova Tatiana, Chen Lei, Wahed Amer, Wang Iris, et al. Automated diagnosis of lymphoma with digital pathology images using deep learning. Ann Clin Lab Sci 2019;49(2):153–60.
[23] Klein Christophe, Zeng Qinghe, Arbaretaz Floriane, Devêvre Estelle, Calderaro Julien, et al. Artificial intelligence for solid tumour diagnosis in digital pathology. Br J Pharm 2021;178(21):4291–315.
[24] Baxi Vipul, Edwards Robin, Montalto Michael, Saha Saurabh. Digital pathology and artificial intelligence in translational medicine and clinical practice. Mod Pathol 2022;35(1):23–32.
[25] Acs B, Rantalainen M, Hartman J. Artificial intelligence as the next step towards precision pathology. J Intern Med 2020;288(1):62–81.
[26] Wong Alex Ngai Nick, He Zebang, Leung Ka Long, To Curtis Chun Kit, Wong Chun Yin, et al. Current developments of artificial intelligence in digital pathology and its future clinical applications in gastrointestinal cancers. Cancers 2022;14(15):3780.
[27] Bera Kaustav, Schalper Kurt A, Rimm David L, Velcheti Vamsidhar, Madabhushi Anant. Artificial intelligence in digital pathology — new tools for diagnosis and precision oncology. Nat Rev Clin Oncol 2019;16(11):703–15.
[28] Borowsky Alexander D, Glassy Eric F, Wallace William Dean, Kallichanda Nathash S, Behling Cynthia A, et al. Digital whole slide imaging compared with light microscopy for primary diagnosis in surgical pathology. Arch Pathol Lab Med 2020;144(10):1245–53.
[29] Aloqaily Ayat, Polonia Antonio, Campelos Sofia, Alrefae Nusaiba, Vale Joao, et al. Digital versus optical diagnosis of follicular patterned thyroid lesions. Head Neck Pathol 2021;15(2):537–43.
[30] Mukhopadhyay Sanjay, Feldman Michael D, Abels Esther, Ashfaq Raheela, Beltaifa Senda, et al. Whole slide imaging versus microscopy for primary diagnosis in surgical pathology: a multicenter blinded randomized noninferiority study of 1992 cases (pivotal study). Am J Surg Pathol 2018;42(1):39–52.
[31] Aeffner Famke, Zarella Mark D, Buchbinder Nathan, Bui Marilyn M, Goodman Matthew R, et al. Introduction to digital image analysis in whole-slide imaging: a white paper from the digital pathology association. J Pathol Inf 2019;10:9.
[32] Yuan Jianda, Hegde Priti S, Clynes Raphael, Foukas Periklis G, Harari Alexandre, et al. Novel technologies and emerging biomarkers for personalized cancer immunotherapy. J Immunother Cancer 2016;4:3.
[33] Heindl Andreas, Nawaz Sidra, Yuan Yinyin. Mapping spatial heterogeneity in the tumor microenvironment: a new era for digital pathology. Lab Invest 2015;95(4):377–84.
[34] McCarthy J, Minsky ML, Rochester N, Shannon CE. A proposal for the dartmouth summer research project on artificial intelligence. August 31, 1955 AI Mag 2006;27(4):12.
[35] Zhao Xue, Bai Jing-Wen, Guo Qiu, Ren Ke, Zhang Guo-Jun. Clinical applications of deep learning in breast MRI. Biochim Biophys Acta Rev Cancer 2023;1878(2):188864.
[36] BenTaieb Aïcha, Li-Chang Hector, Huntsman David, Hamarneh Ghassan. A structured latent model for ovarian carcinoma subtyping from histopathology slides. Med Image Anal 2017;39:194–205.
[37] Jiang Jun, Tekin Burak, Guo Ruifeng, Liu Hongfang, Huang Yajue, et al. Digital pathology-based study of cell-and tissue-level morphologic features in serous borderline ovarian tumor and high-grade serous ovarian cancer. J Pathol Inf 2021;12:24.
[38] Wu Miao, Yan Chuanbo, Liu Huiqiang, Liu Qian. Automatic classification of ovarian cancer types from cytological images using deep convolutional neural networks. Biosci Rep 2018;38(3):BSR20180289.
[39] Farahani Hossein, Boschman Jeffrey, Farnell David, Darbandsari Amirali, Zhang Allen, et al. Deep learning-based histotype diagnosis of ovarian carcinoma whole-slide pathology images. Mod Pathol 2022;35(12):1983–90.
[40] Hong Runyu, Liu Wenke, DeLair Deborah, Razavian Narges, Fenyö David. Predicting endometrial cancer subtypes and molecular features from histopathology images using multi-resolution deep learning models. Cell Rep Med 2021;2(9):100400.
[41] Song JaeYen, Im Soyoung, Lee Sung Hak, Jang Hyun-Jong. Deep learning-based classification of uterine cervical and endometrial cancer subtypes from whole-slide histopathology images. Diagnostics 2022;12(11):2623.

[42] Li Yi-Xin, Chen Feng, Shi Jiao-Jiao, Huang Yu-Li, Wang Mei. Convolutional neural networks for classifying cervical cancer types using histological images. J Digit Imaging 2023;36(2):441–9.
[43] Habtemariam Lidiya Wubshet, Zewde Elbetel Taye, Simegn Gizeaddis Lamesgin. Cervix type and cervical cancer classification system using deep learning techniques. Med Devices 2022;15:163–76.
[44] Shin Seo Jeong, You Seng Chan, Jeon Hokyun, Jung Ji Won, An Min Ho, et al. Style transfer strategy for developing a generalizable deep learning application in digital pathology. Comput Methods Prog Biomed 2021;198:105815.
[45] Sengupta Duhita, Ali Sk Nishan, Bhattacharya Aditya, Mustafi Joy, Mukhopadhyay Asima, et al. A deep hybrid learning pipeline for accurate diagnosis of ovarian cancer based on nuclear morphology. PLoS One 2022;17(1):e0261181.
[46] Mohammadi Mahnaz, Cooper Jessica, Arandelović Ognjen, Fell Christina, Morrison David, et al. Weakly supervised learning and interpretability for endometrial whole slide image diagnosis. Exp Biol Med 2022;247(22):2025–37.
[47] Zhang Xiaobo, Ba Wei, Zhao Xiaoya, Wang Chen, Li Qiting, et al. Clinical-grade endometrial cancer detection system via whole-slide images using deep learning. Front Oncol 2022;12:1040238.
[48] Sun Hao, Zeng Xianxu, Xu Tao, Peng Gang, Ma Yutao. Computer-aided diagnosis in histopathological images of the endometrium using a convolutional neural network and attention mechanisms. IEEE J Biomed Health Inf 2020;24(6):1664–76.
[49] Fick Rutger HJ, Tayart Brice, Bertrand Capucine, Chan Lang Solene, Rey Tina, et al. A partial label-based machine learning approach for cervical whole-slide image classification: the winning tissuenet solution. Annu Int Conf IEEE Eng Med Biol Soc 2021;2021:2127–31.
[50] Du Yue, Zhang Roy, Zargari Abolfazl, Thai Theresa C, Gunderson Camille C, et al. Classification of tumor epithelium and stroma by exploiting image features learned by deep convolutional neural networks. Ann Biomed Eng 2018;46(12):1988–99.
[51] Laury Anna Ray, Blom Sami, Ropponen Tuomas, Virtanen Anni, Carpén Olli Mikael. Artifcial intelligence-based image analysis can predict outcome in high-grade serous carcinoma via histology alone. Sci Rep 2021;11(1):19165.
[52] Zeng Hao, Chen Linyan, Zhang Mingxuan, Luo Yuling, Ma Xuelei. Integration of histopathological images and multi-dimensional omics analyses predicts molecular features and prognosis in high-grade serous ovarian cancer. Gynecol Oncol 2021;163(1):171–80.
[53] Nero Camilla, Boldrini Luca, Lenkowicz Jacopo, Giudice Maria Teresa, Piermattei Alessia, et al. Deep-learning to predict BRCA mutation and survival from digital H&E slides of epithelial ovarian cancer. Int J Mol Sci 2022;23(19):11326.
[54] Fremond Sarah, Andani Sonali, Wolf Jurriaan Barkey, Dijkstra Jouke, Melsbach Sinéad, et al. Interpretable deep learning model to predict the molecular classification of endometrial cancer from haematoxylin and eosin-stained whole-slide images: a combined analysis of the PORTEC randomised trials and clinical cohorts. Lancet Digit Health 2022;S2589-7500(22):00210–2.
[55] Chen Chi, Cao Yuye, Li Weili, Liu Zhenyu, Liu Ping, et al. The pathological risk score: a new deep learning-based signature for predicting survival in cervical cancer. Cancer Med 2023;12(2):1051–63.
[56] Wang Ching-Wei, Chang Cheng-Chang, Lee Yu-Ching, Lin Yi-Jia, Shih-Chang Lo, et al. Weakly supervised deep learning for prediction of treatment effectiveness on ovarian cancer from histopathology images. Comput Med Imaging Graph 2022;99:102093.
[57] Heindl Andreas, Khan Adnan Mujahid, Rodrigues Daniel Nava, Eason Katherine, Sadanandam Anguraj, et al. Microenvironmental niche divergence shapes BRCA1-dysregulated ovarian cancer morphological plasticity. Nat Commun 2018;9(1):3917.
[58] Desbois Mélanie, Udyavar Akshata R, Ryner Lisa, Kozlowski Cleopatra, Yinghui Guan, et al. Integrated digital pathology and transcriptome analysis identifies molecular mediators of T-cell exclusion in ovarian cancer. Nat Commun 2020;11(1):5583.
[59] Tani Shinsuke, Fukunaga Yasuhiro, Shimizu Saori, Fukunishi Munenori, Ishii Kensuke, et al. Color standardization method and system for whole slide imaging based on spectral sensing. Anal Cell Pathol 2012;35(2):107–15.
[60] Salvi Massimo, Michielli Nicola, Molinari Filippo. Stain color adaptive normalization (SCAN) algorithm: separation and standardization of histological stains in digital pathology. Comput Methods Prog Biomed 2020;193:105506.
[61] Zheng Yushan, Jiang Zhiguo, Zhang Haopeng, Xie Fengying, Shi Jun, et al. Adaptive color deconvolution for histological WSI normalization. Comput Methods Prog Biomed 2019;170:107–20.
[62] Huang Bo, Bates Mark, Zhuang Xiaowei. Super-resolution fluorescence microscopy. Annu Rev Biochem 2009;78:993–1016.