

RESEARCH

Open Access



# Ultrasound-based radiomics for predicting the five major histological subtypes of epithelial ovarian cancer

Yang Yang<sup>1</sup>, Xinyu Ji<sup>2</sup>, Sen Li<sup>3</sup>, Xuemeng Gao<sup>2</sup>, Yitong Wang<sup>2</sup> and Ying Huang<sup>1,2\*</sup>

## Abstract

**Background** Computational approaches have been proposed using radiomics in order to assess tumour heterogeneity, which is motivated by the concept that biomedical images may contain underlying pathophysiology information and has the potential to quantitatively measure the heterogeneity of intra- and intertumours. Ovarian cancer has the highest mortality among malignant tumours of female reproductive system and can be further divided into many subtypes with different management strategies and prognosis. The purpose of our study is to develop and validate ultrasound-based radiomics models to distinguish the five major histological subtypes of epithelial ovarian cancer.

**Methods** From January 2018 to August 2022, 1209 eligible ovarian cancer patients were enrolled. There were two subjects in this study: all patients ( $n = 1209$ ) and patients with the five major histological subtypes ( $n = 1039$ ). After image segmentation manually, radiomics features were extracted and some clinical characteristics were added. Nine feature selection methods were used to select the optimal predictive features. Seven classifiers were carried out to construct models. Choose the combination with the best predictive performance as the final result.

**Results** As for low-grade serous carcinoma, endometrioid carcinoma, and clear cell carcinoma, the models yields AUCs below 0.80 in the 10-fold cross-validation in the two groups. As for mucinous carcinoma, the AUCs were 0.83(95%CI, 0.74–0.93) and 0.89(95%CI, 0.83–0.95) in the validation cohorts and 0.80(95%CI, 0.73–0.87) and 0.86(95%CI, 0.78–0.94) in the 10-fold cross-validation in the two groups, respectively. As for high-grade serous carcinoma (HGSC), the models showed AUCs of 0.87(95%CI, 0.83–0.91) and 0.85(95%CI, 0.81–0.89) in the validation cohorts and 0.87(95%CI, 0.85–0.89) and 0.84(95%CI, 0.81–0.87) in the 10-fold cross-validation in the two groups, respectively, and exhibited high consistency between the predicted results and the actual outcomes, and brought great net benefits for patients.

**Conclusions** The ultrasound-based radiomics models in discriminating HGSC and non-HGSC showed good predictive performance, as well as high consistency between the predicted results and the actual outcomes, and brought significant net benefits for patients.

**Keywords** Ultrasound-based, Radiomics, Histological subtype, Epithelial ovarian cancer

\*Correspondence:

Ying Huang  
huangying712@163.com

<sup>1</sup>Department of Ultrasound, Beijing Tiantan Hospital, Capital Medical University, Beijing, China

<sup>2</sup>Department of Ultrasound, Shengjing Hospital, China Medical University, No. 36 Sanhao Street, Heping District, Shenyang, Liaoning Province 110004, China

<sup>3</sup>Yizhun Medical AI, Beijing, China



## Background

Ovarian cancer (OC) has the highest mortality among malignant tumours of female reproductive system [1], with about 52, 100 new cases and 22, 500 cancer-related deaths annually in China [2]. Epithelial ovarian cancers (EOCs) account for more than 85% of all OCs, while germ cell and sex cord stromal cell derived OCs account for the remaining cases. EOCs are not a singular disease entity which can be further divided into five major histological subtypes, including high-grade serous carcinoma (HGSC), mucinous carcinoma (MC), clear cell carcinoma (CC), endometrioid carcinoma (EN), and low-grade serous carcinoma (LGSC), based on their histopathological and immunohistochemical characteristics, as well as the inherent molecular characteristics [3]. Their management and prognosis vary widely by subtype and stage [4]. The current treatment standard for EOCs is primary debulking surgery (PDS) or interval debulking surgery (IDS) following neoadjuvant chemotherapy (NACT) [5, 6]. The chemotherapy regimens and biological behaviors differ among different subtypes. MC, CC, EN, and LGSC are characterized by low invasiveness, insensitivity to chemotherapy, and better prognosis, while HGSC is highly invasive, sensitive to chemotherapy, and associated with poorer prognosis [7, 8]. The distinction between specific subtypes of ovarian cancer is beneficial to the development of precise medicine.

Nowadays, the diagnosis of ovarian cancer subtypes relies on pathology and still requires experienced pathologists, making it a challenging task. To reduce dependence on pathology, especially in the absence of pathological results, interventions targeting subtypes are necessary for ovarian cancer patients. Patients with ovarian cancer typically undergo imaging examinations and serum tumour biomarker detection. Ultrasound (US) is the most basic imaging modality, with advantages such as ease of use, radiation-free, and low cost. Currently, young ultrasound physicians, aided by guidelines and ultrasound experts, find it no longer difficult to distinguish benign and malignant ovarian adnexal masses [9, 10]. However, more imaging information is difficult to be detected by the naked eyes, and accurately identifying the subtypes of ovarian cancer is challenging. Scholars are attempting to utilize limited data to extract more information and have made progress in studying the differentiation of ovarian cancer subtypes from macroscopic and microscopic perspectives. Seven supervised machine learning classifiers based on 32 parameters commonly available from peripheral blood tests and age were able to predict the five major histological subtypes [11]. Distinct transcriptional programs stratify ovarian cancer cell lines into the five major histological subtypes [12]. Klein, O. et al. used Matrix-assisted laser desorption/ionization imaging mass spectrometry combined with

morphological features of protein expression to distinguish histological subtypes of EOCs from tissue microarrays by linear and nonlinear machine learning methods by analyzing Formalin-fixed-paraffin-embedded tissues [13].

Computational approaches have been proposed using radiomics in order to assess tumour heterogeneity [14]. The quantitative features extracted from digital medical images are mineable data, and the process of converting images to high-throughput data is termed radiomics, which is motivated by the concept that biomedical images may contain underlying pathophysiology information and has the potential to quantitatively measure the heterogeneity of intra- and intertumours [15]. A study on renal cancer has showed that computed tomography (CT) texture features, such as entropy, mean of the positive pixels, and standard deviation were significantly associated with the subtype (clear vs. non-clear,  $P < 0.001$ ) [16]. In patients with primary lung adenocarcinoma, there is indisputable evidence for intratumoral heterogeneity on routinely obtained diagnostic CT images. For example, shape complexity can be captured by convexity and density variation can be captured by entropy ratio [17]. As for ovarian cancer, US-based radiomics have achieved good diagnostic performance in the differentiation between type I and type II ovarian cancers [18, 19]. Specifically, type I ovarian cancer includes MC, CC, EN, LGSC, and malignant Brenner tumour, while type II ovarian cancer comprises HGSC, carcinosarcoma, and undifferentiated carcinoma, according to the World Health Organization classification of tumours of female reproductive organs published in 2014 [20]. To our knowledge, no studies have been conducted to distinguish specific subtypes of ovarian cancer by far, whether radiomics based on US, CT, or magnetic resonance imaging (MRI).

This research developed and validated US-based radiomics models to distinguish specific subtypes of ovarian cancer. In addition, we assessed the consistency between the prediction models and the actuality through calibration curves and evaluated the clinical value of the models based on calculating the net benefit (NB) at different threshold probabilities by decision curve analysis (DCA).

## Methods

### Patients

This retrospective study was conducted using consecutive data from Shengjing Hospital of China Medical University from January 2018 to August 2022. It was conducted in accordance with the Declaration of Helsinki and approved by the ethics committee of Shengjing hospital. Clinical trial number: not applicable. The requirement for patients' informed consent was waived owing

to the retrospective study design. The inclusion criteria were as follows: (1) patients with pathological confirmation of ovarian cancer after biopsy or surgery and histological subtypes of ovarian cancer identified; and (2) patients whose US and serum tumour biomarker examination were performed within 14 days before biopsy or surgery. The exclusion criteria were as follows: (1) no available US images ( $n=129$ ); (2) poor images quality ( $n=40$ ); (3) not detected by US ( $n=8$ ); (4) no available serum tumour biomarkers ( $n=33$ ); (5) with a history of chemotherapy or radiotherapy ( $n=15$ ); (6) accompanied with other tumour ( $n=40$ ); and (7) pregnant ( $n=3$ ). All histopathological findings were confirmed by pathologists with more than 10 years of experience in ovarian pathology. The gold standard references of this study were based on the results of histopathological findings. Finally, 1209 patients enrolled in this study.

There were two subjects in this study: all patients ( $n=1209$ ) and patients with the five major histological subtypes ( $n=1039$ ). Eligible patients were randomly divided into a training cohort ( $n=846$ ) and a validation cohort ( $n=363$ ) in a 7:3 ratio for all patients, and a training cohort ( $n=727$ ) and a validation cohort ( $n=312$ ) in a 7:3 ratio for patients with the five major histological subtypes. Histological subtypes of all patients include: HGSC, MC, CC, EN, LGSC, and others. The five major histological subtypes include: HGSC, MC, CC, EN, and LGSC. The study of histological subtypes was in the form of binary classification.

**US images acquisition and characteristics collection**

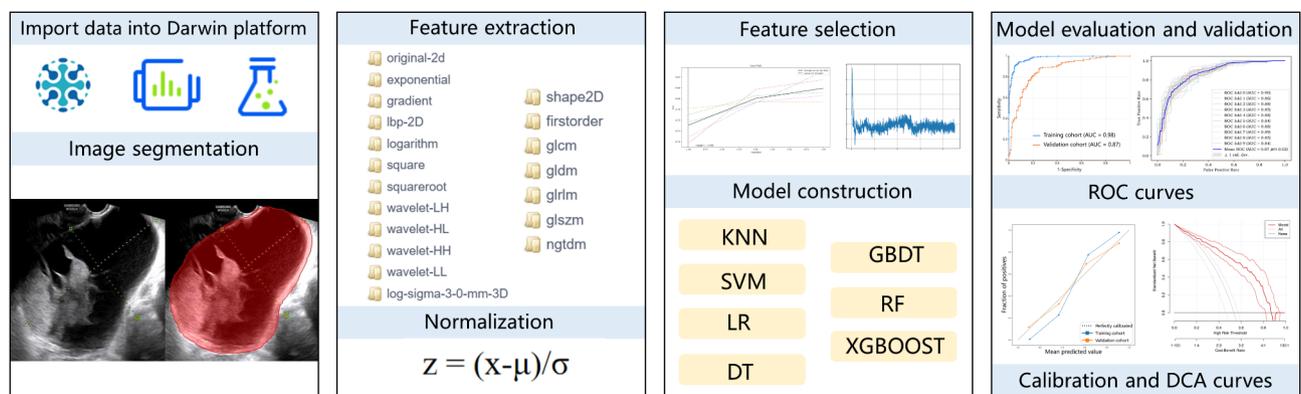
US examinations of ovarian adnexal masses were performed using Philips iU22 (Netherlands), Philips EPIQ 5 (Netherlands), Toshiba Aplio 400 (Japan), Toshiba Aplio 500 (Japan), GE LOGIQ E9 (USA), and SuperSonic Imagine Aixplorer (France). Trans-vaginal ultrasonography was preferred, with trans-abdominal ultrasonography used if the patient had no sexual history or if the mass was huge. When multiple masses were present, the most

complex one was analysed. All images were recorded in the Picture Archiving and Communication System (PACS) in BMP format.

Our selection of parameters is based on guideline recommendations and current research evidences. According to NCCN clinical practice guidelines in oncology [21], tumour biomarkers such as CA-125, CEA and CA-19-9 are recommended. HE4 is considered the most promising ovarian cancer tumour biomarker, which is not expressed in normal ovarian epithelial cells, but is expressed 100% in endometrioid epithelial ovarian cancer and 93% in serous epithelial ovarian cancer [22]. CA-72-4 is now widely used in the detection of gastric and ovarian cancers [23–25]. AFP is a useful tumour biomarker, because it is elevated in almost 100% of malignant yolk sac tumour, and may also be present in other germ cell tumour [26], which may suggesting that the tumour originated from germ cells rather than epithelial cells. In addition, our hospital is a general hospital with excellent medical resources to provide the above-mentioned ovarian tumour biomarkers for patients with suspected ovarian cancer, so as to facilitate the early detection and treatment of the disease. Clinical characteristics including age, histological subtypes, alpha-fetoprotein (AFP), carcinoembryonic antigen (CEA), carbohydrate antigen 125 (CA-125), carbohydrate antigen 19–9 (CA-19-9); carbohydrate antigen 72–4 (CA-72-4), and human epididymis protein 4 (HE4) levels were collected from the Hospital Information System (HIS).

**Image segmentation**

We chose gray scale US images showing the plane with maximal dimension and its orthogonal plane (two images per patient). Use PyCharm to convert BMP format into JPEG format. Then, the patient’s images and clinical data were imported into the Darwin Scientific Research Platform (Beijing Yizhun Intelligent Technology Co., China) and used the platform to delineate the Region of interests (ROIs). The work flow is show Fig. 1. The ROIs were



**Fig. 1** The workflow of this study

manually delineated by a junior radiologist with five years of experience and reviewed by a senior radiologist with more than 20 years of experience. Both radiologists were blinded to patients' information. To ensure reproducibility and accuracy, discrepancies were solved through consultation and resegmented the image until a consensus was reached.

### Feature extraction and normalization

The platform mentioned above was used to extract radiomics features. The original features contain first-order, shape, and texture features extracted from original images. Furthermore, eight filters, including exponential, gradient, local binary pattern, logarithm, square, square root, wavelet, and Laplacian of Gaussian, were used to generate transformed images. Except shape features, both first-order and texture features can be extracted from transformed images. Texture features can describe the heterogeneity of the tumours, including gray level co-occurrence matrix (GLCM), gray level dependence matrix (GLDM), gray level run-length matrix (GLRLM), gray level size zone matrix (GLSZM), and neighbouring gray tone difference matrix (NGTDM). Z-score normalization was used to reduce potential effects related to various parameters. The formula was as follows:  $z = (x - \mu) / \sigma$ , where  $x$  refers to the original value,  $\mu$  refers to the mean value, and  $\sigma$  refers to the standard deviation.

### Feature selection and model construction

To reduce computational complexity and improve classification accuracy, nine feature selection methods, including variance threshold filter, optimal feature filter (number), optimal feature filter (percentage), selecting based on saliency, selecting from model, recursive elimination, stability selection, minimum redundancy maximum relevance (MRMR), and fast correlation-based filter were used to select the optimal predictive features from all features. Seven classifiers were carried out to construct models, including K-nearest neighbor (KNN), support vector machine (SVM), logistic regression (LR), decision tree (DT), gradient boosting decision tree (GBDT), random forest (RF), and extreme gradient boosting (XGBOOST). The nine feature selection methods and the seven classifiers were combined to make a total of 63 combinations. Choose the combination with the best predictive performance as the final result.

### Model evaluation and validation

The predictive performance of the models was further tested in the internal validation cohort using the same thresholds determined in the training cohort. The receiver operating characteristic (ROC) curves were plotted. The area under the curve (AUC), sensitivity, specificity, and accuracy of the models were calculated.

Calibration curves were applied to assess the consistency between the prediction models and the actuality, both in the training and the validation cohorts. DCA was carried out to evaluate the clinical value of the models based on calculating the net benefit at different threshold probabilities. The 10-fold cross-validation method was used to verify the predictive performance of the models.

### Statistical analysis

R software version 4.1.3 and IBM SPSS Statistics version 26.0 were used for analysis. The counting data were expressed in  $n$  (%). The metrological data of this study were shown as skewed distribution by Kolmogorov-Smirnov tests, and were displayed as  $M$  (Q1, Q3). Chi-square tests were used to analyze the categorical variables. Mann-Whitney  $U$  tests were used to analyze the continuous variables. A two-sided  $P$  value  $< 0.05$  was considered statistically significant.

## Results

### Clinical and histopathological characteristics

The results of clinical and histopathological characteristics in the group of all patients and the group of patients with the five major histological subtypes are shown in Table 1. The number of patients included in the two groups were 1209 and 1039, respectively. The most common subtypes in the five major histological subtypes were HGSC, with a number of 645, followed by CC with 133, EN with 110, MC with 79, and LGSC with 72 in this study. The median age in both groups was 54 years. Patients with the five major histological subtypes had higher expression of serum tumour biomarkers than all patients in terms of CEA, CA-125, CA-19-9, CA-72-4, and HE4, except AFP. The comparisons of clinical characteristics among different subtypes are shown in additional file 1. All clinical characteristics were significantly different between HGSC and non-HGSC ( $P < 0.05$ ), except AFP in the group of patients with the five major histological subtypes ( $P > 0.05$ ). All clinical characteristics were significantly different between MC and non-MC ( $P < 0.05$ ), except AFP and CA-72-4 in the two groups ( $P > 0.20$ ). All clinical characteristics were significantly different between CC and non-CC ( $P < 0.05$ ), except age and CEA in the two groups ( $P > 0.05$ ). Age, CEA, and CA-19-9 differed significantly between EN and non-EN in the two groups ( $P < 0.05$ ). There were significant differences between LGSC and non-LGSC in terms of age and CA-72-4 in the two groups and HE4 in the group of patients with the five major histological subtypes ( $P < 0.05$ ).

### Predictive performance of the models

The distribution of subtypes was rational between the training cohort and the validation cohort, both in the

**Table 1** Clinical and histopathological characteristics

Characteristics	Normal range	Expressed	All patients (n = 1209)	Patients with the five major histological subtypes (n = 1039)
Age	-	M (Q1, Q3)	54.00 (47.00, 61.00)	54.00 (48.00, 61.00)
AFP	0–9 ng/mL	M (Q1, Q3)	3.06 (1.95, 4.13)	2.94 (1.90, 3.98)
CEA	0–5 ng/mL	M (Q1, Q3)	1.39 (0.88, 2.28)	1.41 (0.89, 2.30)
CA-125	0–35 U/mL	M (Q1, Q3)	118.10 (34.60, 562.85)	158.30 (44.87, 705.60)
CA-19-9	0–37 U/mL	M (Q1, Q3)	13.52 (7.13, 27.60)	14.39 (7.48, 29.29)
CA-72-4	0–6.9 U/mL	M (Q1, Q3)	5.77 (2.02, 17.82)	6.98 (2.34, 21.83)
HE4	< 140 pmol/L	M (Q1, Q3)	110.50 (59.54, 352.50)	137.70 (68.57, 387.74)
Subtypes	-	n (%)	1209 (100.00%)	1039 (100.00%)
HGSC	-	n (%)	645 (53.35%)	645 (62.08%)
MC	-	n (%)	79 (6.53%)	79 (7.60%)
CC	-	n (%)	133 (11.00%)	133 (12.80%)
EN	-	n (%)	110 (9.10%)	110 (10.59%)
LGSC	-	n (%)	72 (5.96%)	72 (6.93%)
others	-	n (%)	170 (14.06%)	0 (0.00%)

Notes: AFP, alpha-fetoprotein; CEA, carcinoembryonic antigen; CA-125, carbohydrate antigen 125; CA-19-9, carbohydrate antigen 19–9; CA-72-4, carbohydrate antigen 72–4; HE4, human epididymis protein 4; HGSC, high-grade serous carcinoma; MC, mucinous carcinoma; CC, clear cell carcinoma; EN, endometrioid carcinoma; LGSC, low-grade serous carcinoma

**Table 2** The distribution of subtypes between the training cohort and the validation cohort

Group	All patients (n = 1209)		P	Patients with the five major histological subtypes (n = 1039)		P
	Training cohort	Validation cohort		Training cohort	Validation cohort	
HGSC	451 (53.31%)	194 (53.44%)	0.966	451 (62.04%)	194 (62.18%)	0.965
Non-HGSC	395 (46.69%)	169 (46.56%)		276 (37.96%)	118 (37.82%)	
MC	55 (6.50%)	24 (6.61%)	0.943	55 (7.57%)	24 (7.69%)	0.944
Non-MC	791 (93.50%)	339 (93.39%)		672 (92.43%)	288 (92.31%)	
CC	93 (10.99%)	40 (11.02%)	0.989	93 (12.79%)	40 (12.82%)	0.990
Non-CC	753 (89.01%)	323 (88.98%)		634 (87.21%)	272 (87.18%)	
EN	77 (9.10%)	33 (9.09%)	0.995	77 (10.59%)	33 (10.58%)	0.994
Non-EN	769 (90.90%)	330 (90.91%)		650 (89.41%)	279 (89.42%)	
LGSC	50 (5.91%)	22 (6.06%)	0.919	50 (6.88%)	22 (7.05%)	0.920
Non-LGSC	796 (94.09%)	341 (93.94%)		677 (93.12%)	290 (92.95%)	

Notes: HGSC, high-grade serous carcinoma; MC, mucinous carcinoma; CC, clear cell carcinoma; EN, endometrioid carcinoma; LGSC, low-grade serous carcinoma

group of all patients and the group of patients with the five major histological subtypes (all  $P > 0.20$ ), as is shown in Table 2. Construction and the predictive performance of the models is shown in Table 3. As for LGSC, in the training cohort, the combination of stability selection and LR yielded AUCs of 0.71 and 0.76, respectively, in the two groups, showing a poor predictive performance. As for EN, models showed AUCs of 0.83 and 0.80, respectively, in the training cohort, but showed AUCs of 0.67 and 0.69, respectively, in the validation cohort. As for CC, the AUCs were 0.80 and above both in the training cohort and the validation cohort, with relatively lower AUCs of 0.76 and 0.77, respectively, in the 10-fold cross-validation. Furthermore, the specificity of the group of all patients was only 0.70 and the sensitivity of the group of patients with the five major histological subtypes was only 0.70 in the validation cohort, showing an unbalanced diagnostic efficiency. As for MC and HGSC, the AUCs in the training cohort were 0.9 and above, and in the validation cohort and 10-fold cross-validation were

0.8 or above, with a relatively higher sensitivity, specificity, and accuracy. Moreover, the combination of the feature selection method and the classifier was the same which was recursive elimination and GBDT that could distinct HGSC and non-HGSC effectively in the two groups. ROC curves of the models in the training cohort and the validation cohort are shown in Fig. 2 and ROC curves in the 10-fold cross-validation are shown in Fig. 3.

**Calibration and DCA of the models**

Calibration curves of the models in the training cohort and the validation cohort are depicted in Fig. 4. Calibration curves displayed the consistency between the prediction models and the actuality in discriminating HGSC and non-HGSC was favorable, both in the group of all patients (Fig. 4A) and the group of patients with the five major histological subtypes (Fig. 4C). Compared with the calibration of the model in discriminating HGSC and non-HGSC, the calibration of the model in discriminating MC and non-MC in the group of all patients (Fig. 4B)

**Table 3** Construction and the predictive performance of the models

group	subtypes	feature selection methods	classifiers	training cohort			validation cohort			10-fold cross-validation		
				SEN	SPE	ACC	AUC	SEN	SPE	ACC	AUC	AUC
1	HGSC	recursive elimination	GBDT	0.91	0.95	0.93	0.98	0.88	0.75	0.82	0.87	0.87
2	HGSC	recursive elimination	GBDT	0.94	0.93	0.94	0.98	0.74	0.84	0.78	0.85	0.84
1	MC	minimum redundancy maximum relevance	SVM	0.80	0.97	0.96	0.93	0.75	0.89	0.88	0.83	0.80
2	MC	stability selection	LR	0.86	0.80	0.81	0.90	0.92	0.80	0.81	0.89	0.86
1	CC	optimal feature filter (percentage)	LR	0.77	0.77	0.77	0.85	0.83	0.70	0.71	0.80	0.76
2	CC	variance threshold filter	LR	0.84	0.83	0.84	0.90	0.70	0.82	0.80	0.80	0.77
1	EN	optimal feature filter (percentage)	LR	0.79	0.74	0.74	0.83	0.55	0.78	0.76	0.67	0.67
2	EN	stability selection	LR	0.73	0.76	0.75	0.80	0.73	0.68	0.68	0.69	0.70
1	LGSC	stability selection	LR	0.62	0.71	0.70	0.71	0.64	0.66	0.66	0.63	0.64
2	LGSC	stability selection	LR	0.64	0.78	0.77	0.76	0.77	0.67	0.68	0.72	0.70

Notes: 1, all patients; 2, patients with the five major histological subtypes; HGSC, high-grade serous carcinoma; MC, mucinous carcinoma; CC, clear cell carcinoma; EN, endometrioid carcinoma; LGSC, low-grade serous carcinoma; GBDT, gradient boosting decision tree; SVM, support vector machine; LR, logistic regression; SEN, sensitivity; SPE, specificity; ACC, accuracy; AUC, area under the curve

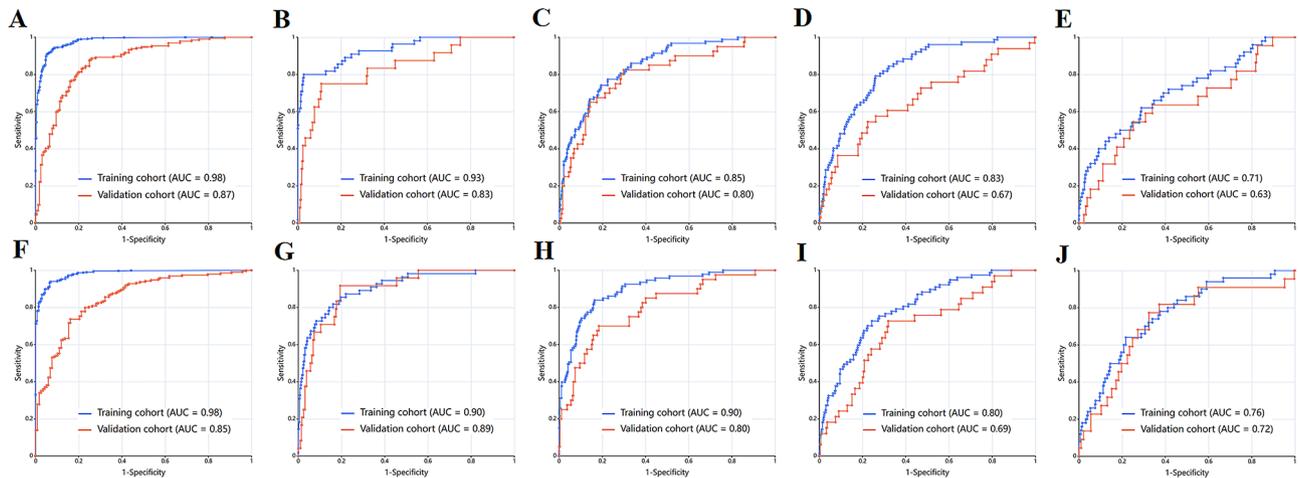
and the group of patients with the five major histological subtypes (Fig. 4D) demonstrated poor agreement. DCA curves of the models in the training cohort and the validation cohort are presented in Fig. 5. The model in discriminating HGSC and non-HGSC could bring more net benefits than the model in discriminating MC and non-MC both in the two groups.

**Feature importance**

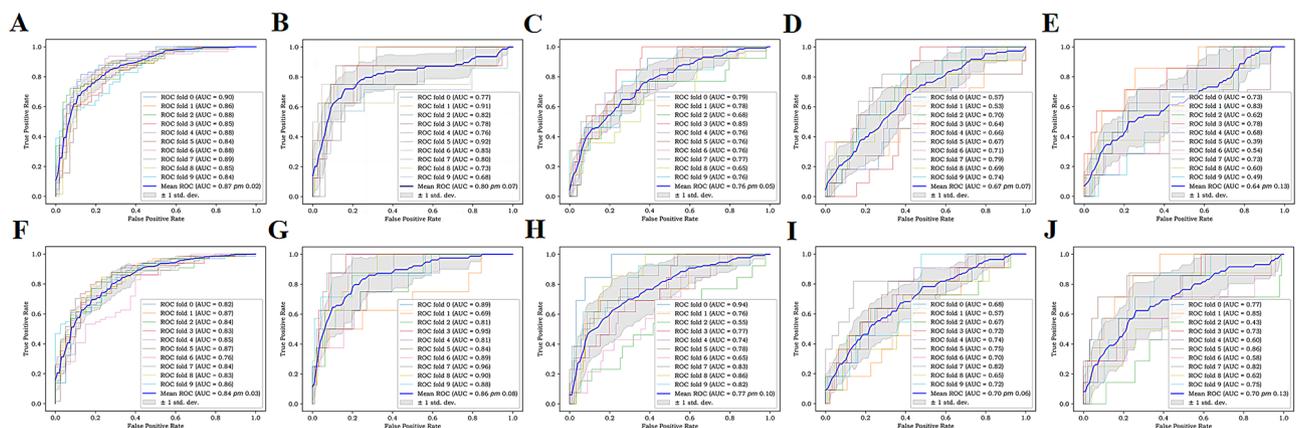
The importance of the features in the models in discriminating HGSC and non-HGSC are shown in Fig. 6. Of the 2250 radiomics features and 7 clinical features, 7 potential predictors that were most relevant to the prediction of HGSC and non-HGSC were selected based on recursive elimination. The radiomics features included Lbp-2D\_glrIm\_LongRunHighGrayLevelEmphasis, Lbp-2D\_glrIm\_RunVariance, and Lbp-2D\_glrIm\_LongRunEmphasis and clinical features included HE4, CA-19-9, AFP, and CA-125 in the group of all patients (Fig. 6A). The radiomics features included original\_shape2D\_MaximumDiameter, original\_shape2D\_Sphericity, exponential\_glszm\_SizeZoneNonUniformity, and exponential\_gldm\_SmallDependenceHighGrayLevelEmphasis and clinical features included HE4, CA-19-9, and AFP in the group of patients with the five major histological subtypes (Fig. 6B).

**Discussion**

There were two subjects in this study: all patients ( $n=1209$ ) and patients with the five major histological subtypes ( $n=1039$ ). The results showed that the conclusions of the two subjects were consistent. As for LGSC, the AUC in the training cohort is low. As for EN, the AUC in the training cohort is 0.8 or higher, but the AUC in the validation cohort is low. The diagnostic efficiency of the training and validation cohort is good for the discrimination of CC, MC, and HGSC. However, the results of 10-fold cross-validation showed that the AUC of the model for the discrimination of CC was below 0.8. Finally, considering the diagnostic efficiency, the model for the discrimination of MC and HGSC is relatively successful. Further considering the consistency between the prediction models and reality, as well as the NBs brought to patients, the performance of MC is poor, while HGSC is excellent in comparison. Therefore, after comprehensive analysis, this study concludes that the model for discriminating HGSC is successful. The model used to distinguish between HGSC and non-HGSC achieved AUCs of 0.98 and 0.87 in the training and validation cohorts, respectively, in the group of all patients, and the model achieved AUCs of 0.98 and 0.85 in the training and validation cohorts, respectively, in the group of patients with the five major histological subtypes. These results are superior to those reported by Wang et al.



**Fig. 2** ROC curves of the models in the training cohort and the validation cohort. Notes: **A-E**, the subject is all patients; **F-J**, the subject is patients with the five major histological subtypes. **A, F**: ROC curves of the models in discriminating HGSC and non-HGSC. **B, G**: ROC curves of the models in discriminating MC and non-MC. **C, H**: ROC curves of the models in discriminating CC and non-CC. **D, J**: ROC curves of the models in discriminating EN and non-EN. **E, J**: ROC curves of the models in discriminating LGSC and non-LGSC. HGSC, high-grade serous carcinoma; MC, mucinous carcinoma; CC, clear cell carcinoma; EN, endometrioid carcinoma; LGSC, low-grade serous carcinoma

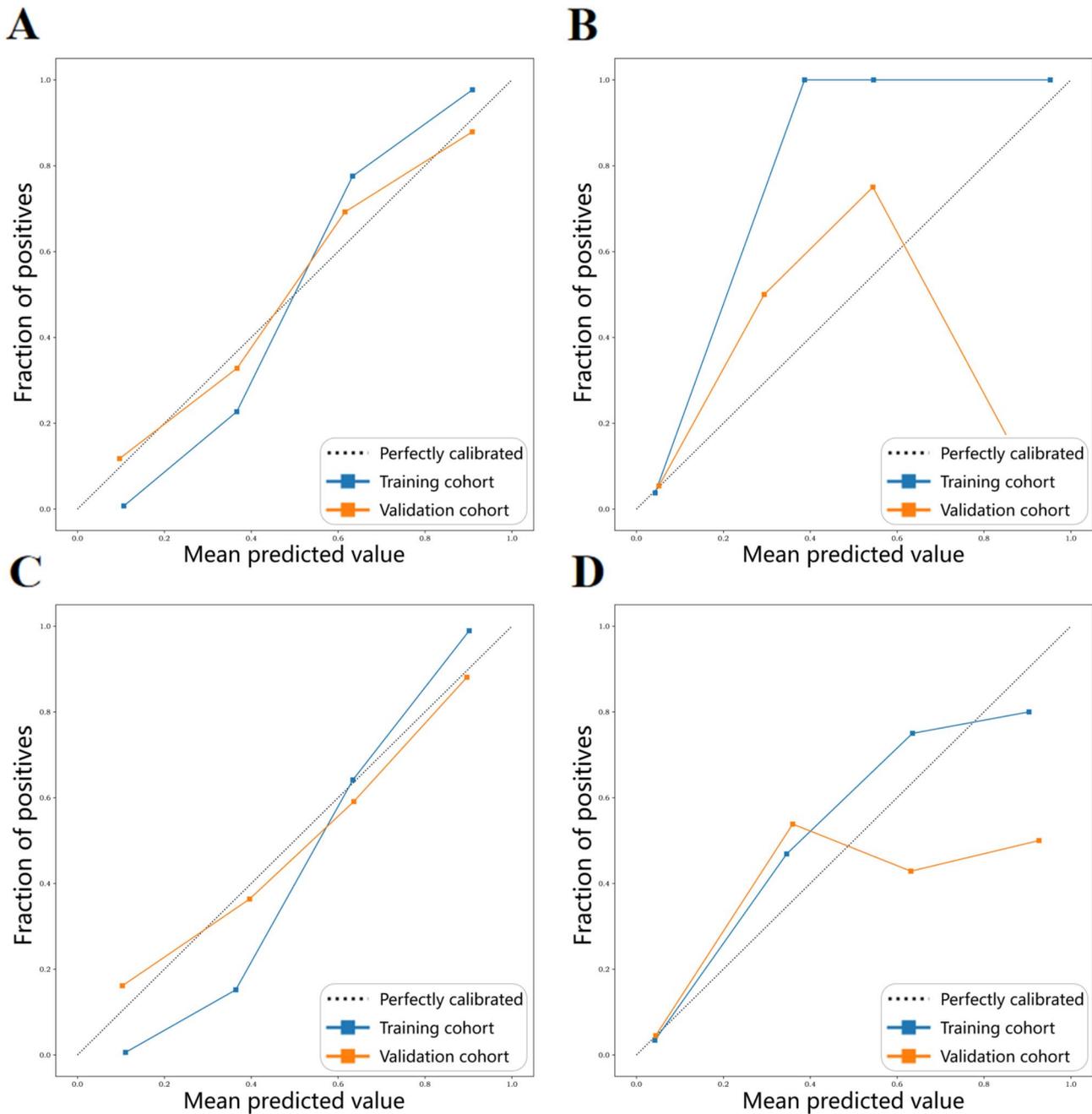


**Fig. 3** ROC curves of the models in the 10-fold cross-validation. Notes: **A-E**, the subject is all patients; **F-J**, the subject is patients with the five major histological subtypes. **A, F**: ROC curves of the models in discriminating HGSC and non-HGSC. **B, G**: ROC curves of the models in discriminating MC and non-MC. **C, H**: ROC curves of the models in discriminating CC and non-CC. **D, J**: ROC curves of the models in discriminating EN and non-EN. **E, J**: ROC curves of the models in discriminating LGSC and non-LGSC. HGSC, high-grade serous carcinoma; MC, mucinous carcinoma; CC, clear cell carcinoma; EN, endometrioid carcinoma; LGSC, low-grade serous carcinoma

[27], whose logistic regression model based on contrast-enhanced CT in discriminating HGSC and non-HGSC achieved AUCs of 0.837 and 0.836 for the training and testing cohorts, respectively. Currently, there have been many studies reporting the use of machine learning to distinguish between type I and type II ovarian cancers has achieved satisfactory outcomes. Zhang, H. et al. [28] retrospectively analyzed preoperative MRI images from 286 patients with pathologically proven ovarian tumours to distinguish type I and type II EOC and obtained satisfactory results with an accuracy of 93% in the leave-one-out cross-validation cohort and 84% in the independent validation cohort. Jian, J. et al. [29] conducted a MR image-based radiomics multicenter study

for differentiating between type I and type II EOC. The results exhibited advantages of utilizing multi-parametric MRI and the combined model performed well both in the internal and external validation cohorts with AUCs of 0.806 and 0.847, respectively. Tang, Z. P. et al. [18] conducted a study about ultrasound-based radiomics for differentiating between type I and type II EOC. The AUCs of the training cohort and testing cohort in the radiomics model and comprehensive model were 0.817 and 0.731 and 0.982 and 0.886, respectively.

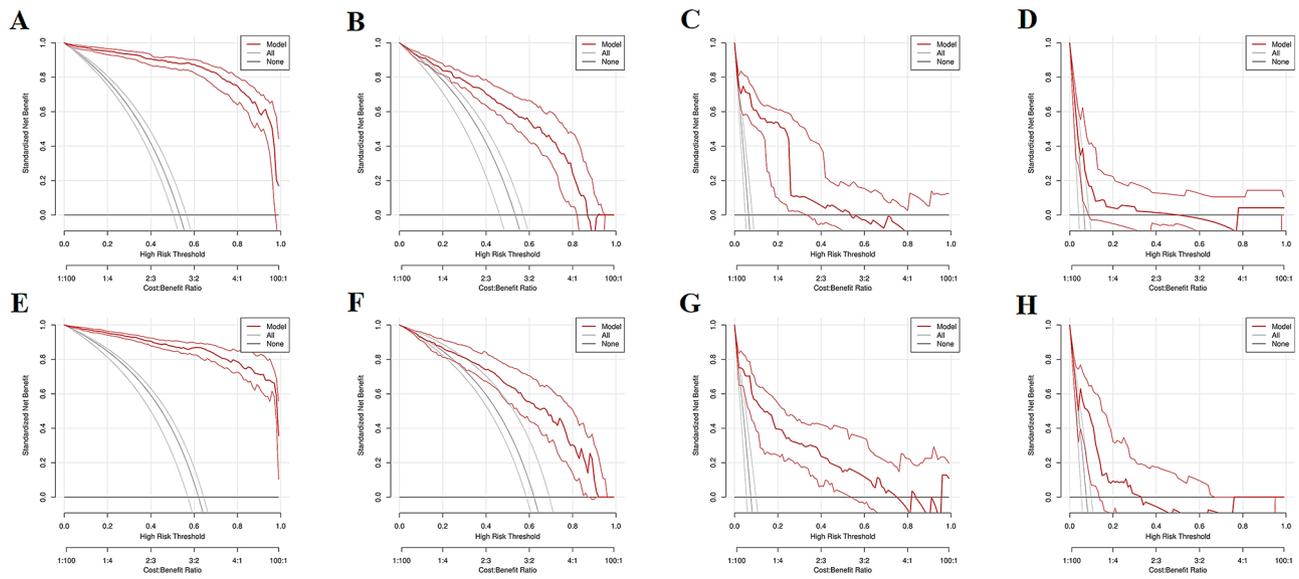
To the best of our knowledge, there is no machine learning research to distinguish specific subtypes of ovarian cancer based on radiomics, no matter whether US, CT, or MRI. But there have been many deep learning



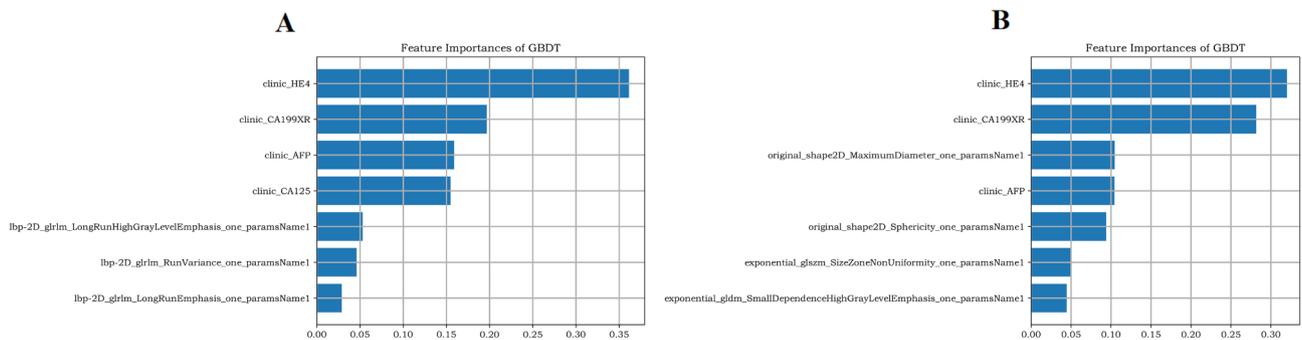
**Fig. 4** Calibration curves of the models in the training cohort and the validation cohort. Notes: **A, B:** the subject is all patients. **C, D:** the subject is patients with the five major histological subtypes. **A, C:** Calibration curves of the models in discriminating HGSC and non-HGSC. **B, D:** Calibration curves of the models in discriminating MC and non-MC. HGSC, high-grade serous carcinoma; MC, mucinous carcinoma

studies aiming to differentiate specific subtypes of ovarian cancer. Some studies focused on US, CT or MRI images, while others concentrated on Hematoxylin-Eosin stained whole-slide pathology images. For examples, the results of a study by Jung, Y. et al. [30] showed a convolutional neural network with a convolutional autoencoder to remove disturbances such as calipers and annotations on the ultrasound images, which cannot be removed manually, and generate new pixels which are compared

with the surrounding pixels without a sense of heterogeneity. Then, the convolutional neural network sorts ovarian neoplasms into five classes: normal, cystadenoma, mature cystic teratoma, endometrioma, and malignant tumor. A structured support vector machine formulation is defined and used to combine the salient slide images information from multiple magnifications while simultaneously operating within the latent variable framework [31]. It can automatically recognized HGSC, CC and MC



**Fig. 5** DCA curves of the models in the training cohort and the validation cohort. Notes: **A-D**: the subject is all patients; **E-H**: the subject is patients with the five major histological subtypes. **A, B, E, F**: DCA curves of the models in discriminating HGSC and non-HGSC. **C, D, G, H**: DCA curves of the models in discriminating MC and non-MC. **A, C, E, G**: DCA curves of the models in the training cohort. **B, D, F, H**: DCA curves of the models in the validation cohort. HGSC, high-grade serous carcinoma; MC, mucinous carcinoma



**Fig. 6** The importance of the features in the models. Notes: **A**: The importance of the features in the model in discriminating HGSC and non-HGSC in all patients. **B**: The importance of the features in the model in discriminating HGSC and non-HGSC in the patients with the five major histological subtypes. HGSC, high-grade serous carcinoma

cases but misclassified EN and LGSC as HGSC which is the most prevalent subtype, achieving an average multi-class classification accuracy of 90%. BenTaieb, A. et al. observed that salient regions from CC cases often contain papillary-looking areas while salient patches from MC cases often show tissue foldings. Wu, M. et al. [32] employed a Deep Convolutional Neural Networks (DCNN) which consists of five convolutional layers, three max pooling layers, and two full reconnect layers to automatically classify the different subtypes (serous carcinoma, mucinous carcinoma, endometrioid carcinoma, and clear cell carcinoma) of ovarian cancers through the Hematoxylin-Eosin stained tissue sections at the cytological level. The testing results are obtained by the method of 10-fold cross-validation, showing that the accuracy of the classification models has reached to 78.20% by using augmented images as training data. Farahani, H. et al. [33]

trained four different artificial intelligence algorithms based on convolutional neural networks to automatically classify histopathology subtypes of ovarian tumors according to the Hematoxylin-Eosin stained whole-slide pathology images. The best-performing model achieved a concordance of 81.38% and Cohen’s kappa of 0.7378 on the training dataset, and a concordance of 80.97% and Cohen’s kappa of 0.7547 on the external dataset.

The model for discriminating between HGSC and non-HGSC is composed of a feature selection method using recursive elimination and a GBDT classifier. Recursive elimination is a commonly used feature selection method that reduces the risk of model overfitting and improves the model’s generalization ability by iteratively removing unimportant features [34]. GBDT is an ensemble learning method based on decision trees, which builds a strong classifier by iteratively training weak classifiers

and combining them. GBDT can capture complex non-linear relationships in datasets and has high accuracy in classification and regression tasks. It is also robust to outliers and noise in the dataset, making it less susceptible to interference [35]. The radiomics features of the model in discriminating HGSC and non-HGSC included two original features and five texture features. Specifically, the original features in this study were shape-based and included MaximumDiameter and Sphericity. This finding could be associated with the different growth patterns. Non-HGSC ovarian cancer is characterized by a slower-growing and less aggressive behavior, typically confined to the ovary. In contrast, HGSC ovarian cancer is known for its highly aggressive nature, resulting in rapid and irregular tumor growth patterns that are less likely to be small and spherical. In a study of Wang, M. et al. [27], the values of sphericity were significantly higher in non-HGSC than HGSC ( $P < 0.001$ ). Similarly, in an MRI radiomics study [29], sphericity was the optimal selected feature for discriminating HGSC and non-HGSC. The reason for excluding first-order features may be that they describe the distribution of values of individual voxels without considering their spatial relationships. In contrast, texture features describe relationships between voxels and may explain the heterogeneity of intra- and intertumours. In this study, texture features, including LongRunHighGrayLevelEmphasis, RunVariance, LongRunEmphasis, SizeZoneNonUniformity, and SmallDependenceHighGrayLevelEmphasis, were extracted from transformed images by two filters, which were exponential and local binary pattern. The clinical features used in the model to differentiate between HGSC and non-HGSC included HE4, CA-19-9, AFP, and CA-125. These results are partly consistent with findings from previous studies by other researchers. For instance, seven supervised machine learning classifiers based on 32 parameters commonly available from peripheral blood tests and age were able to predict HGSC of EOC with an accuracy of 75.8% and an AUC of 0.785 which is manifested as a relatively higher CA-125 and a relatively lower CA-19-9, and could also predict MC of EOC with an accuracy of 96.0% and an AUC of 0.728, showing a relatively higher CEA. The underperformance with regard to CC with an accuracy of 67.7% and an AUC of 0.650 and EN with an accuracy of 55.6% and an AUC of 0.597 may result from the lack of distinct characteristics at the level of serum biomarkers [11]. The results of a study by Yao, F. et al. [19] showed that the clinical model that distinguished between type I and type II EOC including age, CA125, and CA199 suggested that patients with type I were younger than those with type II, with lower levels of CA125 and higher levels of CA199. In the training cohort, the AUCs of the combined model was 0.83. In the testing cohort, the AUC was 0.82.

Although the purpose of this study is to predict the five major histological subtypes of epithelial ovarian cancer, the purpose of the differential diagnosis is to better treat ovarian cancer. When it comes to the treatment of ovarian cancer, one topic that we have to discuss is lymphadenectomy in ovarian cancer. Due to the lack of early screening, more than 80% of women are diagnosed with EOC when it is already at stages III and IV [36, 37]. Lymphatic vessels represent one of the main pathways for the spread of most gynecological malignancies, and patients with advanced ovarian cancer have a high incidence of both pelvic and para-aortic lymph node metastases [38, 39]. Although there is no first level evidence in literature that lymphadenectomy has survival benefits for early ovarian cancer, trials [40, 41] have indicated a significant survival advantage in patients undergoing lymphadenectomy as part of surgical debulking in patients with advanced ovarian cancer. Results from a randomized prospective trial [42] showed that systematic lymphadenectomy significantly improved progression-free survival and reduced recurrence rates, but did not improve overall survival. The lymphadenectomy in ovarian neoplasms trial [43] reports no better outcomes and higher complication and mortality rates associated with lymphadenectomy. Even if performed by expert hands, lymphadenectomy has a cost in terms of longer operative time, blood loss, higher rates of transfusions, and intensive unit care. In the absence of high-level study on nodes, the authoritative guidelines [21, 44] recommended that systematic lymphadenectomy should not be regarded as a standard procedure. Currently, the removal of bulky lymph nodes is carried out as part of an attempt to achieve maximum cytoreduction. Resection of clinically negative nodes is not required.

To the best of our knowledge, this is the first research to utilize machine learning to distinguish specific subtypes of ovarian cancer based on radiomics, no matter whether US, CT, or MRI. In addition to analyzing AUC, sensitivity, specificity, and accuracy of the models, we also assessed the consistency between the prediction models and the actuality through calibration curves and evaluated the clinical value of the models based on calculating the net benefit at different threshold probabilities by DCA. Furthermore, there were two subjects in our study: all patients and patients with the five major histological subtypes, and conclusions drawn from the two subjects were the same, indirectly demonstrating the reliability of the models. However, there were several limitations in our study. First, it included a relatively small number of patients retrospectively collected from a single centre with inevitable selection bias. And this study used 10-fold cross-validation instead of external validation to verify the predictive performance of the models. This is a preliminary exploratory study, and the

model in discriminating HGSC and non-HGSC could be developed using prospectively collected data from multiple centres and externally validated to demonstrate its generalizability and robustness. Second, the delineation of ROI was manually performed, which was time-consuming and prone to error. It could be overcome by an automated segmentation artificial intelligence system. Third, most patients did not undergo genetic testing, so we did not consider this feature of genetics owing to the retrospective study design. However, it is currently known that genes such as *BRCA1/2*, *RAD51C*, *RAD51D*, *BRIP1*, *PALB2*, *ATM*, and Lynch syndrome-related genes (*MLH1*, *MSH2*, *MSH6*, *PMS2*, *EPCAM*) that increase the risk of epithelial ovarian cancers, and *STK11* germline mutations are mainly associated with the development of ovarian Sertoli-Leydig cell tumours [45, 46]. We hope that further research could incorporate genes related to ovarian cancer subtyping.

## Conclusions

This research developed and validated US-based radiomics models to distinguish specific subtypes of ovarian cancer. The results demonstrated it seems difficult for US-based radiomics models to distinguish CC, EN, and LGSC. The diagnostic performances of the models in discriminating HGSC and MC were satisfactory, both in the internal validation and 10-fold cross-validation. Especially in discriminating HGSC, the models exhibited high consistency between the predicted results and the actual outcomes, and brought great net benefits for patients.

## Abbreviations

OC	Ovarian cancer
EOC	Epithelial ovarian cancer
HGSC	High-grade serous carcinoma
MC	Mucinous carcinoma
CC	Clear cell carcinoma
EN	Endometrioid carcinoma
LGSC	Low-grade serous carcinoma
PDS	Primary debulking surgery
IDS	Internal debulking surgery
NACT	Neoadjuvant chemotherapy
US	Ultrasound
CT	Computed tomography
MRI	Magnetic resonance imaging
NB	Net benefit
DCA	Decision curve analysis
PACS	The Picture Archiving and Communication System
AFP	Alpha-fetoprotein
CEA	Carcinoembryonic antigen
CA-125	Carbohydrate antigen 125
CA-19-9	Carbohydrate antigen 19-9
CA-72-4	Carbohydrate antigen 72-4
HE4	Human epididymis protein 4
HIS	Hospital Information System
ROI	Region of interest
GLCM	Gray level co-occurrence matrix
GLDM	Gray level dependence matrix
GLRLM	Gray level run-length matrix
GLSZM	Gray level size zone matrix

NGTDM	Neighbouring gray tone difference matrix
MRMR	Minimum redundancy maximum relevance
KNN	K-nearest neighbor
SVM	Support vector machine
LR	Logistic regression
DT	Decision tree
GBDT	Gradient boosting decision tree
RF	Random forest
XGBOOST	Extreme gradient boosting
ROC	Receiver operating characteristic
AUC	Area under the curve
DCNN	Deep Convolutional Neural Networks

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12880-025-01624-1>.

Supplementary Material 1

## Acknowledgements

Not applicable.

## Author contributions

YY participated in the entire process as the first author, including the conception and design of the study, acquisition of data, analysis and interpretation of data, and drafting the article. XJ spent a lot of time acquiring images. Sen Li is a PhD graduated from the Computer Science department of Tsinghua University. For the methodological part of this study, especially regarding feature extraction methods and classifiers, he ensured the professionalism and correctness of the methodology. XG and YW participated in searching for references. YH played a crucial role as the corresponding author, initiating the research project, supervising the process, providing funding, and revising the manuscript.

## Funding

This study was supported by grants from the Guide Project for Key Research and Development Foundation of Liaoning Province (No. 2019JH8/10300008), 345 Talent Project and Liaoning BaiQianWan Talents Program, the Liaoning Provincial Science and Technology Program Joint Project (Key Research and Development Program, No. 2023JH2/101800018), Xingliao Talents Program-Medical Eminent Experts Project (No. YXMJ-LJ-10), and Shenyang Science and Technology Plan Special Project (Key Technology Research Project for Diagnosis and Treatment of Major Diseases): Research on intelligent management of ultrasound image report (No. 24-214-3-09). It did not involve in the conceptualization, design, data collection, analysis, decision to publish, or preparation of the manuscript.

## Data availability

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

## Declarations

### Ethics approval and consent to participate

The requirement for patients' informed consent was waived owing to the retrospective study design. It was approved by the ethics committee of Shengjing hospital and the committee's reference number was 2023PS1208K.

### Consent for publication

Not applicable.

### Competing interests

The authors declare no competing interests.

Received: 13 January 2024 / Accepted: 3 March 2025

Published online: 15 April 2025

## References

- Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, 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. <https://doi.org/10.3322/caac.21660>.
- Chen W, Zheng R, Baade PD, Zhang S, Zeng H, Bray F, et al. Cancer statistics in China, 2015. *CA cancer J Clin*. 2016;66(2):115–32. <https://doi.org/10.3322/caac.21338>.
- Prat J. Staging classification for Cancer of the ovary, fallopian tube, and peritoneum: abridged republication of guidelines from the international federation of gynecology and obstetrics (FIGO). *FIGO Comm Gynecologic Oncol*. 2015;126(1):171–4. <https://doi.org/10.1097/AOG.0000000000000917>.
- Makar AP, Baekelandt M, Tropé CG, Kristensen GB. The prognostic significance of residual disease, FIGO substage, tumor histology, and grade in patients with FIGO stage III ovarian cancer. *Gynecol Oncol*. 1995;56(2):175–80. <https://doi.org/10.1006/gyno.1995.1027>.
- Lheureux S, Gourley C, Vergote I, Oza AM. Epithelial ovarian cancer. *Lancet*. 2019;393(10177):1240–53. [https://doi.org/10.1016/S0140-6736\(18\)32552-2](https://doi.org/10.1016/S0140-6736(18)32552-2).
- Chiofalo B, Bruni S, Certelli C, Sperduti I, Baiocco E, Vizza E. Primary debulking surgery vs. interval debulking surgery for advanced ovarian cancer: review of the literature and meta-analysis. *Minerva Med*. 2019;110(4):330–40. <https://doi.org/10.23736/S0026-4806.19.06078-6>.
- Landen CN Jr., Birrer MJ, Sood AK. Early events in the pathogenesis of epithelial ovarian cancer. *J Clin Oncol*. 2008;26(6):995–1005. <https://doi.org/10.1200/JCO.2006.07.9970>.
- Kurman RJ, Shih le M. Pathogenesis of ovarian cancer: lessons from morphology and molecular biology and their clinical implications. *Int J Gynecol Pathol*. 2008;27(2):151–60. <https://doi.org/10.1097/PGP.0b013e318161e4f5>.
- Hack K, Gandhi N, Bouchard-Fortier G, Chawla TP, Ferguson SE, Li S, et al. External validation of O-RADS US risk stratification and management system. *Radiology*. 2022;304(1):114–20. <https://doi.org/10.1148/radiol.211868>.
- Hidalgo JJ, Ros F, Aubá M, Errasti T, Olartecoechea B, Ruiz-Zambrana Á, et al. Prospective external validation of IOTA three-step strategy for characterizing and classifying adnexal masses and retrospective assessment of alternative two-step strategy using simple-rules risk. *Ultrasound Obstet Gynecol*. 2019;53(5):693–700. <https://doi.org/10.1002/uog.20163>.
- Kawakami E, Tabata J, Yanaihana N, Ishikawa T, Koseki K, Iida Y, et al. Application of artificial intelligence for preoperative diagnostic and prognostic prediction in epithelial ovarian Cancer based on blood biomarkers. *Clin Cancer Res*. 2019;25(10):3006–15. <https://doi.org/10.1158/1078-0432.CCR-18-3378>.
- Barnes BM, Nelson L, Tighe A, Burghel GJ, Lin IH, Desai S, et al. Distinct transcriptional programs stratify ovarian cancer cell lines into the five major histological subtypes. *Genome Med*. 2021;13(1):140. <https://doi.org/10.1186/s13073-021-00952-5>.
- Klein O, Kanter F, Kulbe H, Jank P, Denkert C, Nebrich G, et al. MALDI-Imaging for classification of epithelial ovarian Cancer histotypes from a tissue microarray using machine learning methods. *Proteom Clin Appl*. 2019;13(1):e1700181. <https://doi.org/10.1002/prca.201700181>.
- Nougaret S, Tardieu M, Vargas HA, Reinhold C, Vande Perre S, Bonanno N, et al. Ovarian cancer: an update on imaging in the era of radiomics. *Diagn Interv Imaging*. 2019;100(10):647–55. <https://doi.org/10.1016/j.diii.2018.11.007>.
- Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. *Radiology*. 2016;278(2):563–77. <https://doi.org/10.1148/radiol.201511169>.
- Scrima AT, Lubner MG, Abel EJ, Havighurst TC, Shapiro DD, Huang W, et al. Texture analysis of small renal cell carcinomas at MDCT for predicting relevant histologic and protein biomarkers. *Abdom Radiol*. 2019;44(6):1999–2008. <https://doi.org/10.1007/s00261-018-1649-2>.
- Grove O, Berglund AE, Schabath MB, Aerts HJ, Dekker A, Wang H, et al. Quantitative computed tomographic descriptors associate tumor shape complexity and intratumor heterogeneity with prognosis in lung adenocarcinoma. *PLoS ONE*. 2015;10(3):e0118261. <https://doi.org/10.1371/journal.pone.0118261>.
- Tang ZP, Ma Z, He Y, Liu RC, Jin BB, Wen DY, et al. Ultrasound-based radiomics for predicting different pathological subtypes of epithelial ovarian cancer before surgery. *BMC Med Imaging*. 2022;22(1):147. <https://doi.org/10.1186/s12880-022-00879-2>.
- Yao F, Ding J, Lin F, Xu X, Jiang Q, Zhang L, et al. Nomogram based on ultrasound radiomics score and clinical variables for predicting histologic subtypes of epithelial ovarian cancer. *Br J Radiol*. 2022;95(1136):20211332. <https://doi.org/10.1259/bjr.20211332>.
- Kurman RJ, Carcangiu ML, Herrington CS, Young RH. WHO classification of tumours of female reproductive organs. WHO classification of tumours. 4th ed. Lyon: WHO; 2014.
- Armstrong DK, Alvarez RD, Bakkum-Gamez JN, Barroillet L, Behbakht K, Berchuck A, et al. Ovarian cancer, version 2.2020, NCCN clinical practice guidelines in oncology. *J Natl Compr Canc Netw*. 2021;19(2):191–226. <https://doi.org/10.6004/jnccn.2021.0007>.
- Drapkin R, von Horsten HH, Lin Y, Mok SC, Crum CP, Welch WR, et al. Human epididymis protein 4 (HE4) is a secreted glycoprotein that is overexpressed by serous and endometrioid ovarian carcinomas. *Cancer Res*. 2005;65(6):2162–9. <https://doi.org/10.1158/0008-5472.CAN-04-3924>.
- Xu Y, Zhang P, Zhang K, Huang C. The application of CA72-4 in the diagnosis, prognosis, and treatment of gastric cancer. *Biochim Biophys Acta Rev Cancer*. 2021;1876(2):188634. <https://doi.org/10.1016/j.bbcan.2021.188634>.
- Margoni A, Gargalionis AN, Papavassiliou AG. CA-125: CA72-4 ratio-towards a promising cost-effective tool in ovarian cancer diagnosis and monitoring of post-menopausal women under hormone treatment. *J Ovarian Res*. 2024;17(1):164. <https://doi.org/10.1186/s13048-024-01487-0>.
- Landolfo C, Achten ETL, Ceusters J, Baert T, Froyman W, Heremans R, et al. Assessment of protein biomarkers for preoperative differential diagnosis between benign and malignant ovarian tumors. *Gynecol Oncol*. 2020;159(3):811–9. <https://doi.org/10.1016/j.ygyyno.2020.09.025>.
- Anfelter P, Testa A, Chiappa V, Froyman W, Fruscio R, Guerriero S, et al. Imaging in gynecological disease (17): ultrasound features of malignant ovarian yolk sac tumors (endodermal sinus tumors). *Ultrasound Obstet Gynecol*. 2020;56(2):276–84. <https://doi.org/10.1002/uog.22002>.
- Wang M, Perucho JAU, Hu Y, Choi MH, Han L, Wong EMF, et al. Computed tomographic radiomics in differentiating histologic subtypes of epithelial ovarian carcinoma. *JAMA Netw Open*. 2022;5(12):e2245141. <https://doi.org/10.1001/jamanetworkopen.2022.45141>.
- Zhang H, Mao Y, Chen X, Wu G, Liu X, Zhang P, et al. Magnetic resonance imaging radiomics in categorizing ovarian masses and predicting clinical outcome: a preliminary study. *Eur Radiol*. 2019;29(7):3358–71. <https://doi.org/10.1007/s00330-019-06124-9>.
- Jian J, Li Y, Pickhardt PJ, Xia W, He Z, Zhang R, et al. MR image-based radiomics to differentiate type I and type II epithelial ovarian cancers. *Eur Radiol*. 2021;31(1):403–10. <https://doi.org/10.1007/s00330-020-07091-2>.
- Jung Y, Kim T, Han MR, Kim S, Kim G, Lee S, et al. Ovarian tumor diagnosis using deep convolutional neural networks and a denoising convolutional autoencoder. *Sci Rep*. 2022;12(1):17024. <https://doi.org/10.1038/s41598-022-20653-2>.
- BenTaieb A, Li-Chang H, Huntsman D, Hamarneh G. A structured latent model for ovarian carcinoma subtyping from histopathology slides. *Med Image Anal*. 2017;39:194–205. <https://doi.org/10.1016/j.media.2017.04.008>.
- Wu M, Yan C, Liu H, Liu Q. Automatic classification of ovarian cancer types from cytological images using deep convolutional neural networks. *Biosci Rep*. 2018;38(3). <https://doi.org/10.1042/BSR20180289>.
- Farahani H, Boschman J, Farnell D, Darbandsari A, Zhang A, Ahmadvand P, et al. Deep learning-based histotype diagnosis of ovarian carcinoma whole-slide pathology images. *Mod Pathol*. 2022;35(12):1983–90. <https://doi.org/10.1038/s41379-022-01146-z>.
- Ding X, Yang F, Ma F. An efficient model selection for linear discriminant function-based recursive feature elimination. *J Biomed Inf*. 2022;129:104070. <https://doi.org/10.1016/j.jbi.2022.104070>.
- Zhang Z, Jung C. GBDT-MO: Gradient-Boosted decision trees for multiple outputs. *IEEE Trans Neural Netw Learn Syst*. 2021;32(7):3156–67. <https://doi.org/10.1109/TNNLS.2020.3009776>.
- Siegel RL, Giaquinto AN, Jemal A. Cancer statistics, 2024. *CA Cancer J Clin*. 2024;74(1):12–49. <https://doi.org/10.3322/caac.21820>.
- Ries LAG, Eisner MP, Kosary CL, Hankey BF, Miller BA, Clegg L et al. SEER cancer statistics review, 1975–2001. Bethesda (MD): National Cancer Institute.
- Zhou J, Zhang WW, Zhang QH, He ZY, Sun JY, Chen QH, et al. The effect of lymphadenectomy in advanced ovarian cancer according to residual tumor status: A population-based study. *Int J Surg*. 2018;52:11–5. <https://doi.org/10.1016/j.ijsu.2018.02.006>.
- Harter P, Gnauer K, Hils R, Lehmann TG, Fissler-Eckhoff A, Traut A, et al. Pattern and clinical predictors of lymph node metastases in epithelial ovarian cancer. *Int J Gynecol Cancer*. 2007;17(6):1238–44. <https://doi.org/10.1111/j.1525-1438.2007.00931.x>.
- Eisenkop SM, Spiro NM. The clinical significance of occult macroscopically positive retroperitoneal nodes in patients with epithelial ovarian cancer. *Gynecol Oncol*. 2001;82(1):143–9. <https://doi.org/10.1006/gyno.2001.6232>.

41. Ferrero A, Ditto A, Giorda G, Gadducci A, Greggi S, Daniele A, et al. Secondary cytoreductive surgery for isolated lymph node recurrence of epithelial ovarian cancer: a multicenter study. *Eur J Surg Oncol*. 2014;40(7):891–8. <https://doi.org/10.1016/j.ejso.2013.11.026>.
42. Panici PB, Maggioni A, Hacker N, Landoni F, Ackermann S, Campagnutta E, et al. Systematic aortic and pelvic lymphadenectomy versus resection of bulky nodes only in optimally debulked advanced ovarian cancer: a randomized clinical trial. *J Natl Cancer Inst*. 2005;97(8):560–6. <https://doi.org/10.1093/jnci/dji102>.
43. Harter P, Sehouli J, Lorusso D, Reuss A, Vergote I, Marth C, et al. A randomized trial of lymphadenectomy in patients with advanced ovarian neoplasms. *N Engl J Med*. 2019;380(9):822–32. <https://doi.org/10.1056/NEJMoa1808424>.
44. Ledermann JA, Raja FA, Fotopoulou C, Gonzalez-Martin A, Colombo N, Sessa C, et al. Newly diagnosed and relapsed epithelial ovarian carcinoma: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. *Ann Oncol*. 2013;24(Suppl 6):24–32. <https://doi.org/10.1093/annonc/mdt333>.
45. Witjes VM, van Bommel MHD, Ligtenberg MJL, Vos JR, Mourits MJE, Ausems M, et al. Probability of detecting germline BRCA1/2 pathogenic variants in histological subtypes of ovarian carcinoma. A meta-analysis. *Gynecol Oncol*. 2022;164(1):221–30. <https://doi.org/10.1016/j.ygyno.2021.10.072>.
46. Moretta J, Berthet P, Bonadona V, Caron O, Cohen-Haguenauer O, Colas C, et al. [The French genetic and Cancer consortium guidelines for multigene panel analysis in hereditary breast and ovarian cancer predisposition]. *Bull Cancer*. 2018;105(10):907–17. <https://doi.org/10.1016/j.bulcan.2018.08.003>.

### Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.