



Identification of serous ovarian tumors based on polarization imaging and correlation analysis with clinicopathological features

Yulu Huang^{*,†,§§}, Anli Hou^{‡,§,§§}, Jing Wang[‡], Yue Yao[‡], Wenbin Miao[§],
Xuewu Tian^{**}, Jiawen Yu^{**}, Cheng Li^{††}, Hui Ma^{‡,¶,||} and Yujuan Fan^{*,§,‡‡}

**Jinan University, Guangzhou, Guangdong 510632, P. R. China*

*†Department of Gynaecology, Wuzhou Red Cross Hospital, Wuzhou
Guangxi 543002, P. R. China*

*‡Shenzhen Key Laboratory for Minimal Invasive Medical Technologies
Guangdong Engineering Center of Polarization Imaging and
Sensing Technology, Tsinghua Shenzhen International Graduate School
Tsinghua University, Shenzhen, Guangdong 518055, P. R. China*

*§Department of Gynaecology, University of Chinese Academy of
Sciences Shenzhen Hospital, Shenzhen
Guangdong 518106, P. R. China*

*¶Tsinghua-Berkeley Shenzhen Institute
Tsinghua University, Shenzhen, Guangdong 518071, P. R. China*

*||Department of Physics, Tsinghua University
Beijing 100084, P. R. China*

***Department of Pathology, University of Chinese Academy of
Sciences Shenzhen Hospital, Shenzhen
Guangdong 518106, P. R. China*

*††Department of Pathology
Wuzhou Red Cross Hospital, Wuzhou
Guangxi 543002, P. R. China*

‡‡yjfan530@163.com

Received 7 April 2022

Accepted 10 August 2022

Published 31 October 2022

Ovarian cancer is one of the most aggressive and heterogeneous female tumors in the world, and serous ovarian cancer (SOC) is of particular concern for being the leading cause of ovarian cancer

^{‡‡}Corresponding author.

^{§§}These authors contributed equally to this work.

death. Due to its clinical and biological complexities, ovarian cancer is still considered one of the most difficult tumors to diagnose and manage. In this study, three datasets were assembled, including 30 cases of serous cystadenoma (SCA), 30 cases of serous borderline tumor (SBT), and 45 cases of serous adenocarcinoma (SAC). Mueller matrix microscopy is used to obtain the polarimetry basis parameters (PBPs) of each case, combined with a machine learning (ML) model to derive the polarimetry feature parameters (PFPs) for distinguishing serous ovarian tumor (SOT). The correlation between the mean values of PBPs and the clinicopathological features of serous ovarian cancer was analyzed. The accuracies of PFPs obtained from three types of SOT for identifying dichotomous groups (SCA versus SAC, SCA versus SBT, and SBT versus SAC) were 0.91, 0.92, and 0.8, respectively. The accuracy of PFP for identifying triadic groups (SCA versus SBT versus SAC) was 0.75. Correlation analysis between PBPs and the clinicopathological features of SOC was performed. There were correlations between some PBPs (δ , β , q_L , E_2 , $r_{q_{cross}}$, P_2 , P_3 , P_4 , and P_5) and clinicopathological features, including the International Federation of Gynecology and Obstetrics (FIGO) stage, pathological grading, preoperative ascites, malignant ascites, and peritoneal implantation. The research showed that PFPs extracted from polarization images have potential applications in quantitatively differentiating the SOTs. These polarimetry basis parameters related to the clinicopathological features of SOC can be used as prognostic factors.

Keywords: Serous ovarian tumor (SOT); polarimetry basis parameter (PBP); polarimetry feature parameter (PFP); polarization imaging; machine learning (ML).

1. Introduction

Ovarian cancer (OC) occupies third rank among gynecologic cancers¹ and is the most lethal among gynecological malignancies, which seriously threatens the life and health of women worldwide.² Gynecologists carry out postoperative comprehensive treatment according to the postoperative pathological stage of ovarian cancer and improve prognosis. Therefore, the rapid and accurate diagnosis of ovarian cancer is particularly important. However, the existing diagnosis is time-consuming and mainly based on the experience of pathologists. The diagnosis and prognosis of ovarian cancer patients remains a significant challenge. The traditional pathological examination based on texture features of histopathological images is regarded as the “gold standard” in the diagnosis of ovarian cancer. The current status of histopathological images for the classification of ovarian cancer is as follows: (1) The process of histopathological image acquisition includes histological sectioning, staining, and the pathologist’s visual judgment by high-resolution white-light microscopy. With a rapid increase in the incidence rates for tumors, the pathologist’s diagnosis becomes more important as this workload continues to increase.³ (2) In the pathological diagnosis process, the pathologist who needs to identify a large number of cells and observes the relationship between the cells on each slice,

relies on superior training and long experience. Even if there are norms and guidelines, there is still discordance in the diagnosis of the same pathological slice by different pathologists. With the development of precision medicine, the demand for accuracy in pathological diagnosis is increasing.⁴ (3) Ovarian tumors have two main characteristics namely affluent histological types and complicated histological morphology, thus making it more difficult to distinguish ovarian tumors.⁵ In a recent study involving 114 patients and three pathologists, it was found that ovarian cancer subtype diagnosis was discordant on an average at 15.8% (18/114), with the disagreement on serous ovarian carcinoma being 10.9% (6/55).⁶ Therefore, there is an urgent need to find a rapid, efficient, objective, quantitative, and reproducible pathological slice evaluation method, rendering more accurate differential diagnoses of ovarian tumors.

Biological tissue is an intrinsically turbid medium. Multiple scattering tends to lead to losses of optical information obtained by optical imaging by reducing the image quality and spatial resolution. Polarization imaging can be used to improve the image quality⁷ and provide far more information to characterize the samples by reducing the effects of unwanted light reflection.⁸ Recently, polarization imaging technology has attracted more and more attention due to its unique ability to detect the microstructural

information of complex biological samples, and has shown great application potential in the biomedical field.^{9–12} Moreover, machine learning (ML) has been an important method for image processing and classification.¹³ Characterization of polarimetry basis parameters (PBPs) has shown them to be associated with the microstructure changes in female cancer. Dong *et al.* proposed a linear discriminant analysis (LDA)-based approach for obtaining the polarimetry feature parameters (PFPs) used to identify the target microstructure in different breast tissues.⁹ In a previous study, polarization imaging and data extraction were carried out on cervical cancer and precancerous lesions, which could be preliminarily quantitatively classified with an accuracy of 85%.¹¹

In this study, polarization detection of ovarian tumors was performed to find a method that can quantitatively distinguish different ovarian lesions, and further analyze the relationship between the clinicopathological characteristics of ovarian cancer and polarization parameters. We collected Mueller matrix images dataset from the H&E pathological slices of serous ovarian tumor (SOT), including serous cystadenoma (SCA), serous borderline tumor (SBT), and serous adenocarcinoma (SAC). We obtained the PBPs calculated from the processing of Mueller matrix polarization decomposition (MMPD) and Mueller matrix transformation (MMT), among others, and then we derived a new PFP from the linearity of PBP by using LDA with machine learning. A total of 105 ovarian specimens were analyzed in this study and diagnosed as SOTs (30 cases of SCA, 30 cases of SBT, and 45 cases of SAC). Training and testing were completed in 315 regions of interest (ROIs). By comparing the results of PFP and the corresponding situation of ground truth in the labeled H&E image, it can be proved that the PFP derived by this technology can quantitatively characterize epithelioglandular microstructure features (EMFs) in three SOTs. This technology is expected to become a powerful complementary diagnostic tool for ovarian cancer. In addition, the characterization of PBPs was obtained by polarization imaging to clarify the correlation between the mean values of PBPs and the clinicopathological features of patients with serous ovarian cancer (SOC), to provide reference indices for evaluating the exasperation degree and prognosis of patients. So far, there has been no research to quantitatively characterize and distinguish SOTs by polarization

imaging and the correlation analysis of clinicopathological features of patients with SOC.

2. Materials and Methods

2.1. Patients and ROI selection

We collected 105 H&E pathological slides (4- μm -thick) from patients from the Shenzhen Hospital of the University of Chinese Academy of Sciences and Wuzhou Red Cross Hospital (from 2012 to 2022). Each H&E pathological slide was diagnosed by an experienced pathologist, wherein the total number of evaluated samples included 30 SCA, 30 SBT, and 45 SAC cases. For quantitative analysis of the EMFs, we chose three ROIs from each case. Hence, 315 ROIs were selected and analyzed in all 105 tissue samples. The ROIs of each SOT were marked and labeled manually by the pathologist through Matlab and open-source software Ilastik. This work was approved by the Institutional Review Boards at Shenzhen Hospital of the University of Chinese Academy of Sciences and Wuzhou Red Cross Hospital.

2.2. Experimental setup

Figure 1 shows the experimental setup. The Mueller matrix microscope based on dual DoFP polarimeters was developed by adding the polarization states generator (PSG) module and polarization states analyzer (PSA) module to a commercial transmission optical microscope (Yuexian L2050, Guangzhou LISS Optical Instrument Co., Ltd., China). As shown in Fig. 1, the PSG consists of an aligned and fixed linear polarizer (P_1 , extinction ratio 500:1, Daheng Optics, China) and a rotatable quarter-wave retarder (R_1 , Daheng Optics, China). PSA consists of two 16-bit DoFP polarimeters (PHX050s-PC, Lucid Vision Labs Inc., Canada, DoFP-CCD₁ and DoFP-CCD₂), a 50:50 non-polarized beam splitter prism (CCM1-BS013/M, Thorlabs, Inc., USA), and a fixed-angle quarter-wave plate (R_2 , Daheng Optics, China). Two DoFP polarimeters, DoFP-CCD₁ and DoFP-CCD₂, are fixed on the transmission end and the reflection end of the nonpolarized beam splitter prism respectively, and R_2 is installed between the transmission end of the splitting prism and DoFP-CCD₁. Before the test, the instrument was calibrated with air to ensure the measurement accuracy, the maximum

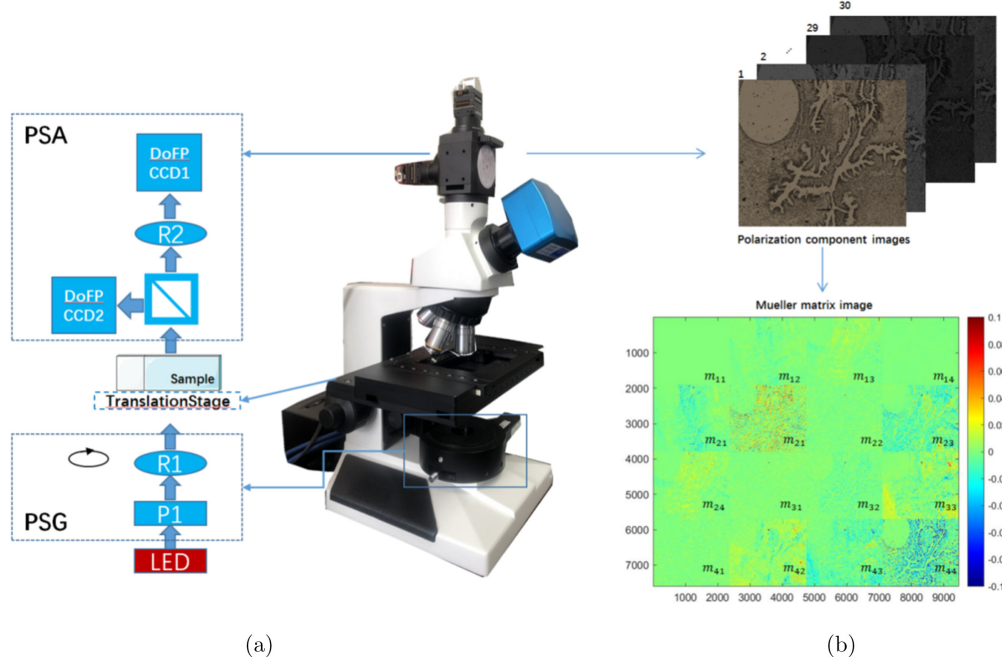


Fig. 1. (a) Schematic of transmission Mueller matrix microscope.¹⁴ (b) An example of Mueller matrix image of the H&E pathological slide of serous ovarian tumor. All the elements are normalized by m_{11} . We subtract the identity matrix from the Mueller matrix for display. The color bar of all elements is from -0.1 to 0.1 .

errors are about 0.01. The purpose of system calibration is to determine W_p . Since we have the equation

$$I = W_{\text{PSA}} M S_{\text{in}},$$

where W_{PSA} denotes the instrument matrix of PSA, when measuring air as the standard sample, the Mueller matrix of air can be regarded as an identity matrix, then W_{PSA} can be calculated by

$$W_{\text{PSA}} = I S_{\text{in}}^{-1},$$

where the superscript -1 represents the inverse or pseudoinverse of the matrix. I is an 8×4 matrix consisting of light intensity measured by dual DoFP polarimeters during four exposures, and S_{in} is a 4×4 matrix including four input Stokes vectors. Then the q th column of polarimetric measurement matrix W_p can be calculated by

$$W_p^T(q) = \text{vec}(S_{\text{in}}^T(q) \otimes W_{\text{PSA}}^T(q)),$$

where \otimes is the Kronecker product.

After the calibration, we use Mueller matrix microscope to measure standard polarization samples including the linear polarizer and quarter-wave plate, and calculate average root-mean-square error (RMSE) to evaluate the calibration performance; the mean RMSEs of the linear polarizer and quarter-wave plate are less than 1%, which means the

Mueller matrix microscope has high measurement accuracy.

After calibration, red light (633 nm, $\Delta\lambda = 20$ nm) was focused onto the sample through an objective lens (4 \times). According to the polarization states of the incident light and the emergent light, the DoFP polarimeter records the polarization component images with independent polarization orientations and 16 images can be calculated. The Mueller matrix measurement can be expressed as

$$I = W_p M, \quad \text{or } M = W_p^{-1} I, \quad (1)$$

$$I = \begin{bmatrix} I_{0^\circ_CCD1} \\ I_{45^\circ_CCD1} \\ I_{90^\circ_CCD1} \\ I_{135^\circ_CCD1} \\ I_{0^\circ_CCD2} \\ I_{45^\circ_CCD2} \\ I_{90^\circ_CCD2} \\ I_{135^\circ_CCD2} \end{bmatrix}, \quad M = \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \\ m_{41} & m_{42} & m_{43} & m_{44} \end{bmatrix}, \quad (2)$$

where I denotes a column vector containing intensity images corresponding to the four polarization orientations of DoFP-CCD₁ and DoFP-CCD₂ under four different incident polarization states. M contains

16 elements of Mueller matrix, W_p represents the polarimetric measurement matrix, and W_p^{-1} denotes the pseudoinverse of W_p . More details can be found in Refs. 14 and 15.

2.3. Polarimetry basis parameters

Polarimetry basis parameters are calculated and obtained by MMPD¹⁶ and MMT,^{8,17} among others, which can be used to deduce the new polarimetry feature parameters. More details about the calculation process and the explanations of physical meanings can be found in Refs. 12, and 17–21. The PBPs used in the study are shown in Table 1.

2.4. Input data and training labels

(1) Histological criteria of distinctive pathological features in typical ovarian tissue

SCA comprises a simple layer of epithelium, lacking cytological atypia, structural complexity, or stromal invasion. The epithelium in SBT is multilayered, architecturally complex, demonstrating cellular atypia but lacking destructive stromal invasion, whereas the epithelium in SAC shows prominent cytological atypia, bizarre cytomorphology, and increased pathological nuclear division.²² Three representative examples of EMFs in the H&E pathological slide of the three SOTs are given in Figs. 3(a)–3(c).

(2) Data preprocessing

In each type of ovarian tissue, machine learning classifier was established to identify the epithelium region from the complex biological tissue, and 100,000 pixels were randomly selected from each type to learn and characterize the target epithelium region. In this study, the input data of the machine learning classifier was pixel values of PBPs used for training and testing, while the H&E images directly marked by pathologists served as the ground truth. Therefore, we need to map the mask labeled by the pathologist on the H&E image with the corresponding two-dimensional images of the PBPs pixel by pixel, in order to correctly distinguish whether the input data belongs to positive or negative class. In this study, we adopted affine transformation for the polarization and H&E image accurate registration.

2.5. Machine learning algorithms

We employed linear discrimination analysis-based machine learning algorithms in this research. The ML methods were implemented using Python

(version 3.8.0) with the open-source library Scikit-learn. We derived new PFPs from the linear combination of existing PBPs based on the LDA machine learning algorithm, which had better performance in distinguishing epithelium regions in the histological sections of different SOTs. The algorithm process of deriving PFPs to quantitatively characterize the target microstructure features is as follows: First, we take 65 PBPs of the training sample as the input data to distinguish the target features and calculate the accuracy of the LDA classifier, and screen out PBPs with a weight of more than one-hundredth of the maximum weight, which have similar accuracy compared to 65 PBPs. The filtered PBPs are denoted as Y_m . Then enter the number of PBPs, namely n ($n = m - i$, $i = 1, 2, \dots, m - 1$), which is decreased one by one. The LDA model is trained by optimizing the correlation coefficients in the combinations of n input PBPs, which minimizes the variability within each class and maximizes the discrepancy between different classes. Therefore, the output of the LDA model is a PFP, i.e., a linear combination of n PBPs with optimized coefficients. In consequence, when the number of input PBPs is n , all possible combinations are traversed, and each one has the accuracy of target characterization. Choosing the highest accuracy value (Y_n) and the corresponding combination of n PBPs with optimized coefficients (X_n), the process moves to the next step. If the margin between Y_m and Y_n is less than 0.01, it indicates that the characterization capability of the new output PFP is still maintained for the target feature, despite a decreased number of PBPs being used, then it can return the step of reducing the number of input PBPs. Otherwise, it means that the LDA model has a poor ability to identify the target microstructures. Therefore, the program stops the loop and sets X_{n+1} in the previous loop to output PFP, which is the optimal linear combination for the classification, while Y_{n+1} remains as accurate as Y_m . Figure 2 shows the flowchart of the machine learning algorithm.

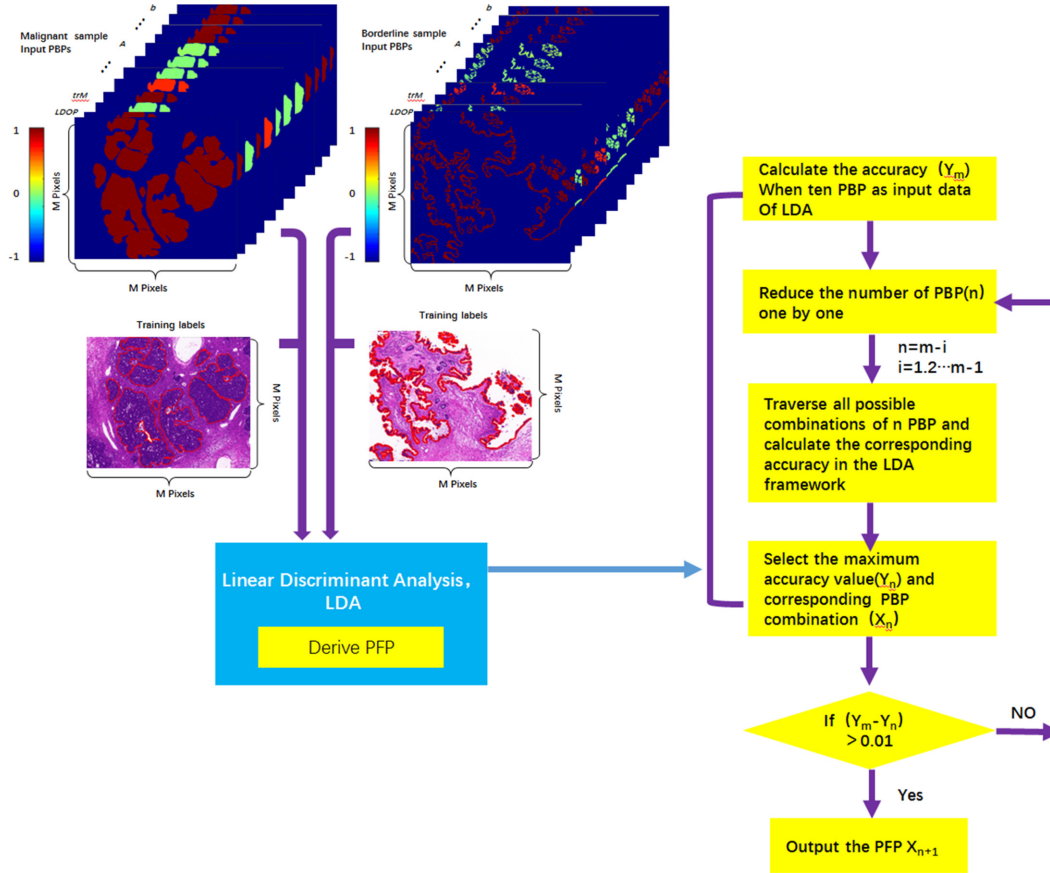
3. Results and Discussion

3.1. Quantitative characterization results of EMFs by PBPs

The PBPs of each pixel on the EMFs can be measured and calculated by Muller matrix microscope,

Table 1. PBPs used in the study.

PBP	Physical meaning	Formula
Delta (Δ)	Depolarization	$\Delta = 1 - \frac{ \text{tr}(M_{\Delta}-1) }{3} \in [0, 1]$
Delta (δ)	Linear retardation	$\delta = \cos^{-1}\{[(M_R(2, 2) + M_R(3, 3))^2 + (M_R(3, 2) + M_R(2, 3))^2]^{\frac{1}{2}} - 1\}$
D	Diattenuation	$D = \sqrt{(m_{12})^2 + (m_{13})^2 + (m_{14})^2} \in [0, 1]$
Theta (θ)	Orientation of fast axis	$\theta = \frac{1}{2} \tan^{-1}\left(\frac{r_2}{r_1}\right)$
Beta (β)	Circular retardance or overlap of multiple anisotropies	$\beta = \frac{m_{23}-m_{32}}{2}$
$t_{2131}(P_L)$	Linear polarizance	$P_L = \sqrt{(m_{21})^2 + (m_{31})^2} \in [0, 1]$
$t_{1213}(D_L)$	Linear diattenuation	$D_L = \sqrt{(m_{12})^2 + (m_{13})^2} \in [0, 1]$
$t_{4243}(q_L)$	Linear birefringence	$q_L = \sqrt{(m_{42})^2 + (m_{43})^2} \in [0, 1]$
$t_{2434}(r_L)$	Linear birefringence	$r_L = \sqrt{(m_{24})^2 + (m_{34})^2} \in [0, 1]$
$\text{phi}_{3121}(\alpha_P)$	Orientation parameter corresponding to linear polarizance	$\alpha_P = \frac{1}{2} \tan^{-1}\left(\frac{m_{31}}{m_{21}}\right)$
$\text{phi}_{4243}(\alpha_q)$	Orientation parameter corresponding to linear birefringence	$\alpha_q = \frac{1}{2} \tan^{-1}\left(\frac{m_{42}}{m_{43}}\right)$
$\text{phi}_{1312}(\alpha_D)$	Orientation parameter corresponding to linear birefringence	$\alpha_D = \frac{1}{2} \tan^{-1}\left(\frac{m_{13}}{m_{12}}\right)$
$\text{phi}_{2434}(\alpha_r)$	Orientation parameter corresponding to linear diattenuation	$\alpha_r = \frac{1}{2} \tan^{-1}\left(\frac{-m_{24}}{m_{34}}\right)$
A	Normalized anisotropy	$A = \frac{2bt}{b^2+t^2}$
b	Depolarization	$b = \frac{m_{22}+m_{33}}{2}$
t_1	Anisotropy degree	$t_1 = \frac{\sqrt{(m_{22}-m_{33})^2 + (m_{23}-m_{32})^2}}{2}$
Bhls ($ B $)	Determinant of Mueller matrix's central 2×2 block	$ B = m_{22}m_{33} - m_{23}m_{32}$
Bfs ($\ B\ $)	Frobenius norm of Mueller matrix's central 2×2 block	$\ B\ = \sqrt{m_{22}^2 + m_{23}^2 + m_{32}^2 + m_{33}^2}$
Trace ($\text{tr}M$)	Trace of Mueller matrix	$\text{tr}M = m_{11} + m_{22} + m_{33} + m_{44}$
Norm ($\ M\ $)	Frobenius norm of Mueller matrix	$\ M\ = m_{11}^2 + m_{12}^2 + m_{13}^2 + m_{14}^2 + m_{21}^2 + m_{22}^2 + m_{23}^2 + m_{24}^2 + m_{31}^2 + m_{32}^2 + m_{33}^2 + m_{34}^2 + m_{41}^2 + m_{42}^2 + m_{43}^2 + m_{44}^2$
LDOP	Linear degree of polarization	$\text{LDOP} = \frac{m_{21}+m_{22}}{m_{11}+m_{12}} \in [0, 1]$
k_C	Circular polarization maintenance and cosine of retardation in pure depolarization and pure linear retarder systems, respectively	$k_C = m_{44} \in [-1, 1]$
P_2	Parameter related to linear birefringence	$P_2 = \sqrt{m_{34}^2 + m_{43}^2}$
P_3	Parameter related to linear birefringence	$P_3 = \sqrt{ m_{34} \times m_{43} }$
TA ₄₂₂₄ (P_4)	Transverse mirror asymmetry	$P_4 = m_{42} + m_{24} $
P_5	Parameter related to linear birefringence	$P_5 = \sqrt{m_{42}^2 + m_{24}^2}$
P_8	Asymmetry of linear birefringence	$P_8 = \frac{m_{42} \times m_{24} + m_{22} + m_{34} \times m_{43} + m_{33}}{3}$
E_2	Equality parameter related to depolarization	$E_2 = (m_{11} - m_{22})^2 - (m_{12} - m_{21})^2 - (m_{33} - m_{44})^2 - (m_{34} + m_{43})^2$
E_3	Equality parameter related to depolarization	$E_3 = (m_{11} + m_{21})^2 - (m_{12} + m_{22})^2 - (m_{13} + m_{23})^2 - (m_{14} + m_{24})^2$
E_4	Equality parameter related to depolarization	$E_4 = (m_{11} - m_{21})^2 - (m_{12} - m_{22})^2 - (m_{13} - m_{23})^2 - (m_{14} - m_{24})^2$
E_6	Equality parameter related to depolarization	$E_6 = (m_{11} - m_{12})^2 - (m_{21} - m_{22})^2 - (m_{31} - m_{32})^2 - (m_{41} - m_{42})^2$
rq_{cross}	Parameter related to transpose symmetry	$\text{rq}_{\text{cross}} = \frac{(m_{24} \times m_{43} - m_{34} \times m_{42})}{(r_L \times q_L)}$



Notes: The linear model is employed: LDA is developed to quantitatively characterize and distinguish the different microstructure features. Here, the EMFs of SBT and SAC samples as an example are recognized by the PFP. The areas inside the red solid line in the H&E images are the target features labeled by the pathologist.

Fig. 2. The flowchart of deriving PFPs by using a machine learning algorithm.

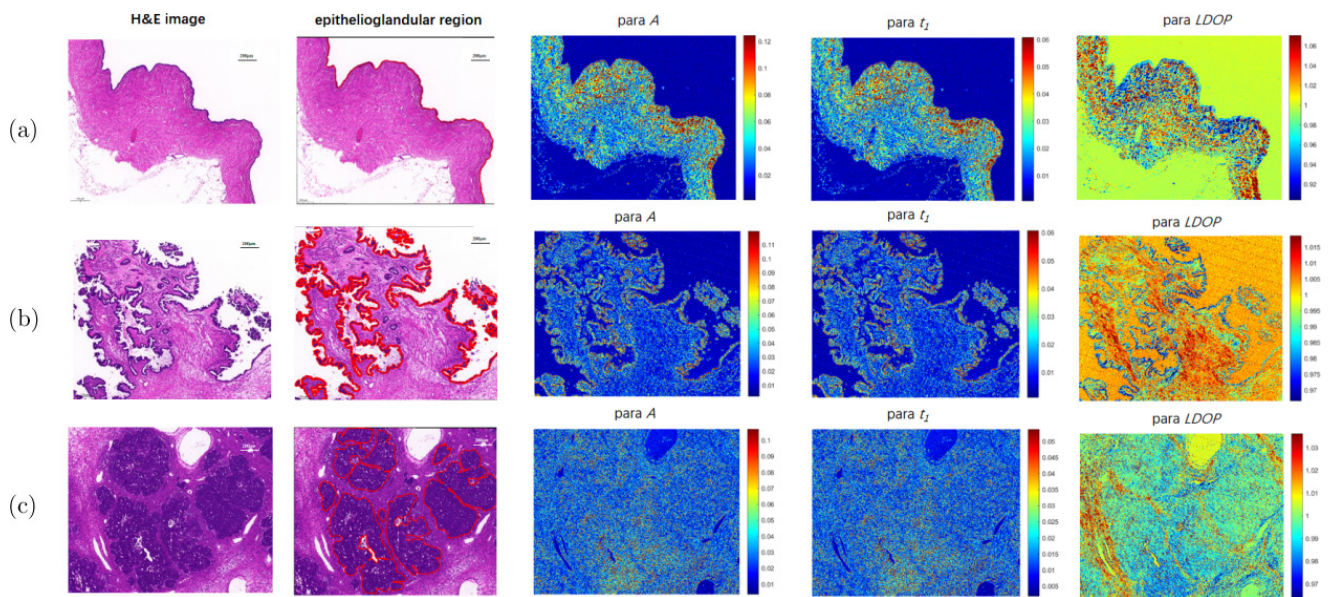


Fig. 3. Representative PBP images of SOTs, i.e., parameter A , parameter t_1 , and parameter LDOP: (a) SCA, (b) SBT, and (c) SAC. In the H&E image, the areas inside the red solid line are the epithelioglandular region marked by the pathologist.

then the microstructure information of SOTs can be quantitatively characterized, which provided a method basis for the quantitative diagnosis of SOTs. As shown in Fig. 3, the epithelioglandular structures in different SOTs were characterized by the parameter A that is related to anisotropy, parameter t_1 related to anisotropy degree, and parameter LDOP related to orientation, respectively. Because PBPs correspond to some physical features, the contrast difference of different PBP images in the lesion area shows the pathological characteristics of the lesion, thus providing a new tool for revealing the mechanism of the lesion. In the EMFs on the SCA sample, the values of parameter A are larger than those in EMFs of the SAC sample, indicating that benign tissues have good anisotropic structures compared with cancerous tissues, as shown in Figs. 3(a) and 3(c). The values of parameter t_1 in the EMFs on the SBT sample are less than those of the SCA sample, indicating that borderline tissues have disordered anisotropic degrees compared with benign tissues, as shown in Figs. 3(a) and 3(b). The values of parameter LDOP in EMFs of the SAC sample are larger than those of the SBT sample, reflecting cancerous tissues have stronger linear polarization maintaining capability than borderline tissues, as shown in Figs. 3(b) and 3(c).

Because the polarization parameters are not sensitive to image resolution, the PBPs measured by low resolution can still provide abundant pathological structure information. To confront pathological tissues with more complicated structures and difficulty in using the existing PBPs to characterize, the new PFP derived by PBPs and ML method can be used to characterize the target microstructure.

3.2. Deriving PFPs from PBPs

After measuring the existing PBPs from each pixel of the SOT sample, we derive four PFPs from the existing PBPs to classify two or three categories of SOTs. Each PFP is mainly used to classify two or three classes of target microstructure features. The performance of PFP derived from the LDA method was tested by validating it with a 10-fold cross-validation method. According to the values of the four PFPs, the classification target microstructure features of the image pixels can be determined. For example, if the PFP value exceeds a certain pre-determined threshold, that pixel is predicted to be

the target feature of that type of SOT in dichotomies, otherwise it is predicted to be another one or two. For the three categories, LDA algorithm records one category of SOTs as class 1, and the other two categories as class 0 for binary classification, and trains the weight of the model via three iterations (i.e., three binary classifications). Then we selected the greatest probability of PFPs to characterize the three classifications by inputting the test samples into the three classification models. Figure 4 shows the process of deriving the PFPs for the characterization of EMFs in the H&E pathological sections of SOTs.

PFPs can be used to quantitatively and rapidly assess the proportion and distribution of EMFs related to pathogenesis and carcinogenesis in typical ovarian pathological tissues, which has great potential in the process of quantitatively characterizing microstructure features and thus accelerating the automatic classification of histopathological images in pathological diagnosis.

Quantitative extraction of polarization information can detect subtle microstructural differences not visible to the naked eye. Ahmad *et al.*²³ found that the derived polarization parameters such as diattenuation, retardance and depolarization can differentiate normal and adenocarcinoma colon tissue samples. Du *et al.*²⁴ confirmed that the parameters LDOP, D , Δ , and b exhibit contrast differences between the skin cancerous and normal tissues. In our study, the PFP, as an optimal linear combination of PBPs, has physical interpretability: (1) The parameter A (with the largest coefficient value of 50.98) represents that sample anisotropy plays an important role in differentiating SAC and SCA, which may be related to changes of tumor cell arrangement with the lesion development. (2) The parameter t_1 (with the largest coefficient value of 88.12) represents that sample overall degree of anisotropy is more sensitive to differentiating SCA and SBT, which may be related to tumor cell proliferation. (3) The parameter LDOP (with the largest coefficient value of -1439.34) represents that linear polarization maintainance capability is more significant for differentiating SAC and SBT, which may be related to the increased number density of cell organelles. (4) The parameter E_6 correlates to the fact that sample depolarization also contributes a lot to differentiate SCA, SBT, and SAC, which may be associated with the density of tumor cells at different types of SOT.

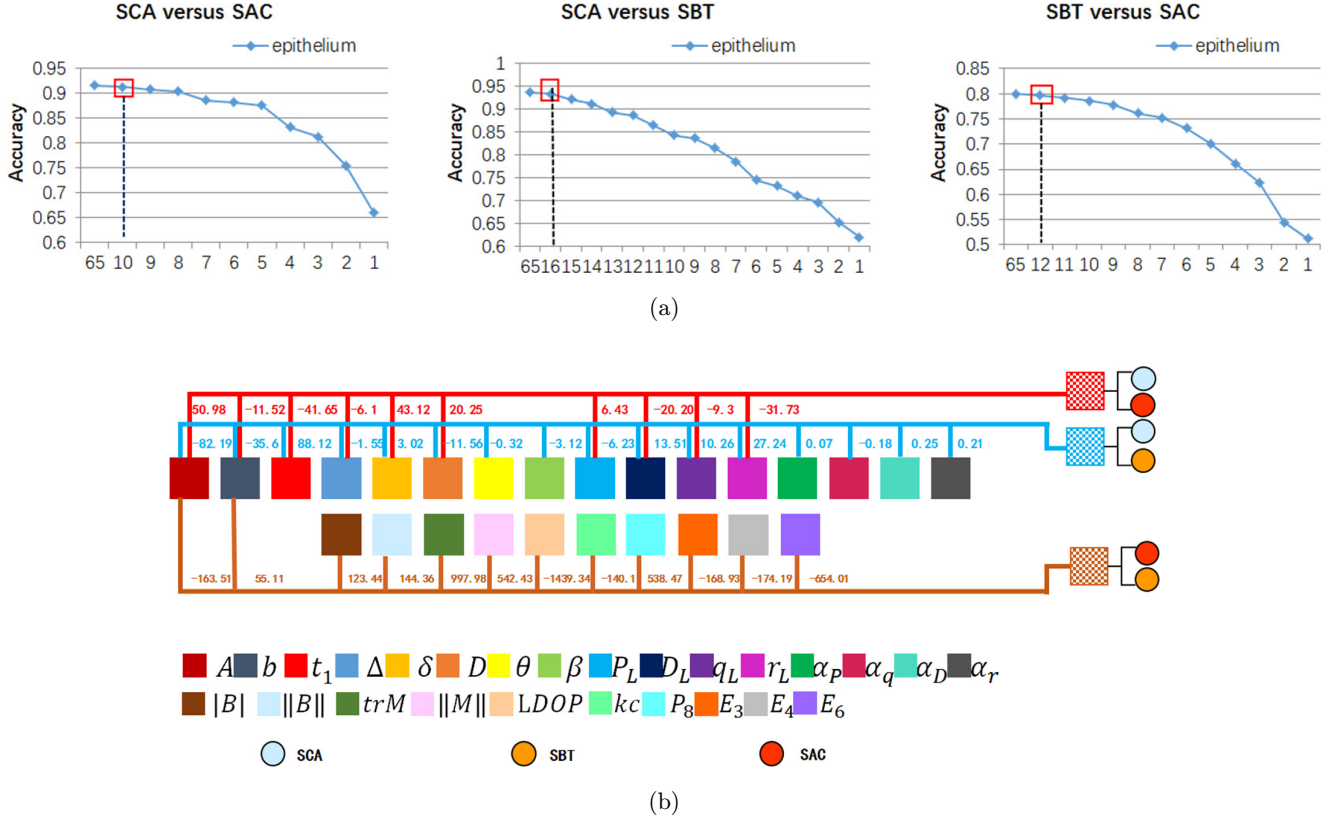


Fig. 4. (a) The process of deriving PFPs using the machine learning method to classify the characterization of EMFs in the three SOTs: SCA versus SAC, SCA versus SBT, and SBT versus SAC. (b) The linear combinations of 26 PBPs (represented by rectangles in different colors without borders) as the three PFPs (represented by rectangles in different textures with borders) which can quantitatively characterize the EMFs in each SOT (represented by the solid circle in different colors with borders). The numbers above PBPs are the optimized coefficients acquired by the machine learning classifier, and the PFP can be obtained from a weighted sum of the PBPs with optimized coefficients.

The characterization is consistent with the growth behavior of serous ovarian tumors.

3.3. Characterization performance of the polarization imaging-based ML model

In this study, the LDA algorithm is used to derive PFPs, and 10-fold cross-validation is used to verify the stability of the classification model. Table 2 presents the average accuracy, precision, recall, and AUC values of the four PFPs.

Table 2. Classification results on the testing set.

Classification	Accuracy	Precision	Recall	AUC
SCA versus SAC	0.91	0.92	0.88	0.85
SCA versus SBT	0.92	0.90	0.93	0.86
SBT versus SAC	0.80	0.84	0.85	0.86
SCA versus SBT versus SAC	0.75	0.81	0.81	0.88

The validation results in Table 2 demonstrate the recognition of EMFs in the three SOT tissue sections: (1) The AUC value of PFPs is higher than 0.85, indicating preferable classification performance. (2) PFPs have a high recall value (0.81–0.93), i.e., clinical diagnostic sensitivity, and are used to quantitatively characterize EMFs in various type of SOTs. (3) The classification of SBT versus SAC has a slightly lower accuracy among all the binary classifications, resulting in the recognition values of different microstructure characteristics ranging from 0.80 to 0.92. This result is consistent with clinical experience that there are more disagreements in distinguishing between borderline tumor and malignant tumor in pathological diagnosis, and the previous studies demonstrated that no single histologic or cellular feature except stromal invasion was absolute in differentiating borderline tumor from ovarian carcinoma.^{25,26} Therefore, although there are still certain misjudgments

in the classification of SBT versus SAC, the accuracy is within an acceptable range. (4) The ternary classification results of SCA versus SBT versus SAC report an accuracy of 0.75, a precision of 0.81, a recall of 0.81, and an AUC of 0.88. The overall prediction performance of PFPs was satisfactory in the characterization of EMFs out of the three SOTs. Precision pathology requires characteristics analysis and an optimized ML model with high precision to assist the clinician in making diagnostic decisions in practical scenarios.²⁷ Here, PFPs are highly sensitive in quantitatively characterizing specific microstructures in low-resolution and wide-field-of-view systems. Therefore, PFPs may play an important role in rapid and extensive initial screening for SOTs.

3.4. The relationship between PBP and clinicopathological features

The Spearman correlation analysis was performed between PBPs and clinicopathological features. The clinicopathological features included the pathological grading (low and high grades), the International Federation of Gynecology and Obstetrics (FIGO) stages (stages I–II, stages III–IV), preoperative

ascites (yes, no), malignant ascites (yes, no), serum CA125 expression level, tumor size (greater than 5 cm, m5 cm), lymph node metastasis (yes, no), vascular invasion (yes, no), and peritoneal implantation (yes, no). The software of SPSS version 26.0 and GraphPad Prism 9 were used for statistical analysis. Two nonnormally distributed groups were compared using the Mann–Whitney *U*-test. Differences were considered statistically significant when the *p*-value was less than 0.05. Figure 5 shows the Spearman correlation coefficients between PBPs and clinicopathological features, and Table 3 shows the association between the mean values of PBPs and clinicopathological characteristics of serous ovarian cancer. The results showed that there were positive correlations between the parameter E_2 and FIGO stage and malignant ascites ($p < 0.05$), while there were negative correlations between the parameters P_2 , P_3 and pathological grading ($p < 0.05$). There were negative correlations between the parameters δ , β , q_L , rq_{cross} , P_2 , P_3 , P_4 , P_5 and the FIGO stage ($p < 0.05$). The parameter rq_{cross} was negatively correlated with preoperative ascites ($p < 0.05$); parameters rq_{cross} , P_4 were negatively correlated with malignant ascites ($p < 0.05$); and the parameters q_L , rq_{cross} were also negatively

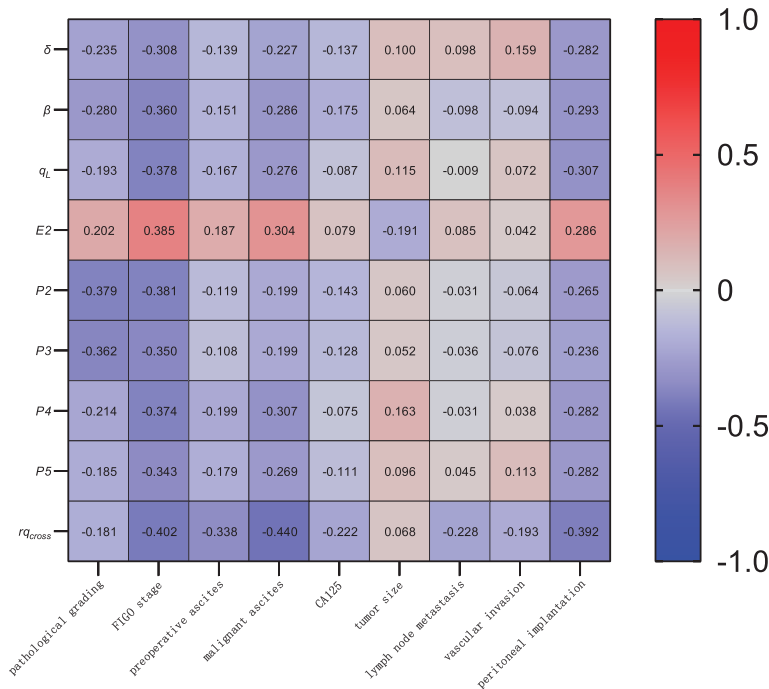


Fig. 5. Correlation heatmap. Spearman correlation coefficients between the mean values of PBPs (*y*-axis) and clinicopathological features (*x*-axis). Red indicates positive correlation, blue indicates negative correlation, and white means no correlation or non-statistically significant correlation.

Table 3. Comparison of the mean values of PBPs in SOC with different clinicopathological features [median (interquartile range)].

Clinicopathological features	<i>n</i>	Para_δ	Para_β	Para_γ _L	Para_E ₂	Para_rq _{cross}
FIGO stage						
Stage I ~ II	19	0.0175 (0.0217 ~ 0.0155)	0.0078 (0.0055 ~ 0.0091)	0.0343 (0.0291 ~ 0.0423)	-0.0044 (-0.0071 ~ -0.0031)	0.5418 (0.3928 ~ 0.6010)
Stage III ~ IV	26	0.0150 (0.0181 ~ 0.0119)	0.0055 (0.0043 ~ 0.0076)	0.0257 (0.0148 ~ 0.0352)	-0.0023 (-0.0042 ~ -0.0008)	0.3884 (0.1739 ~ 0.5417)
Z		-2.045	-2.390	-2.505	-2.551	-2.666
P		0.041*	0.017*	0.012*	0.011*	0.008*
Pathological grading						
Low grade	10	0.0191 (0.0138 ~ 0.0236)	0.0068 (0.0055 ~ 0.0112)	0.0349 (0.0217 ~ 0.0459)	-0.0048 (-0.0072 ~ -0.0019)	0.5117 (0.2510 ~ 0.6463)
High grade	35	0.0157 (0.0123 ~ 0.0183)	0.0056 (0.0046 ~ 0.080)	0.0301 (0.0187 ~ 0.0370)	-0.0031 (-0.0049 ~ -0.0009)	0.4484 (0.2757 ~ 0.5491)
Z		-1.556	-1.856	-1.283	-1.338	-1.201
P		0.120	0.063	0.199	0.181	0.230
Preoperative ascites						
No	11	0.0165 (0.0145 ~ 0.0231)	0.0078 (0.0044 ~ 0.0087)	0.0312 (0.0291 ~ 0.0448)	-0.0044 (-0.0071 ~ -0.0022)	0.5645 (0.4768 ~ 0.6010)
Yes	34	0.0157 (0.0127 ~ 0.0186)	0.0056 (0.0050 ~ 0.0080)	0.0301 (0.0207 ~ 0.0374)	-0.0028 (-0.0055 ~ -0.0012)	0.4263 (0.2405 ~ 0.5427)
Z		-0.924	-1.004	-1.109	-1.241	-2.245
P		0.355	0.316	0.267	0.214	0.025*
Malignant ascites						
No	18	0.0165 (0.0152 ~ 0.0195)	0.0077 (0.0054 ~ 0.0083)	0.0322 (0.0298 ~ 0.0390)	-0.0040 (-0.0067 ~ -0.0029)	0.5530 (0.4558 ~ 0.5956)
Yes	27	0.0155 (0.0121 ~ 0.0185)	0.0055 (0.0040 ~ 0.0080)	0.0260 (0.0127 ~ 0.0386)	-0.0026 (-0.0055 ~ -0.0005)	0.3666 (0.1787 ~ 0.5293)
Z		-1.506	-1.900	-1.830	-2.016	-2.919
P		0.132	0.057	0.067	0.044*	0.004*
Peritoneal implantation						
No	17	0.0175 (0.0156 ~ 0.0217)	0.0078 (0.0053 ~ 0.0089)	0.0333 (0.0290 ~ 0.0423)	-0.0040 (-0.0071 ~ -0.0030)	0.5418 (0.4348 ~ 0.6198)
Yes	28	0.0150 (0.0123 ~ 0.0182)	0.0056 (0.0045 ~ 0.0076)	0.0277 (0.0163 ~ 0.0369)	-0.0025 (-0.0048 ~ -0.0010)	0.3884 (0.1828 ~ 0.5441)
Z		-1.873	-1.943	-2.037	-1.896	-2.599
P		0.061	0.052	0.042*	0.058	0.009*

Clinicopathological features	<i>n</i>	Para_P ₂	Para_P ₃	Para_P ₄	Para_P ₅
FIGO stage					
Stage I ~ II	19	0.0156 (0.0132 ~ 0.0179)	0.0087 (0.0073 ~ 0.0096)	0.0336 (0.0302 ~ 0.0470)	0.0309 (0.0271 ~ 0.0407)
Stage III ~ IV	26	0.0128 (0.0116 ~ 0.0147)	0.0070 (0.0063 ~ 0.0080)	0.0256 (0.0123 ~ 0.0374)	0.0263 (0.0181 ~ 0.0339)
Z		-2.528	-2.321	-2.482	-2.275
P		0.011*	0.020*	0.013*	0.023*
Pathological grading					
Low grade	10	0.0183 (0.0127 ~ 0.0227)	0.0103 (0.0068 ~ 0.0134)	0.0354 (0.0212 ~ 0.0487)	0.0328 (0.0197 ~ 0.0439)
High grade	35	0.0134 (0.0118 ~ 0.0154)	0.0073 (0.0063 ~ 0.0080)	0.0294 (0.0128 ~ 0.0380)	0.0274 (0.0212 ~ 0.0353)
Z		-2.512	-2.402	-1.420	-1.229
P		0.012*	0.016*	0.156	0.219
Preoperative ascites					
no	11	0.0146 (0.0124 ~ 0.0177)	0.0084 (0.0063 ~ 0.0096)	0.0332 (0.0302 ~ 0.0479)	0.0302 (0.0271 ~ 0.0427)
yes	34	0.0138 (0.0118 ~ 0.0167)	0.0074 (0.0064 ~ 0.0088)	0.0286 (0.0128 ~ 0.0384)	0.0279 (0.0197 ~ 0.0355)
Z		-0.792	-0.713	-1.321	-1.188
P		0.428	0.476	0.187	0.235
Malignant ascites					
No	18	0.0155 (0.0123 ~ 0.0178)	0.0086 (0.0064 ~ 0.0098)	0.0323 (0.0300 ~ 0.0414)	0.0304 (0.0273 ~ 0.0380)
Yes	27	0.0138 (0.0118 ~ 0.0154)	0.0073 (0.0063 ~ 0.0085)	0.0259 (0.0110 ~ 0.0397)	0.0260 (0.0160 ~ 0.0358)
Z		-1.321	-1.321	-2.039	-1.784
P		0.187	0.187	0.041*	0.074
Peritoneal implantation					
No	17	0.0156 (0.0124 ~ 0.0178)	0.0087 (0.0067 ~ 0.0094)	0.0319 (0.0290 ~ 0.0462)	0.0309 (0.0266 ~ 0.0411)
Yes	28	0.0131 (0.0119 ~ 0.0151)	0.0072 (0.0063 ~ 0.0083)	0.0269 (0.0127 ~ 0.0379)	0.0266 (0.0189 ~ 0.0353)
Z		-1.756	-1.568	-1.873	-1.873
P		0.079	0.117	0.061	0.061

Note: **p* < 0.05 was defined as statistically significant.

correlated with peritoneal implantation (*p* < 0.05). There were no statistically significant differences in the mean values of existing PBPs among patients with tumor size, lymph node metastasis, vascular invasion, and serum CA125 expression level (*p* > 0.05).

The significant parameters from the previous correlation analysis were then analyzed by the receiver operating characteristic (ROC) regression (The results are shown in Table 4 and Fig. 6). The polarization parameters with the highest Youden

Table 4. Performance analysis of the mean values of PBPs in predicting the risk of clinicopathological features of SOC patients.

Clinicopathological features	Parameter	Sensitivity	Specificity	Youden index	AUC	Cut-off value
FIGO stage	E_2	0.654	0.842	0.495	0.725	-0.0028
Pathological grading	P_3	0.914	0.700	0.614	0.753	0.0092
Preoperative ascites	rq_{cross}	0.882	0.636	0.518	0.727	0.5530
Malignant ascites	rq_{cross}	0.926	0.611	0.537	0.759	0.5479
Peritoneal plant	rq_{cross}	0.643	0.765	0.408	0.733	0.4626

index and $AUC > 0.7$ corresponding to clinicopathological features, FIGO stage, pathological grading, preoperative ascites, malignant ascites, and peritoneal implantation were E_2 , P_3 , and rq_{cross} . The ROC curve indicates that the E_2 -FIGO classification specificity was 84.2% with an AUC of 0.725. The polarization parameter E_2 is related to sample depolarization. The mean value of

parameter E_2 was lower for early stages (FIGO stages I-II). It may be due to the small changes of cell organelles in the early stage of tumors. The sensitivity of P_3 -pathological grading was 91.4% with an AUC of 0.753. The polarization parameter P_3 is related to the linear birefringence of the sample. It shows a significant decrease in the mean value for high grade of pathological grading. It may

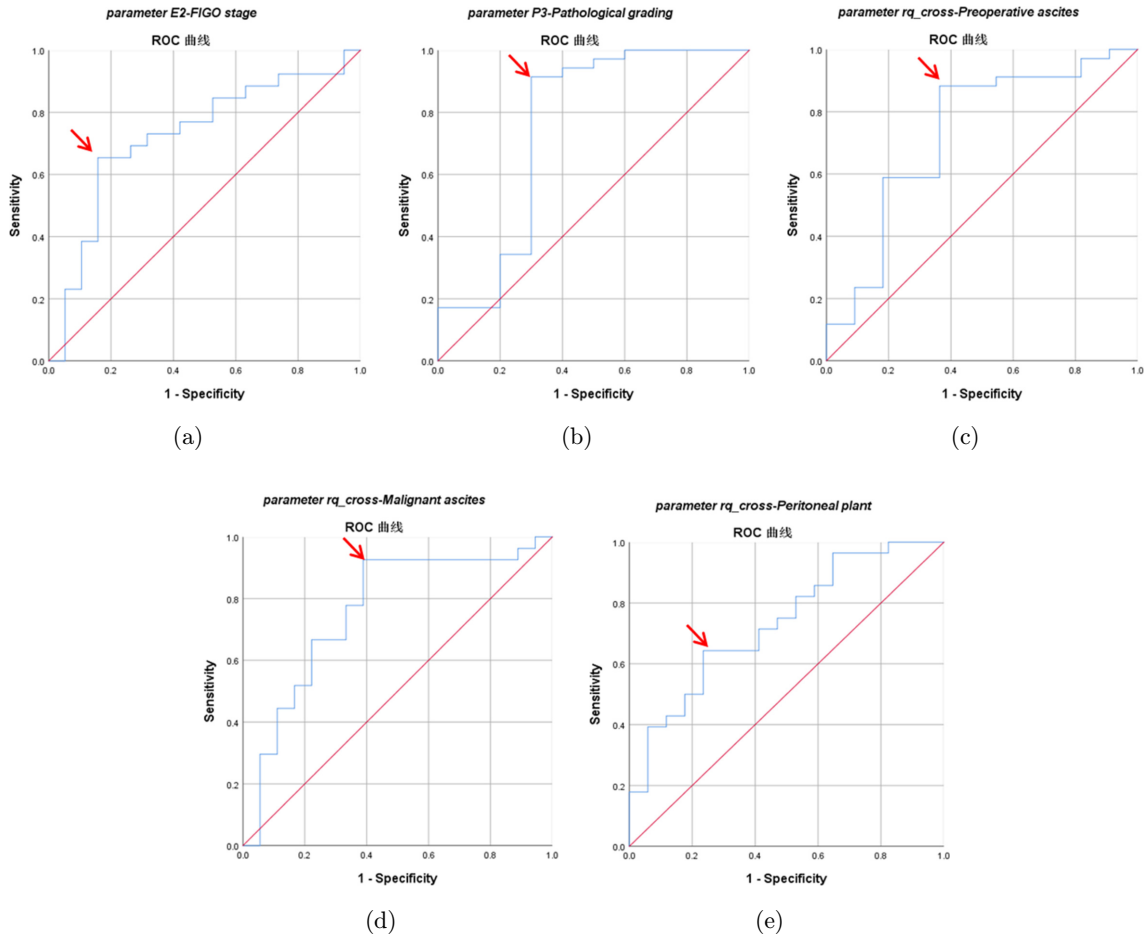


Fig. 6. ROC curves for the mean values of PBPs in predicting the risk of clinicopathological features of SOC patients. (a) The mean value of parameter E_2 predicts the risk of high FIGO stage in SOC. (b) The mean value of parameter P_3 predicts the risk of high grade in SOC. (c) The mean value of parameter rq_{cross} predicts the risk of preoperative ascites in SOC. (d) The mean value of parameter rq_{cross} predicts the risk of malignant ascites in SOC. (e) The mean value of parameter rq_{cross} predicts the risk of peritoneal implantation in SOC.

related to the change in the tumor cell structure reducing the birefringence of the tissue. The rq_{cross} -preoperative ascites and rq_{cross} -malignant ascites have sensitivities of 88.2% and 92.6% with AUCs of 0.727 and 0.759, respectively. The parameter rq_{cross} is related to the destruction of transpose symmetry. Its mean value decreases with the appearance of preoperative ascites or malignant ascites. The specificity for rq_{cross} -peritoneal implantation was 76.5%. The mean value of rq_{cross} increases for tumors with nonoccurrence of peritoneal implantation. Therefore, the results may indicate that the transpose symmetry changes of certain microstructures of tumor cells may lead to the enhancement of tumor cell proliferation activity and cause tumor cells to invade and metastasize.

The mean value of parameter E_2 predicts the early stage of FIGO in SOC patients with $\text{AUC} = 0.725$. The mean value of parameter P_3 predicts high-grade risk in SOC patients with an $\text{AUC} = 0.763$. Parameter rq_{cross} predicts the risk of preoperative ascites, malignant ascites, and peritoneal implantation with all AUCs being greater than 0.72, which indicates an ideal predictive model. The results demonstrate that the parameters E_2 , P_3 , and rq_{cross} may be the risk factors or protective factors for the clinicopathological features of patients with ovarian serous cancer, and they can also be used as predictive markers for clinical diagnosis and prognosis. The results suggest that in the process of clinical treatment, more attention can be paid to patients with low values of P_3 and rq_{cross} to improve the effectiveness of treatment and reduce the incidence of poor prognosis.

4. Conclusion

The research demonstrated that polarization imaging technology and data extraction method by ML can differentiate SOTs, which can also quantitatively characterize microstructure features in the histological sections. We calculated 65 PBP for each ROI which was marked by the pathologist, and derived PFPs which are the optimal linear combinations of PBPs by the ML algorithm. The performance results of PFPs in characterizing EMFs are satisfactory. The average accuracy of PFP could reach 87.7% on the binary classification, while the accuracy of PFP on ternary classification was 75%. Polarization imaging technology may be an

assistive tool for pathological diagnoses to provide multidimensional microstructure information for H&E staining and diagnosis more accurately. Therefore, the integration of polarization imaging technology with the ML method represents an accurate and promising methodology to characterize the complexity of biological tissues concerning human diseases, including cancer.

The pathogenesis and mechanism of ovarian cancer are complex and still under debate. It is difficult to diagnose until in advanced states. If patients are combined with risk factors such as tumor metastasis, high FIGO stage, and high grade, it will often lead to poor treatment effect and unsatisfactory survival rate. Hence, the early prediction of disease progression in patients with ovarian cancer is of positive significance in promoting survival and improving prognosis. In the study, the polarization parameters E_2 , P_3 , and rq_{cross} are closely related to the clinicopathological features and prognosis of patients with serous ovarian carcinoma. These correlations may come from the changes in tumor cell microstructure in the process of occurrence, development, invasion, and metastasis of serous ovarian carcinoma. The mean values of these polarization parameters show prominent potential for the prognosis of patients with serous ovarian cancer and provide valuable new research ideas to improve the accuracy and effectiveness of clinical diagnosis and prognosis. Normal human ovarian tissue samples cannot be easily obtained for research due to ethical considerations and limitations of the sample source.²⁸ Therefore, we only identified three types of SOTs in this study. In the future, when conditions permit, we will incorporate normal ovarian tissue cases into the study and construct a four-classification model of SOTs. Moreover, in the future, further investigations are necessary to reveal the relationships between these polarization parameters and the pathogenesis of ovarian cancers.

Conflict of Interest

The authors have no conflicts of interest relevant to this paper.

Acknowledgments

This work was supported by the Guangming District Economic Development Special Fund (2020R01043).

References

1. K. Tanha *et al.*, “Investigation on factors associated with ovarian cancer: an umbrella review of systematic review and meta-analyses,” *J. Ovarian Res.* **14**(1), 153 (2021).
2. D. Žilovič *et al.*, “Future screening prospects for ovarian cancer,” *Cancers* **13**(15), 3840 (2021).
3. F. Aeffner *et al.*, “The gold standard paradox in digital image analysis: manual versus automated scoring as ground truth,” *Arch. Pathol. Lab. Med.* **141**(9), 1267–1275 (2017).
4. M. K. Santos *et al.*, “Artificial intelligence, machine learning, computer-aided diagnosis, and radiomics: advances in imaging towards to precision medicine,” *Radiol. Bras.* **52**, 387–396 (2019).
5. P. T. Kroegeer, Jr., R. Drapkin, “Pathogenesis and heterogeneity of ovarian cancer,” *Curr. Opin. Obstet. Gynecol.* **29**(1), 26–34 (2017).
6. M. A. Gavrielides *et al.*, “Selection of representative histologic slides in interobserver reproducibility studies: insights from expert review for ovarian carcinoma subtype classification,” *J. Pathol. Inform.* **12**(1), 15 (2021).
7. E. P. Wibowo *et al.*, “An improved calibration technique for polarization images,” *IEEE Access* **7**, 28651–28662 (2019).
8. H. He *et al.*, “Mueller matrix polarimetry — an emerging new tool for characterizing the microstructural feature of complex biological specimen,” *J. Lightw. Technol.* **37**(11), 2534–2548 (2019).
9. Y. Dong *et al.*, “Quantitatively characterizing the microstructural features of breast ductal carcinoma tissues in different progression stages by Mueller matrix microscope,” *Biomed. Opt. Express* **8**(8), 3643–3655 (2017).
10. W. Sheng *et al.*, “Quantitatively characterizing microstructural variations of skin tissues during ultraviolet radiation damaging process based on Mueller matrix polarimetry,” *Proc. SPIE* **10467**, 104670Y (2018).
11. A. Hou *et al.*, “Polarimetry feature parameter deriving from Mueller matrix imaging and auto-diagnostic significance to distinguish HSIL and CSCC,” *J. Innov. Opt. Health Sci.* **15**(1), 2142008 (2022).
12. Y. Dong *et al.*, “A polarization-imaging-based machine learning framework for quantitative pathological diagnosis of cervical precancerous lesions,” *IEEE Trans. Med. Imaging* **40**(12), 3728–3738 (2021).
13. Z. Gong, P. Zhong, W. Hu, “Diversity in machine learning,” *IEEE Access* **7**, 64323–64350 (2019).
14. T. Huang *et al.*, “Fast Mueller matrix microscope based on dual DoFP polarimeters,” *Opt. Lett.* **46**(7), 1676–1679 (2021).
15. J. Chang *et al.*, “Division of focal plane polarimeter-based 3l DoFP polarimeterstitative pathological diagnosis of cervical precancerous lesionsCCMueller,” *J. Biomed. Opt.* **21**(5), 056002 (2016).
16. S. Y. Lu, R. A. Chipman, “Interpretation of Mueller matrices based on polar decomposition,” *J. Opt. Soc. Am. A* **13**(5), 1106–1113 (1996).
17. H. He *et al.*, “A possible quantitative Mueller matrix transformation technique for anisotropic scattering media/Eine mögliche quantitative Müller-Matrix-Transformations-Technik für anisotrope streuende Medien,” *Photonics Lasers Med.* **2**(2), 129–137 (2013).
18. L. V. Wang, D. A. Zimnyakov, *Optical Polarization in Biomedical Applications*, Springer, New York (2006).
19. P. Li *et al.*, “Polaromics: deriving polarization parameters from a Mueller matrix for quantitative characterization of biomedical specimen,” *J. Phys. D, Appl. Phys.* **55**, 034002 (2021).
20. P. Li *et al.*, “Separating azimuthal orientation dependence in polarization measurements of anisotropic media,” *Opt. Express* **26**(4), 3791–3800 (2018).
21. E. S. Fry, G. W. Kattawar, “Relationships between elements of the Stokes matrix,” *Appl. Opt.* **20**(16), 2811–2814 (1981).
22. L. Pillay, R. Wadee, “A retrospective study of the epidemiology and histological subtypes of ovarian epithelial neoplasms at Charlotte Maxeke Johannesburg Academic Hospital,” *South. Afr. J. Gynaecol. Oncol.* **13**(1), 26–35 (2021).
23. I. Ahmad *et al.*, “*Ex vivo* characterization of normal and adenocarcinoma colon samples by Mueller matrix polarimetry,” *J. Biomed. Opt.* **20**(5), 056012 (2015).
24. E. Du *et al.*, “Mueller matrix polarimetry for differentiating characteristic features of cancerous tissues,” *J. Biomed. Opt.* **19**(7), 076013 (2014).
25. N. Gica *et al.*, “Management of borderline ovarian tumors: Series of case report and review of the literature,” *Indian J. Surg.* **83**(3), 617–624 (2021).
26. W. G. McCluggage, “The pathology of and controversial aspects of ovarian borderline tumours,” *Curr. Opin. Oncol.* **22**(5), 462–472 (2010).
27. B. Acs, M. Rantalainen, J. Hartman, “Artificial intelligence as the next step towards precision pathology,” *J. Intern. Med.* **288**(1), 62–81 (2020).
28. W. H. Parker, Indications for oophorectomy and adnexectomy, *Hysterectomy*, I. Alkatout, L. Mettler, Eds., pp. 155–159, Springer, Cham (2018).