

## 基于全偏振显微成像的数字病理技术

姚悦<sup>1,2</sup>, 裴浩杰<sup>1,2</sup>, 李浩<sup>3</sup>, 万嘉晨<sup>1,2</sup>, 陶丽丽<sup>3</sup>, 马辉<sup>1,2\*</sup><sup>1</sup>清华大学深圳国际研究生院, 广东 深圳 518055;<sup>2</sup>广东省偏振光学检测与成像工程技术研究中心, 广东 深圳 518055;<sup>3</sup>北京大学深圳医院病理科, 广东 深圳 518036

**摘要** 数字病理技术利用经过数字化的病理样本显微图像及其特征, 并配合人工智能技术, 实现生物组织病变特征的定量评估和判定, 辅助临床医生做出诊断结论。利用偏振光照明和偏振探测可以实现全偏振成像, 图像每个像素的偏振特征都包含更加丰富的信息, 特别是普通光学成像难以获得的亚细胞超分辨微观结构特征信息, 可为病变组织的识别和定量评估提供更为有效的手段。本文总结了全偏振成像技术, 并结合典型临床应用归纳总结了全偏振显微图像的数据分析方法和最新进展。

**关键词** 医用光学; 全偏振显微成像; 偏振数字病理; 偏振特征提取; 机器学习

中图分类号 O436 文献标志码 A

DOI: 10.3788/CJL231462

## 1 引言

临床医学中病理学检查是许多疾病诊断的金标准。传统病理学检查需要病理医生在显微镜下观察经过染色的实体病理样本, 评估其特异性结构特征, 提供诊断疾病和了解病变机制的依据。但医生对病理样本进行显微镜下的观察和评估往往具有主观性, 会受到个人训练水平、工作经验和现场状态的影响。正在兴起的数字病理技术旨在利用数据技术对病理切片样本进行数字化信息采集、分析和应用, 从而降低病理医生的工作强度, 提升工作效率和病理诊断准确性<sup>[1-4]</sup>。

数字病理系统应包含三个基本功能: 病理样本的数字化、病变特征数字化和人工智能辅助诊断<sup>[1]</sup>。目前的数字病理系统主要基于图像技术: 经过染色的病理样本首先由高分辨扫描仪分块进行彩色成像, 并通过图像算法拼接成一整张全玻片图像(WSI)。这种数字化病理图像方便进行存储、网上传输, 以及进行图像操作和显示, 病理医生可以身处不同应用现场, 通过网络和显示器进行阅片, 实现远程交流和会诊。利用快速发展的数据技术, 如深度学习, 可以预先对大量经过人工标注的数字化病理样本进行深入细致的图像分析, 准确识别特定细胞类型、组织学特征和形态学模式, 获得对应特定病变的特征信息, 并分割出具有病理诊断意义的特征区域。最后, 在人工智能技术的协助下, 病理医生针对待测样本图像, 利用已经获得的病变

特征指标和模型, 结合医生自己的专业知识和临床经验, 识别疑似病变区域, 形成诊断结论<sup>[1, 5]</sup>。

偏振成像是一种正在迅速发展的光学成像技术<sup>[6-10]</sup>。利用包含 4 个分量的全偏振光进行照明和全偏振态检测, 可以获得由 16 幅图像组成的全偏振图像, 图像中每个像素都对应一个 16 维的  $4 \times 4$  偏振变换矩阵, 称之为缪勒矩阵, 它可以完备描述样本在该像素位置的偏振特征, 因此缪勒矩阵图像也称为全偏振图像。偏振特征同样本的光学性质和微观结构特征密切相关<sup>[7]</sup>, 对于生物组织这类具有散射性质的复杂样本, 全偏振图像所包含的信息量远远大于普通非偏振图像。缪勒矩阵不但可以获得散射颗粒和间质的各种光学参数, 如折射率、双折射、吸收系数和二向色性等, 而且更重要的是可以获得各种散射颗粒的几何特征信息, 如粒径、形状、取向、排列、表面形貌、内部结构, 以及浓度分布等, 并且散射颗粒的尺度可以远低于探测波长<sup>[6-7, 11-12]</sup>。由于缪勒图像的每个像素都包含亚细胞层次超分辨微观结构信息, 缪勒成像获得的信息量急剧增加, 并且在单像素和图像等不同层次都隐藏样本特征, 可以大幅度提升医生对生物组织物理特征进行细致表征和定量评估的能力。并且, 全偏振成像不依赖染色, 不对样本产生损伤, 适用于各种染色和非染色样本, 甚至可以对活细胞进行动态监测。尽管仍在萌芽阶段, 基于偏振成像的数字病理技术已经展示出一系列独特的优势和诱人的应用潜力<sup>[6-7, 13-14]</sup>。本文将

收稿日期: 2023-12-01; 修回日期: 2024-01-27; 录用日期: 2024-01-31; 网络首发日期: 2024-02-20

基金项目: 国家自然科学基金(62375152, 11974206, 61527826, 11374179, 11174178, 10974114, 60778044)、深圳市基础研究学科布局项目(JCYJ20170412170814624)

通信作者: \*mahui@tsinghua.edu.cn

介绍全偏振显微成像方法和偏振特征提取方法,并通过一些初步应用实例展示该方法的临床辅助诊断潜力。

## 2 全偏振显微成像技术

### 2.1 缪勒成像基础知识

偏振是光的基本属性之一。任意光的偏振态可用斯托克斯矢量  $S$  来表达,其由四个光强度量组成,其中,  $I$  为总光强,  $Q$ 、 $U$  和  $V$  分别代表水平和垂直线偏振分量的强度差、 $\pm 45^\circ$  方向线偏振分量的强度差,以及右旋和左旋圆偏振分量的强度差。

$$S = \begin{pmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \end{pmatrix} = \begin{pmatrix} I \\ Q \\ U \\ V \end{pmatrix} = \begin{pmatrix} I_H + I_V \\ I_H - I_V \\ I_{45^\circ} - I_{-45^\circ} \\ I_R - I_L \end{pmatrix}. \quad (1)$$

缪勒显微成像中,常用斯托克斯矢量-缪勒矩阵表示法描述光的偏振态及其变化。其中,斯托克斯矢量  $S$  描述光的偏振态,缪勒矩阵  $M$  描述介质的偏振光学特性<sup>[7]</sup>。

$$S_{out} = M \cdot S_{in} \quad (2)$$

$$\begin{pmatrix} I_{out} \\ Q_{out} \\ U_{out} \\ V_{out} \end{pmatrix} = \begin{pmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \\ m_{41} & m_{42} & m_{43} & m_{44} \end{pmatrix} \begin{pmatrix} I_{in} \\ Q_{in} \\ U_{in} \\ V_{in} \end{pmatrix},$$

式中:  $S_{in}$  描述入射光偏振态;  $S_{out}$  描述与样本作用后出射光的偏振态。

### 2.2 缪勒显微镜

缪勒矩阵的测量需要改变入射光偏振态,记录散射光的不同偏振分量,通过至少 16 个偏振分量计算缪勒矩阵的 16 个阵元。由于偏振片和波片等偏振器件

不改变光的传播方向,可通过在显微镜的照明光路中加入起偏器(PSG)进行偏振态调制,在散射光的成像光路中加入检偏器(PSA)记录,从而获得一系列对应不同起偏和检偏组合的偏振分量图像,并进一步计算获得缪勒矩阵图像。

采用不同起偏器、检偏器和探测器模块可以实现不同类型的缪勒成像方法,例如:利用固定偏振片和旋转波片组合<sup>[15-16]</sup>,或利用液晶<sup>[17]</sup>和光弹<sup>[18]</sup>等无运动部件的不同类型起偏器和检偏器组合;利用普通单色 CCD 进行光强成像<sup>[15-16]</sup>,以及利用单个<sup>[19-20]</sup>或者一对<sup>[12]</sup>线偏振相机实现多个偏振分量同时成像。不同硬件配置和偏振调控模式会对应不同的误差传递规律,需要发展相应的系统优化和降噪、校准方法<sup>[21-22]</sup>,提高对应各个像素的缪勒矩阵精度。

Oldenbourg<sup>[17]</sup>在偏振显微镜中引入电控液晶相位延迟器,该装置可精确测量样本的双折射性质,后期经持续改进推出了商业化的 Abrio 系统<sup>[23-24]</sup>。Gribble 等<sup>[18]</sup>在检偏器中使用双光弹搭建了缪勒矩阵成像装置,并通过实验证实了测量获取最稳定缪勒矩阵的最佳输入斯托克斯矢量。Arteaga 等<sup>[16]</sup>设计实现了一种全偏振缪勒矩阵显微成像装置,该装置连续旋转起偏器和检偏器的波片,通过数字解调重建样品缪勒矩阵。在国家自然科学基金委员会仪器项目支持下,本团队基于国产商业显微镜发展了一系列模块化缪勒显微镜<sup>[25-26]</sup>,包括透射正置<sup>[12, 15]</sup>、透射倒置和背反式<sup>[25, 27]</sup>,可分别用于静止薄样本、活细胞和厚组织样本。

图 1 为本团队研发的两种透射式缪勒显微镜结构图,分别基于双旋转波片<sup>[15]</sup>和双线偏振相机<sup>[12]</sup>。如图 1(a)所示,基于双旋转波片的缪勒显微镜起偏器采用固定偏振片(P1)和可旋转的四分之一波片

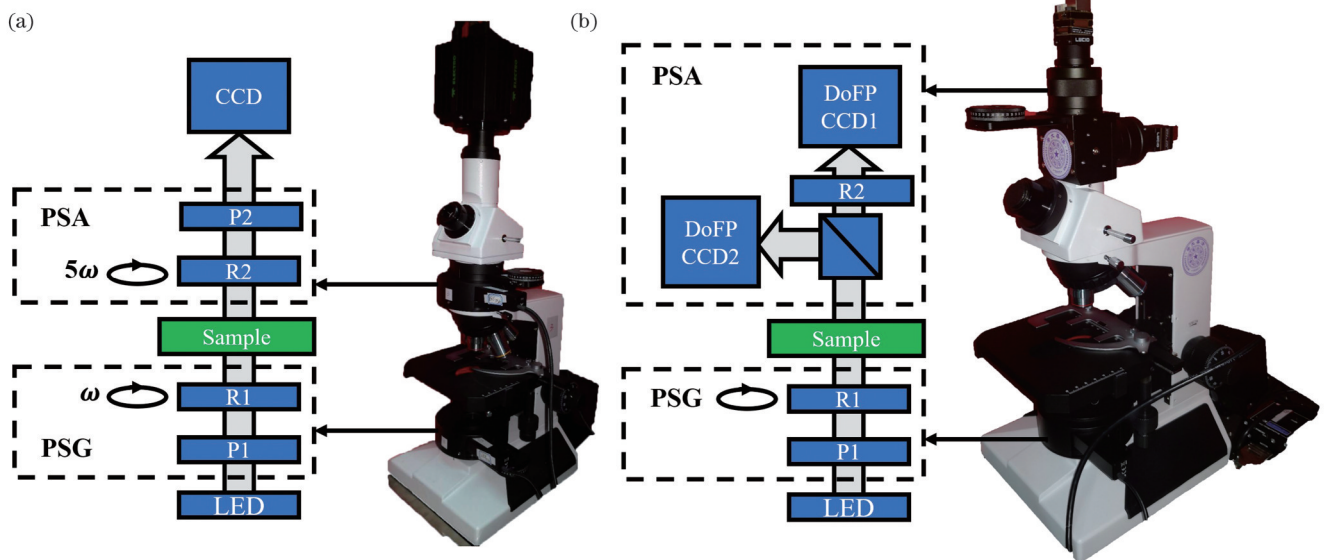


图 1 两种透射式缪勒显微镜结构图<sup>[12]</sup>。(a)基于双旋转波片的缪勒显微镜;(b)基于双分焦平面线偏振相机的缪勒显微镜  
Fig. 1 Configuration of two types of transmission Mueller matrix microscopes<sup>[12]</sup>. (a) Dual rotating retarders-based Mueller matrix microscope; (b) dual division of focal plane polarimeters-based Mueller matrix microscope

(R1), 检偏器采用另一个可旋转四分之一波片(R2)和固定偏振片(P2), 探测模块使用灰度 CCD 记录偏振分量图像。该装置通过记录对应 PSG 和 PSA 不同组合的 30 个偏振分量图像, 进一步计算获得缪勒矩阵的 16 个阵元。基于双旋转波片的缪勒显微镜采用非同时性偏振态测量方法, 其散射光的各个偏振分量需通过多次测量获取。分焦平面(DoFP)线偏振相机可通过在成像探测器每组四个像素前覆盖不同角度的微型偏振片实现线偏振态的同时性测量<sup>[12,28]</sup>。为提升缪勒成像速度, 本团队研发了基于双线偏振相机的缪勒显微镜。如图 1(b)所示, 该显微镜起偏器采用固定偏振片(P1)和可旋转的四分之一波片(R1), 检偏器采用两个分焦平面线偏振相机(DoFP CCD)和一个固定的四分之一波片(R2), 从而实现全偏振态的实时测量与快速缪勒矩阵成像。现有基于双线偏振相机的快速缪勒显微镜<sup>[12]</sup>实际可以用不低于 1 frame/s 速度进行缪勒成像, 随着 PSG 调制速度的提升, 其缪勒成像速度可以不断接近理论最大 5 frame/s 的成像速度。该显微镜针对标准样本缪勒阵元最大误差不超过 1%。针对 4  $\mu\text{m}$  病理切片样本和细胞涂片的缪勒成像结果显示, 信噪比足以凸显微弱偏振特征。

### 2.3 偏振参数及其图像

缪勒矩阵的 16 个阵元虽然能完备表征样本的偏振特征, 但单个阵元与样本的光学性质和微观结构特征之间缺乏显性联系, 用于描述样本复杂结构特征时其物理可解释性不强。同时一些缪勒阵元同方位角密切相关<sup>[29]</sup>, 样本取向的变化会直接影响测量结果。因此, 利用缪勒矩阵的物理特性, 可以从缪勒阵元中推导出一系列偏振参数, 它们是基于缪勒阵元的函数, 但具有更明确的物理可解释性<sup>[7,29-30]</sup>。

偏振参数中应用最广的为 Lu-Chipman<sup>[31-32]</sup>提出的一组缪勒矩阵极化分解(MMPD)参数。缪勒矩阵分解方法将缪勒矩阵  $M$  分解为代表相位延迟( $M_R$ )、二向色性( $M_D$ )和散射退偏( $M_\Delta$ )的三个基本子矩阵, 并经过一系列代数运算获得对应物理意义相对明确的偏振参数, 如二向色性相关参数  $D$ 、退偏相关参数  $\Delta$ 、线性相位延迟参数  $\delta$  及其对应角度参数  $\theta$ <sup>[30]</sup>。这些偏振参数具有相对明确的物理意义, 但由于矩阵相乘的结果受矩阵排列顺序影响, 三个子矩阵的不同排序会导致不同的参数值, 实际应用中需要根据样本的情况进行尝试, 选择最接近真实样本状态的排序。

Morio 等<sup>[33]</sup>研究了三个子矩阵可能出现的六种分解顺序之间的关系, 并根据二向色性和散射退偏出现的顺序将其分为两组。Ossikovski 等<sup>[34]</sup>和 Anastasiadou 等<sup>[35]</sup>进一步研究发现  $M=M_\Delta M_R M_D$  和  $M=M_D M_R M_\Delta$  这两种分解顺序在物理表现上更有意义。Ghosh 等<sup>[36]</sup>通过模拟和实验研究了在同时存在双折射-旋光-散射效

应的复杂介质中, 不同缪勒矩阵分解顺序对偏振参数的影响。结果显示, 对于弱二向色性的介质, 缪勒矩阵分解的不同顺序对偏振参数影响不大。实验结果还显示在生物组织中, 二向色性的效应普遍低于其他两个分解参数, 因此目前针对生物组织的研究大多数采用 Lu-Chipman<sup>[31-32]</sup>的  $M=M_\Delta M_R M_D$  分解顺序获取偏振参数。

偏振测量易于获取样本各向异性特征信息, 但部分缪勒阵元会随着样本摆放方位角的变化而变化, 这使得其应用受到影响。研究发现利用缪勒矩阵分解获得的散射退偏、相位延迟和二向色性等参数<sup>[31-32]</sup>, 以及其他许多偏振特征量<sup>[30]</sup>对组织切片的旋转角并不敏感<sup>[37]</sup>。针对缪勒矩阵对称性特征的深入研究表明, 在照明与接收光共线的情况下, 存在一系列对样本旋转角不敏感的偏振参数, 它们代表的各向异性特征的幅度和方位角可以分离<sup>[29]</sup>。近期一篇综述文章总结了 104 个常用的偏振参数<sup>[30]</sup>。但针对不同偏振测量方法与应用, 还会有一些文章提出其他偏振参数, 随着对偏振特征的了解不断加深, 这个偏振特征参数的清单将不断扩展。

缪勒图像的每个像素都能提供样本的特征信息, 并且每一个偏振参数都可以形成一幅偏振图像, 进一步凸显该参数所对应的微观结构的分布特征。如图 2 所示, 退偏相关的参数如  $\Delta$  和  $b$  可以凸显乳腺癌组织细胞核结构, 相位延迟相关参数如  $\delta$ 、 $r_L$  和  $q_L$  可以凸显乳腺癌组织纤维结构<sup>[7,30,38]</sup>。熟悉组织结构染色图像的病理医生可以直接利用不同偏振图像, 识别特定组织结构和病变, 并在单像素层次进行定量计算<sup>[39]</sup>, 也可以利用数据技术对体现病变特征的不同偏振参数进行图像层次分析<sup>[40]</sup>, 进一步获得病变特征, 用于辅助诊断<sup>[41-43]</sup>。

由于散射过程中偏振态的变化同散射颗粒的几何特征密切相关, 尤其对尺度低于波长的小颗粒散射更敏感, 从偏振特征的变化可以获得超分辨尺度的结构特征, 并且这些特征并不一定会随着成像分辨率的降低迅速消失<sup>[38]</sup>。因此, 偏振成像能获得大量传统光学成像难以获得的微观结构特征信息, 它们都可以用于生物组织的特征表征和细分。

## 3 偏振特征提取方法

### 3.1 基于监督学习的偏振特征提取方法

全偏振显微图像中包含病理组织丰富的微观结构信息, 但如何从各个像素的 16 维缪勒矩阵数据中提取偏振特征, 并识别其对应的组织结构特征, 仍然是偏振技术应用于病理诊断的首要挑战。近年来有越来越多的课题组开始探索偏振数字病理在临床诊断中的应用前景。偏振数字病理通过机器学习提取病变组织的偏振特征并定量表征病变的微观结构, 辅助医生提升临床诊断与预后评估的准确率与效率。偏振数字病理所

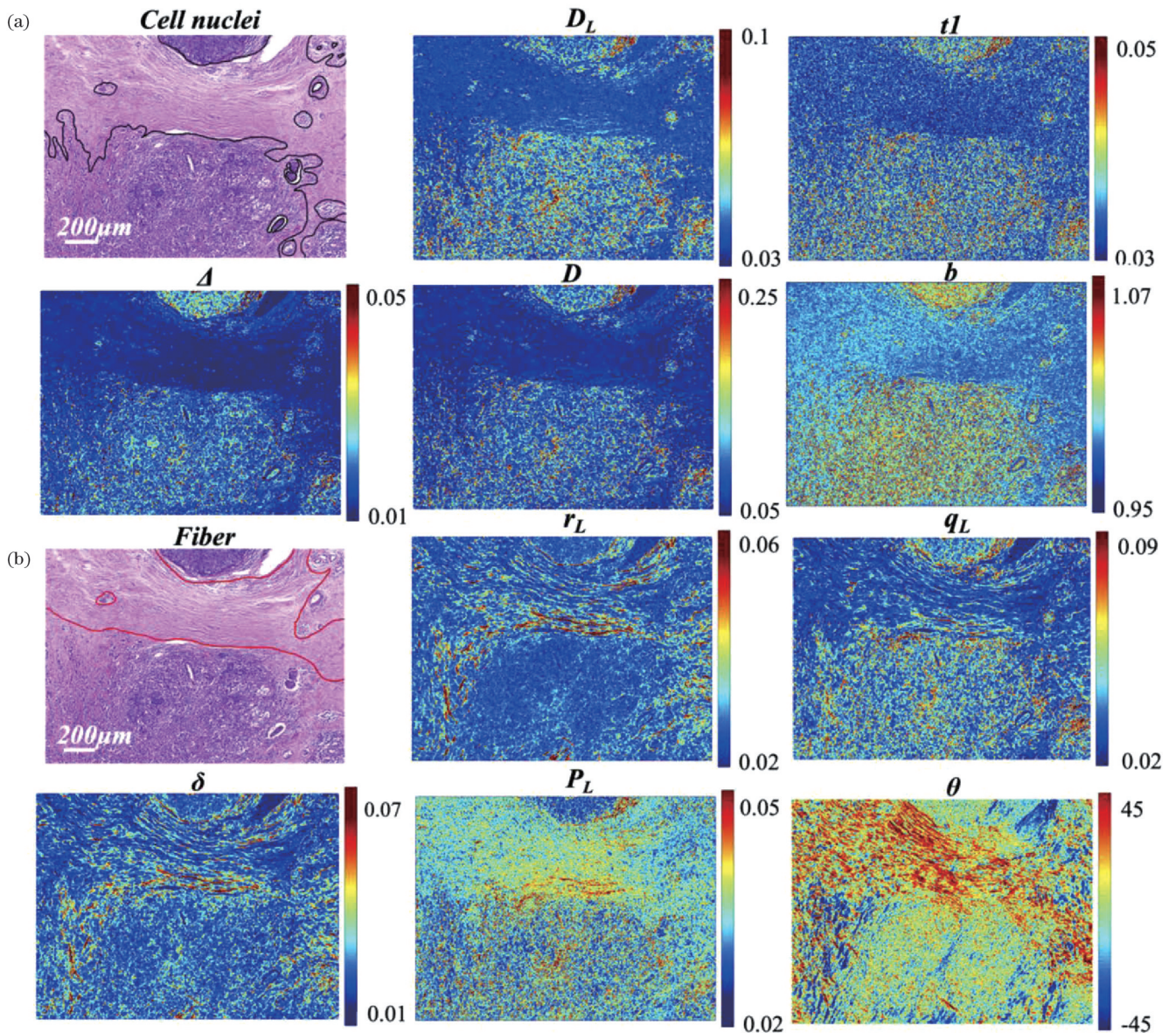


图 2 乳腺癌组织的偏振参数图像<sup>[38]</sup>。(a)偏振参数  $D_L$ 、 $t_1$ 、 $\Delta$ 、 $D$  和  $b$  与组织中细胞核相关(H&E 染色图中黑线标记区域);(b)偏振参数  $r_L$ 、 $q_L$ 、 $\delta$ 、 $P_L$  和  $\theta$  与组织中纤维相关(H&E 染色图中红线标记区域)

Fig. 2 Images of polarization parameters of breast cancer tissues<sup>[38]</sup>. (a) Parameters  $D_L$ ,  $t_1$ ,  $\Delta$ ,  $D$ , and  $b$  correspond to the cell nuclei (labeled by the black solid line in the H&E stained image); (b) parameters  $r_L$ ,  $q_L$ ,  $\delta$ ,  $P_L$ , and  $\theta$  correspond to the fiber tissue (marked by the red solid line in the H&E stained image)

采用的分析手段之一是监督学习,通过对病变组织样本进行全偏振显微成像,并对照病理医生的清晰标注,提取病变所对应的像素级或图像级偏振特征,用于定量识别病变微观结构特征。

病理切片的穆勒矩阵图像中每个像素都可以看作包含病理组织微观结构信息的独立数据,因此可以从像素角度出发,利用各种不同类型的监督学习方法,对医生标记病变区域内偏振像素进行分类,提取病变对应的偏振特征,实现对复杂病变结构的特异性定量表征。例如:Luu等<sup>[44]</sup>使用机器学习分类器如随机森林、多层感知机等对来自 32 张病理切片,4 种皮肤癌组织的 669 个位点的穆勒阵元像素进行分

类,获得平均 93% 的 4 分类准确率;Ivanov等<sup>[45]</sup>使用主成分分析、逻辑回归等对正常与结肠癌组织的 670 个位点的偏振参数像素进行分类,最佳组合获得的准确率为 97%,灵敏度为 100%,特异度为 93%;Dong等<sup>[46]</sup>将穆勒阵元和其他偏振参数看作用于描述不同简单微观结构和光学性质的偏振基底参数(PBP),然后使用线性判别模型最终筛选出 8 个 PBP,并将其线性组合为一个对特定微观结构敏感的偏振特征参数(PFP),从而实现对不同乳腺癌组织中细胞核、有序纤维和无序纤维结构对应像素的特异性识别。针对健康与 3 种乳腺癌组织,用 PFPs 识别细胞核、有序纤维和无序纤维的准确率为 82%~91%,灵敏度为

86%~98%。此外,提取的 PFP 由 8 个 PBP 线性组成,可针对微观结构提供物理可解释性。还可如图 3 所示,Dong 等<sup>[47]</sup>使用双模态模型,首先用 U-Net 分割宫颈上皮区域辅助偏振特征训练,然后提出基于统

计距离的机器学习模型推导出新的 PFP,用于对正常与 3 期宫颈上皮不典型增生对应像素的病变特征识别,获得的准确率、精准率和召回率分别为 90%、92% 和 89.9%。

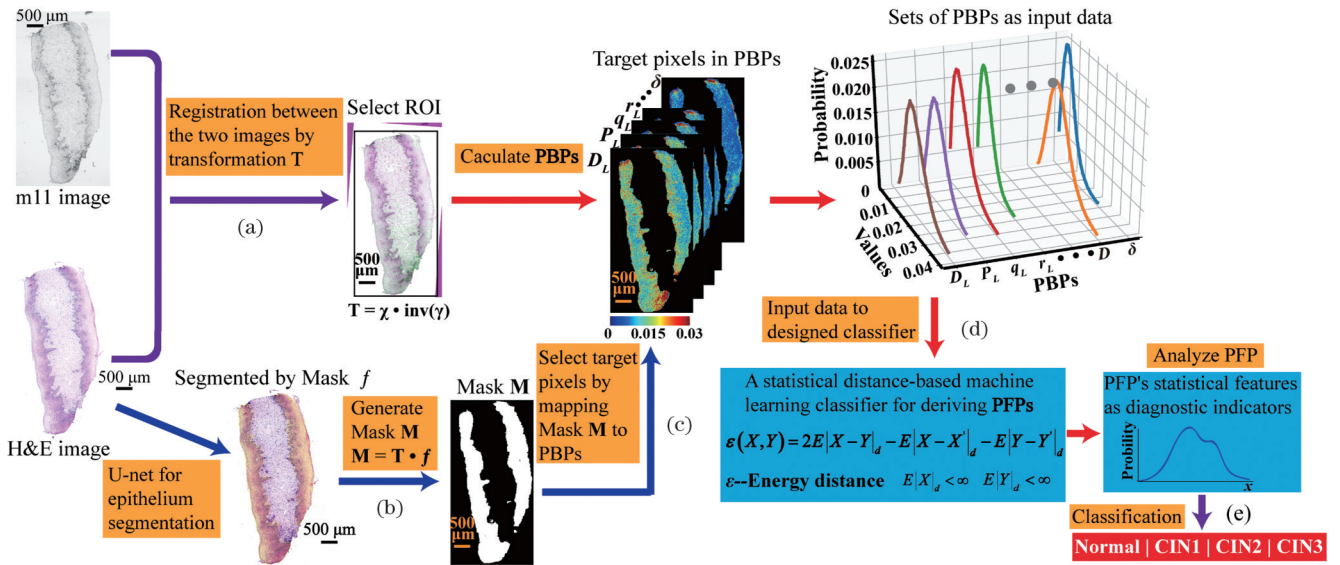


图 3 双模态模型提取宫颈癌前病变像素偏振特征<sup>[47]</sup>。(a) 穆勒矩阵光强图 m11 和 H&E 图像配准;(b) U-Net 分割 H&E 图像中宫颈上皮区域,生成掩模 M;(c) 将 M 映射到偏振基底参数上选择目标偏振像素;(d) 输入目标偏振像素到基于统计距离的机器学习模型中,提取偏振特征参数;(e) 使用 PFP 区分正常与 3 期宫颈癌前病变

Fig. 3 Dual-modality machine learning framework for pixel level polarization feature extraction of cervical precancerous tissues<sup>[47]</sup>. (a) Mueller matrix light intensity image m11 and H&E image registration; (b) U-Net segmentation of cervical epithelial region in H&E image to generate mask M; (c) mapping M to PBPs to select target pixels; (d) derive a PFP by inputting target pixels in to a statistical distance-based machine learning model; (e) using PFP to distinguish normal and 3 stages of cervical precancerous lesions

具有特定偏振特征的像素都可以构成新的图像,利用各种图像分析技术可以提取偏振图像的特征来进一步提高全偏振成像定量识别和表征病变结构的能力,包括“以深度学习为核心的分析方法”和“基于图像组学的特征提取和分析方法”。例如:以深度学习为核心,Roa 等<sup>[48]</sup>使用卷积神经网络和 K-nearest neighbor (K-NN)对偏振基底参数图像进行分类,实现对小鼠宫颈中弹性蛋白的识别准确率达 90% 左右,胶原纤维的识别准确率达 80% 以上,该结果可辅助宫颈重塑与早产的研究;McKinley 等<sup>[49]</sup>结合卷积神经网络与机器学习算法提取偏振基底参数图像特征,用于识别健康与肿瘤脑组织的白质纤维束,希望辅助外科医生在手术中最大限度切除肿瘤的同时保持神经功能<sup>[50]</sup>;Yao 等<sup>[42]</sup>使用逻辑回归、岭回归和两种卷积神经网络对增生期与分泌期子宫内膜的不同 PBP 图像进行特征识别,结果显示卷积神经网络识别二向色性相关偏振参数  $\alpha_p$  可获得最高准确率 87%,精准率和召回率分别为 79% 和 100%;Chen 等<sup>[51]</sup>研究基于偏振特征参数图像的融合网络,用于肺癌、乳腺癌和肝癌共 3 种癌变数据的识别,准确率分别为 98.4%、87.09% 和 84.89%。该研究显示,针对肺癌,训练集缩小为原始数据的 20% 时,使用同样模型,H&E 染色图像的分类准确率由

94.46% 降为 87.58%,偏振参数图像的准确率由 98.4% 降为 94.23%,显示出偏振在小样本量时有辅助诊断优势。相比于深度学习模型,基于“偏振图像特征提取”的方法可提供病变结构的物理可解释性。因此还可如图 4 所示,Yao 等<sup>[43]</sup>使用基于偏振图像特征的组学流程来获取丰富且相对可解释的 4 期肝纤维化微观结构特征用于识别病变,可获得 94.6% 的 4 分类准确率<sup>[43]</sup>。该研究还通过初步对比偏振数字病理与图像数字病理,显示出偏振在低分辨、宽视场应用场景中的优势。

当采用监督学习方法提取全偏振图像的像素或图像级特征时,需要在对大量精确标注的样本进行学习的基础上建立模型或神经网络,而病理标记过程需耗费临床医生大量人工,并且标注质量同病变组织特征、样本制备质量,以及标注医生本人素质和现场发挥都有密切关系,获得高质量样本精确标注的病变组织数据库往往是限制监督学习方法发挥效果的瓶颈。

### 3.2 无监督学习与偏振像素聚类

生物组织非常复杂,涉及不同类型和尺度的微观结构成分,它们分别对应不同的偏振特征。在利用监督学习提取偏振特征时,获取精细且高质量标注数据的过程需要医生耗费大量时间与人工。当训

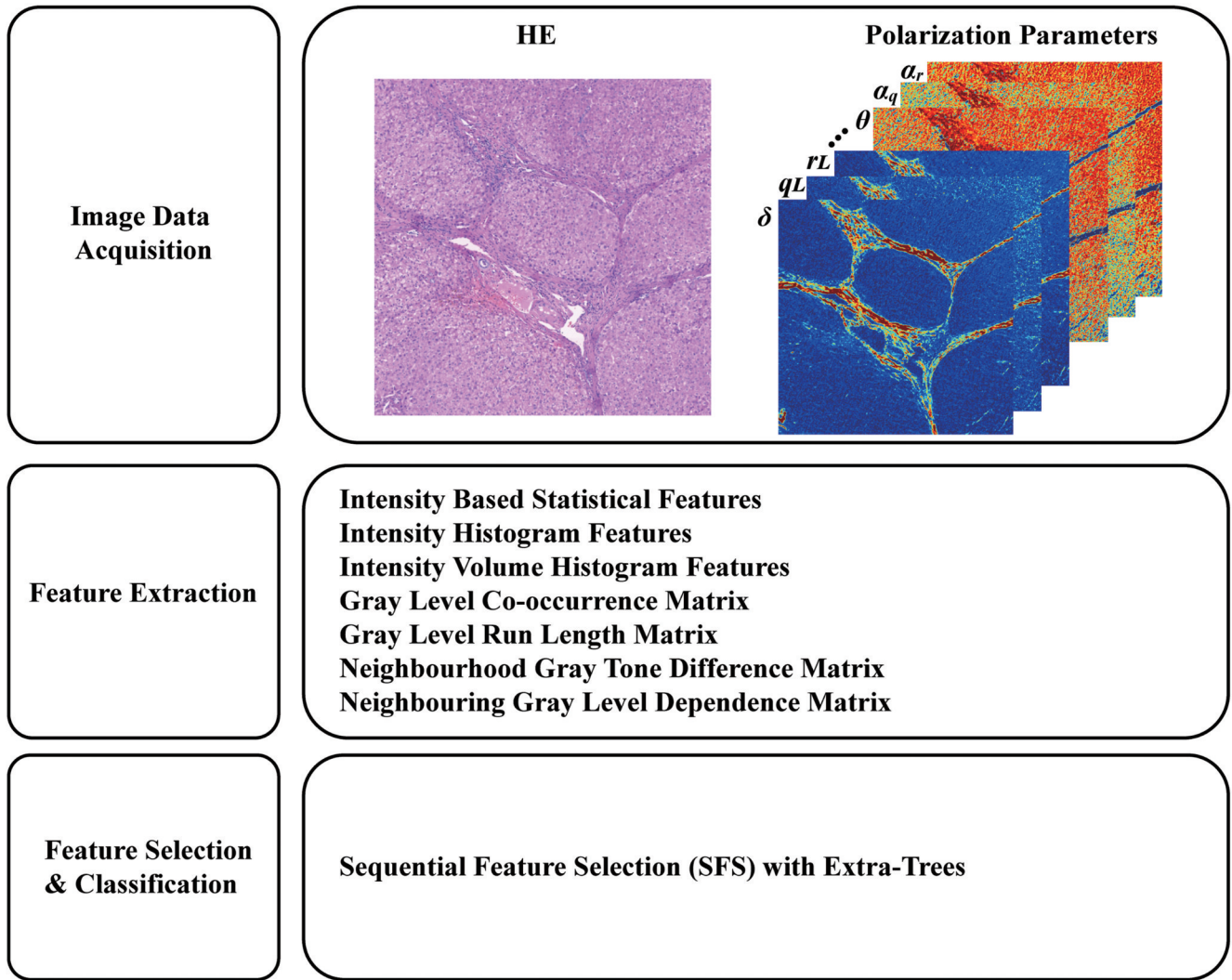


图 4 基于偏振图像特征的组学流程<sup>[43]</sup>

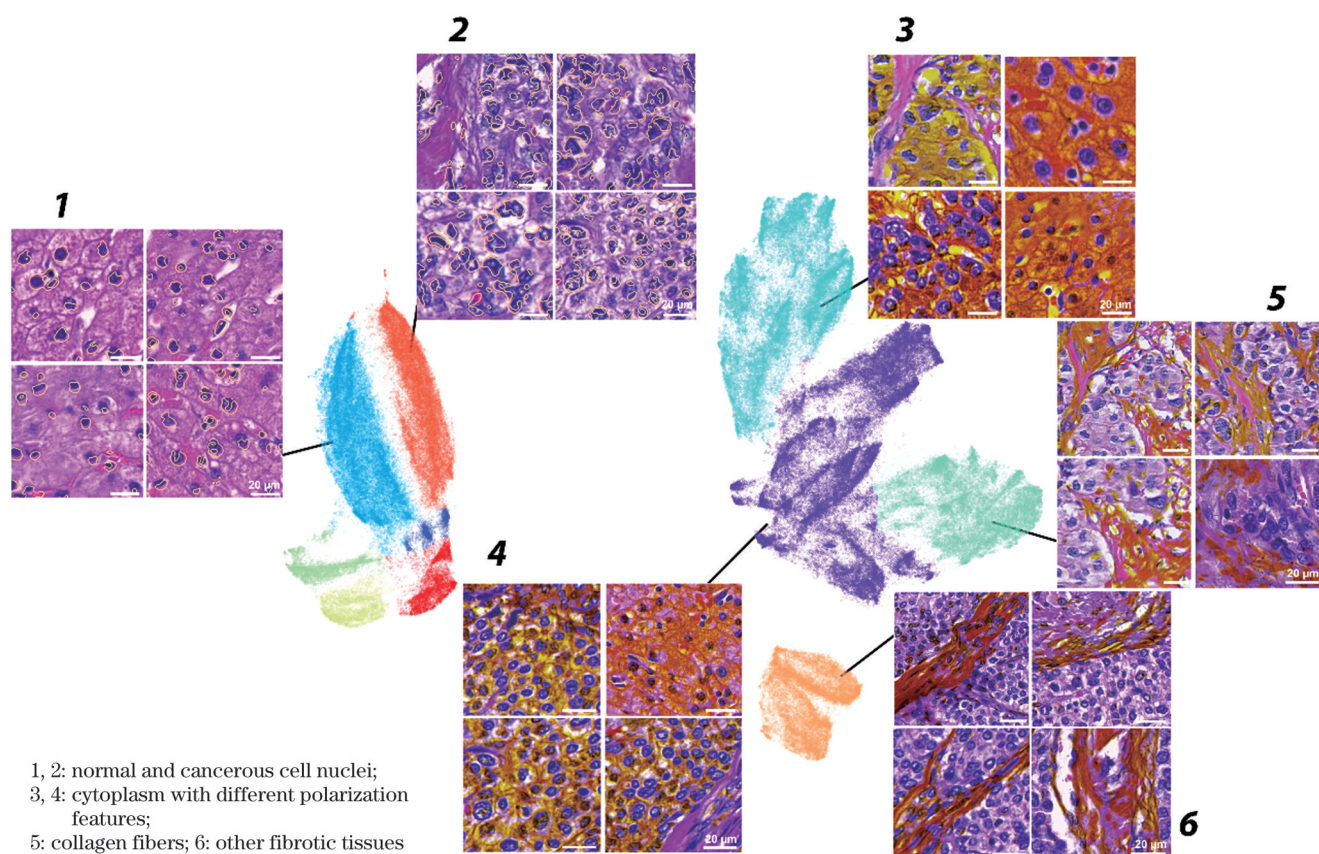
Fig. 4 Polarization parameter images based radiomics approach<sup>[43]</sup>

练数据有限,或者医生的标注不够精细和准确时,监督学习容易产生过拟合,所得病变偏振特征虽然在训练数据上表现相对较好,但对未知样本病变特征进行像素级识别时难以达到理想的效果。无监督学习不需要预先标注训练数据,有利于挖掘偏振像素中存在的隐藏特征。利用无监督学习对缪勒图像的像素进行聚类分析,可以分离并识别不同偏振特征及其对应的微观结构特征,从而把复杂生物组织拆分成一系列具有简单特征结构的基本组分,并在不同成分的基础上分析并凸显不同类型病变和不同病变分期的特征<sup>[52]</sup>。

Wan 等<sup>[52]</sup>对来自 41 位肝癌患者病理组织的 222 个感兴趣区域(ROI)的缪勒图像进行了像素聚类。在 20×物镜下,每幅缪勒图像的像素量超过 400 万。该研究利用统一流形逼近与投影(UMAP)<sup>[53]</sup>方法对所有缪勒像素做非线性降维和聚类,从而在一个二维空间得到可视化聚类散点图显示(图 5)。UMAP 是一种新的降维、聚类和可视化技术,其理论基础是黎曼几何和代数拓扑,它能够尽可能多地保留数据的全局结构,

易于应用于较大维度的数据集,并且速度较快,对于处理高维偏振空间大量像素数据有独特的优势。该研究选择 H&E 染色切片,做缪勒图像并从像素的 UMAP 结果中选择不同聚类像素映射到对应 H&E 染色图像上。病理医生可以根据高分辨 H&E 染色图像识别出不同偏振像素聚类所对应的组织结构特征,如图 5 所示,聚类 1 和 2 分别对应肝组织正常和病变细胞核,3 和 4 对应不同偏振特征的细胞浆,5 和 6 则对应两类不同的胶原纤维结构等。

该研究显示利用这些不同聚类特征及其所对应的微观结构,有利于抑制病变不敏感的成分在图像中的贡献,有效凸显病变特征,从而明显提升识别和分类的准确度。如图 6(a)所示,癌变组织中类别 2 与类别 6 的密度明显大于正常肝组织。如图 6(b)与图 6(c)所示,肝组织不同 ROI 中类别 2 与 6 的面积占比可以区分病变区域,并且其占比与细胞分化程度相关,有潜力以它们作为偏振标志物识别癌变组织。使用类别 2 作为潜在偏振标志物,可以有效对正常与癌变 ROI 区域进行区分,其曲线下面积(AUC)为

图 5 肝癌组织 222 个区域穆勒图像的像素聚类<sup>[52]</sup>Fig. 5 Pixel clustering of Mueller matrix images from 222 ROIs of liver cancer tissues<sup>[52]</sup>

94.84%。图 6(d) 还显示, 高分化与中分化癌细胞样本也可以使用特定细胞质结构进行区分识别。该研究说明利用无监督像素聚类可以将组织拆分成细致的微观结构亚类, 并通过分析分离出与病变相关的细致偏振特征亚类, 辅助医生进行病变识别与定量诊断。

进一步研究发现, 类似的微观结构拆分分析可应用于区分不同癌变亚类, 如肝细胞癌与肝内胆管癌。使用胶原纤维亚类作为偏振标志物可以有效区分肝细胞癌与肝内胆管癌, 其 AUC 为 84.94%。同时, 两种肝癌亚类在细胞质结构上也有一定区分度, 其 AUC 为 71.69%。此外, 特定偏振特征的聚类像素还可以再进行聚类分析, 拆分出更加细致的偏振特征, 并获得更加细致的微观结构特征信息。图 5 中类别 1 和 2 的像素对应细胞核结构。该研究对细胞核所对应的偏振像素进一步进行聚类拆分, 细分为细胞核的亚成分。初步结果显示病变细胞核与正常细胞核有不同的偏振特征, 展示了无监督微观结构拆分方法的潜力, 有望进一步分离出对病变敏感的细胞核偏振特征亚类, 提高病变诊断的灵敏度<sup>[52]</sup>。

现有针对不同病理组织和细胞样本进行穆勒成像和像素聚类分析的结果显示, 将生物样本的微观结构拆分成不同特征基本单元的方法具有普适性, 拆分结

果相对稳定。穆勒成像包含的信息远比普通成像方法更为丰富, 并且对超分辨率散射颗粒的微观结构特征十分敏感。这些信息对了解组织微观结构和生物动态过程有重要意义, 但很难通过临床病理科现有成像方法获取。因此, 利用穆勒像素的无监督聚类十分有利于分离生命体系中复杂的微观结构特征信息, 从而识别病变并为进一步提供病理变化的来源和机制提供依据。

### 3.3 偏振超像素和偏振特征模板

穆勒图像所代表的样本偏振特征对应 16 维偏振空间中像素的分布, 图 5 所显示的 UMAP 像素聚类结果代表这些像素在一个经过非线性变换的二维空间上的投影, 这种高维空间特征的低维投影对数据的选取相对敏感。借用图像超像素<sup>[54-55]</sup>的概念, 我们可以利用  $K$ -means 聚类<sup>[56]</sup>方法, 合并偏振空间中特征接近的像素, 从而把穆勒图像的所有像素压缩成一组偏振超像素的集合<sup>[52]</sup>, 在保留偏振信息的同时大幅度降低偏振数据量, 提高偏振特征的计算效率。更重要的是, 目前针对多种病变组织样本的初步分析显示, 对于特定病变组织的穆勒图像, 它们在偏振空间超像素的质心位置基本保持稳定, 因此偏振超像素集合可以近似表达穆勒图像在高维偏振空间的偏振特征。

数字病理的发展需要用大量的标注数据进行模型

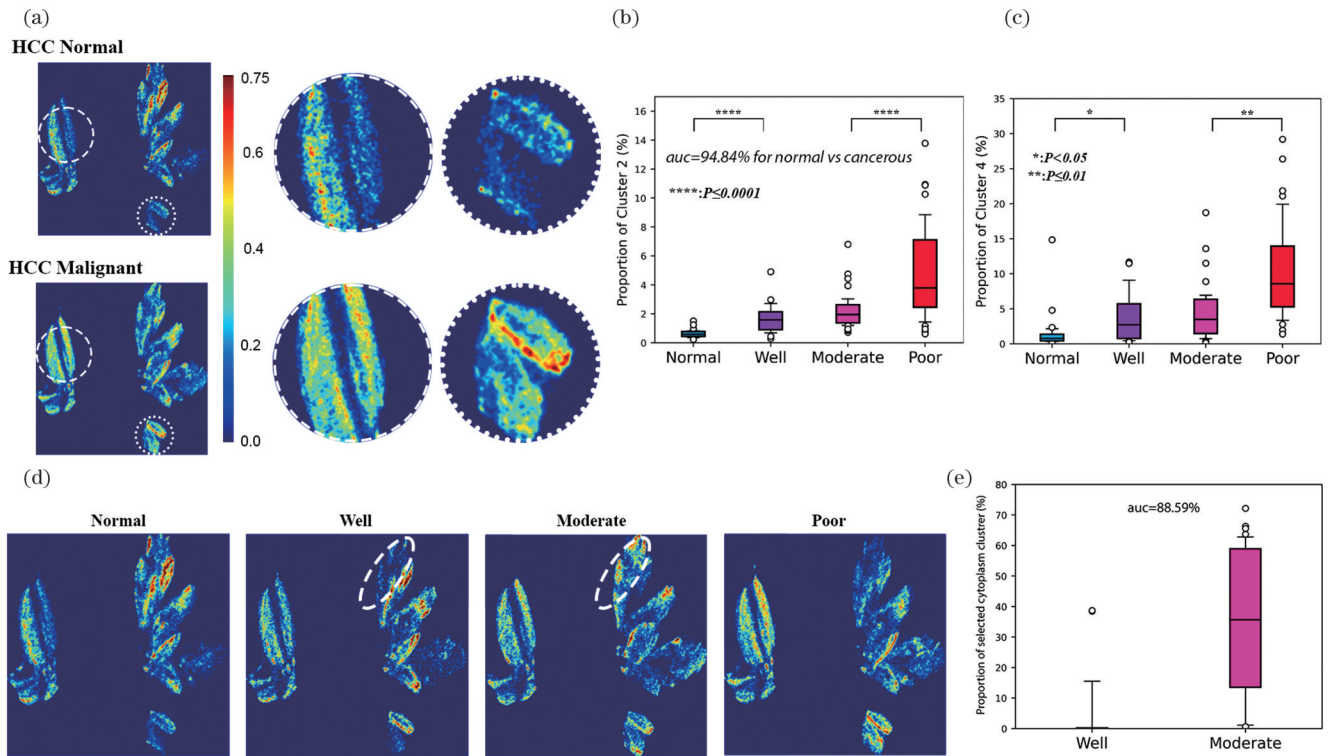


图 6 肝细胞癌中与病变和细胞分化程度相关的偏振标志物提取<sup>[52]</sup>。(a)正常与癌变组织密度热图,放大了类别 2 与 6 两个潜在偏振标志物;(b)针对不同分化程度组织的每个 ROI,算类别 2 像素比例;(c)针对不同分化程度组织的每个 ROI,算类别 6 像素比例;(d)不同分化程度组织热图,虚线为细胞质簇中选择的偏振标志物,用于识别高分化和中分化样本;(e)高分化和中分化样本中,算选择的细胞质簇像素比例

Fig. 6 Identification of polarization markers that correlates with hepatocellular carcinoma (HCC) differentiation degree<sup>[52]</sup>. (a) Density heatmap of normal and malignant HCC, zoomed in on two potential polarization markers, cluster 2 and 6; (b) proportion of pixels belongs to cluster 2 in each ROI are calculated for each differentiation degree; (c) proportion of pixels belongs to cluster 6 in each ROI are calculated for each differentiation degree; (d) density heatmap for each differentiation degree, the dashed line indicates the selected polarization marker in the cytoplasm cluster to differentiate well and moderately differentiated samples; (e) in well and moderately differentiated samples, calculate the pixel proportion of the selected cytoplasm cluster

训练,需要病理医生对 WSI 染色图像进行不同级别的标注,如病理级、区域级和更精细的细胞级别标注。该过程中病理医生需要根据经验在 WSI 染色图像上重复寻找、标注相同的病理特征,如正常和癌变区域,以及更精细的癌变细胞等,非常耗费人工与时间。偏振超像素提供一种细致表征病变组织偏振特征的新方法,可以利用少量标注数据和病理医生动态干预,迅速实现使用偏振特征对其对应 H&E 染色图像的扩展标注,减少医生标注数据工作量、提高精细标注数据质量。

图 7 以肺癌病变特征为例,显示偏振超像素扩展标注流程:1) 选取一块缪勒图像区域,利用 Mini Batch KMeans<sup>[57]</sup>对缪勒像素进行聚类计算,形成偏振超像素集合。通常采用 512~1024 个超像素,每个超像素都包含偏振空间的质心位置、标准差和像素数三个指标。2) 医生针对区域内对应某种待测结构的小区域做初始标注[图 7(a)],计算每个超像素中包含的像素对标注区域的贡献率[图 7(a)红框中,标注区域内像素数/区域内总像素数],贡献度大的超像素意味

着初始标注区域中该偏振特征发挥作用大。以此贡献率作为该超像素及其所含所有偏振像素的权重。3) 对整个区域内所有像素及其权重作散点图,并进一步做平滑和阈值分割,这时初始标注区域以及具有同初始标注区域相似偏振特征的区域会凸显出来,实现拓展标注[图 7(b)]。4) 经过病理医生识别,保留扩大正确标注区域,进一步对上述过程循环迭代,直到区域内所有与初始标注区域具有类似结构和偏振特征的区域都被正确标注。5) 然后扩展到新的缪勒图像区域,利用前面正确标注的区域做初始标注,进行扩展标注[图 7(c)]。6) 迭代循环,直至 ROI 内所有待测结构都被正确标注[图 7(d)],并进一步拓展到其他 ROI 和其他样本[图 7(e)和图 7(f)]。

利用上述方法,可以从少量精确标注的样本出发,在病理医生的干预下,在大量缪勒图像中以初始标注区域为模板迅速找到具有类似偏振特征和微观结构特征的区域,快速实现利用偏振特征对相应 H&E 染色图像的高质量扩展标注。进而在具有大量精确标注数据的情况下,进一步形成对应标注像素的偏振特征模



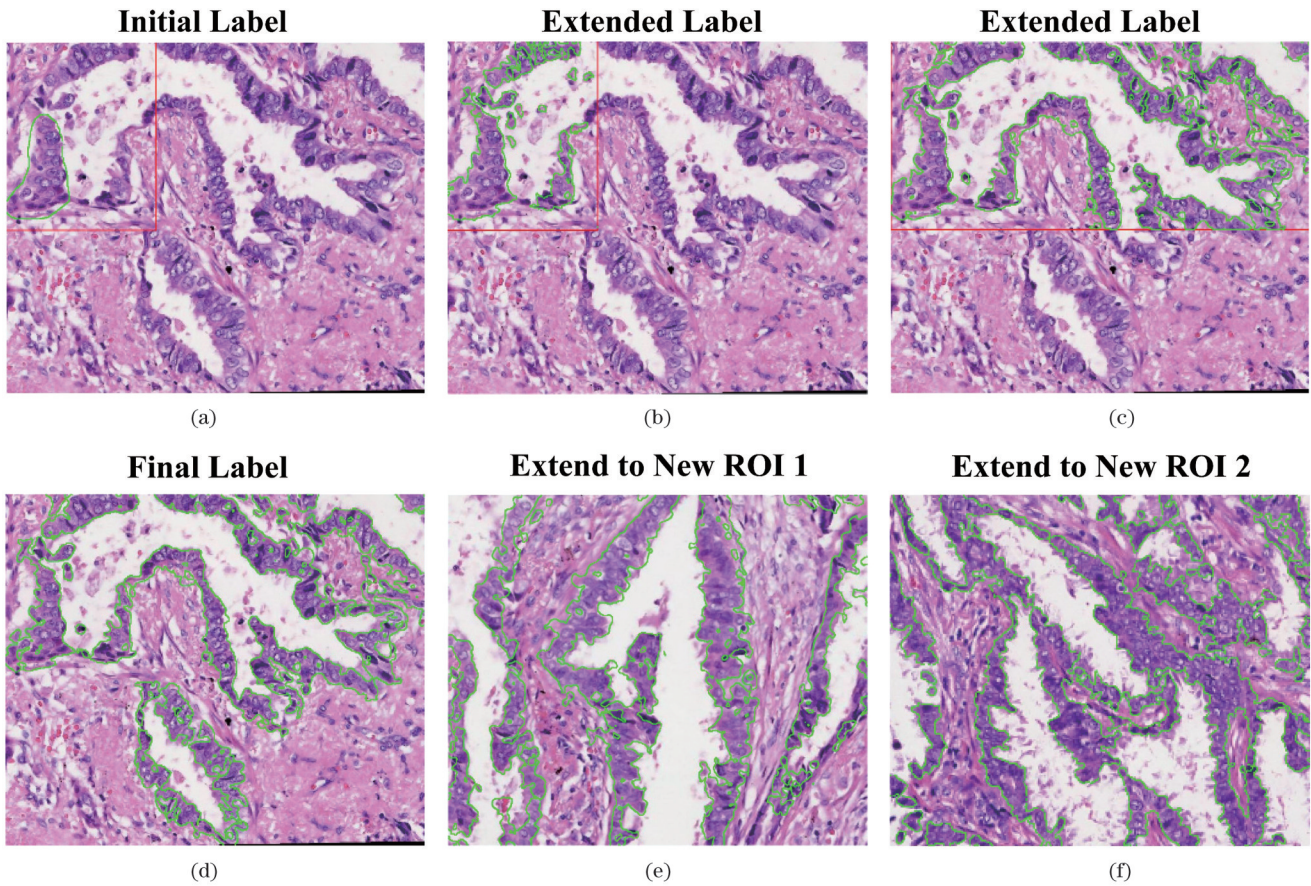


图 7 肺癌癌区区域偏振超像素扩展标注流程。(a)医生在红框内提供少量初始标注;(b)通过筛选贡献度大的偏振超像素实现红框内扩展标注,医生筛选保留正确标注区域;(c)扩大红框范围,医生重复偏振超像素贡献度计算与筛选过程,实现扩展标注;(d)通过迭代上述过程,实现整个 ROI 扩展标注;(e)将已知标注的 ROI 扩展到新未知 ROI 1;(f)将已知标注的 ROI 扩展到新未知 ROI 2

Fig. 7 Polarization super-pixels based label extension process for lung cancer ROIs. (a) Pathologist provides a small initial label within the red box; (b) obtain extended label for the whole red box by selecting super-pixels with high contribution, pathologist identify and retain the correct label; (c) expand the red box, pathologist repeat polarization super-pixels calculation and selection to obtain extended label; (d) by iterating the above process, the entire extended label for the ROI is obtained; (e) extend the label to a new ROI 1; (f) extend the label to a new ROI 2

板,根据模板所代表的偏振特征,在未知样本繆勒图像中凸显具有病变特征的区域,并实现对应 H&E 染色图像的扩展标注,为病理医生快速识别病变区域、细致分析病变特征和解释病变机制提供依据。本文

介绍了基于监督学习、无监督偏振像素聚类,以及偏振超像素和偏振特征模板的偏振特征提取方法。表 1 总结了不同偏振特征提取方法的适用条件、优势以及缺点。

表 1 基于机器学习的偏振特征提取方法总结

Table 1 Summary of machine learning based polarization feature extraction methods

Method	Applicable condition	Advantage	Disadvantage
Supervised learning	1) Have large amount of well labeled polarization data 2) Target on clear structure recognition	1) Superior performance in recognition tasks	1) Time and labor consuming to acquire labels 2) Performance degrades with unseen data
Unsupervised learning	1) Lack of well labeled polarization data 2) Aim to uncover underlying polarization features	1) No need for labeling 2) Can explore hidden polarization features	1) Cannot provide direct predictive results 2) Requires more computational power to handle large datasets
Polarization super-pixel	1) Have small number of initial labels, and large amount of unlabeled data 2) Aim to uncover underlying polarization features	1) Overcome the needs for large amount of well labeled data 2) Fully utilizes polarization features to achieve fast labeling, enhances model generalization	1) May not outperform supervised learning for specific tasks 2) Performance remains constrained by the quality of initial labels

## 4 结束语

缪勒矩阵包含丰富的复杂生物样本光学性质和微观结构特征信息,如样本中散射颗粒的浓度、大小、形状、取向、排列、表面形貌、内部结构等。因此,缪勒图像可通过偏振特征提供超分辨层次丰富的微观结构信息,帮助我们区分不同类型和不同状态的细胞和组织。缪勒成像具有无标记、无损伤、跨尺度、多模态和定量测量的能力,在复杂生物和医学样本的静态和动态测量中已显示出诱人的应用前景。偏振图像的每个像素都是缪勒矩阵,包含样本的局域光学性质和微观结构特征信息。定量提取像素偏振特征并据此表征样本的物质和结构特征是偏振成像应用的核心问题。本文介绍了全偏振显微成像技术,并结合临床应用总结了现有偏振特征提取方法和最新进展,包括基于监督学习的偏振像素与图像特征提取方法、基于无监督学习的偏振像素聚类,以及偏振超像素和偏振特征模板扩展标注方法。

相比基于染色图像分析的数字病理,基于监督学习的偏振数字病理能够提供更加丰富的微观结构特征信息,已突显出其临床定量辅助诊断的优势和潜力。不过,监督学习需要针对已知特征进行训练,容易限制可用的偏振信息。利用无监督学习对缪勒像素进行聚类,有利于挖掘丰富的偏振特征信息,凸显通常“不可见”或超分辨尺度微观结构特征,发挥复杂结构特征识别、分类和动态监测的潜力。无监督偏振像素聚类有潜力把病变组织拆分成一系列不同偏振特征成分,对应染色图像中医生可识别的不同组织和亚细胞结构,如细胞核、细胞浆、细胞膜和细胞外间质等,甚至可以区分正常和病变的细胞核、细胞浆和不同类型的纤维等。偏振像素聚类可在高维偏振空间中提取和定量表征微观结构特征,把复杂生物样本拆分成亚细胞层次的基本组分,提供一种识别、细分、表征复杂结构生物体系的新型工具。结合偏振成像所具有无标记、无损伤潜力,有望发展一种利用偏振特征拆分并定量表征生物样本亚细胞层次光学性质与微观结构特征的新型偏振空间组学技术。

### 参 考 文 献

- [1] Bera K, Schalper K A, Rimm D L, et al. Artificial intelligence in digital pathology - new tools for diagnosis and precision oncology [J]. *Nature Reviews: Clinical Oncology*, 2019, 16(11): 703-715.
- [2] Niazi M K K, Parwani A V, Gurcan M N. Digital pathology and artificial intelligence[J]. *The Lancet: Oncology*, 2019, 20(5): e253-e261.
- [3] Marini N, Marchesin S, Otálora S, et al. Unleashing the potential of digital pathology data by training computer-aided diagnosis models without human annotations[J]. *NPJ Digital Medicine*, 2022, 5(1): 102.
- [4] Rashid R, Chen Y A, Hoffer J, et al. Narrative online guides for the interpretation of digital-pathology images and tissue-atlas data [J]. *Nature Biomedical Engineering*, 2022, 6(5): 515-526.
- [5] van der Laak J, Litjens G, Ciompi F. Deep learning in histopathology: the path to the clinic[J]. *Nature Medicine*, 2021, 27(5): 775-784.
- [6] He C, He H H, Chang J T, et al. Polarisation optics for biomedical and clinical applications: a review[J]. *Light, Science & Applications*, 2021, 10(1): 194.
- [7] He H H, Liao R, Zeng N, et al. Mueller matrix polarimetry: an emerging new tool for characterizing the microstructural feature of complex biological specimen[J]. *Journal of Lightwave Technology*, 2019, 37(11): 2534-2548.
- [8] 沈元星, 姚悦, 何宏辉, 等. 非标记、量化缪勒矩阵偏振成像在辅助临床诊断中的应用[J]. *中国激光*, 2020, 47(2): 0207001. Shen Y X, Yao Y, He H H, et al. Mueller matrix polarimetry: a label-free, quantitative optical method for clinical diagnosis[J]. *Chinese Journal of Lasers*, 2020, 47(2): 0207001.
- [9] 张钰新, 樊志鹏, 翟好宇, 等. 基于偏振成像的面部皮肤结构和特征非接触在体测量方法[J]. *中国激光*, 2023, 50(3): 0307110. Zhang Y X, Fan Z P, Zhai H Y, et al. Non-contact and in vivo polarization imaging method for measuring facial skin structures and characteristics[J]. *Chinese Journal of Lasers*, 2023, 50(3): 0307110.
- [10] 张迎, 吴晓君, 何宏辉, 等. 手持偏振光声计算层析成像探头设计及应用研究[J]. *激光与光电子学进展*, 2022, 59(8): 0817001. Zhang Y, Wu X J, He H H, et al. Design and application of handheld polarized photoacoustic computational tomography probe [J]. *Laser & Optoelectronics Progress*, 2022, 59(8): 0817001.
- [11] Ramella-Roman J C, Saytashev I, Piccini M. A review of polarization-based imaging technologies for clinical and preclinical applications[J]. *Journal of Optics*, 2020, 22(12): 123001.
- [12] Huang T Y, Meng R Y, Qi J, et al. Fast Mueller matrix microscope based on dual DoFP polarimeters[J]. *Optics Letters*, 2021, 46(7): 1676-1679.
- [13] Alali S, Vitkin A. Polarized light imaging in biomedicine: emerging Mueller matrix methodologies for bulk tissue assessment [J]. *Journal of Biomedical Optics*, 2015, 20(6): 061104.
- [14] Qi J, Elson D S. Mueller polarimetric imaging for surgical and diagnostic applications: a review[J]. *Journal of Biophotonics*, 2017, 10(8): 950-982.
- [15] Wang Y, He H H, Chang J T, et al. Differentiating characteristic microstructural features of cancerous tissues using Mueller matrix microscope[J]. *Micron*, 2015, 79: 8-15.
- [16] Arteaga O, Baldris M, Antó J, et al. Mueller matrix microscope with a dual continuous rotating compensator setup and digital demodulation[J]. *Applied Optics*, 2014, 53(10): 2236-2245.
- [17] Oldenbourg R. A new view on polarization microscopy[J]. *Nature*, 1996, 381(6585): 811-812.
- [18] Gribble A, Layden D, Vitkin I A. Experimental validation of optimum input polarization states for Mueller matrix determination with a dual photoelastic modulator polarimeter[J]. *Optics Letters*, 2013, 38(24): 5272-5275.
- [19] Chang J T, He H H, Wang Y, et al. Division of focal plane polarimeter-based 3×4 Mueller matrix microscope: a potential tool for quick diagnosis of human carcinoma tissues[J]. *Journal of Biomedical Optics*, 2016, 21(5): 056002.
- [20] Qi J, Ye M L, Singh M, et al. Narrow band 3×3 Mueller polarimetric endoscopy[J]. *Biomedical Optics Express*, 2013, 4(11): 2433-2449.
- [21] Meng R Y, Chen Z H, Wang X J, et al. Comparison of different calibration methods for Mueller matrix microscopy of cells[J]. *Applied Optics*, 2021, 60(5): 1380-1386.
- [22] Azzam R M. Photopolarimetric measurement of the Mueller matrix by Fourier analysis of a single detected signal[J]. *Optics Letters*, 1978, 2(6): 148-150.
- [23] Liu L, Oldenbourg R, Trimarchi J R, et al. A reliable, noninvasive technique for spindle imaging and enucleation of mammalian oocytes[J]. *Nature Biotechnology*, 2000, 18(2): 223-225.

- [24] Wang W H, Meng L, Hackett R J, et al. Developmental ability of human oocytes with or without birefringent spindles imaged by Polscope before insemination[J]. Human Reproduction, 2001, 16 (7): 1464-1468.
- [25] 黄彤宇, 孟若愚, 赵千皓, 等. 模块化全偏振显微镜[J]. 中国激光, 2021, 48(15): 1517002.  
Huang T Y, Meng R Y, Zhao Q H, et al. Modular full-polarization microscope[J]. Chinese Journal of Lasers, 2021, 48 (15): 1517002.
- [26] Zhao Q H, Huang T Y, Hu Z, et al. Geometric optimization method for a polarization state generator of a Mueller matrix microscope[J]. Optics Letters, 2021, 46(22): 5631-5634.
- [27] Chen Z H, Meng R Y, Zhu Y H, et al. A collinear reflection Mueller matrix microscope for backscattering Mueller matrix imaging[J]. Optics and Lasers in Engineering, 2020, 129: 106055.
- [28] Nordin G P, Meier J T, Deguzman P C, et al. Micropolarizer array for infrared imaging polarimetry[J]. Journal of the Optical Society of America A, 1999, 16(5): 1168-1174.
- [29] Li P C, Lü D H, He H H, et al. Separating azimuthal orientation dependence in polarization measurements of anisotropic media[J]. Optics Express, 2018, 26(4): 3791-3800.
- [30] Li P C, Dong Y, Wan J C, et al. Polaromics: deriving polarization parameters from a Mueller matrix for quantitative characterization of biomedical specimen[J]. Journal of Physics D: Applied Physics, 2022, 55(3): 034002.
- [31] Lu S Y, Chipman R A. Interpretation of Mueller matrices based on polar decomposition[J]. Journal of the Optical Society of America A, 1996, 13(5): 1106-1113.
- [32] Ghosh N, Wood M F G, Vitkin I A. Mueller matrix decomposition for extraction of individual polarization parameters from complex turbid media exhibiting multiple scattering, optical activity, and linear birefringence[J]. Journal of Biomedical Optics, 2008, 13(4): 044036.
- [33] Morio J, Goudail F. Influence of the order of diattenuator, retarder, and polarizer in polar decomposition of Mueller matrices [J]. Optics Letters, 2004, 29(19): 2234-2236.
- [34] Ossikovski R, De Martino A, Guyot S. Forward and reverse product decompositions of depolarizing Mueller matrices[J]. Optics Letters, 2007, 32(6): 689-691.
- [35] Anastasiadou M, Ben Hatit S, Ossikovski R, et al. Experimental validation of the reverse polar decomposition of depolarizing Mueller matrices[J]. Journal of the European Optical Society: Rapid Publications, 2007, 2: 07018.
- [36] Ghosh N, Wood M F G, Vitkin I A. Influence of the order of the constituent basis matrices on the Mueller matrix decomposition-derived polarization parameters in complex turbid media such as biological tissues[J]. Optics Communications, 2010, 283(6): 1200-1208.
- [37] Sun M H, He H H, Zeng N, et al. Probing microstructural information of anisotropic scattering media using rotation-independent polarization parameters[J]. Applied Optics, 2014, 53 (14): 2949-2955.
- [38] Liu Y D, Dong Y, Si L, et al. Comparison between image texture and polarization features in histopathology[J]. Biomedical Optics Express, 2021, 12(3): 1593-1608.
- [39] Wang Y, He H H, Chang J T, et al. Mueller matrix microscope: a quantitative tool to facilitate detections and fibrosis scorings of liver cirrhosis and cancer tissues[J]. Journal of Biomedical Optics, 2016, 21(7): 071112.
- [40] Dong Y, Qi J, He H H, et al. Quantitatively characterizing the microstructural features of breast ductal carcinoma tissues in different progression stages by Mueller matrix microscope[J]. Biomedical Optics Express, 2017, 8(8): 3643-3655.
- [41] Majumdar A, Lad J, Tumanova K, et al. Machine learning based local recurrence prediction in colorectal cancer using polarized light imaging[J]. Journal of Biomedical Optics, 2024, 29(5): 052915.
- [42] Yao Y, Zuo M, Dong Y, et al. Polarization imaging feature characterization of different endometrium phases by machine learning[J]. OSA Continuum, 2021, 4(6): 1776-1791.
- [43] Yao Y, Zhang F D, Wang B, et al. Polarization imaging-based radiomics approach for the staging of liver fibrosis[J]. Biomedical Optics Express, 2022, 13(3): 1564-1580.
- [44] Luu N T, Le T H, Phan Q H, et al. Characterization of Mueller matrix elements for classifying human skin cancer utilizing random forest algorithm[J]. Journal of Biomedical Optics, 2021, 26(7): 075001.
- [45] Ivanov D, Dremine V, Genova T, et al. Polarization-based histopathology classification of *ex vivo* colon samples supported by machine learning[J]. Frontiers in Physics, 2022, 9: 814787.
- [46] Dong Y, Wan J C, Si L, et al. Deriving polarimetry feature parameters to characterize microstructural features in histological sections of breast tissues[J]. IEEE Transactions on Bio-Medical Engineering, 2021, 68(3): 881-892.
- [47] Dong Y, Wan J C, Wang X J, et al. A polarization-imaging-based machine learning framework for quantitative pathological diagnosis of cervical precancerous lesions[J]. IEEE Transactions on Medical Imaging, 2021, 40(12): 3728-3738.
- [48] Roa C, Du Le V N, Mahendroo M, et al. Auto-detection of cervical collagen and elastin in Mueller matrix polarimetry microscopic images using K-NN and semantic segmentation classification[J]. Biomedical Optics Express, 2021, 12(4): 2236-2249.
- [49] McKinley R I, Felger L A, Hewer E, et al. Machine learning for white matter fibre tract visualization in the human brain via Mueller matrix polarimetric data[J]. Proceedings of SPIE, 2022, 12136: 121360G.
- [50] Schucht P, Lee H R, Mezouar H M, et al. Visualization of white matter fiber tracts of brain tissue sections with wide-field imaging mueller polarimetry[J]. IEEE Transactions on Medical Imaging, 2020, 39(12): 4376-4382.
- [51] Chen Y, Dong Y, Si L, et al. Dual polarization modality fusion network for assisting pathological diagnosis[J]. IEEE Transactions on Medical Imaging, 2023, 42(1): 304-316.
- [52] Wan J C, Dong Y, Yao Y, et al. Unsupervised learning of pixel clustering in Mueller matrix images for mapping microstructural features in pathological tissues[J]. Communications Engineering, 2023, 2: 88.
- [53] McInnes L, Healy J, Melville J. UMAP: uniform manifold approximation and projection for dimension reduction[EB/OL]. (2018-02-09)[2023-11-03]. <http://arxiv.org/abs/1802.03426>.
- [54] Achanta R, Shaji A, Smith K, et al. SLIC superpixels compared to state-of-the-art superpixel methods[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34(11): 2274-2282.
- [55] Subudhi S, Patro R N, Biswal P K, et al. A survey on superpixel segmentation as a preprocessing step in hyperspectral image analysis[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2021, 14: 5015-5035.
- [56] Sinaga K P, Yang M S. Unsupervised K-means clustering algorithm[J]. IEEE Access, 2020, 8: 80716-80727.
- [57] Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: machine learning in Python[J]. The Journal of Machine Learning Research, 2011, 12: 2825-2830.

# Digital Pathology Based on Fully Polarized Microscopic Imaging

Yao Yue<sup>1,2</sup>, Pei Haojie<sup>1,2</sup>, Li Hao<sup>3</sup>, Wan Jiachen<sup>1,2</sup>, Tao Lili<sup>3</sup>, Ma Hui<sup>1,2\*</sup>

<sup>1</sup>*Shenzhen International Graduate School, Tsinghua University, Shenzhen 518055, Guangdong, China;*

<sup>2</sup>*Guangdong Engineering Center of Polarization Imaging and Sensing Technology, Shenzhen 518055, Guangdong, China;*

<sup>3</sup>*Department of Pathology, Peking University Shenzhen Hospital, Shenzhen 518036, Guangdong, China*

## Abstract

**Significance** Digital pathology uses digitized pathological images and their features in conjunction with artificial intelligence technology to achieve quantitative characterization of cancerous tissues and assist pathologists in clinical diagnoses. The use of polarized light illumination and polarized light detection can achieve full polarization imaging. Accordingly, the polarization characteristics of each pixel of the image contain abundant microstructural information, especially subcellular super-resolution information, that is difficult to obtain with nonpolarization imaging. Polarization imaging can provide a more effective means for the identification and quantitative characterization of cancerous tissues. This paper introduces Mueller matrix microscopic imaging techniques and comprehensively reviews the latest methods for polarization feature extraction, including supervised learning-based polarization pixel and image feature extraction, unsupervised learning-based polarization pixel clustering, and the extension of annotations through polarization feature templates based on super-pixels, highlighting their potential clinical applications.

**Progress** Mueller matrix imaging provides abundant subcellular-level information on tissue microstructures. The quantitative extraction of polarization features from Mueller pixels is crucial for the clinical application of polarization imaging. In contrast to stain image-based digital pathology, polarization feature extraction through supervised learning offers more abundant microstructural information. However, the reliance on extensive, well-annotated data poses time and labor challenges. Moreover, supervised learning is dependent on pathologists' prior knowledge, limiting the comprehensive utilization of information from the polarization space. Unsupervised clustering methods facilitate the decomposition of pathological tissues into distinct microstructural subtypes, enhancing the exploration of the rich information embedded in Mueller pixels. Additionally, this approach provides evidence for the ongoing discovery of new physical properties, structural characteristics, and dynamic processes at all levels above the subcellular scale in organisms, including living entities.

**Conclusions and Prospects** Following advancements in molecular biology techniques, the specific identification of molecular components in biological entities is becoming a pivotal tool in biomedical research, thus leading to diverse omics approaches. Polarization-based digital pathology can leverage feature extraction methods developed in various omics approaches. The unsupervised clustering of Mueller pixels quantitatively extracts information at various levels above the subwavelength scale, enabling the integration of label-free, noninvasive, abundant information features of Mueller matrix imaging into novel spatiotemporal omics methods.

**Key words** medical optics; fully polarized microscopic imaging; polarization-based digital pathology; polarization feature extraction; machine learning