

# Opto-Electronic Science

ISSN 2097-0382

CN 51-1800/O4

## Data-driven polarimetric imaging: a review

Kui Yang, Fei Liu, Shiyang Liang, Meng Xiang, Pingli Han, Jinpeng Liu, Xue Dong, Yi Wei, Bingjian Wang, Koichi Shimizu and Xiaopeng Shao

**Citation:** Yang K, Liu F, Liang SY, et al. Data-driven polarimetric imaging: a review. *Opto-Electron Sci* **3**, 230042 (2024).

<https://doi.org/10.29026/oes.2024.230042>

Received: 14 November 2023; Accepted: 19 January 2024; Published online: 29 February 2024

## Related articles

### Crosstalk-free achromatic full Stokes imaging polarimetry metasurface enabled by polarization-dependent phase optimization

Yaxin Zhang, Mingbo Pu, Jinjin Jin, Xinjian Lu, Yinghui Guo, Jixiang Cai, Fei Zhang, Yingli Ha, Qiong He, Mingfeng Xu, Xiong Li, Xiaoliang Ma, Xiangang Luo

*Opto-Electronic Advances* 2022 **5**, 220058 doi: [10.29026/oea.2022.220058](https://doi.org/10.29026/oea.2022.220058)

### Towards the performance limit of catenary meta-optics via field-driven optimization

Siran Chen, Yingli Ha, Fei Zhang, Mingbo Pu, Hanlin Bao, Mingfeng Xu, Yinghui Guo, Yue Shen, Xiaoliang Ma, Xiong Li, Xiangang Luo

*Opto-Electronic Advances* 2024 , doi: [10.29026/oea.2024.230145](https://doi.org/10.29026/oea.2024.230145)

### Multi-foci metalens for spectra and polarization ellipticity recognition and reconstruction

Hui Gao, Xuhao Fan, Yuxi Wang, Yuncheng Liu, Xinger Wang, Ke Xu, Leimin Deng, Cheng Zeng, Tingan Li, Jinsong Xia, Wei Xiong

*Opto-Electronic Science* 2023 **2**, 220026 doi: [10.29026/oes.2023.220026](https://doi.org/10.29026/oes.2023.220026)

### Linear polarization holography

Jinyu Wang, Xiaodi Tan, Peiliang Qi, Chenhao Wu, Lu Huang, Xianmiao Xu, Zhiyun Huang, Lili Zhu, Yuanying Zhang, Xiao Lin, Jinliang Zang, Kazuo Kuroda

*Opto-Electronic Science* 2022 **1**, 210009 doi: [10.29026/oes.2022.210009](https://doi.org/10.29026/oes.2022.210009)

More related article in Opto-Electronic Journals Group website 



Opto-Electronic  
Science

<http://www.ojournal.org/oes>



 OE\_Journal



Website

DOI: [10.29026/oes.2024.230042](https://doi.org/10.29026/oes.2024.230042)

# Data-driven polarimetric imaging: a review

Kui Yang<sup>1†</sup>, Fei Liu<sup>1†\*</sup>, Shiyang Liang<sup>3†</sup>, Meng Xiang<sup>1</sup>, Pingli Han<sup>1</sup>,  
Jinpeng Liu<sup>1</sup>, Xue Dong<sup>1</sup>, Yi Wei<sup>4</sup>, Bingjian Wang<sup>2</sup>, Koichi Shimizu<sup>3\*</sup> and  
Xiaopeng Shao<sup>1,5\*</sup>

This study reviews the recent advances in data-driven polarimetric imaging technologies based on a wide range of practical applications. The widespread international research and activity in polarimetric imaging techniques demonstrate their broad applications and interest. Polarization information is increasingly incorporated into convolutional neural networks (CNN) as a supplemental feature of objects to improve performance in computer vision task applications. Polarimetric imaging and deep learning can extract abundant information to address various challenges. Therefore, this article briefly reviews recent developments in data-driven polarimetric imaging, including polarimetric descattering, 3D imaging, reflection removal, target detection, and biomedical imaging. Furthermore, we synthetically analyze the input, datasets, and loss functions and list the existing datasets and loss functions with an evaluation of their advantages and disadvantages. We also highlight the significance of data-driven polarimetric imaging in future research and development.

**Keywords:** deep learning; polarimetric imaging; image processing

Yang K, Liu F, Liang SY et al. Data-driven polarimetric imaging: a review. *Opto-Electron Sci* **3**, 230042 (2024).

## Introduction

Polarization is a fundamental physical property of light that expresses the characteristics of vector shear wave<sup>1,2</sup>. When light interacts with objects or media, it exhibits various characteristics and representations of polarization that correspond to the intrinsic characteristics of the material. This unique quality is an additional dimension of information that has many applications in various fields, such as polarimetric descattering<sup>3-9</sup>, 3D shape reconstruction<sup>10-13</sup>, reflection removal<sup>14,15</sup>, target detection<sup>16-20</sup>, biomedical imaging<sup>21-23</sup>, pathological dia-

gnosis<sup>24-29</sup>, remote sensing<sup>30-33</sup>, and semantic segmentation<sup>34-36</sup>. Furthermore, harnessing the features of polarization opens new possibilities for research and technology development.

The conventional polarization method is limited in accurately capturing and utilizing information owing to complicated interactions in the transmission process. Moreover, convolutional neural networks (CNN) excel at nonlinear expression and information extraction based on large datasets, making them better suited for modeling and interpreting polarization information than traditional algorithms, which can bridge the gap between

<sup>1</sup>School of Optoelectronic Engineering, Xidian University, Xi'an 710071, China; <sup>2</sup>School of Physics, Xidian University, Xi'an 710071, China; <sup>3</sup>Graduate School of Information, Production and Systems, Waseda University, Kitakyushu 808-0135, Japan; <sup>4</sup>Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA; <sup>5</sup>Hangzhou Institute of Technology, Xidian University, Hangzhou 311200, China.

<sup>†</sup>These authors contributed equally to this work.

\*Correspondence: F Liu, E-mail: feiliu@xidian.edu.cn; K Shimizu, E-mail: shimizu@xidian.edu.cn;

XP Shao, E-mail: xpshao@xidian.edu.cn

Received: 14 November 2023; Accepted: 19 January 2024; Published online: 29 February 2024



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License.

To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024. Published by Institute of Optics and Electronics, Chinese Academy of Sciences.

theory and practice<sup>37–41</sup>.

To overcome the lack of a physical theoretical model, the use of deep learning technology to optimize the information processing procedure has been proposed. Several experiments have proven its advantages in extracting potential features from images and determining the relationship between information transmission systems based on massive datasets. However, intensity information is commonly used in existing data-driven methods, which lose other three-dimensional information. In addition, the unicity of information would make the imaging performance suffer from various challenges; for example, in target detection tasks, spoofing targets and camouflaged targets can reduce the accuracy rate; in semantic segmentation, water hazards and metallic surfaces are the key challenges in the segmentation of road scenes; in pathological diagnosis, the only color information in medical images would increase the risk of misdiagnosis; and for transparent objects, parts of imaging processes would be challenging to implement, such as 3D reconstruction and segmentation in the intensity field. Therefore, the introduction of additional light-field physical information, such as polarization, is gaining increasing attention as a supplemental feature of objects for improved performance in higher-level visual tasks, expanding beyond intensity-only coverage. Therefore, polarimetric imaging and deep learning will contribute to future research and development. This review summarizes the existing methods of combining polarimetric imaging and deep learning and demonstrates the current gains visually and comprehensively.

The remainder of this study is organized as follows: First, we review the recent trends in data-driven polarimetric imaging from four aspects. Second, seven existing research fields are categorized and analyzed with their corresponding algorithms. Third, we discuss various algorithmic approaches used in each field. Next, we discuss three critical aspects of their existing datasets and loss functions and their advantages and disadvantages. Next, we discuss the strengths and weaknesses of the practical applications of data-driven polarimetric imaging and possible opportunities to address these challenges. Finally, we provide conclusions, highlighting the importance and potential of data-driven polarimetric imaging in various applications and research areas that could extend the current uses of data-driven polarimetric imaging and provide insights to advance its future development.

## Data-driven polarimetric imaging

### Short history

Polarimetric imaging, first observed by Sir Isaac Newton and Christiaan Huygens in the 17th century, is a technique used to capture and analyze the polarization properties of light. Moreover, they noticed that light could be separated into two polarized lights when interacting with calcite crystals. In the 19th century, scientists such as Malus and Fresnel made significant contributions to the understanding of polarization. Since its inception, polarimetric imaging has progressed and been applied in three essential fields in the 20th century: polarization microscopy, which allows scientists to study the birefringence properties of materials under polarized light, remote sensing and polarimetry, enhancing the detection and discrimination of objects based on their polarization characteristics; medical imaging; and tissue optical properties. In recent years, digital cameras and other innovative imaging techniques have become capable of capturing valuable information on surfaces, materials, and biological samples. Recent developments in computer vision have enabled the incorporation of polarization information into various tasks, offering potential advantages over traditional RGB imaging.

Deep learning, a subfield of machine learning, is inspired by the human brain. Artificial neural networks emerged during the 1940s and 1950s. Research on neural networks continued during the 1960s and 1980s, and perceptron and backpropagation algorithms were developed. However, neural networks face limitations owing to computational constraints and insufficient data. In the early 2000s, traditional machine learning techniques, such as support vector machines and decision trees, outperformed deep learning in many tasks, leading to a decline in interest in deep learning. After 2010, the proposal of a new network structure and improvements in computational power propelled the advancement of deep learning. Key developments during this decade include the use of CNN for computer vision and recurrent neural networks for natural language processing. With ongoing research, deep learning continues to evolve into more efficient training algorithms, model architectures, and applications across various domains, making it a central part of the AI field of artificial intelligence.

Until the 21st century, the combination of deep learning and polarimetric imaging has become more pronounced. In the mid-2010s, data-driven techniques were

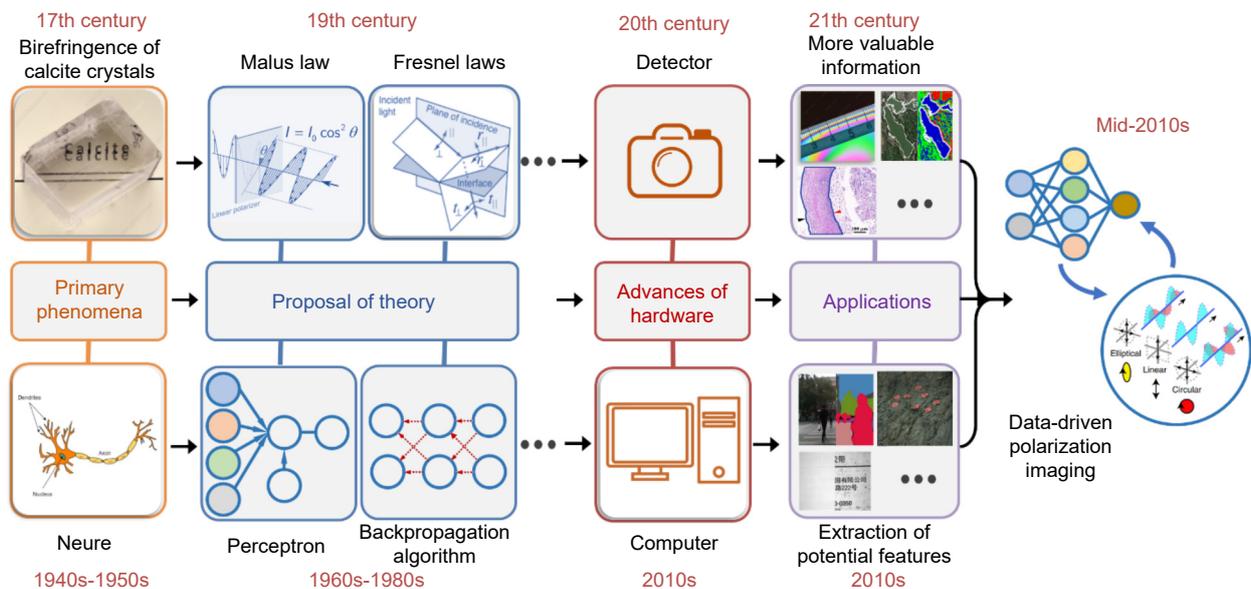
applied to polarimetric imaging. Machines and deep learning have been used for tasks, including object recognition in polarimetric images. Subsequently, researchers explored the potential of deep learning for processing and analyzing polarimetric data in more diverse domains because deep learning helps extract valuable information from polarization data. Additionally, the synergy between deep learning and polarimetric imaging continues to evolve with advancements in models, algorithms, and applications, forming the significant field called “data-driven polarimetric imaging”, which is the novel method that combines the learning method and polarimetric imaging. Data-driven polarimetric imaging can revolutionize various fields and represents a promising interdisciplinary research area and technological innovation. A diagram of this short history is shown in Fig. 1.

**Trends**

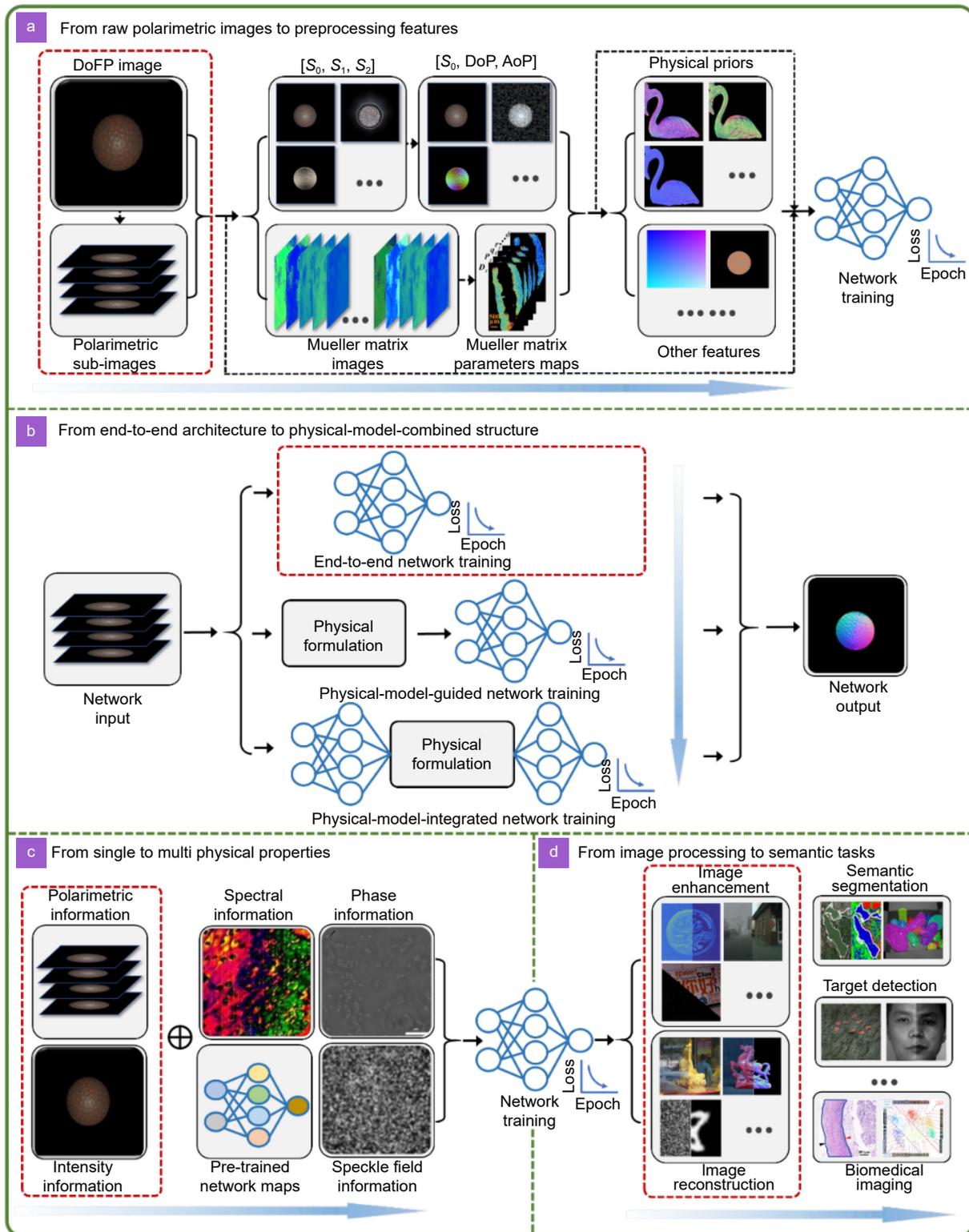
Data-driven polarimetric imaging is a novel approach aimed at compensating for the defects and difficulties of a single-information interpretation model. Recently, the advantages of combining polarimetric imaging and deep learning have been applied in several fields, and dozens of algorithms based on data-driven polarimetric imaging have been proposed, covering several application areas. This section describes more recent trends in data-driven polarimetric imaging from four perspectives. Schematics of the trends in existing data-driven polarimetric imaging are shown in Fig. 2.

With the exploration of data-driven polarimetric imaging, the application fields and utilization of polarization information have gradually increased. Regarding the input and use of polarization information, directly captured images from detectors, such as division of focal plane (DoFP) images<sup>42–45</sup> or 0°, 45°, 90°, and 135°<sup>46–52</sup> polarization images, are the most common inputs. It should be noted that, despite having four components, the last one is often neglected in practical applications, leading to the predominant use of only the initial three parameters. This is because, theoretically, the last parameter can be derived from the first three, thereby reducing the difficulty in image acquisition. In addition, polarimetric parameter features were calculated to represent the polarization information more intuitively, such as the degree of polarization (DoP) and angle of polarization(AoP)<sup>52–57</sup>,  $[S_0, S_1, S_2]$ <sup>52,55,58,59</sup>, Mueller matrix images<sup>60–63</sup> and other combinations of these elements. Features based on the corresponding physical model during different tasks were also computed to guide network training, such as zenith and azimuth angle maps derived from specular and diffuse reflection<sup>64,65</sup> in 3D shape reconstruction tasks.

Furthermore, physical modes are crucial during network training, such as polarimetric descattering models, three-dimensional imaging models, Fresnel equations, Mueller matrix interpretation models, and other traditional polarization models. The preliminary methods are often end-to-end architectures<sup>43,46,47,56,64,66–70</sup>, suggesting that the polarization information is input into the



**Fig. 1 | Schematic of short history about data-driven polarimetric imaging.**



**Fig. 2 | Schematics of the trends of existing data-driven polarimetric imaging.** The solid black arrows refer to the data flow, the dotted black arrows refer to that the data may or not flow. The gradient blue arrows mean the trends of corresponding aspects. (a) The utilization of polarimetric information has been gradually deepened from raw polarimetric images to preprocessing features. (b) The physical modes have been crucial during the training of network from end-to-end architecture to physical-model-combined structure. (c) The various physical properties have been gradually fused to network. (d) The tasks have been developed from image processing to semantic tasks.

network to generate the desired outputs directly; however, the physical models are gradually guided<sup>46,47,49,71</sup> or integrated into network training<sup>48,72–74</sup>. Furthermore, the physical model and its inverse process can form a self-supervised closed loop to achieve better performance. Furthermore, data-driven polarimetric imaging enables the physical interpretability of networks compared with conventional deep network methods. Polarimetric parameters were initially adopted to perform different tasks without exploring the hidden mechanisms. In more in-depth research on data-driven polarimetric imaging, the estimation of specific physical parameters based on the nonlinear representation of deep networks<sup>36,58</sup> and the physical interpretation of network layers<sup>53</sup> have been gradually studied. This process contributes to future research and development. Based on the polarimetric information fused into the network, other physical properties of light have also been introduced in network training, such as spectrum<sup>172</sup> and phase<sup>71,75</sup>. Finally, the application fields tend to be semantic tasks based on convolutional enhancement or restoration imaging. Initially, image processing was the main field applied by data-driven polarimetric imaging, such as descattering imaging<sup>46–49,66,70,72,76</sup>, denoising<sup>45,77,78</sup>, demosaicing<sup>43,44,68</sup>, dynamic range enhancement<sup>79</sup>, reflection removal<sup>53,73,74</sup>, low-light imaging<sup>42,50</sup>, and even 3D reconstruction shape<sup>64,65,67,71,80,81</sup>. Next, semantic tasks appeared gradually similar to semantic segmentation<sup>56,57,69,82,83</sup>, camouflage object detection<sup>84</sup>, classification<sup>60,61</sup>, pathological diagnosis<sup>62,63,85,86,87</sup>.

## Applications of data-driven polarimetric imaging

According to the application field, existing data-driven polarimetric imaging methods can be classified into seven categories: polarimetric descattering, 3D shape reconstruction, reflection removal, restoration and enhancement of polarization information, target detection, biomedical imaging and pathological diagnosis, and semantic segmentation, as shown in Fig. 3.

### Restoration and enhancement of accurate polarization information

Accurate polarization information is the foundation of imaging and its applications. In the real world, theoretical constraints and technological limitations lead to distortion of polarization information. Additionally, the po-

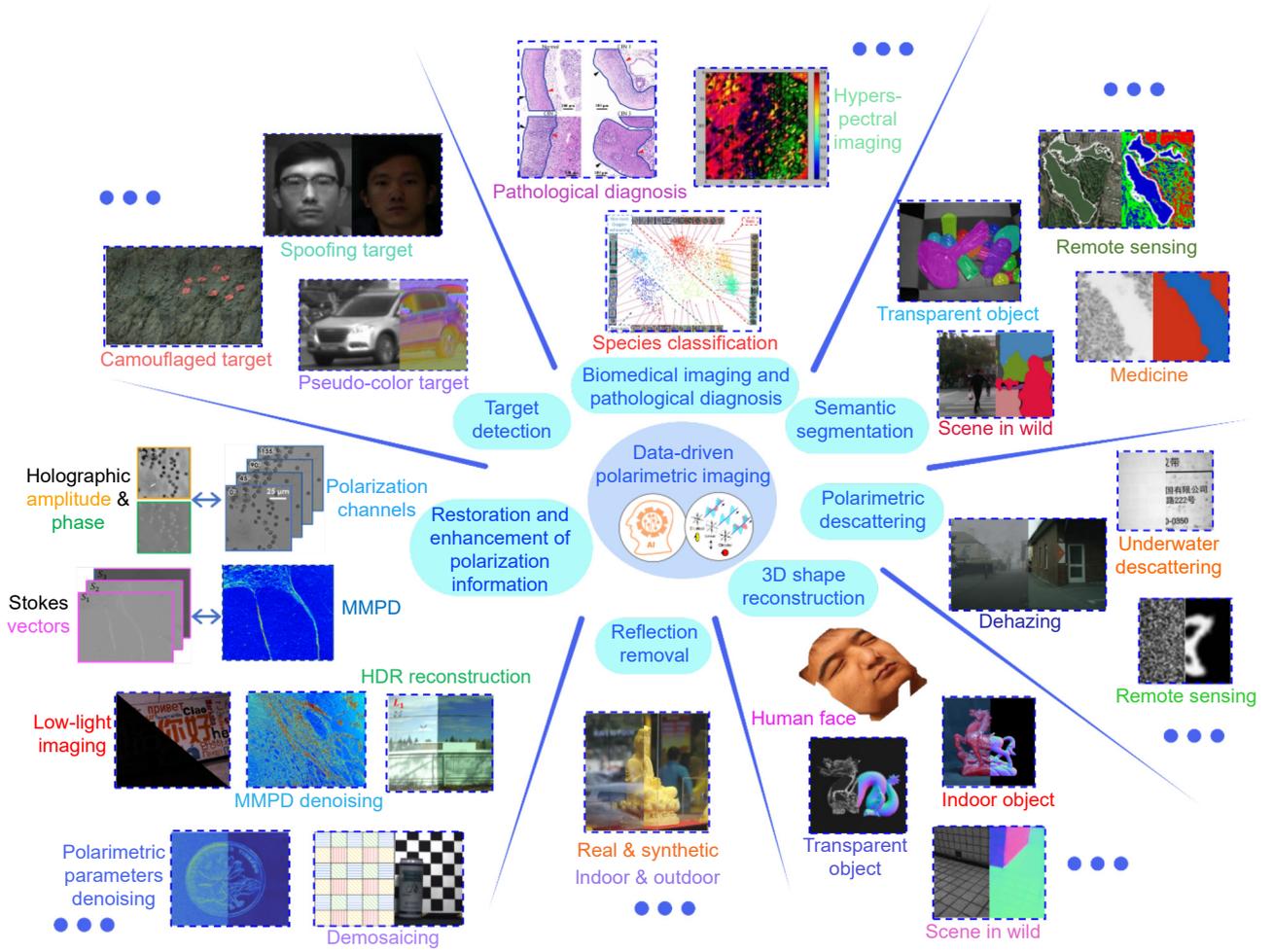
larization parameters of the nonlinear operators are sensitive to noise. Consequently, the effective restoration and enhancement of polarization information are crucial for subsequent applications. In this section, we analyze the limitations of polarimetry techniques and review the existing restoration and enhancement of accurate polarization information methods.

### Polarimetry techniques

Polarimetry techniques are crucial in obtaining polarization information. Several sub-polarized direction images ( $0^\circ$ ,  $45^\circ$ ,  $90^\circ$ , and  $135^\circ$  or  $0^\circ$ ,  $60^\circ$ , and  $120^\circ$ ) must be captured to obtain the polarization characteristics based on the Stokes vector model. However, the obtaining and calculating process, which acquires polarization information indirectly, introduces extra imaging noise to sharply reduce the accuracy of the polarization information. To date, there are four typical methods for measuring polarization images: division of time/rotating elements, division of amplitude<sup>88–90</sup>, division of aperture<sup>91,92</sup>, and division of the focal plane<sup>93–95</sup>.

The division of time/rotating elements is the most common method, depending on the time-sequential activity. The polarizers and retarders were rotated, and measurements were made at different positions of the polarimetry elements. Furthermore, the time gap between operations may cause misregistration of polarization images in a dynamic scene or the motion of the camera case. The division of amplitude could capture multiple images simultaneously; however, the inherent drawback is that the intensity of images would decrease to less than a quarter of the intensity of the original signal, which causes the image contrast to decrease sharply and amplify the image noise. Moreover, misregistration must be solved. The aperture captures images simultaneously using four coaxial cameras with different polarization directions. Distinctly, it is expensive and has a fixed misregistration. The division of the focal plane is the method used in real-time polarimetric imaging, even in dynamic scenes. However, instantaneous field-of-view errors (i.e., mosaicking and low-resolution problems) affect the calculation of polarization parameters. Table 1 presents a comparison of the polarimetric elements.

After capturing the images using the aforementioned polarimetric techniques, Stokes vectors were adopted to display the polarization characteristics. The relationship between them can be characterized by the follow equation:



**Fig. 3 | Applications of data-driven polarimetric imaging.**

$$I_\varphi = (S_0 + S_1 \cos 2\varphi + S_2 \sin 2\varphi) / 2, \quad (1)$$

where  $I_\varphi$  is the image with the polarization direction  $\varphi$ . Then, the Stokes vectors can be calculated from the sub-polarization direction images<sup>96,97</sup>:

$$\begin{aligned} S_0 &= \frac{1}{2} (I_0 + I_{45} + I_{90} + I_{135}) = \frac{2}{3} (I_0 + I_{60} + I_{120}), \\ S_1 &= I_0 - I_{90} = \frac{4}{3} \left( I_0 - \frac{1}{2} I_{60} - \frac{1}{2} I_{120} \right), \\ S_2 &= I_{45} - I_{135} = \frac{2\sqrt{3}}{3} (I_0 - I_{120}), \\ S_3 &= I_{45, \frac{\pi}{2}} - I_{135, \frac{\pi}{2}} = \frac{2\sqrt{3}}{3} (I_{0, \frac{\pi}{2}} - I_{120, \frac{\pi}{2}}), \end{aligned} \quad (2)$$

where  $S_0, S_1, S_2, S_3$  are the components of Stokes vectors. The parameters  $[S_0, S_1, S_2]$  are the most common representations of the linear polarization components.  $S_3$  shows the circular polarization component. Therefore, the polarization characteristic parameters can be obtained using Eq. (3):

$$\begin{aligned} P &= \frac{\sqrt{S_1^2 + S_2^2 + S_3^2}}{S_0}, \quad P_L = \frac{\sqrt{S_1^2 + S_2^2}}{S_0}, \\ P_C &= \frac{S_3}{S_0}, \quad \phi_{\text{pol}} = \frac{1}{2} \arctan \left( \frac{S_2}{S_1} \right), \end{aligned} \quad (3)$$

where  $P, P_L, P_C$  are DoP, degree of linear polarization (DoLP), degree of circular polarization (DoCP), respectively.  $\phi_{\text{pol}}$  is the AoP.

Furthermore, the Mueller matrix is another common but comprehensive parameter that describes the modulation of light after interaction with a material or media<sup>98-100</sup>. The Mueller matrix contained 16 elements using a 4x4 matrix. The Stokes vector of the output light  $S^{\text{out}}$  can be expressed by the Mueller matrix after the incident light  $S^{\text{in}}$  propagates through the medium as follows:

$$\begin{pmatrix} S_0^{\text{out}} \\ S_1^{\text{out}} \\ S_2^{\text{out}} \\ S_3^{\text{out}} \end{pmatrix} = \begin{pmatrix} m_{00} & m_{01} & m_{02} & m_{03} \\ m_{10} & m_{11} & m_{12} & m_{13} \\ m_{20} & m_{21} & m_{22} & m_{23} \\ m_{30} & m_{31} & m_{32} & m_{33} \end{pmatrix} \begin{pmatrix} S_0^{\text{in}} \\ S_1^{\text{in}} \\ S_2^{\text{in}} \\ S_3^{\text{in}} \end{pmatrix}, \quad (4)$$

where  $m_{00}$  represents the transformation of intensity and

**Table 1 | Comparisons of various polarimetric elements.**

	Advantages	Disadvantages
Division of time	<ul style="list-style-type: none"> <li>• Robust</li> <li>• Small system size and inexpensive</li> <li>• Easiest to implement</li> </ul>	<ul style="list-style-type: none"> <li>• Not suitable for dynamic scenes</li> <li>• Not simultaneously imaging</li> <li>• Probable linear misregistration</li> </ul>
Division of amplitude	<ul style="list-style-type: none"> <li>• Simultaneous imaging</li> </ul>	<ul style="list-style-type: none"> <li>• Largest system size</li> <li>• Expensive and complex</li> <li>• Nonlinear misregistration</li> <li>• Loss of image intensity</li> <li>• Highly mechanical flexibility required</li> </ul>
Division of aperture	<ul style="list-style-type: none"> <li>• Simultaneous imaging</li> <li>• Robust</li> </ul>	<ul style="list-style-type: none"> <li>• Expensive</li> <li>• Loss of field of view</li> <li>• Fixed misregistration</li> </ul>
Division of focal plane	<ul style="list-style-type: none"> <li>• Simultaneous imaging</li> <li>• No misregistration</li> <li>• Small and portable</li> </ul>	<ul style="list-style-type: none"> <li>• Loss of spatial resolution</li> <li>• Mosaicking problem</li> <li>• Fabrication difficult</li> <li>• Expensive</li> </ul>

the other 15 elements encode the vectorial properties of the object. Furthermore, Mueller matrix polar decomposition (MMPD)<sup>101</sup>, Mueller matrix transformation (MMT)<sup>102</sup>, Mueller matrix anisotropy coefficients (MMAC)<sup>103</sup>, and other decompositions of the Mueller matrix<sup>104–106</sup> have been proposed to quantitatively characterize the properties of an object.

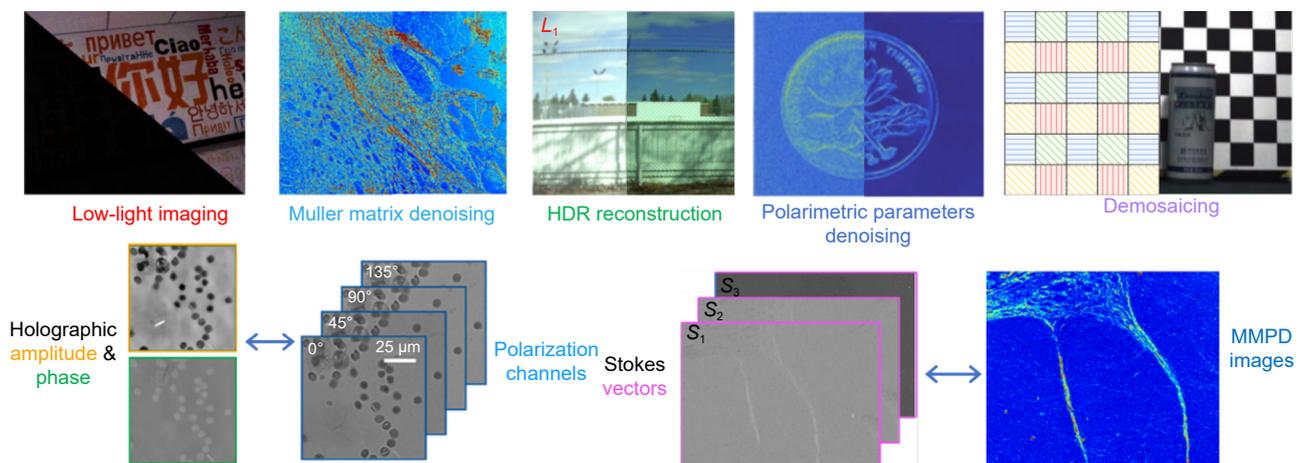
The estimation of these parameters, which represent different polarizations, is always based on nonlinear operations, which are unavoidable when introducing errors and noise. To achieve the desired performance, the calculated polarization parameters must be refined or restored to obtain highly accurate information. In addition, obtaining precise polarization information can be challenging when operating in special imaging environments, such as low light and noise, or with special imaging devices. In these scenarios, the disturbed polariza-

tion parameters amplify errors and negatively affect subsequent imaging performance.

In summary, the acquisition device for raw sub-polarization images, the use of nonlinear operators to calculate polarization characteristic parameters, and special imaging environments also decrease the inaccuracy of the polarization information. Therefore, polarization information with high accuracy must be restored and enhanced. In this section, we examine photostarvation, mosaicking, and noising to demonstrate the restoration and enhancement of polarization information, as depicted in Fig. 4.

### Restoration and enhancement of polarization information methods

In photostarved environments, imaging always suffers

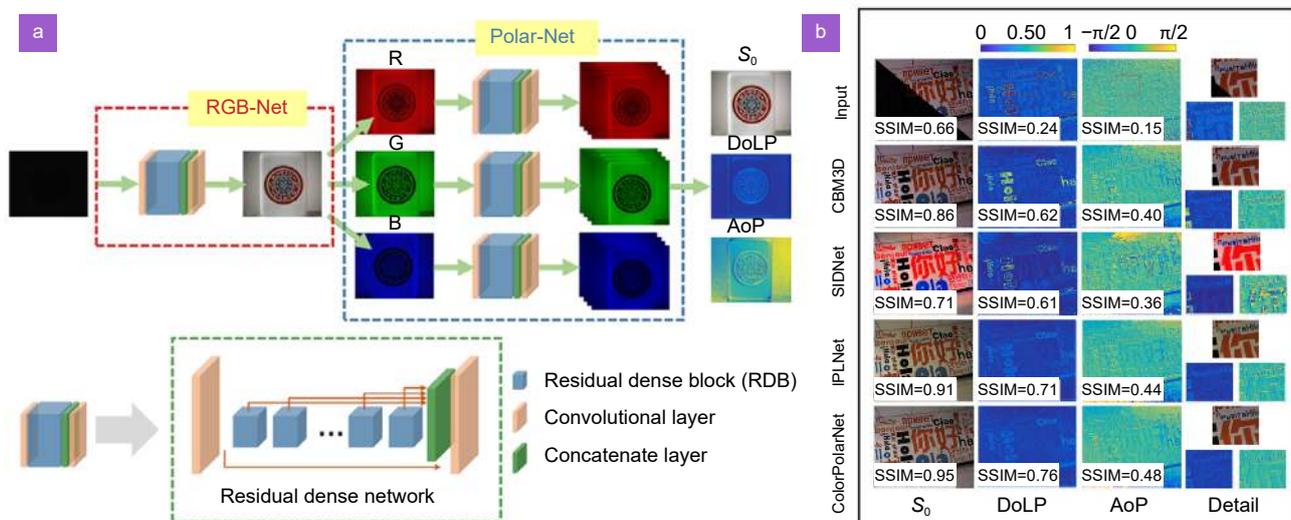


**Fig. 4 | The considered tasks of restoration and enhancement of polarization information, including low-light imaging, Muller matrix denoising, high dynamic range (HDR) reconstruction, polarimetric parameters denoising, demosaicing, transformation between holographic amplitude & phase and polarization channels, transformation between Stokes vectors and MMPD images.**

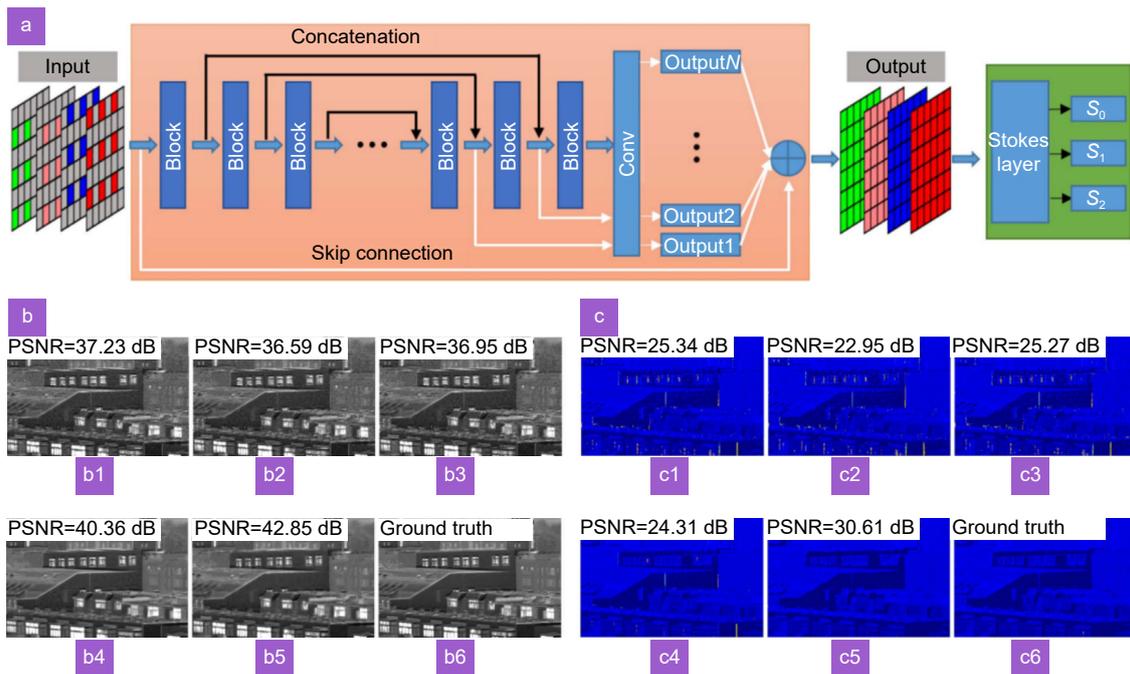
from a low signal-to-noise ratio, which affects the imaging quality, making low-light imaging challenging. In polarimetric imaging, the accuracy of polarization information is degraded. Existing methods consider denoising, correction of color bias, and exposure time in intensity imaging<sup>107,108</sup>. However, the difference between polarimetric and conventional intensity imaging has not been fully considered. Based on the power of being data driven, Hu et al. presented a one-to-three (intensity, DoLP, and AoP) hybrid network called IPLNet to simultaneously enhance the image quality of intensity and polarization information<sup>42</sup>, as shown in Fig. 5(a). The enhanced RGB image generated by the chromatic RGB subnetwork is divided into three channels, and each channel is fed into the polarization subnetwork to predict the polarization information. Enhanced results and corresponding comparisons with mainstream methods are shown in Fig. 5(b) with corresponding values of structural similarity index (SSIM) of each image. However, many parameters are generated in this network, which sharply reduces the operational efficiency. The image color is also inaccurate. Therefore, Xu et al. first performed initial denoising and color deviation correction of four polarization orientation images by network named ColorPolarNet and used the polarization difference network to enhance intensity details, DoLP, and AoP maps<sup>50</sup>. The results demonstrate that the proposed methods have faster processing speed and better performance regarding signal fidelity, contrast enhancement, and color reproduction<sup>50,109</sup>. Compared with IPLNet, ColorPolarNet demonstrates slightly superior per-

formance and achieves higher peak signal-to-noise ratio (PSNR), SSIM and patch-based contrast quality index (PCQI) in terms of  $S_0$ , DoLP, and AoP. Additionally, ColorPolarNet achieves notably lower color difference (CD), indicating reduced distortion compared to IPLNet. In processing speed, ColorPolarNet (2.88 s) is more than twice as fast as the IPLNet model (6.10 s).

Division of the focal plane polarimeter is one of the most common polarimetric imaging sensors that can instantaneously capture dynamic polarization information. Each individual pixel with different polarization orientations is situated in a  $2 \times 2$  superpixel, recording only one from the four essential intensity measurements; therefore, demosaicing and reconstruction of full-resolution and accurate polarization information are indispensable. Zhang et al. proposed a convolutional demosaicing network called the PDCNN to learn end-to-end mapping between coarse interpolation results and full-resolution polarization images<sup>68</sup>, which is the first typical demosaicing architecture, as shown in Fig. 6(a). The bicubic-interpolated results are used as the input, which introduces an interpolation bias, resulting in inaccurate reconstructed results. After comparison with several mainstream methods, the reconstruction results of DoLP and AoP outperformed those of other methods, as shown in Fig. 6(b) and 6(c) with corresponding values of PSNR of each image<sup>110–113</sup>. Zeng et al. proposed a four-layer end-to-end fully convolutional neural network that directly learns mapping from DoFP to three polarization properties: intensity, DoLP, and AoP<sup>43</sup>. However, the noise in the AoP images remained significant. Wu et al. provided



**Fig. 5 | Enhanced method of polarization information.** (a) Architecture of the Hu et al.'s method. (b) Enhanced results and corresponding comparisons with mainstream methods. Figure reproduced with permission from ref.<sup>42</sup>, Optical Society of America Publishing AG.



**Fig. 6 | The convolutional demosaicing network proposed by Zhang et al. and results<sup>68</sup>.** (a) Architecture of the Zhang et al. method. (b) Reconstructed DoLP images of different methods on the test images. (b1) The results of method<sup>113</sup>. (b2) The results of method<sup>110</sup>. (b3) The results of method<sup>111</sup>. (b4) The results of method<sup>112</sup>. (b5) and (b6) are respectively PDCNN and ground truth. (c) Reconstructed AoP images of different methods on the test images. (c5) and (c6) are respectively PDCNN and ground truth. Figure reproduced with permission from ref.<sup>68</sup>, Optical Society of America Publishing AG.

a more physically relevant loss function for the angle of linear polarization (AoLP) reconstruction, establishing a two-stage lightweight approach for reconstructing the intensity and polarization information in real time<sup>44</sup>. The improved version meets the demand for real-time inference. Wen et al.<sup>114</sup>, Sargent et al.<sup>115</sup>, Sun et al.<sup>116</sup> and Pistellato et al.<sup>117</sup> also proposed the data-driven demosaicing methods to ensure the fidelity of polarization signatures and enhance image resolution. Besides deep learning-based methods, other approaches also yield favorable results in this domain, such as the sparse tensor factorization-based model, which introduces a combination of tensor factorization and sparse coding for the first time<sup>118</sup>.

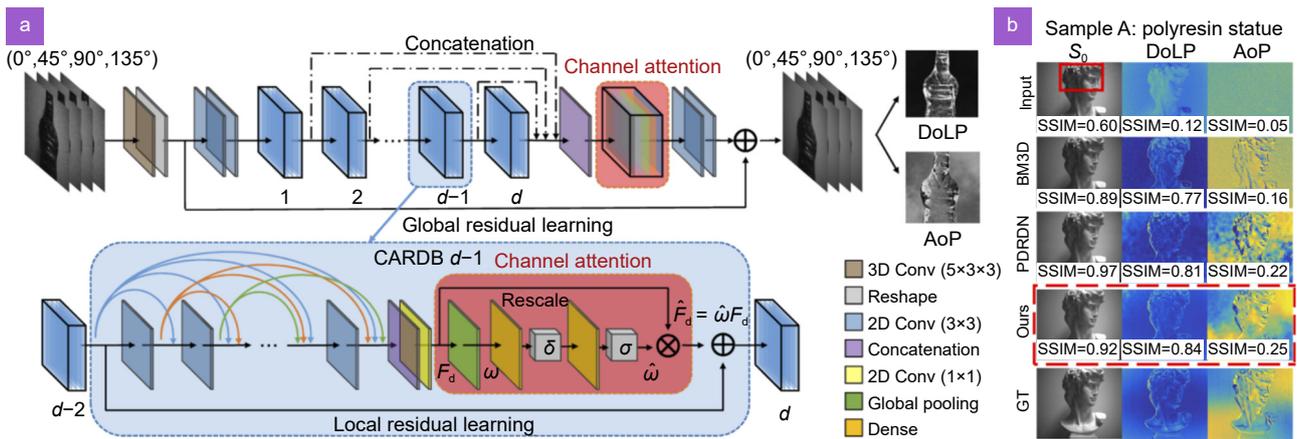
Because the polarimetric parameters are always derived from the measured intensities through nonlinear operators, which would amplify the noise for the AoP, removing noise to precisely restore the polarization information is a significant task<sup>107,119,120</sup>. CNN have distinct advantages regarding extracting image features and hidden structures; thus, they are suitable for image restoration and enhancement in noisy environments. Li et al. employed deep neural networks to significantly suppress the noise in polarimetric images and enhance the image

quality<sup>45</sup>. However, all channel-wise features are treated equally, resulting in a lack of flexibility. Inspired by the attention mechanism, Liu et al. proposed an attention-based residual neural network to remove noise and restore the polarization information of polarimetric images<sup>77</sup>, as Fig. 7. Therefore, the proposed methods can suppress noise more effectively and restore polarization information more accurately, as shown in Fig. 7(b)<sup>109</sup>. Additionally, SSIM is used to compare the quality of images obtained by different methods. Focusing on the denoising of Mueller matrix images, Yang et al. built a deep residual U-Net that incorporated channel attention with many paired low- and high-SNR Mueller matrix images<sup>78</sup>. The ground truth is obtained based on the low equally weighted variance (EWV), which can be expressed as:

$$EWV = \frac{220}{N} \left( \sigma^2 + \frac{m_{00}}{4} \right), \quad (5)$$

where  $\sigma^2$  is the variance of Gaussian noise and  $N$  denotes the number of states of polarization. The larger the value of  $N$ , the higher is the signal-to-noise ratio. The proposed method can effectively resolve the conflict between the measurement accuracy and acquisition time.

The fused images contain more information than



**Fig. 7 | Liu et al. method.** (a) Architecture of the proposed method. (b) Enhanced results and corresponding comparisons with mainstream methods. Figure reproduced with permission from ref.<sup>77</sup>, Optical Society of America Publishing AG.

single physical properties because the mixed information describes various characteristics. The fusion of polarization and intensity is the most common method used in practical applications<sup>121–125</sup>. The intensity image describes the reflectivity and transmissivity of the object, whereas the polarization image describes the texture details, material properties, shape, shading, and roughness. These two types of images provide complementary information from different aspects to obtain images with rich physical features and improve the performance of practical tasks. Conventional fusion methods are challenging to handle various scenes because a manually designed fused factor is adopted. Based on the excellent performance of CNN, Zhang et al. proposed an unsupervised deep network called PFNet to fuse the intensity and DoLP images<sup>126</sup>. The feature extraction module transforms the images of  $S_0$  and DoLP into high-dimensional nonlinear feature maps using two Dense Blocks, and the concatenation operator is used to fuse the feature maps to reconstruct the fused image using the reconstruction module. The architecture of PFNet is shown in Fig. 8(a). Therefore, the method based on deep learning outperforms other state-of-the-art methods, as shown in Fig. 8(b)<sup>122–125</sup>. SSIM is used to compare the quality of images obtained by different methods. Contrarily, they modified the architecture to enhance performance. A Dense Block is used to encode the input images and the fusion subnetwork rather than a concatenation operator to fuse the feature maps<sup>127</sup>. New loss function strategies are adopted, such as the loss between fused and input images and the loss between fused and encoded features. The proposed architecture can also be used for infrared and visible image fusion, and multi-focus image fusion.

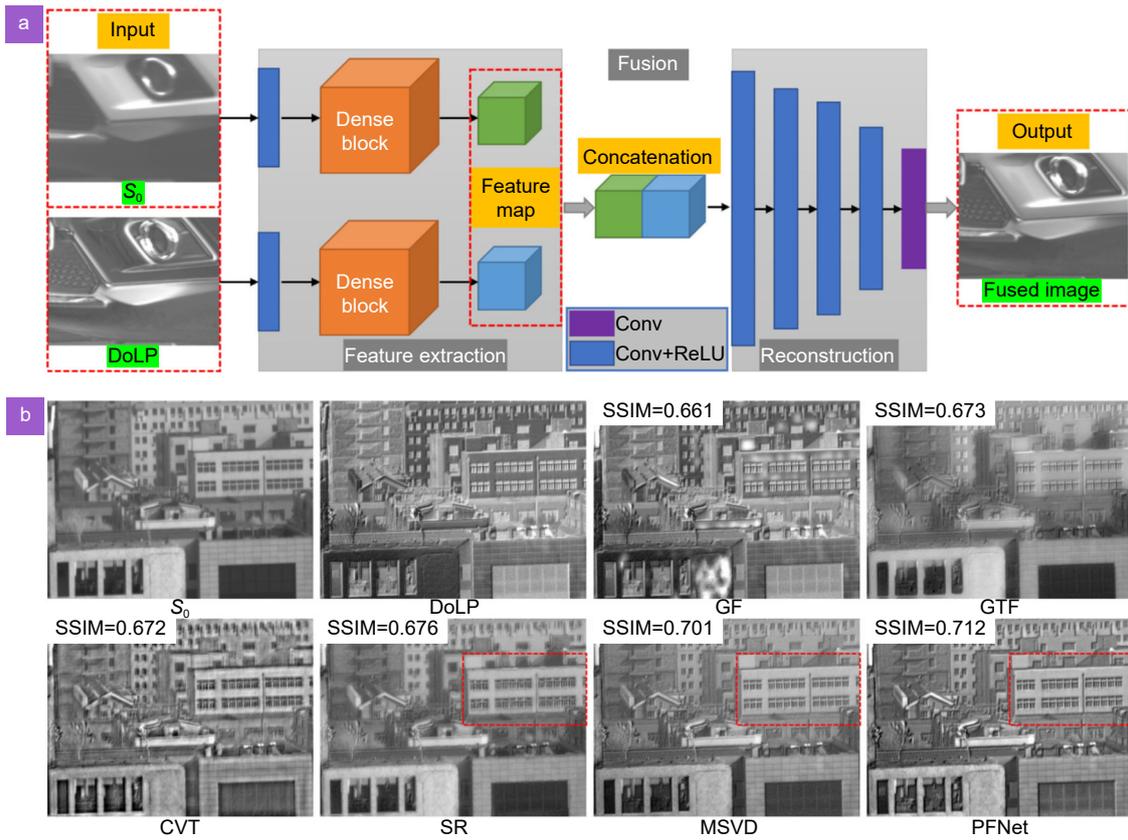
Because of the limitations of the response function of

the camera, a digital camera always captures only a limited fraction of the range, resulting in low-dynamic-range images with over- or underexposed areas that cannot reflect real-world scenes in high-dynamic-range images<sup>128,129</sup>. Ting et al. studied the relationship between polarization parameters and the exposure time of a pixel in a polarization image and trained the reconstruction framework to recover a high-dynamic-range image using polarization images<sup>79</sup>.

Other applications of restoration and enhancement are based on data-driven polarimetric imaging. The polarization parameters and other physical properties are interconvertible. Liu et al. used a deep neural network to transform holographic images reconstructed from a single state of polarization into images equivalent to those captured using a single-shot computational polarized light microscope<sup>75</sup>. Si et al. fed Stokes images to a well-designed deep learning network to generate Mueller matrix-based parameter images, such as linear retardance and diattenuation parameters<sup>58</sup>.

### Polarimetric descattering

Clear vision in scattering media is critical for various applications such as industrial and civil fields<sup>130</sup>, traffic surveillance systems<sup>131</sup>, automatic drives<sup>132</sup>, remote sensing<sup>133</sup>, rescue operations<sup>134</sup>, seabed mapping<sup>135</sup>, monitoring of marine species migration and coral reefs, and scene analysis<sup>136</sup>. However, when capturing images in a scattered environment, the visibility of objects is typically sharply degraded, which is caused by the scattering of suspended particles such as clouds, water, haze, smoke, smog, fog, and mist in the air, soil particles, floating excrement of marine animals, algae, and mineral salt in underwater scenes. The backscattered light was mixed



**Fig. 8 | The unsupervised deep network PFNet proposed by Zhang et al.<sup>126</sup>** (a) The architecture of PFNet. (b) Fused results compared with conventional methods. Figure reproduced with permission from ref.<sup>126</sup>, Optical Society of America Publishing AG.

with the object signal during its propagation towards the camera. Because the physical properties of imaging in haze and underwater environments are similar, we analyzed the dehazing and descattering processes based on deep learning and polarimetric imaging.

**Polarimetric imaging model in scattering media**

Based on the atmospheric transport model<sup>137,138</sup>, image  $I$  captured by the camera after propagation in scattering media consists of two components: (1) direct transmission signal  $D$ , which represents the light that an object reflection is scattered by the media during propagation towards the camera; and (2) backscattered light  $A$ , which denotes the light backscattered by the particles in the object light line of sight without an object signal<sup>139</sup>. The model can be expressed as:

$$I = D + A . \tag{6}$$

First, as the light reflected from the object propagates towards the camera via the medium, the object radiance suffers from absorption and scattering, yielding a degraded signal. This process is described as follows:

$$D = L_{\text{object}} e^{-\beta z} , \tag{7}$$

where  $z$  is the distance between the object and the camera;  $\beta$  is the degraded coefficient, and  $L_{\text{object}}$  is the original object signal not attenuated by the scattering media along the line of sight. The transmitted process  $e^{-\beta z}$  is the delay of exponential function, which also expressed as  $t$ .

Second, backscattered light  $A$  is an undesired component that veils the object light to reduce the contrast of the image. The backscattered light can be expressed as:

$$A = A_{\infty} (1 - e^{-\beta z}) , \tag{8}$$

where  $A_{\infty}$  is the saturated backscattering light as the distance increases. We aimed to reconstruct the original object signal  $L_{\text{object}}$  by combining Eq. (6) and (7). Then,  $L_{\text{object}}$  can be expressed as:

$$L_{\text{object}} = \frac{I - A}{1 - A/A_{\infty}} , \tag{9}$$

as a result, the estimation of two unknown components  $A$  and  $A_{\infty}$  is key to reconstructed  $L_{\text{object}}$ . Schechner et al. proposed a polarization descattering model based on an atmospheric scattering model. The proposed method obtains two orthogonally polarized images,  $I_{\perp}$  and  $I_{\parallel}$  through two orthogonal polarization states of the polarizer<sup>139</sup>. This method assumes that the air light is usually

partially polarized, whereas the object is not polarized. Thus, the captured images can be described by:

$$I_i = \frac{D}{2} + A_i, \quad i \in [\perp, \parallel]. \quad (10)$$

Because the object is not polarized, the DoP of the images is equal to that of the backscattered light. Furthermore, the backscattered light is estimated using the polarization model:

$$\hat{A} = \frac{IP}{P_A}, \quad (11)$$

where  $P_A$  and  $P$  are the DoPs of the backscattered light and captured images, respectively. The saturated backscattering light  $A_\infty$  is the mean of the background region selected in the image where there is no object and the  $P_A$  is calculated by the selected region which is the DoP of total backscattering light. Finally, the reconstructed  $L_{\text{object}}$  can be expressed as:

$$L_{\text{object}} = \frac{I - A}{1 - A/A_\infty} = \frac{I(1 - P)}{1 - IP/A_\infty P_A}. \quad (12)$$

A polarimetric imaging model is a physical, low-cost, and effective method for restoring clear images in a scattering environment. However, the method is based on the following assumptions:

- 1) The backscattered light at infinity is uniform, but in the real world, clouds, the solar radiation angle, and other factors may influence the distribution.
- 2) The DoP of the object is not polarized, but it does not make sense.
- 3) The DoP of the backscattered light is constant; however, it must be spatially variable.
- 4) The polarized direction of the image is equal to that of the backscattered light and the object signal, but there are possible differences.
- 5) The degraded coefficients of backscattering light and object signal are similar; however, Akkaynak et al.'s study has proved that attenuation coefficient of object signal depends on the distance  $z$ , reflectance  $\rho$ , spectrum of ambient light  $E$ , spectral response of the camera  $S_c$ , and beam attenuation coefficients of the water body  $\beta_b$ ; however, the backscattering light is related to  $E$ ,  $S_c$ ,  $\beta_b$  and scattering of the water body  $b^{84,140,141}$ .

Several improved methods focusing on this insufficiency have been proposed. For example, Huang et al. estimated the polarization difference image (PDI) of an object signal using feasible region fitting to overcome the limitations of the second assumption. Hu et al. estimated the spatial distributions of the DoP of an object and

backscattered light by extrapolation fitting to overcome the first and third assumptions. However, selecting the fitting function when the light was irregular was challenging<sup>142</sup>. Wei et al. considered the difference between the AoP of backscattering light and the object signal, using independent component analysis (ICA) to estimate the object signal with nonuniform polarization characteristics to avoid the limitation of assumptions (3) and (4)<sup>143</sup>. These methods are based on physical models that lack the robustness and effectiveness of a single method in complex scenes because accurate estimation of key parameters may not be achieved. However, the latent factors influencing image quality have not been explored.

Deep learning methods based on CNN are adept at extracting hidden features and fitting the nonlinear relationships between backscattering light and object signals. Thus, it is a promising choice for descattering and dehazing. Polarimetric data-driven descattering methods combine polarization information and deep learning. Moreover, the existing model can guide the training of the descattering network to combine its advantages. Three different pipelines were used in the data-driven polarization descattering method. First, the end-to-end architecture without the physical model, which always uses the polarization images to feed into the network. Second, the physical-model-guided network methods, which are guided by the existing or proposed but non-participation in network training; Third, the physical-model-integrated network methods, which integrate the physical model into the network to train the descattering network together. The following section provides a detailed explanation of these three perspectives.

### End-to-end descattering network

The end-to-end architecture is a common structure that enhances image quality and relies on higher-order nonlinear representations. The descattering process of a network is the fitting of the descattering transmission of a high-order function, which has been successfully applied to complete scattering removal using intensity information<sup>136,144–146</sup>. In recent studies, the introduction of polarization information has proven that the input of polarization images into the network can improve the quality of qualitative and quantitative evaluations. This section discusses the polarization end-to-end descattering network.

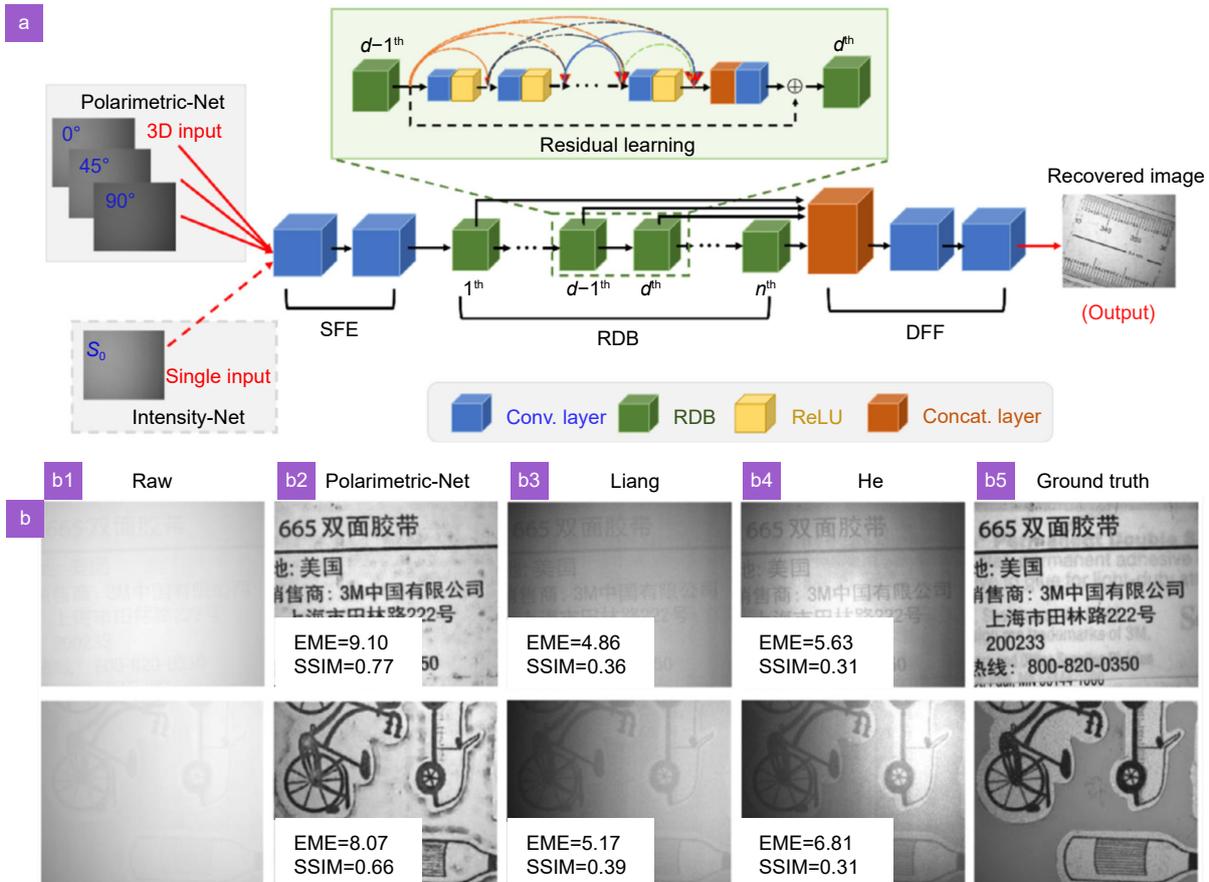
In 2020, Hu et al. first employed a deep learning technique in polarimetric underwater imaging<sup>16</sup>, as shown in Fig. 9(a), which is a typical end-to-end descattering

architecture. The three dimensional inputs of different linear polarization orientations with  $0^\circ, 45^\circ, 90^\circ$  are fed into an end-to-end polarimetric dense network (PDN) rather than intensity images, which contains three main components: Shallow Feature Extraction (SFE) used to extract shallow polarization features, Residual Dense Block (RDB) as the basic structure to connect with each other, and concatenated by Dense Feature Fusion (DFF). The third crucial component was utilized to fuse all the features and output the descattering results. A dataset containing abundant polarization image pairs was built using a commercial DoFP camera, and a water tank filled with milk was used to capture turbid and clear object signals. The same network structure, based only on intensity images, was trained to verify the significance of the polarization information. Compared with the intensity network and existing methods, the polarization network demonstrated higher values of image contrast (IC), measure of enhancement (EME), PSNR, and SSIM, which indicate higher image quality<sup>6,147</sup>, as shown in Fig. 9(b).

Another end-to-end architecture was proposed by

Zhang et al., which contained four pairs of networks consisting of polarized and gray versions. Furthermore, this method assumes that the intensity and polarization images are information streams, two types of information flows in their respective networks, and join together at the end of the model. Furthermore, the addition of polarization information to gray information and feeding these two parts into fused networks, called DENSE-U-NET BLOCK, at the forefront of the network could achieve better results than the front or end of the network<sup>70</sup>. The experiments with different turbidity levels performed better than the other methods and demonstrated the excellent robustness of the proposed method.

In remote sensing, the scattering medium significantly impacts on the results of object reconstruction, even producing speckle patterns. Obtaining the correspondence between the original object and the imaging process is challenging and crucial. Li et al. combined an object's polarization information with a modified U-net-based deep learning network (MU-DLN) to retrieve the original object's information influenced by the scattering



**Fig. 9 | The flow chart and results of the Hu et al. method, which is a typical end-to-end descattering architecture<sup>46</sup>. (a) The architecture of proposed method. (b) Comparison of the enhanced images by different methods. Figure reproduced with permission from ref.<sup>46</sup>, Elsevier BV.**

medium<sup>66</sup>. Data were acquired using a Monte Carlo simulation system and deep learning technology to learn a physical model of the scattering process. The experimental results show that the object's information for the Q-component can be reconstructed very well because of the suppression of scattering light and highlighting of ballistic light. Several fixed optical thickness environments are tested to reflect the superiority of the trained MU-DLN.

### Physical-model-guided descattering network

The physical-model-guided descattering network is trained to remove the scattering effect guided by the physical model but does not participate in network training. Based on the physical model, the theoretical feasibility was proven before guiding the design of the network architecture. Therefore, the middle stage of combining the physical priors and models is instructive for further fusing the physical model into the design of the network pipeline.

Guided by a physical imaging model, Ren et al. trained a lightweight dehazing CNN to rapidly process turbid images, comparing it with conventional dehazing methods and introducing additional circular polarization information<sup>148</sup>. Furthermore, this is the first time circular polarization information has been fed into a network. Two unknown parameters in the polarimetric descattering model resulted in an underdetermined function by directly generating these parameters. Therefore, the proposed method combines these two parameters into a single formula to avoid an underdetermination problem and minimize the reconstruction error. The new parameter  $K$  is expressed as:<sup>149</sup>

$$K = \frac{(I - A_\infty)/t + A_\infty - 1}{I - 1}. \quad (13)$$

The Eq. (12) can be rewritten as an  $L_{\text{object}} = KI - K + 1$ , after which the descattering process is viable once the parameter  $K$  is obtained. Tests were conducted in different turbid environments to verify the feasibility of the proposed method, and the results indicated the effectiveness and high efficiency of the lightweight architecture. Subsequently, Ding et al. adopted a multi-polarization fusion adversarial generative network to enhance turbid images<sup>47</sup>. Compared with the conventional model, the proposed method introduces an angle of polarization to calculate the backscattered light, expressed as:

$$\hat{A} = \frac{I_0 - I[1 - P]/2}{P_A \cos^2 \phi_A}, \quad (14)$$

where the  $\phi_A$  is AoP of backscattering light computed by the selected background region. They built the first color polarization image datasets in the natural underwater environment, which selected the visually better enhanced results among the results produced by several conventional methods<sup>150-154</sup>. Compared with the underwater results by four supervised data-driven polarimetric methods mentioned above, the experimental results in laboratory simulated by the milk have huge improvement; however, the natural results are improved to a lesser extent because the complicated environment would increase the diversities of known or hidden parameters and then result in the networks become more generalized with higher robustness but lower performance in a specific example.

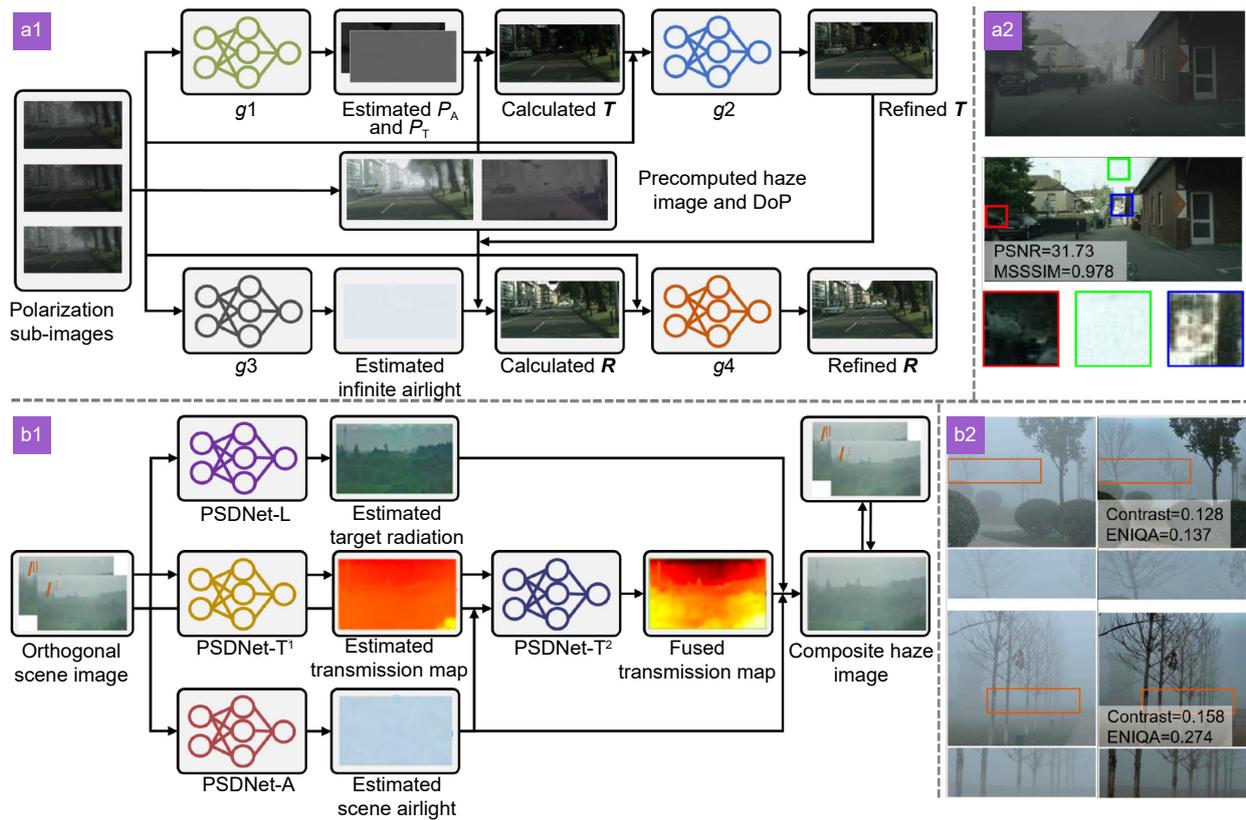
### Physical-model-integrated descattering network

The physical-model-integrated descattering network integrates the descattering model into the network as the backbone to guide the descattering process, which introduces constraints compared to the physical-model-guided descattering network. Therefore, the main task of the network is to generate or refine specific parameters before they are utilized to generate improved results. Furthermore, the physical model and its inverse process can form a self-supervised closed loop to achieve improved performance.

To further combine the physical formation model with deep learning methods, some researchers have embedded existing dehazing approaches into the proposed pipeline. Zhou et al. proposed a robust polarization-based dehazing architecture with a generalized physical formation model that requires no specific clues to estimate the required physical parameters or handcrafted priors<sup>48</sup>. Figure 10(a1) and 10(a2) show the network architecture and corresponding example with evaluation index (PSNR and multi-scale SSIM). The transmitted light  $D$  ( $T$  in Fig. 10(a1)) and original scene radiance  $L_{\text{object}}$  ( $R$  in Fig. 10(a2)) can be calculated using the following equations:

$$D = \frac{PI - IP_A}{P_T - P_A} \quad \text{and} \quad L_{\text{object}} = \frac{DA_\infty}{A_\infty - (I - D)}, \quad (15)$$

where  $P_T$  and  $P_A$  define the DoP of the transmitted light and backscattered light, respectively, which are estimated by the subnetworks. The symbols  $I$  and  $P$  are the



**Fig. 10 | The methods combining the physical formation model with deep learning methods and corresponding examples. (a1) Zhou et al. method and (a2) the corresponding example. (b1) Shi et al. method and (b2) the corresponding examples.**

haze image and its DoP, respectively, which can be calculated from the polarized images. The DoPs of the object signal and scattered light can be generated by subnetwork  $g_1$  before estimating the transmitted light using the imaging model. The refined light transmitted by subnetwork  $g_2$  is utilized to calculate the original scene radiance, in which the saturated backscattered light is obtained by subnetwork  $g_3$ . Finally, the refined subnetwork  $g_4$  was adopted to generate the refined results. The raw polarization direction images join each subnetwork.

The generation of the synthetic dataset was instructive. Clear images with depth and semantic segmentation maps must be provided for the generation process. The Foggy Cityscapes-DBF dataset was eligible, and reasonable values of the corresponding parameters were set to generate the synthetic dataset<sup>155–157</sup>. Gaussian noise was introduced to make them spatially variant and conform to real-world scattering conditions to improve the robustness of the network<sup>158,159</sup>.

Contrarily, Shi et al. processed a polarization-based self-supervised dehazing network called PSDNet to eliminate the influence of haze on images<sup>72</sup>. Figure 10(b1) and 10(b2) show the proposed network architecture and

corresponding results, with assessment criteria (IC and entropy-based no-reference image quality assessment (ENIQA)) which consist of three subnetworks to compute the object radiation, transmitted light, and backscattered light. The pipeline processed a self-supervised closed loop to optimize the network. The end-to-end descattering network is part of the total pipeline, which effectively reduces the scale of the network and enhances performance. Additionally, a secondary product with an accurate transmission map was produced, which may be helpful for other computer vision tasks, such as 3D reconstruction. Several experiments demonstrated that the proposed architecture can effectively improve the visibility of object details and is highly robust for the scene.

Because it is hard to capture the ground truth corresponding to the object in the underwater environment, unsupervised learning polarimetric underwater methods were proposed. Zhu et al. synergistically make use of an untrained network and polarimetric imaging formation model to recover images from scattering in underwater scenarios without requiring additional datasets<sup>49</sup>. There are two stages during the network training. First, the raw input images are input to the network  $\mathcal{R}(\cdot)$  to generate

the optimized image. Furthermore, using the imaging formation model  $\mathcal{M}(\cdot)$  to recalculate the degraded image. After the circular process, the loss of raw image and calculated image is adopted to optimize the network:

$$[\mathcal{R}(\cdot; \mathbf{w}^*), \Theta^*] = \underset{\mathbf{w}, \Theta}{\operatorname{argmin}} \| \mathcal{M}\{\mathcal{R}(I(x, y); \mathbf{w}), \Theta\} - I(x, y) \|_2^2, \quad (16)$$

where  $\mathbf{w}$  is the weights,  $\Theta$  is parameters of image formation model including  $\{\alpha, \eta, p_s^\circ, S_\infty\}$ . These parameters are estimated using a neural network, where the initial value is selected by traditional methods. Specifically,  $\alpha$  is a compensation factor for the water absorption is set to 0.99 as the subject was placed in a comparably shallow position with pure water in the underwater environment, where the influence on the absorption of natural light could be ignored.  $\eta$  is a bias factor and  $p_s^\circ$  is degree of the scattering light, where the range of  $\eta$  is from 1 to  $1/p_s^\circ$ . The measurement value of  $p_s^\circ$  is set to 0.8333 and  $\eta=1.13$  to meet the range of  $1 \leq \eta \leq p_s^\circ$ .  $S_\infty$  is the scattered light radiation at an infinite distance estimated by the intensities of the brightest region of the no object.

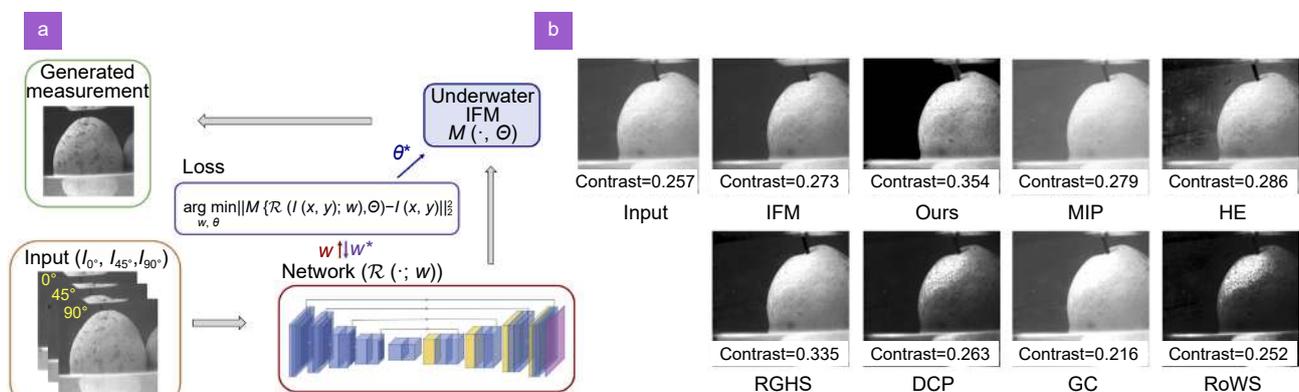
The proposed methods not only overcome the acquisition of polarization characteristics of the environment and object in the conventional process but also minimize the dependency on datasets even when training on only one image. Moreover, the mismatch between the model and a real scene lacking environmental priors is significantly reduced. Figure 11(a) and 11(b)<sup>49</sup> show the proposed architecture and a visual comparison among the different descattering methods<sup>147,160–163</sup> together with statistical index of image contrast. The method represents a pioneering attempt in the realm of unsupervised descattering imaging. However, its capacity for enhancing imaging outcomes remains somewhat limited.

Yang et al. trained a network to inpaint backscattered light with different polarization orientations, which was used to calculate the DoP and AoP of the background light<sup>76</sup>. Furthermore, this is another unsupervised method that does not require clear ground truth. The primary task of the proposed method is to calculate the complete backscattered light. After removing the region of the object in the captured polarized images using the GrabCut algorithm, the incomplete image, randomly erasing a region, is input to the network to generate a mission region before using a filtering method that compares the gray value of one pixel with those in other nearby pixels and replacing the singular point with the averaged value in a  $7 \times 7$  square. Furthermore, a clear image is calculated using the following modified function:

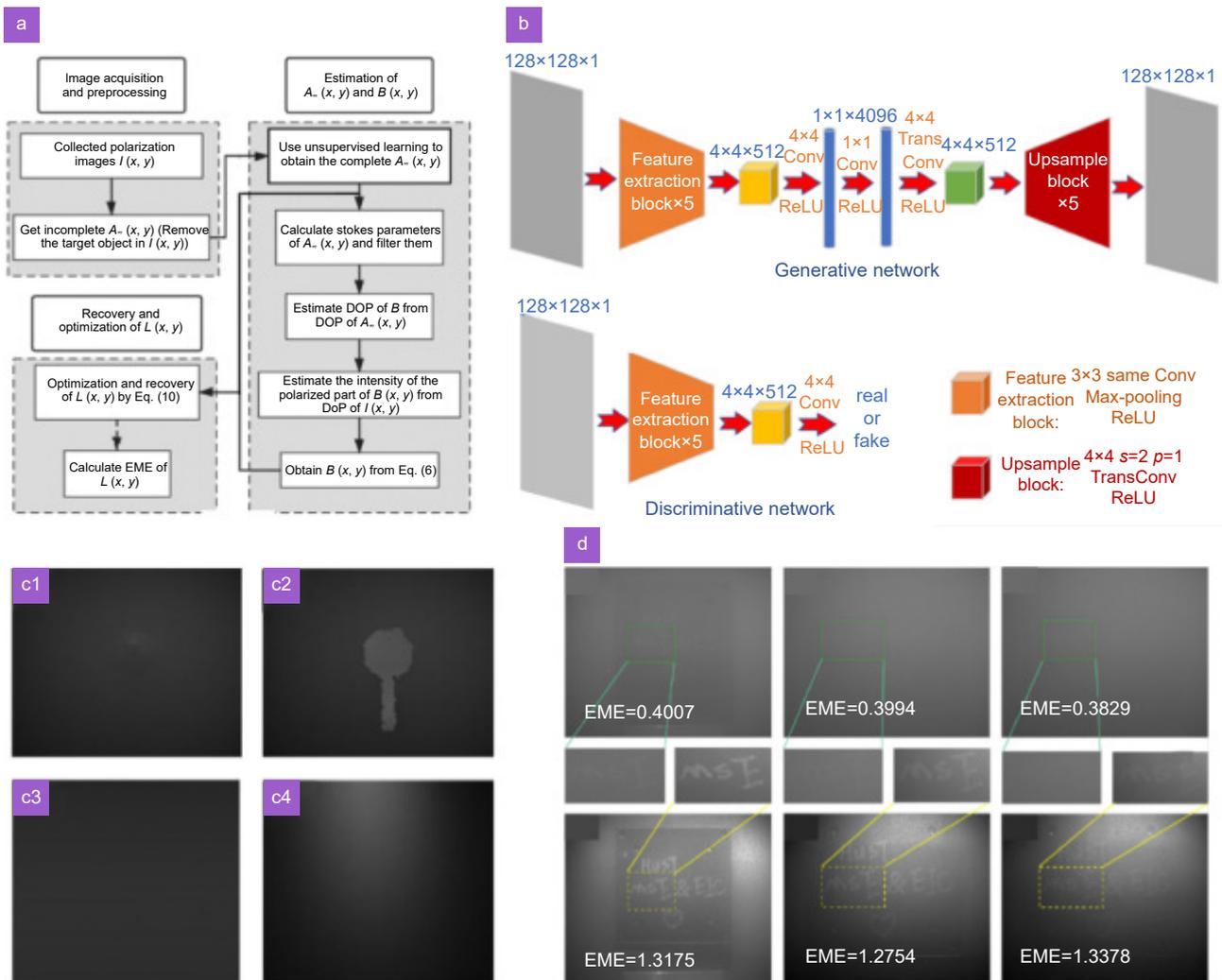
$$L_{\text{object}} = \frac{I - (A/\varepsilon)}{1 - (A/\varepsilon A_\infty)}, \quad (17)$$

where the  $\varepsilon$  is bias factor. Consequently, the object radiance was optimized and recovered based on the underwater image recovery process. The proposed method has a much lower cost for preparing the training datasets and demonstrates the capability of recovering underwater images under different nonuniform optical fields. The flow chart, proposed architecture of the Yang et al. method<sup>76</sup>, corresponding results, and comparisons with existing methods<sup>142,147,164,165</sup> are shown in Fig. 12.

Data-driven polarization descattering methods have been gradually introduced into physical models to guide network training, which resolves the limitations of traditional methods. Data-driven methods learn more comprehensive features and adjust them for more complex media. Furthermore, it improves the training performance and provides another training process for a self-supervised closed loop to optimize the network.



**Fig. 11 | Unsupervised underwater descattering method.** (a) Architecture of Zhu et al. method<sup>49</sup> and (b) visual comparison among different de-scattering methods. Figure reproduced with permission from ref.<sup>49</sup>, Optical Society of America Publishing AG.



**Fig. 12 | Unsupervised underwater descattering methods and corresponding results.** For the equations and variables, see the ref.<sup>76</sup>. (a) The flow chart. (b) Proposed architecture of Yang et al. method. (c) The results of generated backscattering light; (c1) the captured polarized image; (c2) the generated backscattering image of proposed trained network; (c3) the backscattering image after smooth filtering; (c4) the corresponding ground truth. (d) Final descattering results and comparisons with other methods. Figure reproduced with permission from ref.<sup>76</sup>, Optical Society of America Publishing AG.

Additionally, the physical model in the network may generate extra parameters that are probably helpful for other tasks, such as depth maps, backgrounds, and inherent coefficients of media. The acquisition of dataset is key for network training in scattering media. The synthetic and generated methods may solve the problem despite remaining crucial and challenging in the future.

### Three-dimensional shape reconstruction

By analyzing the interactions between light and surface geometry, we can reconstruct the 3D shapes of objects<sup>166,167</sup>, where polarization is crucial. Natural illumination becomes partially polarized after reflection from an object's surface. Polarized reflection implies shape information because the Fresnel equations relate

the DoP, AoP, and micro-surface zenith and azimuth. An intrinsic drawback of deriving a shape from polarization is the ambiguous estimation of surface orientation. The suitable arctangent function in the model results in a multivalued azimuth, commonly known as azimuth ambiguity. Cues from various aspects, such as geometry<sup>168–170</sup>, spectrum<sup>171,172</sup>, photometry<sup>173–176</sup>. However, relying only on a physical-based imaging model, recovering the shape with high accuracy remains challenging under nonlaboratory conditions. The excellent nonlinear representation ability of deep neural networks can narrow the gap between ideal and real-world conditions. This section reviews the existing 3D shapes from polarization methods combined with deep learning (DL).

### Principles of polarization 3D shape reconstruction

The surface shape changed the polarization states of the incident illumination, providing the possibility of recovering the shape from polarization. In polarization detection, polarization information can be obtained using a camera and a rotated linear polarizer mounted in front of it or a camera with a pixelated polarizer. The captured image intensity varied sinusoidally.

$$I(\phi_{\text{pol}}) = \frac{I_{\text{max}} + I_{\text{min}}}{2} + \frac{I_{\text{max}} - I_{\text{min}}}{2} \cos(2\phi_{\text{pol}} - 2\phi), \quad (18)$$

where  $\phi_{\text{pol}}$  denotes the angle of polarizer axis relative to a chosen reference orientation,  $\phi$  denotes the azimuth angle of micro-surface of an object,  $I_{\text{max}}$  and  $I_{\text{min}}$  refer to the observed maximum and minimum intensity during  $\phi_{\text{pol}}$  from 0 to  $\pi$ . A whole  $2\pi$  period of sinusoidal function results in a  $\pi$  ambiguity. For instance, if  $\phi = \phi_{\text{pol}}$ , the maximum intensity is obtained. However, the minimum intensity corresponds to two azimuth angles, i.e.  $\phi \pm \pi/2$ . The  $\pi$ -ambiguity problem is the one of shape from polarization.

The polarization state of the reflected light directly depends on the reflection type occurring over the surface, which is primarily specular or diffuse reflection, as shown in Fig. 13.

The Fresnel equations describe how incident light changes when propagating in media with different refractive indices. When specular reflection dominates, the DoP of the specular reflection is calculated using the Fresnel reflection coefficients:

$$P_s = \frac{R_{\perp} - R_{\parallel}}{R_{\perp} + R_{\parallel}}. \quad (19)$$

Combined with the Fresnel function, it has the expression in Eq.(20)<sup>177</sup>:

$$P_s = \frac{2\sin^2\theta\cos\theta\sqrt{n^2 - \sin^2\theta}}{n^2 - \sin^2\theta - n^2\sin^2\theta + 2\sin^4\theta}, \quad (20)$$

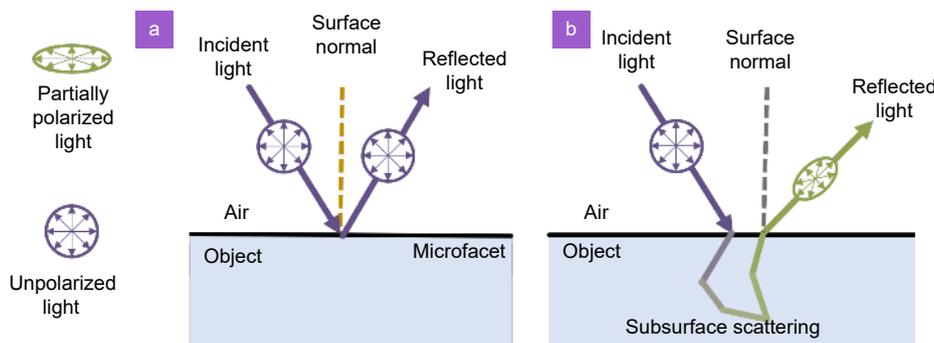


Fig. 13 | Polarization of specular reflection and diffuse reflection. (a) Specular reflection. (b) Diffuse reflection.

where  $n$  denotes the refractive index,  $\theta$  refers to the zenith angle, assuming that  $\eta_i=1$  because in most conditions light is incident from air. The refractive index of specular surface is denoted  $\eta_t = n$ . Because the azimuth angle is perpendicular to the phase of the specular polarization<sup>178</sup>, leading to the  $\pi/2$  shift of azimuth angle. This is another ambiguity problem regarding shape owing to polarization.

Material object reconstruction is more complicated than that of regular specular surfaces. The refraction index of metal is a complex number defined as  $\hat{n} = n(1 + i\kappa)$ , where  $\kappa$  is the attenuation coefficient. And then Eq. (20) can be derived as:

$$P_{\text{sm}} = \frac{2n \tan\theta \sin\theta}{\tan^2\theta \sin^2\theta + |\hat{n}|^2}. \quad (21)$$

Diffuse reflection originates from the light refracted by the shallow surface of an object, in which it is partially polarized owing to the irregular interactions between the light and the interior particles. Therefore, the DoP of diffuse reflection is determined by the Fresnel transmission coefficients. The relationship between the DoP and the Fresnel coefficients of diffuse reflection is defined as:

$$P_d = \frac{T_{\parallel} - T_{\perp}}{T_{\parallel} + T_{\perp}} = \frac{R_{\perp} - R_{\parallel}}{2 - R_{\perp} - R_{\parallel}}. \quad (22)$$

Under these conditions, light is refracted to the air in the object. Therefore  $\eta_t=1$ , i.e. the refraction index of air, and  $\eta_i=n$ , i.e., the refraction index of diffuse surface. Eq. (23) can then be derived as:

$$P_d = \frac{\left(n - \frac{1}{n}\right)^2 \sin^2\theta}{2 + 2n^2 - \left(n + \frac{1}{n}\right)^2 \sin^2\theta + 4\cos\theta\sqrt{n^2 - \sin^2\theta}}. \quad (23)$$

Given the azimuth angle  $\phi$  and zenith angles  $\theta$ , the normal vector of shape surface at any point could be expressed as<sup>179</sup>:

$$\begin{aligned} \mathbf{n}(\mathbf{u}) &= [n_x(\mathbf{u}), n_y(\mathbf{u}), n_z(\mathbf{u})]^T \\ &= [\sin\phi(\mathbf{u})\sin\theta(\mathbf{u}), \cos\phi(\mathbf{u})\sin\theta(\mathbf{u}), \cos\theta(\mathbf{u})]^T, \end{aligned} \tag{24}$$

where  $n_x(\mathbf{u}), n_y(\mathbf{u}), n_z(\mathbf{u})$  are the normal vectors of element surface  $\mathbf{u}$ . The normal vector can be expressed using the surface gradient, as shown in Eq. (25)

$$\mathbf{n}(\mathbf{u}) = \frac{1}{\sqrt{p(\mathbf{u})^2 + q(\mathbf{u})^2 + 1}} [-p(\mathbf{u})^2 \quad -q(\mathbf{u})^2 \quad 1]^T, \tag{25}$$

where  $p(\mathbf{u}) = \partial z_x(\mathbf{u}), q(\mathbf{u}) = \partial z_y(\mathbf{u})$ , i.e.,  $\nabla z(\mathbf{u}) = [p(\mathbf{u}) \quad q(\mathbf{u})]^T$ . Finally, the surface shape  $z(\mathbf{u})$  was reconstructed.

The shape of the polarization differs according to the type of reflection. Compared to specular reflection, the DoP and zenith angles for diffuse reflection are one-to-one maps. However, diffuse reflection makes the reconstruction process challenging because of the lower signal-to-noise ratio and higher dependence of the DoP on the refractive index. For specular reflection, the mirror-like surface maintains a relatively uniform direction and phase, thus avoiding the influence of random noise. However, specular reflection leads to more ambiguous problems than diffuse reflection does. Moreover, it is challenging to find an object when the viewing direction exceeds the reflection direction. Table 2 compares the advantages and disadvantages of the shape from polarization based on both specular and diffuse reflections.

The shape of the polarization depends on the estimated azimuth and zenith angles. Ambiguity is a critical challenge. Regarding the azimuth angle, two phase angles with a  $\pi$  shift are derived from the period of the sinusoid function. For specular reflection, the azimuth angle would be retrieved with  $\pm \pi/2$  operation. However, the two zenith angles of the ambiguous solution are determined for a given DoP in the specular reflection, which cannot be excluded without other information. These contributing factors result in high error rates and limitations in the generalization to mixed materials and light-

ing conditions using only polarization images.

In addition to the ambiguity problem, other limitations to the shape owing to polarization exist. For example, estimating the zenith angle requires an unknown prior for the refractive index, which limits the reconstruction of complex objects and natural scenes. Second, when the zenith angle is close to zero, the influence of noise increases because the DoP is small. Third, mixed reflections were common in real-world scenarios. Moreover, achieving satisfactory reconstruction results in complex scenarios by using the linear superposition of a single physical model is challenging. Finally, the discontinuous depth is also a significant challenge for recovering the shape from the derived surface normal by integration. Consequently, the introduction of other information is essential to avoid the problems mentioned above and expand the application fields.

Typical methods include a combination of heuristic priors, such as the boundary and convexity of objects<sup>180</sup>, shading<sup>181</sup>, and photometric stereo<sup>64</sup>, but noise is a major limitation. This complicated calculation amplifies the noise, leading to a degraded texture or profile in the recovered shapes. Data-driven imaging is powerful for 3D imaging with a nonlinear modeling ability. Furthermore, this primarily depends on the semantic information of the image. Under the guidance of the physical model, this brings new possibilities for shapes from polarization.

### Data-driven shape from polarization in single reflection

In certain situations, the reflection can be purely specular or diffuse. For instance, in human face recognition or clothed body reconstruction tasks, the skin, clothes, and other human tissues are diffuse surfaces, and specular reflection is negligible. In transparent object reconstructions, specular reflection dominates.

For 3D clothed human shape reconstruction with clothing details, Zou et al. introduced polarization images and two ambiguous normal maps into the designed

**Table 2 | Comparisons of specular reflection and diffuse reflection.**

	Advantages	Disadvantages
Specular reflection	<ul style="list-style-type: none"> <li>• Higher signal-to-noise ratio</li> <li>• Lesser influence of the refraction index</li> </ul>	<ul style="list-style-type: none"> <li>• Suffer from more ambiguity problem</li> <li>• Lower viewing angle</li> <li>• Lesser application fields</li> </ul>
Diffuse reflection	<ul style="list-style-type: none"> <li>• Monotony of degree of polarization about refraction index</li> <li>• Wide viewing angle</li> <li>• More application fields even for scene in wild</li> </ul>	<ul style="list-style-type: none"> <li>• More influence of noise</li> <li>• Worse visual effects</li> <li>• Stronger influence of refraction index</li> </ul>

network<sup>65</sup>, as shown in Fig. 14. Specular reflection was omitted because of the rough surface of the clothing. Owing to the azimuth ambiguity problem, two possible maps were resolved and input into the network as physical priors. The two ambiguous normal maps,  $\mathbf{n}_1$  and  $\mathbf{n}_2$  are classified into three categories:  $\mathbf{n}_1$ ,  $\mathbf{n}_2$ , and background. Each pixel point was classified as belonging to one of these three categories and then merged into  $\mathbf{n}_3$  using Eq. (26) with probabilities  $p_0, p_1$  and  $p_2$ .

$$\mathbf{n}_3 = (1 - p_0) \cdot \frac{p_1 \mathbf{n}_1 + p_2 \mathbf{n}_2}{\|p_1 \mathbf{n}_1 + p_2 \mathbf{n}_2\|_2} \quad (26)$$

The final surface-normal prediction is refined using a denoising network. The smoothed normal concatenates the fused normal and raw polarization direction images as the input to accurately estimate the surface normal. Subsequently, the skinned multi-person linear (SMPL) representation<sup>182</sup> and deformation stage were used to reconstruct the refined 3D human shape with clothing details rather than naked.

Regarding diffuse-reflection-dominated cases, such as

human face reconstruction, Han et al. proposed a learning-based method for passive 3D face reconstruction from polarization<sup>183</sup>, as shown in Fig. 15(a). Furthermore, it derives the ambiguous normal of each microfacet over the face at the pixel level based on the polarization of the diffuse reflection. The CNN-based 3D morphable model (3DMM) generates a rough depth map of the face based on a directly captured polarization image, and is used to amend the ambiguous polarization normal and further reconstruct an accurate 3D face using Frankot–Chellappa 3D surface restoration functions. Figure 15(b) illustrates the final results, including a male face under indoor lighting, a male face under natural outdoor illumination, and an indoor plaster statue. The 3D rendering features fit well with the original appearance, and the lighting conditions had little influence. The experiments also demonstrate the benefits of introducing deep learning into 3D polarization reconstruction.

The transparent objects exhibited typical specular reflections. Shao et al. proposed a multibranch fusion

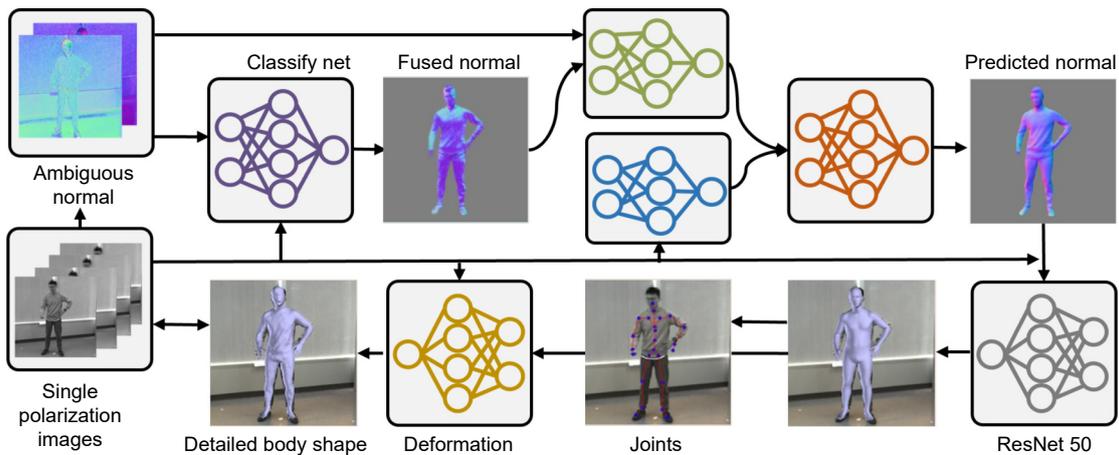


Fig. 14 | The proposed architecture of the Zou et al. method.

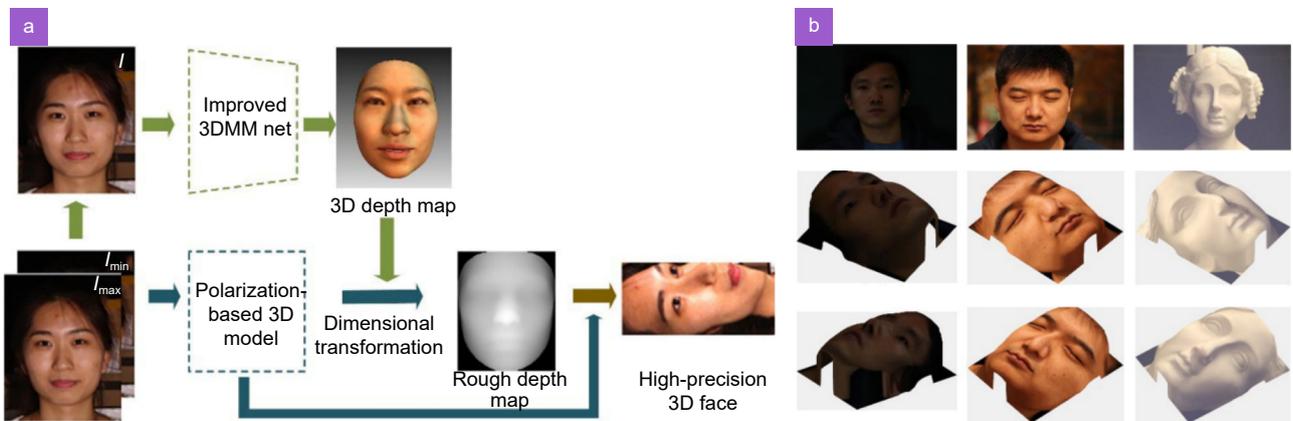


Fig. 15 | Han et al. passive 3D polarization face reconstruction method. (a) Overall schematic of the proposed method. (b) 3D polarization face reconstruction results. Figure reproduced with permission from ref.<sup>183</sup>, Multidisciplinary Digital Publishing Institute.

network to reconstruct 3D shape-transparent objects from specular reflections<sup>81</sup>. However, transmitted light is often a diffuse reflection. Therefore, separating the transmitted light is critical. The AoLP features indicate stronger background noise in areas with higher transmittance. The closer it is to the center area, the stronger the noise, as shown in Fig. 16. The physical prior confidence concept is based on intrinsic faults in the AoLP maps of transparent objects.

$$Confidence_{K,m}(i,j) = \frac{\sum_{p \in P_{i,j}} |p - \bar{p}_{i,j}|^m}{\max_{\substack{0 \leq x \leq W \\ 0 \leq y \leq H}} \sum_{p \in P_{x,y}} |p - \bar{p}_{x,y}|^m}, \quad (27)$$

where  $p_{i,j}$  represents the pixel values in the  $K \times K$  neighborhood of the point  $(x, y)$ ,  $\bar{p}_{x,y}$  is the mean of the pixel values in this area,  $m$  is the smoothing exponential term.  $H$  and  $W$  denote the height and width of the map, respectively. This physics-based prior confidence map is then input into the network as an attention map to guide the fusion of the original-polar (DoLP and AoLP maps) and physics-based prior (four ambiguous maps). The proposed method achieves optimal performance and provides a new perspective for further transparent shapes from polarization research.

### Data-driven shape from polarization in mixed reflections

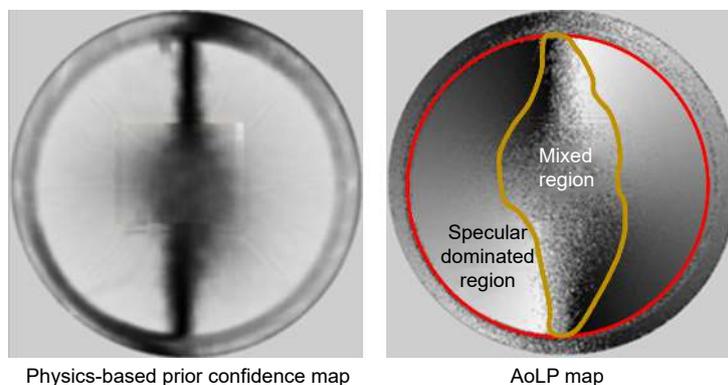
Mixed reflections stemming from the two primary conditions were prevalent in the natural scenarios. First, the reflectance of the surface determines the type of reflection: specular reflection or diffuse reflection. For materials such as ceramics, plastics, and lacquers, specular reflection dominates the highlighted areas, whereas diffuse reflection dominates the other areas. Second, objects made of different materials create reflections that vary in

each segmented area. Moreover, this problem can also be solved by reconstructing each area separately; however, the segmentation algorithm and stitching of the 3D shape of an object are huge challenges. Neural networks provide solutions to fuse explainable or inexplicable features with mixed reflections, relying on their excellent nonlinear representation ability.

The first method combines deep learning and polarization-reconstructed models; Ba et al. fed polarization images and ambiguous normal maps into the network and trained the network to learn the effective inputs from training data automatically<sup>64</sup>. The inputs were four images captured with a polarizer, and the ambiguous normal maps consisted of one diffuse and two specular ambiguous maps. The proposed method achieved the lowest test error on the tested data under the three lighting conditions compared with conventional methods.

Based on a polarimetric Bidirectional Reflectance Distribution Function (pBRDF) model and real polarization scene rendering, Kondo et al. applied rendered polarized images to train a network for an accurate surface normal estimation<sup>71</sup>. A physics-based renderer was built to simulate the polarization behavior of the rays based on the proposed pBRDF model for each material. Furthermore, it can correctly reproduce the polarization property, including the inter-reflection effect, in real-life scenes. Therefore, the synthetic-colored image and simulated polarization information, such as the phase and DoP, were fed into the CNN to estimate the surface normal. The detailed process and reconstruction results are shown in Fig. 17(a).

The proposed pBRDF model is described by the angle of incident light, plane, reflection angle, camera direction, and half-vector, which allows accurate transmission Mueller matrix modeling for arbitrary cameras and



**Fig. 16 | The physical-based prior confidence map according to the differences of polarization characteristics between transparent object and background.**



BRDF<sup>71,80</sup> will suffer poor recovery, and accurate simulation of mixed polarization remains an open challenge.

Mixed reflection dominated the outdoor scenario with multiple objects with different refractive indices. Lei et al. constructed the first real-world SfP dataset for complex scenes to train a network<sup>67</sup>. The proposed network input includes three parts: 1) four raw captured polarization images and 2) polarization feature images, including intensity, DoP, and encoded AoP by sine and cosine operations. Encoded AoP solves the problem regarding raw AoP maps being similar at two given polarization angles with  $\pi$  difference; 3) The viewing encoding to account for non-orthographic projection in scene-level SfP. The introduction of viewing encoding effectively calibrates the polarization parameters influenced by the spatially varying viewing directions.

In summary, existing data-driven polarization 3D shape reconstruction methods are always end-to-end structures guided by a physical model. The basic structure is input to the raw polarization images and other prior images, which are always ambiguous, to optimize the network, and the refined normal image is generated. In addition, other prior images are crucial in the corresponding methods for enhancing the reconstructed effects. Different inputs may also result in various network architectural designs. Therefore, as shown in Fig. 18, the inputs in the 3D shape reconstruction task are visually

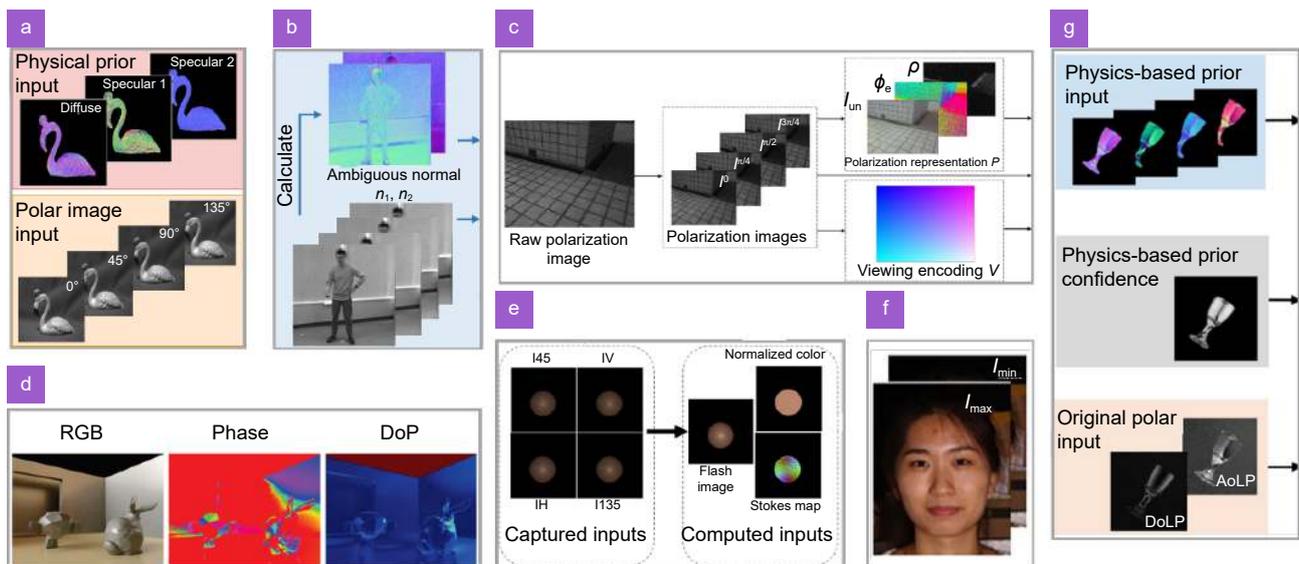
displayed.

The data-driven dataset for the ground truths was obtained in two ways. First, a Kinect depth camera is the most commonly used equipment for capturing a coarse depth map as the ground truth. Other operations may be conducted to refine the captured depth map, such as denoising, exclusion of inaccurate values, and sparse point cloud<sup>67</sup>. Second, a simulation method based on the BRDF is adopted to generate the synthetic datasets. Plausible results could be produced, and other features such as roughness and depth maps<sup>80</sup> could also be generated, which may be helpful for other computer vision tasks. However, robustness will decrease in a complicated environment.

### Reflection removal

#### Limitations of reflection removal based on polarization

Removal of reflection contamination is a challenging but critical and frequently encountered task because it may contaminate image quality. Several studies have been conducted based on diverse physical and image characteristics, but this remains an unsubstantial task<sup>14,184,185</sup>. Because the transmission image through the surface and the image reflected by the surface are simultaneously captured by a photographer, recovering two images from a single-mixture image is a highly ill-posed problem.



**Fig. 18 | The inputs of network in 3D shape reconstruction task.** (a) The raw polarization images as well as the diffuse and specular normal prior. (b) The raw polarization images and two diffuse ambiguous normal. (c) The DoFP image consists of four polarization sub-images, viewing encoding, intensity, DoP image and encoded AoP image. (d) The chromatic intensity, phase and DoP image. (e) The raw polarization images, flash image, normalized color and Stokes map. (f) The maximum and minimum polarization images. (g) The raw polarization images, physics-based prior confidence, DoLP and AoLP.

Constraints are essential to solve this problem.

The reflected light is polarized, and polarization has proven to be a feasible solution to this problem. Based on the Fresnel functions, the number of unknowns was similar to that of the input images. The reflection and transmission components can be separated by viewing the Brewster angle. Moreover, the closer to the Brewster angle, the better the reflection removal performance. Despite the incident angle having to be known, obtaining it in the real world is challenging. Additionally, the robustness and generalization of existing polarization-based methods cannot meet the requirements of real-world high-quality imaging, considering the various viewing angles, complex refractive indices, smoothness, and local curvature of the surface.

Deep learning methods based on CNN are excellent for extracting hidden features. Furthermore, it enables the prediction of potential prior information from captured images and demonstrates good performance in the reflection removal task. Introducing polarization and imaging models into the network can improve removal performance and expand the diversity of datasets and architectures. However, acquiring the ground truth dataset is crucial. Thus, the generation of a synthetic dataset based on the proposed method and refined real dataset methods was proposed. Table 3 lists the elements that must be considered in synthetic and real-world dataset acquisition and their advantages and disadvantages.

Therefore, artificial manipulations must be added to generate synthetic images or collect real-world images. In this section, based on the acquired aspects of the training datasets, we review the existing reflection removal methods using synthetic and real-world datasets.

### Data-driven polarization reflection removal based on synthetic datasets

Synthetic datasets are commonly used to train networks because they are hard-accessible real-world datasets. The traditional method directly sums the candidate image as a reflection and transmission using normalized weights. Real-world tests exhibit poor performance. Therefore, based on the polarimetric imaging model, synthetic datasets of high quality and robustness were generated.

By leveraging the properties of light polarization and residual representation, Wieschollek et al. presented the first deep learning approach to separate reflected and transmitted components<sup>73</sup>, as shown in Fig. 19(a1). The proposed network architecture uses three polarization orientation images to calculate the parallel and perpendicular components  $I_{//}, I_{\perp}$  as the input into the network. The output of network comprises the residual images  $T', R'$  and the two single-channel weights  $\xi_{//}, \xi_{\perp}$ . The final estimates of the reflection and transmission can be computed as follows:

$$I_r = \xi_{\perp} R' + (1 - \xi_{\perp}) I_{\perp} \text{ and } I_t = \xi_{//} T' + (1 - \xi_{//}) I_{//} , \tag{30}$$

where  $I_r$  and  $I_t$  denote the estimation of reflected and transmitted light, respectively, represented as predicted transmission and reflection in Fig. 19(a1).

An accurate synthetic data generation pipeline is introduced, including the simulation of realistic reflections, such as high-dynamic-range scenes, nonstatic scenes, and curved and nonideal surfaces, to enhance the robustness of the proposed method, as shown in Fig. 19(a2). First, because the world consists of high-dynamic-range elements, the light intensity naturally diminishes as it

**Table 3 | The comparison of synthetic and real-world datasets.**

	Synthetic dataset	Real-world dataset
<ul style="list-style-type: none"> <li>• Elements should be considered in acquisition of datasets</li> </ul>	<ul style="list-style-type: none"> <li>• High dynamic range in the real-world.</li> <li>• Disappearance or interruption of the reflection.</li> <li>• Discontinuous scene of the reflection or transmission.                             <ul style="list-style-type: none"> <li>• Dynamic scenes.</li> <li>• Semi-reflective surface.</li> </ul> </li> <li>• Curved or irregular surface.</li> <li>• Rotating of reflection.</li> </ul>	<ul style="list-style-type: none"> <li>• Acquisition of ground truth.</li> <li>• Dynamic scenes.</li> <li>• Misalignment of refraction and transmission light.                             <ul style="list-style-type: none"> <li>• Luminousness and thickness of glass.</li> <li>• Image registration with or without glass.</li> <li>• Overexposure areas.</li> </ul> </li> </ul>
<ul style="list-style-type: none"> <li>• Advantages</li> </ul>	<ul style="list-style-type: none"> <li>• Abundant samples.</li> <li>• Easily generation.</li> </ul>	<ul style="list-style-type: none"> <li>• Real.</li> <li>• Complicated enough to cover the real-world conditions.</li> </ul>
<ul style="list-style-type: none"> <li>• Disadvantages</li> </ul>	<ul style="list-style-type: none"> <li>• Illusory.</li> <li>• Difficult to take fully account all conditions.</li> <li>• Domain gap between synthesis and real-world data.</li> </ul>	<ul style="list-style-type: none"> <li>• Difficult capturing process.</li> <li>• Limited samples.</li> </ul>

travels. Consequently, artificial adjustments match this phenomenon in real-world environments. The proposed method separately manipulates the dynamic range of the transmitted and reflected input images using a random factor. Second, for nonstatic scenes, such as cases where a swaying tree branch occurs during capture, local and nonrigid deformations are adopted by perturbing each grid over a patch. Third, for curved and non-ideal surfaces, a parabola was utilized to simulate unconstrained surface curvatures with four variable parameters.

Lyu et al. exploited the physical constraints from a pair of unpolarized and polarized images to separate reflection and transmission<sup>74</sup>, as shown in Fig. 19(b). The coefficients of the glass plane are predicted by the semireflector orientation module to compute the reflection and transmission based on the proposed physical image formation, denoted as:

$$\begin{aligned} \xi &= R_{\perp} + R_{\parallel} \\ \zeta &= R_{\perp} \cos^2(\phi_{\text{pol}} - \phi) + R_{\parallel} \sin^2(\phi_{\text{pol}} - \phi), \end{aligned} \quad (31)$$

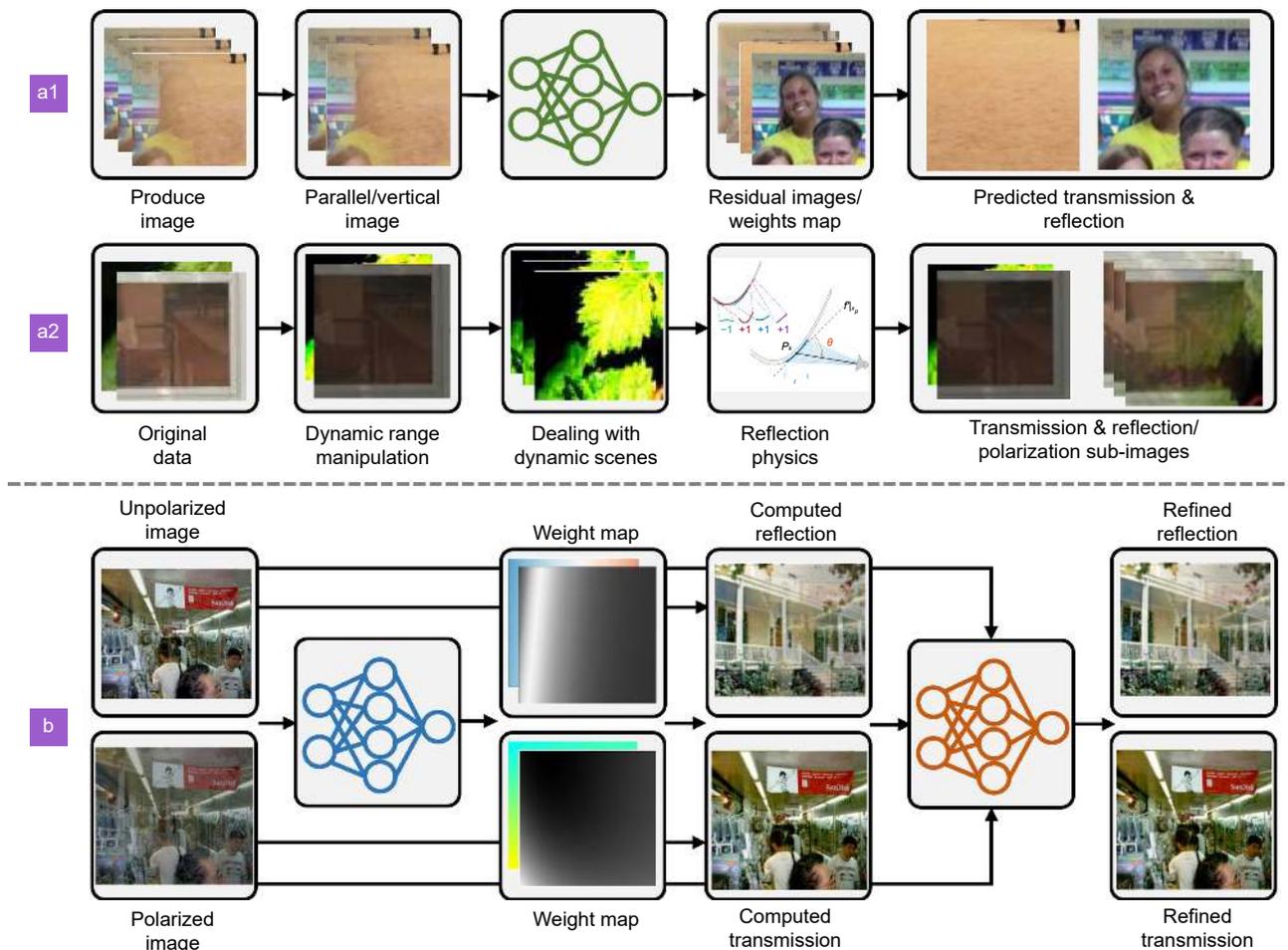
where  $\phi_{\text{pol}}$  denotes the polarization angle and  $\phi$  denotes the azimuth angle. The unpolarized and polarized images were then calculated using:

$$\begin{aligned} I_{\text{unpol}} &= \frac{\xi}{2} \cdot I_r + \frac{2 - \xi}{2} \cdot I_t \\ I_{\text{pol}} &= \frac{\zeta}{2} \cdot I_t + \frac{2 - \zeta}{2} \cdot I_r, \end{aligned} \quad (32)$$

where  $I_{\text{unpol}}$ ,  $I_{\text{pol}}$ ,  $\xi$  and  $\zeta$  denotes the unpolarized, polarized images and weights for reflection and transmission, respectively. Next, the reflection and transmission images can be computed as:

$$\begin{aligned} I_r &= 2 \cdot \frac{(2 - \xi) \cdot I_{\text{pol}} - (1 - \zeta) \cdot I_{\text{unpol}}}{2\zeta - \xi}, \\ I_t &= 2 \cdot \frac{\zeta \cdot I_{\text{unpol}} - \xi \cdot I_{\text{pol}}}{2\zeta - \xi}. \end{aligned} \quad (33)$$

Finally, to close the gap between the physical model and real data, a refined module was adopted to improve the initial estimation. Additionally, the proposed capturing setup can potentially be integrated into smartphones



**Fig. 19 | The typical data-driven polarimetric methods of reflection removal. (a1) P Wieschollek et al. method. (a2) The image-based data generation procedure of (a1). (b) Y Lyu et al. method.**

without affecting the original photography quality or achieving reliable results. The dataset was generated based on the PLACE2<sup>186</sup> dataset, in which two random images were selected as the original reflection and transmission images. Reflection is blurred by Gaussian smoothing based on the assumption that people take photos of the transmitted light.

Pang et al. proposed a novel progressive polarization-based reflection removal network (P<sup>2</sup>R<sup>2</sup>Net) to generate a preliminary estimation of coarse transmission images before guiding the final reflection removal<sup>187</sup>. The input images for the network consisted of two parts: a reflection-obstructed image and high-column features from a pre-trained VGG19, which is a successful example of using pre-trained features as prior information. Reflection-obstructed images are synthetic, based on physical function Eq. (33). The high dynamic range in the real-world, considered as light intensity, is nonlinearly compressed in the captured image through the power function of gamma encoding<sup>73</sup>. Two independent parameters were used to simulate diverse practical imaging environments. Additionally, flat and parabolic surface models were adopted to simulate curved surfaces, which can be calculated using the camera position ( $x_c, y_c$ ) and incidence point ( $x_p, y_p$ ):

$$\theta = \arccos \frac{x_c y_p - x_p y_c - x_p y_p}{\sqrt{x_p^2 + y_p^2} \sqrt{(x_p^2 - x_c^2) + (y_p^2 - y_c^2)}}, \quad (34)$$

where the  $\theta$  is the angle of incidence. Additionally, random deformation, rotation, and wrapping expand the scale of the dataset and improve the robustness when synthesizing the reflected obstructed image.

### Data-driven polarization reflection removal under real-world datasets

Despite synthetic datasets not being easy to obtain, they are often too ideal, and complex conditions cannot be fully considered in real-world environments. Real-world datasets are also crucial but challenging to obtain with the influence of glass and misalignment issues.

Real-world datasets are typically collected using removable glass. Reflection-obstructed images were captured using a camera with glass in front of the detector. The ground truth of the transmission was captured after the glass was removed. However, the difference between transmission and refractive transmission cannot be ignored. Intensity delay caused by attenuation and color distortion caused by colored glass are also common. To

eliminate the misalignment between these two elements, a loss of image similarity at the perceptual level, such as perceptual loss and contextual loss, was designed. However, the intensity delay and color distortion persist.

The collected reflection-free images are not perfectly aligned with the input mixed images owing to glass refraction. To avoid misalignment issues, Lei et al. used a piece of black cloth to cover the back of the glass to block all transmissions for clear reflection<sup>53</sup>. This dataset includes approximately 100 types of glass in the real world, which guides the proposed method to handle different types of reflections without introducing artifacts. The reflection removal network uses multi-polarization direction images as input. Furthermore, the calculated intensity, degree, and angle of polarization, and the overexposure mask eliminating the overexposed areas, were combined into the network. Two stages were adopted to estimate reflection and transmission. This design significantly improves the performance of the proposed method by a large margin.

In summary, reflection removal is crucial because obtaining analytical solutions to ill-posed problems is challenging. The introduction of polarization information can guide reflection removal using Fresnel functions despite the unknown incident angle. Thus, combining deep learning to learn prior parameters is a feasible method. Acquiring datasets determines the effectiveness and robustness of the parameter estimation of the proposed methods. In this study, synthetic and real-world methods are proposed. In the future, more comprehensive environments and complete theories must be developed to solve reflection-removal tasks effectively.

### Target detection

For target classification or detection, polarimetric data-driven methods can improve efficiency and do not require manual extraction of image features compared to traditional methods. However, existing methods use only intensity information images, resulting in a reduction in the accuracy rate for low-light environments or camouflaged targets<sup>188–193</sup>. The targets and backgrounds also differ in their polarization characteristics. Polarimetric imaging can effectively reveal these differences and assist in target detection<sup>194,195</sup>. Therefore, we can expect positive results by introducing polarization into data-driven target detection.

Fan et al. first proposed the use of polarization complementary to intensity-based information to improve

car detection accuracy<sup>54</sup>. A feature-selection process was performed to select the most informative polarization feature. Final detection is based on a fusion rule that takes the polarization-based model to confirm the color-based one. Gao et al. presented a similar work<sup>51</sup>. Blin et al. proved that polarimetric imaging is useful for target detection in road scenes<sup>52</sup>. Sun et al. adopted three-dimensional convolutions to consider the relationship among  $S_0$ ,  $S_1$ , and  $S_2$  images to improve the detection rate with limited polarimetric images<sup>59</sup>. Xie et al. used the Stokes vector to obtain four different configurations of polarization parameter image datasets:  $I$ , DoP,  $[I, \text{DoP}]$ , AoP, and  $[S_0, S_1, S_2]$  and trained different polarization image detection models, indicating that increased polarization information fusion enabled more learned target features and better target detection<sup>55</sup>. Tian et al. proposed a human face anti-spoofing method for real-life scenarios, which extracts and classifies the unique polarized features of faces using a CNN and an SVM together<sup>196</sup>. Experiments covering diverse face-spoofing attacks (print, replay, and mask) under uncontrolled indoor and outdoor conditions were conducted. Usmani et al. proposed unified polarimetric target detection and classification in degraded environments using 3D polarimetric integral imaging data<sup>197</sup>. 3D polarimetric images with deep neural networks can effectively detect and classify polarimetric targets under different low-light conditions and in the presence of occlusions. Shen et al. combined the advantages of polarimetric imaging and deep learning for rapid target detection of artificial targets camouflaged in natural scenes<sup>198</sup>, as shown in Fig. 20. The color difference of each image is calculated to prove the proposed method can highlight the camouflaged artificial targets to a greater extent.

### Biomedical imaging and pathological diagnosis

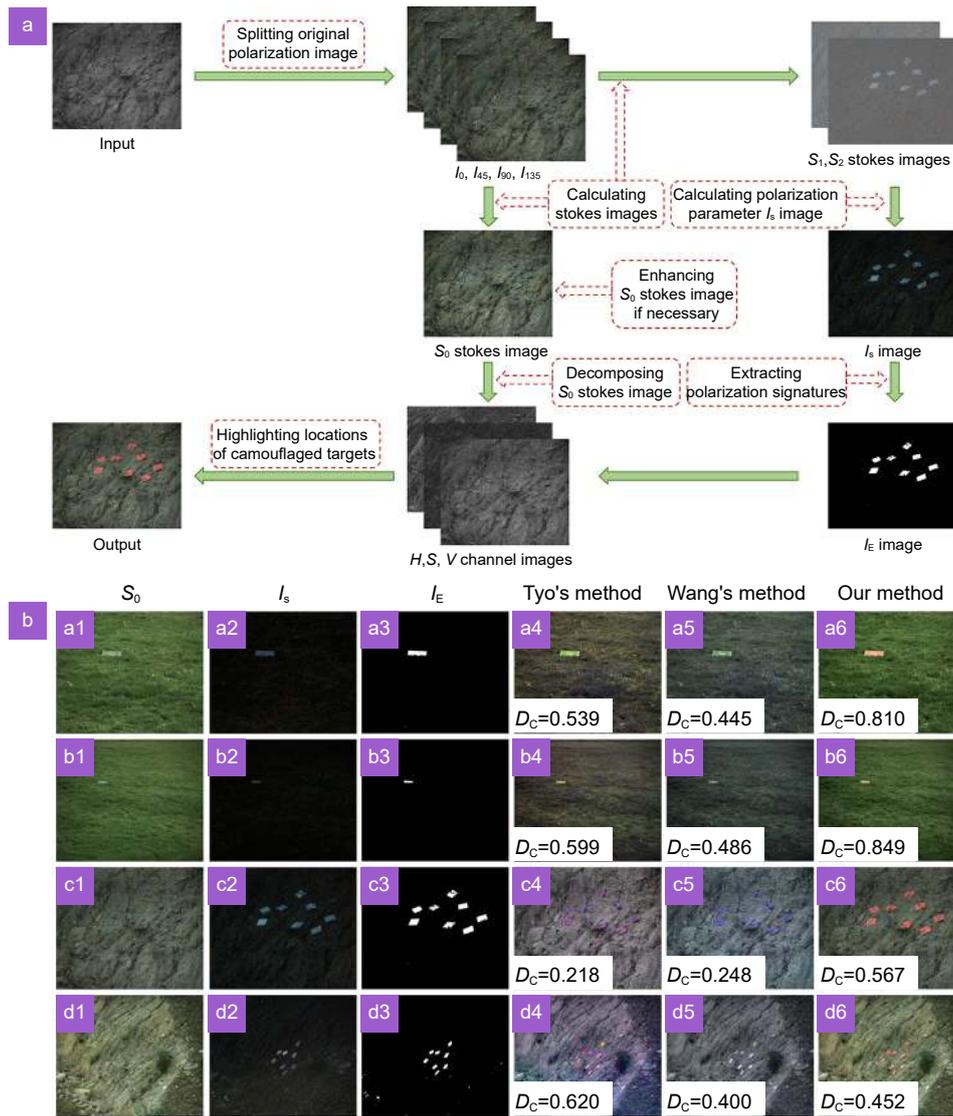
Biomedical imaging and pathological diagnosis methods based on Mueller matrix features, a typical polarization feature, are emerging label-free and noninvasive techniques suitable for characterizing the microstructures of biological tissues with anisotropic properties. Recently, results have been published based on Mueller matrix imaging for digital pathology<sup>25,28,29,199–207</sup>. However, achieving accurate pathological diagnosis by observing and evaluating stained pathological sections for interns is challenging. Furthermore, pathological diagnosis is a classification problem; therefore, learning-based methods are crucial in achieving fast and accurate digital

pathology. This section reviews the existing data-driven biomedical imaging and pathological diagnosis methods and applications. Next, we discuss the interpretation of physical properties of the network layers based on the distance-based learning classifier.

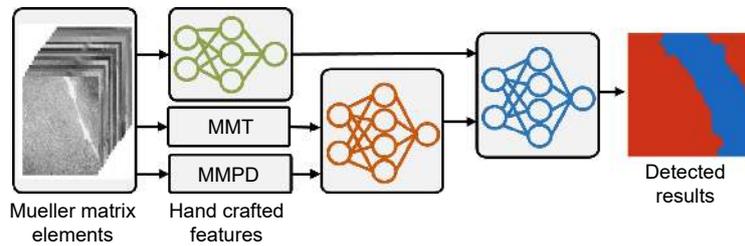
### Existing biomedical imaging methods

Li et al. first presented a Mueller matrix imaging system to classify morphologically similar algae using a CNN<sup>60</sup>. Because of the low contrast in the polarimetric signals of algae based on previous measurements of the algal Mueller matrix, performing classification without high-precision instruments is challenging. The proposed methods compare the performances of various stacks of network layers to identify the number of convolution layers. The classifier network was trained to extract features from the Mueller matrix and achieved a classification accuracy of 97%. Subsequently, they introduced a distance metric learning method called the Siamese network, which aimed to learn good distance metrics of algal Mueller matrix images in low-dimensional feature spaces<sup>61</sup>. Compared to the convolutional CNN method, in the Siamese approach, data pairs are generated stochastically as inputs to train the network to determine if they belong to the same category. The experiments demonstrated that the coupling of Mueller matrix imaging and CNN of the Siamese approach may be an efficient solution for the automatic classification of morphologically similar algae.

Zhao et al. proposed a giant cell tumor bone detection method using Mueller matrix polarization microscopic imaging and a multi-parameter fusion network (MPFN) that combines three extracted polarimetric features: deep micro-Pol features, MMPD features, and MMT features<sup>62</sup>, as shown in Fig. 21. Wang et al. and Zhou et al. used polarized speckle images for in vivo skin cancer detection<sup>85</sup> and polarized hyperspectral images for head and neck squamous cell carcinoma detection<sup>86</sup>. Yao et al. characterized the microstructures of endometrial samples at the typical proliferative and secretory phases using Müller matrix polar decomposition and a set of rotation-invariant parameters and their corresponding angular parameters<sup>87</sup>. In this study, polarimetric imaging was combined with a digital pathology technique to quantitatively study the microstructural features of endometrial samples. Furthermore, the incorporation of local image texture information through Local Binary Pattern (LBP) analysis improves the characterization



**Fig. 20 | Shen et al. method.** (a) The flow chart of proposed method. (b) The detected results. Figure reproduced with permission from ref.<sup>198</sup>, Institute of Electrical and Electronics Engineers.



**Fig. 21 | The architecture and results of multi-parameters fusion network (MPFN) and the corresponding results.**

ability of the polarization parameter images. The experiments demonstrated the feasibility of combining polarimetric imaging with digital pathological techniques in typical proliferative and secretory phases.

**Physical interpretation of network layers**

However, the physical properties of network layers re-

main unclear. In data-driven polarimetric imaging, the Muller matrix provides the most comprehensive information representing the polarization information, and most decomposition methods that provide raw fundamental parameters have been proposed. Thus, the Muller matrix is crucial in exploring the interpretation of network layers.

In ref.<sup>61</sup>, the authors calculated the Pearson correlation coefficients between the elements of the algal Mueller matrix and features extracted by the CNN from  $f_0$  to  $f_{15}$ . The experiments demonstrate that features  $f_2$ ,  $f_3$ ,  $f_6$ , and  $f_7$  are positively correlated with depolarization-related elements; however,  $f_1$ ,  $f_4$ ,  $f_{10}$ ,  $f_{12}$ , and  $f_{13}$  are negatively correlated. In addition, the fast-axis-orientation-dependent periodic variations were preserved in  $f_0$ ,  $f_5$ ,  $f_9$ , and  $f_{15}$ . Dong et al. proposed a data-driven polarimetric imaging framework and constructed a dual-modality machine-learning framework for the quantitative diagnosis of cervical precancerous lesions<sup>63</sup>, as shown in Fig. 22. The U-net architecture was adopted to segment the epithelium in digitized cervical hematoxylin-eosin-stained images and mask the corresponding cervical sample's polarimetry basis parameters (PBPs), which were decomposed based on the MMPD, MMT, and other Mueller matrix rotation-invariant parameters. Furthermore, these masked parameters are processed by the designed statistical distance-based learning classifier for deriving a polarimetry feature parameter (PFP). The classifier of the negative class and those of the CIN1 (mild dysplasia) samples can be expressed as:

$$D = \begin{bmatrix} \mathbf{x}_1^T \\ \mathbf{x}_2^T \\ \vdots \\ \mathbf{x}_N^T \end{bmatrix}_{N \times M}, \quad X = D\omega = \begin{bmatrix} \mathbf{x}_1^T \cdot \omega \\ \mathbf{x}_2^T \cdot \omega \\ \vdots \\ \mathbf{x}_N^T \cdot \omega \end{bmatrix}_{N \times 1},$$

$$L(\omega) = d(P_{\text{Normal}}(X), P_{\text{CIN1}}(X))^{-1}, \quad (35)$$

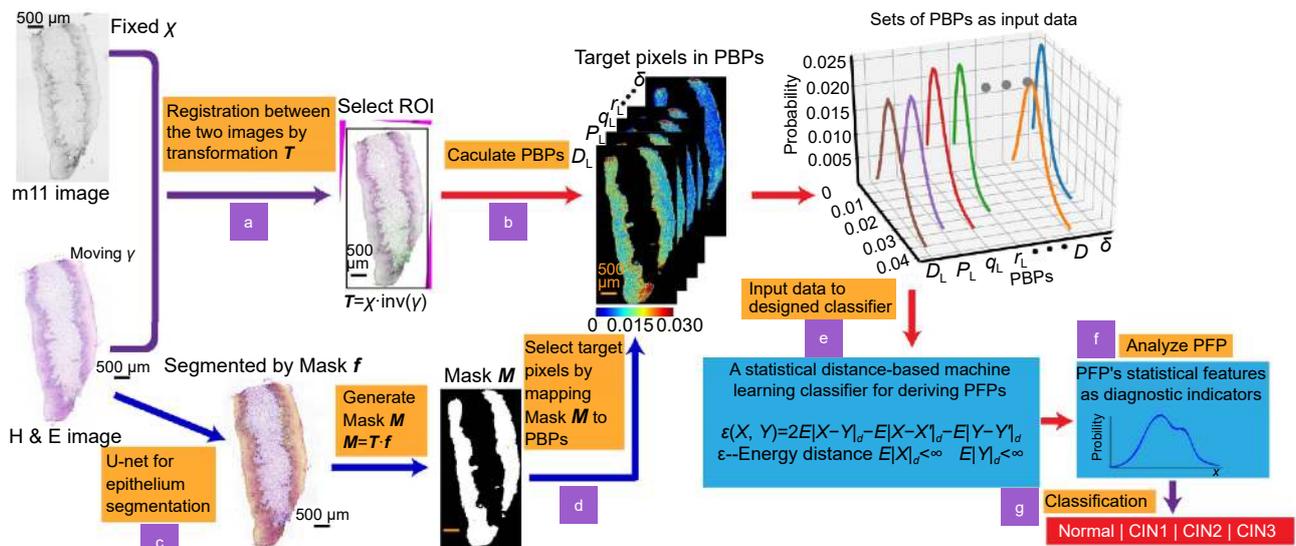
where  $\mathbf{x}_i$  is an  $M \times 1$  vector representing PBPs elements.

$M$  and  $N$  are the number of PBPs and target pixels, respectively.  $X$  is an  $N \times 1$  vector, calculated as a linear projection of the input PBPs.  $P_{\text{Normal}}(X)$  is the probability distribution of  $X$  from Normal cervical pathological tissues, whereas  $P_{\text{CIN1}}(X)$  represents CIN1 tissues.  $\omega$  represents weight coefficients of PBPs.  $L(\omega)$  is the energy distance between  $P_{\text{Normal}}(X)$  and  $P_{\text{CIN1}}(X)$  by energy distance function  $d$ . PFP can be represented a simplified linear combination of the PBPs, which is similar to the distribution of specific microstructural variations. The different weights indicate the significance of the elements of the PFP feature.

The results demonstrate the physical interpretability of the polarimetry feature parameters. For example, complex cervical precancerous samples exhibit polarization characteristics of various types of anisotropic superpositions. The depolarization ability of precancerous cervical samples changed with the development of lesions. In addition, changes in retardation and depolarization occur during the propagation and scattering of pathological cervical samples at different stages. Therefore, the proposed method has high sensitivity and precision for the screening of cervical lesion pathological tissues, and may bring physical interpretability to the CNN.

### Semantic segmentation

Segmentation is a popular topic for understanding the scene in remote sensing and automatic navigation fields. According to learning from different types of massive data, the data-driven segmentation method achieved



**Fig. 22 | Polarization-imaging-based ML framework for quantitative diagnosis of cervical precancerous lesions.** Figure reproduced with permission from ref.<sup>174</sup>, Institute of Electrical and Electronics Engineers.

good performance. However, intensity-based methods always suffer from degradation in scenes with similar colors, clutter, or reflective areas<sup>208–212</sup>. This choice is oriented towards polarimetric imaging, which provides the ability to distinguish and recover from changes in complex scenes. Several approaches have been proposed to achieve segmentation of remote sensing, road scenes, and transparent objects via polarimetric imaging and deep learning.

Shaunak et al. transformed the information in an augmented dataset into a compact representation of polarimetric synthetic aperture radar data to classify and segment urban areas<sup>82</sup>. The segmentation of road scenes is a typical application in which water hazards, transparent glass, and metallic surfaces are key challenges. Yang et al. proposed the prediction of polarization information from monocular RGB images as a complement to RGB-based pixel-wise semantic segmentation for applications in real-world wearable assistive navigation systems<sup>69</sup>, as shown in Fig. 23. Similarly, Zhang et al.<sup>56</sup> and Blanchon et al.<sup>83</sup> used different architectures to achieve the same goals: robust and accurate scene parsing of outdoor environments paves the way for autonomous navigation and relationship inference. Focusing on transparent object segmentation, the polarization textures of transparent objects provide extra but very different information than the background. Therefore, a polarized CNN frame-

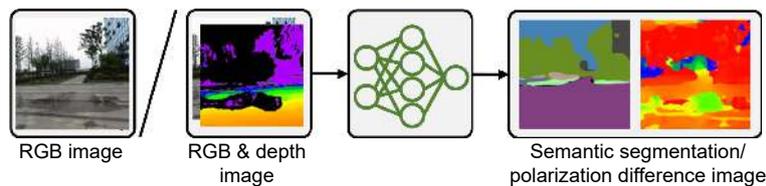
work can be trained based on the intensity and polarization information<sup>57</sup>, which will be helpful for applications in broad areas such as robotics, autonomous driving, and face authentication, as shown in Fig. 24. The mean average precision(mAP) is used to measure accuracy.

## Discussion

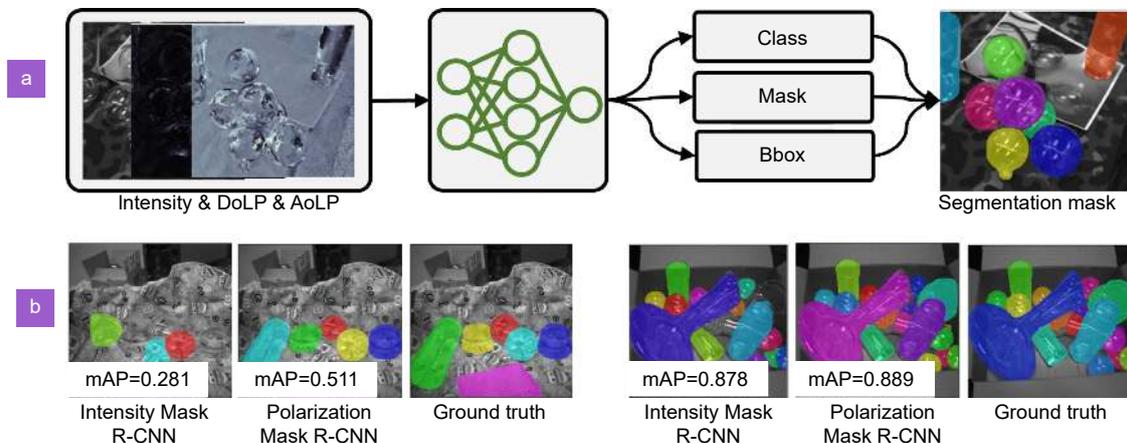
### Input and utilization of polarization information

Input information is a crucial element in network training with polarimetric imaging and deep learning. However, various inputs and polarization parameters exist for different tasks. Three perspectives were considered as network inputs: original polarization images (OPI), polarimetric parameter feature maps (PPFM), and associated parameter maps (APM), as shown in Fig. 25.

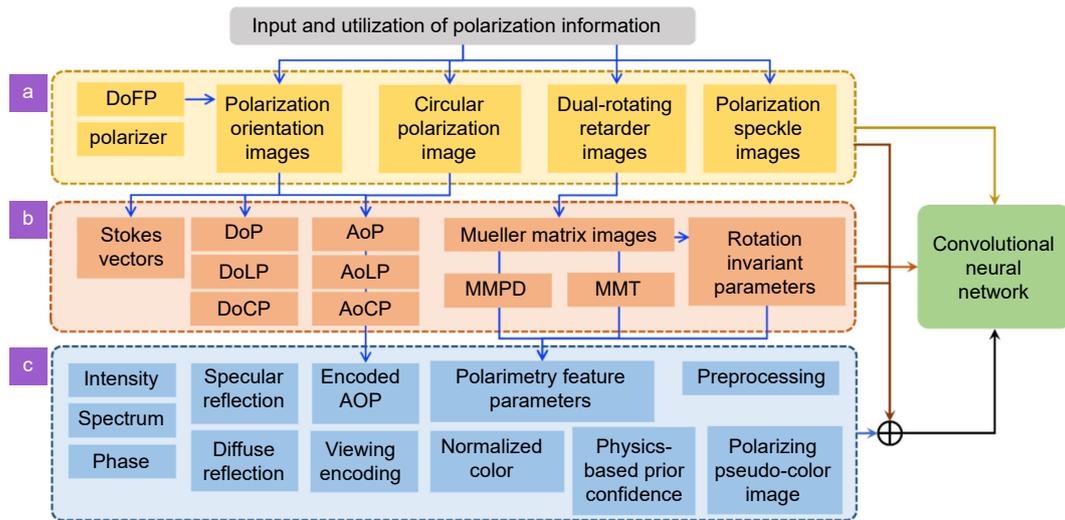
Original polarization images are among the most widely used inputs. OPI refers to images captured directly using a DoFP, camera, Mueller Matrix Polarization Microscope, or other equipment. Because of the rapid response and comprehensive polarization information captured in one shot, raw super-pixel images captured by DoFP are common input<sup>42–45</sup>. Other common inputs are polarization-oriented images, usually set to 0°, 45°, and 90°<sup>46–49,52</sup>. Furthermore, there are variants with the addition of 135°<sup>50,51</sup> or circular polarization information<sup>46</sup>. Parallel and perpendicular polarization components were used to train the dehazing network<sup>72</sup>. Polarization



**Fig. 23 | Yang et al. method:** proposed architecture and produced results of depth and segmented results.



**Fig. 24 | Transparent object segmentation.** (a) The designed architecture. (b) The segmented results of intensity Mask R-CNN and polarized Mask R-CNN in several dataset.



**Fig. 25 | Input and utilization of polarization information.** (a) Original polarization images (OPI). (b) Polarimetric parameter feature maps (PPFM). (c) Associated parameters maps (APM).

speckle images are another type of OPI<sup>66,85</sup> captured by the detector after scattering the media.

The polarimetric parameter feature maps were calculated using the OPI. Based on the Stokes representation,  $[S_0, S_1, S_2]$  is a common set widely used in polarization network training<sup>52,55,58,59</sup>. Furthermore, the DoP and AoP computed by the OPI or Stokes vector images are a type of new material for network training<sup>52–57</sup>. In biomedical diagnosis, Mueller matrix images are the most common input to the network<sup>91,95</sup>, and Mueller matrix parameters decomposed by Mueller matrix images can also be fed into network training<sup>62,208</sup>. Table 4 lists the existing decomposition elements of Mueller matrix methods.

Associated parameter maps are images in which OPI, and PPFM are combined with other information or pre-processing for different tasks. Intensity information is the most common complement based on the OPI and PPFM<sup>53,55–57,66,70,74,196</sup>. Similarly, the spectrum<sup>86</sup> and phase<sup>71,75</sup> are general additions to the polarization information. In 3D shape reconstruction tasks, there are different complements such as zenith and azimuth angle maps derived from specular and diffuse reflection<sup>64,65</sup>, viewing encoding, encoded AoP<sup>67</sup>, normalized color<sup>80</sup>, and physics-based prior confidence<sup>81</sup> based on different conditions. The raw images were interpolated using the bicubic interpolation method in the demosaicing task<sup>68</sup>. An overexposure mask is used in the network input to avoid overexposed areas during reflection removal<sup>53</sup>. The scene segmentation network utilizes HSL color space representation by incorporating a polarizing pseudo-color image<sup>83</sup>.

### Datasets

Owing to the demand for different tasks, the number of data-driven polarimetric imaging datasets has gradually increased, as listed in Table 5. Seven types of datasets were associated with the corresponding tasks described in previous section. Three strategies for building the datasets were considered. First, ground truths corresponding to the inputs exist; therefore, the real ground truth is captured directly to rectify the outputs. Second, the transfer function of the imaging system was considered, and the generative process was simulated to generate the ground truth. In addition, comparing different traditional methods and selecting the best results regarding the ground truth is another method that combines the advantages of existing methods.

The overall datasets have a large gap between each other in terms of number and size, even in the same task, as shown in Table 5, which would not avoid the difference in the extraction of features by the CNN. Furthermore, researchers use self-collected training and test datasets; however, evaluating and comparing different methods is challenging. Therefore, authoritative datasets must be built for this task.

### Loss function

Loss functions are critical elements, and their selection is crucial in guiding network training. Each loss function has advantages and disadvantages. Therefore, a specific loss function was adopted based on the given task and imaging environment. Table 6 lists the latest quality-loss functions for data-driven polarimetric imaging. The

**Table 4 | Existing decomposition of Mueller matrix.**

Decomposition of Mueller matrix	Physical meanings
Mueller matrix polar decomposition (MMPD) <sup>101</sup>	
$M = M_{\Delta} M_R M_D$	
$\delta = \arccos \left\{ \left[ \frac{(M_R(2,2) + M_R(3,3))^2 + (M_R(3,2) + M_R(2,3))^2}{2} - 1 \right] \right\}$	Linear retardation
$D = \sqrt{m_{12}^2 + m_{13}^2 + m_{14}^2} \quad (0 \leq D \leq 1)$	Diattenuation
$\Delta = 1 - \frac{ \text{tr}(M_{\Delta} - 1) }{3} \quad (0 \leq \Delta \leq 1)$	Depolarization
$\theta = 0.5 \arctan \left( \frac{r_2}{r_1} \right)$	Orientation of fast axis
$\psi = \frac{1}{2} \arctan \left[ \frac{M_R(3,2) - M_R(2,3)}{M_R(2,2) - M_R(3,3)} \right]$	Optical rotation
Mueller matrix transformation (MMT) <sup>102</sup>	
$t_1 = \frac{\sqrt{(m_{22} - m_{33})^2 + (m_{23} + m_{32})^2}}{2}$	Anisotropy degree
$t_2 = \frac{\sqrt{m_{21}^2 + m_{31}^2}}{2}$	
$b = \frac{m_{22} + m_{33}}{2}$	Polarizance
$A = \frac{2b \cdot t_1}{b^2 + t_1^2} \in [0, 1]$	Normalized anisotropy
$t_3 = \frac{\sqrt{m_{42}^2 + m_{43}^2}}{2}$	Linear retardation
Rotation invariant parameters <sup>213-215</sup>	
$P_L = \sqrt{m_{21}^2 + m_{31}^2} \in [0, 1]$	Linear polarizance
$P_C = m_{41} \in [-1, 1]$	Circular polarizance
$D_L = \sqrt{m_{12}^2 + m_{13}^2} \in [0, 1]$	Linear diattenuation
$D_C = m_{14} \in [-1, 1]$	Circular diattenuation
$q_L = \sqrt{m_{42}^2 + m_{43}^2} \in [0, 1]$	Capability of transforming between linear and circular polarizations
$r_L = \sqrt{m_{24}^2 + m_{34}^2} \in [0, 1]$	
$\alpha_P = \frac{1}{2} \arctan \left( \frac{m_{31}}{m_{21}} \right)$	Orientation of maximum transmittance
$\alpha_D = \frac{1}{2} \arctan \left( \frac{m_{13}}{m_{12}} \right)$	
$\alpha_q = \frac{1}{2} \arctan \left( -\frac{m_{42}}{m_{43}} \right)$	Orientation of fast axis
$\alpha_r = \frac{1}{2} \arctan \left( -\frac{m_{24}}{m_{34}} \right)$	
Mueller matrix anisotropy coefficients <sup>103</sup>	
$\alpha = \sqrt{\frac{1}{\Sigma} [(m_{01} + m_{10})^2 + (m_{23} - m_{32})^2]}$ $\Sigma = 3m_{00}^2 - (m_{11}^2 + m_{22}^2 + m_{33}^2) + m_{01}m_{10} + m_{02}m_{20} + m_{03}m_{30} - (m_{23}m_{32} + m_{13}m_{31} + m_{12}m_{21})$	Horizontal linear anisotropy
$\beta = \sqrt{\frac{1}{\Sigma} [(m_{02} + m_{20})^2 + (m_{13} - m_{31})^2]}$	45° linear anisotropy
$\gamma = \sqrt{\frac{1}{\Sigma} [(m_{03} + m_{30})^2 + (m_{12} - m_{21})^2]}$	Circular anisotropy

**Table 4 (Continued)**

Decomposition of Mueller matrix	Physical meanings
Differential decomposition <sup>104</sup>	
$\mathbf{m} = \sum_{n=1}^{16} \mathbf{m}_n = \begin{bmatrix} \kappa_i & \kappa_q + \kappa'_q & \kappa_u + \kappa'_u & \kappa_v + \kappa'_v \\ \kappa_q - \kappa'_q & \kappa_i - \kappa'_{i,q} & \eta_v + \eta'_v & \eta_u + \eta'_u \\ \kappa_u - \kappa'_u & -\eta_v + \eta'_v & \kappa_i - \kappa'_{i,u} & \eta_q + \eta'_q \\ \kappa_v - \kappa'_v & -\eta_u + \eta'_u & -\eta_q + \eta'_q & \kappa_i - \kappa'_{i,v} \end{bmatrix}$	Differential Mueller matrix for depolarizing anisotropic media
Symmetric decomposition <sup>105</sup>	
$\mathbf{M} = \mathbf{M}_{D2} \mathbf{M}_{R2} \mathbf{M}_{\Delta d} \mathbf{M}_{R1}^T \mathbf{M}_{D1}$	Symmetric decomposition of Mueller matrix
$\mathbf{M} = \mathbf{M}_{R2} \mathbf{M}'_{D2} \mathbf{M}_{\Delta d} \mathbf{M}'_{D1} \mathbf{M}_{R1}^T$	
$\mathbf{M} = \mathbf{M}_{D2} \mathbf{M}_{R2} \mathbf{M}_{\Delta d} \mathbf{M}'_{D1} \mathbf{M}_{R1}^T$	
$\mathbf{M} = \mathbf{M}_{R2} \mathbf{M}'_{D2} \mathbf{M}_{\Delta d} \mathbf{M}_{R1}^T \mathbf{M}_{D1}$	
Cloude decomposition <sup>106</sup>	
$m_{ij} = \frac{1}{2} \text{Tr} (T_c \eta_{4i+j})$	Cloude decomposition

following is a detailed description of several functions that differ from the intensity loss functions.

i) Mean squared error (MSE) is the most widely used indicator in deep learning, which standards the difference between the output images and the ground truth. In data-driven polarimetric imaging, in addition to the difference in intensity, the polarization parameters also ensure the accuracy of the polarimetric representations. The loss can be expressed as:

$$L_{MSE} = \|y - \hat{y}\|_2, \tag{36}$$

where  $y$  is the ground truth, and  $\hat{y}$  is the output of the network. Therefore, the measured information consists of polarimetric representations, such as DoP(DoLP), AoP(AoLP), Stokes vectors, Mueller matrix, spectrum, HSV color space, and their perceptual representations computed by a network model VGG.

ii) The mean absolute error (MAE) is a widely used indicator. Compared to the MSE, the MAE has less blurriness and noise; however, it is more unstable. The utilization of the MAE loss function was similar to that of MSE because the range of AoLP is 0 to  $\pi$  that always maps into 0–1<sup>77</sup>. However, 0 and 1 indicate the same physical meaning where the error is the largest. Therefore, the HSV spatial display rule was introduced to design a closer distance on the circle of the AoLP. The loss function is defined as follows:

$$L_{MAE,AoLP} = \min (\|y - \hat{y}\|_1, 1 - \|y - \hat{y}\|_1). \tag{37}$$

iii) Cosine similarity (CS) is commonly used in the surface-normal estimation of 3D reconstructions based on polarimetric imaging. In the 3D polarimetric imaging task, the surface normal map is calculated using Fresnel's formula to generate the normal vectors, which

is different from the other information forms, ensuring that the CS loss becomes the most suitable indicator. The loss function is defined as follows:

$$L_{CS} = 1 - \frac{y \cdot \hat{y}}{\|y\|_2 \|\hat{y}\|_2}. \tag{38}$$

iv) SSIM is a widely used indicator of end-to-end networks. SSIM focuses on the brightness, contrast, and structural similarity between two images. The SSIM improves as the value increases within the range of [0, 1], opposite to the goal of minimizing the similarity. The loss function is defined as follows:

$$L_{SSIM} = 1 - \frac{2\mu_y \mu_{\hat{y}} + c_1}{\mu_y^2 + \mu_{\hat{y}}^2 + c_1} \cdot \frac{2\sigma_{y\hat{y}} + c_2}{\sigma_y^2 + \sigma_{\hat{y}}^2 + c_2}, \tag{39}$$

where  $\mu_*$  and  $\sigma_*$  denote the mean and standard deviations of the image, respectively.  $\sigma_{y\hat{y}}$  is the cross-covariance computed from the images of  $y$  and  $\hat{y}$ . where  $c_1$  and  $c_2$  are constants.

### Future of data-driven polarimetric imaging

The field of polarimetric imaging has been influenced by deep learning, which has recently become one of the most disruptive technologies. First, we analyze the trends in data-driven polarimetric imaging, focusing on the application of data-driven polarimetric imaging and review existing research achievements. Furthermore, the acquisition of high-accuracy polarization information is the foundation for subsequent imaging and semantic processing. Descattering, 3D shape reconstruction, reflection removal, biomedical imaging, pathological diagnosis, target detection, and semantic segmentation are crucial in the application of data-driven polarimetric imaging. A comprehensive discussion is essential owing

**Table 5 | Summary of data-driven polarimetric imaging datasets.**

Task	Data	Dataset	Paired or not	No. of images	Size	Synthetic/Real
Section Restoration and enhancement of accurate polarization information	2020	Hu <sup>42</sup>	Paired	300000	64×64	Real
	2022	LLCP <sup>50</sup>	Paired	—	64×64	Real
	2018	Zhang <sup>68</sup>	Paired	215	640×480	Synthetic
	2019	Zeng <sup>43</sup>	Paired	76890	40×40	Synthetic
	2021	Wu <sup>44</sup>	Paired	—	64×64	Synthetic
	2020	Li <sup>45</sup>	Paired	137000	64×64	Synthetic
	2022	Yang <sup>78</sup>	Paired	337	1961×2381	Real
	2022	Liu <sup>77</sup>	Paired	140000	64×64	Real
	2021	Ting <sup>79</sup>	Paired	42228	512×512	Real
Section Polarimetric descattering	2022	Si <sup>58</sup>	Paired	—	256×256	Real
	2020	Hu <sup>46</sup>	Paired	103000	64×64	Real
	2021	Zhang <sup>70</sup>	Paired	2000	256×256	Real
	2021	Ren <sup>148</sup>	Paired	—	—	Real
	2021	Zhou <sup>48</sup>	Paired	—	240×240	Synthetic
	2022	Shi <sup>72</sup>	Paired	60000	960×576	Synthetic
	2022	Ding <sup>47</sup>	Paired	700	512×512	Real+Synthetic
Section Three-dimensional shape reconstruction	2022	Li <sup>66</sup>	Paired	5000	256×256	Real+Synthetic
	2019	Ba <sup>64</sup>	Paired	300	256×256	Real
	2020	SURREAL+PHSPD <sup>65</sup>	Paired	312915	—	Real
	2020	Kondo <sup>71</sup>	Paired	—	—	Real+Synthetic
	2021	Lei <sup>67</sup>	Paired	522	1224×1024	Real
	2021	Deschaintre <sup>80</sup>	Paired	100000	512×512	Synthetic
Section Reflection removal	2022	TransSfP <sup>81</sup>	Paired	936	512×512	Real+Synthetic
	2018	URD <sup>73</sup>	Paired	—	128×128	Synthetic
	2019	Lyu <sup>74</sup>	Paired	—	256×256	Synthetic
Section Target detection	2020	Lei <sup>53</sup>	Paired	100107	—	Real
	2018	Fan <sup>54</sup>	Paired	153	—	Real
	2019	Blin <sup>52</sup>	Paired	2730	—	Real
	2020	Sun <sup>59</sup>	Paired	—	96×96	Real
	2020	CASIA-DOLP <sup>196</sup>	Unpaired	10697	224×224	Real
	2021	USMANI <sup>197</sup>	Paired	240	—	Real
	2022	Gao <sup>51</sup>	Paired	60	1224×1024	Real
Section Biomedical imaging and pathological diagnosis	2017	Li <sup>60</sup>	Paired	10463	—	Real
	2018	Li <sup>61</sup>	Paired	12162	—	Real
	2020	Zhao <sup>62</sup>	Paired	100000	—	Real
	2021	Zhou <sup>66</sup>	Paired	4500	40×40	Real
	2021	Yao <sup>87</sup>	Paired	400	64×64	Real
	2021	Dong <sup>63</sup>	Unpaired	49870	128×128	Real
Section Semantic segmentation	2018	Yang <sup>69</sup>	Paired	9736	320×240	Real
	2019	POLABOT <sup>56</sup>	Paired	700	256×256	Real
	2020	Kalra <sup>57</sup>	Paired	1000	—	Real

to the input and utilization of polarization information, datasets, and loss functions.

This section evaluates the future of data-driven polarimetric imaging based on its strengths, weaknesses, and opportunities. Furthermore, this approach is suitable for developing instructive strategies for further studies that combine deep learning and polarization information.

### Strengths

Data-driven polarimetric imaging represents a new paradigm that combines deep learning and traditional physical properties, including altering the patterns of physical properties to achieve better results and balancing the physical model and information extraction between the traditional physical and high-order nonlinear

**Table 6 | Summary of loss functions.**

Loss function	Applications
Frobenius matrix norm <sup>45,46,68</sup>	Section <i>Restoration and enhancement of accurate polarization information</i> , Section <i>Polarimetric descattering</i>
Weighted mean squared error (WMSE) <sup>70</sup>	Section <i>Polarimetric descattering</i>
Preceptual loss <sup>53,70,77</sup>	Section <i>Polarimetric descattering</i>
Mean squared error (MSE) <sup>46,48,49,61,65,71-74,76-78</sup>	Section <i>Restoration and enhancement of accurate polarization information</i> , Section <i>Polarimetric descattering</i> , Section <i>Three-dimensional shape reconstruction</i> , Section <i>Reflection removal</i>
Mean absolute Error(MAE) <sup>42-44,47,48,50,66,75,79,80,148</sup>	Section <i>Restoration and enhancement of accurate polarization information</i> , Section <i>Polarimetric descattering</i> , Section <i>Three-dimensional shape reconstruction</i>
Negative pearson correlation coefficient (NPCC) <sup>66</sup>	Section <i>Polarimetric descattering</i>
Cosine similarity (CS) <sup>64,65,67,81</sup>	Section <i>Three-dimensional shape reconstruction</i>
Perceptual normalized cross-correlation (PNCC) <sup>53</sup>	Section <i>Reflection removal</i>
Structural similarity index (SSIM) <sup>43,79</sup>	Section <i>Restoration and enhancement of accurate polarization information</i>
Total variation (TV) <sup>58,75</sup>	Section <i>Restoration and enhancement of accurate polarization information</i>
Contrastive loss <sup>196</sup>	Section <i>Target detection</i>
Logistic regression <sup>87</sup>	Section <i>Biomedical imaging and pathological diagnosis</i>
Ridge loss <sup>87</sup>	Section <i>Biomedical imaging and pathological diagnosis</i>
Focal loss <sup>69</sup>	Section <i>Semantic segmentation</i>
Cross entropy loss (CEL) <sup>83</sup>	Section <i>Semantic segmentation</i>

representation domains. Applying the information represented by the physical model and the deep network layers enables researchers to exploit the potential features embedded in all information-transmitted paths.

During the imaging process, the light source, transmitted media, imaging system, and image processing method influence visual performance. Moreover, modeling these complex processes using physical functions or traditional methods is challenging. Network layers are high-order polynomials expressed by the convolution neural network. Therefore, nonlinear representations of the network layers may be promising for simulating the process.

The introduction of deep learning into conventional polarimetric imaging and pattern recognition generates more accurate coefficients for the physical model, simulating a complicated processing approach not modeled by physical functions. However, introducing a polarimetric imaging model into the deep network would add physical constraints and polarization information to guide the network training to achieve better performance compared with the intensity network. Therefore, data-driven polarimetric imaging probably enables capabilities that cannot be realized using traditional methods.

### Weaknesses

To achieve a better performance by training a polarimet-

ric imaging network, researchers must weigh the costs associated with data-driven polarimetric imaging and conventional approaches, which include the establishment of polarimetric imaging datasets, storage of data always four times that of traditional datasets, and imaging systems. Because balancing these costs and benefits is not exact, some uncertainties are considered in this process.

A fundamental element of data-driven polarimetric imaging is the availability of comprehensive datasets. Most data-driven polarimetric imaging methods have focused on supervised deep learning. However, based on polarimetric imaging methods and imaging environments, the corresponding ground truth is more challenging to capture than in an intensity-based network. Additionally, different polarimetric imaging techniques introduce various errors and influence the visual performance of the network. For example, the division of time always suffers from mismatching in dynamic scenes, and the division of the focal plane has a mosaicking problem in principle. Existing methods have established their datasets to respond to specific tasks. However, ensuring similar performance from other datasets is challenging; moreover, the existing dataset is insufficient to cover all conditions, and its generalization ability is insufficient. Real data changes over time, indicating increasing volume and improper handling of methods.

The loss function, which is crucial in guiding the

training of the network, is always similar to the intensity-based methods used in data-driven polarimetric imaging. General operators replace intensity images with polarization parameters. However, the intensity and polarization information have entirely different optical properties. The intensity image describes the reflectivity and transmissivity of the object; however, the polarization image describes the texture details, material properties, shape, shading, and roughness. These differences determine the disparate designs of loss functions. However, there are insignificant loss functions that focus on polarization parameters.

Black boxes and their acceptance by applied researchers are inherent drawbacks of deep learning methods by health professionals. Most researchers in practical applied fields are wary because deep learning theories have not yet provided a complete and reasonable answer. In addition, further development and optimization would depend only on the performance of tasks without any guidance from theories, which results in indeterminacy in the study. Moreover, the legal implications of black box functionality could be another challenge. For example, who would be responsible if the results were incorrect in pathological diagnosis or target detection? In data-driven polarimetric imaging, the introduction of polarization information may be helpful for the interpretability of deep learning. Dong et al. attempted to use a linear projection of input PBPs to interpret their significance by learning the factors of each parameter. However, many studies have been conducted to achieve these goals<sup>63</sup>.

The combination and utilization of polarization information in deep learning are still in their infancy. In most existing methods, polarization parameter images are the only approach that uses polarization information as the input into the network. The extraction of polarization features relies on the automatic processing of network layers, which remains a challenge to utilize polarization information despite the unlimited opportunity for improvement yet to be explored.

### Opportunities

Based on the weaknesses of data-driven polarimetric imaging, solutions have been proposed to address these gaps. Furthermore, many novel training methods and physical models exist, such as unsupervised or semi-supervised training, transfer learning, and computational imaging. Therefore, it must be combined with other ima-

ging or training theories to guide the optimization process. In addition, three broad application areas: descattering imaging, even high-scattering media, detection of camouflage, spoofing targets, and enhancement and fusion of information assess the potential of data-driven polarimetric imaging in future applications.

### The opportunities of methods

*The assistance of physical model:* A synthetic dataset solves the scarcity of datasets. An accurate physical model that simulates information transmission is crucial in dataset generation. With an additional physical constraint on the CNN, fewer training data are required to achieve a more generalized result than conventional methods. In addition, to obtain a synthetic dataset, various parameters covering different imaging conditions are crucial. Thus, the development of conventional polarimetric imaging methods is suitable for designing network architectures.

*Unsupervised or semi-supervised learning*<sup>216–218</sup>: Obtaining the ground truth for a large polarization dataset is challenging. Therefore, unsupervised or semi-supervised learning is required to reduce the dependence on the ground truth. However, image enhancement or image processing is an end-to-end task; thus, existing learning methods without ground truth achieve poor performance. A more comprehensive physical model must be established, and more effective loss functions designed to guide pipelines. In addition, the middle parameter may be generated without the ground truth, which is also a feasible way to improve the performance.

*Transfer learning*<sup>219,220</sup>: Transfer learning allows the optimized parameters for one dataset to train a new network as initialization values for another dataset, which is a feasible approach for reducing the dependence on datasets in data-driven polarimetric imaging because the learned features can be promptly transferred from a trained network to a new network for another task. The fine-tuning technique is a typical method used in transfer learning, which is faster and easier than training a network from scratch. Therefore, the extracted features were similar to the shallow layers; moreover, the shallow layers in a trained network can be copied to the new network for another task to reduce the cost of training time.

*Multi-dimensional learning*<sup>221,222</sup>: The introduction of polarization information that displays different physical properties into a traditional intensity network can provide more constraints and information sources to

promote network inference. Similarly, the phase, spectrum, and other physical properties can be embedded into a network to enhance performance. Phase is a representative generation of the change in light, and the spectrum describes the characteristics of the wavelength of light. These properties would fill the gap in which a single domain cannot fully represent the light domain for all physical properties.

*Federated learning*<sup>223</sup>: To provide a practical training set for deep learning in polarimetric imaging applications, obtaining various available datasets from different institutes or corporations may be a possible solution. There are several datasets for different tasks, as shown in Table 5; however, the datasets collected by different groups are not uniform because it is difficult to guarantee a similar performance from other datasets. Therefore, different datasets were beneficial, increasing the diversity of the collected samples. In addition, different imaging systems, detectors, environments, and observation directions are challenging to simulate using existing physical functions, which can improve the generalization ability of the network and avoid overfitting. Therefore, this is instructive for building and optimizing the evaluation criterion, focusing on polarization images.

*Emergence of metasurface and metalens*<sup>224–227</sup>: The utilization of lenses and metasurfaces allows for tailored control over light with specific polarization states, achieved through deliberate design. This deliberate control enables superior capture, separation, and analysis of polarized light signals, thereby significantly enhancing the sensitivity and accuracy of acquiring polarization datasets. These advancements not only amplify the potential of data-driven polarimetric imaging but also complement the capabilities of deep learning methodologies, promising refined insights and higher precision in polarization imaging applications.

### The opportunities of applications

In the future, the optimization of methods will aim for better visual performance in more widely applied fields. For polarimetric imaging, the methods that depend on polarization properties include descattering imaging, high-scattering media, detection of camouflage, spoofing targets, and enhancement and fusion of information. Deep learning, nonlinear representation ability, and potential feature extraction improve the accuracy of estimating parameters and feasible transmitted functions compared to conventional methods.

For descattering imaging in high-scattering media, such as clouds, water, haze, smoke, smog, fog, mist in the air, soil particles, algae, and mineral salt, in underwater scenes, there are potential opportunities for data-driven polarimetric imaging. However, further development of physical models has amplified this ability. However, traditional model functions are challenging to handle in complex imaging environments and always use simple assumptions to simulate real parameters. The introduction of deep learning can model complicated conditions in nonlinear cases using convolutional neural layers. Future opportunities will arise from the development of more accurate parameters for forming improved imaging functions generated by deep learning.

Spoofing targets are another opportunity for camouflage detection. Target detection is widely applied in polarimetric imaging because the polarization information can describe the material of an object, which is suitable for camouflaging and spoofing targets of the same color that the intensity information cannot distinguish. Next, more comprehensive extraction of special features by the neural network may further improve the success rate of target detection.

The material surface, texture, and contrast are the main characteristics described by polarization information for enhancing and fusing information. Polarization parameters are observable in low- or hard-light environments because they are unaffected by intensity. Consequently, the fusion of polarization and other images can extend the feature domain of an object. The network generates a fusion based on data-driven polarization fusion, enhancing performance by extracting more features and providing more information on imaging objects or scenes compared with artificial coefficients. Moreover, complementary features from various domains are advantageous for other computer vision tasks, such as object detection.

### Conclusion

This review provides an overview of recent efforts to summarize data-driven polarimetric imaging based on seven classifications and discusses them comprehensively from three perspectives. Based on the application fields, the classifications consist of polarimetric descattering, 3D shape reconstruction, reflection removal, restoration, enhancement of polarization information, target detection, biomedical imaging and pathological diagnosis, and semantic segmentation. Subsequently, we

synthetically analyze the input, datasets, and loss function, which are crucial in data-driven polarimetric imaging, listing the existing datasets and loss functions with an evaluation of their advantages and disadvantages. In conclusion, deep-learning-based polarimetric imaging introduces polarization information into the convolutional neural network to achieve better performance than traditional intensity imaging, bringing physical interpretability to CNN through physical models. Through research on existing data-driven polarimetric imaging, the study of the corresponding fields can be improved to a higher level, enabling them to enhance high-level visual tasks.

## References

- Ronchi V, Barocas V. *The Nature of Light: An Historical Survey* (Harvard University Press, Cambridge, 1970).
- Huard S. *Polarization of Light* (Wiley, Chichester, 1997).
- Schechner YY, Karpel N. Recovery of underwater visibility and structure by polarization analysis. *IEEE J Oceanic Eng* **30**, 570–587 (2005).
- Li XB, Hu HF, Zhao L et al. Polarimetric image recovery method combining histogram stretching for underwater imaging. *Sci Rep* **8**, 12430 (2018).
- Liu TG, Guan ZJ, Li XB et al. Polarimetric underwater image recovery for color image with crosstalk compensation. *Opt Laser Eng* **124**, 105833 (2020).
- Liang J, Ren LY, Qu ES et al. Method for enhancing visibility of hazy images based on polarimetric imaging. *Photonics Res* **2**, 38–44 (2014).
- Liu F, Han PL, Wei Y et al. Deeply seeing through highly turbid water by active polarization imaging. *Opt Lett* **43**, 4903–4906 (2018).
- Liu F, Wei Y, Han PL et al. Polarization-based exploration for clear underwater vision in natural illumination. *Opt Express* **27**, 3629–3641 (2019).
- Wei Y, Han PL, Liu F et al. Enhancement of underwater vision by fully exploiting the polarization information from the Stokes vector. *Opt Express* **29**, 22275–22287 (2021).
- Li X, Liu F, Han PL et al. Near-infrared monocular 3D computational polarization imaging of surfaces exhibiting nonuniform reflectance. *Opt Express* **29**, 15616–15630 (2021).
- Han PL, Cai YD, Liu F et al. Computational polarization 3D: new solution for monocular shape recovery in natural conditions. *Opt Laser Eng* **151**, 106925 (2022).
- Cui ZP, Gu JW, Shi BX et al. Polarimetric multi-view stereo. In *Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition* (IEEE, 2017); <http://doi.org/10.1109/CVPR.2017.47>.
- Kadambi A, Taamazyan V, Shi BX et al. Depth sensing using geometrically constrained polarization normals. *Int J Comput Vis* **125**, 34–51 (2017).
- Kong N, Tai YW, Shin JS. A physically-based approach to reflection separation: from physical modeling to constrained optimization. *IEEE Trans Pattern Anal Mach Intell* **36**, 209–221 (2014).
- Bronstein AM, Bronstein MM, Zibulevsky M et al. Sparse ICA for blind separation of transmitted and reflected images. *Int J Imag Syst Technol* **15**, 84–91 (2005).
- Forsell G. Test and analysis of the detectability of personnel mines in a realistic minefield by polarization in the infrared LW region. *Proc SPIE* **5415**, 187–195 (2004).
- Forsell G. Passive IR polarization measurements applied to covered surface landmines. *Proc SPIE* **5089**, 547–557 (2003).
- Cremer F, de Jong W, Schutte K. Infrared polarization measurements and modelling applied to surface laid anti-personnel landmines. *Opt Eng* **41**, 1021–1032 (2002).
- Aron Y, Gronau Y. Polarization in the LWIR: a method to improve target acquisition. *Proc SPIE* **5783**, 653–661 (2005).
- Ratliff BM, Lemaster DA, Mack RT et al. Detection and tracking of RC model aircraft in LWIR microgrid polarimeter data. *Proc SPIE* **8160**, 816002 (2011).
- Voss KJ, Fry ES. Measurement of the Mueller matrix for ocean water. *Appl Opt* **23**, 4427–4439 (1984).
- Fry ES, Voss KJ. Measurement of the Mueller matrix for phytoplankton. *Limnol Oceanogr* **30**, 1322–1326 (1985).
- Svensen Ø, Stamnes JJ, Kildemo M et al. Mueller matrix measurements of algae with different shape and size distributions. *Appl Opt* **50**, 5149–5157 (2011).
- Wang WF, Lim LG, Srivastava S et al. Investigation on the potential of Mueller matrix imaging for digital staining. *J Biophotonics* **9**, 364–375 (2016).
- Du E, He HH, Zeng N et al. Mueller matrix polarimetry for differentiating characteristic features of cancerous tissues. *J Biomed Opt* **19**, 076013 (2014).
- Le DL, Huynh TN, Nguyen DT et al. Characterization of healthy and nonmelanoma-induced mouse utilizing the Stokes-Mueller decomposition. *J Biomed Opt* **23**, 125003 (2018).
- Pierangelo A, Manhas S, Benali A et al. Multispectral Mueller polarimetric imaging detecting residual cancer and cancer regression after neoadjuvant treatment for colorectal carcinomas. *J Biomed Opt* **18**, 046014 (2013).
- Shukla P, Pradhan A. Mueller decomposition images for cervical tissue: potential for discriminating normal and dysplastic states. *Opt Express* **17**, 1600–1609 (2009).
- Pierangelo A, Nazac A, Benali A et al. Polarimetric imaging of uterine cervix: a case study. *Opt Express* **21**, 14120–14130 (2013).
- Egan WG. Polarization in remote sensing. *Proc SPIE* **0891** (1988). <https://doi.org/10.1117/12.944289>
- David G, Thomas B, Dupart Y et al. UV polarization lidar for remote sensing new particles formation in the atmosphere. *Opt Express* **22**, A1009–A1022 (2014).
- Carotenuto V, de Maio A, Clemente C et al. Invariant rules for multipolarization SAR change detection. *IEEE Trans Geosci Remote Sens* **53**, 3294–3311 (2015).
- Nagdimunov L, Kolokolova L, Mackowski D. Characterization and remote sensing of biological particles using circular polarization. *J Quant Spectrosc Radiat Transfer* **131**, 59–65 (2013).
- Wang F, Ainouz S, Lian CF et al. Multimodality semantic segmentation based on polarization and color images. *Neurocomputing* **253**, 193–200 (2017).
- Xiang KT, Yang KL, Wang KW. Polarization-driven semantic segmentation via efficient attention-bridged fusion. *Opt Express* **29**, 4802–4820 (2021).
- Ni J, Zhang F, Ma F et al. Random region matting for the high-

- resolution PolSAR image semantic segmentation. *IEEE J Sel Top Appl Earth Obs Remote Sens* **14**, 3040–3051 (2021).
37. Li ZS, Sun JS, Fan Y et al. Deep learning assisted variational Hilbert quantitative phase imaging. *Opto-Electron Sci* **2**, 220023 (2023).
38. Wang YYD, Wang H, Gu M. High performance “non-local” generic face reconstruction model using the lightweight Speckle-Transformer (SpT) UNet. *Opto-Electron Adv* **6**, 220049 (2023).
39. Guo YM, Zhong LB, Min L et al. Adaptive optics based on machine learning: a review. *Opto-Electron Adv* **5**, 200082 (2022).
40. Li YX, Qian JM, Feng SJ et al. Deep-learning-enabled dual-frequency composite fringe projection profilometry for single-shot absolute 3D shape measurement. *Opto-Electron Adv* **5**, 210021 (2022).
41. Chen YX, Zhang FY, Dang ZB et al. Chiral detection of biomolecules based on reinforcement learning. *Opto-Electron Sci* **2**, 220019 (2023).
42. Hu HF, Lin Y, Li XB et al. IPLNet: a neural network for intensity-polarization imaging in low light. *Opt Lett* **45**, 6162–6165 (2020).
43. Zeng XL, Luo Y, Zhao XJ et al. An end-to-end fully-convolutional neural network for division of focal plane sensors to reconstruct  $S_0$ , DoLP, and AoP. *Opt Express* **27**, 8566–8577 (2019).
44. Wu RY, Zhao YQ, Li N et al. Real-time division-of-focal-plane polarization imaging system with progressive networks. arXiv: 2110.13823 (2021). <https://doi.org/10.48550/arXiv.2110.13823>
45. Li XB, Li HY, Lin Y et al. Learning-based denoising for polarimetric images. *Opt Express* **28**, 16309–16321 (2020).
46. Hu HF, Zhang YB, Li XB et al. Polarimetric underwater image recovery via deep learning. *Opt Laser Eng* **133**, 106152 (2020).
47. Ding XY, Wang YF, Fu XP. Multi-polarization fusion generative adversarial networks for clear underwater imaging. *Opt Laser Eng* **152**, 106971 (2022).
48. Zhou C, Teng MG, Han YF et al. Learning to dehaze with polarization. In *Proceedings of the 35th Conference on Neural Information Processing Systems* (NeurIPS, 2021).
49. Zhu YM, Zeng TJ, Liu KW et al. Full scene underwater imaging with polarization and an untrained network. *Opt Express* **29**, 41865–41881 (2021).
50. Xu XY, Wan MG, Ge JY et al. ColorPolarNet: residual dense network-based chromatic intensity-polarization imaging in low-light environment. *IEEE Trans Instrum Meas* **71**, 5025210 (2022).
51. Gao SH, Cao Y, Zhang WJ et al. Learning feature fusion for target detection based on polarimetric imaging. *Appl Opt* **61**, D15–D21 (2022).
52. Blin R, Ainouz S, Canu S et al. Road scenes analysis in adverse weather conditions by polarization-encoded images and adapted deep learning. In *Proceedings of 2019 IEEE Intelligent Transportation Systems Conference* 27–32 (IEEE, 2019); <https://doi.org/10.1109/ITSC.2019.8916853>.
53. Lei CY, Huang XH, Zhang MD et al. Polarized reflection removal with perfect alignment in the wild. In *Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition* 1747–1755 (IEEE, 2020); <https://doi.org/10.1109/CVPR42600.2020.00182>.
54. Fan W, Ainouz S, Meriaudeau F et al. Polarization-based car detection. In *Proceedings of the 2018 25th IEEE International Conference on Image Processing* 3069–3073 (IEEE, 2018); <https://doi.org/10.1109/ICIP.2018.8451397>.
55. Xie RC, Zu HY, Xue Y et al. Target detection method for polarization imaging based on convolutional neural network. *Proc SPIE* **11455**, 114557Z (2020).
56. Zhang Y, Morel O, Blanchon M et al. Exploration of deep learning-based multimodal fusion for semantic road scene segmentation. In *Proceedings of the 14th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications* 336–343 (SciTePress, 2019); <https://doi.org/10.5220/0007360403360343>.
57. Kalra A, Taamazyan V, Rao SK et al. Deep polarization cues for transparent object segmentation. In *Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition* 8599–8608 (IEEE, 2020); <https://doi.org/10.1109/CVPR42600.2020.00863>.
58. Si L, Huang TY, Wang XJ et al. Deep learning-based polarization feature retrieval from a single stokes vector. *Proc SPIE* **11963**, 1196307 (2022).
59. Sun R, Sun XB, Chen FN et al. Polarimetric imaging detection using a convolutional neural network with three-dimensional and two-dimensional convolutional layers. *Appl Opt* **59**, 151–155 (2020).
60. Li XP, Liao R, Zhou JL et al. Classification of morphologically similar algae and cyanobacteria using Mueller matrix imaging and convolutional neural networks. *Appl Opt* **56**, 6520–6530 (2017).
61. Li XP, Liao R, Ma H et al. Polarimetric learning: a Siamese approach to learning distance metrics of algal Mueller matrix images. *Appl Opt* **57**, 3829–3837 (2018).
62. Zhao YQ, Reda M, Feng K et al. Detecting giant cell tumor of bone lesions using Mueller matrix polarization microscopic imaging and multi-parameters fusion network. *IEEE Sens J* **20**, 7208–7215 (2020).
63. Dong Y, Wan JC, Wang XJ et al. A polarization-imaging-based machine learning framework for quantitative pathological diagnosis of cervical precancerous lesions. *IEEE Trans Med Imaging* **40**, 3728–3738 (2021).
64. Ba YH, Gilbert A, Wang F et al. Deep shape from polarization. In *Proceedings of the 16th European Conference on Computer Vision* 554–571 (Springer, 2020); [https://doi.org/10.1007/978-3-030-58586-0\\_33](https://doi.org/10.1007/978-3-030-58586-0_33).
65. Zou SH, Zuo ZX, Qian YM et al. 3D human shape reconstruction from a polarization image. In *Proceedings of the 16th European Conference on Computer Vision* 351–368 (Springer, 2020); [https://doi.org/10.1007/978-3-030-58568-6\\_21](https://doi.org/10.1007/978-3-030-58568-6_21)
66. Li DK, Lin B, Wang XY et al. High-performance polarization remote sensing with the modified U-Net based deep-learning network. *IEEE Trans Geosci Remote Sens* **60**, 5621110 (2022).
67. Lei CY, Qi CY, Xie JX et al. Shape from polarization for complex scenes in the wild. In *Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition* 12622–12631 (IEEE, 2021); <https://doi.org/10.1109/CVPR52688.2022.01230>.
68. Zhang JC, Shao JB, Luo HB et al. Learning a convolutional demosaicing network for microgrid polarimeter imagery. *Opt Lett* **43**, 4534–4537 (2018).
69. Yang KL, Bergasa LM, Romera E et al. Predicting polarization

- beyond semantics for wearable robotics. In *Proceedings of the 2018 IEEE-RAS 18th International Conference on Humanoid Robots* 96–103 (IEEE, 2018); <https://doi.org/10.1109/HUMANOIDS.2018.8625005>.
70. Zhang R, Gui XY, Cheng HY et al. Underwater image recovery utilizing polarimetric imaging based on neural networks. *Appl Opt* **60**, 8419–8425 (2021).
  71. Kondo Y, Ono T, Sun LG et al. Accurate polarimetric BRDF for real polarization scene rendering. In *Proceedings of the 16th European Conference on Computer Vision* 220–236 (Springer, 2020); [https://doi.org/10.1007/978-3-030-58529-7\\_14](https://doi.org/10.1007/978-3-030-58529-7_14).
  72. Shi YJ, Guo EL, Bai LF et al. Polarization-based haze removal using self-supervised network. *Front Phys* **9**, 789232 (2022).
  73. Wieschollek P, Gallo O, Gu JW et al. Separating reflection and transmission images in the wild. In *Proceedings of the 15th European Conference on Computer Vision* 90–105 (Springer, 2018); [https://doi.org/10.1007/978-3-030-01261-8\\_6](https://doi.org/10.1007/978-3-030-01261-8_6).
  74. Lyu YW, Cui ZP, Li S et al. Reflection separation using a pair of unpolarized and polarized images. In *Proceedings of the 33rd International Conference on Neural Information Processing Systems* 1304 (Curran Associates Inc., 2019); <https://doi.org/10.5555/3454287.3455591>.
  75. Liu TR, de Haan K, Bai BJ et al. Deep learning-based holographic polarization microscopy. *ACS Photonics* **7**, 3023–3034 (2020).
  76. Yang SL, Qu BW, Liu GS et al. Unsupervised learning polarimetric underwater image recovery under nonuniform optical fields. *Appl Opt* **60**, 8198–8205 (2021).
  77. Liu HD, Zhang YZ, Cheng ZZ et al. Attention-based neural network for polarimetric image denoising. *Opt Lett* **47**, 2726–2729 (2022).
  78. Yang XJ, Zhao QH, Huang TY et al. Deep learning for denoising in a Mueller matrix microscope. *Opt Express* **13**, 3535–3551 (2022).
  79. Wu XS, Zhang H, Hu XP et al. HDR reconstruction based on the polarization camera. *IEEE Robotics Autom Lett* **5**, 5113–5119 (2020).
  80. Deschaintre V, Lin YM, Ghosh A. Deep polarization imaging for 3D shape and SVBRDF acquisition. In *Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition* 15562–15571 (IEEE, 2021); <https://doi.org/10.1109/CVPR46437.2021.01531>.
  81. Shao MQ, Xia CK, Yang ZD et al. Transparent shape from a single view polarization image. arXiv: 2204.06331 (2023) <https://doi.org/10.48550/arXiv.2204.06331>.
  82. De S, Bruzzone L, Bhattacharya A et al. A novel technique based on deep learning and a synthetic target database for classification of urban areas in PolSAR data. *IEEE J Sel Top Appl Earth Obs Remote Sens* **11**, 154–170 (2018).
  83. Blanchon M, Morel O, Seulin R et al. Outdoor scenes pixel-wise semantic segmentation using polarimetry and fully convolutional network. In *Proceedings of the 14th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications* 328–335 (SciTePress, 2019); <https://doi.org/10.5220/0007360203280335>.
  84. Akkaynak D, Treibitz T, Shlesinger T et al. What is the space of attenuation coefficients in underwater computer vision? In *Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition* 568–577 (IEEE, 2017); <http://doi.org/10.1109/CVPR.2017.68>.
  85. Wang YH, Louie DC, Cai JY et al. Deep learning enhances polarization speckle for *in vivo* skin cancer detection. *Opt Laser Technol* **140**, 107006 (2021).
  86. Zhou XM, Ma L, Brown W et al. Automatic detection of head and neck squamous cell carcinoma on pathologic slides using polarized hyperspectral imaging and machine learning. *Proc SPIE* **11603**, 116030Q (2021).
  87. Yao Y, Zuo M, Dong Y et al. Polarization imaging feature characterization of different endometrium phases by machine learning. *OSA Continuum* **4**, 1776–1791 (2021).
  88. Azzam RMA. Division-of-amplitude photopolarimeter (DOAP) for the simultaneous measurement of all four Stokes parameters of light. *Opt Acta Int J Opt* **29**, 685–689 (1982).
  89. Jellison GE. Four-channel polarimeter for time-resolved ellipsometry. *Opt Lett* **12**, 766–768 (1987).
  90. Compain E, Drevillon B. Broadband division-of-amplitude polarimeter based on uncoated prisms. *Appl Opt* **37**, 5938–5944 (1998).
  91. Ju HJ, Ren LY, Liang J et al. A Mueller matrix measurement technique based on a division-of-aperture polarimetric camera. *Proc SPIE* **10839**, 108391F (2019).
  92. Pezzaniti JL, Chenault DB. A division of aperture MWIR imaging polarimeter. *Proc SPIE* **5888**, 58880V (2005).
  93. Gao SK, Gruev V. Bilinear and bicubic interpolation methods for division of focal plane polarimeters. *Opt Express* **19**, 26161–26173 (2011).
  94. York T, Gruev V. Calibration method for division of focal plane polarimeters in the optical and near-infrared regime. *Proc SPIE* **8012**, 80120H (2011).
  95. Hsu WL, Davis J, Balakrishnan K et al. Polarization microscope using a near infrared full-Stokes imaging polarimeter. *Opt Express* **23**, 4357–4368 (2015).
  96. Goldstein DH. *Polarized Light* 3rd ed (CRC Press, Boca Raton, 2017).
  97. Li XB, Liu TG, Huang BJ et al. Optimal distribution of integration time for intensity measurements in Stokes polarimetry. *Opt Express* **23**, 27690–27699 (2015).
  98. Mueller H. On the theory of scattering of light. *Proc Roy Soc A Math Phys Eng Sci* **166**, 425–449 (1938).
  99. Liu F, Zhang SC, Han PL et al. Depolarization index from Mueller matrix descatters imaging in turbid water. *Chin Opt Lett* **20**, 022601 (2022).
  100. Ortega-Quijano N, Haj-Ibrahim B, Garcia-Caurel E et al. Experimental validation of Mueller matrix differential decomposition. *Opt Express* **20**, 1151–1163 (2012).
  101. Lu SY, Chipman RA. Interpretation of Mueller matrices based on polar decomposition. *J Opt Soc Am A* **13**, 1106–1113 (1996).
  102. He HH, Zeng N, Du E 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**, 129–137 (2013).
  103. Arteaga O, Garcia-Caurel E, Ossikovski R. Anisotropy coefficients of a Mueller matrix. *J Opt Soc Am A* **28**, 548–553 (2011).
  104. Ortega-Quijano N, Arce-Diego JL. Mueller matrix differential decomposition. *Opt Lett* **36**, 1942–1944 (2011).
  105. Ossikovski R. Analysis of depolarizing Mueller matrices through a symmetric decomposition. *J Opt Soc Am A* **26**,

- 1109–1118 (2009).
106. Cloude SR. Group theory and polarisation algebra. *Optik* **75**, 26–36 (1985).
  107. Carnicer A, Javidi B. Polarimetric 3D integral imaging in photon-starved conditions. *Opt Express* **23**, 6408–6417 (2015).
  108. Chen C, Chen QF, Xu J et al. Learning to see in the dark. In *Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition* 3291–3300 (IEEE, 2018); <https://doi.org/10.1109/CVPR.2018.00347>.
  109. Tibbs AB, Daly IM, Roberts NW et al. Denoising imaging polarimetry by adapted BM3D method. *J Opt Soc Am A* **35**, 690–701 (2018).
  110. Gao SK, Gruev V. Gradient-based interpolation method for division-of-focal-plane polarimeters. *Opt Express* **21**, 1137–1151 (2013).
  111. Zhang JC, Luo HB, Hui B et al. Image interpolation for division of focal plane polarimeters with intensity correlation. *Opt Express* **24**, 20799–20807 (2016).
  112. Zhang JC, Luo HB, Liang RG et al. Sparse representation-based demosaicing method for microgrid polarimeter imagery. *Opt Lett* **43**, 3265–3268 (2018).
  113. Ratliff BM, LaCasse CF, Tyo JC. Interpolation strategies for reducing IFOV artifacts in microgrid polarimeter imagery. *Opt Express* **17**, 9112–9125 (2009).
  114. Wen SJ, Zheng YQ, Lu F et al. Convolutional demosaicing network for joint chromatic and polarimetric imagery. *Opt Lett* **44**, 5646–5649 (2019).
  115. Sargent GC, Ratliff BM, Asari VK. Conditional generative adversarial network demosaicing strategy for division of focal plane polarimeters. *Opt Express* **28**, 38419–38443 (2020).
  116. Sun YY, Zhang JC, Liang RG. Color polarization demosaicking by a convolutional neural network. *Opt Lett* **46**, 4338–4341 (2021).
  117. Pistellato M, Bergamasco F, Fatima T et al. Deep demosaicing for polarimetric filter array cameras. *IEEE Trans Image Process* **31**, 2017–2026 (2022).
  118. Zhang JC, Chen JL, Yu HW et al. Polarization image demosaicking via nonlocal sparse tensor factorization. *IEEE Trans Geosci Remote Sens* **60**, 5607210 (2021).
  119. Hagen N, Otani Y. Stokes polarimeter performance: general noise model and analysis. *Appl Opt* **57**, 4283–4296 (2018).
  120. Li XB, Hu HF, Liu TG et al. Optimal distribution of integration time for intensity measurements in degree of linear polarization polarimetry. *Opt Express* **24**, 7191–7200 (2016).
  121. Lewis JJ, O'Callaghan RJ, Nikolov SG et al. Pixel- and region-based image fusion with complex wavelets. *Inf Fusion* **8**, 119–130 (2007).
  122. Nencini F, Garzelli A, Baronti S et al. Remote sensing image fusion using the curvelet transform. *Inf Fusion* **8**, 143–156 (2007).
  123. Li ST, Kang XD, Hu JW. Image fusion with guided filtering. *IEEE Trans Image Process* **22**, 2864–2875 (2013).
  124. Liu Y, Liu SP, Wang ZF. A general framework for image fusion based on multi-scale transform and sparse representation. *Inf Fusion* **24**, 147–164 (2015).
  125. Li ST, Kang XD, Fang LY et al. Pixel-level image fusion: a survey of the state of the art. *Inf Fusion* **33**, 100–112 (2017).
  126. Zhang JC, Shao JB, Chen JL et al. PFNet: an unsupervised deep network for polarization image fusion. *Opt Lett* **45**, 1507–1510 (2020).
  127. Zhang JC, Shao JB, Chen JL et al. Polarization image fusion with self-learned fusion strategy. *Pattern Recognit* **118**, 108045 (2021).
  128. Lin TY, Dollár P, Girshick R et al. Feature pyramid networks for object detection. In *Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition* 936–944 (IEEE, 2017); <http://doi.org/10.1109/CVPR.2017.106>.
  129. Song YB, Ma C, Gong LJ et al. CREST: convolutional residual learning for visual tracking. In *Proceedings of 2017 IEEE International Conference on Computer Vision* 2574–2583 (IEEE, 2017); <https://doi.org/10.1109/iccv.2017.279>.
  130. Xu Y, Wen J, Fei LK, Zhang Z. Review of video and image defogging algorithms and related studies on image restoration and enhancement. *IEEE Access* **4**, 165–188 (2015).
  131. Guo Y, Liu RW, Lu YX et al. Haze visibility enhancement for promoting traffic situational awareness in vision-enabled intelligent transportation. *IEEE Trans Veh Technol* **72**, 15421–15435 (2023).
  132. Nguyen K, Nguyen P, Bui DC et al. Analysis of the influence of de-hazing methods on vehicle detection in aerial images. *Int J Adv Comput Sci Appl* **13**, 846–856 (2022).
  133. Liu JP, Wang SJ, Wang X et al. A review of remote sensing image dehazing. *Sensors* **21**, 3926 (2021).
  134. Karavarsamis S, Gkika I, Gkitsas V et al. A survey of deep learning-based image restoration methods for enhancing situational awareness at disaster sites: the cases of rain, snow and haze. *Sensors* **22**, 4707 (2022).
  135. Song YF, Nakath D, She MK et al. Optical imaging and image restoration techniques for deep ocean mapping: a comprehensive survey. *PFG J Photogramm Remote Sens Geoinf Sci* **90**, 243–267 (2022).
  136. Islam J, Xia YY, Sattar J. Fast underwater image enhancement for improved visual perception. *IEEE Robot Autom Lett* **5**, 3227–3234 (2020).
  137. Negi A, Chauhan P, Kumar K et al. Face mask detection classifier and model pruning with keras-surgeon. In *Proceedings of the 2020 5th IEEE International Conference on Recent Advances and Innovations in Engineering* 1–6 (IEEE, 2020); <http://doi.org/10.1109/ICRAIE51050.2020.9358337>.
  138. Narasimhan SG, Nayar SK. Vision and the atmosphere. *Int J Comput Vis* **48**, 233–254 (2002).
  139. Schechner YY, Narasimhan SG, Nayar SK. Polarization-based vision through haze. *Appl Opt* **42**, 511–525 (2003).
  140. Akkaynak D, Treibitz T. A revised underwater image formation model. In *Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition* 6723–6732 (IEEE, 2018); <http://doi.org/10.1109/CVPR.2018.00703>.
  141. Akkaynak D, Treibitz T. Sea-Thru: a method for removing water from underwater images. In *Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition* 1682–1691 (IEEE, 2019); <http://doi.org/10.1109/CVPR.2019.00178>.
  142. Hu HF, Zhao L, Li XB et al. Underwater image recovery under the nonuniform optical field based on polarimetric imaging. *IEEE Photonics J* **10**, 6900309 (2018).
  143. Wei Y, Han PL, Liu F et al. Polarization descattering imaging: a solution for nonuniform polarization characteristics of a target surface. *Chin Opt Lett* **19**, 111101 (2021).
  144. Fabbri C, Islam J, Sattar J. Enhancing underwater imagery using generative adversarial networks. In *Proceedings of 2018*

- IEEE International Conference on Robotics and Automation* 7159–7165 (IEEE, 2018);  
<http://doi.org/10.1109/ICRA.2018.8460552>.
145. Anwar S, Li CY, Porikli F. Deep underwater image enhancement, arXiv: 1807.03528 (2018). <https://doi.org/10.48550/arXiv.1807.03528>
  146. Wang N, Zhou YB, Han FL et al. UWGAN: underwater GAN for real-world underwater color restoration and dehazing. arXiv: 1912.10269 (2019). <https://doi.org/10.48550/arXiv.1912.10269>
  147. He KM, Sun J, Tang XO. Single image haze removal using dark channel prior. *IEEE Trans Pattern Anal Mach Intell* **33**, 2341–2353 (2011).
  148. Ren QM, Xiang YF, Wang GC et al. The underwater polarization dehazing imaging with a lightweight convolutional neural network. *Optik* **251**, 168381 (2022).
  149. Li BY, Peng XL, Wang ZY et al. AOD-Net: all-in-one dehazing network. In *Proceedings of 2017 IEEE International Conference on Computer Vision* 4780–4788 (IEEE, 2017);  
<http://doi.org/10.1109/ICCV.2017.511>.
  150. Peng YT, Cao KM, Cosman PC. Generalization of the dark channel prior for single image restoration. *IEEE Trans Image Process* **27**, 2856–2868 (2018).
  151. Fu XP, Liang Z, Ding XY et al. Image descattering and absorption compensation in underwater polarimetric imaging. *Opt Laser Eng* **132**, 106115 (2020).
  152. Drews PLJ, Nascimento ER, Botelho SSC et al. Underwater depth estimation and image restoration based on single images. *IEEE Comput Graph Appl* **36**, 24–35 (2016).
  153. Fu XY, Zhuang PX, Huang Y et al. A retinex-based enhancing approach for single underwater image. In *Proceedings of 2014 IEEE International Conference on Image Processing* 4572–4576 (IEEE, 2014);  
<http://doi.org/10.1109/ICIP.2014.7025927>.
  154. Fu XY, Fan ZW, Ling M et al. Two-step approach for single underwater image enhancement. In *Proceedings of 2017 International Symposium on Intelligent Signal Processing and Communication Systems* 789–794 (IEEE, 2017);  
<http://doi.org/10.1109/ISPACS.2017.8266583>.
  155. Li BY, Ren WQ, Fu DP et al. Benchmarking single-image dehazing and beyond. *IEEE Trans Image Process* **28**, 492–505 (2019).
  156. Sakaridis C, Dai DX, Van Gool L. Semantic foggy scene understanding with synthetic data. *Int J Comput Vis* **126**, 973–992 (2018).
  157. Sakaridis C, Dai DX, Van Gool L. Model adaptation with synthetic and real data for semantic dense foggy scene understanding. In *Proceedings of the 15th European Conference on Computer Vision* (Springer, 2018);  
[https://doi.org/10.1007/978-3-030-01261-8\\_42](https://doi.org/10.1007/978-3-030-01261-8_42).
  158. Zhang YF, Ding L, Sharma G. HazeRD: an outdoor scene dataset and benchmark for single image dehazing. In *Proceedings of 2017 IEEE International Conference on Image Processing* 3205–3209 (IEEE, 2017);  
<http://doi.org/10.1109/ICIP.2017.8296874>.
  159. Zhao SY, Zhang L, Huang SY et al. Dehazing evaluation: real-world benchmark datasets, criteria, and baselines. *IEEE Trans Image Process* **29**, 6947–6962 (2020).
  160. Carlevaris-Bianco N, Mohan A, Eustice RM. Initial results in underwater single image dehazing. In *Proceedings of OCEANS 2010 MTS/IEEE SEATTLE* 1-8 (IEEE, 2010);  
<http://doi.org/10.1109/OCEANS.2010.5664428>.
  161. Huang DM, Wang Y, Song W et al. Shallow-water image enhancement using relative global histogram stretching based on adaptive parameter acquisition. In *Proceedings of the 24th International Conference on Multimedia Modeling* 453–465 (Springer, 2018);  
[https://doi.org/10.1007/978-3-319-73603-7\\_37](https://doi.org/10.1007/978-3-319-73603-7_37).
  162. Huang SC, Cheng FC, Chiu YS. Efficient contrast enhancement using adaptive gamma correction with weighting distribution. *IEEE Trans Image Process* **22**, 1032–1041 (2013).
  163. Chao L, Wang M. Removal of water scattering. In *Proceedings of the 2010 2nd International Conference on Computer Engineering and Technology* V2-35–V2-39 (IEEE, 2010);  
<https://doi.org/10.1109/ICCET.2010.5485339>.
  164. Liang J, Ren LY, Ju HJ et al. Visibility enhancement of hazy images based on a universal polarimetric imaging method. *J Appl Phys* **116**, 173107 (2014).
  165. Hitam MS, Awalludin EA, Yussof WNJHW et al. Mixture contrast limited adaptive histogram equalization for underwater image enhancement. In *Proceedings of 2013 International Conference on Computer Applications Technology* 1–5 (IEEE, 2013); <http://doi.org/10.1109/ICCAT.2013.6522017>.
  166. Fan R, Wang HL, Cai PD et al. SNE-RoadSeg: incorporating surface normal information into semantic segmentation for accurate freespace detection. In *Proceedings of the 16th European Conference on Computer Vision* 340–356 (Springer, 2020); [https://doi.org/10.1007/978-3-030-58577-8\\_21](https://doi.org/10.1007/978-3-030-58577-8_21).
  167. Kazhdan M, Bolitho M, Hoppe H. Poisson surface reconstruction. In *Proceedings of the Fourth Eurographics Symposium on Geometry Processing* 61–70 (2006);  
<https://dl.acm.org/doi/10.5555/1281957.1281965>.
  168. Qi XJ, Liao RJ, Liu ZZ et al. Geonet: geometric neural network for joint depth and surface normal estimation. In *Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition* 283–291 (IEEE, 2018);  
<http://doi.org/10.1109/CVPR.2018.00037>.
  169. Huang JW, Zhou YC, Funkhouser T, Guibas L. FrameNet: learning local canonical frames of 3D surfaces from a single RGB image. In *Proceedings of 2019 IEEE/CVF International Conference on Computer Vision* 8637–8646 (IEEE, 2019);  
<http://doi.org/10.1109/ICCV.2019.00873>.
  170. Wang R, Geraghty D, Matzen K et al. VPLNet: deep single view normal estimation with vanishing points and lines. In *Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition* 686–695 (IEEE, 2020);  
<http://doi.org/10.1109/CVPR42600.2020.00077>.
  171. Bansal A, Chen XL, Russell B et al. PixelNet: towards a general pixel-level architecture. arXiv: 1609.06694 (2016).  
<https://arxiv.org/abs/1609.06694>
  172. Li B, Shen CH, Dai YC et al. Depth and surface normal estimation from monocular images using regression on deep features and hierarchical CRFs. In *Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition* 1119–1127 (IEEE, 2015);  
<http://doi.org/10.1109/CVPR.2015.7298715>.
  173. Atkinson GA. Polarisation photometric stereo. *Comput Vis Image Underst* **160**, 158–167 (2017).
  174. Fukao Y, Kawahara R, Nobuhara S et al. Polarimetric normal stereo. In *Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition* 682–690 (IEEE,

- 2021); <http://doi.org/10.1109/CVPR46437.2021.00074>
175. Zhu DZ, Smith WAP. Depth from a polarisation + RGB stereo pair. In *Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition* 7578–7587 (IEEE, 2019); <http://doi.org/10.1109/CVPR.2019.00777>.
176. Ding YQ, Ji Y, Zhou MY et al. Polarimetric helmholtz stereopsis. In *Proceedings of 2021 IEEE/CVF International Conference on Computer Vision* 5017–5026 (IEEE, 2021); <http://doi.org/10.1109/ICCV48922.2021.00499>.
177. Atkinson GA, Hancock ER. Recovery of surface orientation from diffuse polarization. *IEEE Trans Image Process* 15, 1653–1664 (2006).
178. Robles-Kelly A, Huynh CP. *Imaging Spectroscopy for Scene Analysis* (Springer, London, 2013).
179. Frankot RT, Chellappa R. A method for enforcing integrability in shape from shading algorithms. *IEEE Trans Pattern Anal Mach Intell* 10, 439–451 (1988).
180. Miyazaki D, Tan RT, Hara K et al. Polarization-based inverse rendering from a single view. In *Proceedings of the Ninth IEEE International Conference on Computer Vision* 982–987 (IEEE, 2003); <https://doi.org/10.1109/ICCV.2003.1238455>.
181. Mahmoud AH, El-Melegy MT, Farag AA. Direct method for shape recovery from polarization and shading. In *Proceedings of the 2012 19th IEEE International Conference on Image Processing* 1769–1772 (IEEE, 2012); <https://doi.org/10.1109/ICIP.2012.6467223>.
182. Loper M, Mahmood N, Romero J et al. SMPL: a skinned multi-person linear model. *ACM Trans Graph* 34, 248 (2015).
183. Han PL, Li X, Liu F et al. Accurate passive 3D polarization face reconstruction under complex conditions assisted with deep learning. *Photonics* 9, 924 (2022).
184. Levin A, Weiss Y. User assisted separation of reflections from a single image using a sparsity prior. *IEEE Trans Pattern Anal Mach Intell* 29, 1647–1654 (2007).
185. Li Y, Brown MS. Exploiting reflection change for automatic reflection removal. In *Proceedings of 2013 IEEE International Conference on Computer Vision* 2432–2439 (IEEE, 2013); <https://doi.org/10.1109/ICCV.2013.302>.
186. Zhou BL, Lapedriza A, Khosla A et al. Places: a 10 million image database for scene recognition. *IEEE Trans Pattern Anal Mach Intell* 40, 1452–1464 (2018).
187. Pang YX, Yuan MK, Fu Q et al. Progressive polarization based reflection removal via realistic training data generation. *Pattern Recognit* 124, 108497 (2022).
188. Girshick R, Donahue J, Darrell T et al. Rich feature hierarchies for accurate object detection and semantic segmentation. In *Proceedings of 2014 IEEE Conference on Computer Vision and Pattern Recognition* 580–587 (IEEE, 2014); <https://doi.org/10.1109/CVPR.2014.81>.
189. Girshick R. Fast R-CNN. In *Proceedings of 2015 IEEE International Conference on Computer Vision* 1440–1448 (IEEE, 2015); <https://doi.org/10.1109/ICCV.2015.169>.
190. Ren SQ, He KM, Girshick R et al. Faster R-CNN: towards real-time object detection with region proposal networks. *IEEE Trans Pattern Anal Mach Intell* 39, 1137–1149 (2017).
191. Redmon J, Divvala S, Girshick R et al. You only look once: unified, real-time object detection. In *Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition* 779–788 (IEEE, 2016); <https://doi.org/10.1109/CVPR.2016.91>.
192. Liu W, Anguelov D, Erhan D et al. SSD: single shot multibox detector. In *Proceedings of the 14th European Conference on Computer Vision* 21–37 (Springer, 2016); [https://doi.org/10.1007/978-3-319-46448-0\\_2](https://doi.org/10.1007/978-3-319-46448-0_2).
193. Law H, Deng J. CornerNet: detecting objects as paired keypoints. In *Proceedings of the 15th European Conference on Computer Vision* 765–781 (Springer 2018); [https://doi.org/10.1007/978-3-030-01264-9\\_45](https://doi.org/10.1007/978-3-030-01264-9_45)
194. Wolff LB. Polarization-based material classification from specular reflection. *IEEE Trans Pattern Anal Mach Intell* 12, 1059–1071 (1990).
195. Wolff LB. Surface orientation from polarization images. *Proc SPIE* 850, 110–121 (1988).
196. Tian Y, Zhang KB, Wang LY et al. Face anti-spoofing by learning polarization cues in a real-world scenario. In *Proceedings of the 4th International Conference on Advances in Image Processing* 129–137 (ACM, 2020); <https://doi.org/10.1145/3441250.3441254>.
197. Usmani K, Krishnan G, O'Connor T et al. Deep learning polarimetric three-dimensional integral imaging object recognition in adverse environmental conditions. *Opt Express* 29, 12215–12228 (2021).
198. Shen Y, Lin WF, Wang ZF et al. Rapid detection of camouflaged artificial target based on polarization imaging and deep learning. *IEEE Photonics J* 13, 7800309 (2021).
199. Liu T, Lu M, Chen BG et al. Distinguishing structural features between Crohn's disease and gastrointestinal luminal tuberculosis using Mueller matrix derived parameters. *J Biophotonics* 12, e201900151 (2019).
200. Shen YX, Huang RR, He HH et al. Comparative study of the influence of imaging resolution on linear retardance parameters derived from the Mueller matrix. *Biomed Opt Express* 12, 211–225 (2021).
201. Sun T, Liu T, He HH et al. Distinguishing anisotropy orientations originated from scattering and birefringence of turbid media using Mueller matrix derived parameters. *Opt Lett* 43, 4092–4095 (2018).
202. Reh binder J, Haddad H, Deby S et al. *Ex vivo* Mueller polarimetric imaging of the uterine cervix: a first statistical evaluation. *J Biomed Opt* 21, 071113 (2016).
203. Chue-Sang J, Bai YQ, Stoff S et al. Use of combined polarization-sensitive optical coherence tomography and Mueller matrix imaging for the polarimetric characterization of excised biological tissue. *J Biomed Opt* 21, 071109 (2016).
204. Novikova T, Pierangelo A, Manhas S et al. The origins of polarimetric image contrast between healthy and cancerous human colon tissue. *Appl Phys Lett* 102, 241103 (2013).
205. Pierangelo A, Manhas S, Benali A et al. *Ex vivo* photometric and polarimetric multilayer characterization of human healthy colon by multispectral Mueller imaging. *J Biomed Opt* 17, 066009 (2012).
206. Dubreuil M, Babilotte P, Martin L et al. Mueller matrix polarimetry for improved liver fibrosis diagnosis. *Opt Lett* 37, 1061–1063 (2012).
207. Wang WF, Lim LG, Srivastava S et al. Roles of linear and circular polarization properties and effect of wavelength choice on differentiation between *ex vivo* normal and cancerous gastric samples. *J Biomed Opt* 19, 046020 (2014).
208. Chen GH, Wang JY, Zhang AJ. Transparent object detection and location based on RGB-D camera. *J Phys Conf Ser* 1183,

- 012011 (2019).
209. Liu YJ, Jourabloo A, Liu XM. Learning deep models for face anti-spoofing: binary or auxiliary supervision. In *Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition* 389–398 (IEEE, 2018); <https://doi.org/10.1109/CVPR.2018.00048>.
210. Wang T, He XM, Barnes N. Glass object localization by joint inference of boundary and depth. In *Proceedings of the 21st International Conference on Pattern Recognition* 3783–3786 (IEEE, 2012).
211. Xu YC, Nagahara H, Shimada A et al. TransCut: transparent object segmentation from a light-field image. In *Proceedings of 2015 IEEE International Conference on Computer Vision* 3442–3450 (IEEE, 2015); <https://doi.org/10.1109/ICCV.2015.393>.
212. Zhang SF, Wang XB, Liu AJ et al. A dataset and benchmark for large-scale multi-modal face anti-spoofing. In *Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition* 919–928 (IEEE, 2019); <https://doi.org/10.1109/CVPR.2019.00101>.
213. Sun MH, He HH, Zeng N et al. Characterizing the microstructures of biological tissues using Mueller matrix and transformed polarization parameters. *Biomed Opt Express* 5, 4223–4234 (2014).
214. Li PC, Lv DH, He HH et al. Separating azimuthal orientation dependence in polarization measurements of anisotropic media. *Opt Express* 26, 3791–3800 (2018).
215. Gil JJ. Invariant quantities of a Mueller matrix under rotation and retarder transformations. *J Opt Soc Am A* 33, 52–58 (2016).
216. Le QV. Building high-level features using large scale unsupervised learning. In *Proceedings of 2013 IEEE International Conference on Acoustics, Speech and Signal Processing* 8595–8598 (IEEE, 2013); <https://doi.org/10.1109/ICASSP.2013.6639343>.
217. Dy JG, Brodley CE. Feature selection for unsupervised learning. *J Mach Learn Res* 5, 845–889 (2004).
218. Dike HU, Zhou YM, Deveerasetty KK et al. Unsupervised learning based on artificial neural network: a review. In *Proceedings of 2018 IEEE International Conference on Cyborg and Bionic Systems* 322–327 (IEEE, 2018); <https://doi.org/10.1109/CBS.2018.8612259>.
219. Pan SJ, Yang Q. A survey on transfer learning. *IEEE Trans Knowl Data Eng* 22, 1345–1359 (2010).
220. Pan SJ, Tsang IW, Kwok JT et al. Domain adaptation via transfer component analysis. *IEEE Trans Neural Netw* 22, 199–210 (2011).
221. Stevens SY, Delgado C, Krajcik JS. Developing a hypothetical multi-dimensional learning progression for the nature of matter. *J Res Sci Teach* 47, 687–715 (2010).
222. Xin M, Kundu S. Multi-task learning with high-dimensional noisy images. *J Am Stat Assoc* (2021).
223. Li T, Sahu AK, Talwalkar A et al. Federated learning: challenges, methods, and future directions. *IEEE Signal Process Mag* 37, 50–60 (2020).
224. Rubin NA, D'Aversa G, Chevalier P et al. Matrix Fourier optics enables a compact full-Stokes polarization camera. *Science* 365, eaax1839 (2019).
225. Zhang YX, Pu MB, Jin JJ et al. Crosstalk-free achromatic full Stokes imaging polarimetry metasurface enabled by polarization-dependent phase optimization. *Opto-Electron Adv* 5, 220058 (2022).
226. Tang DL, Shao ZL, Xie X et al. Flat multifunctional liquid crystal elements through multi-dimensional information multiplexing. *Opto-Electron Adv* 6, 220063 (2023).
227. Chen J, Wang DP, Si GY et al. Planar peristrophic multiplexing metasurfaces. *Opto-Electron Adv* 6, 220141 (2023).

## Acknowledgements

We are grateful for the financial support from the National Natural Science Foundation of China (Nos. 62205259, 62075175, 61975254, 62375212, 62005203 and 62105254), the Open Research Fund of CAS Key Laboratory of Space Precision Measurement Technology (No. B022420004), and the Fundamental Research Funds for the Central Universities (No. ZYTS23125).

## Author contributions

K Yang, F Liu and SY Liang contributed equally to this work and drafted the manuscript. M Xiang contributed to the part on restoration and enhancement of accurate polarization information. PL Han contributed on the part on polarimetric descattering and three-dimensional shape reconstruction. JP Liu contributed to the part on reflection removal. X Dong contributed to the part on target detection and semantic segmentation. Y Wei contributed to the part on biomedical imaging and pathological diagnosis. BJ Wang, K Shimizu, XP Shao provided resource support and supervised the project. All authors read, corrected and approved the manuscript.

## Competing interests

The authors declare no competing financial interests.



Scan for Article PDF