UAB
Universitat Autònoma de Barcelona

Polarimetry meets biophotonics: featuring new applications supported by advanced image processing and machine learning

Carla Rodríguez

Polarimetry meets biophotonics: featuring new applications supported by advanced image processing and machine learning

Carla Rodríguez

Submitted in fulfillment of the requirements for the degree of Doctor of Physics

Thesis supervisors:

Dr. Angel Lizana
Prof. Dr. Juan Campos

Department of Physics
Universitat Autònoma de Barcelona

Universitat Autònoma de Barcelona

Bellaterra, 2023

To my mum, the brightest light,
and my family, the warmest embrace

This work has been supported by the research project of "*Development of advanced polarimetric instrumentation for the characterization of materials and tissues; and surface metrology by deflectometry*"; RTI2018-097107-B-C31 and PID2021-126509OB-C21, Ministerio de Economía y Competitividad (Spanish MINECO).

Carla Rodríguez thanks the Ministerio de Economía y Competitividad (Spanish MINECO) for the grant "*Ayudas para contratos predoctorales para la formación de doctores 2019*", PRE2019-087613.

Acknowledgements

At the end of this journey, I find myself filled with an overwhelming sense of gratitude towards those who made it possible for me to complete this PhD.

I would like to begin by expressing my appreciation to my supervisors, Angel and Juan. As an undergraduate student, you trusted me without hesitation and provided invaluable guidance and support that helped me grow into the good scientist I am today. Thank you for accompanying me for so many years. I will be forever grateful for your trust.

Next, I would like to thank all my scientific collaborators. Enrique, Emilio, Teresa and Jordi, this could not have been possible without you. Special thanks to Mireia, who helped me going through all the plant measurements in just a few weeks. To my Optics group colleagues, Albert, Haolin and Irene. Even though we haven't met each other so much lately, I am grateful for your support during these years, specially in the beginning of my scientific career. Marifí, you are a constant source of inspiration. I will always hold dear your guidance, support, and friendship.

To the *cool corridor* people, thank you for all the coffee-break moments filled with joy. My heartfelt thanks to Mónica. I will forever keep in my mind the insane amount of laughs we have had, the door decoration, the incredible experiences in San Francisco and the customized map of the town, the forbidden-to-eat chocolate Xmas decoration and your absolutely delicious cakes and *Pokémon* cookies. Thank you for being such a very good friend, you made this journey so much brighter.

Much obliged to the *Theory division* of the *Max Planck Institute for the Science of Light*. Special thanks to the *Artificial Scientist Lab* team, I could not be more grateful. Since the very beginning I felt Erlangen as my new home. Carlos, Sören and Xuemei, thank you for taking me in as part of the crew. Mario, thank you for trusting me, specially when it comes to the aim of revolutionizing science. I feel incredibly lucky to be here.

I would also like to thank the *CrossFit Live n' Lift* team. One of the best experiences here in Germany has been joining the most supportive community I have ever met. Thank you for providing me with insane workouts full of heavy weights, for letting me push my own boundaries and improve myself every day. Special thanks to Dimi, for all the cheers in weightlifting class, the good times playing cards and eating burgers, and for being the best judge in the *CrossFit Open 2023*.

To my lifelong friends, Violeta and Martí. No words can do justice to our unbreakable bond. Despite the distance, your warm hugs transcend borders. To Sílvia and the family, thank you for welcoming me with open arms. My special thanks to Sami, for

all the fur you left on my hoodies and the *bobba-tea* eyes sneaking behind my computer screen.

To my family, I feel incredibly fortunate of having you. Thank you for sparking my curiosity and for giving me the strength to soar to great heights. Mum, we both know how tough life can be. Your unwavering love and support have shaped me into the woman I am today. You are the wind beneath my wings.

Finally, this is for my soulmate Jan. I met love when I met you.

Abstract

This thesis presents a comprehensive study on the potential of using imaging polarimetry methods for the inspection of biological tissues, both plant and animal origin. The suitability of polarimetric methods for the inspection of diseased plant tissues is demonstrated by measuring the experimental Mueller matrices of various plant specimens with different disease symptoms and infection stages using a complete image Mueller polarimeter. The ability to retrieve the depolarizing content of the plant samples is shown to enhance image contrast and reveal wounded regions not visible with regular inspections. In addition, to combine the discriminatory potential of polarimetric observables in a single image, we propose two new methods: (1) using the Euclidean distance between polarimetric values of different tissues, and (2) using a Normal (Gaussian) function based on polarimetric data to estimate the probability of belonging to a particular tissue. These two methods are tested on four biological samples of animal and plant origin and show potential for use in biomedical and botanical applications. We also present a predictive optical model method for tissue recognition using polarimetric indicators and Mueller matrices to recognize the tissue type of a sample in a single measurement. This method is successfully applied to discriminate between four different animal tissues: tendon, muscle, bone and myotendinous junction. Finally, using machine learning, we compare 12 classification models based on different polarimetric datasets to determine the most effective method for tissue classification in *ex-vivo* chicken samples. The results show that raw Mueller matrix elements are the most effective for the design of classification models. The provided results pave the way of imaging and classification methods based on depolarizing observables for applications in biophotonics, as for the early detection of pathologies or to implement automatic assisted procedures.

Preface

This thesis consists of a collection of research articles I published in peer-reviewed international scientific journals. The papers contained in this thesis are listed below.

A. **Carla Rodríguez**, Enrique Garcia-Caurel, Teresa Garnatje, Mireia Serra i Ribas, Jordi Luque, Juan Campos, and Angel Lizana, "Polarimetric observables for the enhanced visualization of plant diseases," *Scientific Reports*, vol. 12, pp. 14743, 2022. DOI: 10.1038/s41598-022-19088-6.

B. **Carla Rodríguez**, Albert Van Eeckhout, Enrique Garcia-Caurel, Angel Lizana and Juan Campos, "Automatic pseudo-coloring approaches to improve visual perception and contrast in polarimetric images of biological tissues," *Scientific Reports*, vol. 12, pp. 18479, 2022. DOI: 10.1038/s41598-022-23330-6.

C. **Carla Rodríguez**, Albert Van Eeckhout, Laia Ferrer, Enrique Garcia-Caurel, Emilio González-Arnay, Juan Campos and Angel Lizana, "Polarimetric data-based model for tissue recognition," *Biomedical Optics Express*, vol. 12, no. 8, pp. 4852–4872, 2021. DOI: 10.1364/BOE.426387.

D. **Carla Rodríguez**, Irene Estévez, Emilio González-Arnay, Angel Lizana and Juan Campos, "Optimizing the classification of biological tissues using Machine Learning models based on polarized data," *Journal of Biophotonics*, e202200308, 2022. DOI 10.1002/jbio.202200308.

Contents

PART I

INTRODUCTION

1 Lighting the way: a look at the history and applications of light in biology and biomedical research

Since the earliest times, humankind has been learning from nature's phenomena. When it comes to the science of optics, the first discoveries involve the ancient Egyptians and Mesopotamians (Babylonians and Assyrians), who polished quartz lenses in order to replicate the behavior of light in water. Subsequently, the ancient Greeks described the light as a ray and proposed reflection laws. In turn, various vision theories were developed in classical antiquity, based either on the eyes of the beholder emitting light rays or contrarily supporting that the objects were particular light sources. Later, during the Islamic Golden Age (from the 8th century to the 14th century), Ibn al-Haytham conducted a wide variety of optical experiments and wrote the *Kitab al-Manathir* (in English, "Book of Optics") [1], which gathered all the results and proposed a new and more accurate description of the nature of light. He also laid the foundations of photography with the design of the *camera obscura*.

After several centuries of developments, the strong scientific understanding in light science and technology has given place to spectacular contemporary discoveries. Some examples are the Laser Interferometer Gravitational-Wave Observatory (LIGO) [2], capable to detect gravitational waves using the basic principle of light interferometry, and the James Webb Space Telescope (JWST), which revealed galaxies once invisible to us thanks to its sensibility to capture infrared radiation. Importantly, the strong interest on exploiting the fundamental properties of light-matter interaction has accelerated the progress in the design of new light-based technologies. This innovative environment has allowed great advances in the biomedical research field as, for instance, the non-invasive techniques which are used to treat tumors. A very popular practice is the so-called photodynamic therapy (PDT), which takes advantage of the natural aggregation of photosensitizers to tumoral tissue. Accordingly, the molecules are externally induced to chemically react and destroy the cancer cells by means of controlled light [3].

With the aim of enhancing the image contrast and the resolution of the different biological structures, a wide variety of new and advanced imaging techniques have been developed [4, 5]. All of these exploit the different light-matter interactions of scattering and absorption processes. Particular studies demonstrate the rapid detection and location of *Knightia* spp. fish fossils from the Eocene epoch in Green River Formation (United States), by combining various optical techniques, such as fluorescence, Raman spectroscopy and scanning electron microscopy [6]. Another example is given by the so-called multi-spectral and hyper-spectral imaging, which are based on the optical phenomena of different wavelengths reaching different tissue depths [7, 8]. Accordingly, the

intensity response of a sample under different light wavelengths (e.g., infrared, visible range, etc.) provides information of various tissue layers. This may lead to a differentiation between two or more structures which were unable to be distinguished under the same wavelength [9]. These techniques have demonstrated some advantages in dermatology and cancer inspection [10–14] and in plant science; this last, for instance, to diagnose some plant diseases and stress [15, 16] or plant phenotyping [17], among others. When it comes to the microscopic world, a wide variety of optical methods are being applied for inspection of biological samples. In this context, the measurement of their physical properties and the proper visualization of different inherent structures is of special interest. These approaches include optical coherence tomography (OCT) [18–20], electron microscopy [21–23], phase contrast microscopy [24–27] and Raman spectroscopy [28–30], among others. Particular examples are the optical tweezers [31], which have been used to measure the DNA elastic parameters and its phase transitions when subjected to different forces. Furthermore, super-resolution techniques such as STED microscopy (Stimulated Emission Depletion) [32, 33], overcome the diffraction limit in spatial resolution, thus allowing the precise measurements of the length of DNA fragments [34] or to image intracellular proteins [35]. As a representative example, Fig. 1 shows the STED images of denditric structures of mouse neurons of less than 70 nm in size (Fig. 1d), which indicates that the achieved resolution is at least of that order of magnitude.

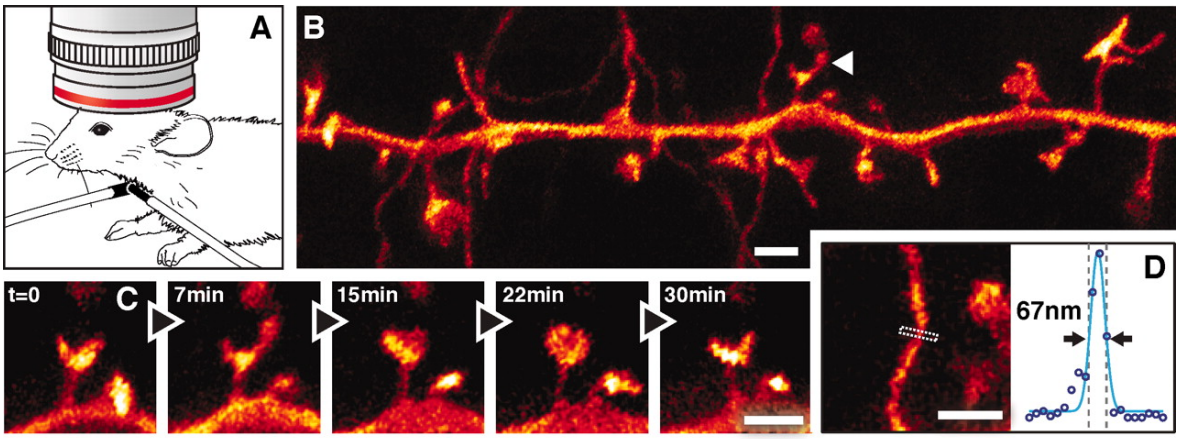


Figure 1: STED microscopy in the molecular layer of the somatosensory cortex of a mouse with EYFP-labeled neurons. (A) Anesthetized mouse under the objective lens (63X, NA 1.3, glycerol immersion) with tracheal tube. (B) Projected volumes of dendritic and axonal structures reveal (C) temporal dynamics of spine morphology with (D) an approximately fourfold improved resolution compared with diffraction-limited imaging. Curve is a three-pixel-wide line profile fitted to raw data with a Gaussian. Scale bars, 1 μm. [Reproduced from [33]].

Another widely-used technique is confocal microscopy [36], which plays a potential role in medical applications, such as for melanoma diagnosis [37] or cardiovascular solutions, as it provides very precise measurements of the surface coatings of stents [38, 39]. Importantly, countless biological structures are fluorescent [40, 41], specially in plant tissues [42, 43]. The strong understanding on how this phenomena is produced has allowed its combination with other optical techniques such as confocal microscopy [44, 45]: as a representative example, in Ref. [46] both techniques allow to keep track of the growth direction of root tips and the migrating cell dynamics in developing embryo of zebrafish.

Some of the above-mentioned methods are invasive and require sample preparation (e.g., the injection of some dyes to generate fluorescence [47]). In contrast, image polarimetry exploits both the non-invasive and non-contact optical measurement techniques. Polarimetry is based on the study of polarization (i.e., property of light, describes the orientation of the oscillation of the electrical field of EM waves) and how it is modified due to their interaction with matter [48, 49]. Those testing techniques provide information about the optical properties of the probed sample and appear as very useful tools to be implemented in different research fields. Some particular examples of their interest are related to astronomy [50–53], for instance to detect planets outside the solar system [51]. Also for material characterization [54–56], as the strain characterisation in transparent films [56]; for security and remote sensing [57–60], as to detect antipersonnel mines [59], and atmospheric pollution studies [61–64], such as the detection of atmospheric aerosols [62], among others.

Biological tissues are turbid media that strongly scatter light. That is, the incoming photons take different optical paths across the tissue, which induces changes in their polarization state. In fact, the different ultra-structure properties produce certain polarimetric responses which can be understood as specific polarimetric signatures of the media. Accordingly, polarimetry takes advantage of these different polarimetric properties of biological tissues and defines a set of useful tools to enhance the image contrast of some organic structures and to provide information of certain structures invisible by using regular (non-polarimetric) images [65–83]. As for the case of some biological samples presenting certain characteristic degree of anisotropy due to their basic components. In botany, the birefringence signature of some structures and macro-molecules (e.g., cellulose) allowed the study of the cell wall composition in plant samples [75]. Similarly, in animal samples, the strong birefringent signature of dense collagen and elastin matrix due to their oriented fibers in animal samples [74, 77, 83] allowed to reveal inhomogeneous regions within intermediate layers of articular cartilages in chicken knees [76]. As a representative example, Fig. 2 shows the comparison

between intensity and polarized images taken from a sheep optic nerve sample [74]. Clearly, the combined use of polarized light and the association of different colors with a given collagen fiber orientation, eases the visualization of the various types of fiber architectures, bundles and interactions.

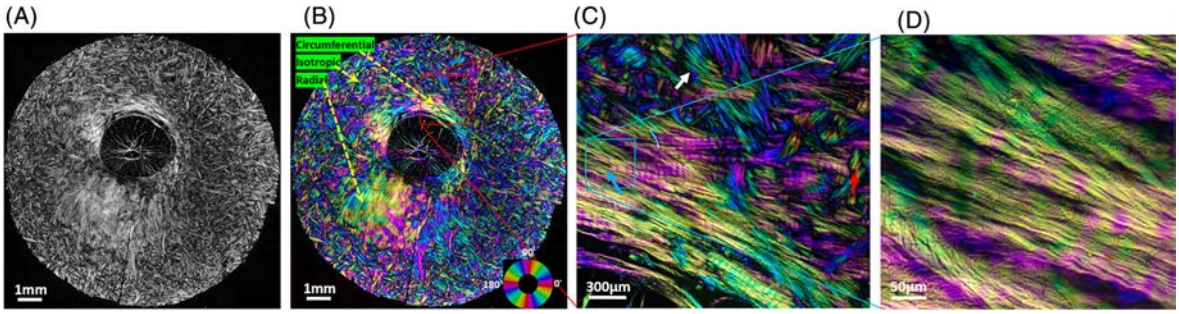
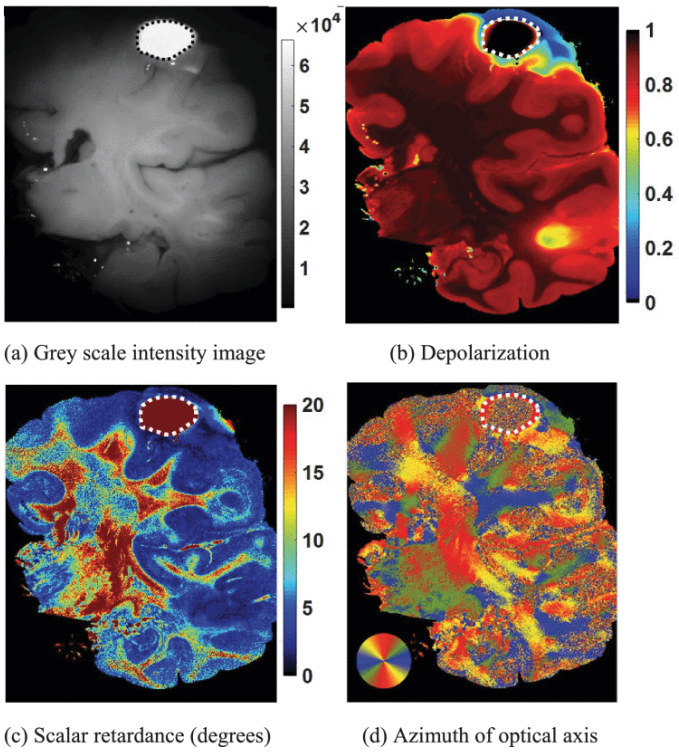


Figure 2: a) Intensity and b) instant polarized light microscopy (IPOL) image of a sheep optic nerve head section. Three major organizational components of the sclera are discernible. c) Close-up of peripapillary sclera revealed highly detailed collagen fiber features: two families of crossing fiber (white arrow), collagen fiber undulations or crimp (red arrow), and sub-bundle fiber composition (cyan arrow). d) Close-up of a bundle shows its fiber composition. [Reproduced from [74]].

Polarimetric techniques can be combined with other optical methods to take advantage of different light-matter interaction features. As for the case of applications in biophotonics, when it comes to dermatology [65, 84–88], polarimetric measurements have demonstrated to ease the delimitation of the diseased tissue region [85] as well as to assist for a more reliable diagnose even at early stages [65, 84]. Furthermore, the use of polarimetric measurements and multispectral imaging has allowed the better visualization, in vivo, of skin pathologies in dark skin types [65], as well as their earlier diagnose [84]. Also it is useful for monitoring the hemoglobin concentration and oxygen saturation in the superficial layer of tissue [73]. Other studies found the advantage of using polarization gating techniques [89] combined with fluorescence microscopy for the diagnosis of oral cancer [66] and breast cancer [67] at cellular level. Second harmonic polarimetry methods allow to characterize the orientation of 3D collagen fibrils [69, 70] and skeletal muscle [71], among others.

One important polarimetric signature of biological samples is their inherent capability to induce depolarization to the incident light, which has been considered as a deletory effect which tends to screen or to erase the polarimetric properties of the samples. Depolarization is defined as a statistical concept originated by the incoherent addition (either temporal or spatial) of different light polarizations at the level of the detector. In other words, it can be understood as the degree of polarization disorder (or polarimetric randomness) introduced by a given tissue structure to an input polarization. It is of particular interest for the inspection of biological samples, as they result in

scattering media, and the non-uniform distribution of polarization properties along the tissues are the main causes of depolarization. However, the relation between the depolarization response of the sample and its intrinsic physical structure and properties are still not well understood [90–94]. Advantageously, a widespread number of macroscopic measurements are based on Mueller polarimeters [58, 95] which compound a useful tool to analyze the depolarization content of samples. For instance, some studies related to glaucoma, tracked the depolarization response of both healthy and diseased eyes. They found that patients suffering from glaucoma presented higher depolarization response than individuals with healthy eyes [82, 83]. When it comes to cancer-related research fields [3, 18, 19, 66, 67], there is an important workforce focused on providing optical solutions for its early detection and diagnosis [10–14, 37, 96–112]. For instance, to inspect the polarimetric signatures of brain tissue, could allow to better delineate the tumor border in brain cancer [97]. As shown in Fig. 3, the contrast between the cortex and white matter of the brain was increased when inspecting the total depolarization response of the human brain sample [97].



(a) Grey scale intensity image (b) Depolarization

(c) Scalar retardance (degrees) (d) Azimuth of optical axis

Figure 3: Images of a thick coronal section of fixed human brain immersed in water and measured at 550 nm. The dashed line delineates the area of specular reflection on the air–water interface. [Reproduced from [97]].

Also, polarimetric imaging has been proved to be useful to characterize biological signatures related to liver cancer [98], uterine and ovarian cancer [99, 104–106, 113],

breast cancer [67, 100, 108, 109], skin cancer [101, 102, 114, 115] and colon cancer [103, 110–112], among others [66, 105, 107]. Furthermore, the recently use of Artificial Intelligence (AI) - assisted inspection of biological tissues [111, 112, 114–119] has allowed the use of polarimetric information for tissue classification tasks. Some studies use some widely-used supervised and unsupervised machine learning (ML) techniques [120], such as convolutional neural networks [121], Principal Components Analysis [122], Logistic Regression [123], Random Forest [124], and Support Vector Machines [125], among others [126]. In particular, it has been demonstrated how polarimetric observables constitute reliable metrics capable to distinguish between healthy and cancerous ex-vivo colon samples [111, 112] as well as basal cell carcinoma in human skin [114].

The study of polarimetric properties of animal and human tissues is a well established field of work in constant development [10–14, 37, 65–88, 96–112, 114, 116–119]. However, the application of polarimetric methods for the study of plant samples is less common and, in the last decade, there is a growing interest of exploring more complex (and rich in terms of information) polarimetric solutions for applications in plant science [57, 58, 75, 127–130]. It has been demonstrated that the inspection of different polarimetric signatures of plant samples (i.e., dichroism [57, 131–135], birefringence [57, 75, 134, 136–139] and depolarization [57, 140–148]) lead to relevant information that contribute to the further knowledge and understanding of biological processes. Some interesting results demonstrate that the light reflected / scattered by the leaf surface and its internal structure can be well differentiated through polarimetric measurements [131, 132]. Furthermore, the optical response from the internal leaf structure was dependent on the leaf pigmentation, among others [131]. Other studies demonstrate distinctive polarimetric features for the leaf tissues and leaf veins [57], thus allowing to better differentiate those structures when inspecting their optical response to different incident states of polarization. Furthermore, dichroism measurements reveal the spatial organization and concentration of some plant organelles such as pigment–protein complexes in plant thylakoid membranes [133], chloroplasts from onion and tobacco [134] or spinach-extracted quantasomes [135]. The polarimetric signature of birefringence is commonly inspected in plant structures and organelles [75, 136–138]. For instance, it is useful to study the cellulose micro-fibril orientation in developing guard cells of *Allium* and Cotton [138, 139], as well as the structure of the trichromes in *Arabidopsis thaliana* [136] and stomata [137], or the cell wall composition of different phylogenetic groups [75], among others. As a representative example, Fig. 4 shows how polarization unveils the wide variety of cellulose microfibril orientations in stomatal complex of plant samples.

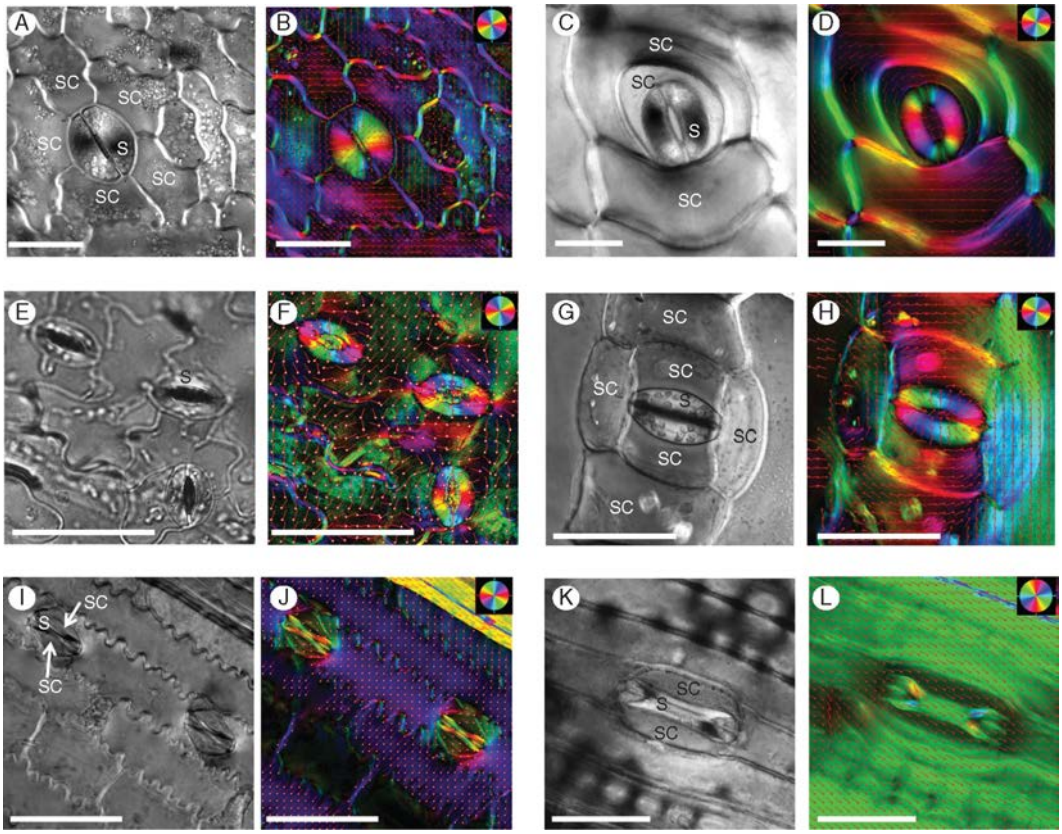


Figure 4: PolScope images of stomata showing crystalline cellulose orientation. Representative polarized light (left) and colour-coded images (right) of cellulose microfibril orientation are presented for each species. (A, B) *Asplenium*, (C, D) *Platycerium*, (E, F) *Arabidopsis*, (G, H) *Commelina* (note the birefringent crystals in the epidermis), (I, J) *Sorghum*, (K, L) *Triticum*. The orientation colour pie-chart codes the cellulose microfibril orientation for every image. The red vector arrows also show the orientation direction for a clearer view. S, stoma; SC, subsidiary cell. Size bars =50 μm. [Reproduced from [75]].

When it comes to depolarization [57, 140–148], it has been underused in plant science [140], although it provides relevant information. Alike for the case in animal tissues, the constituent units of plant samples (i.e., cellulose, pectin, water, etc.) present different polarimetric features, density, shape and spatial organization. Accordingly, this environment induces depolarization at macroscopic scale. In particular, depolarization signatures in plant specimen revealed different plant stress and leaf pollution levels in leaves and shoots of woody and agricultural plants [140, 141] such as corn and soybean crop canopies [142]. Importantly, the wide variety of polarimetric responses in plant samples is of interest, for instance, to be used in plant classification and characterization tasks [57, 143–148]. For instance, the depolarization content of vegetation is used in remote sensing applications. Particular studies differentiate among plant species by inspecting variations in surface features which, in turn, modify the incident

polarization [143]. Other studies use the depolarization signature to discriminate land mines from background vegetation [144] or to detect and identify extraterrestrial life forms of plants and bacteria [145], among others [146–148]. More recent studies [57, 127, 128] demonstrate that the inspection of depolarization signatures in plant samples involves the further physical understanding and the enhance of image contrast among the different plant structures. Few examples of interest are the midrib, stomata and raphides in *Epipremnum aureum* [128] and vascular webs in *Hedera marocanna* [127, 149]. Still, in botanical applications, the specific inspection of depolarization channel is not as common as the use of other polarimetric signatures. In this sense, some perspectives include the research on new imaging techniques that track the polarimetric responses of different plant diseases, whose characterization may be of interest for preventing yield and economic losses [150] due to pests and diseases [151].

1.1 Goals of this thesis

The main objectives of this thesis are related to the use of imaging polarimetry in biological applications. A variety of polarimetric techniques have been developed and applied for image enhancement in biological applications, including the use of polarimetric observables to study biological tissues for the inspection of pathologies, enhancement of image contrast of biological samples, and its classification. It is noteworthy that the same optical instrumentation has been used throughout the study to measure the experimental Mueller matrix of the samples, and the same polarimetric observables have been retrieved from the wide variety of studied organic samples.

The main goals of this thesis are:

1. Mueller polarimetry for plant pathology inspection: to use imaging polarimetry for the analysis of plant tissues, specifically to study the influence of depolarizing metrics in plant pathology inspection.

2. Design new image processing methods using Mueller polarimetry: to design new pseudo-coloration methods for enhancing image contrast in polarimetric images of biological samples corresponding to tissues with different characteristics that cannot be detected with regular intensity images.

3. Use statistics and machine learning based on polarimetric information: to design classification algorithms for animal tissue recognition that can be applied to human tissue research. Also, to extend the recognition models to more sophisticated algorithms using machine learning techniques and analyze tendencies in the polarimetric databases that could enhance the efficiency of the classification.

2 Theoretical framework

Light beams can be distinguished in terms of their state of polarization (SoP): (1) partially or fully unpolarized, when the electrical field oscillates randomly varying its its polarization ellipse in time [48]; (2) linearly polarized, when the electrical field oscillates in a particular (single) direction; or (3) circularly/elliptically polarized, for light beams showing their electrical field rotating at a constant rate in a given plane. Different mathematical formalism describe polarized light under different assumptions thus allowing the expression of physical characteristics being useful at different frameworks. For instance, R. C. Jones developed a formalism based on the electrical field components of light and how anisotropic materials modify their properties [152]; however, it is restricted to fully polarized light beams. When it comes to the description of polarized light when interacting with biological tissues, the Mueller-Stokes (M-S) is the formalism which suits better [153, 154]: it allows to characterize the state of polarization of light beams (Stokes vector) and the polarimetric properties of samples (Mueller matrices, $\mathbf{M}$). In particular, biological tissues are strong depolarizers: they induce randomness on the amplitude and relative phase of the incident light wave components. From this perspective, M-S includes the description of fully, partially and unpolarized light beams. Moreover, the M-S formalism is based on radiometric measurements and therefore is easy to be implemented experimentally and to interpret the data. Accordingly, any state of polarization (fully or partially) of a light beam is defined by means of the Stokes vector $\mathbf{S}$ as [153]:

$$S = \begin{pmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \end{pmatrix} = \begin{pmatrix} I_{0^\circ} + I_{90^\circ} \\ I_{0^\circ} - I_{90^\circ} \\ I_{45^\circ} - I_{135^\circ} \\ I_{CL} - I_{CR} \end{pmatrix}. \tag{1}$$

where the four Stokes parameters (i.e., the S_0, S_1, S_2 and S_3) are related to the total intensity of the light beam (S_0) and quantify the amount of horizontal/vertical linear polarization (S_1), the diagonal/anti-diagonal polarization (S_2) or the right/left-handed circular polarization (S_3). These parameters are defined by means of the intensity measurements when analyzing the horizontal/vertical polarization (i.e., I_{0° and I_{90°, respectively), the diagonal/anti-diagonal (i.e., I_{45° and I_{135°, respectively) and the right/left-handed circular polarization (i.e., I_{CR} and I_{CL}, respectively). Fully unpolarized light is encoded by $\mathbf{S}$ when all the intensity measurements are equal to the same amount, i.e., $S = (S_0, 0, 0, 0)^T$. Related with this, the degree of polarization (DoP) [48, 49], which characterizes the depolarization behavior of light beams, is defined as:

$$DoP = \frac{1}{S_0}\sqrt{S_1^2 + S_2^2 + S_3^2}, \tag{2}$$

which allows to directly retrieve if the state corresponds to fully polarized ($DoP = 1$), unpolarized ($DoP = 0$) or partially polarized light ($0 < DoP < 1$).

When describing light-matter interaction, Hans Mueller included the sample description to the Stokes formalism by means of the Mueller matrix $\mathbf{M}$ [154], a real 4×4 matrix. Importantly, the Mueller matrix of a sample does not only depend on their intrinsic characteristics (e.g., biological components, structures, etc.) but on the incident wavelength and direction of both the incident and the sample's output light beams [48, 49]. Thus, the sample ($\mathbf{M}$) modifies the SoP of the incident light beam (S_{in}) and outputs the state S_{out} by means of the following linear process:

$$S_{out} = M \cdot S_{in}. \tag{3}$$

An important advantage of polarimetric-based optical methods is that the measurement of the physical properties of the samples is independent of the intensity of the illuminating light beam. Importantly, different structures may disperse or transmit the same amount of light; however, these objects may modify the SoP of the incident light beam in a different way. This dissimilarity allows to distinguish structures using polarimetric measurements [155]. Some materials induce a phase-shift between the electromagnetic components of the incident light wave without modifying its intensity or the degree of polarization (DoP). We refer to the medium which is capable to produce such phenomena as birefringent. In optics, we take advantage of the wide variety of materials that are mainly characterized by this property, and use them as optical retarders (such as wave-plates). On the other hand, some materials, in particular biological samples or sky clouds, are characterized by its capability to reduce the DoP of the incident light beam. Those are called depolarizers. Furthermore, dichroic materials are characterized by the transmittance anisotropy, the dependence of the output intensity with the incident state of polarization.

In this way, samples are classified under the criteria of their polarimetric behavior as their physical properties are encoded into the elements of $\mathbf{M}$, m_{ij} (where $i, j = 0, ..., 3$) [48, 49]:

$$M = \begin{pmatrix} m_{00} & m_{01} & m_{02} & m_{03} \\ m_{10} & m_{11} & m_{12} & m_{13} \\ m_{20} & m_{21} & m_{22} & m_{23} \\ m_{30} & m_{31} & m_{32} & m_{33} \end{pmatrix}. \tag{4}$$

The Mueller matrix provides the intensity and the complete set of polarimetric properties (retardance, depolarization and dichroism) which are sensitive to the microarchitecture of the tissue [156]. These polarimetric characteristics are not related to a particular element of $\mathbf{M}$ but they appear entangled in multiple $\mathbf{M}$ elements in a complex way. Particularly, the dichroic content of the sample is defined by the combination

of (1) diattenuation, D, which describes the property of the sample by which the intensity of the exiting beam, S_{out}, depends on the polarization state of the incident beam S_{in}, and (2) polarizance, P, which describes the polarization capability of the sample when illuminated by an unpolarized input state S_{in}. For this reason, it is suitable to write the Mueller matrix $\mathbf{M}$ in a more compact (block) form [48]:

$$M = m_{00} \begin{pmatrix} 1 & D^T \\ P & m \end{pmatrix}, \tag{5}$$

where m_{00} denotes for the mean intensity of transmittance or reflectance of unpolarized input states, D and P correspond to the 3-dimensional diattenuation and polarizance vectors, respectively, and the 3×3 sub-matrix m encodes the sample information about retardance (i.e., birefringence) and depolarization. Importantly, the so-called Components of Purity (CPs) [48] are defined by the triplet composed by polarizance, P, dichroism, D and the sphericity degree P_S. This last defines the contribution on depolarization which differs from dichroic origin. The CPs can be directly computed from $\mathbf{M}$ as [48, 49]:

$$D = \frac{\sqrt{m_{01}^2 + m_{02}^2 + m_{03}^2}}{m_{00}}, \qquad P = \frac{\sqrt{m_{10}^2 + m_{20}^2 + m_{30}^2}}{m_{00}}, \qquad P_S = \frac{||m||_2}{\sqrt{3}}, \tag{6}$$

where m_{ij} $(i, j = 0, ..., 3)$ are elements of $\mathbf{M}$ and m denotes for the the 3×3 sub-matrix of $\mathbf{M}$. However, further mathematical treatment is required to disentangle the birefringent and depolarizing behavior of the sample.

These polarimetric properties (retardance and depolarization) can be evaluated after decomposing the original $\mathbf{M}$ into a set of matrices of easier physical interpretation of the medium. For this purpose, a wide variety of matrix decompositions can be conducted, either based on the product, sum or differential decomposition strategies [48]. Importantly, the choice of the method is not trivial. In fact, although in some cases different decompositions may lead to comparable results, in general, the method strongly depends on the specific sample. For the studies conducted in this thesis, we chose the commonly used Lu-Chipman decomposition [157] due to its high numerical efficiency. The Lu-Chipman formalism describes any Mueller matrix $\mathbf{M}$ as the product of three 4×4 pure Mueller matrices [157] as follows:

$$M = m_{00} \hat{M}_\Delta \hat{M}_R \hat{M}_D, \tag{7}$$

where m_{00} denotes for the unpolarized transmission / scattering of the sample, and $\hat{M}_\Delta$, $\hat{M}_R$ and $\hat{M}_D$ correspond to the well-defined polarimetric observables of pure depolarizers, pure retarders and pure diattenuators, respectively. Therefore, the information

related to these properties of samples is encoded in different scalar polarimetric observables derived from those pure matrices. In particular, retardance parameters can be easily retrieved from the pure retarder matrix $\hat{M}_R$:

$$\hat{M}_R = \begin{pmatrix} 1 & 0^T \\ 0 & m_R \end{pmatrix}, \tag{8}$$

where m_R corresponds to a 3×3 sub-matrix. Accordingly, the total retardance R (which describes the global behavior of a general retarder [158]), the linear retardance, δ, and the optical rotation, Ψ (i.e., the rotation applied to the linear retarder) can be defined as [159]:

$$R = cos^{-1} \left| \frac{Tr(\hat{M}_R)}{2} - 1 \right|, \tag{9}$$

$$\delta = cos^{-1} \left(\sqrt{(M_{R11} + M_{R22})^2 + (M_{R21} - M_{R12})^2} - 1 \right), \tag{10}$$

and

$$\Psi = tan^{-1} \left(\frac{M_{R21} - M_{R12}}{M_{R11} + M_{R22}} \right), \tag{11}$$

where Tr denotes for the trace of the matrix and $M_{Ri,j}$ are the matrix elements of $\hat{M}_R$.

When it comes to depolarization, the pure depolarizer matrix defined from Lu-Chipman decomposition is written as:

$$\hat{M}_\Delta = \begin{pmatrix} 1 & 0^T \\ P_\Delta & m_\Delta \end{pmatrix}, \tag{12}$$

where P_Δ corresponds to the polarizance vector, and m_Δ is the non-diagonal 3×3 sub-matrix. Similarly to the definition of above-described retardance parameters, the depolarization coefficient (or depolarization power) Δ, that indicates the average depolarizing capability of the Mueller matrix $\mathbf{M}$, is defined in the following way:

$$\Delta = 1 - \frac{|Tr(m_\Delta)|}{3}, \quad \text{where} \quad 0 \leq \Delta \leq 1. \tag{13}$$

Importantly, depolarization has been considered as a deletory effect which tends to screen or to erase the polarimetric properties (traditionally considered as the desired observables). However, novel depolarization metrics introduced in the literature [126, 160–167], demonstrate that there is a potential interest of considering both polarization and depolarization effects of media of equal importance. These depolarization-related metrics describe the process in which the purity of polarization of the incoming light beam is lost after interacting with the sample. For the particular case of biological tissues, depolarization effects are mainly induced by the light scattering (e.g., be due

to the water contained in tissues) and the non uniformity distribution of different polarimetric properties within the sample, directly related to different sample regions showing a wide variety of biological compositions, symmetric / aligned structures, etc. Since the information about these ultra-structures present in the tissues (i.e., structural and chemical variety) is encoded into the depolarization response, the depolarization metrics may unveil differences or particularities, overlooked when probed with non polarized light, which can present a true interest for biological / medical purposes or characterization.

To describe the depolarization behavior of media, the depolarization index[48], P_Δ, is widely used. This metric is similar to the depolarization coefficient Δ in the sense that both describe, quantitatively, the overall depolarizing power of the sample. However, P_Δ is directly computed from the Mueller matrix elements as:

$$P_\Delta \equiv \sqrt{\frac{\left(\sum_{i,j=0}^{3} m_{ij}^2\right) - m_{00}^2}{3m_{00}^2}}, \quad \text{being} \quad 0 \leq P_\Delta \leq 1, \tag{14}$$

where m_{ij} $(i,j = 0,...,3)$ are elements of $\mathbf{M}$. Furthermore, the depolarization information contained in P_Δ can also be split in terms of the Components of Purity (P, D and P_S) as,

$$P_\Delta = \frac{1}{3}\sqrt{D^2 + P^2 + 3P_S^2}, \qquad 0 \leq P_\Delta \leq 1. \tag{15}$$

It is important to remark that P_Δ is suitable to represent homogeneous depolarization. However, it does not provide enough information regarding the situations where depolarization actually depends on the state of polarization of the illuminating beam. In this way, an alternative to the depolarization approach by Lu-Chipman [157] was proposed by I. San José and J. J. Gil [168]. They defined the Indices of Polarimetric Purity (IPPs) [168], which are three invariant depolarizing indicators that contain complete and detailed information of the capability of samples to induce randomness to input polarization states. Since the Mueller matrix is not Hermitian - thus we can not ensure its diagonalization -, to define the IPPs it is worth using the covariance matrix $\mathbf{H}$, a positive semi-definite Hermitian matrix associated with the Mueller matrix, $\mathbf{M}$ [169]:

$$H(M) = \frac{1}{4}\sum_{i,j=0}^{3} m_{ij}(\sigma_i \otimes \sigma_j), \tag{16}$$

where m_{ij} $(i,j = 0,...,3)$ are coefficients of $\mathbf{M}$, σ are the Pauli matrices and $\otimes$ the Kronecker product. Accordingly, the IPPs are defined as a set of three real magnitudes, P_1, P_2 and P_3, which are directly obtained from the four eigenvalues λ_i of the covariance

matrix **H** as:

$$P_1 \equiv \frac{\lambda_0 - \lambda_1}{Tr(H)}, \qquad P_2 \equiv \frac{\lambda_0 + \lambda_1 - 2\lambda_2}{Tr(H)}, \qquad P_3 \equiv \frac{\lambda_0 + \lambda_1 + \lambda_2 - 3\lambda_3}{Tr(H)}, \tag{17}$$

where the covariance matrix eigenvalues λ_i are taken in decrease order as $\lambda_0 \geq \lambda_1 \geq \lambda_2 \geq \lambda_3$ and the IPPs are restricted to $0 \leq P_1 \leq P_2 \leq P_3 \leq 1$.

Importantly, the feasible region for the IPPs is defined as the so-called *purity space*, represented in Fig. 5. This purity space is a tridimensional space whose axis correspond to a given IPP (i.e., P_1, P_2 and P_3). The boundaries imposed by the inequalities $0 \leq P_1 \leq P_2 \leq P_3 \leq 1$, lead to any physically realizable depolarizer to be represented within the restricted space of the tetrahedron region in Fig. 5.

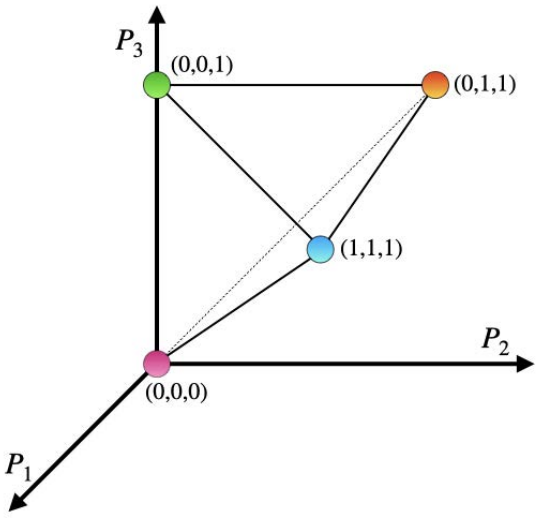


Figure 5: Purity space representing the feasible region for the Indices of Polarimetric Purity, P_1, P_2 and P_3.

In the purity space, the equality $P_1 = P_2 = P_3 = 0$ (see the pink dot in Fig. 5), corresponding to the coordinate $(0, 0, 0)$, describes the ideal depolarizer characterized by the Mueller matrix $M_{ideal} = diag(m_{00}, 0, 0, 0)$. On the contrary, the non-depolarizing (or pure) systems are characterized by $P_1 = P_2 = P_3 = 1$, which corresponds to the coordinate $(1, 1, 1)$ (blue dot in Fig. 5). What is more, the physical interpretation of the purity space, in addition to how much light is depolarized, is related with different depolarizing mechanisms in the sample [48]. Therefore, the depolarization spaces can be potentially used to discriminate among structures with different depolarization signatures due to their properties and structure. Accordingly, the definition of the IPPs lies in the statement that any depolarizer response can be synthesized as an incoherent sum of four non-depolarizing components whose relative statistical weights are

performed by combinations of IPPs. In this way, different types of depolarizers can be distinguished by only looking at the statistical weight of each pure component. In particular, the Mueller matrix of the depolarizer can be decomposed by means of the so-called trivial decomposition as [169, 170]:

$$M = m_{00}[P_1 M(\hat{H}_1) + (P_2 - P_1)M(\hat{H}_2) + (P_3 - P_2)M(\hat{H}_3) + (1 - P_3)M(\hat{H}_4)], \quad (18)$$

where P_1 quantifies the relative portion of pure non-depolarizing component, $M(\hat{H}_1)$, P_2-P_1 is the relative weight of a bi-dimensional depolarizer, $M(\hat{H}_2)$, P_3-P_2 is the relative portion of an equiprobable mixture of three pure components (tridimensional depolarizer, $M(\hat{H}_3)$) and 1-P_3 is associated with an ideal depolarizer, $M(\hat{H}_4)$.

Finally, the already introduced depolarization index, P_Δ, can also be defined in terms of the IPPs as [48]:

$$P_\Delta = \frac{1}{\sqrt{3}}\sqrt{2P_1^2 + \frac{2}{3}P_2^2 + \frac{1}{3}P_3^2}, \quad \text{where} \quad 0 \leq P_\Delta \leq 1. \quad (19)$$

It is necessary to remark that different combinations of IPPs could lead to equal values of P_Δ [168]. Unlike P_Δ, the IPPs are sensitive to the different polarimetric anisotropies (i.e., depolarizing mechanisms) inherent to the probed sample. In this way, the IPPs synthesize the global information provided by P_Δ, thus allowing to transform from 1-dimensional space into a 3-dimensional space of richer polarimetric information [168].

3 Methods

This section is devoted to the description of the complete image Mueller polarimeter used for the experimental Mueller matrix measurements and the samples corresponding to the *ex-vivo* animal tissues and plants used in this study.

3.1 Complete image Mueller polarimeter

The acquisition of experimental Mueller matrices of the different biological samples is performed by means of a complete image Mueller polarimeter based on Parallel Aligned Liquid Crystals (PA-LC) retarders. It consists of two compact and mobile arms: The Polarization State Generator (PSG), which is composed by a linear polarizer oriented at 0^{o}, followed by two PA-LC at 45^{o} and 0^{o}, respectively, with respect to the laboratory vertical, and the Polarization State Analyzer (PSA), whose internal elements are the same as those in PSG but arranged in the inverse order, but additionally placing a CCD camera for the capture of the sample intensity. This PSG-PSA architecture leads to the capability of generating and analyzing, respectively for the PSG and PSA systems, any fully polarized state [171]. The visual representation of the inner components of both arms is shown in Fig. 6.

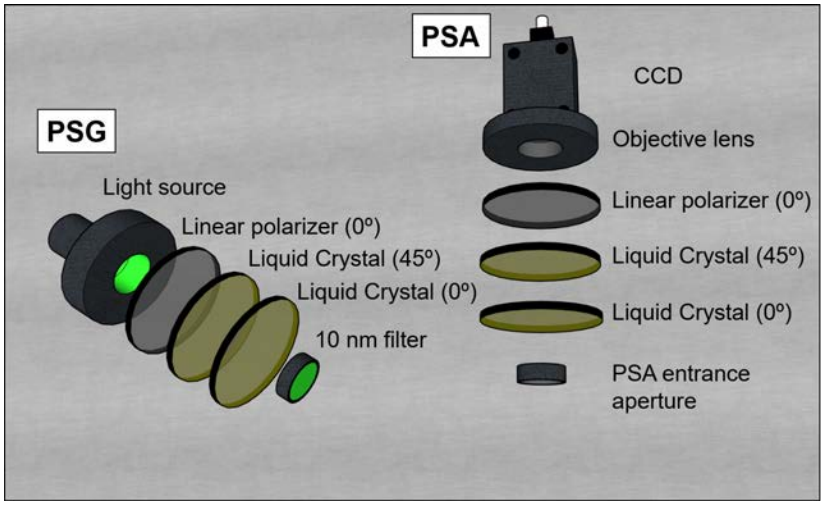


Figure 6: Detailed 3D representation of Polarization State Generator (PSG) and Polarization State Analyzer (PSA) and their constituent optical components. [Adapted from [149]].

We usually conduct two different optical configuration measurements. Scattering measurements 7b) are performed by illuminating the sample placing the PSG at 34º with respect to the laboratory horizontal and the PSA in vertical position (90º) to avoid the ballistic reflection. Transmission configuration (Fig. 7a) is characterized by placing both the PSG and the PSA at 0º with respect to laboratory horizontal.

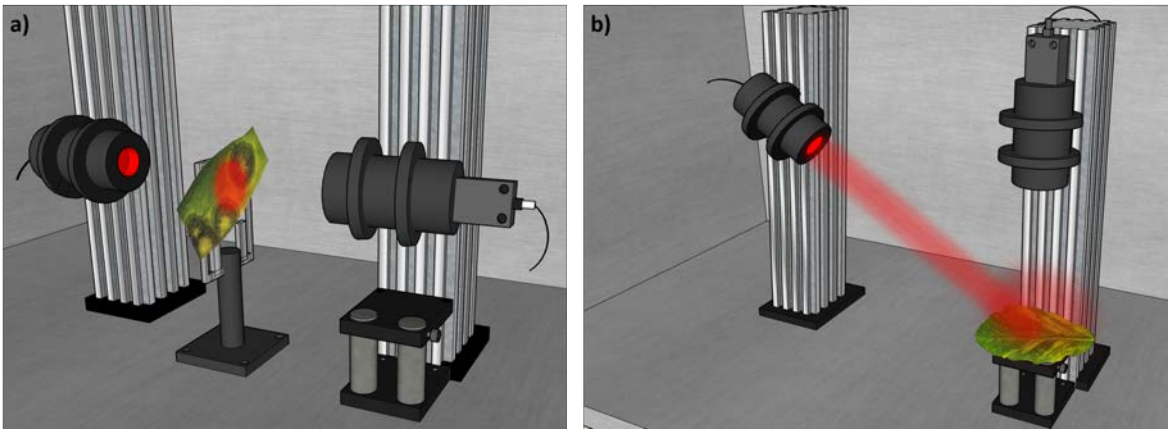


Figure 7: 3D representation of the complete image Mueller polarimeter used in this study at a) transmission configuration and b) scattering configuration. [Reproduced from [172]].

In addition, different illuminating wavelengths are also used, covering the visible range (625 nm, 530 nm and 470 nm), this allowing us to inspect different depths into the sample [173]. To build the experimental Mueller matrix, 36 images of the region of interest (1.1 x 1.1 cm^2) are taken in order to minimize the measurement noise (at least 16 images are needed): we use the 6 illumination (generators) states of polarization and

the 6 analyzers proposed in Ref. [171]. Complete Mueller matrix measurement lasts approximately 3.6 seconds, considering the time of calculation of the Mueller matrix by using the inverse method (0.4 to 2.2 s) [49], PA-LC molecules orientation process from one polarization state to the following one ($\approx$ 100 ms) and CCD exposure time.

Particularly, the illumination is performed by means of a four-wavelength high-power Thorlabs LED source (LED4D211, operated by DC4104 drivers distributed by Thorlabs) complemented with 10 nm dielectric bandwidth filters distributed by Thorlabs: FB530-10 and FB470-10 for green and blue wavelengths, respectively. About the linear polarizers: the one arranged on PSG is a Glan-Thompson prism-based CASIX meanwhile the one in PSA is a dichroic sheet polarizer distributed by Meadowlark Optics. The four Parallel Aligned Liquid Crystals are Variable Retarders with Temperature Control (LVR-200-400-700-1LTSC distributed by Meadowlark Optics). Imaging is performed by means of a 35mm focal length Edmund Optics TECHSPEC ®high-resolution objective followed by an Allied Vision Manta G-504B CCD camera, with 5 Megapixel GigE Vision and Sony ICX655 CCD sensor, 2452(H) x 2056(V) resolution and cell size of 3.45 μm x 3.45 μm, so a spatial resolution of 22 μm is achieved.

3.2 Sample description

To conduct the polarimetric analysis of *ex-vivo* animal tissues in *Paper C* [174] and *Paper D* [175], we inspected a collection of 33 different chicken thighs which were dissected and split into soft tissues under the supervision of a pathologist. In particular, from each thigh, we dissected the bone and two sample units per type of the soft tissues standing for skeletal muscle, tendon and myotendinous junction (Fig. 8). The aforementioned soft tissues present different physiological function and structural properties: the tendon is composed by dense type-I collagen fibers (up to 60 - 80%) disposed, by following the same orientation as muscle bundles, in parallel and compact fascicles [176, 177]. Skeletal muscle comprises bundled fascicles of contractile myofibril chains (in counterpart to tendon, whose non-contractile properties allow it to withstand tensions) sheathed with type-I collagen rich tissues [178] which protect muscle fibers from friction. Myotendinous junction is a combination of the previously mentioned muscle and tendon tissues: fiber fascicles of contractile myofibrils and collagen are progressively mixed and covered by fasciae [179–181].

For completeness, a histological analysis from the same tissue regions was conducted. The histological sections of muscle, tendon, bone and myotendinous junction were collected and stained using a Masson Trichromic technique. This allowed a clear-cut differentiation of all connective tissue (and otherwise mesodermal as bone and muscle). The resulting histological images are presented in Figs. 9 and 10.

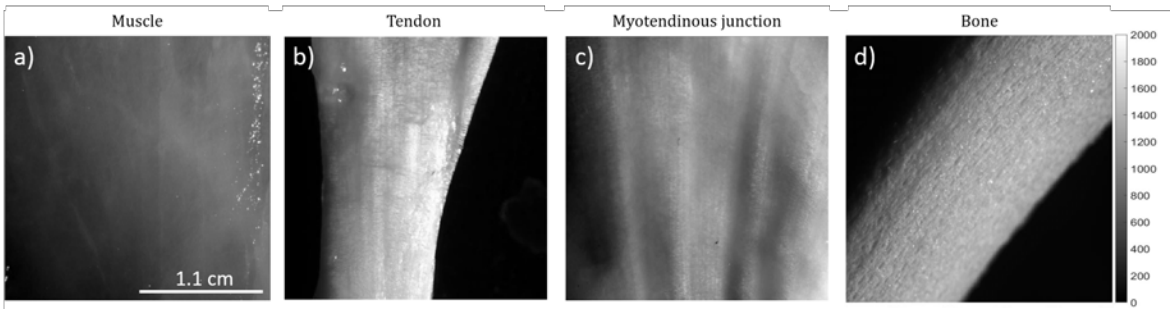


Figure 8: Intensity image measured at 625 nm of a) muscle, b) tendon, c) myotendinous junction and d) bone tissues of a given chicken thigh. Images correspond to an area of 2.2 × 2.2 cm^2.

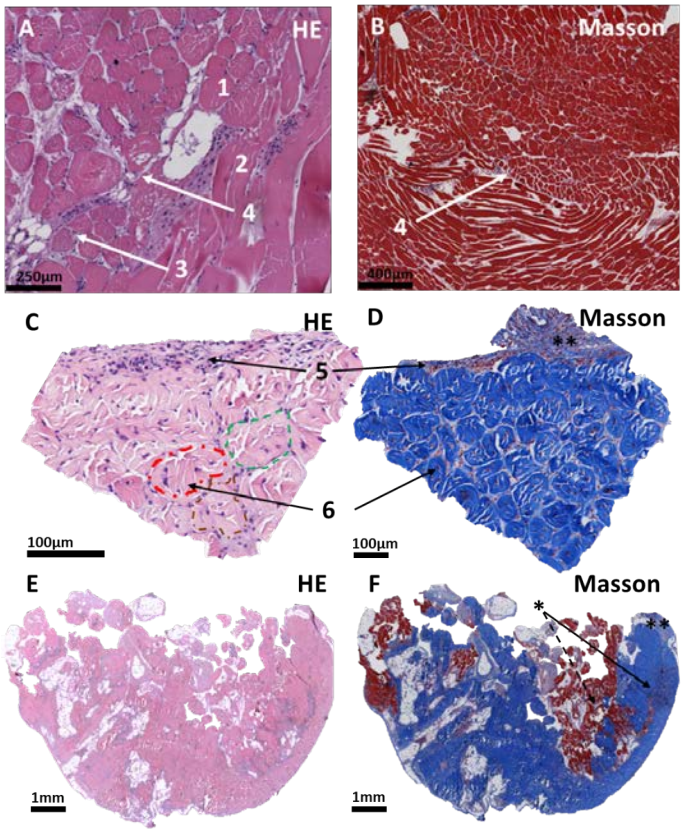


Figure 9: Histochemical stainings of muscle (A, B), tendon (C, D) and myotendinous junction (E, F). Bundles (1, 2, 6) of either contractile (A, B, E, F) or dense fibrous (collagen) tissue (C-F) are surrounded by sheets of lax connective tissue (collagen) concentrically organized as epimysium (**)/epitenon (**), perimysium/peritenon (5) and endomysium/endotenon (6). In E, F a transverse section of the myotendinous junction is shown, where the Masson staining reveals the intermixed (*) fascicles of contractile fibers (discontinuous arrow in (F) and collagenous fibers (arrow in F)). Different fascicles of tendinous collagen are highlighted in red, green and gold (C). [Reproduced from [174]].

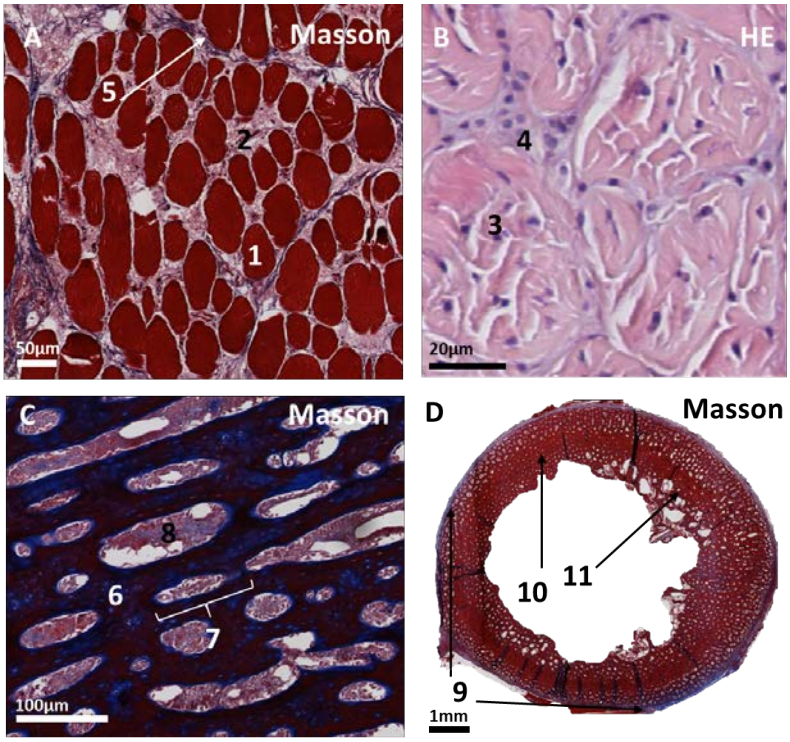


Figure 10: Transversal section of striated muscle (A) showing bundles of contractile fibers (red colored, 1) surrounded by endomysium (2, compare to 3, 4) and included in a single fascicle surrounded by collagen (perimysium, 5). This structure is analogous to the tendinous structure (B), where the collagen fibers (3) are densely packed into a sheet of peritenon (4). Longitudinal section of cortical bone (C), showing bony matrix (6) and lacunae (7) containing osteocytes, vessels and collagen that are heavily artifacted due to pre-treatment (8); Transverse section through the diaphysis of a long bone (D), where collagenic periosteum (9), cortical bone (10) and rests of trabecular bone (11) are arranged in a concentrically layers. Note the regular arrangement of lacunae, both in longitudinal and transverse sections. Color (from reddish to purple) depends on the amount of mineral deposits in any given region of the bone. [Reproduced from [174]].

For the study conducted in *Paper B* [182], we used sections of a trachea, tongue and heart dissected from a lamb (Figs. 11a, 11b and 11c, respectively). Particularly, the trachea composition is based on hyaline cartilages rings (collagen type-II and chondroitin sulphate rich) covered and joined by annular ligament (collagen type-I and fibroblast cells rich fibrous membrane). The undissected tongue is compsoed by skeletal muscle. As a protective surface, the tongue has stratified squamous epithelium, heavily keratinised with long papillae. Heart sample corresponds to an endocardial view, particularly on endocardium-covered muscle and sub-valvular apparatus on connective tissue.

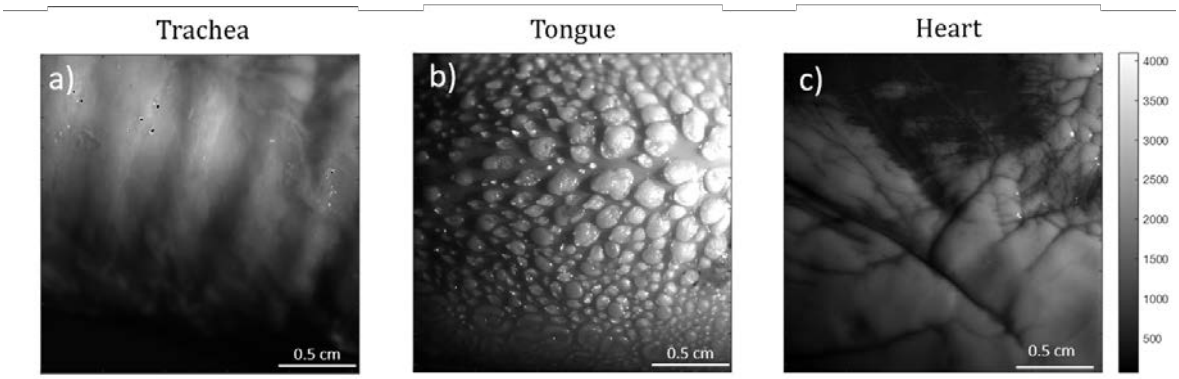


Figure 11: Intensity image measured at 470 nm of a) trachea, b) tongue and c) heart of a lamb. Images correspond to an area of 2.2 × 2.2 cm^2.

The plant samples inspected in *Paper A* and *Paper B* [172, 182] correspond to a set of 73 leaves from a collection of 18 different plant specimens. To not to extent the content of the manuscripts, we restricted the presented results to the most relevant plant specimens. We analyzed (1) a leaf of *Medicago sativa* specimen infected with the alfalfa mosaic virus (AMV); (2) a leaf of*Olea europaea* specimen infected with fungus Venturia oleaginea (causal agent of the olive leaf spot); (3) a leaf of *Prunus dulcis* specimen showing disease symptoms of infection from fungus Transzchelia discolor, (4) a leaf of *Quercus pubescens* specimen infected with fungus E. alphitoides, which causes powdery mildew lesions on leaf surface, and (5) a leaf of *Vitis vinifera* specimen showing no symptoms of disease.

The *M. sativa* specimen (Figs. 12a and 12b) is native from warmer-temperate climates of south-central Asia. It belongs to the Fabaceae family (i.e., legumes) and it is cultivated worldwide for livestock feeding purposes. Despite the toxicity of unsprouted alfalfa, it is also suitable for human consumption in sprout stage or dehydrated. The main characteristics of the infection caused by AMV are the emergence of wilting or white flecks to necrotic wounds and chlorotic mosaics on leaves. Regarding *O. europaea* (Figs. 12c and 12d), this species belongs to the Oleaceae family. Although the native species were found in eastern land regions around the Mediterranean Sea, his production is not limited to Mediterranean countries: *O. europaea* is cultivated in several countries such as South Africa, New Zealand, North and South America, and Australia. In addition to olive oil production and fruit consumption (olives), *O. europaea* trees are also grown for fine wood manufacturing. Regarding the olive leaf spot, this worldwide spread disease of olive may cause severe tree defoliation and a delay in fruit ripening, among other symptoms, thus leading to relevant yield losses.

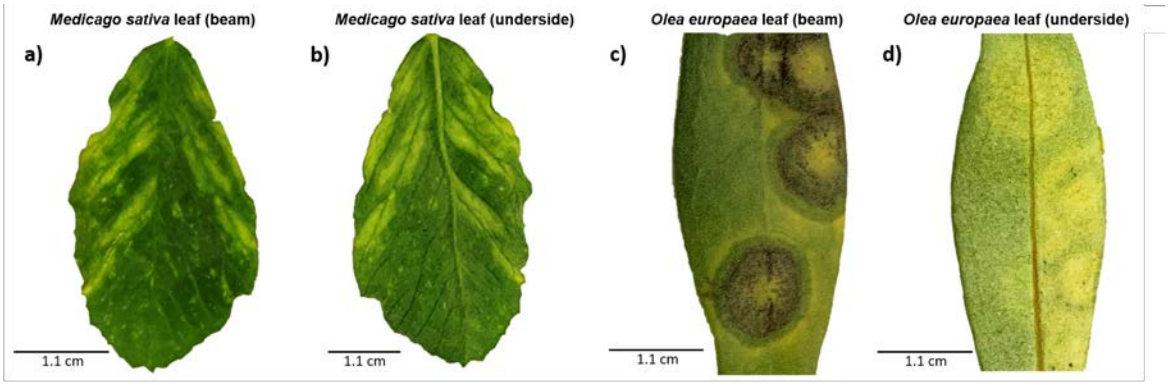


Figure 12: Picture of the *Medicago sativa* leaf a) viewed from the beam and b) underside, and picture of the *Olea europaea* leaf a) viewed from the beam and b) underside.

The *P. dulcis* (Figs. 13a and b), commonly known as almond tree, is a species native from Iran that belongs to the Rosaceae family. It is common from Mediterranean climate regions and produces almond fruits which are widely used for nutritional and cosmetic purposes. The *Q. pubescens* (Figs. 13c and 13d), a species of white oak, belongs to the Fagaceae family and it is commonly found in central and southern Europe. It produces acorns (oak nuts) which can be consumed or extract their oil.

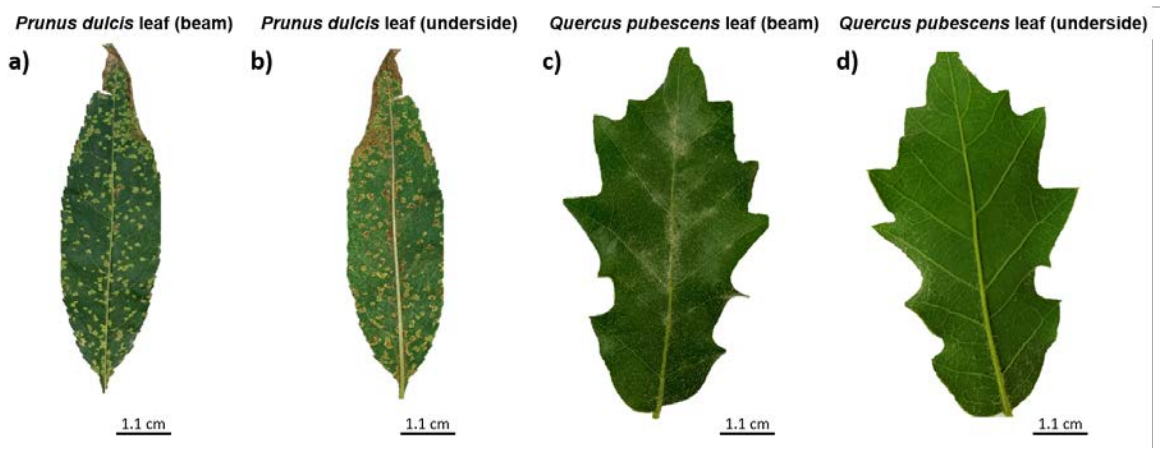


Figure 13: Picture of the *Prunus dulcis* leaf a) viewed from the beam and b) underside, and picture of the *Quercus pubescens* leaf a) viewed from the beam and b) underside.

The *V. vinifera* (Fig. 14), commonly known as grape vine, belongs to the Vitaceae family. Native from the central Europe, the land regions around the Mediterranean Sea and southwestern Asia, *V. vinifera* is cultivated worldwide for both grape (fresh or dried) consuming, and vinegar and wine production.

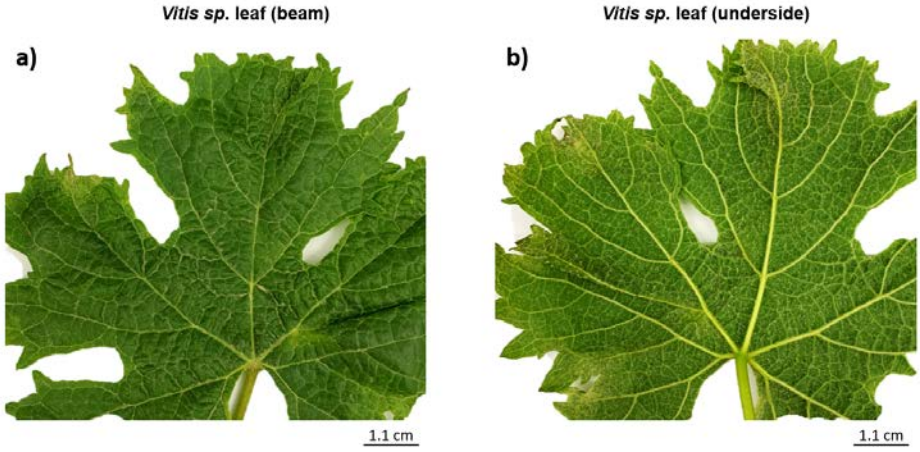


Figure 14: Picture of the *Vitis vinifera* leaf a) viewed from the beam and b) underside.

Dr. T. Garnatje and Dr. J. Luque undertook the formal identification of the plant material used in these studies. An herbarium voucher of *M. sativa*, *O. europaea* and *Q. pubescens* are deposited in the Herbarium of the Botanical Institute of Barcelona (BC-983007, BC-983006 and BC-983018 respectively).

4 Results and discussion

In this thesis we explore the use of depolarization-related polarimetric observables in joint with the commonly-used set of enpolarizing metrics for biophotonics applications. Through four different research works, henceforth labeled as *Paper A* [172], *Paper B* [182], *Paper C* [174] and *Paper D* [175], we inspect the polarimetric response of a wide variety of biological tissues and address the extracted information in different ways, either related to specific imaging techniques (*Paper A* and *Paper B*) or classification purposes (*Paper C* and *Paper D*). In the following, the results and discussions related to each work are presented.

4.1 Paper A – *Polarimetric observables for the enhanced visualization of plant diseases*

As demonstrated in the literature [89, 149, 160, 183], the inspection of depolarization-related observables allow to enhance the image contrast between structures contained in biological tissues or even to reveal patterns or organelles which were hidden in regular intensity images. Despite of its potential, the analysis of depolarization is underused in plant science. In this work (*Paper A* [172]) we report the use of depolarization observables applied for the first time on the visualization and characterization of plant diseases. We tested the polarimetric methods in a set of 73 leaves corresponding to 18

plant diseased species. However, we bounded the study in two representative samples due to their impact and utility in humans: (1) a leaf of *Medicago sativa* showing the particular disease symptoms due to the infection of alfalfa mosaic virus (AMV) and (2) a leaf of *Olea europaea* presenting the characteristic lesions due to the infection of the fungus *Venturia oleaginea*. In particular, *M. sativa* showed chlorotic areas surrounding the affected vascular structure. In turn, *O. europaea* showed alternating necrotic and chlorotic ring-like lesions surrounding a chlorotic spot. A picture of the *M. sativa* and the *O. europaea* leaves is shown in Fig. 12 of the previous section 3.2. The obtained results clearly show the suitability of polarimetric methods for the recognition of these two pathologies in the *M. sativa* and *O. europaea* specimens. Importantly, the procedure was validated in the remaining set of 16 different diseased species used in this study and all together highlight the potential of the polarimetric methods. The complete list of studied plants and diseases can be consulted in Ref. [172].

We first provide the complete polarimetric analysis of the leaves of *M. sativa* and *O. europaea*. Particularly, we measure their experimental Mueller matrices (MMs) by means of the complete image Mueller polarimeter described in the previous section 3.1. From the acquired MMs, we calculate a set of different depolarization-related observables corresponding to the Indices of Polarimetric Purity (IPPs: P_1, P_2 and P_3), the Components of Purity (CPs: P, D and P_S) and the depolarization index, P_Δ. The standard non-polarized image and the obtained polarimetric images corresponding to *M. sativa* and *O. europaea* are shown in Figs. 15 and 16, respectively. These polarimetric images were obtained from certain regions of interest (ROIs) in *M. sativa* and *O. europaea*, which are indicated with white and black squares in Figs. 15a and 16a, respectively. Regarding the *M. sativa* sample, the depolarization metrics of P_1, P_2, P_3, P_Δ and P_S (Figs. 15c-f and 15i, respectively), clearly manifest an overall enhancement of image contrast and help to unveil and better spatially locate the wounded zones and vascular structures within the leaf, which are invisible by using standard non-polarized images. In particular, up to seven chlorotic spots as well as some vascular structures which are barely observable by standard intensity images, are well visible by using polarimetric means in *M. sativa* (see the yellow arrows and the numeric labels in Figs. 15c and 15d, respectively). Similarly, the *O. europaea* sample presents a necrotic ring surrounding a chlorotic spot which is hardly visible in non-polarized scattering images (16b). In contrast, the depolarization-related observables of P_1, P_2, P_3, P_Δ and P_S (Figs. 16c-f and 16i, respectively) unveil the presence of vascular structures (see the yellow arrows in Fig. 16d) and provide a more accurate visualization and spatial delimitation of the lesions in the leaf: the visual contrast of the wounded region is increased when compared to that of non-polarized image (Fig. 16b).

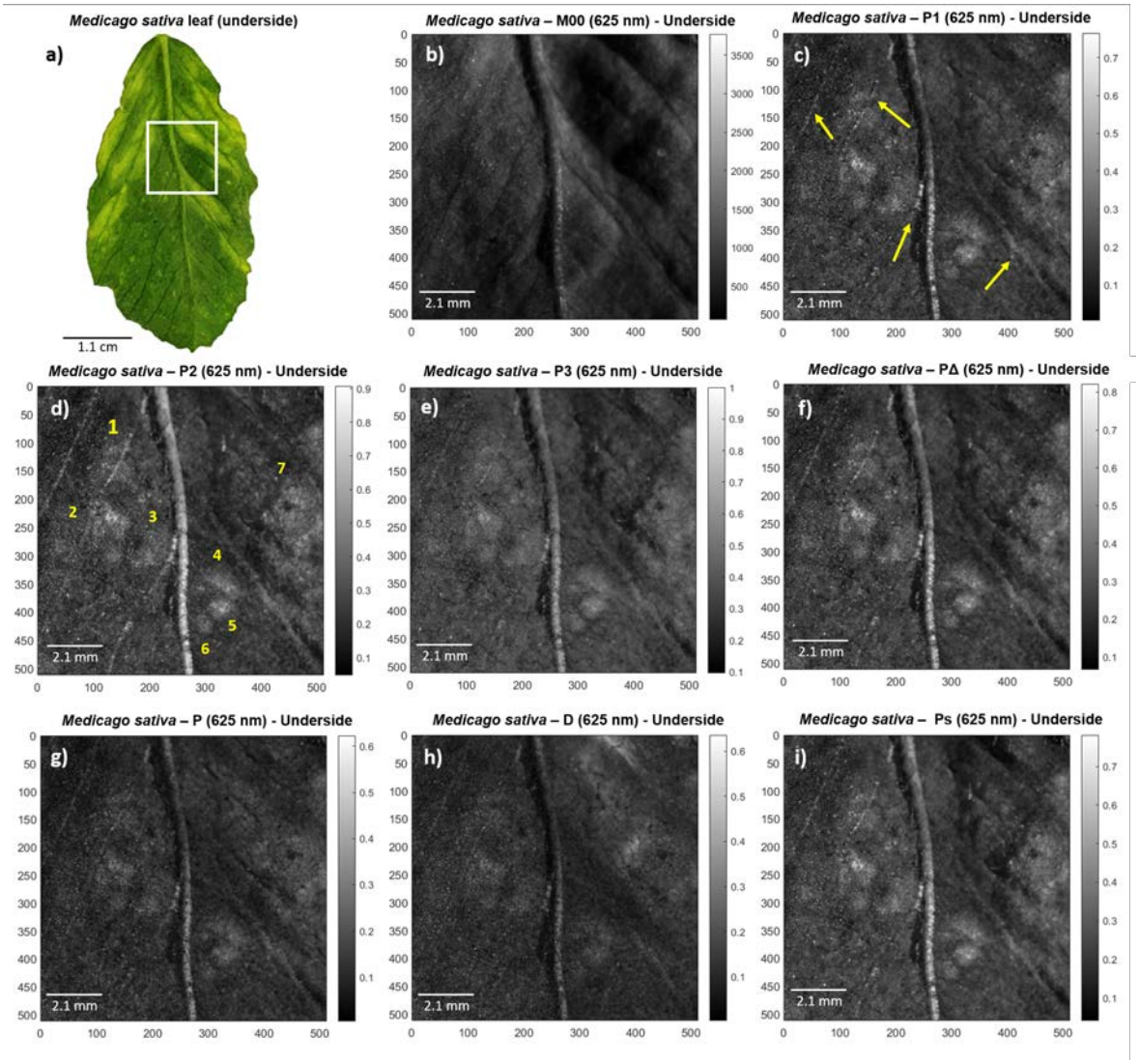


Figure 15: Polarimetric images of *Medicago sativa* leaf used in this study. a) Picture of the underside part of the *M. sativa* leaf. White square denotes for selected region of interest (ROI) analyzed in remaining images, b) regular intensity image (M00) of the *M. sativa* underside ROI and its corresponding polarimetric observables c) P_1, d) P_2, e) P_3, f) P_Δ, g) P, h) D and i) P_S for visual comparison. All images correspond to 625 nm illumination wavelength measurements performed at scattering setup configuration. Yellow arrows correspond to the enhanced vascular structures within the sample, whereas numeric labels (from 1 to 7) indicate the number of chlorotic spots unraveled by means of polarimetric observables. [Reproduced from [172]].

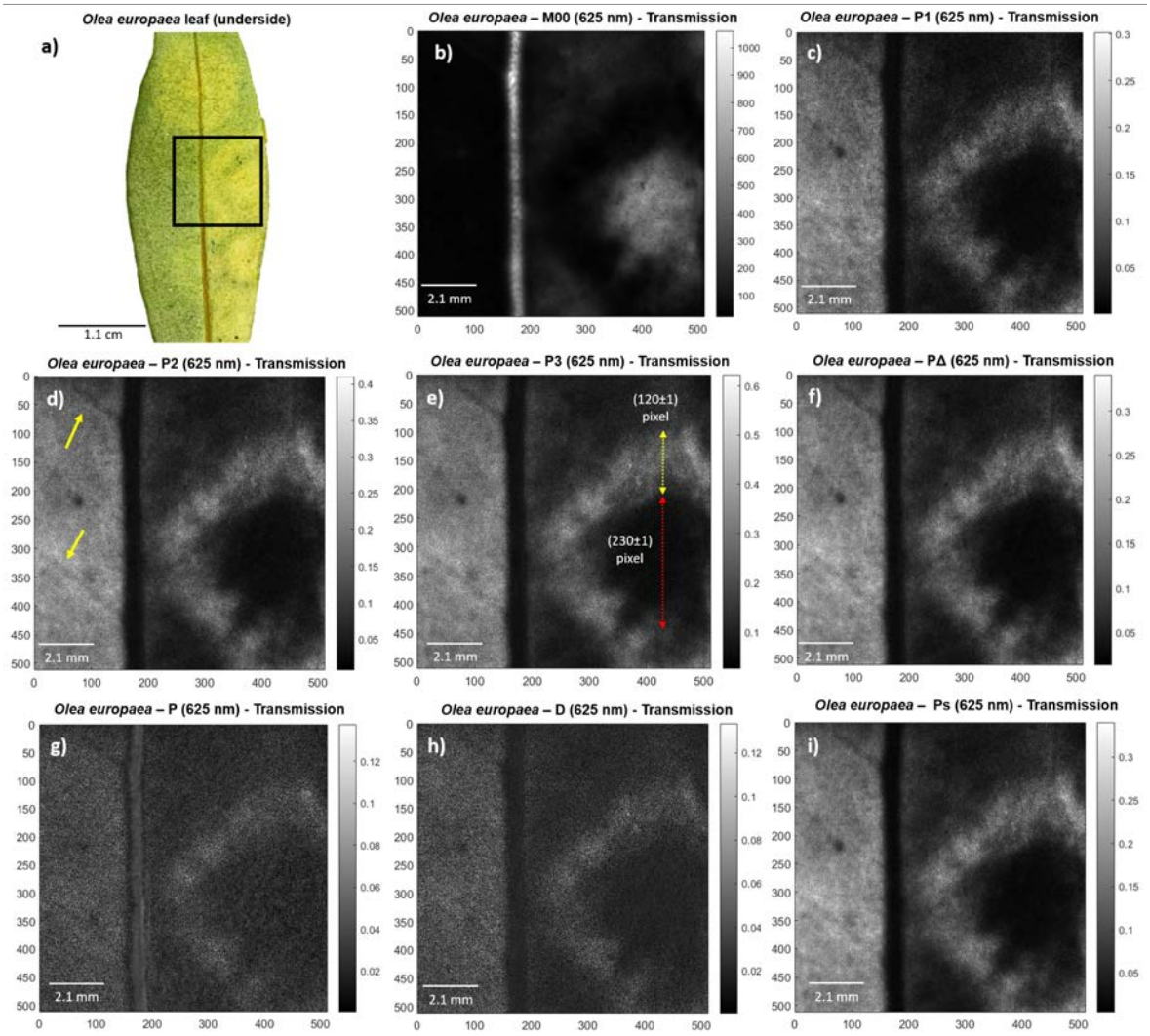


Figure 16: Polarimetric images of Olea europaea leaf used in this study. a) Picture of the underside part of *O. europaea* leaf. Black square denotes for selected region of interest (ROI) analyzed in remaining images, b) regular intensity image (M00) of the *O. europaea* transmission ROI and its corresponding polarimetric observables c) P_1, d) P_2, e) P_3, f) P_Δ, g) P, h) D and i) P_S for visual comparison. All images correspond to 625 nm illumination wavelength measurements performed at transmission set-up configuration. Red and yellow dotted lines correspond to diameter and width measurements for chlorotic spot and necrotic ring, respectively. Yellow arrows indicate the unveiled vascular structures. [Reproduced from [172]].

Furthermore, the image enhancement related to depolarizing observables for both studied plant specimens is quantified by means of various analysis. We first compare the values of the depolarizing metrics versus the regular intensity of the cross-section along tissue transitions corresponding to the healthy-chlorotic and the healthy-chlorotic-necrotic on *M. sativa* and *O. europaea*, respectively. The particular cross-sections analyzed, highlighted in yellow lines, together with the corresponding values are shown in Fig. 17. For the case of *M. sativa*, the difference between the chlorotic and the healthy areas is hardly visible when inspecting the regular intensity (M_{00}, Fig. 17a). However, this difference becomes quite apparent when inspecting the image corresponding to the index of polarimetric purity P_2 (Fig. 17). In such case, healthy areas show P_2 values around 0.20 ± 0.01 while chlorotic areas show characteristic values around 0.50 ± 0.01. Analogously, the regular intensity image (M_{00}, Fig. 17c) of *O. europaea* sample is not sensitive to the necrotic ring border. Conversely, the index of polarimetric purity P_3 shows clear differences between healthy, necrotic and chlorotic tissue regions (Fig. 17d). For instance, the value of P_3 for healthy, necrotic and chlorotic areas is around 0.20 ± 0.01, 0.45 ± 0.01 and 0.10 ± 0.01, respectively. In contrast, regular intensity (M_{00} in Fig. 17c) shows the chlorotic area as a bright spot with values close to 0.50 ± 0.01, the necrotic ring in dark with typical values of 0.05 ± 0.01, and the healthy area in dark grey with typical values of 0.90 ± 0.01. Accordingly, the difference between healthy and necrotic areas is less net in the intensity image than in the P_3 one.

We further demonstrate this phenomena by computing the visibility, $V = [I_{max} - I_{min}]/[I_{max} + I_{min}]$, where I_{max} and I_{min} correspond to the maximum and minimum value of the pixel) of the cross-section along healthy-diseased tissue transitions. The visibility features a value of $V = (0.41 \pm 0.05)$ for the non-polarized intensity and of $V = (0.61 \pm 0.03)$ for the P_2 channel in *M. sativa*. Similarly, we measure the contrast improvement in polarimetric images compared to the regular intensity. As a representative example, the visibility of the necrotic ring is estimated to $V = (0.84 \pm 0.03)$ and $V = (0.00 \pm 0.03)$ for the index of polarimetric purity P_3 and the regular intensity M_{00}, respectively.

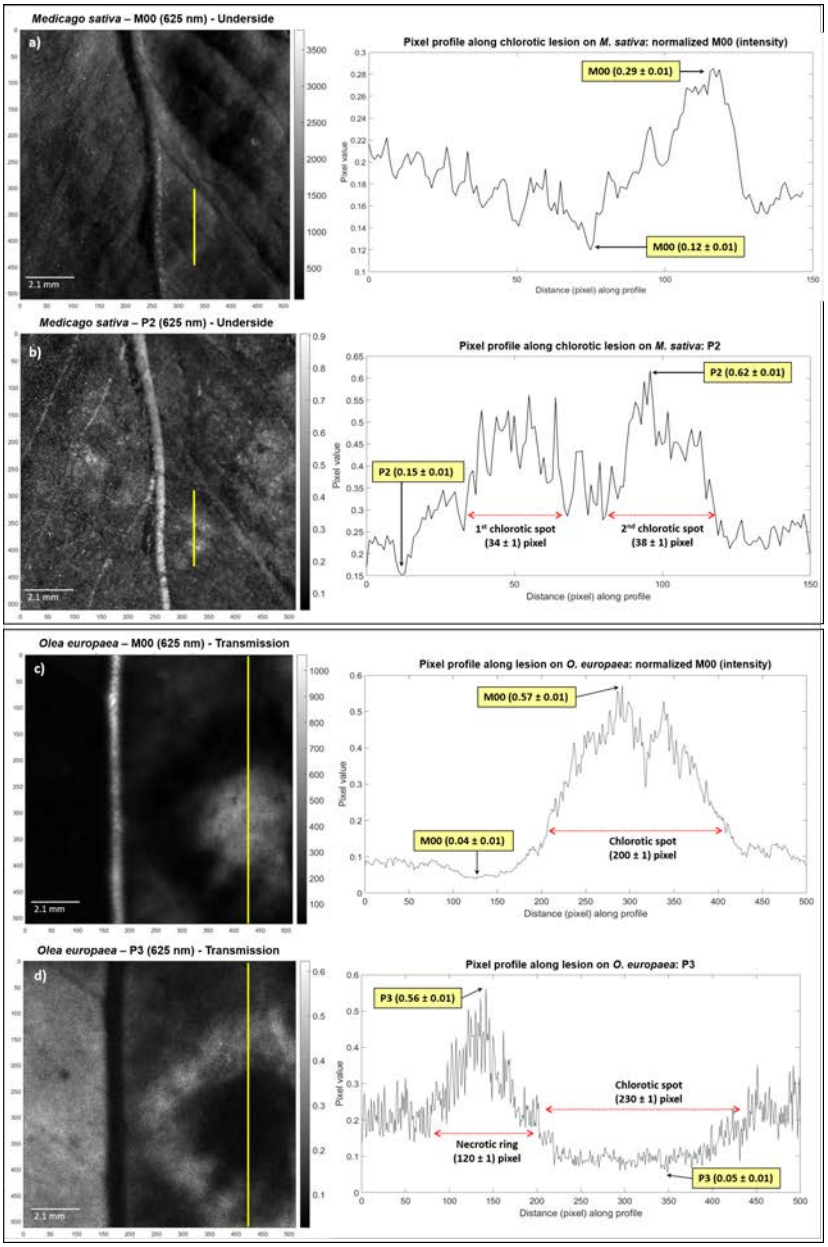


Figure 17: Pixel profile comparison for *Medicago sativa* and *Olea europaea*. a) Intensity image at 625 nm of the underside part of *M. sativa* sample and its corresponding pixel profile, b) polarimetric purity index P_2 and its corresponding pixel profile, c) intensity image at 625 nm of the underside part of *O. europaea* sample and the corresponding healthy-necrotic-chlorotic transition pixel profile, and d) polarimetric purity index P_3 and its corresponding pixel profile. The vertical yellow lines on polarimetric images indicate the plotted pixel profile segments. The squared numeric labels for M00, P_2 and P_3 indicate their respective highest and lowest pixel values within the inspected pixel regions. Red-dotted horizontal lines on plots indicate the diameter of the two chlorotic spots of *M. sativa* and width measurements for chlorotic spot and necrotic ring of *O. europaea*. [Reproduced from [172]].

Further analysis of the potential of those observables to discriminate between different typologies (healthy or diseased tissues) of plant structures is conducted. In particular, we select the regions of interest corresponding to the healthy, chlorotic and necrotic tissue locations within the samples and represent the measured data in two different polarimetric spaces built from the IPPs and the CPs metrics (the so-called Purity space and the Components of Purity space, respectively). This leads us to a very intuitive visualization of data, and also providing quantitative information of the structures (or tissue types) that may be present in the images of the probed samples. In a data cloud figure, pixels corresponding to different regions should group in separate clouds. Therefore, non-overlapping clouds indicate that the related regions are well discriminated. Conversely, either fully or partially overlapping clouds indicate a poor discrimination of nominally different zones. Figure 18 presents the data clouds from selected ROIs (marked in Figs. 18a and 18d as colored rectangles) of healthy (green), chlorotic (yellow) and necrotic (dark blue) tissue regions of *M. sativa* (first row in Fig. 18) and *O. europaea* (second row in Fig. 18) leaves.

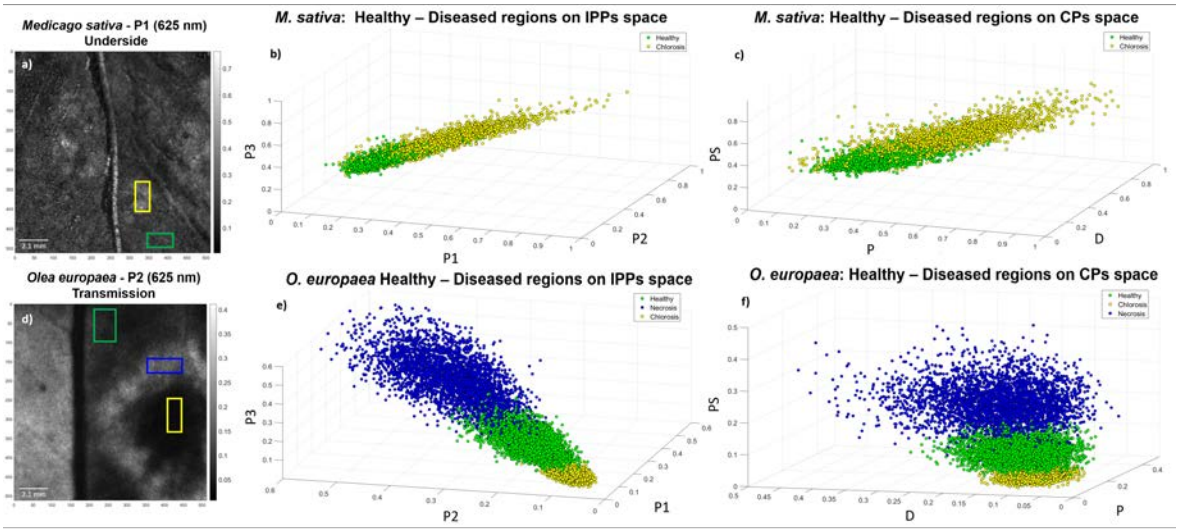


Figure 18: Scatter data plots of healthy and diseased tissue regions of *Medicago sativa* and *Olea europaea*. a) Visual indicative for healthy (green) and chlorotic (yellow) tissue selected region of interest (ROI) for *M. sativa*, b) corresponding IPPs space (P_1, P_2, P_3) for healthy and chlorotic data clouds representation, c) components of purity (P, D, P_S) space for healthy and chlorotic data clouds representation, d) visual indicative for selected healthy (green), chlorotic (yellow) and necrotic (dark blue) tissue ROIs for *O. europaea*, e) corresponding IPPs space (P_1, P_2, P_3) for healthy, chlorotic and necrotic data clouds representation and f) components of purity (P, D, P_S) space for healthy, chlorotic and necrotic data clouds representation.[Reproduced from [172]].

Interestingly, from Fig. 18 we conclude that the polarimetric data from different tissue conditions (i.e. healthy, chlorotic and necrotic) is clearly located in different

spatial positions within the spaces, showing no data mixing, which is directly related with the high discriminatory potential of the depolarizing metrics when differentiating the studied tissue characteristics. For the particular case of *M. sativa*, healthy tissue data points tend to group close to higher depolarization, while chlorotic regions tend to group to areas related with less depolarization (green and yellow squares in Fig. 18a, respectively). This behavior occurs when the data is represented in either the IPPs or the CPs space (Figs. 18b and 18c, respectively). Therefore, these two tissue conditions are well-discriminated, as practically no data mixing between tissues with different health condition is produced. Importantly, a stronger depolarization response occurs when the leaves contain non-organized spatial structures or an important number of microstructures which efficiently scatter light. In this context, either the biological or the structural changes caused by the chlorotic symptoms of infected regions are translated into a less depolarizing effect on incident light when compared with healthy tissues, thus increasing the sensitivity of depolarizing channels to chlorosis detection. Similarly, data clouds corresponding to healthy, chlorotic, and necrotic tissue regions in *O. europaea* (green, yellow and blue squares, respectively, in Fig. 18d) are clearly spatially separated in both IPPs and CPs spaces (Figs. 18e and 18f, respectively). In this case, however, the *O. europaea* sample demonstrates that chlorotic tissue is more depolarizing (i.e., yellow dots are quite close to the coordinate (0,0,0) in Figs. 18e and 18f) than healthy tissues. Accordingly, the inspection of the different disease symptomatology suggest that the different leaf tissue structure, in addition to the type of pathogen which caused the chlorotic symptoms on both inspected leaves, may also play a role in this differential depolarizing response.

One particular interest is to include the discriminatory potential related to different polarimetric signatures in a single image. Recent works [149, 183, 184] suggest the definition of a pseudo-colored image whose layers contain the different polarimetric features of the sample. In other words, different polarimetric information origins are included, all together, in a single and enhanced image. The main idea in these works is to design a pseudo-colored function based on the weighted combination of three different polarimetric observables showing the highest image contrast between the structures of interest within the sample and associate each chosen observable to a primary color (i.e., red, green and blue). So far, the methods proposed in previous studies to build polarimetric based pseudo-colored images are based on simple linear combinations of polarimetric observables whose weights are set *ad-hoc*, and thus, far from optimal approaches. In this work we propose two general pseudo-colored functions which are based on the statistical analysis of the polarimetric data retrieved from the different tissue conditions of the inspected samples. The novelty of this approach is that the chosen depolarizing observables are not restricted to the IPPs space, but

extended to an optimized selection of polarimetric observables within the IPPs and CPs spaces. The key advantage of this approach is that we select the polarimetric observables showing the highest discriminatory potential between the healthy regions and the different lesions and combine them in such a manner that the output results in a colored image. To estimate the differences between these observables when applied to image healthy and wounded areas, we conducted a Boxplot analysis [185] for the regions of interest highlighted in in Figs. 18a and 18d. As a representative example, Fig. 19 shows the resulting analysis of the IPPs and CPs for the *M. sativa* sample.

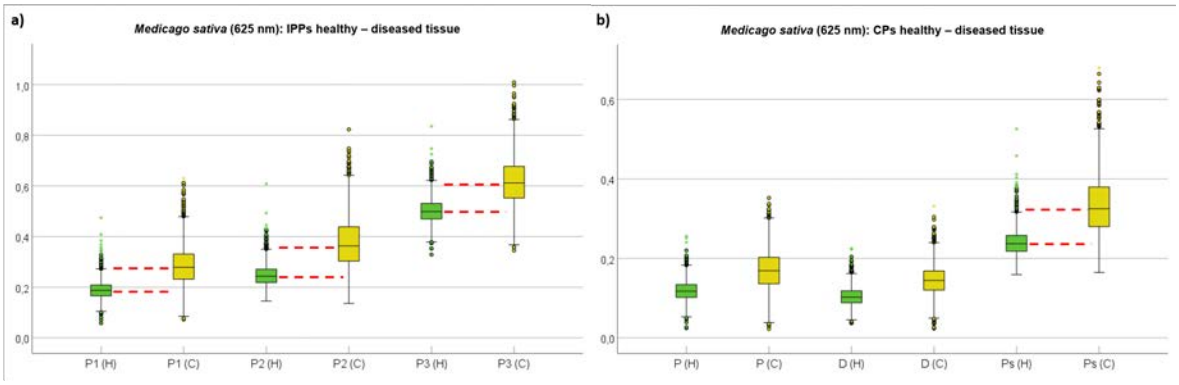


Figure 19: Boxplot charts for healthy and diseased regions for *Medicago sativa*. a) Indices of polarimetric purity (P_1, P_2, P_3) boxplot for healthy and chlorotic locations on *M. sativa*, b) Components of purity (P, D, P_S) Boxplot for healthy and chlorotic locations on *M. sativa*. The corresponding healthy and chlorotic data distributions are labeled and colored as H (green) and C (yellow), respectively. Red-dashed lines indicate the locations of median values and illustrate they do not fit within the boxes of different tissue conditions (healthy or chlorotic), allowing discrimination. Circles and stars correspond to mild and extreme outlier values, respectively. [Adapted from [172]].

Importantly, to evaluate such differences, the median values in Fig. 19 are highlighted with dotted red lines. These indicate the case of a given polarimetric observable that demonstrates the largest difference between both tissue conditions (healthy / chlorotic). In the case of the *M. sativa*, the healthy and chlorotic tissue regions are clearly separated by the whole set of IPPs as well as by the sphericity degree, P_S (Figs. 19a and 19b, respectively). Similarly, in the case of the *O. europaea* which presents chlorotic, necrotic and healthy tissues, the same polarimetric observables (i.e., the whole set of IPPs and the sphericity degree P_S) are the metrics that lead to the larger distances between different tissue conditions. Accordingly, note that the different tissue conditions are well differentiated when represented within the boxplot charts (Fig. 19). Taking advantage of this situation, we can set a numeric threshold able to discriminate between each pair of tissue conditions in a binary way. To do so, we computed for each polarimetric observable the distance between the median values corresponding to different colored boxes for the *M. sativa* and the *O. europaea* samples.

The three largest median differences between chlorotic and healthy tissue regions on the *M. sativa* leaf were demonstrated by P_2, P_3 and P_1, with values of 0.119 ± 0.104, 0.113 ± 0.106 and 0.091 ± 0.085, respectively. In the case of *O. europaea*, P_3, P_2 and P_S demonstrate the largest median differences between the chlorotic and healthy tissue regions, with values of 0.064 ± 0.029, 0.039 ± 0.020 and 0.037 ± 0.015, respectively. Similarly, this polarimetric triplet demonstrates the largest values for the case of necrotic and healthy tissues comparison, with 0.117 ± 0.059, 0.080 ± 0.041 and 0.069 ± 0.032 for P_3, P_2 and P_S, respectively.

Therefore, by considering the largest differences between medians, we selected two triplets of polarimetric observables for the pseudo-colored images construction: the first one consists of the mix of observables from the IPPs and CPs spaces (P_2, P_3 and P_S) and the other is based on the IPPs by themselves (P_1, P_2 and P_3). Importantly, from these median differences, we set a given threshold value for each selected polarimetric observable which allows us to numerically discriminate between different tissue conditions: healthy/chlorotic for *M. sativa* and healthy/chlorotic/necrotic, for *O. europaea*. Afterwards, each separated condition is assigned to a primary color (red, green and blue, for chlorotic, healthy, and necrotic tissues, respectively). As a representative example, for the case of the *M. sativa* leaf sample, given a particular polarimetric observable (e.g., P_2), a first image carrying the chlorotic information of the plant is binarized with black and red colors according to the pixel values above or below this certain threshold. Then, the image containing the healthy information of the plant is similarly binarized this time with black and green color conditions. Accordingly, a given pseudo-colored function is constructed for each polarimetric observable as follows:

$$P_{i,approach}(x,y) = P_{i,R-Chlorotic}(x,y) + P_{i,G-Healthy}(x,y) + P_{i,B-Necrotic}(x,y), \quad (20)$$

where the sub-index i denotes for the particular depolarizing observable ($i = 1, 2, 3, S$), R, G and B denote for red, green and blue, respectively; and each term in corresponds to a binary-colored image. Importantly, for the case of the *M. sativa* we can consider the term $P_{i,Blue-Necrotic}(x,y) = 0$ because the sample has no necrotic content. Finally, we consider the content of the whole triplet of polarimetric observables of [P_1, P_2, P_3] and [P_2, P_3, P_S]. This information is put together by constructing a linear combination of the pseudo-colored observables implemented as defined in Eq. 20. The resulting pseudo-colored functions valid for the *M. sativa* and the *O. europaea* are defined as follows:

$$\begin{aligned} \text{Pseudo\#1}(x,y) &= P_{2,approach}(x,y) + P_{3,approach}(x,y) + P_{S,approach}(x,y)], \\ \text{Pseudo\#2}(x,y) &= P_{1,approach}(x,y) + P_{2,approach}(x,y) + P_{3,approach}(x,y)]. \end{aligned} \quad (21)$$

The visual representation of the effect of pseudo-colored approaches in Eq. 21 is shown in Fig. 20. In particular, for comparative purposes, Fig. 20 shows the images of

the classical non-polarizing intensity (M_{00}, Figs. 20a and 20d) and the representative polarimetric observables of P_2 (Fig. 20b) and P_1 (Fig. 20e) and the corresponding pseudo-colored image for *M. sativa* and *O. europaea*, respectively.

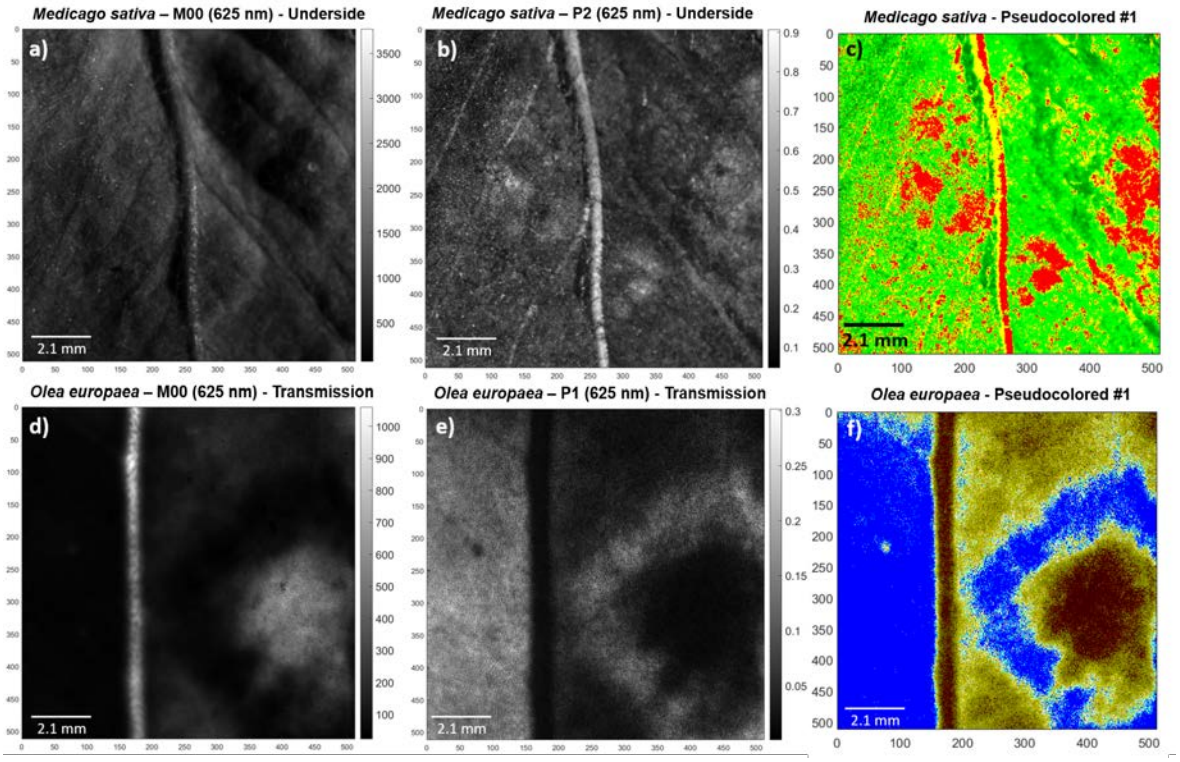


Figure 20: Visual comparison of *Medicago sativa* leaf: a) 625 nm intensity image (M00), b) polarimetric purity index P_2, c) processed image by means of #1 pseudo-coloration; Visual comparison of *Olea europaea* leaf: d) 625 nm intensity image (M00), e) polarimetric purity index P_2, f) processed image by means of #1 pseudo-coloration. [Reproduced from [172]].

The pseudo-colored images resulting from the triplet P_2, P_3 and P_S (i.e., *Pseudo#1*) demonstrate to be the most suitable one to construct the pseudo-colored functions for both *M. sativa* and *O. europaea* samples (Figs. 20c and 20f, respectively). However, the pseudo-colored images obtained based on the IPPs (P_1, P_2 and P_3) lead to similar results. Importantly, both approaches (*Pseudo#1* and *Pseudo#2*) demonstrate an overall visual enhancement of disease symptoms. In particular, the extreme different coloration of the chlorotic lesions on *M. sativa* with respect to the healthy tissue of the leaf lamina (red and green-colored regions in Fig. 20c, respectively) leads to a more accurate location of the diseased area. Similar phenomena occurs for the lesions located on *O. europaea* leaf. In particular, the pseudo-coloration approach lead to a better delimitation of the different transitions from chlorotic spot to necrotic ring and healthy tissue of leaf lamina (dark/brown, blue and yellowish-colored regions in Fig.

20f, respectively). It is important to remark that the pseudo-colored images enhance the contrast between different tissue conditions, even more than the performance of isolated imaging of polarimetric observables for *M. sativa* and *O. europaea* samples. The encoded functions successfully delimit the lesions and other structures in both plant species and demonstrate the further enhance of the visual image contrast between healthy and diseased regions in inspected plants. This results highlight the suitability of using this pseudo-coloration methods for biological samples analysis. Particularly, for the inspection and estimation of direct lesions and the characterization and early detection of infection processes on plant tissues.

4.2 Paper B – *Automatic pseudo-coloring approaches to improve visual perception and contrast in polarimetric images of biological tissues*

Further research on new pseudo-colored approaches is presented in this work (*Paper B* [182]). We report two new polarization-based pseudo-coloration methods devised for the optimized visualization and spatial location of different structures within biological samples. Importantly, the previously reported pseudo-colored methods [149, 183, 184] are based on basic linear combinations of polarization observables whose weights are not generalized but set ad-hoc for each particular sample in a Heuristic approach. Unlike this, here we propose two new definitions for the relative weights based on: (1) the use of the Euclidean distances of actual values of pixels and an average value taken over a given region of interest in the considered image, and (2) the likelihood for each pixel to belong to a given class, defined on the basis of a statistical model that describes the statistical distribution of values of the pixels in the considered image. These two robust methods are applied on two different polarimetric spaces derived from the measured experimental Mueller matrix of inspected samples: the so-called Indices of Polarimetric Purity (IPPs: P_1, P_2 and P_3) and the Components of Purity (CPs: P, D and P_S). These spaces are chosen because they provide a complete description of depolarizing samples and compound a suitable set of metrics for tissues discrimination, as demonstrated in previous works [126, 149, 183, 184]. Therefore, the use of depolarizing spaces define an ideal framework to implement into pseudo-colored approaches addressed for tissue discrimination. Accordingly, four different models are hereby proposed: the Euclidean distance and the Normal-based approaches for the IPPs and the CPs polarimetric spaces. In particular, we test these approaches by analyzing different structures within four representative biological samples with the aim of enhancing the image contrast between them: two of the samples are animal-origin tissues and the other two correspond to vegetal samples. The two animal samples are biopsies from (1) a lamb trachea, (2) a lamb tongue; the vegetal tissues consist of (3) a leaf of *Quercus pubescens* showing powdery mildew lesions caused by the fungus

Erysiphe alphitoides and (4) a leaf of *Vitis vinifera* presenting no symptoms of disease. The complete sample description is provided in section 3.2.

The procedure to obtain the pseudo-colored images is conducted as follows. We start by measuring the experimental Mueller matrix (MM) of the samples above mentioned by means of the complete image Mueller polarimeter described in section 3.1. It is important to remark that, for the case of the *V. vinifera*, the experimental Mueller matrix was acquired by means of a multimodal microscope polarimeter [186]. From the MM we retrieve the standard intensity image (i.e., M_{00}), as well as the depolarization-related observables of interest: the IPPs - P_1, P_2 and P_3 -, and the CPs - P, D and P_S -. The acquired polarimetric images for each sample clearly demonstrate the potential of using depolarization observables. In particular, they show an overall enhancement of image contrast and the revelation of various structures which are unable to be seen by means of regular intensity images. As a representative example, in Fig. 21 we show the non-polarized intensity image (Fig. 21a) and the polarimetric observables corresponding to the IPPs and CPs (Figs. 21b-g) from the *V. vinifera* leaf sample.

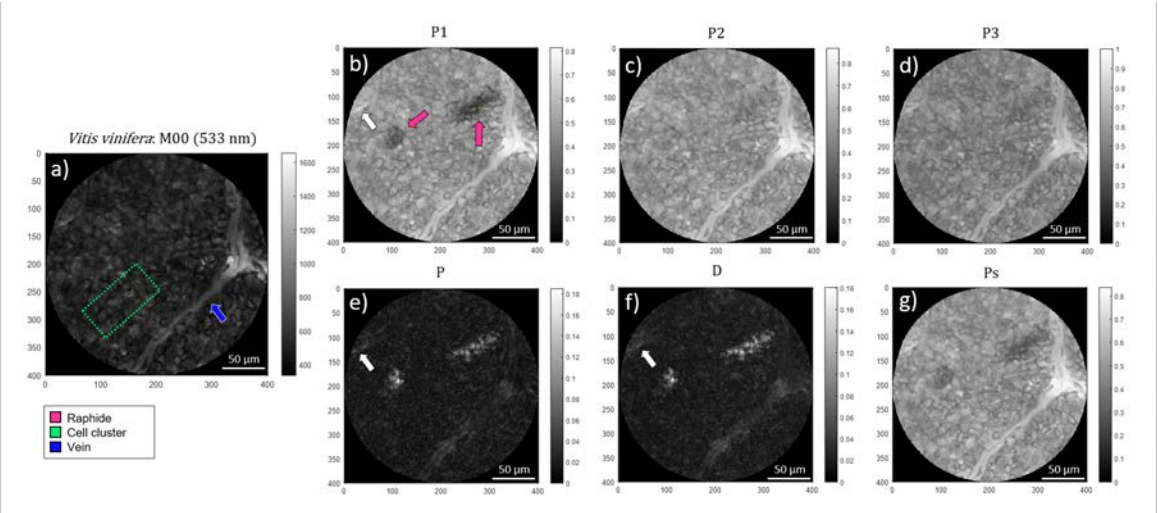


Figure 21: Polarimetric images of the *Vitis vinifera* leaf measured under the microscope for 533 nm illumination wavelength: a) non-polarized transmission image (M00), the Indices of Polarimetric Purity b) P_1, c) P_2 and d) P_3 and the Components of Purity e) P, f) D and g) P_S. The blue and pink arrows indicate the location of the vein and the raphides, respectively. The white arrows indicate the secondary vascular structure. The lime-green dotted box indicates an illustrative region comprising a cell cluster. [Reproduced from [182]].

Interestingly, in the non-polarized transmission image (Fig. 21a) we can observe two main leaf features: the vein (indicated by the blue arrow in Fig. 21a), and the leaf lamina cell clusters (indicated by the green rectangle in Fig. 21a). In addition to that, the polarimetric images reveal the presence of a third structure consisting of a

raphide. In vine leaves, raphides are made of calcium oxalate needle shaped crystals packed together forming prorated clusters of typically 80 μm (long axis) x 30 μm (short axes). Remarkably, raphides are completely invisible in non-polarized transmission images (Fig. 21a), but their spatial location becomes clearly visible in polarization-based images. For instance, they are well visible in the images corresponding to the index of polarimetric purity P_1, polarizance P, diattenuation D and spherical purity P_S (see the pink arrows in Fig. 21b). Furthermore, these polarimetric images reveal another structure that has a polarimetric signature similar to the vein, which is located at the upper-left part of the sample (see the white arrows in Figs. 21b and 21e-f, respectively). This structure may correspond to a secondary vascular structure unable to be detected by means of regular intensity images (Fig. 21a). Therefore, it is clear how polarimetric observables allow the recognition of plant structures not visible in non-polarimetric intensity image M_{00}. This situation highlights the importance of studying the polarimetric channels for plant structures imaging not only for increasing the visual image contrast, but also because they reveal structures hidden in regular intensity images.

To give some numbers, the mean values of the IPPs (Figs. 21b–d) corresponding to the raphide are $P_1 = 0.24 \pm 0.01$, $P_2 = 0.41 \pm 0.01$ and $P_3 = 0.47 \pm 0.01$, followed by these corresponding to the cell clusters, $P_1 = 0.47 \pm 0.01$, $P_2 = 0.50 \pm 0.01$ and $P_3 = 0.56 \pm 0.01$, and finally the ones corresponding the leaf vein, $P_1 = 0.68 \pm 0.01$, $P_2 = 0.72 \pm 0.01$ and $P_3 = 0.81 \pm 0.01$. According to the differences in the polarimetric observables values, the raphide possess an individual signature which differs to that of the vein and the cluster of cells. The reasons that may explain the elevated depolarization of the raphide (i.e., lowest IPPs values) are the scattering and spatial heterogeneity's in the polarimetric properties, which may be higher in the raphide that in the vein or in the cluster of cells. Concerning scattering, it is expected that the refractive index mismatch between a given region and the surrounding media should be higher for the raphide, made of a solid inorganic component, than for veins or clusters of cells which are essentially made of a liquid similar to the surrounding media contained by the membranes forming the cell walls. The largest IPPs values are demonstrated for the vein structure. Accordingly, it induces less depolarization to the incident light than other structures because it is a structure essentially filled with a liquid with low scattering and no polarimetric properties. Regarding the CPs observables, the highest polarizance mean values are demonstrated for the raphides, $P = 0.03 \pm 0.01$, followed by the cell cluster, $P = 0.003 \pm 0.001$, and the leaf vein, $P = 0.001 \pm 0.001$ (see Fig. 21e). Likewise, the raphides show the highest mean diattenuation ($D = 0.07 \pm 0.01$, in Fig. 21f), compared with $D = 0.04 \pm 0.01$ and $D = 0.02 \pm 0.01$, corresponding to the vein and the cell cluster, respectively. Considering these low polarizance and diattenuation

values, the *V. vinifera* leaf can be understood as a non-dichroic sample. Conversely, the highest mean values of the spherical purity P_S (Fig. 21g) are demonstrated for the leaf vein ($P_S = 0.70 \pm 0.01$), this being a direct consequence of the strong alignment of the cellulose filaments within the vein structure. In turn, the P_S values for the cell cluster and the raphides are reduced to $P_S = 0.49 \pm 0.01$, and $P_S = 0.31 \pm 0.01$, respectively. Analogously to the situation in P_1, the differences in the P_S values for the raphide, the cell cluster and the vein, lead to a well-contrasted image with well-differentiated regions (Fig. 21e).

The above-discussed enhancement of image contrast between biological structures is similarly demonstrated when inspecting the polarimetric images of three further examples: the lamb trachea and tongue samples as well as for the *Q. pubescens* leaf. As a representative example, Fig. 22 shows the comparison between non-polarimetric image with the polarimetric observables for the lamb trachea (Fig. 22a-c) and the *Q. pubescens* leaf sample (Fig. 22d-f).

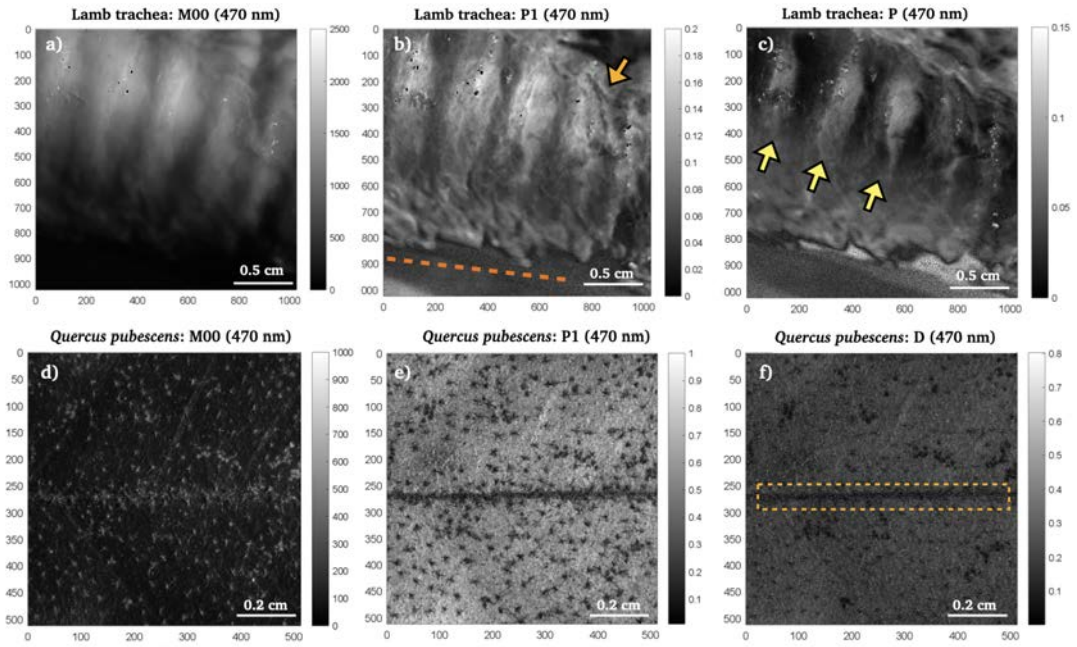


Figure 22: Polarimetric images of the lamb trachea measured at 470 nm: a) non-polarized transmission image (M00), b) the Index of Polarimetric Purity P_1, and c) polarizance, P. Polarimetric images of the *Q. pubescens* leaf measured at 470 nm: d) non-polarized transmission image (M00), b) the Index of Polarimetric Purity P_1, and c) diattenuation, D. The orange arrow and the orange dotted line show the vascular structure within the external trachea sheath and the sample border, respectively. The yellow arrows indicate the location of the cartilaginous rings of the trachea. The orange-dotted box indicates the location of the *Q. pubescens* leaf vein. [Adapted from [182]].

For the particular case of the trachea, P_1 (Fig. 22b) provides information about the surface details of the trachea sheath (e.g. some vascular structures within the external sheath, indicated with the orange arrow in Fig. 22b). Interestingly, the cartilaginous rings demonstrate higher mean values for $P_1 = 0.14 \pm 0.01$ when compared with the trachea sheath ($P_1 = 0.05 \pm 0.01$). This means that the cartilaginous rings are less depolarizing than the trachea sheath. Accordingly, the cartilaginous rings demonstrate higher mean polarizance, P (Fig. 22c) values ($P = 0.06 \pm 0.01$) than the sheath ($P = 0.02 \pm 0.01$). When it comes to the *Q. pubescens*, the evident differences in biological structure and chemical composition between a plant tissue and a fungus, in this case, the *E. alphitoides*, lead to different polarimetric responses. The polarimetric image of P_1 (Fig. 22e) demonstrate the enhancement of the overall image contrast, thus allowing a proper spatial localization of the leaf lesions. Particularly, P_1 is the observable that leads to larger contrast between classes, and demonstrates a higher capability to discriminate between features. Regarding the depolarization content of the inspected sample, the healthy leaf lamina shows high P_1 mean values ($P_1 = 0.82 \pm 0.01$). Conversely, the powdery mildew shows lower mean values ($P_1 = 0.27 \pm 0.01$). Accordingly, it can be said that the leaf lamina induces less depolarization to the incident light than the powdery mildew lesions. The low depolarizing performance of the leaf lamina suggests that it presents a well-organized cell layout within the leaf, and a homogeneous polarization response through the structure. In contrast, the effect of the fungus seems to modify the cell layout structure of the leave, leading to an evident modification of the polarimetric response of the regions with lesions. Regarding to the diattenuation response (Fig. 22f), the sample demonstrates, overall, low mean values ($D = 0.13 \pm 0.01$ and $D = 0.23 \pm 0.01$ for leaf lamina and powdery mildew, respectively).

Once the potential of polarimetric observable images is demonstrated, polarimetric based pseudo-colored approaches are constructed to maximize the visual contrast of different tissues in the sample. As previously stated, two different approaches are following studied: the Euclidean distance-based method and the Normal-based approach. Importantly, the Euclidean distance-based method is limited to the selection of two features, but the Normal-based approach is able to discriminate an unlimited number of characteristics within a single image. Therefore, from the obtained polarimetric images for each inspected sample, we chose to distinguish the following pair of features: for the lamb trachea we aim to enhance the tissue differentiation between the trachea ring and the trachea sheath (see the yellow and blue dashed squares in Fig. 23a, respectively). For the the lamb tongue, we chose the lingual papillae and the tongue epithelial tissue (see the pink and green squares in Fig. 23b, respectively). For the *Q. pubescens* plant sample we select to discriminate between the powdery mildew and the leaf lamina (see

the yellow and blue squares in Fig. 23c, respectively). Complementary, to highlight the potential of the Normal-based model, we select three characteristics within the *V. vinifera* sample corresponding to the leaf vein, the raphide and the cell cluster (see the blue, pink and green squares in Fig. 23d, respectively).

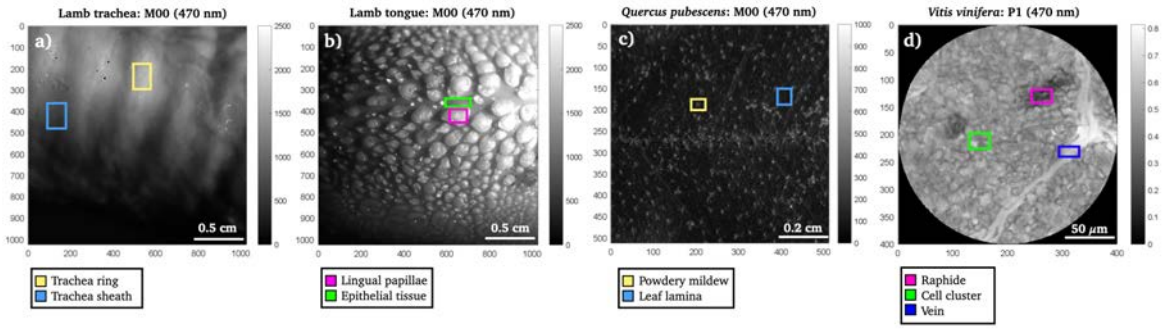


Figure 23: Non-polarized intensity image of a) the lamb trachea, b) the lamb tongue, c) the *Q. pubescens* and the polarimetric image corresponding to the index of polarimetric purity P_1 of d) the *V. vinifera* leaf sample. Each colored square indicates the location of the regions of interest corresponding to the types of tissue to be distinguished within the resulting pseudo-coloration image. [Adapted from [182]].

In the following we define the the parameters of interest involved in the pseudo-coloring models and the general expressions for both the Euclidean-distance and Normal-based approaches. From the experimental Mueller matrix measurement of a sample, we extract n polarimetric observables, $\vec{p} = [p_1, \ldots, p_n]$. Furthermore, we define i classes corresponding to the pair (or more) different organic tissues to be distinguished. Each kind of organic tissues is characterized by the $j = 1, \ldots, n$ (where n is the number of polarimetric observables) means, m_j^i, and standard deviations, σ_j^i, corresponding to the p_j^i observables calculated from a Region of Interest (ROIs) within the specific class i. For image coloring purposes we also define the vector $\vec{C}^i = [R^i, G^i, B^i]$ (where $i = 1, \ldots, k$) as the standard RGB color space coordinates associated with a particular tissue class, i.

The pseudo-colored approach based on the Euclidean distance method consists on computing the Euclidean distance from the values of the n polarimetric observables, $\vec{p} = [p_1, \ldots, p_n]$, from a given image pixel to the mean values $\vec{m}^i = [m_1^i, \ldots, m_n^i]$ of the polarimetric observables of a given class i. Assuming the discrimination between two ($k = 2$) tissue classes (e.g., healthy / infected tissue), the normalized distance is given by,

$$\hat{d}^i = \sqrt{\sum_{j=1}^{n} \left(\frac{m_j^i - p_j}{m_j^{i=1} - m_j^{i=2}} \right)^2}, \qquad i = 1, 2, \tag{22}$$

where $m_j^{i=1} - m_j^{i=2}$ corresponds to the distance between the means of the two classes for a given polarimetric observable, p_j ($j = 1, \ldots, n$). Accordingly, we associate a given RGB color coordinate, $\vec{C}_i$, to each particular tissue class, i. Consequently, the larger the distance from a given pixel to a given class, the lower the corresponding weight to its particular class color coordinate, $\vec{C}_i$. Therefore, each particular class color $\vec{C}_i$ is pondered as,

$$\mathcal{R}^i = 1 - \frac{\hat{d}^i}{d}, \qquad i = 1, 2, \tag{23}$$

where each distance $\hat{d}^i$ is normalized by the sum of the distances, $d = \hat{d}^{i=1} + \hat{d}^{i=2}$, such that $\mathcal{R}$ is positive-definite and ranges between 0 and 1. Thus, the larger the distance $\hat{d}^i$, the lower the amount of i-color level included in the pixel. The final pixel color for the two classes ($i = 1, 2$), is given by,

$$\vec{C} = [R, G, B] = \mathcal{R}^1[R^1, G^1, B^1] + \mathcal{R}^2[R^2, G^2, B^2]. \tag{24}$$

This approach outputs a colored polarimetric image, $\vec{C} = [R, G, B]$, which is based on the linear combination of the $\mathcal{R}^i \cdot \vec{C}^i$ terms of the two classes ($i = 1, 2$), involved.

It is difficult to generalize the Euclidean distance method to more than two tissue classes to be distinguished (i.e., $k > 2$). Therefore, we propose a method to construct the pseudo-colored functions based on the Normal (Gaussian) probability distribution of the n polarimetric observables extracted from the experimental Mueller matrix of the sample. Contrary to the Euclidean method, the Normal approach allows to inspect an unlimited number of classes k. The probability P^i of a pixel corresponding to the jth polarimetric observable ($j = 1, .., n$) of being part of a given class i ($i = 1, .., k$) is defined as,

$$P^i = \prod_j exp\left[-\left(\frac{p_j - m_{i,j}}{\sigma_{i,j}}\right)^2\right], \qquad 0 \leq P^i \leq 1. \tag{25}$$

Accordingly, for each pixel we get as many probability functions, P^i, as classes, $i = 1, \ldots, k$, we want to distinguish. The pseudo-colored image is constructed similarly as for the Euclidean distance-based method. Each probability P^i is associated with a particular RGB color coordinate as:

$$\vec{C} = [R, G, B] = \sum_i^k P^i \cdot \vec{C}^i = \sum_i^k P^i \cdot [R^i, G^i, B^i]. \tag{26}$$

This approach outputs a colored polarimetric image, $\vec{C} = [R, G, B]$ such that the amount of i-color level within a pixel is pondered by the probability of the particular pixel to be recognized as belonging to the class i.

In this study, from the experimental Mueller matrix measurement of four biological samples (lamb trachea, lamb tongue, *Q. pubescens* and *V. vinifera* leaves), we extract the polarimetric observables corresponding to the Indices of Polarimetric Purity (IPPs) and the Components of Purity (CPs). Considering the definition of i classes corresponding to the pair (or more) different organic tissues to be distinguished, we chose the following: the ring and sheath in the lamb trachea sample, the papillae and epithelial tissues in the lamb tongue sample, the powdery mildew and leaf lamina in *Q. pubescens*, and the vein, raphides and cell cluster in *V. vinifera* plant sample.

Therefore, for each one of the polarimetric observables triplet (i.e., the IPPs and the CPs), we build four pseudo-colored images per sample, according to the two pseudo-coloration approaches above-described. These four pseudo-colored images, correspond to the Euclidean-based or Normal-based designs, for the IPP and CP spaces cases. In the following, we provide the resulting pseudo-colored images obtained for the lamb trachea (Fig. 24) and tongue (Fig. 25), as well as for the plant samples corresponding to the *Q. pubescens* (Fig. 26) and *V. vinifera* (Fig. 27) leaves.

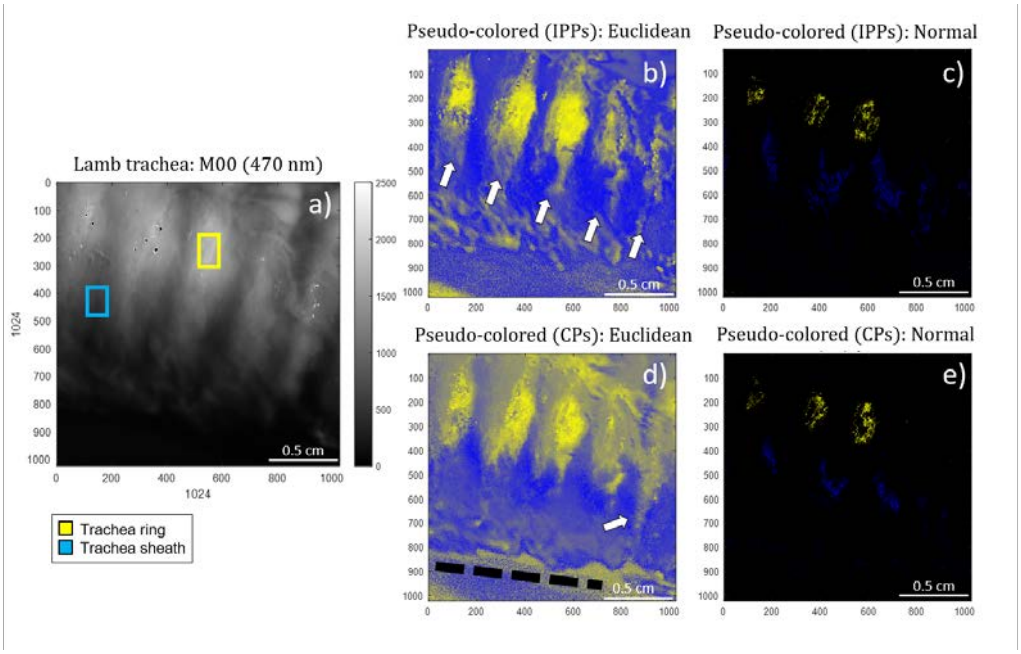


Figure 24: Raw and pseudo-colored images of the lamb trachea: a) Unpolarized intensity image (M00) taken at 470 nm, b) Euclidean and c) Normal pseudo-colored images based on the IPP triplet, d) Euclidean and e) Normal pseudo-coloring based on the CP triplet. Yellow and blue squares show reference areas corresponding to the trachea ring and sheath, respectively. The white arrows (Figs. 2b and 2d) and the black dotted line (Fig. 2d) denote for the cartilaginous rings and the trachea border, respectively. [Reproduced from [182]].

In particular, for the lamb trachea sample, we associate the trachea rings and the

sheath with yellow and blue color coordinates, respectively. The Euclidean distance-based method for both the IPPs and the CPs polarimetric triplets (see Figs. 24b and 24d, respectively) enhances the perceived contrast between cartilaginous rings and trachea sheath. The performance of both polarimetric spaces is quite similar. However, the use of the IPPs set of observables seem to better differentiate the five trachea rings from the sheath (see Fig. 24b), thus allowing a clear an accurate identification and spatial localization of the cartilaginous tissue within the sample. Despite this, the fifth trachea ring and the full trachea structure border are both better spatially located when using the CPs (see Fig. 24d). Conversely, the Normal-based approach (see the Figs. 24c and 24e) has not enough capability to improve visualization, as most pixels in the image are not recognized as part of any of the two tissue classes. Furthermore, the set of IPPs observables seem to well differentiate the five rings (indicated with white arrows in Fig. 24b) from the sheath, thus allowing a clear an accurate identification and spatial localization of the cartilaginous tissue within the sample. In turn, the transition between the cartilaginous rings ans the trachea sheath is less accurate when using the CPs than in the IPPs case: in Fig. 24d, the trachea sheath occupying the space between rings is misrecognized as cartilaginous tissue too.

The resulting pseudo-colored images for the lamb tongue sample are presented in Fig. 25. The lingual papillae and the epithelial tissue were associated with pink and lime-green color coordinates, respectively. The best results are obtained for the Euclidean pseudo-coloring method based on the CPs triplet (Fig. 25d). Importantly, note that the tongue structures placed at the bottom part of the non-polarized diffuse reflectance image (Fig. 25a) are very difficult, or even impossible to be seen in some parts (check for instance the region between the Y axis pixels 800 and 1000) due to the low contrast between them. Same effect can be observed at the very top of the image. This is mostly due to intensity losses or defocusing introduced by the measure of a non-planar sample (the tongue), where only the central region is properly illuminated and in focus. In contrast to that, a clear visual enhancement of those regions is provided by the pseudo-colored images in Figs. 25c and 25d. In the case of the pseudo-coloring based on the Normal method (Figs. 25c and 25e), the best visualization is demonstrated by the implementation based on the IPPs (Fig. 25c). However, the class recognition rate is lower than for the Euclidean distance case (see the white arrows in Figs. 25c and 25e, pointing the lingual papillae recognition for the Normal method, compared with correct recognition of the epithelial tissue in Fig. 25b, by the Euclidean distance method). In the case of the pseudo-coloring based on the Normal method, the mean values selected for the epithelial tissue (which are obtained from the region of interest corresponding to the green rectangle in Fig. 25a) are not representative enough of the properties of the tissue as they present a large variance across the image. For this

reason, an important part the epithelial tissue could not be successfully assigned to the correct class by the method and therefore appears in black (Figs. 25c and 25e). Therefore, the pseudo-coloring based on the Euclidean approach performs better than the Normal approach.

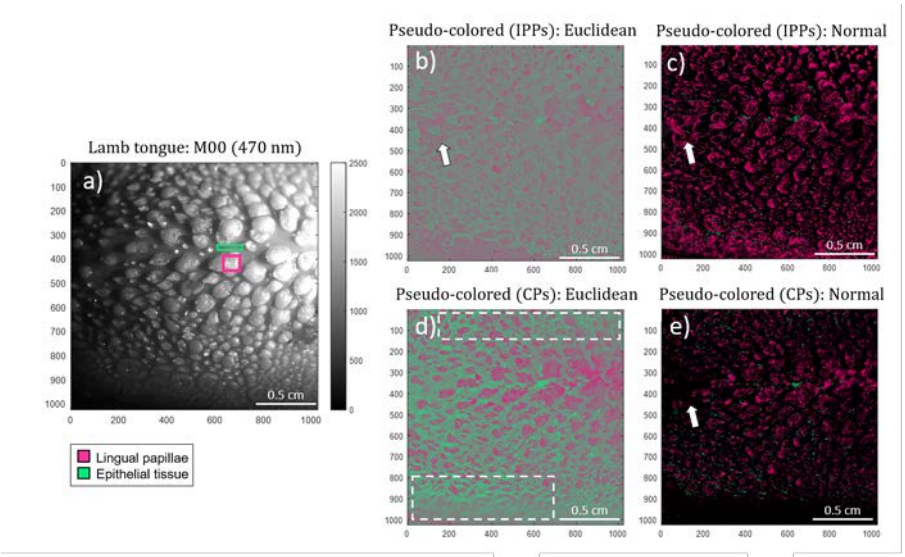


Figure 25: Intensity and pseudo-colored images for the inspected lamb tongue: a) regular intensity image (M00) captured at 470 nm illumination wavelength, b) Euclidean and c) Normal pseudo-colored images for IPPs, d) Euclidean and e) Normal pseudo-coloring for CPs. Pink and green squares (Fig. 4a) denote for the selected regions of interest (ROI) of lingual papillae and epithelial tissue, respectively. White arrows (Figs. 4b, 4c and 4e) indicate a particular region only containing epithelial tissue. The dotted-squares indicate the unseen / out of focus region (by means of the unpolarized image, M00) of the tongue. [Reproduced from [182]].

Switching to the inspection of plant samples, the pseudo-colored images of the *Q. pubescens* leaf are presented in Fig. 26. In this case, we associate the yellow and blue color coordinates to the powdery mildew and the leaf lamina, respectively. The Euclidean distance-based approach (Figs. 26b and 26d) demonstrates an obvious visual enhancement of the two tissue types to be distinguished. The infected regions by the powdery mildew clearly appear in a shiny yellow over a blue background which corresponds to the leaf lamina. Importantly, the use of the IPPs observables (Fig. 26b) seem to provide a better discrimination than the CPs set (Fig. 26d). The latter can be seen, because in the pseudo-colored image using the CPs set (Fig. 26d) there are some areas (highlighted with white rectangles) where some pixels are misclassified while the same pixels are correctly classified in the image pseudo-colored using the IPPs set. Moreover, the Euclidean distance-based method shows a better performance than the Normal function-based approach.

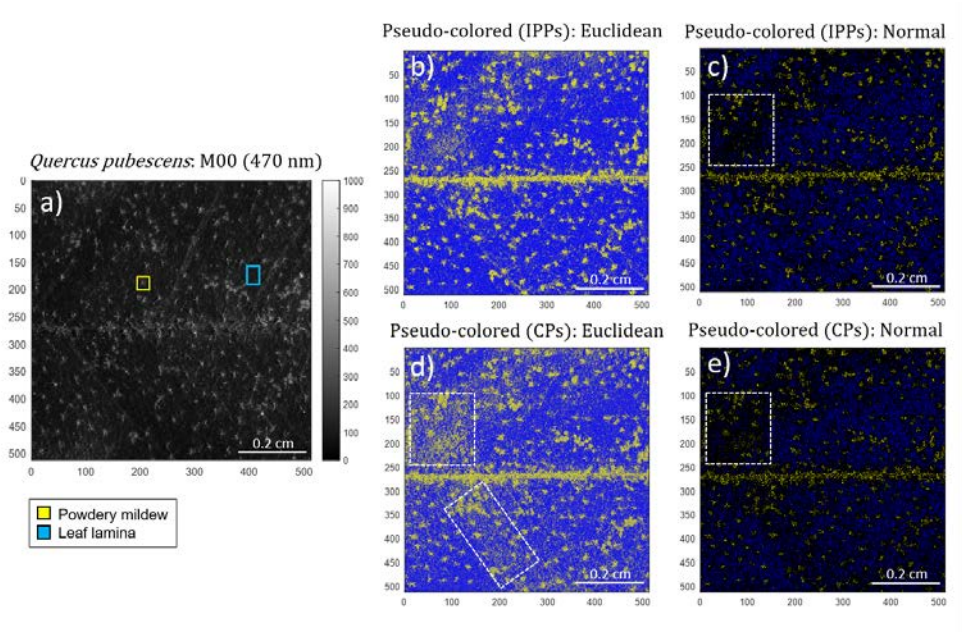


Figure 26: Intensity and pseudo-colored images for the inspected *Q. pubescens* leaf: a) non-polarized (transmission – diffuse reflection) image (M00) captured at 470 nm illumination wavelength, b) Euclidean and c) Normal pseudo-colored images for IPPs, d) Euclidean and e) Normal pseudo-coloring for CPs. Yellow and blue squares (Fig. 6a) denote for the selected regions of interest (ROI) of powdery mildew lesion caused by the fungus Erysiphe alphitoides and the healthy leaf lamina, respectively. White-dotted squares (Figs. 6c and 6e) denote for the misrecognized pixel regions of both Euclidean and Normal-based methods. [Reproduced from [182]].

Finally, the *V. vinifera* leaf sample contains three tissue classes to be distinguished. Accordingly, the pseudo-coloring based on Normal distribution may have an advantage in this particular situation. Unlike for the examples previously discussed, we select three classes of structures to be simultaneously visualized in the Normal-based approach: the raphides, a cell cluster and the leaf vein. The selected regions of interest, corresponding to the three classes, are indicated with pink, lime-green and blue squares within the purity index P_1 (Fig. 27a). In the case of the Euclidean distance method, as it is restricted to handle only two classes, we chose the raphide and the cell cluster (pink and lime-green in Fig. 27a). Note that unlike in the previous samples, instead of using the non-polarized transmission image (M00; Fig. 21a) to design the ROIs for the classes we use the P_1 image (Fig. 27a) since the raphides are not visible in the non-polarized transmission intensity channel. The Euclidean distance-based pseudo-coloring images of IPPs and CPs are shown in Fig. 27b and Fig. 27d, respectively; and the Normal-based models implemented for IPPs and CPs are shown in Fig. 27c and Fig. 27e, respectively.

Regarding the performance of the two methods in terms of class coloring and

visual discrimination, we observe some differences. On the one hand, the IPPs are not sensitive enough to correctly identify the location of raphides when implementing the Euclidean approach (Fig. 27b). In particular, some pixels belonging to raphides are not well-colored in pink, but other pixels that do not belong to raphides, they are incorrectly painted in pink. Unlike this, the Euclidean method applied with CPs observables is much more efficient and quite accurate discriminating between raphides and cell cluster (Fig. 27d). On the other hand, the IPPs observables applied with the Normal-based approach are able to correctly identify and localize all the studied classes: raphides (pink regions in Fig. 8c), the leaf vein (blue region in Fig. 27c) and the cell cluster (lime-green pixels in Fig. 27c). Furthermore, the vascular structure located on the upper-left part of the sample image is also colored in blue (indicated with a white arrow in Fig. 27c), as it is recognized as part of a vein. Finally, when applying the CPs observables with the Normal-based approach (Fig. 27e), all the classes are correctly discriminated as well, but due to they are more affected by distances between pixel-values and mean classes-values, when applying the Gaussian probability function, more pixels tends to zero probability of belonging to any class, and then, painted in black, this darkening the whole image.

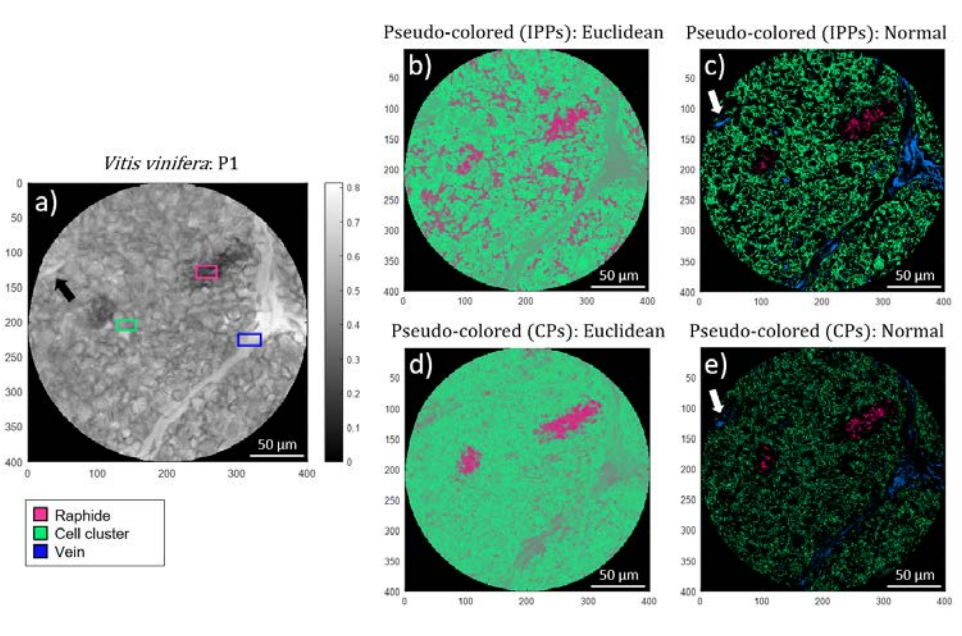


Figure 27: Non-polarized transmission (M00) and pseudo-colored microscopic images from a *Vitis vinifera* leaf section captured at 533 nm illumination wavelength: a) polarimetric purity index P1, b) Euclidean and c) Normal pseudo-colored images for IPPs, d) Euclidean and e) Normal pseudo-coloring for CPs. Pink, lime-green and blue squares denote for the selected regions of interest (ROI) of raphides, leaf lamina cells and vein, respectively. The white arrows indicate the location of the secondary vascular structure. [Reproduced from [182]].

Summarizing, the obtained pseudo-colored images clearly overcome the standard intensity images as well as individual polarimetric observable images, in terms of spatial location, visualization and recognition of the selected structures within the inspected biological samples. It is important to note that the pseudo-colored images include within a single image the relevant information of multiple polarimetric observables. Therefore, the proposed models demonstrate more potential - in tissue discrimination - than individual polarimetric images.

Importantly, with regards to the application of the IPPs or the CPs observables within the proposed methods, we have chosen these metrics because previous works highlight the suitability of depolarization observables for the discrimination of tissues and because these two bases, all together, completely describe the depolarization response of a samples [149, 183, 184]. Therefore, both the IPPs and CPs represent an ideal framework to implement pseudo-colored functions for tissue discrimination. Furthermore, for each biological sample to be analyzed, we recommend the use to both the IPPs and the CPs: whereas the IPPs are more sensitive to the depolarization anisotropies, the CPs are more related to the physical properties of the constituents of the sample being at the roots of depolarization (retardance, polarizance, diattenuation). At the end, the characteristics of each particular biological sample will determine which one of the two basis will provide the more vivid contrast.

4.3 Paper C – *Polarimetric data-based model for tissue recognition*

The use of polarimetric observables for biological tissue classification has gained significant attention in recent years due to its potential in cancer detection and recognition. Many studies have shown the effectiveness of polarimetry in detecting and characterizing various types of cancer [10–14, 37, 96–112], making it an important tool in the field of biomedical research.

In this work (*Paper C* [174]), we aim to further explore the potential of polarimetry in tissue classification and provide a predictive optical model for tissue recognition based on polarimetric indicators. The development of accurate and reliable tissue classification methods is crucial for early detection and diagnosis of diseases, as well as for guiding surgical procedures and improving patient outcomes. We believe that our work will contribute to the advancement of polarimetry as a valuable tool in the field of biomedical research and have the potential to improve the diagnosis and treatment of various diseases.

The study is conducted on the experimental Mueller matrices of four biological tissues corresponding to bone, tendon, muscle and myotendinous junction (see Fig. 8

in section 3.2), measured from a collection of 157 *ex-vivo* chicken samples. A total of 23 bones, 50 muscles, 50 myotendinous junctions and 34 tendons are analyzed. The experimental Mueller matrices are measured at three different wavelengths covering the visible range (625 nm, 530 nm and 470 nm) using the complete Mueller imaging polarimeter described in section 3.1. These experimental Mueller images correspond to a region of interest (ROI) of 512 $\times$ 512 pixels which correlates to an area of 1.1 $\times$ 1.1 cm^2. By using the measured experimental Mueller images, we retrieve the polarimetric images corresponding to the IPPs (P_1, P_2 and P_3), the depolarization index P_Δ, polarizance P, diattenuation D, global and linear retardance R and δ, respectively, and the optical rotation Ψ. To construct the database, from each complete 512 $\times$ 512 image we compute the mean value for the different polarimetric observables from an image sub-ROI of 150 $\times$ 150 pixels (which does not include the saturated -specular reflection originated- regions). Accordingly, the data matrix used in this study is composed by 157 $\times$ (27 + 1) items: 157 tissues cases, analysed through 27 predictor variables (9 polarimetric metrics -P_Δ, P_1, P_2, P_3, P, D, R δ and Ψ- measured at 3 wavelength channels; 625 nm, 530 nm, and 470 nm) and one classifier variable with four categories (the type of biological tissue: muscle, tendon, myotendinous junction and bone).

We start by conducting an exploratory statistical analysis on the data matrix of polarimetric indicators to evaluate the normality of the data distributions. We examined measures of central tendency, dispersion, and asymmetry, as well as conducted Kolmogorov-Smirnov and Shapiro-Wilk tests to assess the normal distribution fit assumptions [187, 188]. The analysis revealed the non-normal behavior in the measured data. Therefore, we employed non-parametric statistics procedures, specifically the Kruskal-Wallis test [189], to determine the discriminatory potential of the polarimetric indicators for different tissue types: the analysis revealed that P_Δ, the IPPs (P_1, P_2, and P_3), and polarizance P are the most sensitive and informative indicators for tissue characterization. Notably, P appears to be the only metric capable of discriminating between myotendinous junction and bone tissue. Additionally, retardance indicators also demonstrated some predictive potential. Based on these findings, we conclude that the polarimetric characteristics of selective absorption, retardance, and depolarization exhibit distinct signatures in the studied tissues, and are suitable to be used in predictive optical models.

Importantly, techniques such as discriminant analysis, which require normally-distributed data and equal variances for each tissue type, cannot be used for our model. As an alternative, we propose a more robust model based on the logistic regression, which does not have these requirements. Additionally, the number of variables in the

original dataset is aimed to be reduced in order to minimize the multicollinearity and improve the model stability. Specifically, our solution involves: (i) a factor analysis with Principal Component Analysis (PCA) [122] to identify a set of independent predictor variables that retain the information from the original variables and explain a significant portion of the variance in the data. These principal components will be used as parameters in our multivariate predictive models based on a binary logistic regression; and (2) the use of Receiver Operating Characteristic (ROC) curves [190] to evaluate, compare, and optimize the predictive capacity of these models for classifying different types of biological tissue.

The Principal Component Analysis (PCA) is a tool used to reduce the dimensions of a data set. These principal components, $\mathcal{C}$, are mathematical constructs that help to structure the data and are determined by the original variables. These components can be analyzed in relation to other variables, and geometrically, they correspond to the best fitting axes for the data points in the original data set. Before performing the PCA, we assessed the suitability of this method using Bartlett's test of Sphericity and the Kaiser-Meyer-Olkin Measure of Sampling Adequacy (KMO) [191, 192]. The results of these tests indicated the appropriateness of our data; thus we calculated the principal components using the *SPSS software*. We retain the first 10 principal components, because they explain more than 90% of the variance in the original metrics. These 10 selected principal components encode the polarimetric information and can be written as linear combinations of the 27 studied polarimetric metrics, weighed by constants provided by the component score coefficient matrix [122]. Figure 28 shows a two-dimensional space plot of the 27 indicators as a function of the two main principal components, $\mathcal{C}_1$ and $\mathcal{C}_2$.

Plotting principal components against each other allows to see how they are connected to physical information of samples and which polarimetric indicators contribute most to data variability. It is important to note that the metrics in Fig. 28 are clustered into four main areas based on physical characteristics regardless the wavelength. These clusters included indicators of retardance (R and δ), depolarization (IPPs and P_Δ), dichroism (D and P), and optical rotation (Ψ). Analysis of polarimetric indicators found that $\mathcal{C}_1$ was primarily responsible for describing scattering and depolarization (IPPs and P_Δ) with weights around 0.5 and 1.0, while $\mathcal{C}_2$ had weights between 0 and 0.5. Retardance indicators are largely influenced by the $\mathcal{C}_2$ principal component (between 0.5 and 1.0) and less so by the $\mathcal{C}_1$ component (close to 0). Diattenuation and polarizance are influenced by both main principal components, with a greater representation by the $\mathcal{C}_2$ component. Optical rotation is not significantly described by the first two principal components, with both $\mathcal{C}_1$ and $\mathcal{C}_2$ values near 0. The grouping of

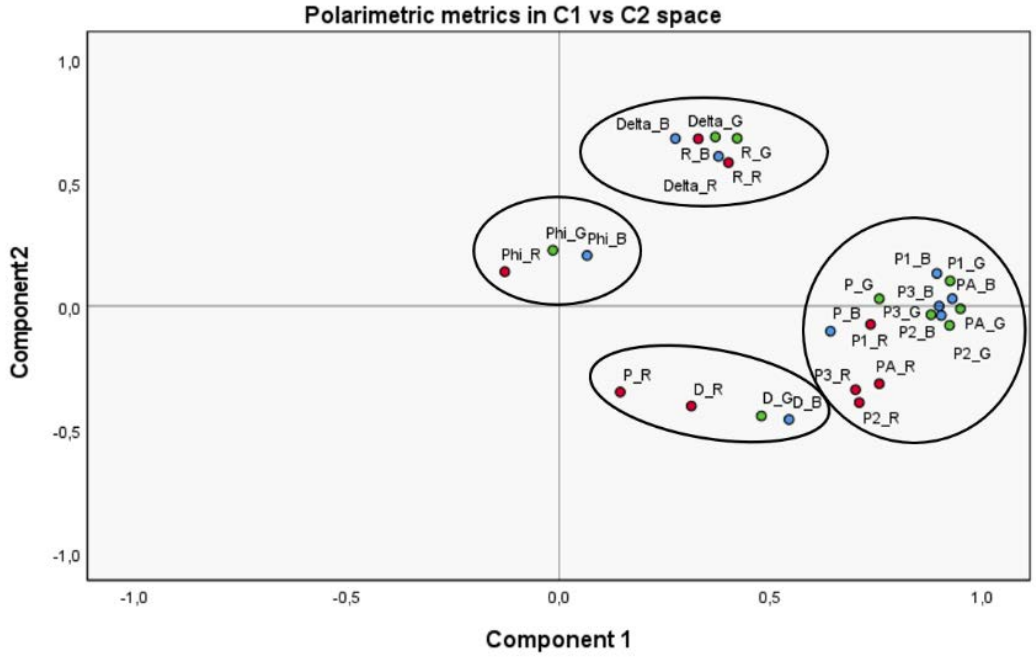


Figure 28: The plot showing principal components $\mathcal{C}_1$ against $\mathcal{C}_2$ represents the correlation coefficients between the 27 polarimetric indicators and the two first extracted principal components. The notation of the polarimetric indicators names is composed of the word that represents the measured parameter: P_1, P_2, P_3 and P_A (IPPs and P_Δ, respectively), D and P (diattenuation and polarizance), R and *Delta* (global and linear δ retardance) and *Phi* (optical rotation Ψ), followed by R, G or B (corresponding to red, green and blue measured wavelength, respectively). [Reproduced from [174]].

physical variables in defined areas of the $\mathcal{C}_1$ - $\mathcal{C}_2$ space suggests that different origins of data variability are related to distinct physical structures in the tissue samples.

We further examined the performance of the 10 principal components as tissue classifiers by using ROC curve analysis [190]. This method plots the true positive rate (TPR) against the false positive rate (FPR) for different thresholds of the classifier, allowing the evaluation of its ability to classify samples into two categories. The TPR and FPR are based on the combination of the sample's actual value and how it is classified by the model, resulting in four possible outcomes. These outcomes are used to calculate the sensitivity and specificity, which form the basis for ROC curves. As a representative example, Fig. 29 shows the ROC curve for the principal component $\mathcal{C}_1$ and each tissue type.

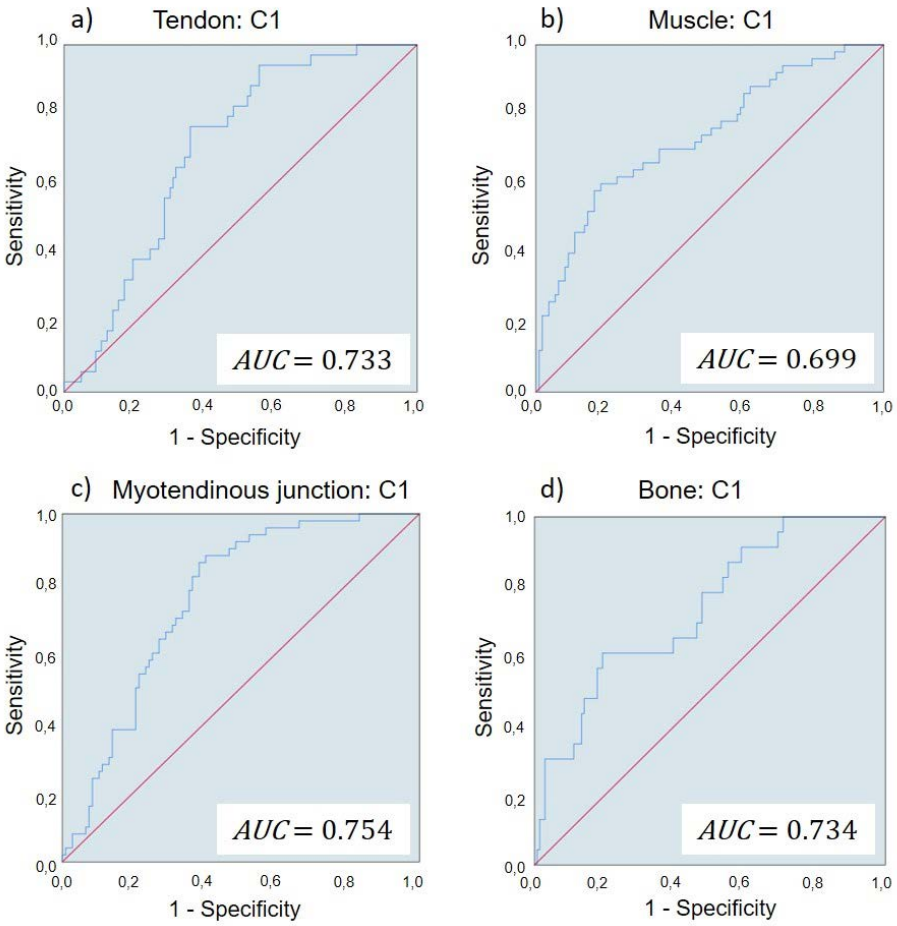


Figure 29: ROC curve of the principal component C_1 for (a) muscle, (b) tendon, (c) myotendinous junction and (d) bone. [Reproduced from [174]]

.

In addition, to compare the performance of different classifiers, we use the area under the curve (AUC): AUC values of 0.5 indicate the low predictive capability of the classifier; AUC values of 1 indicate the 100% sensitiviy and specificity of the classifier. Using *SPSS software*, we calculated the ROC curve for all 10 principal components and for each of the four tissue types. All principal components showed statistical significance in the ROC curves for one or more tissue types, indicating their potential usefulness in the classification process. The AUC values of all principal components when discriminating each biological tissue against the remaining three are summarized in Table 1. The results in Table 1 indicate that the majority of principal components possess discriminative potential, as evidenced by values of AUC greater than 0.5. The components with the highest AUC for each tissue are highlighted in gray in Table 1. Notably, $\mathcal{C}_1$ demonstrates significant discriminatory values across all tissues, while $\mathcal{C}_2$ and $\mathcal{C}_3$ primarily provide discriminatory information for tendon. Components $\mathcal{C}_4$ and

$\mathcal{C}_5$ demonstrate a focus on discrimination between muscle and tendon, and $\mathcal{C}_6$ between tendon and bone; $\mathcal{C}_7$ primarily provides information on bone discrimination, while $\mathcal{C}_8$ maintains information on discrimination of muscle, tendon, and bone. Finally, $\mathcal{C}_9$ and $\mathcal{C}_{10}$ provide discrimination for the myotendinous junction.

Table 1: AUC of 10 first principal components for all tissues. [Reproduced from [174]].

	Muscle	**Tendon**	**Myotendinous junction**	**Bone**
C_1	0.733	0.699	0.754	0.734
C_2	0.559	0.717	0.567	0.577
C_3	0.528	0.651	0.583	0.512
C_4	0.731	0.745	0.563	0.542
C_5	0.667	0.733	0.500	0.527
C_6	0.516	0.714	0.521	0.725
C_7	0.567	0.567	0.508	0.722
C_8	0.703	0.706	0.544	0.649
C_9	0.581	0.540	0.616	0.507
C_{10}	0.569	0.495	0.623	0.601

The probabilistic model is based on a probabilistic function fit to the experimental data using a logistic function - which transforms the non-linear data into a linear model through a logistic regression fit -. Importantly, we have chosen to use a logistic function for our model because it is suitable for non-parametric data and does not require a linear relationship between the predictors and the probability of the target outcome. The probability function depends on the values of certain principal components, which are calculated based on the polarimetric indicator values measured from the samples. We used *SPSS software* to conduct a step-wise regression with backward elimination and the Wald estimator [193] to determine the most significant principal components. This routine allow us to obtain the following probabilistic functions for the four studied tissues:

$$P_{Muscle} = \frac{1}{1 + e^{-(-2.183C_1 + 1.535C_4 - 1.280C_5 - 1.077C_8 + 0.629C_9 - 1.879)}}, \tag{27}$$

$$P_{Tendon} = \frac{1}{1 + e^{-(-2.213C_1 - 2.358C_4 + 1.625C_5 + 1.257C_6 + 1.059C_8 - 3.128)}}, \tag{28}$$

$$P_{Myotendinous} = \frac{1}{1 + e^{-(0.900C_1 - 0.573C_9 + 0.589C_{10} - 0.934)}}, \tag{29}$$

$$P_{Bone} = \frac{1}{1 + e^{-(1.018C_1 + 0.715C_2 + 0.621C_4 - 1.452C_6 - 1.437C_7 - 0.688C_{10} - 3.132)}}. \tag{30}$$

The efficiency of the four probabilistic models was studied by computing the associated ROC curves, represented in Fig. 30. The AUC values were significantly larger than those obtained when representing the principal components alone (see Table 1). The AUC values for the muscle, tendon and bone tissue models are high (0.92, 0.95 and 0.89 in Fig. 30, respectively), which highlights the superior discriminatory potential of the probabilistic functions. The myotendinous junction model has the lowest discriminatory potential ($AUC = 0.79$).

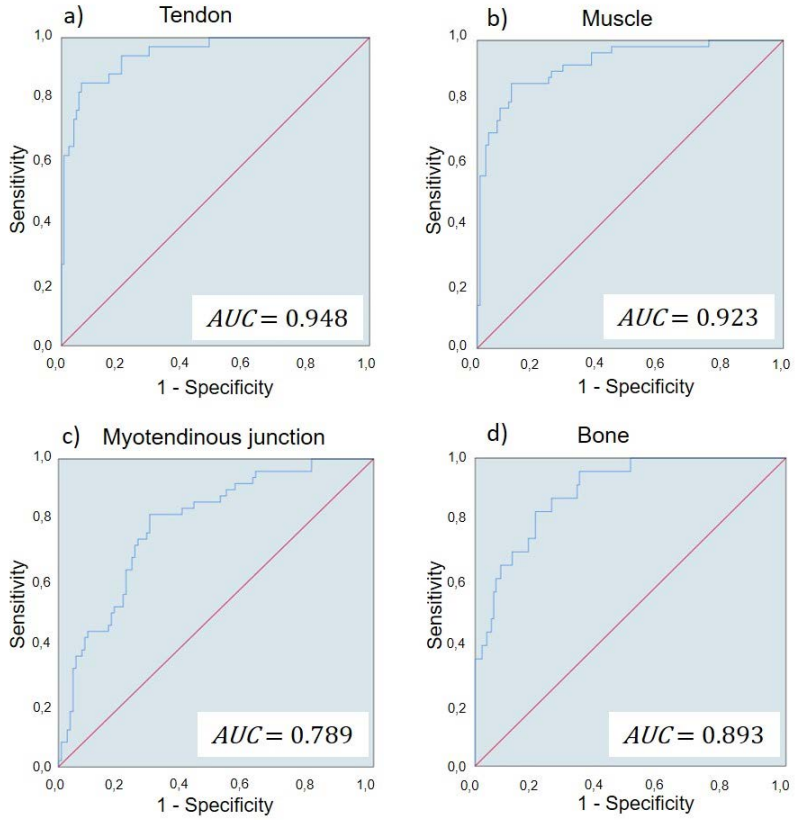


Figure 30: ROC curve of the probabilistic model for (a) tendon, (b) muscle, (c) myotendinous junction and (d) bone. [Reproduced from [174]].

Models in equations (27-30) output a probability value between 0 and 1 associated with the likelihood of a tissue belonging to a certain category. To build a dichotomous model, a threshold must be set to distinguish between a "*Yes*" or "*No*" classification. However, each threshold will result in different sensitivity and specificity values. We use the Youden's Index [194] to determine the optimal threshold, which corresponds to the farthest point from the ROC diagonal on the ROC curve, as calculated by d = sensitivity + specificity – 1. Table 2 shows the computed Youden's index, sensitivity and specificity values for each model. The muscle, tendon and bone models have high

specificity and sensitivity values (above 80%). The worst result is for the myotendinous junction tissue with a specificity of 71%.

Table 2: AUC, Youden's index, sensitivity and specificity of each predictive model. [Reproduced from [174]].

	Muscle	**Tendon**	**Myotendinous junction**	**Bone**
AUC	0.923	0.948	0.789	0.893
Threshold (Youden's index)	0.392	0.351	0.285	0.162
Sensitivity (%)	86.0	85.3	82.0	82.6
Specificity (%)	88.8	93.5	71.0	80.6

The sensitivity and specificity results for the predictive models in Table 2 demonstrate their ability to classify the studied organic tissues. Additionally, the study highlights the potential of using predictive models based on polarimetric data to differentiate between animal tissues and suggests potential applications, such as in medical imaging. To further demonstrate the potential of this method, an algorithm was developed such that it applies the probabilistic function to each pixel of the image of a measured sample. Specifically, using a Mueller matrix image of an arbitrary chicken tissue measured at three wavelengths (625 nm, 530 nm, 470 nm), the 27 polarimetric images were calculated. Accordingly, the algorithm calculates the probability of the pixel being recognized as a particular tissue (resulting in a probability image with values ranging between 0 and 1) using the functions in equations (27-30). The predictive model was tested for arbitrary samples of tendon, muscle, myotendinous junction and bone. In all cases, the results of the algorithm demonstrated its high capability and accuracy when classifying the corresponding tissues; thus indicating that the proposed models are suitable for tissue classification and recognition, as shown by the sensitivity and specificity outcomes in Table 2.

As a representative example, Fig. 31 illustrates the probability function images for the four predictive models (equations (27-30)) when analyzing an arbitrary sample of a chicken tendon. The figure includes the intensity image of the tendon (M00, Fig. 31a) and the output images of the probabilistic models for recognizing muscle (Fig. 31b), tendon (Fig. 31c), myotendinous junction (Fig. 31d), and bone (Fig. 31e). The probability images are presented in gray scale, with white representing the maximum probability value (1) and black representing the minimum (0). The tendon predictive model (Fig. 31c) shows the highest probabilistic values, correctly classifying the sample as a tendon. The probability images for muscle (Fig. 31b), myotendinous junction (Fig. 31d), and bone (Fig. 31e) have significantly lower values and correctly

classify the sample as no-belonging to their corresponding tissue classes (values close to zero). The myotendinous junction model has the largest values among these models, but is the least sensitive and specific, as previously discussed. In the same vein, as the myotendinous junction is a transition between tendon and muscle, it is the most difficult tissue to be properly classified.

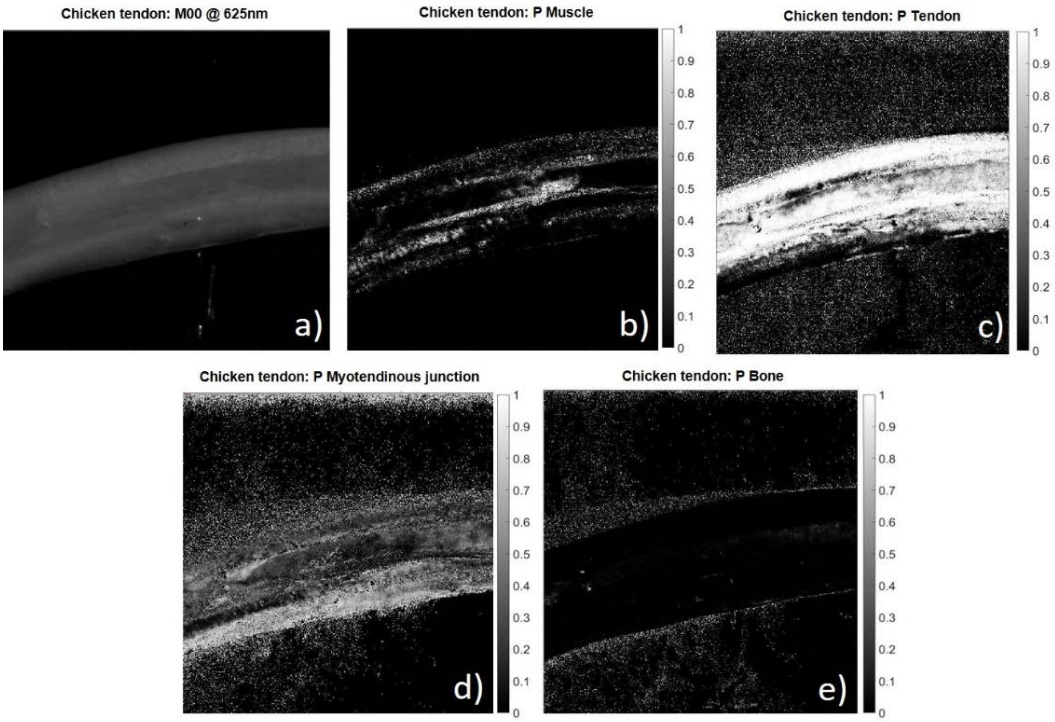


Figure 31: Intensity image M00 (a) and probability image outcome when applying the Muscle-model (b), Tendon-model (c), Myotendinous junction-model (d) and Bone-model (e) on an arbitrary chicken tendon sample. The gray level bars, placed to the right of the corresponding probability function images, defines the probability of the pixel to be recognized as a particular tissue, in a range between one (white) or zero (black). [Reproduced from [174]].

4.4 Paper D – *Optimizing the classification of biological tissues using Machine Learning models based on polarized data*

The potential of polarimetric methods to enhance the visualization of different organic tissue structures is now being used to construct recognition models (as discussed in *Paper C* [174]), which may be useful as a guided assistance for clinical procedures or for the early detection of pathologies in animal and vegetal tissues. In the literature, different proposals using polarimetric data to train tissue recognition models based on regressive [174] or classificatory methods have been found [111, 112, 116]. Machine learning (ML) models trained with polarimetric data have been found to be suitable

for classification purposes [109, 111, 112, 114, 118]. Studies have shown sensitivity to malignant formations in *ex-vivo* colon samples [111, 112], accurate classification of healthy and cancerous tissue in histological sections of breast tissues [109] and early diagnosis of carcinoma in skin [114], among others.

When building recognition models, different polarimetric observables can be used for their physical interpretation of samples and their potential for contrast or classification. These observables are derived from the Mueller matrix of the sample, which includes all measurable information on linear polarimetric light-matter transformations. However, these mathematical transformations may result in loss of relevant information and decreased performance of the recognition models. This is an important issue not discussed in the literature so far, and that becomes the central goal of this work.

In this study (*Paper D* [175]), we compare different sets of polarimetric observables to determine the most suitable framework for biological tissue classification tasks. We aim to evaluate the possible loss of relevant information when transitioning from Mueller matrix elements to physically interpretable polarimetric observables using three different datasets (labeled A, B and C). The first dataset, A, includes 10 polarimetric observables such as the Indices of Polarimetric Purity (IPPs, P_1, P_2 and P_3), the Components of Purity (CPs, comprising the degree of spherical purity P_S, polarizance P and diattenuation D), the depolarization index P_Δ, global and linear retardance (R and δ, respectively) and retarder rotation Ψ, obtained from the corresponding Mueller matrices at the illumination wavelengths of 625, 530 and 470 nm. This results in a total of 30 observables. Note that the selection of the observables we chose for the study is not arbitrary but we selected a set of observables that fully inspect the main polarimetric features of a sample: dichroism, birefringence and depolarization. The second dataset, B, consists of 15 Mueller matrix elements (excluding the coefficient m_{00}) measured at the same three illumination wavelengths. The third dataset, C, contains the 75 features from datasets A and B together. We also consider the influence of different wavelengths within the visible range on recognition model's performance as different wavelengths interact with samples differently. The study is conducted in a set of 33 *ex-vivo* chicken thigh specimens from which we dissected bones and soft tissues (muscle, tendon and myotendinous junction). The complete description of the samples is provided in section 3.2. Particularly, we measured the experimental Mueller matrix - by means of the complete image Mueller polarimeter described in section 3.1 - from a total of 165 samples (50 muscles, 34 tendons, 50 myotendinous junctions, and 31 bones). Our raw database consists of: 3 wavelengths $\times$ 25 observables (corresponding to 15 normalized MM elements excluding m_{00} + 10 polarimetric observables) = 75 features per type of tissue (tendon, muscle, myotendinous junction and bone).

We start by applying a data cleansing process to the raw data. To do so, we selected a reduced region of interest (ROI) of 150 × 150 pixels from the original 512 × 512-pixel images to ensure that the images consist of pixels only from the specific tissue, and to avoid artifacts and misleading information. This resulted in homogeneous tissue sub-region images that exclusively contain the desired tissue information and do not include saturated pixels or background information. Machine learning algorithms are trained using a set of labeled data known as training data and their classification accuracy is evaluated using a separate set of data, referred to as test data, which is not included in the training set. Therefore, we divide our dataset into two mutually exclusive sub-sets, one for training and one for testing the algorithms. This division consists of 80% of the data for training and 20% for testing. To avoid *overfitting*, we created two separated datasets containing tissues from different chicken samples, with 50 muscles, 34 tendons, 50 myotendinous junctions, and 31 bones. To increase the amount of data for training and testing, we divided each 150 × 150 pixels sub-sample image into a grid of 15 × 15 sub-ROIs, with each sub-ROI having a 10 × 10 pixels size. This resulted in an increase of 37,125 samples from the original 165 samples in the raw database, with 75 features per sample: we augmented the original data to 11,250 muscle images (9,000 for training and 2,250 for testing), 7,650 tendon images (6,075 for training and 1,575 for testing), 11,250 myotendinous junction images (9,000 for training and 2,250 for testing), and 6,975 bone images (5,625 for training and 1,350 for testing) respectively.

The importance of the different features within these A, B and C datasets is evaluated using the Boruta algorithm [195]. This algorithm is commonly employed to identify the salient features and eliminate those that are not essential for classification, potentially improving the performance of the classifier. Our results indicate that, in this study, the Boruta algorithm identifies the entire set of features in datasets A, B, and C as important and therefore, none of the polarimetric observables nor the Mueller matrix elements were removed from the datasets. It is worth noting that the Boruta algorithm was utilized in this study to select the most pertinent features, however, alternative approaches such as factor analysis, could have been employed as well.

We trained various algorithms commonly found in the literature to train on our datasets A, B, and C. In order to maintain the focus of the manuscript, we restricted the examination to the four models that yielded the highest accuracy: an Artificial Neural Network (ANN) [121], Support Vector Machines (SVMs) [125], Light Gradient Boosting Machines (LGBMs) [196], and Extreme Gradient Boosting (XGBoost) [197]. This resulted in the creation of 12 final models: (1) ANN with A, (2) ANN with B, (3) ANN with C, (4) SVM with A, (5) SVM with B, (6) SVM with C, (7) LGBM with A,

(8) LGBM with B, (9) LGBM with C, (10) XGBoost with A, (11) XGBoost with B, and (12) XGBoost with C.

We initiate the training process for multiple machine learning (ML) algorithms using a set of pre-determined initial parameters. The model parameters are then fine-tuned using A, B, and C training datasets. This process is iteratively repeated until the classification error on the training data is sufficiently small, as determined by the accuracy score. To optimize our classifiers, we employ the random search cross-validation method (with five-fold partitioning) as described in [198]. Finally, the performance of the various classifiers is evaluated by applying them to test data, which are classified into one of the four classes under investigation: bone, muscle, tendon, and myotendinous junction. The optimization results are presented in terms of the global accuracy score, which is defined as the ratio of correct predictions to the total number of predictions made. The corresponding results are summarized in Table 3.

Table 3: Global accuracy score (in %) obtained from each of the 12 classification models. Labels A, B and C denote for the polarimetric datasets. Gray-colored boxes indicate the highest accuracy score obtained from each classifier. [Adapted from [175]].

Database (Number of features)	**Artificial Neural Network**	**SVM**	**LGBM**	**XGboost**
A(30)	81.980	80.835	79.690	80.741
B(45)	88.754	85.549	86.774	86.626
C(75)	89.172	85.832	86.519	85.253
	Accuracy score (%)			

From Table 3 it is determined that the worst results were obtained for dataset A, regardless of the chosen machine learning classifier. Notably, dataset A is constructed by incorporating data from different polarimetric observables, without considering the raw MM elements. In contrast, dataset B is created by exclusively considering 15 normalized elements of the Mueller matrix. Lastly, dataset C is formed by combining all the features in datasets A and B (i.e., A + B). The inferior performance associated with dataset A implies that the raw information of the Mueller matrix elements is crucial for the development of classification models. This is demonstrated by the fact that, for all the studied classifiers (as seen in Table 3), datasets B and C improved the performance of models based on dataset A by approximately 6%. This suggests that the transformations applied to the MM elements, when calculating different polarimetric observables, result in a loss of vital information for classificatory model development. While polarimetric observables can be useful for physical interpretation of a sample in various scenarios, they may not be the optimal framework for classification purposes.

Furthermore, the Artificial Neural Network (ANN) and support vector machines (SVM) achieved the highest accuracy scores on dataset C, with percentages of 89.172% and 85.832%, in Table 3, respectively. Meanwhile, the light gradient boosting machine (LGBM) and extreme gradient boosting (XGBoost) exhibited the best performance on dataset B, with percentages of 86.774% and 86.626%, in Table 3, respectively. These results indicate that the use of normalized Mueller matrix elements alone or in combination with polarimetric observables may be beneficial depending on the chosen classifier. Notably, only ANN and SVM exhibit a slightly improved performance on dataset C (A + B) compared to dataset B, with an increase greater than 1%. However, in applications where high recognition rate is essential, even a small improvement may be significant. Thus, the use of raw Mueller matrix data in conjunction with transformed polarimetric data may be recommended. Additionally, the results of comparing various machine learning models for fixed datasets (shown in Table 3) indicate that the best performance is consistently achieved with an Artificial Neural Network (ANN) classifier. This suggests that among the machine learning algorithms evaluated in this study, ANNs are the most appropriate for classifying the studied tissues. Moreover, this result is noteworthy given the physiologically similarity between the tissue types corresponding to tendon, muscle and myotendinous junction (see the physiological description of soft tissues, in section 3.2). In particular, the myotendinous junction is a complex structure composed of skeletal muscle and tendon. The transition between muscle and tendon is characterized by a variable proportion of muscle and tendon fascicles. The varying proportion of muscle and tendon at the myotendinous junction highlights the potential for increased accuracy when investigating healthy and pathological tissue samples.

This last hypothesis is reinforced when studying the classifiers in terms of the confusion matrix. The confusion matrix [199] allows us to study the classification performance of our classifiers when comparing the predicted (columns) and real (rows) tissue classes. As a representative example, in Fig. 32 we present the confusion matrix for the ANN, which revealed that this classifier achieved high accuracy for classifying tendon (99.36%) and muscle tissue (90.49%), with lower accuracy for classifying myotendinous junction (80.89%) and bone tissue (88.89%).

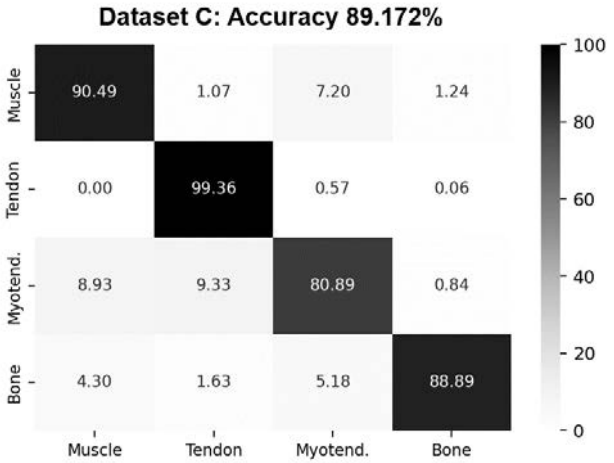


Figure 32: Confusion matrix and accuracy score obtained from testing the Artificial Neural Network with the dataset C, the merged (polarimetric observables + MM elements) – label C. [Reproduced from [175]].

So far, we have demonstrated that the inclusion of raw MM elements in polarimetric-based classification models is crucial for achieving optimal performance. This is significant as many proposals in literature omit the raw MM elements in favor of polarimetric observables, which provide physical interpretation but at the cost of lost information and decreased accuracy in classification. This suggests that the conversion from MM elements to polarimetric observables leads to loss of information that is detrimental to the accuracy of classificatory models. Therefore, we suggest to prioritize the use of databases with explicit information on MM elements for biological tissue classification tasks. Additionally, among the ML algorithms investigated in this work, the use of Artificial Neural Networks (ANN) is identified as the most appropriate approach for global classification tasks.

We now study the applicability of classification models for image classification of tissues. To test this, we construct an artificial image composed of four sub-images of 256×256 pixels each, representing muscle, tendon, myotendinous junction, and bone tissue. These images are randomly selected from tissue measurements not used for training the classifiers, and normalized to a maximum intensity value of 1. Figure 33 shows the resulting composition of these randomly selected tissue samples measured at 625 nm illumination wavelength.

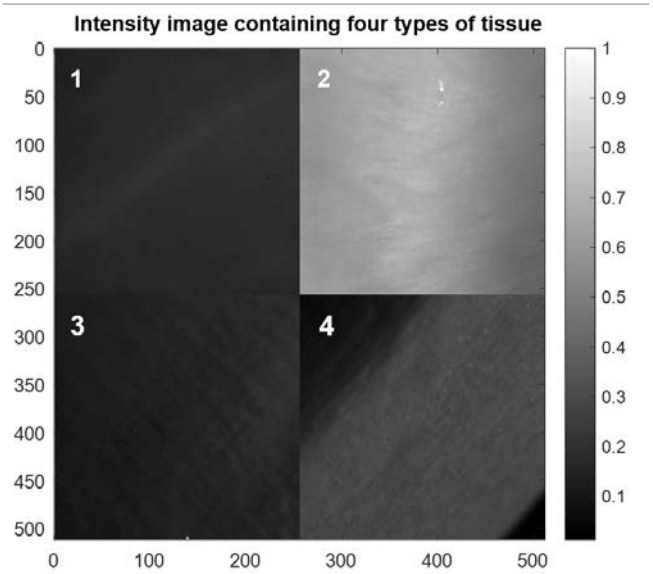


Figure 33: Visual composition containing the normalized intensity images corresponding to random samples of (1) muscle, (2) tendon, (3) myotendinous junction and (4) bone; at 625 nm. [Reproduced from [175]].

We want to conduct a pixel-by-pixel imaging classification test. For this purpose, we utilize the normalized Mueller matrix elements and polarimetric observables at three wavelengths to test the predictive capacity of the methods for tissue composition imaging as shown in Fig. 33. The models are constructed using mean values from 10 × 10 pixel sub-images of tissue. This selection is crucial for the successful recognition of pixel-to-pixel data, as it provides information that is not far from a single pixel. By properly grouping this obtained data, we generated three new polarimetric datasets (like the previous A, B and C), labeled as A', B' and C'. The trained classifiers were evaluated on datasets A', B' and C' to assess their classification performance and a 512 × 512 pixel image was created, with each pixel representing a tissue class, indicated by a corresponding color. This image was compared to the actual tissue image (Figure 33) for visual analysis. The imaging recognition was conducted using the ANN, SVM, LGBM and XGBoost classifiers. However, to not to extent the content, we show the results obtained for the XGBoost as a representative example in Fig. 34. Similar images were demonstrated for the three remaining models of ANN, SVM, and LGBM. It is important to note that Fig. 34 includes three images corresponding to the analysis based on the three datasets (A', B' and C'). The original image consist of four sub-images each one representing the different tissues (muscle, tendon, myotendinous junction, and bone in Fig. 33). The final classification image (Fig. 34) consists of four sectors, each one representing the performance of the ANN in classifying one of these four tissues. Tissue assignment is visualized with a particular color: purple (muscle),

blue (tendon), green (myotendinous junction), and yellow (bone).

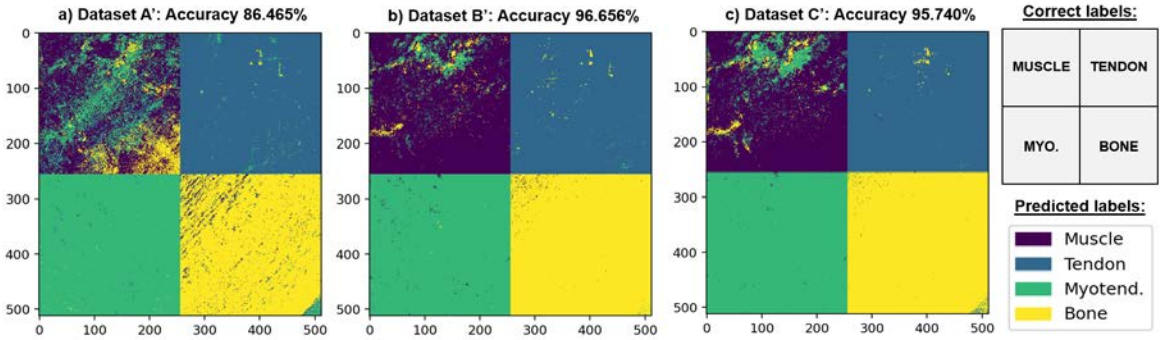


Figure 34: XGBoost imaging tissues classification performance and global accuracy scores of four random samples corresponding to muscle, tendon, myotendinous junction and bone. [Adapted from [175]].

Ideally, the resulting image would feature purple in the upper-left corner, blue in the upper-right corner, green in the bottom-left corner, and yellow in the bottom-right corner, indicating that all pixels in the image have been correctly classified. However, as depicted in Figure 34, there are instances of misclassification, particularly for the muscle class (upper-left block), where some pixels are classified as green (myotendinous junction) or yellow (bone). Conversely, the myotendinous and bone classes (bottom-left and bottom-right corners) exhibit the highest degree of accuracy. It is noteworthy that the original image of the bone (as depicted in the right-bottom sub-image of Figure 33) includes the presence of a background (holder) in which the bone sample is placed. The application of XGBoost model to this image allows for the identification of the edge of the bone sample, as demonstrated in Fig. 34a and 34c. This is achieved by classifying the holder pixels as an arbitrary tissue class (represented by the green small triangle located at the bottom-left of the images), resulting in a visual contrast. These findings highlight the need for proper recognition of the holder through the inclusion of a new class in the training of the models, utilizing polarimetric data specifically related to holder measurements. Furthermore, the muscle class is the most poorly classified, with some pixels being misclassified as bone or myotendinous junction. We suggest that the observed misclassifications in the mean values-based model, which utilizes averages taken over 10×10 pixel images, arise from its application in individual pixel-to-pixel analysis. It is important to note that this analysis was performed on a single example, whereas the models used in the study account for information from a much larger number of images ($50 \times 255 = 12{,}750$ in the case of muscle analysis). Furthermore, the global accuracy scores for the three datasets presented in Fig. 34 demonstrate consistent performance, with scores of 86.465% for the A' dataset, 96.656% for the B'

dataset, and 95.740% for the C' dataset. Notably, the B' dataset exhibited the best results among the three datasets tested (Figure 34b), which highlights the importance of including raw MM elements to improve classification model performance.

On the classification results obtained from the three remaining machine learning models of ANN, SVM and LGBM, they demonstrate that the inclusion of additional raw data (B' and C' datasets) improved the tissue classification performance. To give some numbers, the global accuracy scores for the ANN are of with scores of 88.135% for the A' dataset, 96.649% for the B' dataset, and 96.258% for the C' dataset. SVM model are of 78.785%, 92.703% and 95.736% for datasets A', B' and C', respectively. Regarding the LGBM, the global accuracy scores are of 87.577%, 97.279% and 95.891% for datasets A', B' and C', respectively. It worth to mention that the LGBM and XGBoost models demonstrate excellent classification of tendon, myotendinous junction, and bone tissues, with rates above 99%. However, the muscle tissue is the worst classified class. Overall, the results indicate that LGBM and XGBoost models are the best performing models for the composed image studied, providing similar results with classification differences smaller than 0.5%.

The results hereby presented demonstrate that, when applying the methods for pixel-to-pixel imaging classification, the inclusion of MM-raw data is crucial to achieve optimal classification of biological tissues.

5 Conclusions

The research presented in this thesis has provided a comprehensive examination of the use of polarimetric observables for the inspection of biological tissues. First, we have demonstrated the utility of polarimetric observables, specifically the Indices of Polarimetric Purity (IPPs) and the Components of Purity (P, D, and P_S), for identifying plant disease symptoms. Additionally, we have shown that the implementation of a pseudo-coloration image processing method, which utilizes these polarimetric observables, is an effective tool for enhancing the contrast between different tissue types (healthy, necrotic, and chlorotic) in botanical samples. Our results indicate that depolarizing metrics have the potential to characterize various plant pathologies, as demonstrated in 18 different plant species infected by various pathogens. For the purpose of this study, we have focused on two important agricultural crops: alfalfa (*Medicago sativa*) and olive (*Olea europaea*).

In the study of *M. sativa* leaves, the use of polarimetric imaging techniques, specifically the depolarization-related observables resulted in a significant enhancement of image contrast between healthy and diseased regions as compared to standard unpolar-

ized light intensity images. This improvement in visibility is evident in the observation of chlorotic spots and vascular structures, which were barely observable in the standard intensity images. Moreover, the use of visualization and contrast improvement techniques in polarimetric imaging allowed us for precise measurement of diseased regions in *M. sativa* plant sample. In the case of the *O. europaea* sample, polarimetric observables revealed an increased contrast compared to non-polarized transmission/scattering images, providing a more accurate visualization and spatial delimitation of lesions in the leaf. The use of polarimetric observables also unveiled the presence of vascular structures not visible in non-polarized transmission/scattering images. In addition, the width of the necrotic ring and the diameter of the chlorotic spot became clearly delimited in polarimetric images.

We also utilized polarimetric images to investigate the potential of characterizing disease symptoms in vegetal tissues. By collecting pixels from the original image and representing them in a data cloud figure, we aimed to determine if regions corresponding to different tissue conditions could be discriminated. The data clouds were generated using a three-dimensional representation of the Indices of Polarimetric Purity (IPPs) and the Components of Purity (CPs). The results from the analysis of a *M. sativa* leaf demonstrated clear discrimination between healthy and chlorotic regions, as these tissue conditions were practically spatially separated in both the IPPs and CPs space, indicating the potential utility of this approach in characterizing disease symptoms in vegetal tissues. The same analysis conducted on the *O. europaea* leaf tissue indicated that the depolarizing properties of both IPPs and CPs spaces can effectively discriminate between healthy, chlorotic, and necrotic tissue regions.

To take advantage of this discriminatory behavior, we employed the pseudo-coloration approach to enhance the contrast of polarimetric images for improved visual discrimination of features present in complex scattering media. Our method utilized triplets of polarimetric observables encoded in the R, G, and B color channels. To the best of our knowledge, this is the first use of the polarization based pseudo-coloration approach for the analysis of plant disease symptoms. We improved upon previous methods by incorporating an optimized selection of polarimetric observables within the IPPs and CPs spaces and by including an image filtering process based on data obtained from a Boxplot analysis. These modifications allowed for a more accurate discrimination between healthy and diseased tissue. In particular, the final pseudo-colored images for *M. sativa* and *O. europaea* demonstrate a visual enhancement of disease symptoms, leading to a more accurate location of the damaged area. Importantly, it is noteworthy that the pseudo-colored images enhance the contrast between different tissue conditions, even more than the performance of isolated imaging of po-

larimetric observables for *M. sativa* and *O. europaea* samples. This behavior highlights the suitability of using this image-processing method for analysis of biological samples. Specifically, it can be used for the inspection and estimation of direct lesions, and the characterization and early detection of infection processes on plant tissues.

As previously stated, recent studies have proposed the use of pseudo-colored images to enhance the visualization of tissue imaging, not only in plants, but also in biological tissues of animal origin. These images are constructed by layering different polarimetric features of the sample, allowing for the inclusion of multiple polarimetric information origins in a single image. The methods employed to construct these images typically involve a qualitative selection of polarimetric observables and their relative weights, resulting in a pseudo-colored model that may not be optimal. This suggests a need for further research to develop more robust and objective methods for creating these images. In this thesis, we introduced two novel pseudo-coloring models for optimized enhancement of the visual contrast of different tissues in biological samples. These models are based on the Euclidean distance and Normal (Gaussian) function, and utilize the Indices of Polarimetric Purity (IPPs) and Components of Purity (CPs) spaces to optimize tissue discrimination. Both methods involve associating different colors with specific tissue classes, resulting in pseudo-colored images that incorporate a range of polarimetric observables. These approaches outperform previous methods based on basic linear combinations of polarimetric observables, which were obtained using non-optimal and heuristic approaches. The Euclidean method was found to be effective in discriminating between two classes in the samples studied, but is limited to only two classes. On the other hand, the Normal-based approach could discriminate between more than two classes and provided a probabilistic interpretation of the data, making it useful as an automatic classifier. However, it is sensitive to the statistical distribution, variance, and presence of outliers in the data.

The obtained results demonstrate that these models outperform traditional polarimetric observables in terms of spatial location, visualization, and recognition of structures within the samples. It should be noted that these pseudo-colored functions are not limited to polarimetric spaces and can be applied to other approaches such as image segmentation and coloring. Further comparison is needed to determine the optimal scenario for pseudo-colored imaging. We have chosen to use the IPPs and CPs observables in our proposed methods as previous research has demonstrated their suitability for tissue discrimination. Our results indicate that the choice of observable basis, either IPPs or CPs, depends on the characteristics of the sample being analyzed. Therefore, we recommend using both the IPPs and CPs observables in order to achieve the best possible visualization of structures in tissue samples.

Polarimetric techniques have become increasingly prevalent in biomedical applications due to their ability to provide enhanced contrast of certain structures and reveal information that is not accessible through non-polarimetric methods, as demonstrated in this thesis. The utilization of polarimetric data in clinical applications has led to improved visualization of structures, and the integration of these techniques in automatic recognition models has facilitated the classification of tissues and early detection of certain pathologies.

With this in mind, we designed four predictive models based on polarimetric metrics derived from ex-vivo chicken tissue samples (bone, tendon, muscle and myotendinous junction tissue) for tissue recognition. Our results showed that the data was non-parametric and had significant potential for tissue discrimination. To construct a robust predictive model, we applied Principal Component Analysis on the data, which reduced the dimension of the data space while preserving important information. By analyzing the connection between the Principal Components and the original polarimetric metrics, we found that different variability directions in the principal components space can be associated with different physical origins of tissue structures.

We used principal component analysis to construct a logistic regression-based model for each of the four studied tissue categories. The models achieved high sensitivity and specificity values: 85.3% and 93.5% for the tendon model, 86.0 % and 88.8 % for the muscle model, 82.6 % and 80.6 % for the bone model, and 82.0 % and 71.0 % for the myotendinous junction model. The results provide a complementary visual interpretation for tissue recognition, with a satisfactory level of categorization as per the sensitivities and specificities stated earlier. The proposed non-invasive methods discussed in this study could be applied in various biomedical scenarios, such as early diagnosis of pathologies, with the appropriate statistical data.

The Mueller matrix (MM) is a widely utilized framework for analyzing the polarimetric response of samples. The MM encodes important polarimetric features such as dichroism, birefringence, and depolarization. To gain a physical understanding of a particular sample's features, a variety of polarimetric observables have been derived from experimental Mueller matrices in literature. These polarimetric observables are currently being utilized in various biomedical scenarios to train classification models. In the last work of this thesis, we studied the potential loss of classificatory information during the mathematical transformation from MM elements to derived polarimetric observables. Through an examination of the ideal set of polarimetric features for creating classificatory models, it was determined that the use of raw MM elements, MM-derived polarimetric observables, or a combination of both, may yield the most accurate results. This issue has not been previously reported in the literature and

highlights the importance of considering the preservation of classificatory information during the polarimetric feature extraction process. To do so, we measured the experimental Mueller matrices from a collection of four ex-vivo chicken tissue types (muscle, tendon, myotendinous junction and bone) at three different wavelengths (625, 530 and 470 nm) within the visible range. These measurements were used to create a polarimetric database from which three specific datasets were derived: A, B, and C datasets considering MM-derived observables (A), normalized MM-elements (B) and the combination of both (C). These datasets were utilized to train well-known machine learning classifiers, with a focus on four specific models: Artificial Neural Network (ANN), Support Vector Machine (SVM), Light Gradient Boosting Machine (LGBM) and Extreme Gradient Boosting (XGBoost). Each of these models were trained using the three datasets, resulting in 12 classification models which were then compared in terms of tissue classification performance. We evaluated the suitability of these models based on global accuracy scores. Results showed that the ANN classifier trained with dataset C, which included both MM raw data and polarimetric observables, yielded the highest global accuracy of 89.172%. This suggests that the inclusion of a mix of information improves the classification performance of the models. Additionally, dataset B (raw MM elements alone) also produced a high global accuracy of 88.754%. However, dataset A (MM-derived polarimetric observables alone) resulted in the lowest global accuracy of 81.980%. This trend was consistent across all the classification models tested. These findings emphasize the importance of incorporating raw MM elements in recognition models based on polarimetric data and have not been reported in previous literature. In addition, we also evaluated the performance of the classifiers by computing the corresponding confusion matrix. The results for the best-performing model (ANN with the C model) indicated that the tendon class was excellently classified, with an accuracy of 99.36%. Conversely, the myotendinous junction class was found to be the most challenging to classify, with a best result of 81.16% accuracy. This difficulty in classification can be attributed to the transitional nature of myotendinous junction tissue between muscle and tendon, resulting in misclassification with one of these two classes.

Finally, the potential for using the created models for image classification was evaluated. An artificial image composed of four different tissue images was constructed from the testing data and used to create three new datasets (A', B', and C'). The previously developed models were then applied to these datasets to predict the tissue class for each pixel of the new image. The results showed that the best proposed models achieved global accuracy scores higher than 97% with classification rates above 99% for tendon, myotendinous junction, and bone, as observed for XGBoost and LGBM models when applied to B' and C' datasets (i.e., including raw MM-elements). Among the four

tissues studied, muscle was the worst classified tissue, with the best results obtained using the ANN model (B' dataset), achieving a misclassification rate of less than 6%. The results suggest that optimal classification of different tissues can be achieved by using different algorithms and highlight the potential for creating ensemble methods that combine multiple machine learning algorithms for improved performance.

It is worth to mention that we chose to study chicken tissue due to its similarities to human tissue and the ease of obtaining samples. While further validation on human tissue is necessary, our findings suggest that these methods may have applications in biomedical assistance, such as the early detection of certain cancers or skin pathologies.

References

[1] J. Al-Khalili, "In retrospect: Book of optics," *Nature*, vol. 518, pp. 164–165, 2015. DOI: 10.1038/518164a.

[2] B. P. Abbott *et al.*, "Observation of gravitational waves from a binary black hole merger," *Phys. Rev. Lett.*, vol. 116, 6 2016. DOI: 10.1103/PhysRevLett.116.061102.

[3] D. Dolmans, D. Fukumura, and R. Jain, "Photodynamic therapy for cancer," *Nat. Rev. Cancer*, vol. 3, pp. 380–387, 2003. DOI: 10.1038/nrc1071.

[4] A. Kherlopian, T. Song, Q. Duan, *et al.*, "A review of imaging techniques for systems biology," *BMC. Syst. Biol.*, vol. 2, no. 74, 2008. DOI: 10.1186/1752-0509-2-74.

[5] R. Carriles, D. N. Schafer, K. E. Sheetz, *et al.*, "Invited review article: Imaging techniques for harmonic and multiphoton absorption fluorescence microscopy," *Rev. Sci. Instrum.*, vol. 80, no. 8, 2009. DOI: 10.1063/1.3184828.

[6] A. Misra, S. Rowley, J. Zhou, *et al.*, "Biofinder detects biological remains in green river fish fossils from eocene epoch at video speed," *Sci. Rep.*, vol. 12, p. 10 164, 2022. DOI: 10.1038/s41598-022-14410-8.

[7] S. Jacquemoud and S. Ustin, *Leaf Optical Properties*. Cambridge University Press, 2019.

[8] H. Zhang, D. Salo, D. M. Kim, *et al.*, "Penetration depth of photons in biological tissues from hyperspectral imaging in shortwave infrared in transmission and reflection geometries," *J. Biomed. Opt.*, vol. 21, no. 12, 2016. DOI: 10.1117/1.JBO.21.12.126006.

[9] H. F. Grahn and P. Geladi, *Techniques and Applications of Hyperspectral Image Analysis*. John Wiley and Sons, 2007.

[10] L. Rey-Barroso, F. J. Burgos-Fernández, X. Delpueyo, *et al.*, "Visible and extended near-infrared multispectral imaging for skin cancer diagnosis," *Sensors (Basel)*, vol. 18, no. 5, 2018. DOI: 10.3390/s18051441.

[11] A. Spreinat, G. Selvaggio, L. Erpenbeck, and S. Kruss, "Multispectral near infrared absorption imaging for histology of skin cancer," *J. Biophotonics*, vol. 13, no. 1, 2020. DOI: 10.1002/jbio.201960080.

[12] G. Lu, L. Halig, D. Wang, *et al.*, "Spectral-spatial classification for noninvasive cancer detection using hyperspectral imaging," *J. Biomed. Opt.*, vol. 19, no. 10, 2014. DOI: 10.1117/1.JBO.19.10.106004.

[13] H. Akbari, L. V. Halig, D. M. Schuster, *et al.*, "Hyperspectral imaging and quantitative analysis for prostate cancer detection," *J. Biomed. Opt.*, vol. 17, no. 7, 2012. DOI: 10.1117/1.JBO.17.7.076005.

[14] S. L. P. Aggarwal and F. A. Papay, "Applications of multispectral and hyperspectral imaging in dermatology," *Exp. Dermatol.*, 2022. DOI: 10.1111/exd.1462.

[15] A. Lowe, N. Harrison, and A. French, "Hyperspectral image analysis techniques for the detection and classification of the early onset of plant disease and stress," *Plant Methods*, vol. 13, no. 80, 2017. DOI: 10.1186/s13007-017-0233-z.

[16] C. Veys, F. Chatziavgerinos, A. AlSuwaidi, *et al.*, "Multispectral imaging for presymptomatic analysis of light leaf spot in oilseed rape," *Plant Methods*, vol. 15, no. 4, 2019. DOI: 10.1186/s13007-019-0389-9.

[17] R. Sarić, V. D. Nguyen, T. Burge, *et al.*, "Applications of hyperspectral imaging in plant phenotyping," *Trends. Plant. Sci.*, vol. 27, no. 3, DOI: 10.1016/j.tplants.2021.12.003.

[18] L. Ferrante di Ruffano, J. Dinnes, J. J. Deeks, *et al.*, "Optical coherence tomography for diagnosing skin cancer in adults," *The Cochrane database of systematic reviews*, vol. 12, no. 12, 2018. DOI: 10.1002/14651858.CD013189.

[19] L. van Manen, J. Dijkstra, C. Boccara, *et al.*, "The clinical usefulness of optical coherence tomography during cancer interventions," *J. Cancer Res. Clin. Oncol.*, vol. 144, no. 10, 2018. DOI: 10.1007/s00432-018-2690-9.

[20] R. F. Spaide, J. G. Fujimoto, N. K. Waheed, S. R. Sadda, and G. Staurenghi, "Optical coherence tomography angiography," *Prog. Ret. Eye Res.*, vol. 64, 2018. DOI: 10.1016/j.preteyeres.2017.11.003.

[21] Z. J. Guan, S. B. Zhang, K. Y. Guan, S. Y. Li, and H. Hu, "Leaf anatomical structures of paphiopedilum and cypripedium and their adaptive significance," *J. Plant Res.*, vol. 124, no. 2, 2011. DOI: 10.1007/s10265-010-0372-z.

[22] J. A. Terzakis, E. Santagada, A. Hernandez, and M. Taskin, "Scanning electron microscopy of peripheral blood smears: Comparison of normal blood with some common leukemias," *Ultra. pat.*, vol. 29, no. 1, 2005. DOI: 10.1080/01913120490903459.

[23] R. N. Burton-Smith and K. Murata, "Cryo-electron microscopy of the giant viruses," *Microscopy*, vol. 70, no. 6, 2021. DOI: 10.1093/jmicro/dfab036.

[24] C. J. R. Sheppard, "Microscopy: Phase contrast microscopy," *Encycl. Mod. Opt.*, vol. 3, 2004.

[25] Z. Kong, X. Zhu, S. Zhang, J. Wu, and Y. Luo, "Phase contrast microscopy of living cells within the whole lens: Spatial correlations and morphological dynamics," *Mol. vision*, vol. 18, 2012.

[26] A. Momose, W. Yashiro, K. Kido, *et al.*, "X-ray phase imaging: From synchrotron to hospital," *Philos. Trans. A Math. Phys. Eng. Sci.*, vol. 372, no. 2010, 2014. DOI: 10.1098/rsta.2013.0023.

[27] C. Karunakaran, R. Lahlali, N. Zhu, *et al.*, "Factors influencing real time internal structural visualization and dynamic process monitoring in plants using synchrotron-based phase contrast x-ray imaging," *Sci. Rep.*, vol. 5, 2015. DOI: 10.1038/srep12119.

[28] Y. A. Khristoforova, I. A. Bratchenko, O. O. Myakinin, *et al.*, "Portable spectroscopic system for in vivo skin neoplasms diagnostics by raman and autofluorescence analysis," *J. Biophotonics*, vol. 12, no. 4, 2019. DOI: 10.1002/jbio.201800400.

[29] K. J. I. Ember, M. A. Hoeve, S. L. McAughtrie, *et al.*, "Raman spectroscopy and regenerative medicine: A review," *npj. Regen. Med.*, vol. 2, no. 1, 2017. DOI: 10.1038/s41536-017-0014-3.

[30] T. Ichimura, L.-d. Chiu, K. Fujita, *et al.*, "Visualizing cell state transition using raman spectroscopy," *PLoS ONE*, vol. 9, no. 1, 2014. DOI: 10.1371/journal.pone.0084478.

[31] C. Bustamante, Y. Chemla, S. Liu, *et al.*, "Optical tweezers in single-molecule biophysics," *Nat. Rev. Meth. Primers*, vol. 1, no. 25, 2021. DOI: 10.1038/s43586-021-00021-6.

[32] T. A. Klar, S. Jakobs, M. Dyba, A. Egner, and S. W. Hell, "Fluorescence microscopy with diffraction resolution barrier broken by stimulated emission," *Proc. Natl. Acad. Sci.*, vol. 97, no. 15, 2000. DOI: 10.1073/pnas.97.15.8206.

[33] S. Berning, K. I. Willig, H. Steffens, P. Dibaj, and S. W. Hell, "Nanoscopy in a living mouse brain," *Science*, vol. 335, no. 6068, 2012. DOI: 10.1126/science.1215369.

[34] N. Kim, H. J. Kim, Y. Kim, K. S. Min, and S. K. Kim, "Direct and precise length measurement of single, stretched dna fragments by dynamic molecular combing and sted nanoscopy," *Analytical and bioanalytical chemistry*, vol. 408, no. 23, 2016. DOI: 10.1007/s00216-016-9764-9.

[35] E. Betzig, G. H. Patterson, R. Sougrat, *et al.*, "Imaging intracellular fluorescent proteins at nanometer resolution," *Science*, vol. 313, no. 5793, 2006. DOI: 10.1126/science.1127344.

[36] A. D. Elliott, "Confocal microscopy: Principles and modern practices," *Curr. Protoc. Cytom.*, vol. 92, no. 1, 2020. DOI: 10.1002/cpcy.68.

[37] J. Dinnes, J. J. Deeks, D. Saleh, *et al.*, "Reflectance confocal microscopy for diagnosing cutaneous melanoma in adults," *Coch. Data. Syst. Rev.*, vol. 12, no. 12, 2018. DOI: 10.1002/14651858.CD013190.

[38] A. J. Guerra, P. Cano, M. Rabionet, T. Puig, and J. Ciurana, "3d-printed pcl/pla composite stents: Towards a new solution to cardiovascular problems," *Materials*, vol. 11, no. 9, 2018. DOI: 10.3390/ma11091679.

[39] A. Belu, C. Mahoney, and K. Wormuth, "Chemical imaging of drug eluting coatings: Combining surface analysis and confocal raman microscopy," *J. Cont. Rel.*, vol. 126, no. 2, 2008. DOI: 10.1016/j.jconrel.2007.11.015.

[40] Y. Hontani, F. Xia, and C. Xu, "Multicolor three-photon fluorescence imaging with single-wavelength excitation deep in mouse brain," *Science Advances*, vol. 7, no. 12, 2021. DOI: 10.1126/sciadv.abf3531.

[41] M. Liebel, J. O. Arroyo, V. S. Beltrán, J. Osmond, A. Jo, H. Lee, R. Quidant, and N. F. van Hulst, "3d tracking of extracellular vesicles by holographic fluorescence imaging," *Science Advances*, vol. 6, no. 45, 2020. DOI: 10.1126/sciadv.abc2508.

[42] L. Donaldson, "Autofluorescence in plants," *Molecules*, vol. 10, no. 25, 2020. DOI: 10.3390/molecules25102393.

[43] C. P. Scotson, A. van Veelen, K. A. Williams, *et al.*, "Developing a system for in vivo imaging of maize roots containing iodinated contrast media in soil using synchrotron xct and xrf," *Plant and Soil*, vol. 460, 2021. DOI: 10.1007/s11104-020-04784-x.

[44] P. Hepler and B. Gunning, "Confocal fluorescence microscopy of plant cells," *Protoplasma*, no. 201, 1998. DOI: 10.1007/BF01287411.

[45] S. Nie, D. T. Chiu, and R. N. Zare, "Probing individual molecules with confocal fluorescence microscopy," *Science*, vol. 266, no. 5187, pp. 1018–1021, 1994. DOI: 10.1126/science.7973650.

[46] D. von Wangenheim, R. Hauschild, M. Fendrych, V. Barone, E. Benková, and J. Friml, "Live tracking of moving samples in confocal microscopy for vertically grown roots," *eLife*, vol. 6, 2017. DOI: 10.7554/eLife.26792.

[47] J. W. Lichtman and J. A. Conchello, "Fluorescence microscopy," *Nature methods*, vol. 2, no. 12, 2005. DOI: 10.1038/nmeth817.

[48] J. J. G. Pérez and R. Ossikovski, *Polarized Light and the Mueller Matrix Approach*. CRC Press, 2016.

[49] D. H. Goldstein, *Polarized Light*. CRC Press, 2011.

[50] J. H. Hough, "New opportunities for astronomical polarimetry," *Journal of Quantitative Spectroscopy and Radiative Transfer*, vol. 106, no. 1, pp. 122–132, 2007. DOI: https://doi.org/10.1016/j.jqsrt.2006.12.012.

[51] J. Hough, "Polarimetry: A powerful diagnostic tool in astronomy," *Astron. Geophys.*, vol. 47, no. 13, 2006. DOI: 10.1111/j.1468-4004.2006.47331.x.

[52] D. O. Suárez, *Polarimetric Observations of the Sun BT - Astronomical Polarisation from the Infrared to Gamma Rays*. Springer International Publishing, 2019.

[53] F. Snik and C. U. Keller, *Astronomical Polarimetry: Polarized Views of Stars and Planets BT - Planets, Stars and Stellar Systems*. Springer Netherlands, 2013.

[54] S. A. Hall, M.-A. Hoyle, J. S. Post, and D. K. Hore, "Combined stokes vector and mueller matrix polarimetry for materials characterization," *Analytical Chemistry*, vol. 85, no. 15, pp. 7613–7619, 2013. DOI: 10.1021/ac401864g.

[55] M. Hasan, K. Lyon, L. Trombley, C. Smith, and A. Zakhidov, "Thickness measurement of multilayer film stack in perovskite solar cell using spectroscopic ellipsometry," *AIP Advances*, vol. 9, no. 12, p. 125 107, 2019. DOI: 10.1063/1.5125686.

[56] L. M. S. Aas, P. G. Ellingsen, and M. Kildemo, "Near infra-red Mueller matrix imaging system and application to retardance imaging of strain," *Thin Solid Films*, vol. 519, no. 9, pp. 2737–2741, 2011. DOI: 10.1016/j.tsf.2010.12.093.

[57] C. L. Patty *et al.*, "Imaging linear and circular polarization features in leaves with complete mueller matrix polarimetry," *Biochimica et Biophysica Acta (BBA) - General Subjects*, vol. 1862, no. 6, pp. 1350–1363, 2018. DOI: https://doi.org/10.1016/j.bbagen.2018.03.005.

[58] J. S. Tyo, D. L. Goldstein, D. B. Chenault, and J. A. Shaw, "Review of passive imaging polarimetry for remote sensing applications," *Appl. Opt.*, vol. 45, no. 22, pp. 5453–5469, 2006. DOI: 10.1364/AO.45.005453.

[59] J. Iannarilli Frank J., H. E. Scott, and S. H. Jones, "Passive IR polarimetric hyperspectral imaging contributions to multisensor humanitarian demining," in *Detection and Remediation Technologies for Mines and Minelike Targets VI*, ser. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, vol. 4394, Oct. 2001, pp. 346–352. DOI: 10.1117/12.445486.

[60] R. Hegedüs, A. Barta, B. Bernáth, V. B. Meyer-Rochow, and G. Horváth, "Imaging polarimetry of forest canopies: How the azimuth direction of the sun, occluded by vegetation, can be assessed from the polarization pattern of the sunlit foliage," *Appl. Opt.*, vol. 46, no. 23, pp. 6019–6032, 2007. DOI: 10.1364/AO.46.006019.

[61] O. Dubovik *et al.*, "Polarimetric remote sensing of atmospheric aerosols: Instruments, methodologies, results, and perspectives," *J. Quant. Spectrosc. Radiat. Transfer*, vol. 224, pp. 474–511, 2019. DOI: 10.1016/j.jqsrt.2018.11.024.

[62] S. Ding, J. Wang, and X. Xu, "Polarimetric remote sensing in oxygen a and b bands: Sensitivity study and information content analysis for vertical profile of aerosols," *Atmospheric Measurement Techniques*, vol. 9, no. 5, pp. 2077–2092, 2016. DOI: 10.5194/amt-9-2077-2016.

[63] Z. Kong *et al.*, "A polarization-sensitive imaging lidar for atmospheric remote sensing," *Journal of Quantitative Spectroscopy and Radiative Transfer*, vol. 271, p. 107 747, 2021, ISSN: 0022-4073. DOI: https://doi.org/10.1016/j.jqsrt.2021.107747.

[64] J. L. Deuzé *et al.*, "Remote sensing of aerosols over land surfaces from polder-adeos-1 polarized measurements," *Journal of Geophysical Research: Atmospheres*, vol. 106, no. D5, pp. 4913–4926, 2001. DOI: https://doi.org/10.1029/2000JD900364.

[65] F. Vasefi, N. MacKinnon, R. Saager, *et al.*, "Polarization-sensitive hyperspectral imaging in vivo: A multimode dermoscope for skin analysis," *Sci. Rep.*, vol. 4, 2014. DOI: 10.1038/srep04924.

[66] E. Gnanatheepam, U. Kanniyappan, K. Dornadula, A. Prakasarao, and G. Singaravelu, "Polarization gating technique extracts depth resolved fluorescence redox ratio in oral cancer diagnostics," *Photodiagnosis Photodyn Ther.*, vol. 30, p. 101 757, 2020. DOI: 10.1016/j.pdpdt.2020.101757.

[67] A. N. Yaroslavsky, X. Feng, A. Muzikansky, and M. R. Hamblin., "Fluorescence polarization of methylene blue as a quantitative marker of breast cancer at the cellular level," *Sci. Rep.*, vol. 9, no. 1, p. 940, 2019. DOI: 10.1038/s41598-018-38265-0.

[68] V. V. Tuchin, *Tissue Optics: Light Scattering Methods and Instruments for Medical Diagnosis*. SPIE Press, 2015.

[69] T. Yasui, Y. Tohno, and T. Araki, "Characterization of collagen orientation in human dermis by two-dimensional second-harmonic-generation polarimetry," *Journal of Biomedical Optics*, vol. 9, no. 2, pp. 259–264, 2004. DOI: 10.1117/1.1644116.

[70] K. Reiser, P. Stoller, and A. Knoesen, "Three-dimensional geometry of collagenous tissues by second harmonic polarimetry," *Sci. Rep.*, vol. 7, p. 2642, 2017. DOI: 10.1038/s41598-017-02326-7.

[71] N. Mazumder and F. J. Kao, "Stokes polarimetry-based second harmonic generation microscopy for collagen and skeletal muscle fiber characterization," *Lasers Med. Sci.*, vol. 36, pp. 1161–1167, 2020. DOI: 10.1007/s10103-020-03144-6.

[72] A. D. Silva, C. Deumié, and I. Vanzetta, "Elliptically polarized light for depth resolved optical imaging," *Biomed. Opt. Express*, vol. 3, no. 11, pp. 2907–2915, 2012. DOI: 10.1364/BOE.3.002907.

[73] M. P. Siegel, Y. L. Kim, H. K. Roy, R. K. Wali, and V. Backman, "Assessment of blood supply in superficial tissue by polarization-gated elastic light-scattering spectroscopy," *Appl. Opt.*, vol. 45, no. 2, pp. 335–342, 2006. DOI: 10.1364/AO.45.000335.

[74] B. Yang *et al.*, "Instant polarized light microscopy for imaging collagen microarchitecture and dynamics," *Journal of Biophotonics*, vol. 14, no. 2, e202000326, 2021. DOI: https://doi.org/10.1002/jbio.202000326.

[75] I. Shtein *et al.*, "Stomatal cell wall composition: distinctive structural patterns associated with different phylogenetic groups," *Annals of Botany*, vol. 119, no. 6, pp. 1021–1033, 2017. DOI: 10.1093/aob/mcw275.

[76] P. G. Ellingsen, M. B. Lilledahl, L. M. S. Aas, C. de Lange Davies, and M. Kildemo, "Quantitative characterization of articular cartilage using Mueller matrix imaging and multiphoton microscopy," *Journal of Biomedical Optics*, vol. 16, no. 11, p. 116 002, 2011. DOI: 10.1117/1.3643721.

[77] F. Fanjul-Vélez, N. Ortega-Quijano, and J. L. Arce-Diego, "Polarimetry group theory analysis in biological tissue phantoms by mueller coherency matrix," *Optics Communications*, vol. 283, no. 22, pp. 4525–4530, 2010. DOI: https://doi.org/10.1016/j.optcom.2010.04.074.

[78] J. C. Ramella-Roman, S. A. Mathews, H. Kandimalla, A. Nabili, D. D. Duncan, S. A. D'Anna, S. M. Shah, and Q. D. Nguyen, "Measurement of oxygen saturation in the retina with a spectroscopic sensitive multi aperture camera," *Opt. Express*, vol. 16, no. 9, pp. 6170–6182, 2008. DOI: 10.1364/OE.16.006170.

[79] V. N. D. Le, I. Saytashev, S. Saha, P. F. Lopez, M. Laughrey, and J. C. Ramella-Roman, "Depth-resolved mueller matrix polarimetry microscopy of the rat cornea," *Biomed. Opt. Express*, vol. 11, no. 10, pp. 5982–5994, 2020. DOI: 10.1364/BOE.402201.

[80] K. M. Twietmeyer, R. A. Chipman, A. E. Elsner, Y. Zhao, and D. VanNasdale, "Mueller matrix retinal imager with optimized polarization conditions," *Opt. Express*, vol. 16, no. 26, pp. 21 339–21 354, 2008. DOI: 10.1364/OE.16.021339.

[81] J. M. Bueno and P. Artal, "Polarization and retinal image quality estimates in the human eye," *J. Opt. Soc. Am. A*, vol. 18, no. 3, pp. 489–496, 2001. DOI: 10.1364/JOSAA.18.000489.

[82] J. M. Bueno, "Measurement of parameters of polarization in the living human eye using imaging polarimetry," *Vision Research*, vol. 40, no. 28, pp. 3791–3799, 2000. DOI: https://doi.org/10.1016/S0042-6989(00)00220-0.

[83] O. K. Naoun, V. L. Dorr, P. Allé, J.-C. Sablon, and A.-M. Benoit, "Exploration of the retinal nerve fiber layer thickness by measurement of the linear dichroism," *Appl. Opt.*, vol. 44, no. 33, pp. 7074–7082, 2005. DOI: 10.1364/AO.44.007074.

[84] M. Hasan, K. Lyon, L. Trombley, C. Smith, and A. Zakhidov, "Thickness measurement of multilayer film stack in perovskite solar cell using spectroscopic ellipsometry," *AIP Advances*, vol. 9, no. 12, p. 125 107, 2019. DOI: 10.1063/1.5125686.

[85] L. Graham, Y. Yitzhaky, and I. Abdulhalim, "Classification of skin moles from optical spectropolarimetric images: a pilot study," *Journal of Biomedical Optics*, vol. 18, no. 11, p. 111 403, 2013. DOI: 10.1117/1.JBO.18.11.111403.

[86] S. L. Jacques, S. Roussel, and R. V. Samatham, "Polarized light imaging specifies the anisotropy of light scattering in the superficial layer of a tissue," *Journal of Biomedical Optics*, vol. 21, no. 7, p. 071 115, 2016. DOI: 10.1117/1.JBO.21.7.071115.

[87] D. Fricke, M. Wollweber, and B. Roth, "Mueller matrix measurement system for skin polarimetry as additional module for non-contact dermatoscopy," *2019 Conference on Lasers and Electro-Optics Europe and European Quantum Electronics Conference*, 2019.

[88] Y. Dong *et al.*, "A quantitative and non-contact technique to characterise microstructural variations of skin tissues during photo-damaging process based on Mueller matrix polarimetry," *Sci. Rep.*, vol. 7, p. 14 702, 2017. DOI: 10.1038/s41598-017-14804-z.

[89] A. Lizana, A. Van Eeckhout, K. Adamczyk, C. Rodríguez, J. C. Escalera, E. Garcia-Caurel, I. Moreno, and J. Campos, "Polarization gating based on mueller matrices," *J. Biomed. Optics*, vol. 22, no. 5, 2017. DOI: 10.1117/1.JBO.22.5.056004.

[90] J. C. Ramella-Roman, S. A. Prahl, and S. L. Jacques, "Three monte carlo programs of polarized light transport into scattering media: Part i," *Opt. Express*, vol. 13, no. 12, pp. 4420–4438, 2005. DOI: 10.1364/OPEX.13.004420.

[91] J. C. Ramella-Roman, S. A. Prahl, and S. L. Jacques, "Three monte carlo programs of polarized light transport into scattering media: Part ii," *Opt. Express*, vol. 13, no. 25, pp. 10 392–10 405, 2005. DOI: 10.1364/OPEX.13.010392.

[92] Y. Zhang, B. Chen, and D. Li, "Propagation of polarized light in the biological tissue: A numerical study by polarized geometric monte carlo method," *Appl. Opt.*, vol. 55, no. 10, pp. 2681–2691, 2016. DOI: 10.1364/AO.55.002681.

[93] E. DU *et al.*, "Characteristic features of mueller matrix patterns for polarization scattering model of biological tissues," *Journal of Innovative Optical Health Sciences*, vol. 07, no. 01, p. 1 350 028, 2014. DOI: 10.1142/S1793545813500284.

[94] M. Dubreuil, F. Tissier, S. Rivet, and Y. L. Grand, "Linear diattenuation imaging of biological tissues with near infrared mueller scanning microscopy," *Biomed. Opt. Express*, vol. 12, no. 1, pp. 41–54, 2021. DOI: 10.1364/BOE.408354.

[95] S. K. U., K. Mahato, and N. Mazumder, "Polarization-resolved Stokes-Mueller imaging: a review of technology and applications," *Lasers Med Sci*, vol. 34, pp. 1283–1293, 2019. DOI: 10.1007/s10103-019-02752-1.

[96] S. Kaniyala Melanthota *et al.*, "Types of spectroscopy and microscopy techniques for cancer diagnosis: A review," *Lasers Med Sci*, 2022. DOI: https://doi.org/10.1007/s10103-022-03610-3.

[97] P. Schucht *et al.*, "Visualization of white matter fiber tracts of brain tissue sections with wide-field imaging mueller polarimetry," *IEEE Transactions on Medical Imaging*, vol. 39, no. 12, pp. 4376–4382, 2020. DOI: 10.1109/TMI.2020.3018439.

[98] Y. Wang *et al.*, "Mueller matrix microscope: a quantitative tool to facilitate detections and fibrosis scorings of liver cirrhosis and cancer tissues," *Journal of Biomedical Optics*, vol. 21, no. 7, p. 071 112, 2016. DOI: 10.1117/1.JBO.21.7.071112.

[99] M. Kupinski *et al.*, "Polarimetric measurement utility for pre-cancer detection from uterine cervix specimens," *Biomed. Opt. Express*, vol. 9, no. 11, pp. 5691–5702, 2018. DOI: 10.1364/BOE.9.005691.

[100] M. Sun *et al.*, "Characterizing the microstructures of biological tissues using mueller matrix and transformed polarization parameters," *Biomed. Opt. Express*, vol. 5, no. 12, pp. 4223–4234, 2014. DOI: 10.1364/BOE.5.004223.

[101] D. L. Le, D. T. Nguyen, T. H. Le, Q.-H. Phan, and T.-T.-H. Pham, "Characterization of healthy and cancerous human skin tissue utilizing stokes–mueller polarimetry technique," *Optics Communications*, vol. 480, p. 126 460, 2021, ISSN: 0030-4018. DOI: 10.1016/j.optcom.2020.126460.

[102] E. Du *et al.*, "Mueller matrix polarimetry for differentiating characteristic features of cancerous tissues," *Journal of Biomedical Optics*, vol. 19, no. 7, p. 076 013, 2014. DOI: 10.1117/1.JBO.19.7.076013.

[103] A. Pierangelo *et al.*, "Ex-vivo characterization of human colon cancer by mueller polarimetric imaging," *Opt. Express*, vol. 19, no. 2, pp. 1582–1593, 2011. DOI: 10.1364/OE.19.001582.

[104] ——, "Polarimetric imaging of uterine cervix: A case study," *Opt. Express*, vol. 21, no. 12, pp. 14 120–14 130, 2013. DOI: 10.1364/OE.21.014120.

[105] C. He *et al.*, "Characterizing microstructures of cancerous tissues using multispectral transformed mueller matrix polarization parameters," *Biomed. Opt. Express*, vol. 6, no. 8, pp. 2934–2945, 2015. DOI: 10.1364/BOE.6.002934.

[106] J. Rehbinder *et al.*, "Ex vivo Mueller polarimetric imaging of the uterine cervix: a first statistical evaluation," vol. 21, no. 7, p. 071 113, 2016. DOI: 10.1117/1.JBO.21.7.071113.

[107] B. Kunnen *et al.*, "Application of circularly polarized light for non-invasive diagnosis of cancerous tissues and turbid tissue-like scattering media," *Journal of Biophotonics*, vol. 8, no. 4, pp. 317–323, 2015. DOI: https://doi.org/10.1002/jbio.201400104.

[108] L. Graham, Y. Yitzhaky, and I. Abdulhalim, "Classification of skin moles from optical spectropolarimetric images: a pilot study," *Journal of Biomedical Optics*, vol. 18, no. 11, p. 111 403, 2013. DOI: 10.1117/1.JBO.18.11.111403.

[109] R. Patel, A. Khan, R. Quinlan, and A. N. Yaroslavsky, "Polarization-Sensitive Multimodal Imaging for Detecting Breast Cancer," *Cancer Research*, vol. 74, no. 17, pp. 4685–4693, 2014. DOI: 10.1158/0008-5472.CAN-13-2411.

[110] D. Ivanov *et al.*, "Colon cancer detection by using poincaré sphere and 2d polarimetric mapping of ex vivo colon samples," *Journal of Biophotonics*, vol. 13, no. 8, e202000082, 2020. DOI: https://doi.org/10.1002/jbio.202000082.

[111] ——, "Polarization-based histopathology classification of ex vivo colon samples supported by machine learning," *Frontiers in Physics*, vol. 9, 2022. DOI: 10.3389/fphy.2021.814787.

[112] ——, "Polarization and depolarization metrics as optical markers in support to histopathology of ex vivo colon tissue," *Biomed. Opt. Express*, vol. 12, no. 7, pp. 4560–4572, 2021. DOI: 10.1364/BOE.426713.

[113] Y. Huang *et al.*, "Identification of serous ovarian tumors based on polarization imaging and correlation analysis with clinicopathological features," *Journal of Innovative Optical Health Sciences*, DOI: 10.1142/S1793545822410024.

[114] T. Marvdashti, L. Duan, S. Z. Aasi, J. Y. Tang, and A. K. E. Bowden, "Classification of basal cell carcinoma in human skin using machine learning and quantitative features captured by polarization sensitive optical coherence tomography," *Biomed. Opt. Express*, vol. 7, no. 9, pp. 3721–3735, 2016. DOI: 10.1364/BOE.7.003721.

[115] I. Lihacova *et al.*, "Multi-class cnn for classification of multispectral and autofluorescence skin lesion clinical images," *Journal of Clinical Medicine*, vol. 11, no. 10, 2022. DOI: 10.3390/jcm11102833.

[116] J. Wan *et al.*, "Polarization-based probabilistic discriminative model for quantitative characterization of cancer cells," *Biomed. Opt. Express*, vol. 13, no. 6, pp. 3339–3354, 2022. DOI: 10.1364/BOE.456649.

[117] X. Yang *et al.*, "Deep learning for denoising in a mueller matrix microscope," *Biomed. Opt. Express*, vol. 13, no. 6, pp. 3535–3551, 2022. DOI: 10.1364/BOE.457219.

[118] Y. Yao, M. Zuo, Y. Dong, L. Shi, Y. Zhu, L. Si, X. Ye, and H. Ma, "Polarization imaging feature characterization of different endometrium phases by machine learning," *OSA Continuum*, vol. 4, no. 6, pp. 1776–1791, 2021. DOI: 10.1364/OSAC.414109.

[119] Y. Zhu, Y. Dong, Y. Yao, L. Si, Y. Liu, H. He, and H. Ma, "Probing layered structures by multi-color backscattering polarimetry and machine learning," *Biomed. Opt. Express*, vol. 12, no. 7, pp. 4324–4339, 2021. DOI: 10.1364/BOE.425614.

[120] T. M. Mitchell, *Machine Learning*. McGraw Hill, 1997.

[121] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural Networks*, vol. 61, 2015. DOI: 10.1016/j.neunet.2014.09.003.

[122] S. Gajjar, M. Kulahci, and A. Palazoglu, "Selection of non-zero loadings in sparse principal component analysis," *Chemometrics and Intelligent Laboratory Systems*, vol. 162, pp. 160–171, 2017. DOI: https://doi.org/10.1016/j.chemolab.2017.01.018.

[123] S. de la and F. Fernandez, *Regresión Logística*. Facultad de Ciencias Económicas y Empresariales, Universidad Autónoma de Madrid, 2011.

[124] L. Breiman, "Random forests," *Mach Learn*, vol. 45, 2001. DOI: 10.1023/A:1010933404324.

[125] C. Cortes and V. Vapnik, "Support-vector networks," *Mach Learn*, vol. 20, 1995. DOI: 10.1007/BF00994018.

[126] A. Van Eeckhout *et al.*, "Depolarization metric spaces for biological tissues classification," *Journal of Biophotonics*, vol. 13, no. 8, e202000083, 2020. DOI: https://doi.org/10.1002/jbio.202000083.

[127] ——, "Depolarizing metrics for plant samples imaging," *PLOS ONE*, vol. 14, no. 3, pp. 1–19, Mar. 2019. DOI: 10.1371/journal.pone.0213909.

[128] ——, "Polarimetric imaging microscopy for advanced inspection of vegetal tissues," *Sci. Rep.*, vol. 11, p. 3913, 2021. DOI: 10.1038/s41598-021-83421-8.

[129] Z. Zhuo, H. Wang, R. Liao, and H. Ma, "Machine learning powered microalgae classification by use of polarized light scattering data," *Applied Sciences*, vol. 12, no. 7, 2022. DOI: 10.3390/app12073422.

[130] T. Nguyen, M. Kandel, M. Rubessa, M. Wheeler, and G. Popescu, "Gradient light interference microscopy for 3d imaging of unlabeled specimens," *Nat. Comm.*, vol. 210, 2017. DOI: 10.1038/s41467-017-00190-7.

[131] L. Grant, C. S. T. Daughtry, and V. C. Vanderbilt, "Polarized and non-polarized leaf reflectances of coleus blumei," *Environmental and Experimental Botany*, vol. 27, pp. 139–145, 1987.

[132] V. C. Vanderbilt, C. S. T. Daughtry, and L. L. Biehl, "Is there spectral variation is the polarized reflectance of leaves?" In *Polarization: Measurement, Analysis, and Remote Sensing XI*, vol. 9099, 2014.

[133] T. N. Tóth *et al.*, "Fingerprinting the macro-organisation of pigment-protein complexes in plant thylakoid membranes in vivo by circular-dichroism spectroscopy," *Biochim Biophys Acta*, vol. 1857, no. 9, pp. 1479–1489, 2016. DOI: 10.1016/j.bbabio.2016.04.287.

[134] Z. E and H. P. K., "Production of guard cell protoplasts from onion and tobacco," *Plant Physiol.*, vol. 58, no. 4, 1976. DOI: 10.1104/pp.58.4.492.

[135] K. Sauer and M. Calvin, "Molecular orientation in quantasomes: I. electric dichroism and electric birefringence of quantasomes from spinach chloroplasts," *Journal of Molecular Biology*, vol. 4, no. 6, pp. 451–466, 1962. DOI: https://doi.org/10.1016/S0022-2836(62)80102-8.

[136] V. Bischoff *et al.*, "Trichome birefringence and its homolog at5g01360 encode plant-specific duf231 proteins required for cellulose biosynthesis in arabidopsis," *Plant Physiology*, vol. 153, no. 2, pp. 590–602, 2010. DOI: 10.1104/pp.110.153320.

[137] C.-J. Gao, X.-J. Xia, K. Shi, Y.-H. Zhou, and J. Yu, "Response of stomata to global climate changes and the underlying regulation mechanism of stress responses," *Zhiwu Shengli Xuebao/Plant Physiology Journal*, vol. 48, pp. 19–28, 2012.

[138] B. Palevitz and P. Hepler, "Cellulose microfibril orientation and cell shaping in developing guard cells of allium: The role of microtubules and ion accumulation," *Planta*, vol. 132, pp. 71–93, 1976.

[139] R. L. Long, M. P. Bange, S. G. Gordon, and G. A. Constable, "Measuring the maturity of developing cotton fibers using an automated polarized light microscopy technique," *Textile Research Journal*, vol. 80, no. 5, pp. 463–471, 2010.

[140] V. Vanderbilt, L. Grant, and C. Daughtry, "Polarization of light scattered by vegetation," *Proceedings of the IEEE*, vol. 73, no. 6, pp. 1012–1024, 1985. DOI: 10.1109/PROC.1985.13232.

[141] V. Kharuk and V. Yegorov, "Polarimetric indication of plant stress," *Remote Sensing of Environment*, vol. 33, no. 1, pp. 35–40, 1990. DOI: https://doi.org/10.1016/0034-4257(90)90053-0.

[142] G. Rondeaux and M. Herman, "Polarization of light reflected by crop canopies," *Remote Sensing of Environment*, vol. 38, no. 1, pp. 63–75, 1991. DOI: https://doi.org/10.1016/0034-4257(91)90072-E.

[143] L. Grant, C. Daughtry, and V. Vanderbilt, "Polarized and specular reflectance variation with leaf surface features," *Physiologia Plantarum*, vol. 88, no. 1, pp. 1–9, 1993. DOI: https://doi.org/10.1111/j.1399-3054.1993.tb01753.x.

[144] M. Lindgren, "Discriminating land mines from natural backgrounds by depolarization," 2008.

[145] S. V. Berdyugina, J. R. Kuhn, D. M. Harrington, T. Šantl-Temkiv, and E. J. Messersmith, "Remote sensing of life: polarimetric signatures of photosynthetic pigments as sensitive biomarkers," *International Journal of Astrobiology*, vol. 15, no. 1, pp. 45–56, 2016. DOI: 10.1017/S1473550415000129.

[146] K. Arai and Y. Nishimura, "Polarization model for discrimination of broad and needle shaped leaves and estimation of LAI using polarization measurements," *Advances in Space Research*, vol. 44, no. 4, pp. 510–516, Aug. 2009. DOI: 10.1016/j.asr.2009.04.026.

[147] S. Savenkov, R. Muttiah, E. Oberemok, A. Priezzhev, I. Kolomiets, and A. Klimov, "Measurement and interpretation of mueller matrices of barley leaves," *Quantum Electronics*, vol. 50, no. 1, pp. 55–60, 2020. DOI: 10.1070/qel17178.

[148] P. J. Curran, "Polarized visible light as an aid to vegetation classification," *Remote Sensing of Environment*, vol. 12, no. 6, pp. 491–499, 1982. DOI: https://doi.org/10.1016/0034-4257(82)90023-2.

[149] C. Rodríguez *et al.*, "Indices of polarimetric purity: application in biological tissues," in *Polarized Light and Optical Angular Momentum for Biomedical Diagnostics*, vol. 11646, SPIE, 2021, 116460P. DOI: 10.1117/12.2578254.

[150] S. Savary, A. Ficke, J.-N. Aubertot, and C. A. Hollier, "Crop losses due to diseases and their implications for global food production losses and food security," *Food Security*, vol. 4, pp. 519–537, 2012.

[151] E.-C. OERKE, "Crop losses to pests," *The Journal of Agricultural Science*, vol. 144, no. 1, pp. 31–43, 2006. DOI: 10.1017/S0021859605005708.

[152] R. C. Jones, "A new calculus for the treatment of optical systems. description and discussion of the calculus," *J. Opt. Soc. Am.*, vol. 31, no. 7, 1941. DOI: 10.1364/JOSA.31.000488.

[153] G. G. Stokes, "On the composition and resolution of streams of polarized light from different sources," *Tra. Cam. Phi. Soc.*, no. 9, 1851.

[154] H. Mueller, "The foundation of optics," *J. Opt. Soc. Am.*, vol. 38, 1948.

[155] P. Li, D. Lv, H. He, and H. Ma, "Separating azimuthal orientation dependence in polarization measurements of anisotropic media," *Opt. Express*, vol. 26, no. 4, 2018. DOI: 10.1364/OE.26.003791.

[156] I. Ahmad, M. Ahmad, K. Khan, and M. Ikram, "Polarimetry based partial least square classification of ex vivo healthy and basal cell carcinoma human skin tissues," *Photodiag. Photody. Ther.*, vol. 14, 2016. DOI: 10.1016/j.pdpdt.2016.04.004.

[157] S. Y. Lu and R. A. Chipman, "Interpretation of mueller matrices based on polar decomposition," *J. Opt. Soc. Am. A.*, vol. 13, no. 5, 1996. DOI: 10.1364/JOSAA.13001106.

[158] M. F. W. N. Ghosh and I. A. Vitkin, "Influence of the order of the constituent basis matrices of the mueller matrix decomposition-derived polarization parameters in complex turbid media such as biological tissues," *Opt. Commun.*, vol. 283, no. 6, 2010. DOI: 10.1016/j.optcom.2009.10.111.

[159] K. S. P. Babilotte and V. N. Silva, "Retardance properties of twisted ferroelectric liquid crystal," *Optik*, vol. 125, no. 14, 2014. DOI: 10.1016/j.ijleo.2014.01.070.

[160] A. V. Eeckhout, A. Lizana, E. Garcia-Caurel, J. J. Gil, R. Ossikovski, and J. Campos, "Synthesis and characterization of depolarizing samples based on the indices of polarimetric purity," *Opt. Lett.*, vol. 42, no. 20, pp. 4155–4158, 2017. DOI: 10.1364/OL.42.004155.

[161] A. Tariq, H. He, P. Li, and H. Ma, "Purity-depolarization relations and the components of purity of a mueller matrix," *Opt. Express*, vol. 27, no. 16, pp. 22 645–22 662, 2019. DOI: 10.1364/OE.27.022645.

[162] R. Ossikovski and J. Vizet, "Eigenvalue-based depolarization metric spaces for mueller matrices," *J. Opt. Soc. Am. A*, vol. 36, no. 7, pp. 1173–1186, 2019. DOI: 10.1364/JOSAA.36.001173.

[163] C. J. R. Sheppard, A. Bendandi, A. L. Gratiet, and A. Diaspro, "Eigenvectors of polarization coherency matrices," *J. Opt. Soc. Am. A*, vol. 37, no. 7, pp. 1143–1154, 2020. DOI: 10.1364/JOSAA.391902.

[164] M. Borovkova, M. Peyvasteh, O. Dubolazov, *et al.*, "Complementary analysis of mueller-matrix images of optically anisotropic highly scattering biological tissues," *J. Eur. Opt. Soc.-Rapid Publ.*, vol. 20, p. 20, 2018. DOI: 10.1186/s41476-018-0085-9.

[165] M. Borovkova, A. Bykov, A. Popov, A. Pierangelo, T. Novikova, J. Pahnke, and I. Meglinski, "Evaluating beta-amyloidosis progression in alzheimer's disease with mueller polarimetry," *Biomed. Opt. Express*, vol. 11, no. 8, pp. 4509–4519, 2020. DOI: 10.1364/BOE.396294.

[166] V. A. Ushenko, B. T. Hogan, A. Dubolazov, *et al.*, "Embossed topographic depolarisation maps of biological tissues with different morphological structures," *Biomed. Opt. Express*, no. 11, p. 3871, 2021. DOI: 10.1038/s41598-021-83017-2.

[167] ——, "3d mueller matrix mapping of layered distributions of depolarisation degree for analysis of prostate adenoma and carcinoma diffuse tissues," *Biomed. Opt. Express*, no. 11, p. 5162, 2021. DOI: 10.1038/s41598-021-83986-4.

[168] I. S. José and J. J. Gil, "Invariant indices of polarimetric purity: Generalized indices of purity for n×n covariance matrices," *Optics Communications*, vol. 284, no. 1, pp. 38–47, 2011. DOI: https://doi.org/10.1016/j.optcom.2010.08.077.

[169] Gil, J. J., "Polarimetric characterization of light and media - physical quantities involved in polarimetric phenomena," *Eur. Phys. J. Appl. Phys.*, vol. 40, no. 1, pp. 1–47, 2007. DOI: 10.1051/epjap:2007153.

[170] J. J. Gil, "Review on Mueller matrix algebra for the analysis of polarimetric measurements," *Journal of Applied Remote Sensing*, vol. 8, no. 1, p. 081 599, 2014. DOI: 10.1117/1.JRS.8.081599.

[171] A. Peinado, A. Lizana, J. Vidal, C. Iemmi, and J. Campos, "Optimization and performance criteria of a stokes polarimeter based on two variable retarders," *Opt. Express*, vol. 18, pp. 9815–9830, 2010.

[172] C. Rodríguez, E. Garcia-Caurel, T. Garnatje, M. Serra i Ribas, J. Luque, J. Campos, and A. Lizana, "Polarimetric observables for the enhanced visualization of plant diseases," *Scientific Reports*, vol. 12, p. 14 743, 2022. DOI: https://doi.org/10.1038/s41598-022-19088-6.

[173] F. H. Mustafa and M. S. Jaafar, "Comparison of wavelength-dependent penetration depths of lasers in different types of skin in photodynamic therapy," *Indian J. Phys.*, vol. 87, no. 3, 2013. DOI: 10.1007/s12648-012-0213-0.

[174] C. Rodríguez, A. Van Eeckhout, L. Ferrer, E. Garcia-Caurel, E. González-Arnay, J. Campos, and A. Lizana, "Polarimetric data-based model for tissue recognition," *Biomed. Opt. Express*, vol. 12, no. 8, pp. 4852–4872, Jul. 15, 2021. DOI: 10.1364/BOE.426387.

[175] C. Rodríguez, I. Estévez, A. Lizana, and J. Campos, "Optimizing the classification of biological tissues using machine learning models based on polarized data," *J. Biophotonics*, vol. e202200308, 2022. DOI: 10.1002/jbio.202200308.

[176] D. L. Butler, E. S. Grood, F. R. Noyes, and R. F. Zernicke, "Biomechanics of ligaments and tendons," *Exerc Sport Sci Rev*, vol. 6, pp. 125–181, 1978.

[177] R. James, G. Kesturu, G. Balian, and A. B. Chhabra, "Tendon: Biology, biomechanics, repair, growth factors, and evolving treatment options," *The Journal of Hand Surgery*, vol. 33, pp. 102–112, 2008.

[178] B. Charvet, F. Ruggiero, and D. Le Guellec, "The development of the myotendinous junction. a review.," *Musc. Liga. Tend. J.*, vol. 2, p. 53, 2012.

[179] D. J. Hulmes, "Building collagen molecules, fibrils and suprafibrillar structures," *Journal of Structural Biology*, vol. 137, pp. 2–10, 2002.

[180] S. Schiaffino and C. Reggiani, "Fiber types in mammalian skeletal muscles," *Physiological Reviews*, vol. 91, pp. 1447–1531, 2011.

[181] H. Asahara, M. Inui, and M. K. Lotz, "Tendons and ligaments: Connecting developmental biology to musculoskeletal disease pathogenesis," *Journal of Bone and Mineral Research*, vol. 32, pp. 1773–1782, 2017.

[182] C. Rodríguez, A. Van Eeckhout, E. Garcia-Caurel, A. Lizana, and J. Campos, "Automatic pseudo-coloring approaches to improve visual perception and contrast in polarimetric images of biological tissues," *Scientific Reports*, vol. 12, p. 18 479, 2022. DOI: 10.1038/s41598-022-23330-6.

[183] A. Van Eeckhout, A. Lizana, E. Garcia-Caurel, J. J. Gil, A. Sansa, C. Rodríguez, I. Estévez, E. González-Arnay, J. C. Escalera, I. Moreno, and J. Campos, "Polarimetric imaging of biological tissues based on the indices of polarimetric purity," *J. Biophotonics*, vol. 11, no. 4, 2017. DOI: 10.1002/jbio.201700189.

[184] A. V. Eeckhout, A. Lizana, E. Garcia-Caurel, J. J. Gil, A. Sansa, C. Rodríguez, I. Estévez, E. González, J. C. Escalera, I. Moreno, and J. Campos, "Indices of polarimetric purity for biological tissues inspection," in *Imaging, Manipulation, and Analysis of Biomolecules, Cells, and Tissues XVI*, vol. 10497, 2018, p. 104971V. DOI: 10.1117/12.2290283.

[185] R. McGill, J. W. Tuckey, and W. Larsen, "Variations of boxplots," *Am. Stat.*, vol. 32, no. 1, pp. 12–16, 1978.

[186] A. V. Eeckhout *et al.*, "Polarimetric imaging microscopy for advanced inspection of vegetal tissues," *Sci. Rep.*, vol. 11, no. 1, p. 3913, 2021.

[187] Y. L. Katchanov, Y. V. Markova, and N. A. Shmatko, "The distinction machine: Physics journals from the perspective of the kolmogorov–smirnov statistic," *Journal of Informetrics*, vol. 13, no. 4, p. 100 982, 2019. DOI: https://doi.org/10.1016/j.joi.2019.100982.

[188] M. B. Wilk and R. Gnanadesikan, "Probability plotting methods for the analysis of data," *Biometrika*, vol. 55, no. 1, pp. 1–17, 1968.

[189] S. Guo, S. Zhong, and A. Zhang, "Privacy-preserving kruskal–wallis test," *Computer Methods and Programs in Biomedicine*, vol. 112, no. 1, pp. 135–145, 2013. DOI: https://doi.org/10.1016/j.cmpb.2013.05.023.

[190] T. Fawcett, "An introduction to roc analysis," *Pattern Recognition Letters*, vol. 27, no. 8, pp. 861–874, 2006, ROC Analysis in Pattern Recognition, ISSN: 0167-8655. DOI: https://doi.org/10.1016/j.patrec.2005.10.010.

[191] Y. Yahtali, "Use of factor scores in multiple regression analysis for estimation of body weight by certain body measurements in romanov lamb," *PeerJ*, no. 7, 2019. DOI: https://doi.org/10.7717/peerj.7434.

[192] B. D. Hill, "The sequential kaiser-meyer-olkin procedure as an alternative for determining the number of factors in common-factor analysis: A monte carlo simulation," *Dissertation (Oklahoma State University)*, 2011.

[193] S. de la and F. Fernandez, "Regresión logística," *Facultad de Ciencias Económicas y Empresariales, Universidad Autónoma de Madrid*, 2011.

[194] W. J. Youden, "Index for rating diagnostic tests," *Cancer*, no. 3, 2950.

[195] M. B. Kursa and W. R. Rudnicki, "Feature selection with the boruta package," *Journal of Statistical Software*, vol. 36, no. 11, pp. 1–13, 2010. DOI: 10.18637/jss.v036.i11.

[196] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, and T.-Y. Liu, "Lightgbm: A highly efficient gradient boosting decision tree," in *NIPS*, 2017.

[197] T. Chen and C. Guestrin, "Xgboost: A scalable tree boosting system," ser. KDD '16, San Francisco, California, USA: Association for Computing Machinery, 2016, ISBN: 9781450342322. DOI: 10.1145/2939672.2939785.

[198] P. Refaeilzadeh, L. Tang, and H. Liu, "Cross-validation," in *Encyclopedia of Database Systems*, L. LIU and M. T. ÖZSU, Eds. Boston, MA: Springer US, 2009, pp. 532–538, ISBN: 978-0-387-39940-9. DOI: 10.1007/978-0-387-39940-9_565.

[199] K. M. Ting, "Confusion matrix," in *Encyclopedia of Machine Learning and Data Mining*. Boston, MA: Springer US, 2017, pp. 260–260. DOI: 10.1007/978-1-4899-7687-1_50.

PART II

PUBLICATIONS

PAPER A

Polarimetric observables for the enhanced visualization of plant diseases

Carla Rodríguez[1,*], Enrique Garcia-Caurel[2], Teresa Garnatje[3], Mireia Serra i Ribas[1], Jordi Luque[4], Juan Campos[1] and Angel Lizana[1]

[1] *Grup d'Òptica, Physics Department, Universitat Autònoma de Barcelona, Bellaterra 08193, Spain*
[2] *LPICM, CNRS, Ecole Polytechnique, Institut Polytechnique de Paris, Palaiseau 91120, France*
[3] *Botanical Institute of Barcelona (IBB, CSIC - Ajuntament de Barcelona), Barcelona, 08038, Spain*
[4] *Institute of Agrifood Research and Technology (IRTA), Cabrils, 08348, Spain*

This paper has been published in *Scientific Reports* **12**, pp. 14743, 2022

scientific reports

Check for updates

OPEN

Polarimetric observables for the enhanced visualization of plant diseases

Carla Rodríguez[1]✉, Enrique Garcia-Caurel[2], Teresa Garnatje[3], Mireia Serra i Ribas[1], Jordi Luque[4], Juan Campos[1] & Angel Lizana[1]

This paper highlights the potential of using polarimetric methods for the inspection of plant diseased tissues. We show how depolarizing observables are a suitable tool for the accurate discrimination between healthy and diseased tissues due to the pathogen infection of plant samples. The analysis is conducted on a set of different plant specimens showing various disease symptoms and infection stages. By means of a complete image Mueller polarimeter, we measure the experimental Mueller matrices of the samples, from which we calculate a set of metrics analyzing the depolarization content of the inspected leaves. From calculated metrics, we demonstrate, in a qualitative and quantitative way, how depolarizing information of vegetal tissues leads to the enhancement of image contrast between healthy and diseased tissues, as well as to the revelation of wounded regions which cannot be detected by means of regular visual inspections. Moreover, we also propose a pseudo-colored image method, based on the depolarizing metrics, capable to further enhance the visual image contrast between healthy and diseased regions in plants. The ability of proposed methods to characterize plant diseases (even at early stages of infection) may be of interest for preventing yield losses due to different plant pathogens.

Polarimetric instrumentation and methods are of interest in a wide range of applications, as for instance, in astronomy[1], atmospheric pollution studies[2,3], security and remote sensing[4], materials characterization[5], biomedicine[6], etc. In the case of applications in biophotonics, polarimetric methods have proved to be very useful tools to enhance the image contrast of some organic structures, and/or providing information of certain structures invisible by using regular (non-polarimetric) images. This situation is useful, for instance, for the early detection of some diseases, such as breast cancer[7,8], colon cancer[9], skin cancer[10], or brain cancer[11], among others.

The above-stated use of polarimetric methods for the study of animal[12] or even human[8–10] tissues is a well-established field of work[13] and nowadays continues in constant development. However, the application of polarimetric methods for the study of plant diseases is less common and, in the last decade, there is a growing interest of exploring more complex (and rich in terms of information) polarimetric solutions for applications in plant science. Historically, one of the most widely used polarization-based optical instruments for studying plant structures is the polarimetric microscope[14]. This instrument allows clear observation of vegetal cells, as centrosome-nuclear complexes in cell division[15] or cell suspension culture of plant specimens, as in *Picea glauca*[16]. Note that polarimetric features exploited by polarimetric microscopes are dichroism and birefringence, when they are present in samples. For instance, in the case of vegetal structures, dichroism measurements reveal the concentration and spatial organization of some plant organelles such as pigment–protein complexes in plant thylakoid membranes[17], chloroplasts[18], or quantasomes[19]. Moreover, the birefringence signature of some structures and macromolecules (for instance, cellulose), allows to study the cell wall composition[20], the trichomes structure[21] or stomata[22].

In addition to dichroism and birefringence, depolarization is third polarimetric channel which provide valuable information, although it has been underused in plant characterization. Depolarization is a statistical concept originated by incoherent (temporal or spatial) addition of different light polarizations at the level of the detector, and it can be understood as the degree of polarization disorder (randomness) introduced by a given structure to an input polarization. A suitable tool for studying depolarizing information are Mueller polarimeters[23,24], that commonly performs macroscopic analysis of samples. Recent works[4,25,26] have shown and discussed the

[1]Optics Group, Physics Department, Universitat Autònoma de Barcelona, 08193 Bellaterra, Spain. [2]LPICM, CNRS, Ecole Polytechnique, Institut Polytechnique de Paris, 91120 Palaiseau, France. [3]Botanical Institute of Barcelona (IBB, CSIC-Ajuntament de Barcelona), 08038 Barcelona, Spain. [4]Institute of Agrifood Research and Technology (IRTA), 08348 Cabrils, Spain. ✉email: carla.rodriguez@uab.cat

advantages of the depolarization-related observables to enhance image contrast between different plant structures of interest, as midrib, secondary veins, stomata and raphides. This is because their constituent units (cellulose, pectin, water content, among others) present different polarimetric features (retardance or dichroism) and/or different spatial organization, these situations inducing depolarization at macroscopic scale observation. Importantly, depolarization methods have proved to be a very interesting tool for the inspection of plant tissues, independently or in combination to other optical instrumentation, as spectroscopic instrumentation, phase or fluorescent microscopy, among others[25,27].

A popular figure or metric to study the depolarizing response of plant structures has been the so-called degree of polarization (DoP)[26,28], which measures the global change in the polarization degree of a radiation beam after interacting with a given material media, in our case, a plant tissue. This metric has been used, for instance, to estimate the chlorophyll content[29], the plant stress[30] and for vegetation classification purposes[31]. The description of the depolarization provided by the DoP, can be refined with the use of the so-called depolarization metric sets as it has been discussed in recent works[25,26]. A depolarization space is a combination of at least three depolarization-related metrics which are derived from the Mueller matrix, to enhance the image contrast in plant structures[4,25,26,32]. This is the case of the so-called Indices of Polarimetric Purity (IPPs)[33,34], that are connected with the type of polarimetric randomness that a system induces to incident light (they give information of the depolarization anisotropy of the system, i.e., dependence of the input polarization with the resultant depolarization response).

When it comes to plant pathogens, virial infections in plant specimens can induce, among others, a decrease in photosynthesis through decreases in chlorophyll efficiency and the disruption of cellular processes. In turn, the damaged cellular metabolism induces the development of abnormal substances which are injurious to the functions of the plant. Furthermore, plant diseases caused by fungi infections are generally described by the pathogen consuming cells or secreting toxins. These may lead to plant tissues containing a mix of both necrotic (dead) and healthy tissues intermingled with the fungal mycelium and may induce to microscopic structural changes[35]. Those biological modifications in plants caused by virus and fungi may be the origin of depolarization differences between healthy and pathological tissues in plants, but more study must be developed in this research line to connect microscopic biological changes with macroscopic depolarization measures. In the present paper we discuss for the first time, the use of polarimetry and in particular the use of depolarization-related metrics for the visualization and characterization of plant diseases. In particular we show how depolarizing metrics are suitable for the discrimination between healthy tissues and different type and stages of plant infections. The depolarization set of metrics chosen for the present study is the IPPs set because, as stated before, it provides three suitable metrics for plant visualization[25,26]. In addition to IPPs, we include in the study, for completeness, a second set of depolarizing observables: the so-called Components of Purity (CPs)[36]. The selection of these two polarimetric spaces is not arbitrary. Whereas the IPPs describe the capability of depolarizing samples to introduce polarimetric randomness to incident light, the CPs provide information of the polarimetric characteristics in samples inducing depolarization (diattenuation, polarizance and retardance). Importantly, these two spaces result in complementary analytical tools, and their combined use completely describes the polarimetric behavior of depolarizing samples[37]. We further demonstrate the suitability, from a quantitative and qualitative point of view, of such depolarizing spaces for plant diseases visualization enhancement. The analysis is conducted on a set of different botanical specimens showing different disease symptoms and injury stages due to the particular infection of different pathogen agents. Moreover, the potential of the method is strengthened by applying a pseudo-colored approach that helps to magnify the visual contrast between healthy and diseased tissues, or between different stages of the disease. We believe that the use of polarimetric methods, which are non-invasive and non-contact, for the early detection of plant diseases are of interest because they may contribute to prevent large product (and economical) losses in crops[38,39].

Results

The results presented in this work are related to the study of two leaf specimens suffering different infections: (1) a leaf from a specimen of *Medicago sativa* (alfalfa), which was found infected with *alfalfa mosaic virus* (AMV); and (2) a leaf from a specimen of *Olea europaea* (olive), infected with the fungus *Venturia oleaginea* (olive leaf spot). We have bounded the study in these two representative samples, for the impact and utility of these specimens in humans, but the main conclusions of this works can be extrapolated to other specimens and infectious agents. In particular, we tested the methods in a set of 73 leaves corresponding to 18 plant diseased species (complete list presented in Supplementary Table S1) and the obtained results agree. In the particular case of study, the *M. sativa* and the *O. europaea*, both leaves showed the characteristic lesions of their respective diseases. *Medicago sativa* showed chlorotic areas surrounding the affected vascular structure (see Fig. 1a and Supplementary Fig. S1a). In turn, *O. europaea* showed alternating necrotic and chlorotic ring-like lesions surrounding a chlorotic spot (see Fig. 2a and Supplementary Fig. S2a). A description of the plants used on the present study is included in the "Methods" section.

The potential of polarimetric observables for the characterization of the above-mentioned diseases was studied. The polarimetric and standard intensity images of the *M. sativa* and *O. europaea* leaves were taken by means of a complete image Mueller polarimeter working at three different illumination wavelengths (625 nm, 530 nm and 470 nm) in scattering and transmission configurations. In the context of Mueller matrix-derived images, we refer to the standard intensity (i.e., the M_{00} coefficient) as the non-polarized intensity image due to the fact that it may be interpreted as the image that would have been taken if the sample was illuminated with natural (or unpolarized) light[36]. A representative example of standard intensity images of samples can be seen in Figs. 1b and 2b. In addition, the detailed description of the polarimeter used in this study can be found in "Methods" section. The results presented here correspond to the underside part of the chlorotic *M. sativa* and *O. europaea*

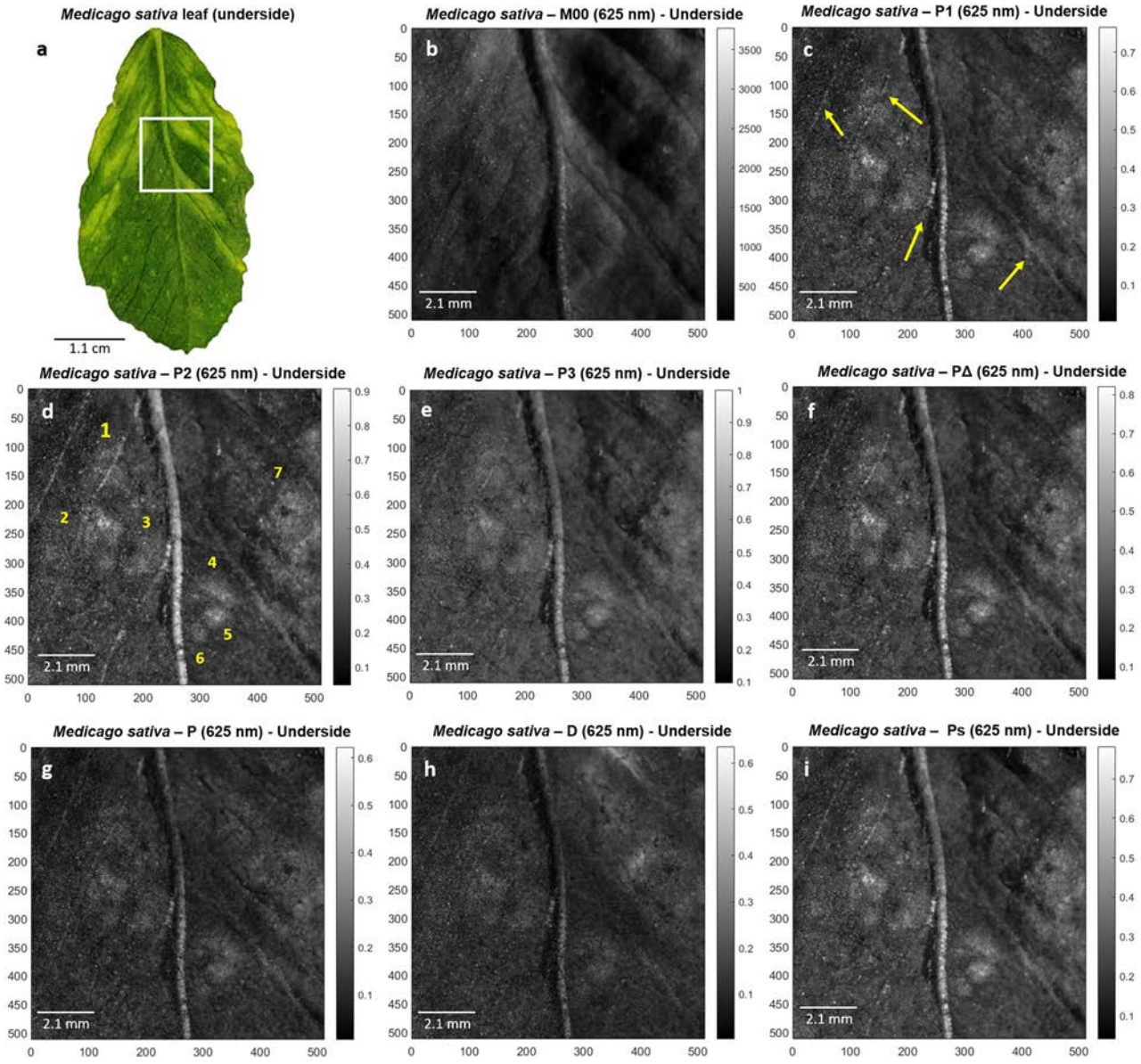


Figure 1. Polarimetric images of *Medicago sativa* leaf used in this study. (**a**) Picture of the underside part of the *Medicago sativa* leaf. White square denotes for selected region of interest (ROI) analyzed in remaining images, (**b**) regular intensity image (M_{00}) of the *M. sativa* underside ROI and its corresponding polarimetric observables (**c**) P_1, (**d**) P_2, (**e**) P_3, (**f**) P_Δ, (**g**) P, (**h**) D and (**i**) P_S for visual comparison. All images correspond to 625 nm illumination wavelength measurements performed at scattering set-up configuration. Yellow arrows correspond to the enhanced vascular structures within the sample, whereas numeric labels (from 1 to 7) indicate the number of chlorotic spots unraveled by means of polarimetric observables.

leaves measured in scattering and transmission configuration, respectively, at a wavelength of 625 nm because they represent the most relevant and interesting findings in terms of image enhancement using the polarimetric approach. With regards to the illumination channel selection, 625 nm wavelength light penetrates more deeply into the sample than shorter wavelengths (530 nm or 470 nm in our case). In this way, in the case of *O. europaea* leaf, light is capable to reach the opposite surface of the leaf, providing information about the diseased tissues at all depth levels. Similar behavior occurs for *M. sativa* leaf, where 625 nm illumination wavelength measurements provide a more accurate description of the lesions. Complementary measurement configurations and their respective obtained polarimetric images are presented in Supplementary Figs. S1 and S2.

For the optimal presentation and interpretation of results, we perform a qualitative analysis by comparing the acquired standard intensity images with the polarimetric observables described in "Methods" section. For consistency, the analysis is complemented with quantitative study regarding the polarimetric behavior on different structures of healthy and diseased tissues of samples. We also present a pseudo-colored approach image technique that allows a better visualization of certain healthy and diseased plant structures.

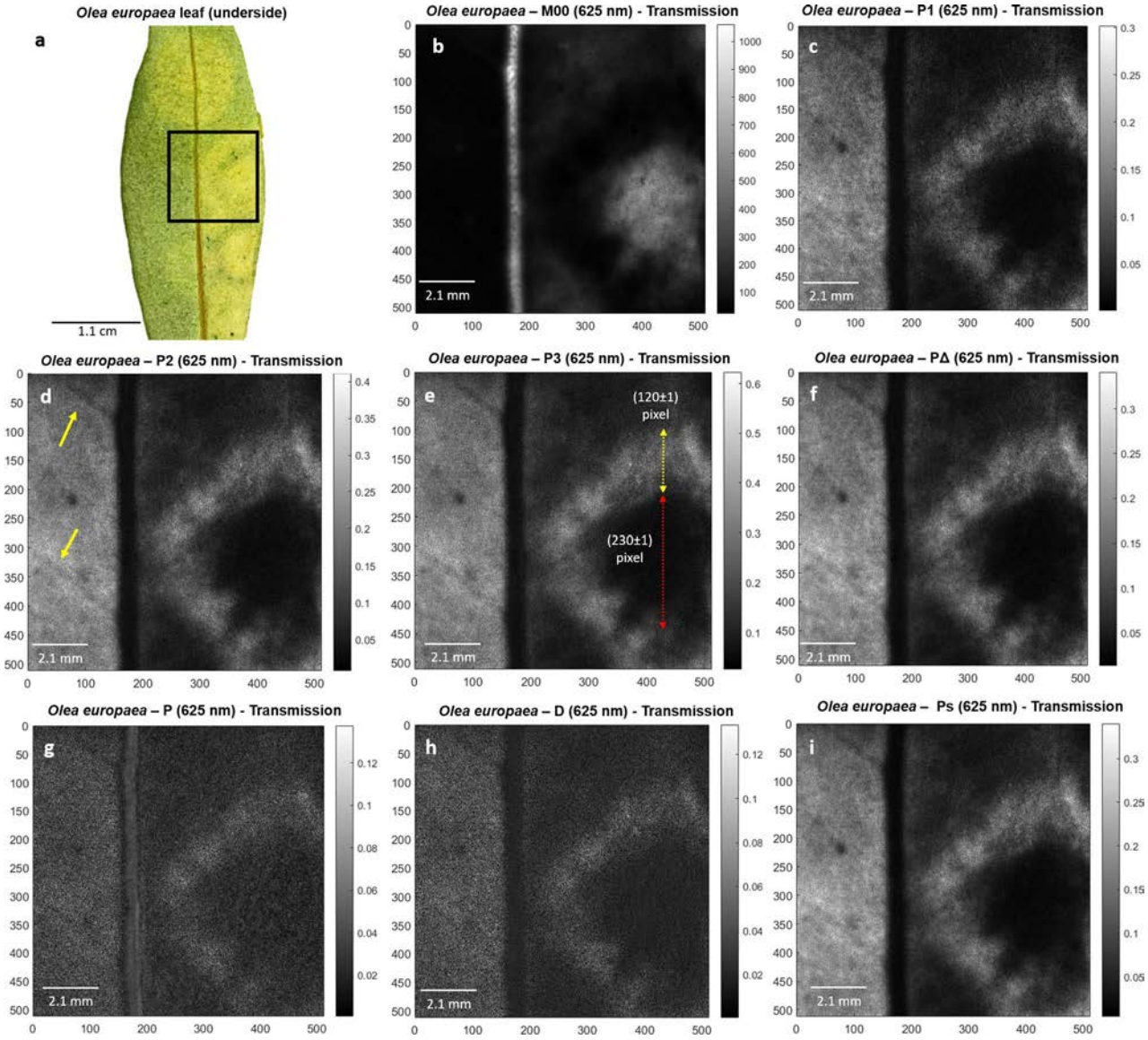


Figure 2. Polarimetric images of *Olea europaea* leaf used in this study. (**a**) Picture of the underside part of *O. europaea* leaf. Black square denotes for selected region of interest (ROI) analyzed in remaining images, (**b**) regular intensity image (M_{00}) of the *O. europaea* transmission ROI and its corresponding polarimetric observables (**c**) P_1, (**d**) P_2, (**e**) P_3, (**f**) P_Δ, (**g**) P, (**h**) D and (**i**) P_S for visual comparison. All images correspond to 625 nm illumination wavelength measurements performed at transmission set-up configuration. Red and yellow dotted lines correspond to diameter and width measurements for chlorotic spot and necrotic ring, respectively. Yellow arrows indicate the unveiled vascular structures.

Polarimetric analysis of chlorosis and necrosis. By taking advantage of the different depolarizing behavior of biological structures in samples[40], we calculated different Mueller matrix-derived polarimetric observables images and we compared them with standard intensity images, for the two specimen leaves analyzed. The selected observables for the analysis of plant samples, which are described in the "Methods" section, provide a complete description of the enpolarizing properties of samples[37]. In particular, we use the indices of polarimetric purity—P_1, P_2, P_3 -, the components of purity—P, D, P_S—and the depolarization index P_Δ, for the polarimetric sample description. The obtained results show that the depolarization metrics P_1, P_2, P_3, P_Δ and P_S clearly manifest an overall enhancement of image contrast and they help to unveil wounded zones or vascular structures that are invisible by using standard non-polarized images. In the following we present two illustrative results. The standard non-polarized transmission images for the *M. sativa* and *O. europaea* leaves are shown in Figs. 1b and 2b, respectively. These images were obtained from certain regions of interest (ROIs) in the leaves, indicated with white and black squares in Figs. 1a and 2a, respectively. For comparison purposes, we provide the polarimetric images corresponding to the analyzed polarimetric observables, identifying the ones that provide the largest image enhancement, i.e., the P_2 channel for the *M. sativa* case (Fig. 1d) and the P_3 channel for the *O. europaea* case (Fig. 2e).

Importantly, thanks to the contrast enhancement of polarimetric observables in the visualization of the *O. europaea* image, we can clearly distinguish the necrotic ring and chlorotic spot edges visible in different polarimetric channels (Fig. 2c–i), the necrotic ring being invisible by using non-polarized light intensity images (Fig. 2b). A thorough discussion of the visualization of disease symptoms by using polarimetric means as well as the current sample chlorotic spot and necrotic ring characteristics (diameter and width, red and yellow dotted line in Fig. 2e, respectively) is provided in the "Discussion" section.

As a complement to the above-presented images, we quantify the potential of depolarizing metrics to characterize plant pathologies. To do so, we consider the values taken along different cross-sections in the images of the observables shown in Figs. 1 and 2. The particular cross-sections analyzed, highlighted in yellow lines, together with the corresponding values are shown in Fig. 3. Figure 3a and b present a comparison between the values of the classical unpolarized light intensity transmission (M_{00}) metric with the depolarization metrics P_2. The values of the cross-section displayed correspond to the transition between a healthy and a chlorotic zone in the *M. sativa* specimen leaf. The difference between the chlorotic and the heathy areas is hardly visible using the classical M_{00} observable, while it becomes quite apparent in the P_2 metric image. Healthy areas show P_2 values around 0.20 ± 0.01 while chlorotic areas show characteristic values around 0.50 ± 0.01. Analogously, Fig. 3c and d show a comparison between the values of the M_{00} metric and the P_3 metric for the *O. europaea* case. The selected cross-section displays the transition between healthy, necrotic, and chlorotic regions. As in the previous example, the classical M_{00} observable (Fig. 3c) is not sensitive to all the features in the plant, for instance it does not delimitate the necrotic ring, while the P_3 observable (Fig. 3d) clearly discriminates between healthy, necrotic and chlorotic. For instance, the value of P_3 for healthy, necrotic and chlorotic areas is around 0.20 ± 0.01, 0.45 ± 0.01 and 0.10 ± 0.01 respectively. Such selectivity results in a clear enhancement of the contrast of P_3 images when compared to M_{00} images. In the example shown in Fig. 3d the necrotic area is clearly visible and appears as a "ring" area delimited between the chlorotic spot and the healthy tissue. The chlorotic area appears as dark spot, whereas the healthy area appears in grey. In contrast, M_{00} observable shows the chlorotic area as a bright spot with values close to 0.50 ± 0.01, the necrotic ring in dark, typical values of 0.05 ± 0.01, and the healthy area in dark grey with typical values of 0.90 ± 0.01. The difference between healthy and necrotic areas is less net in a M_{00} image than in a P_3 one.

We now further analyze the potential of those observables to discriminate between different typologies (healthy or diseased tissues) of plant structures. To do so, we represent the measured data on different polarimetric spaces[41], this leading to a very intuitive visualization of data, and also providing quantitative information of the structures (or tissue types) that may be present in the images of the probed samples. Moreover, the data in each ROI is collected and grouped in what we call a data-cloud which can be used for ulterior statistical data treatment or for graphical representation. For instance, data from a homogeneous region should show close values with little variance, in contrast to data from a heterogeneous region, which should show different values with extended deviation from a mean value. When represented in a graph, data from a homogeneous region may group within a well-defined region, whereas data from an heterogenous region may tend to spread over an irregular volume. Therefore, a graphical representation of data-clouds is a good tool to see at a glimpse the variability and the average values of such data. When a given area in a sample is represented according to various observables, such for instance M_{00}, IPPs, or CPs, then, data clouds corresponding each to a given observable, can be defined. Figure 4 presents the data clouds from selected ROIs (marked in Fig. 4a and d as colored rectangles) of healthy (green), chlorotic (yellow) and necrotic (dark blue) tissue regions of *M. sativa* (first row in Fig. 4) and *O. europaea* (second row in Fig. 4) leaves. The size of the corresponding regions of interest were selected depending on the tissue availability of sample to perform a homogeneous selection of healthy and diseased tissue. In particular, healthy and chlorotic tissue regions on *M. sativa* correspond to a total of 2800 pixels (40×70 and 70×40 pixels, respectively.) In counterpart, healthy, necrotic and chlorotic ROIs for *O. europaea* are of 90×40 (3600 pixels), 30×110 (3300 pixels) and 70×40 (2800 pixels), respectively. The data clouds are represented in the IPPs space for the *M. sativa* (Fig. 4b) and *O. europaea* (Fig. 4e), as well in the CPs space (Fig. 4c,f, respectively). Whereas the *M. sativa* presents two differentiated structures (healthy and chlorotic), the *O. europaea* presents three differentiated structures (healthy, necrotic and chlorotic). We see as in both plant specimens, and for the two studied polarimetric spaces, the different leaf structures (healthy, necrotic and chlorotic) are clearly differentiated and located in different spatial positions without data mixing (the different type of data -different colored dots- practically do not overlap in any region of the spaces). This trend was also observed when selecting other affected regions and the selected examples in Fig. 4 are representative cases that illustrate the discriminatory potential of polarimetric observables for symptom detection and description.

Pseudo-colored approach. In this section, we want to go one step further in plant pathology imaging based on polarimetric observables by using the discussed polarimetric spaces into a pseudo-coloration image method[42–44]. The main idea consists of building a "polarimetric triplet" by selecting, between all the used polarimetric observables, the three of them leading to larger image contrast between healthy and diseased tissues within the inspected sample. Afterwards, each one of the selected polarimetric images of the triplet is associated with a primary color (red, green and blue, respectively), and they are properly combined to build pseudo-colored images providing visual contrast between plant structures of interest. Detailed description of the pseudo-colored approach proposed in this study can be found in Supplementary Sect. 3.3.

To achieve an optimal use of the pseudo-color approach, the three observables used to represent the RGB base must be the ones that better discriminate between healthy and wounded areas. The discrimination ability of a given set of observables can be quantified by measuring the differences of these observables when applied to image healthy and wounded areas. To estimate these differences, we performed a Boxplot analysis[45,46] for the regions of interest (ROIs) shown in Fig. 4. In Fig. 5 we show the split Boxplot analysis of the Indices of

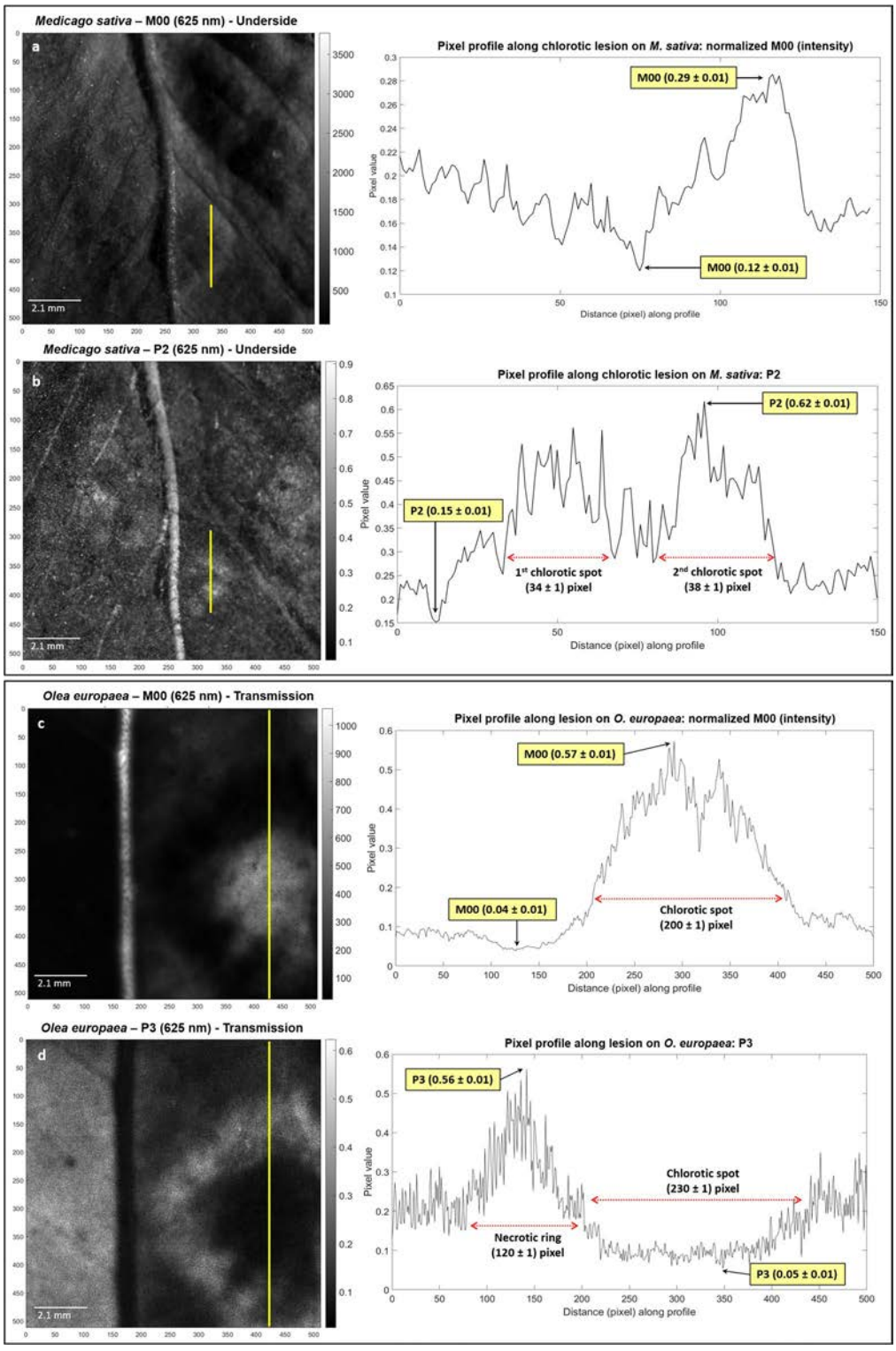


Figure 3. Pixel profile comparison for *Medicago sativa* and *Olea europaea*. (**a**) Intensity image at 625 nm of the underside part of *M. sativa* sample and its corresponding pixel profile, (**b**) polarimetric purity index P_2 and its corresponding pixel profile, (**c**) intensity image at 625 nm of the underside part of *O. europaea* sample and the corresponding healthy-necrotic-chlorotic transition pixel profile, and (**d**) polarimetric purity index P_3 and its corresponding pixel profile. The vertical yellow lines on polarimetric images indicate the plotted pixel profile segments. The squared numeric labels for M_{00}, P_2 and P_3 indicate their respective highest and lowest pixel values within the inspected pixel regions. Red-dotted horizontal lines on plots indicate the diameter of the two chlorotic spots of *M. sativa* and width measurements for chlorotic spot and necrotic ring of *O. europaea*.

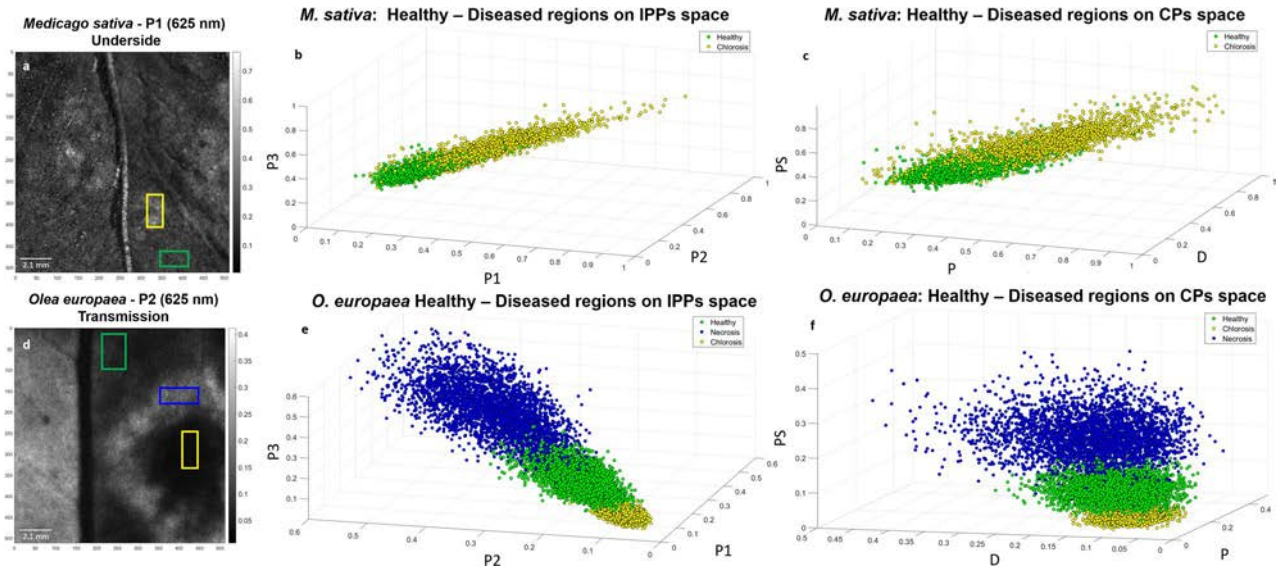


Figure 4. Scatter data plots of healthy and diseased tissue regions of *Medicago sativa* and *Olea europaea*. (**a**) Visual indicative for healthy (green) and chlorotic (yellow) tissue selected region of interest (ROI) for *M. sativa*, (**b**) corresponding IPPs space (P_1, P_2, P_3) for healthy and chlorotic data clouds representation, (**c**) components of purity (P, D, P_S) space for healthy and chlorotic data clouds representation, (**d**) visual indicative for selected healthy (green), chlorotic (yellow) and necrotic (dark blue) tissue ROIs for *O. europaea*, (**e**) corresponding IPPs space (P_1, P_2, P_3) for healthy, chlorotic and necrotic data clouds representation and (**f**) components of purity (P, D, P_S) space for healthy, chlorotic and necrotic data clouds representation.

Polarimetric Purity (P_1, P_2, P_3) and the Components of Purity (P, D, P_S) for healthy, chlorotic and necrotic tissue regions of *M. sativa* (Fig. 5a,b) and *O. europaea* (Fig. 5c,d).

We can see in Fig. 5 that some observables clearly separate the studied necrotic, chlorotic and healthy tissues. To help readers to evaluate such differences, the median values are highlighted with dotted red lines in the case of largest differences. In the case of the *M. sativa,* which presents chlorotic and healthy tissues, they are clearly separated by all the IPPs (P_1, P_2, P_3) metrics (Fig. 5a) as well as by the sphericity degree observable, P_S (Fig. 5b). In the case of the *O. europaea*, which presents chlorotic, necrotic and healthy tissues, we see how the same observables (IPPs and sphericity degree P_S) are those leading to the larger distances between different tissues (Fig. 5c,d). Under these results, the IPPs as well as the sphericity degree, P_S, are considered to be good candidates to implement the pseudo-colored images. Moreover, Boxplot unveils the amount of outlier values (mild and extreme, illustrated as small circles and stars in Fig. 5, respectively) of each data distribution so that the homogeneity of the selected tissue region can be quantified. The current distributions of outliers range from 0.44 to 3.67%. Note that the low percentage of data outliers ensures the homogeneous selection of the tissue conditions (healthy, necrotic or chlorotic). The complete description of outlier values for each analyzed tissue region can be found in Supplementary Table S2.

In addition to the above-presented Boxplot analysis, we computed, for each observable in Fig. 5, the distance between medians in yellow and green boxes (chlorotic and healthy, respectively) for the *M. sativa* case (Fig. 5a,b); and between medians in yellow and green boxes (chlorotic and healthy, respectively) as well as between medians in blue and green boxes (necrotic and healthy, respectively) for the *O. europaea* case (Fig. 5c,d). The values resulting from such comparison are shown in Table 1. The largest values in Table 1 are highlighted in bold. The median values of each polarimetric observable for selected healthy and diseased regions on both plant species are shown in Supplementary Table S3.

By considering the largest differences between means according to Table 1, we selected two triplets of polarimetric observables for the pseudo-colored images construction: one mixing observables of the IPPs and CPs spaces (P_2, P_3, P_S) and another based on the IPPs by themselves (P_1, P_2, P_3). At this stage, for each depolarizing observable selected P_i (where $i = 1, 2, 3, S$) we set different threshold values, which were derived from the Boxplot analysis (see Supplementary Table S5). These thresholds allowed us to numerically discriminate between different tissue conditions: healthy/chlorotic for *M. sativa* and healthy/chlorotic/necrotic, for *O. europaea*, and then, each separated condition is assigned to a primary color (red, green and blue, for chlorotic, healthy, and necrotic tissues, respectively). This process is illustrated by considering particular examples. In the case of the *M. sativa*, and for a particular observable selected P_i, a first binary image (black–red) is constructed according to pixel values above/below a certain threshold, this image carrying the chlorotic information of the plant. Then, a second binary image (black–green) is similarly constructed carrying the healthy information of the plant. The same procedure is conducted for the for *O. europaea,* but now, as this specimen presents an extra condition (necrosis), three binary-images are obtained: (1) for the chlorotic content (black–red), (2) for the healthy content (black–green), and (3) for necrotic content (black–blue). Finally, for each studied metric (P_1, P_2, P_3, P_S), a final pseudo-colored image is obtained by adding all the contributions as following,

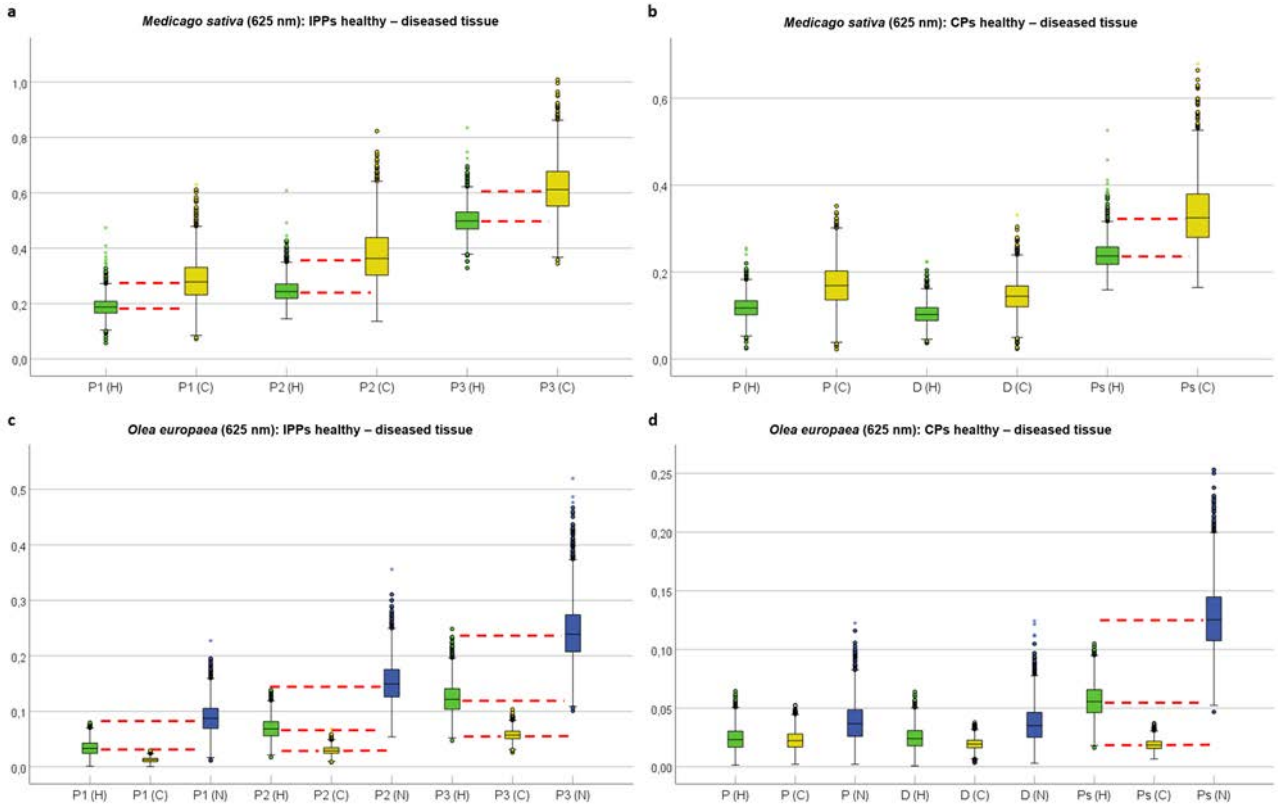


Figure 5. Boxplot charts for healthy and diseased regions for *Medicago sativa* and *Olea europaea*. (**a**) Indices of polarimetric purity (P_1, P_2, P_3) boxplot for healthy and chlorotic locations on *M. sativa*, (**b**) Components of purity (P, D, P_S) Boxplot for healthy and chlorotic locations on *M. sativa*, (**c**) Indices of polarimetric purity (P_1, P_2, P_3) Boxplot for healthy, necrotic and chlorotic locations on *O. europaea* and (**d**) Purity components (P, D, P_S) Boxplot for healthy, necrotic and chlorotic locations on *O. europaea*. The corresponding healthy, chlorotic and necrotic data distributions are labeled and colored as H (green), C (yellow) and N (dark blue), respectively. Red-dashed lines indicate the locations of median values and illustrate they do not fit within the boxes of different tissue conditions (healthy, chlorotic or necrotic), allowing discrimination. Circles and stars correspond to mild and extreme outlier values, respectively.

	P_1	P_2	P_3	P	D	P_S
Medicago sativa						
Chlorotic—Median diff	0.091 ± 0.085	**0.119 ± 0.104**	0.113 ± 0.106	0.051 ± 0.055	0.042 ± 0.043	0.088 ± 0.081
Olea europaea						
Chlorotic—Median diff	0.020 ± 0.014	0.039 ± 0.020	**0.064 ± 0.029**	0.001 ± 0.013	0.004 ± 0.011	0.037 ± 0.015
Necrotic—Median diff	0.054 ± 0.031	0.080 ± 0.041	**0.117 ± 0.059**	0.013 ± 0.019	0.011 ± 0.019	0.069 ± 0.032

Table 1. Polarimetric observables median value difference and propagated errors for healthy-diseased tissue. The corresponding three largest median difference values are highlighted in bold.

$$P_{i,approach}(x,y) = P_{i,Red-Chlorotic}(x,y) + P_{i,Green-Healthy}(x,y) + P_{i,Blue-Necrotic}(x,y) \quad (1)$$

where the sub-index i denotes for the particular depolarizing observable (i.e., $i = 1, 2, 3, S$). Note that in the case of the *M. sativa* we can consider $P_{i,\ Blue\text{-}Necrotic} = 0$ because there is not necrotic content. Importantly, each term in Eq. (1) provides a binary-colored image that has been filtered according to the threshold criteria above-explained. Finally, the content of a full polarimetric triplet, $[P_1, P_2, P_3]$ or $[P_2, P_3, P_S]$, is put together by constructing a linear combination of the pseudo-colored observables implemented according to Eq. (1), so generalized final pseudo-colored images, valid for both *M. sativa* and *O. europaea*, are obtained as:

$$\text{Pseudo\#1}(x,y) = P_{2,approach}(x,y) + P_{3,approach}(x,y) + P_{S,approach}(x,y), \quad (2)$$

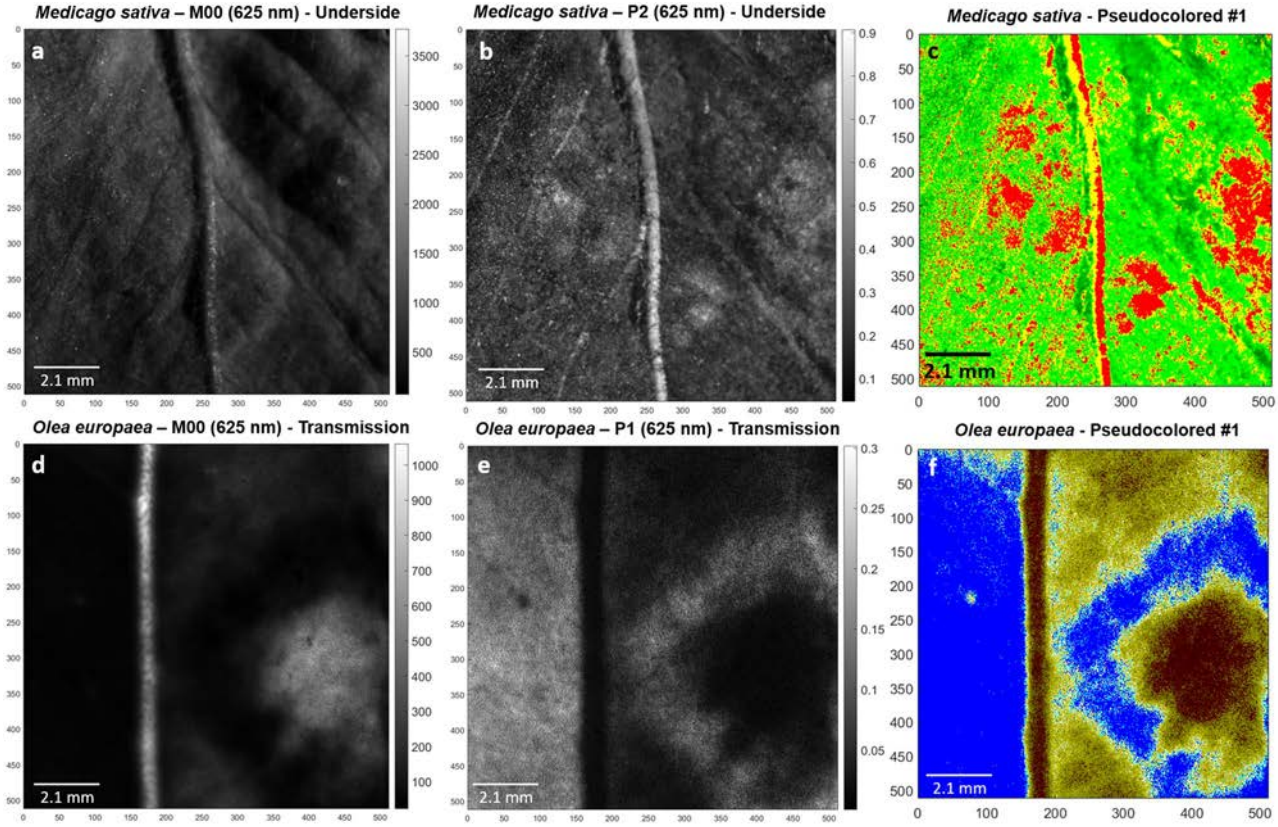


Figure 6. Visual comparison of *Medicago sativa* leaf: (**a**) 625 nm intensity image (M_{00}), (**b**) polarimetric purity index P_2, (**c**) processed image by means of #1 pseudo-coloration; Visual comparison of *Olea europaea* leaf: (**d**) 625 nm intensity image (M_{00}), (**e**) polarimetric purity index P_1, (**f**) processed image by means of #1 pseudo-coloration.

$$\text{Pseudo\#2}(x,y) = P_{1,approach}(x,y) + P_{2,approach}(x,y) + P_{3,approach}(x,y). \tag{3}$$

As shown in Eqs. (2) and (3), we built two general pseudo-colored functions (corresponding to each of the two selected observables triplets) labeled as #1 and #2. Note that both Eqs. (2) and (3) can be applied for the two plant specimens analyzed: The *M. sativa* and the *O. europaea* cases. Note as well that the corresponding weights of each term in Eqs. (2) and (3) were chosen, for simplicity, as the unit, this giving the same weight to all the observables in the triplet, but other weights could be selected to enhance plant structures visualization. In fact, the low values of the selected polarimetric observables in the case of the *O. europaea* (see Supplementary Table S3) lead to a darkened final image. In such a case, to visually improve the pseudo-coloration approach, the weights were pondered by a factor 2 so that the resulting image turned brighter. A more detailed description of the method proposed to build pseudo-colored functions is provided in Supplementary Sect. 3.3.

Figure 6 shows three images that illustrate the effect of pseudo-coloration to enhance the contrast of polarimetric images as well as the interest of this technique to discriminate between different tissues. The figure shows images of classical M_{00}, the representative polarimetric observable P_2 and the corresponding pseudo-colored image for *M. sativa*, and P_1 for *O. europaea*. In particular the first row in Fig. 6 presents the classical non-polarizer transmission (M_{00}, Fig. 6a), the polarimetric purity index P_2 (Fig. 6b), and pseudo-coloration resulting from Eq. (2), labeled as #1 (Fig. 6c), for *M. sativa* leaf sample. The second row in Fig. 6 presents the intensity (Fig. 6d), the polarimetric purity index P_1 (Fig. 6e), and pseudo-coloration resulting from Eq. (2), labeled as #1 (Fig. 6f), for *O. europaea* leaf. We found that the pseudo-colored resulting images for purity-mixed space and isolated purity space both shown very similar results. To not to present redundant information, pseudo-colored images regarding isolated Indices of Polarimetric Purity [Eq. (3)] for *M. sativa* and *O. europaea*, are presented in Supplementary Fig. S4d and h, respectively. The images resulting from polarimetric based pseudo-colored processing are excellent in terms of enhanced visual discrimination of healthy/diseased tissues. Detailed analysis of enhanced structures is provided in "Discussion" section.

Discussion

The present work highlights the suitability of using polarimetric observables, in particular, the Indices of Polarimetric Purity (IPPs) and the Components of Purity (P, D and P_S), for the inspection of plant disease symptoms. In addition, we show that the implementation of a pseudo-coloration image processing method, which is based on the above-mentioned polarimetric observables, is a useful tool to enhance the image contrast between different tissue natures (healthy, necrotic and chlorotic) of botanical samples. Although the potential of depolarizing

metrics to characterize different plant pathologies was observed in 18 different species affected by different infection agents, we focus our attention in two specific plant species because of their relevance in agricultural production: alfalfa (*M. sativa*) and olive (*O. europaea*). The complete list of the specimens used in this study is presented in Supplementary Table S1.

Normal levels of chlorophyll in plant cells are the cause of the characteristic green color of leaves and allow the natural performance of photosynthetic processes, i.e., the synthesis of organic substances like carbohydrates, from the sources of inorganic prime matter and water from soil, CO_2 from the atmosphere, and light[35]. Chlorophyll deficiency in plant foliage triggers the appearance of chlorosis: the lack of green pigments, which leads to the yellowing of tissues, thus hampering the production of nutritional substances. The chlorotic symptoms can appear due to different biotic and abiotic causes, e.g., iron-deficient or alkaline soils (high pH levels), and plant pathogens[35], such as the alfalfa pathosystem (AMV) studied here. With regards to the necrotic symptomatology caused by a pathogen infection, it is characterized by the premature death of affected plant cells and the darkening of tissues. The release of pathogen toxins into the plant cell or, additionally, the release of residual components from surrounding dead cells into the intracellular space typically induces the necrotic lesion appearance. In this study we selected the case of necrosis in leaves of *O. europaea* as an illustrating example.

All plant samples were inspected at different wavelengths (625 nm, 530 nm and 470 nm) and measuring configurations (scattering–measuring both the beam and the underside part of the samples, and transmission). However, to not to extent the content of the manuscript, we only discuss here the cases providing the most interesting results in terms of disease symptom visualization. In particular, we consider *M. sativa* leaves measurements at 625 nm in scattering configuration from the underside, and the *O. europaea* leaves measured in transmission configuration from the outside. Measurements with additional orientations of the leaves were performed and the corresponding results are summarized in the Supplementary document.

Although different optical leaf properties may play an important role in leaves spectral response[40], as for example the spectral signature of a leaf, in the specimens studied in this work we have observed that longer wavelengths penetrate more into samples than shorter wavelengths, as they are less affected by scattering processes than shorter wavelengths. This result agrees with discussion provided in Ref.[47]. In fact, longer wavelengths may carry more information about microstructures present in the bulk of the sample than shorter wavelengths, which are more sensitive to features present in surface in a bulk region near the surface of the leaves. Since chlorosis or necrosis affect the whole bulk of the leaf, a relative long wavelength such as 625 nm provides a more complete picture of the leave than a shorter wavelength and it is for this reason why we chose in the results shown in the main text of the present manuscript. A second reason for our choice, is that chlorophyll has an absorption peak close to 625 nm, and although being weak this particular absorption feature will impact the measured depolarization. In transparent scattering media, photons which contribute to increase the depolarization of light are those who followed a higher number of scattering events before reaching the detector. Those photons are the ones which also show longer paths inside the sample because of the multiple scattering events that they experienced. In absorbing media, photons which account a higher number of scattering events will be preferentially removed from the beam because their paths are longer than those of photons which account a fewer number of scattering events. Therefore, and in general, it can be said that an absorbing media is expected to show less depolarization than analogous transparent media. In the cases discussed here, because of chlorotic and necrotic tissues show reduced chlorophyll concentration, they appear more transparent than healthy ones at 625 nm: reason why depolarization measurements at this particular wavelength may be more sensitive to the presence of chlorophyll (and therefore to the presence of a wounded tissue) than measurements at wavelengths for which chlorophyll does not show any absorption.

In the case of the *M. sativa* leaf (Fig. 1), the polarimetric images; in particular, the ones corresponding to the depolarization-related P_1, P_2, P_3, P_Δ and P_S observables (Fig. 1c–f,i, respectively); show a significant enhancement of image contrast between healthy and diseased regions when compared to the image of the standard M_{00} unpolarized transmitted or scattered light intensity. For instance, up to seven chlorotic spots (labeled from 1 to 7 in Fig. 1d) as well as some vascular structures (yellow arrows in Fig. 1c), which are barely observable by standard intensity images (Fig. 1b), are well visible by using polarimetric means. The improvement is further highlighted after comparing the line cross-sections corresponding to the yellow segments in Fig. 3a (regular intensity image, *M*00 channel) and Fig. 3b (P_2 channel), which are taken within healthy and diseased regions in *M. sativa* sample (diseased spots labeled as 4 and 5 in Fig. 1d). In this sense, considering the maximum and minimum values (peak-to-valley) in the Fig. 3a and b pixel profiles (corresponding to the healthy/diseased tissues regions) we demonstrate a significant increase of the chlorotic symptom visibility. For practical reason we quantify visibility with the figure defined by the following expression: $V = [I_{\max} - I_{\min}]/[I_{\max} + I_{\min}]$, where I_{max} and I_{min} correspond to the maximum and minimum value of the pixel within the selected segment in the cross-section or the region in the image. In the case discussed here, the visibility features a value of $V = (0.41 \pm 0.05)$ for the non-polarized transmission/scattering intensity in Fig. 3a and of $V = (0.61 \pm 0.03)$ for the P_2 channel in Fig. 3b.

Moreover, thanks to the visualization and contrast improvement in polarimetric images, it is possible to measure the width of the diseased regions with a pixel precision limited by the density of pixels in the detector. After calibration we found that one pixel of CCD corresponds to approx. 21 μm in sample. Therefore, the size of the chlorotic regions shown in Fig. 1 are estimated to (34 ± 1) pixel, (730 ± 21) μm, and (38 ± 1) pixel, (820 ± 21) μm, for the first and second chlorotic spots, respectively. On the other hand, note that we can physically interpret the depolarizing behavior of healthy/diseased tissues in *M. sativa*, considering that the higher the value of IPPs within a particular tissue region, the lower depolarization induces to incident light[33]. Under this scenario, and according to images in Fig. 1, the chlorotic spots induce fewer depolarizing effects on incident light (high IPPs values) than healthy tissues (low IPPs values) as expected in connection with the previous discussion, because of the reduced absorption of light in the sample due to low levels of chlorophyll.

In the case of the *O. europaea* sample (Fig. 2), it presents a necrotic ring (surrounding a chlorotic spot) hardly visible in non-polarized transmission/scattering images (Fig. 2b). In contrast, polarimetric observables show an increased contrast compared to that of non-polarized transmission/scattering, which provide a more accurate visualization and spatial delimitation of the lesions in the leaf. Moreover, as in the case of *M. sativa* previously discussed, the use of polarimetric observables unveils the presence of vascular structures (see Fig. 2c–i) non-visible in non-polarized transmission/scattering images. Thanks to this visual improvement, the width of the necrotic ring and the diameter of the chlorotic spot are clearly delimited in polarimetric images (yellow and red dotted lines in Fig. 2e, respectively): (120 ± 1) pixel, (2.580 ± 0.021) mm, for the necrotic ring width and (230 ± 1) pixel, (4.940 ± 0.021) mm for the chlorotic spot diameter. The unveiled vascular structures are pointed out with yellow arrows in Fig. 2d. As in the previous case, this visual discrimination between the diseased and healthy regions within the sample can be quantitatively measured analyzing the pixel profile corresponding to the cross-section represented by yellow segments in Fig. 3c (non-polarized transmission/scattering, M00, image) and Fig. 3d (observable P_3). As in the previous case, the cross-sections are selected crossing healthy and wounded areas of the tissue. The chlorotic spot appears as a broad bell-like feature in the non-polarized transmission/scattering cross-section (Fig. 3c) hiding the presence of the necrotic area. In contrast, both the chlorotic spot and the necrotic ring can be distinguished in the cross-section corresponding to the P_3 observable (indicated with a red dashed segment in Fig. 3d). To measure the contrast improvement in polarimetric images compared to non-polarized transmission/scattering ones, we take for instance the visibility value for the necrotic ring which is estimated to $V = (0.84 \pm 0.03)$ and $V = (0.00 \pm 0.03)$ respectively.

Furthermore, the necrotic ring registers the highest IPPs values (P_3 from 0.15 to 0.56) followed, in descending order, by the healthy leaf lamina (P_3 values ranging from 0.13 to 0.31) and the chlorotic spot (e.g. P_3 from 0.05 to 0.18). This behavior is physically translated as the necrotic ring inducing fewer depolarizing effects on incident light than healthy regions whose response is, in turn, even less depolarizing than the chlorotic spot. Moreover, note that the left side (leaf lamina) of the *O. europaea* sample presents similar values as the necrotic ring for some polarimetric channels (see Fig. 2), although such region does not present necrotic-like tissue condition. This result, which can lead to errors in the physical interpretation of the studied structure, is originated by the curvature of the leaf at that left-region, which leads to out of focus measurements. Under this scenario, the collected polarimetric information of a plant region is affected by out of focus polarimetric information corresponding to other plant regions, all this information being mixed with the actual polarimetric information of the studied structure (the left-side of the leave, in this case). This situation highlights the importance of analyzing well-focused images of plant structures, for the visual inspection of diseased plant samples through polarimetric methods. If due to the non-planar surface of the studied sample a proper image focusing of plant structures of interest cannot be simultaneously ensured by a single image shot, a scanning-based imaging approach is recommended. In this work, we focus on the structures at the right-side of the image, as all structures of interest (healthy and wounded tissues) are in a non-curved and well-focused region of the leaf. To better illustrate the interest of the use of polarimetric images to characterize disease symptoms in vegetal tissues, we chose to collect the pixels from the original image and to represent them as a whole in a data cloud figure. In a data cloud figure, pixels corresponding to different regions should group in separate clouds. Non-overlapping clouds indicate that the related regions are well discriminated. On the contrary, either fully or partially overlapping clouds indicate a poor discrimination of nominally different zones. For the present study, we used three-dimensional data clouds with a selection of variables corresponding to the so-called space of Indices of Polarimetric Purity, (P_1, P_2 and P_3), or the space of the Components of Purity, (P, D and P_S) (see Fig. 4). In the case of the *M. sativa* leaf, the data clouds, corresponding to healthy and chlorotic regions (green and yellow squares in Fig. 4a, respectively), are well-discriminated as these two tissue conditions are clearly spatially separated (i.e., practically no data mixing between tissues with different health condition is produced) when represented in either the: IPPs or the CPs space (Fig. 4b,c, respectively). According to the previous, healthy tissue (green squares) tends to group close to the point (0,0,0) which corresponds to higher depolarization, while chlorotic regions tend to group to areas related with less depolarization. Importantly, a stronger depolarization response occurs when the leaves contain non-organized spatial structures (i.e., neither not aligned nor homogeneously distributed) or an important number of microstructures which efficiently scatter light. In this context, either the biological or the structural changes caused by the chlorotic symptoms of infected regions are translated into a less depolarizing effect on incident light when compared with healthy tissues, thus increasing the sensitivity of depolarizing channels to chlorosis detection. In the case of the *O. europaea* leaf, both depolarizing spaces clearly discriminate between tissue conditions: see how data clouds corresponding to healthy, chlorotic, and necrotic tissue regions in *O. europaea* (green, yellow and blue squares, respectively, in Fig. 4d) are clearly spatially separated in both IPPs and CPs spaces (Fig. 4e,f, respectively). However, in this case, the chlorotic data (yellow squares) are quite close to the point (0,0,0), therefore indicating that, in this case, chlorotic tissue is more depolarizing than healthy tissue (green squares).

The differential depolarizing behavior among the two studied plant species might be related to specific biological characteristics of these species. Regarding the visible chlorotic symptoms that showed a differential polarimetric response between the species, several hypotheses could be considered: first, leaves of *M. sativa* and *O. europaea* correspond to non-sclerophyllous and sclerophyllous species, respectively, thus suggesting that different leaf tissue structure could be involved in this differential depolarizing response. Secondly, the chlorophyll-a and chlorophyll-b leaf content is different for each plant species[48–50] so that the chlorotic symptoms development could be depending on the infected specimen chlorophyllic profile. Additionally, the type of pathogen which caused the chlorotic symptoms on both inspected leaves may also play a role in the differential depolarizing behavior: *M. sativa* was infected with the alfalfa mosaic virus (AMV) and *O. europaea* was infected with the fungus *V. oleaginea*. As far as the infection strategy of viruses is different from fungi, the chlorotic symptoms may manifest in different ways and lead to different physical transformations of plant tissues (thickness, stress

resistance, turgor, structure, organization, etc.) of the unitary scatters (cells), therefore resulting in different depolarizing behavior. Moreover, Lanza et al.[51] described the morphological changes induced by the infection of *V. oleaginea* on *O. europaea*, consisting of a progressive loss and degradation of plastids and chloroplasts in the palisade parenchyma cells. This degradation process leads to a release of cytoplasmic contents of palisade cells at advanced infection stages, which may affect the cuticle by reducing the resistance to water loss and causing stress on leaf tissues. Interestingly, the chlorotic ring on *O. europaea* might be also caused by fungal phytotoxins, as it was previously reported[52]. On the basis of the above comments, it is not surprising that a same symptom, i.e. chlorosis, is a consequence of different biological processes in different plants specimens, and therefore, may correspond to different polarimetric responses. In terms of the enhancement of the image contrast of disease symptoms, we have showed that differential information of polarimetric channels can well discriminate between different infection status (i.e., chlorosis, necrosis and healthy tissue). This has been demonstrated for *M. sativa* and *O. europaea*, and additionally confirmed in other plant specimens (see the complete list of analyzed specimens in Supplementary Table S1).

The pseudo-coloration is the second approach that we chose to better use the contrast enhancement provided by polarimetric images for a visual discrimination of features present in complex scattering media. As previously described, we performed a pseudo-coloration image processing based on triplets of polarimetric observables information codified in three color (R, G, B) channels. At this point we would like to emphasize that we use for the first time the pseudo-colored approach to the analysis of plant disease symptoms. Compared with previous references, the pseudo-coloration approach was improved by conducting two main modifications: (1) the depolarizing observables were not restricted to the IPPs space, but extended to an optimized selection of polarimetric observables within the IPPs and CPs spaces (2) an image filtering process, based on data obtained from a Boxplot analysis (see Fig. 5), was included in the process to largely discriminate between different tissue conditions (healthy/diseased).

In this section we bring the discussion to the pseudo-colored images resulting from the triplet P_2, P_3 and P_S, as it is the most suitable one to construct the pseudo-colored functions for *M. sativa* and *O. europaea* samples (see Table 1). The pseudo-colored images obtained based on the IPPs lead to similar results, therefore, we invite the interested readers to consult results for IPPs in Supplementary Fig. S4. The final pseudo-colored images for *M. sativa* and *O. europaea*, presented in Fig. 6c and f, respectively, demonstrate a visual enhancement of disease symptoms: the extreme different coloration of the chlorotic lesions on *M. sativa* with respect to the healthy tissue of the leaf lamina (red and green regions on Fig. 6c) leads to a more accurate location of the diseased area. Similar behavior occurs for lesions on *O. europaea* leaf, where performed pseudo-colorations lead to a better delimitation of the different transitions from chlorotic spot to necrotic ring and healthy tissue of leaf lamina (Fig. 6f). Importantly, we remark here the fact that pseudo-colored images enhance the contrast between different tissue conditions, even more than the performance of isolated imaging of polarimetric observables for *M. sativa* and *O. europaea* samples (Fig. 6b,e, respectively). This behavior highlights the suitability of using this image-processing method for biological samples analysis. Particularly, the inspection and estimation of direct lesions, and the characterization and early detection of infection processes on plant tissues.

Methods

Sample description. The plant samples used in this work were a leaf of *Medicago sativa* specimen infected with *alfalfa mosaic virus* (AMV, which causes wilting or white flecks to necrotic wounds and chlorotic mosaics on leaves) and a leaf of *Olea europaea* specimen infected with *Venturia oleaginea* (causal agent of the olive leaf spot). This worldwide spread disease of olive may cause severe tree defoliation and a delay in fruit ripening, among other symptoms, thus leading to relevant yield losses.

Native from warmer-temperate climates of south-central Asia, *M. sativa* belongs to the Fabaceae family (legumes), it is cultivated worldwide for livestock feeding purposes. Despite the toxicity of unsprouted alfalfa, it is also suitable for human consumption in sprout stage or dehydrated. Regarding *O. europaea*, this species belongs to the Oleaceae family. Although the native species were found in eastern land regions around the Mediterranean Sea, his production is not limited to Mediterranean countries: *O. europaea* is cultivated in several countries such as South Africa, New Zealand, North and South America, and Australia. In addition to olive oil production and fruit consumption (olives), *O. europaea* trees are also grown for fine wood manufacturing.

T. Garnatje and J. Luque undertook the formal identification of the plant material used in this study. An herbarium voucher of both *M. sativa* and *O. europaea* are deposited in the Herbarium of the Botanical Institute of Barcelona (BC-983007 and BC-983006, respectively). All methods were performed in accordance with relevant guidelines and regulation.

Mueller–Stokes formalism. Different mathematical approaches can be used to describe the polarimetric properties of material media[36]. Among those, in this work we use the Mueller–Stokes (M–S) formalism, because it is especially suited to deal with partially polarized or unpolarized light, as well as with depolarizing samples[53,54]. In this approach, polarization of light is described by means of Stokes vectors[36] which are composed by four real coefficients., S_0, S_1, S_2 and S_3, whose physical interpretation is directly related to the irradiance (total intensity of the beam, S_0) and the amount of light which is linearly polarized in vertical and horizontal (S_1), 45° and 135° (S_2) directions, and right or left-handed circularly polarized (S_3). This Stokes formalism leads to a tridimensional representation of light states of polarization in the so-called Poincaré sphere[36].

In the M–S formalism, polarimetric samples are described by 4×4 real coefficient matrices, the so-called Mueller matrices $\mathbf{M}$, which can be understood as the polarimetric transfer functions of polarimetric systems. In addition, polarization of light exiting from a sample (represented by the $\boldsymbol{S}_{out}$ Stokes vector) is linearly related with the input polarization (represented by the $\boldsymbol{S}_{in}$ vector) through the Mueller matrix describing a sample. The Mueller

matrix of a sample can be experimentally measured by using polarimeters[23,24], and it encodes rich polarimetric information: dichroism (diattenuation and polarizance), retardance and depolarization. In the following, the polarimetric principle of a Mueller matrix experimental determination is briefly reviewed, but it is thoroughly described in Ref.[54]. By means of a complete image Mueller polarimeter, the sample is illuminated by a set of n different controlled light beam polarization states. Accordingly, the polarization of detected (imaged) light emerging from the sample is analyzed. To fully determine the experimental Mueller matrix, at least four input (and analyzed) states are required, which corresponds to 16 radiometric measurements. In the current work, a total of 36 measurements are taken: 6 input states of polarization (generators) and the corresponding 6 analyzers proposed in[55], which are used for minimization of noise in radiometric measurements. The mathematical relationship between the set of incident states and the detected ones is given by the sample 4×4 Mueller matrix, $\mathbf{M}_S$, in the following way:

$$\mathbf{I} = \mathbf{S}_{PSA}\mathbf{M}_S\mathbf{S}_{PSG}, \tag{4}$$

where the detected radiometric measurements are given by the $n\times n$ matrix $\boldsymbol{I}$, $\boldsymbol{S}_{PSG}$ is the $4\times n$ matrix of input set of polarized light beams where the n columns represent the different Stokes vectors used to illuminate the sample, and the $n\times4$ matrix $\boldsymbol{S}_{PSA}$ which corresponds to the transposed input Stokes vectors. In this way, by computing the pseudoinverse of $\boldsymbol{S}_{PSG}$ and $\boldsymbol{S}_{PSA}$, matrices in Eq. (4), the corresponding Mueller matrix $\mathbf{M}_S$ of the sample can be directly deduced.

Polarimetric observables. The complete set of polarimetric properties of the sample can be derived from the experimental Mueller matrix[36,53,54]. While some metrics (as dichroism) can be directly deduced from the Mueller matrix, other information as retardance and depolarization content are entangled in such a way that a straightforward interpretation is not possible and advanced algebraic methods of analysis, known as matrix decomposition methods are needed[56–60]. In this way, Mueller matrix $\mathbf{M}$ can be written as

$$\mathbf{M} = m_{00}\begin{bmatrix} 1 & \mathbf{D}^T \\ \mathbf{P} & m \end{bmatrix}, \tag{5}$$

where m_{00} entails the non-polarized transmission/scattering of the sample, $\boldsymbol{D}$ and $\boldsymbol{P}$ are 3-dimensional vectors encoding diattenuation and polarizance, respectively, and 3×3 submatrix m entangles the retardance and depolarization in a complex way. Whereas $\boldsymbol{D}$ describes the dependency of intensity from emergent light from sample with the input state of polarization, $\boldsymbol{P}$ is related to the capability of said sample to polarize light when illuminated with an unpolarized state[36]. Regarding the m-submatrix, a commonly used formulism to gather entangled polarimetric properties is the so-called Lu-Chipman decomposition[55], which describes any Mueller matrix $\mathbf{M}$ as the product of three 4×4 pure Mueller matrices (pure depolarizers, retarders and diattenuators) that synthesize well-defined polarimetric observables leading to an easier physical interpretation of the medium.

Regarding the depolarization behavior of media, a general quantitative indicator of the overall depolarizing power of the sample, the depolarization index P_Δ, is commonly used[61,62]. Despite of the fact that P_Δ is suitable to represent homogeneous depolarization, it does not provide enough information regarding the situations where depolarization actually depends on the state of polarization of the illuminating beam. In this way, it is worth defining the covariance matrix $\mathbf{H}$ (associated with Mueller matrix, $\mathbf{M}$)[61]:

$$\mathrm{H(M)} = \frac{1}{4}\sum_{i,j=0}^{3} m_{ij}(\sigma_i \otimes \sigma_j), \tag{6}$$

where m_{ij} represent the Mueller matrix coefficients, σ are the Pauli matrices and $\otimes$ the Kronecker product. Since Mueller matrices are not Hermitian and thus we cannot ensure they are diagonalizable, we define the so-called indices of polarimetric purity (IPP)[33] as a set of three real magnitudes, P_1, P_2 and P_3, directly derived from the covariance matrix $\mathbf{H}$ eigenvalues:

$$P_1 \equiv \frac{\lambda_0-\lambda_1}{Tr\mathbf{H}}, \quad P_2 \equiv \frac{\lambda_0+\lambda_1-2\lambda_2}{Tr\mathbf{H}}, \quad P_3 \equiv \frac{\lambda_0+\lambda_1+\lambda_2-3\lambda_3}{Tr\mathbf{H}}, \qquad 0 \le P_i \le 1 \quad (i=1,2,3). \tag{7}$$

where λ-eigenvalues are taken in decreasing order as $\lambda_0 \ge \lambda_1 \ge \lambda_2 \ge \lambda_3$ and IPP values are restricted to $0 \le P_1 \le P_2 \le P_3 \le 1$. Indices of polarimetric purity define a real tridimensional depolarization space whose interpretation, in addition to how much light is depolarized, is related with different depolarizing mechanisms in the sample. Therefore, the depolarization spaces can be potentially used to discriminate among structures which different depolarization signatures due to their properties and structure. Based on the idea of representing depolarization as the incoherent sum of four pure components[33], IPPs correspond to the statistical weights of each component: P_1 quantifies the relative portion of pure non-depolarizing component, P_2–P_1 the relative weight of a bidimensional depolarizer, P_3–P_2 the relative portion of an equiprobable mixture of three pure components (tridimensional depolarizer) and $1-P_3$ is associated with an ideal depolarizer. In consequence, different IPPs values lead to the comprehension of the inherent depolarizing mechanisms of samples[34]. However, we can define depolarization index, P_Δ, by means of IPPs but also eventually splitting depolarization information in the commonly used components of purity P, D and P_S:

$$P_\Delta = \frac{1}{\sqrt{3}}\sqrt{2P_1^2+\frac{2}{3}P_2^2+\frac{1}{3}P_3^2} = \frac{1}{3}\sqrt{D^2+P^2+3P_S^2}, \qquad 0 \le P_\Delta \le 1, \tag{8}$$

where components of purity P and D are the polarizance and diattenuation vector magnitudes, respectively. The sphericity degree, P_S, defines the contribution on depolarization which differs from dichroic origin. Therefore, depolarization index builds a common link between both purity spaces. Pure depolarizing systems are those which entail $P_\Delta = P_1 = P_2 = P_3 = 0$, meanwhile pure non-depolarizing media is characterized by $P_\Delta = P_1 = P_2 = P_3 = 1$. Recently, it has been demonstrated that the combined use of IPPs and components of purity is an ideal framework to describe depolarizing behavior of samples[37].

Complete image Mueller polarimeter. The polarimetric images shown in this work (Figs. 1, 2 and Supplementary Figs. S1 and S2) are gathered from the experimental Mueller matrices of the studied samples, which are acquired by means of a complete image Mueller polarimeter. By taking advantage of the wide spectral response of the light source, which actually covers the visible spectrum (from 400 to 700 nm approx.), we use three different illuminating wavelengths (625 nm, 530 nm and 470 nm) for the consequently inspection of the sample at different depths[47]. The polarimeter consists of two independent optical systems based on Parallel Aligned Liquid Crystals (PA-LC) retarders, mounted into two compact mobile arms respectively. The Polarization State Generator (PSG) optical design leads to generate any fully polarized state. It is composed by a linear polarizer oriented at 0° with respect to the laboratory vertical, followed by two PA-LC at 45° and 0°. Equivalently, the Polarization State Analyzer (PSA) is capable to detect any state of polarization. The PSA uses 6 optimized polarization analysis states[55]. In our instrument, both the PSG and the PSA consist of a linear polarizer followed by two Parallel-Aligned Liquid Crystal cells externally controlled by sending different voltages. The combination of both PSG and PSA, are used to record 36 intensity images which are used to deduce the Mueller matrix[55]. Regarding internal optical set-up, PSA has the same optical elements as PSG but arranged in reverse order. For the acquisition of sample intensity, a CCD camera is placed on the PSA system.

To perform the Mueller matrix measurements of biological samples, two main optical configurations are used. By tilting by 34° the PSG with respect to the horizontal laboratory reference and maintaining the PSA at 0° to avoid the ballistic reflection, we perform what we call scattering measurements. Complementary, by placing both PSG and PSA at 90° we perform transmission measurements. In both configurations, we selected from the whole sample, a region of interest (ROI) of 512×512 pixels, which corresponds to an area of 1.1×1.1 cm^2. The detailed information about optical components and the visualization of measurement configurations (Figs. S5 and S6) is shown in Supplementary document.

Data availability

The datasets generated during and/or analysed during the current study are not publicly available due to the conduction of different research studies but are available from the corresponding author on reasonable request.

Received: 14 February 2022; Accepted: 24 August 2022
Published online: 30 August 2022

References

1. Hough, J. New opportunities for astronomical polarimetry. *J. Quant. Spec. Radiat. Transf.* **106**(1–3), 122–132. https://doi.org/10.1016/j.jqsrt.2006.12.012 (2007).
2. Kong, Z. *et al.* A polarization-sensitive imaging lidar for atmospheric remote sensing. *J. Quant. Radiat. Transf.* **271**, 107747. https://doi.org/10.1016/j.jqsrt.2021.107747 (2021).
3. Deuzé, J. L. *et al.* Remote sensing of aerosols over land surfaces from POLDER-ADEOS-1 polarized measurements. *J. Geophys. Res.* **106**(D5), 4913–4926. https://doi.org/10.1029/2000JD900364 (2021).
4. Patty, C. H. L. *et al.* Imaging linear and circular polarization features in leaves with complete Mueller matrix polarimetry. *Biochim. Biophys. Acta Gen. Subj.* **1862**(6), 1350–1363. https://doi.org/10.1016/j.bbagen.2018.03.005 (2018).
5. Hall, S. A., Hoyle, M. A., Post, J. S. & Hore, D. K. Combined Stokes vector and Mueller matrix polarimetry for materials characterization. *Anal. Chem.* **85**(15), 7613–7619. https://doi.org/10.1021/ac401864g (2013).
6. Tuchin, V. V. *Tissue Optics: Light Scattering Methods and Instruments for Medical Diagnosis* (SPIE Press, 2007).
7. Yaroslavsky, A. N., Feng, X., Muzikansky, A. & Hamblin, M. R. Fluorescence polarization of methylene blue as a quantitative marker of breast cancer at the cellular level. *Sci. Rep.* **9**(1), 940. https://doi.org/10.1038/s41598-018-38265-0 (2019).
8. Patel, A., Khan, A., Quinlan, R. & Yaroslavsky, A. N. Polarization-sensitive multimodal imaging for detecting breast cancer. *Cancer Res.* **74**(17), 4685–4693. https://doi.org/10.1158/0008-5472.CAN-13-2411 (2014).
9. Ivanov, D. *et al.* Colon cancer detection by using Poincaré sphere and 2D polarimetric mapping of ex vivo colon samples. *J. Biophotonics* https://doi.org/10.1002/jbio.202000082 (2020).
10. Le, D. L., Nguyen, D. T., Le, T. H., Phan, Q. H. & Pham, T. T. H. Characterization of healthy and cancerous human skin tissue utilizing Stokes–Mueller polarimetry technique. *Opt. Commun.* **480**, 126460. https://doi.org/10.1016/j.optcom.2020.126460 (2021).
11. Schucht, P. *et al.* Visualization of white matter fiber tracts of brain tissue sections with wide-field imaging Mueller polarimetry. *IEEE Trans. Med. Imaging* **39**(12), 4376–4382. https://doi.org/10.1109/TMI.2020.3018439 (2020).
12. Rodríguez, C. *et al.* Polarimetric data-based model for tissue recognition. *Biomed. Opt. Express* **12**, 4852–4872. https://doi.org/10.1364/BOE.426387 (2021).
13. Ahmad, I., Khaliq, A., Iqbal, M. & Khan, S. Mueller matrix polarimetry for characterization of skin tissue samples: A review. *Photodiagn. Photodyn. Ther.* **30**, 101708. https://doi.org/10.1016/j.pdpdt.2020.101708 (2020).
14. Nguyen, T. H., Kandel, M. E., Rubessa, M., Wheeler, M. B. & Popescu, G. Gradient light interference microscopy for 3D imaging of unlabeled specimens. *Nat. Commun.* **8**, 210. https://doi.org/10.1038/s41467-017-00190-7 (2017).
15. Hyman, A. A. Centrosome movement in the early divisions of *Caenorhabditis elegans*: A cortical site determining centrosome position. *J. Cell Biol.* **9**(3), 1185–1193. https://doi.org/10.1083/jcb.109.3.1185 (1989).
16. Hakman, I. & Fowke, L. C. An embryogenic cell suspension culture of *Picea glauca* (White spruce). *Plant Cell Rep.* **6**(1), 20–22. https://doi.org/10.1007/BF00269730 (1987).
17. Tóth, T. N. *et al.* Fingerprinting the macro-organisation of pigment–protein complexes in plant thylakoid membranes in vivo by circular-dichroism spectroscopy. *Biochim. Biophys. Acta* **1857**(9), 1479–1489. https://doi.org/10.1016/j.bbabio.2016.04.287 (2016).
18. Zeiger, E. & Hepler, P. Production of guard cell protoplasts from onion and tobacco. *Plant Physiol.* **58**(4), 492–498. https://doi.org/10.1104/pp.58.4.492 (1976).

19. Sauer, K. & Calvin, M. Molecular orientation in quantasomes. *J. Mol. Biol.* **4**, 451–466. https://doi.org/10.1016/s0022-2836(62)80102-8 (1962).
20. Shtein, I. S. *et al.* Stomatal cell wall composition: Distinctive structural patterns associated with different phylogenetic groups. *Ann. Bot.* **119**(6), 1021–1033. https://doi.org/10.1093/aob/mcw275 (2017).
21. Bischoff, V. *et al.* Trichome birefringence and its homolog AT5G01360 encode plant-specific DUF231 proteins required for cellulose biosynthesis in arabidopsis. *Plant Phys.* **153**(2), 590–602. https://doi.org/10.1104/pp.110.153320 (2010).
22. Gao, C. J., Xia, X. J., Shi, K., Zhou, Y. H. & Yu, J. Q. Response of stomata to global climate changes and the underlying regulation mechanism of stress responses. *Plant Physiol. J.* **48**(1), 19–28 (2012).
23. Tyo, J. S. Review of passive imaging polarimetry for remote sensing applications. *App. Opt.* **45**(22), 5453–5469. https://doi.org/10.1364/AO.45.005453 (2006).
24. Spandana, K. U. *et al.* Polarization-resolved Stokes–Mueller imaging: A review of technology and applications. *Lasers Med. Sci.* https://doi.org/10.1007/s10103-019-02752-1 (2019).
25. Van Eeckhout, A. *et al.* Polarimetric imaging microscopy for advanced inspection of vegetal tissues. *Sci. Rep.* **11**(1), 3913. https://doi.org/10.1038/s41598-021-83421-8 (2021).
26. Van Eeckhout, A. *et al.* Depolarizing metrics for plant samples imaging. *PLoS ONE* https://doi.org/10.1371/journal.pone.0213909 (2019).
27. Arai, K. & Nishmura, Y. Polarization model for discrimination of broad and needle shaped leaves and estimation of LAI using polarization measurements. *Adv. Sp. Res.* **44**(4), 510–516. https://doi.org/10.1016/j.asr.2009.04.026 (2009).
28. Vanderbilt, V. C., Grant, L. & Daughtry, C. S. T. Polarization of light scattered by vegetation. *Proc. IEEE* **73**(6), 1012–1024. https://doi.org/10.1109/PROC.1985.13232 (1985).
29. Yao, C., Lu, S. & Sun, Z. Estimation of leaf chlorophyll content with polarization measurements: Degree of linear polarization. *Radiat. Transf.* **242**, 106787. https://doi.org/10.1016/j.jqsrt.2019.106787 (2020).
30. Kharuk, V. I. & Yegorov, V. V. Polarimetric indication of plant stress. *Remote Sens. Environ.* **33**(1), 35–40. https://doi.org/10.1016/0034-4257(90)90053-O (1990).
31. Curran, P. J. Polarized visible light as an aid to vegetation classification. *Remote Sens. Environ.* **12**(6), 491–499. https://doi.org/10.1016/0034-4257(82)90023-2 (1982).
32. Savenkov, S. N. *et al.* Measurement and interpretation of Mueller matrices of barley leaves. *Quantum Electron.* **50**, 55. https://doi.org/10.1070/QEL17178 (2020).
33. San José, I. & Gil, J. J. Invariant indices of polarimetric purity: Generalized indices of purity for NXN covariance matrices. *Opt. Commun.* **284**(1), 38–47. https://doi.org/10.1016/j.optcom.2010.08077 (2011).
34. Van Eeckhout, A. *et al.* Synthesis and characterization of depolarizing samples based on the indices of polarimetric purity. *Opt. Lett.* **42**(20), 4155–4158. https://doi.org/10.1364/OL.42.004155 (2017).
35. Agrios, G. N. *Plant Pathology* 5th edn. (Elsevier Academic Press, 2005).
36. Gil, J. J. & Ossikovski, R. *Polarized Light and the Mueller Matrix Approach* (CRC Press, 2016).
37. Van Eeckhout, A. *et al.* Unravelling physical information of depolarizers. *Opt. Express* **29**(23), 38811–38823. https://doi.org/10.1364/OE.438673 (2021).
38. Savary, S., Ficke, A., Aubertot, J. N. & Hollier, C. Crop losses due to diseases and their implications for global food production losses and food security. *Food Secur.* **4**, 519–537. https://doi.org/10.1007/s12571-012-0200-5 (2012).
39. Oerke, E. C. Crop losses to pests. *J. Agric. Sci.* **144**(1), 31–43. https://doi.org/10.1017/S0021859605005708 (2006).
40. Jacquemoud, S. & Ustin, S. *Leaf Optical Properties* (Cambridge University Press, 2019).
41. Van Eeckhout, A. *et al.* Depolarization metric spaces for biological tissues classification. *J. Biophotonics* https://doi.org/10.1002/jbio.202000083 (2020).
42. Rodríguez, C. *et al.* Indices of polarimetric purity: application in biological tissues. I*Proceedings of the SPIE*, 116460P. https://doi.org/10.1117/12.2578254 (2021).
43. Van Eeckhout, A. *et al.* Indices of polarimetric purity for biological tissues inspection. In *Proceedings of the SPIE*, 104971V. https://doi.org/10.1117/12.2290283 (2018).
44. Van Eeckhout, A. *et al.* Polarimetric imaging of biological tissues based on the indices of polarimetric purity. *J. Biophotonics* **11**, e201700189. https://doi.org/10.1002/jbio.201700189 (2017).
45. Tukey, J. W. *Exploratory Data Analysis* (Addison-Wesley, 1977).
46. McGill, R., Tuckey, J. W. & Larsen, W. A. Variations of box plots. *Am. Stat.* **32**(1), 12–16. https://doi.org/10.2307/2683468 (1978).
47. Mustafa, F. H. & Jaafar, M. S. Comparison of wavelength-dependent penetration depths of lasers in different types of skin in photodynamic therapy. *Indian J. Phys.* **87**(3), 203–209. https://doi.org/10.1007/s12648-012-0213-0 (2013).
48. Rekik, I. *et al.* Effects of untreated and treated wastewater at the morphological, physiological and biochemical levels on seed germination and development of sorghum (*Sorghum bicolor* (L.) Moench), alfalfa (*Medicago sativa* L.) and fescue (*Festuca arundinacea* Schreb). *J. Hazard Mater.* **326**, 165–176. https://doi.org/10.1016/j.jhazmat.2016.12.033 (2017).
49. Bahloul, N., Kechaou, N. & Mihoubi, N. B. Comparative investigation of minerals, chlorophylls contents, fatty acid composition and thermal profiles of olive leaves (*Olea europeae* L.) as by-product. *Grasas Aceites* https://doi.org/10.3989/gya.0102141 (2014).
50. Kakade, M. D. & Borse, R. D. Effect of gamma radiation on chlorophyll content of lucerne (*Medicago sativa* L.). *Online Int. Interdiscip. Res. J.* **9**, 2249–9598 (2019).
51. Lanza, B., Ragnelli, A. M., Priore, M. & Aimola, P. Morphological and histochemical investigation of the response of *Olea europaea* leaves to fungal attack by *Spilocaea oleaginea*. *Plant Path.* **66**(8), 1239–1247. https://doi.org/10.1111/ppa.12671 (2017).
52. Yahiaoui, R., Guechi, A., Lukasova, E. & Girre, L. Mutagenic and membranal effect of a phytotoxic molecule isolated from olive leaves parasitized by the fungus *Cycloconium oleaginum* Cast. *Mycopathologia* **126**, 121–129. https://doi.org/10.1007/BF01146204 (1994).
53. Goldstein, D. *Polarized Light* 2nd edn. (Marcel Dekker, 2003).
54. Chipman, R. A. *Polarimetry: Handbook of Optics* 2nd edn. (McGraw Hill, 1995).
55. Peinado, A., Lizana, A., Vidal, J., Iemmi, C. & Campos, J. Optimization and performance criteria of a Stokes polarimeter based on two variable retarders. *Opt. Express* **18**, 9815–9830. https://doi.org/10.1364/OE.18.009815 (2010).
56. Lu, S. Y. & Chipman, R. A. Interpretation of Mueller matrices based on polar decomposition. *J. Opt. Soc. Am. A* **13**(5), 1106–1113. https://doi.org/10.1364/JOSAA.13001106 (1996).
57. Ossikovski, R., De Martino, A. & Guyot, S. Forward and reverse product decompositions of depolarizing Mueller matrices. *Opt. Lett.* **32**(6), 689–691. https://doi.org/10.1364/OL.32.000689 (2007).
58. Gil, J. J. Transmittance constraints in serial decompositions of depolarizing Mueller matrices: The arrow form of a Mueller matrix. *J. Opt. Soc. Am. A* **30**(4), 701–707. https://doi.org/10.1364/JOSAA.30.000701 (2013).
59. Ossikovski, R. Analysis of depolarizing Mueller matrices through a symmetric decomposition. *J. Opt. Soc. Am. A* **26**(5), 1109–1118. https://doi.org/10.1364/JOSAA.26.001109 (2009).
60. Ossikovski, R. Canonical forms of depolarizing Mueller matrices. *J. Opt. Soc. Am. A* **27**(1), 123–130. https://doi.org/10.1364/JOSAA.27.000123 (2010).
61. Gil, J. J. Polarimetric characterization of light and media. *Eur. Phys. J. Appl. Phys.* **40**, 1–47. https://doi.org/10.1051/epjap:2007153 (2007).

62. Gil, J. J. & Bernabéu, E. Polarization and depolarization indices of an optical system. *Opt. Acta* **33**(2), 185–189. https://doi.org/10.1080/713821924 (1986).

Acknowledgements
We acknowledge the financial support of Spanish MINECO (RTI2018-097107-B-C31, and Fondos FEDER); Catalan Government (2017-SGR-001500). J. L. was supported by the CERCA Program, Generalitat de Catalunya.

Author contributions
C.R., T.G., J.L., E.G.C, A.L. and J.C. were involved in the study design. T.G. and J.L. collected samples. C.R. and M.S.R. conducted measurements. J.C. and C.R. developed software for data analysis. C.R., M.S.R., E.G.C., J.C. and A.L. analyzed data. C.R., E.G.C., and A.L. drafted the manuscript. A.L. and J.C. supervised the project. All authors reviewed the manuscript.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-022-19088-6.

Correspondence and requests for materials should be addressed to C.R.

Reprints and permissions information is available at www.nature.com/reprints.

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

Polarimetric observables for the enhanced visualization of plant diseases: supplementary document

Carla Rodríguez[1,*], Enrique Garcia-Caurel[2], Teresa Garnatje[3], Mireia Serra i Ribas[1], Jordi Luque[4], Juan Campos[1] and Angel Lizana[1]

[1]Optics group, Physics Department, Universitat Autònoma de Barcelona, Bellaterra, 08193, Spain
[2]LPICM, CNRS, Ecole Polytechnique, Institut Polytechnique de Paris, Palaiseau, 91120, France
[3]Botanical Institute of Barcelona (IBB, CSIC-Ajuntament de Barcelona), Barcelona, 08038, Spain
[4]Institute of Agrifood Research and Technology (IRTA), Cabrils, 08348, Spain
[*]carla.rodriguez@uab.cat

1. Plant material

In this work we have considered a set of 18 plant specimens and 73 inspected leaves in total. The plant samples were collected at different disease stages. To not to extent the content of the main manuscript, we limited the main discussion to the results related to two sample cases: a leaf of *M. sativa* specimen, showing chlorotic symptomatology due to the infection of alfalfa mosaic virus (AMV) and a leaf of *O. europaea* specimen, showing necrotic ring-like wounds and chlorotic halos due to the infection of *Venturia oleaginea*. In the following Table S1, we present the complete list of the 18 collected plant specimens, as well as their causal agent and short symptom description.

Table S1. Complete name list of the 18 collected specimens and their causal agent and symptom appearance. Different symptoms, which were not identified as chlorosis or necrosis were labeled as "other" discolorations observed on leaves. The selected specimens for this work are highlighted in gray.

Plant species	Pathogen – Organism group	Leaves inspected	Symptoms
Arauija sericifera	Undetermined organism	2	Chlorosis
Arbutus unedo	*Ruptoseptoria unedonis* – Fungi	1	Necrosis
Celtis australis	Undetermined virus	1	Other
	Undetermined phytoplasma	1	Chlorosis
Ficus carica	Undetermined organism	1	Other
	Fig Mosaic Virus – Virus	1	Chlorosis
Hedera helix	*Xanthomonas* – Bacteria	1	Other
	Mycosphaerella hederae-helicis – Fungi	1	Necrosis
	Undetermined organism	1	Chlorosis
Malva sylvestris	Malva Mosaic Virus – Virus	13	Chlorosis
	Undetermined bacterium	2	Necrosis
	Puccinia malvacearum – Fungi	20	Other
Medicago sativa	Alfalfa Mosaic Virus – Virus	1	Chlorosis
Morus sp.	Mulberry Mosaic Virus – Virus	1	Chlorosis
Olea europaea	*Venturia oleaginea* – Fungi	3	Necrosis
Parietaria officinalis	*Ramularia parietariae* – Fungi	1	Necrosis
Plantago coronopus	*Golovinomyces sordidus* – Fungi, powdery mildew	1	Other
Platanus × hybrida	*Erysiphe platani*– Fungi, powdery mildew	1	Other
	Corythucha ciliata – Insect	1	Chlorosis
Prunus dulcis	*Transzchelia discolor* – Fungi, rust	2	Other
	Panonychus ulmi – Arthropoda, mite	1	Chlorosis

	Polystigma amygdalinum – Fungi	1	Other
	Wilsonomyces carpophilus – Fungi	1	Necrosis
Quercus pubescens	Undetermined leaf miner – Arthropoda, insect	2	Other
	Erysiphe alphitoides – Fungi, powdery mildew	1	Other
Rubus idaeus	Undetermined organism	1	Necrosis
	Phragmidium violaceum – Fungi, rust	2	Other
Rumex pulcher	*Ramularia* sp. – Fungi	2	Other
Viburnum tinus	*Septoria* sp. – Fungi	1	Necrosis
	Undetermined fungus	1	Other
Vitis spp.	*Daktulosphaira vitifoliae* – Arthropoda, insect	1	Other
	Undetermined fungus	1	Other
	Uncinula necator – Fungi, powdery mildew	1	Chlorosis
	Guignardia bidlewii – Fungi	1	Necrosis

2. Qualitative analysis: polarimetric observables

In the following we present, as a complement to the material presented in the main manuscript, the non-polarized transmission / scattering images and the Mueller matrix-derived polarimetric observables (P_1, P_2, P_3, P_Δ, P, D and P_S) [1] for the leaf specimens of *Medicago sativa* and *Olea europaea* used in this study, both measured at 625 nm illumination wavelength.

Figure S1 shows a picture of the *M. sativa* specimen (Fig. S1a), the non-polarized transmission / scattering image (Fig. S1b) and the obtained polarimetric images of the Indices of polarimetric purity (P_1, P_2 and P_3), the depolarization index P_Δ, and the components of purity (P, D and P_S) (Figs. S1c to S1i, respectively), by means of the transmission-configuration measurement of the leaf. The depolarization inspection for the *M. sativa* sample clearly demonstrates an overall enhancement of image contrast and the revelation of the different spatial locations of the chlorotic wounds and the vascular structures (clearly seen in Fig. S1c) of the leaf, which are invisible in regular intensity images. Yellow arrows in Figs. S1d, S1e and S1f indicate the chlorotic tissue regions.

Regarding the *O. europaea* sample, Fig. S2 shows a picture of the specimen (Fig. S2a), the standard intensity image (Fig. S2b) and the obtained polarimetric images of the Indices of polarimetric purity (P_1, P_2 and P_3), the depolarization index P_Δ, and the components of purity (P, D and P_S) (Figs. S2c to S2i, respectively), by performing the scattering-configuration measurements of the beam part of the leaf. Inspected polarimetric channels demonstrate, in addition to the overall enhancement of image contrast, the accurate delimitation of the necrotic ring as well as the chlorotic spot (at the center of the lesion) and halo. Yellow and white arrows in Figs. S1d and S1e indicate the limits of the necrotic ring and the revealed diseased regions (chlorotic spot and halo) within the sample, respectively.

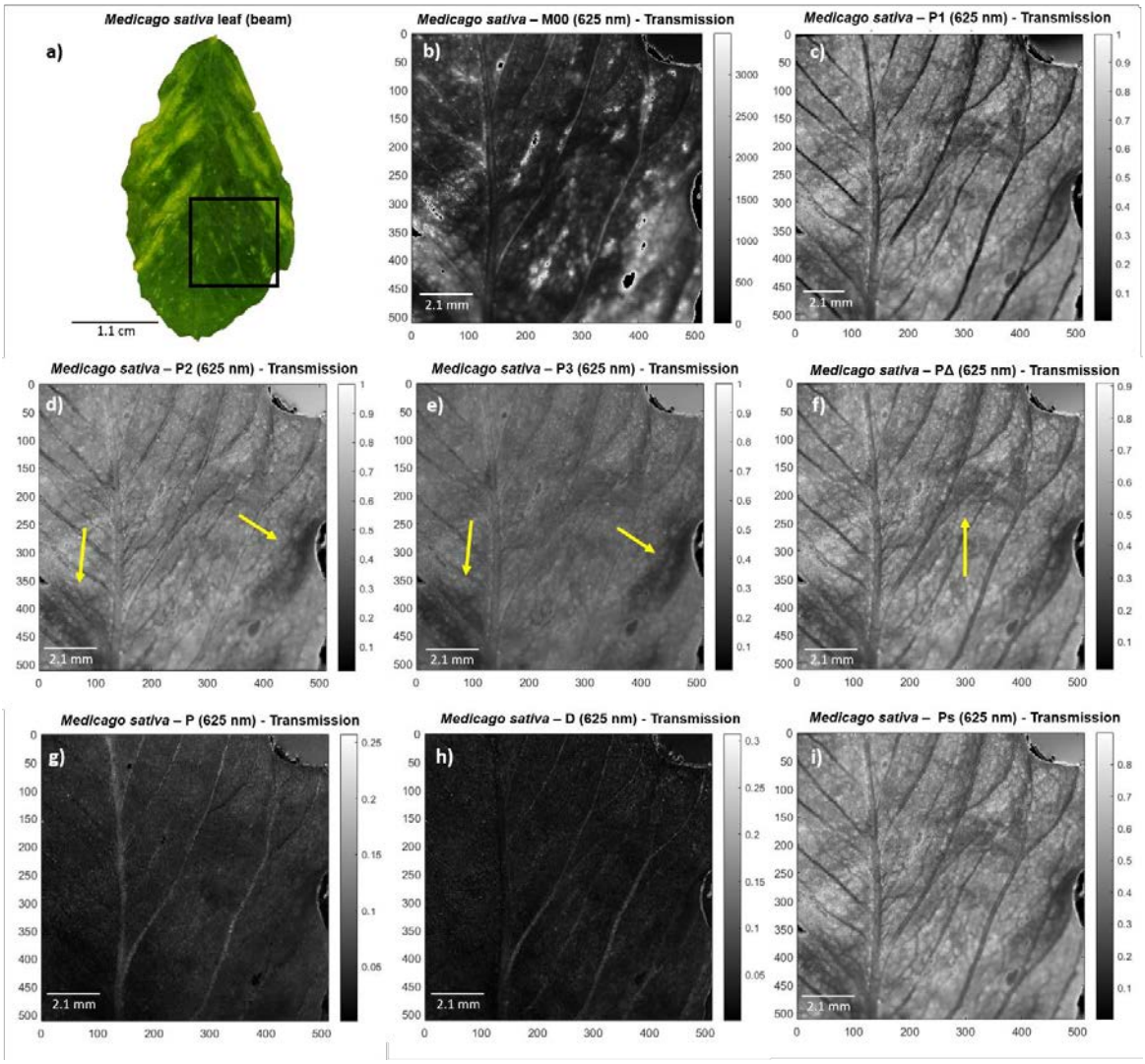


Figure S1. Images of polarimetric observables of *Medicago sativa* leaf. a) Picture of the beam part of the *M. sativa* leaf used in this study. Black square denotes for selected region of interest (ROI) analyzed in remaining images, b) regular intensity image (M_{00}) of the *M. sativa* transmission ROI and its corresponding polarimetric observables c) P_1, d) P_2, e) P_3, f) P_Δ, g) P, h) D and i) P_S for visual comparison. All images correspond to 625 nm illumination wavelength measurements performed at transmission set-up configuration. Yellow arrows indicate the locations of chlorotic tissue regions within the sample.

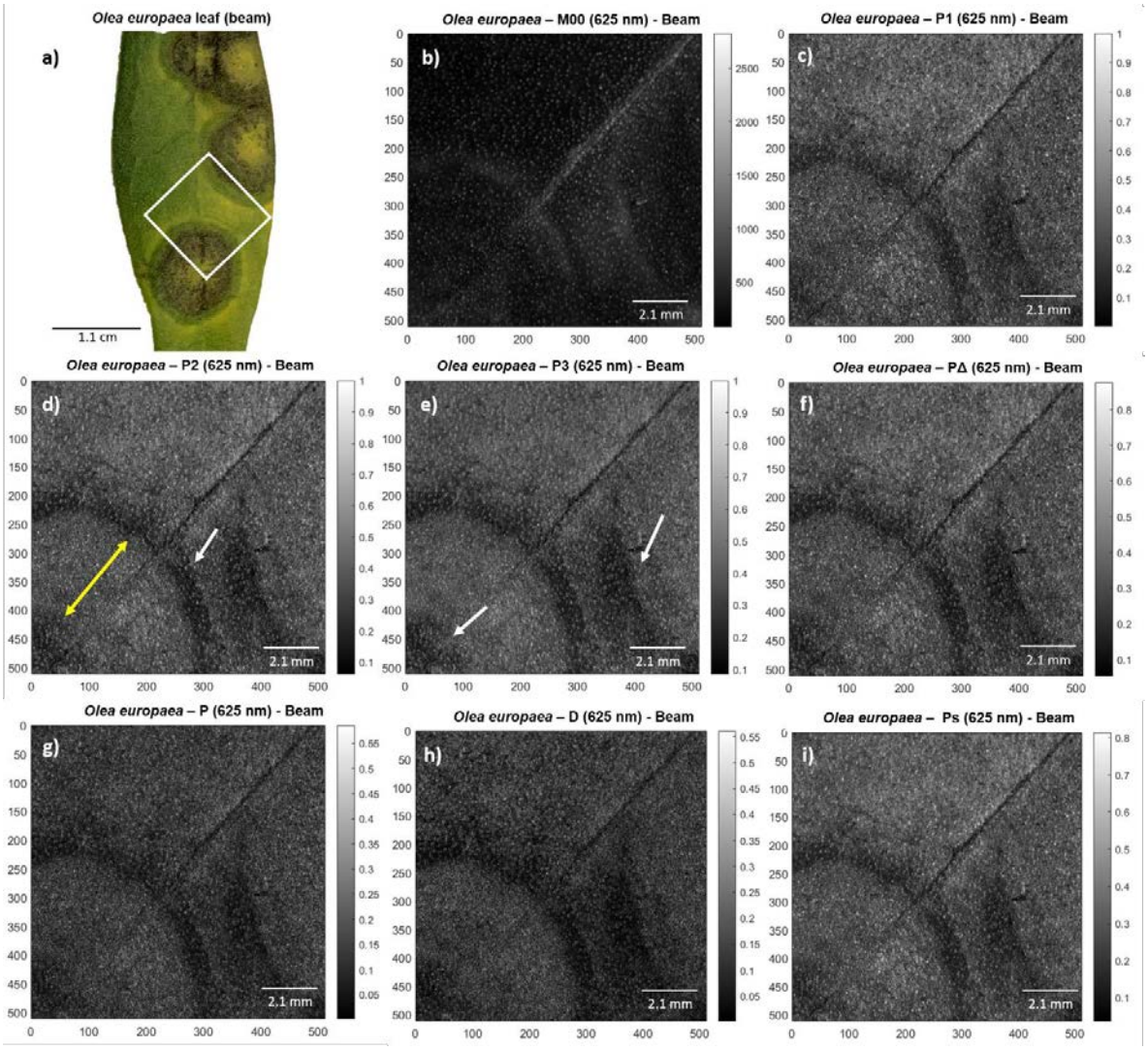


Figure S2. Images of polarimetric observables of *Olea europaea* leaf. a) Picture of the beam part of the *O. europaea* leaf used in this study. White square denotes for selected region of interest (ROI) analyzed in remaining images, b) regular intensity image (M_{00}) of the *O. europaea* beam ROI and its corresponding polarimetric observables c) P_1, d) P_2, e) P_3, f) P_Δ, g) P, h) D and i) P_S for visual comparison. All images correspond to 625 nm illumination wavelength measurements performed at scattering set-up configuration. Yellow and white arrows indicate the locations of necrotic ring-like and other injured tissue regions (chlorotic spot and halo) within the sample.

3. Exploratory analysis of polarimetric observables data

This section is devoted to demonstrate the reliability of the method used for the selection of the polarimetric observables triplet. As described in the main manuscript, our purpose is to achieve a pseudo-colored image which provides significant results in terms of disease symptom visualization, i.e., to enhance visual differentiation between healthy and diseased tissue regions of inspected plants. To this aim, we decided to base the selection of the observable triplet in terms of observable median values leading to largest polarimetric differences between healthy and wounded regions of inspected specimen leaves. Under this scenario, it is important to know how our data distributions behave and if we are dealing with reliable median values. To do so, we performed the Boxplot analysis [2,3] (Fig. 5 of the main manuscript) for the selected purity spaces (P_1,

P_2, P_3, P, D and P_S) regarding different tissue conditions regions (healthy, chlorotic and necrotic, shown in Fig. 4 of the main manuscript) of *M. sativa* and *O. europaea* specimens. In this context, we studied the percentage of outliers as well as the standard errors of the medians for each observable data distribution.

3.1 Outliers inspection

Regarding Boxplot charts (Fig. 5 of the main manuscript), the polarimetric data distributions for each type of tissue condition (healthy, chlorotic and necrotic) present a certain amount of outlier values, which are known as particular observation data points which lie at an abnormal distance (1.5 times –mild– or 3 times –extreme– the distance between first and third quartile) from the first or third quartiles of the distribution (illustrated as circles –mild– and stars –extreme–, respectively, in Fig. 5 of the main manuscript).

The calculated outlier percentages for each polarimetric data observable and for each tissue condition studied are provided in Table S2. Importantly, note that a low outlier percentage from a polarimetric observable distribution within a given tissue condition indicates that only few data points differ, in an abnormal way, from the whole tendency. Consequently, we can demonstrate the homogeneity selection of the different tissue regions (healthy, chlorotic and necrotic). In particular, the minimum percentage of outliers (0.44%) is demonstrated for the spherical degree, P_S, distribution within the healthy region of the *O. europaea* leaf, whereas the largest (3.67%) lies in the distribution of the index of polarimetric purity P_1 for the chlorotic tissue region of the *M. sativa* inspected leaf. From data in Table S2 we can ensure the validity of most polarimetric data used for the study presented in this manuscript.

Table S2. Polarimetric observables outliers in percentage (%) for healthy and diseased regions in *Medicago sativa* and *Olea europaea* inspected leaves. Maximum and minimum outlier % values are highlighted in gray.

			P_1	P_2	P_3	P	D	P_S
Medicago sativa	**% Outliers**	**Healthy**	1.92%	1.39%	1.21%	1.32%	1.21%	1.32%
		Chlorosis	3.67%	1.50%	2.00%	1.16%	2.08%	3.08%
Olea europaea	**% Outliers**	**Healthy**	0.58%	1.03%	0.97%	0.77%	0.58%	0.44%
		Chlorosis	0.82%	0.89%	1.21%	0.57%	1.03%	0.93%
		Necrosis	1.24%	1.03%	1.63%	0.90%	1.39%	1.30%

3.2 Analysis on median values

Once the validity of the selected tissue regions of both inspected samples is demonstrated, we can ensure the correct treatment of polarimetric data as well as its interpretation in terms of discriminative potential. Regarding the last, we visually deduce from median values extended red-dotted lines in Fig. 5 of the main manuscript, the potential of several polarimetric observables to set apart the different type of tissues (healthy, chlorotic and necrotic) within a sample. This is because in the Boxplot analysis, if the median of one tissue does not coincide within the box of another tissue for a given polarimetric observable, those tissues can be interpreted as statistically discriminable within such observable (i.e., the median value of one class, falls out the 75% of data corresponding to the other class).

For the correct quantification of the visual discrimination for both *M. sativa* and *O. europaea* samples, the calculated median values and the corresponding associated errors for each polarimetric observables (P_1, P_2, P_3, P, D and P_S) and type of tissue (healthy, chlorotic and necrotic) are presented in Table S3. In this context, the median values for the studied polarimetric observables present, overall, low standard deviations of the median. This behavior is correlated with reliable polarimetric median values, allowing us to perform comparatives between the discriminative potential of the different metrics (P_1, P_2, P_3, P, D and P_S) through bloxplot analysis.

Table S3. Polarimetric observables median values and standard deviations on healthy, chlorotic and necrotic regions of the studied leaves of *Medicago sativa* and *Olea europaea*.

		P_1	P_2	P_3	P	D	P_S
Medicago sativa	**Healthy**	0.187 ± 0.034	0.243 ± 0.041	0.498 ± 0.047	0.117 ± 0.025	0.102 ± 0.022	0.236 ± 0.031
	Chlorosis	0.278 ± 0.077	0.363 ± 0.095	0.612 ± 0.094	0.169 ± 0.048	0.144 ± 0.036	0.325 ± 0.074
Olea europaea	**Healthy**	0.033 ± 0.013	0.068 ± 0.018	0.012 ± 0.027	0.023 ± 0.010	0.024 ± 0.010	0.055 ± 0.014
	Chlorosis	0.012 ± 0.004	0.029 ± 0.007	0.057 ± 0.010	0.022 ± 0.007	0.019 ± 0.004	0.018 ± 0.004
	Necrosis	0.087 ± 0.027	0.149 ± 0.036	0.238 ± 0.052	0.036 ± 0.016	0.035 ± 0.016	0.125 ± 0.028

3.3 Pseudo-coloration functions implementation

In this section we present the detailed description of the performed steps regarding the pseudo-coloration of *M. sativa* and *O. europaea* inspected samples, they being based on two triplets of polarimetric observables, (P_2, P_3, P_S) and (P_1, P_2, P_3).

The main goal is to achieve an image visual enhancement of different plant tissue conditions (healthy and diseased regions) when compared with non-pseudocolored polarimetric images. To construct the pseudo-colored images, as explained in the main text, we need to define some threshold values discriminating between tissues conditions (chlorosis, healthy and necrosis), which were derived from Boxplot analysis (see Fig. 5 in main manuscript). In the following, we explain in detail the method to set the different thresholds, corresponding to each polarimetric observable contained into a given triplet.

Note that the different tissue conditions are well differentiated when represented within the boxplot charts (see Fig. 5 in main manuscript). Taking adventatge of this situation, we can set a threshold able to discriminate between each pair of tissue conditions in a binary way. In the case of the *M. sativa* sample, we only need to discriminate between chlorotic and healthy tissues, so only one threshold value is required. In turn, in the case of the *O. europaea* sample, we need to discriminate between healthy, chlorotic and nectrotic tissues, so two thresholds are required: (1) chlorotic-healthy and (2) necrotic-healthy thresholds (note that healthy tissues gets intermediate values between chlorotic and nectrotic tissues, so no more threshods are needed – see Fig. 4 in main manuscript). In addition, this threshold assignment must be repeated for each polarimetric observable present in one of the studied triplets (P_1, P_2, P_3, and P_S). Under this scenario, for a particular observable and for a particular case between two tissue conditions to be discriminated, let us call them condtion A and B, we calculated, from the Boxplot chart, the observable value of the first quartile (bottom side of the box), for the tissue condition with the highest mean value between A and B, as well as the observable value of the third quartile (upper side of the box), for the tissue condition with the smallest mean between A and B. Finally, the threshold value is directly obtained by conducting the difference, in each case, between these two quartile values (third and first). Obtained quartile values for the different polarimetric observables (columns) and for the different tissue conditions (rows) present in the two studied samples (*M. sativa* and *O. europea*) are provided in Table S4. Resulting thresholds are presented in following Table S5.

Table S4. First and third quartiles for chlorotic, healthy and necrotic tissue distributions for each polarimetric observable of *Medicago sativa* and *Olea europaea* leaf studied samples.

		P_1		P_2		P_3		P_S	
		Q_1	Q_3	Q_1	Q_3	Q_1	Q_3	Q_1	Q_3
Medicago sativa	**Healthy**	-	0.209	-	0.271	-	0.531	-	0.271
	Chlorotic	0.232	-	0.303	-	0.552	-	0.303	-
Olea europaea	**Chlorotic**	-	0.015	-	0.034	-	0.064	-	0.022
	Healthy	0.024	0.043	0.056	0.082	0.104	0.141	0.046	0.066
	Necrotic	0.069	-	0.126	-	0.207	-	0.108	-

Table S5. Definition of leaf lesion types by means of threshold pixel values for chlorotic, healthy and necrotic tissues for each polarimetric observable of *Medicago sativa* and *Olea europaea* leaf studied samples.

		Threshold: P_1	**Threshold: P_2**	**Threshold: P_3**	**Threshold: P_S**
Medicago sativa	**Healthy**	$P_1 < 0.220$	$P_2 < 0.287$	$P_3 < 0.542$	$P_S < 0.269$
	Chlorotic	$P_1 \geq 0.220$	$P_2 \geq 0.287$	$P_3 \geq 0.542$	$P_S \geq 0.269$
Olea europaea	**Chlorotic**	$P_1 \leq 0.020$	$P_2 \leq 0.045$	$P_3 \leq 0.084$	$P_S \leq 0.034$
	Healthy	$0.056 > P_1 > 0.020$	$0.104 > P_2 > 0.045$	$0.174 > P_3 > 0.084$	$0.087 > P_S > 0.034$
	Necrotic	$P_1 \geq 0.056$	$P_2 \geq 0.104$	$P_3 \geq 0.174$	$P_S \geq 0.087$

Note that in a boxplot representation, the box sides (upper and bottom) define the first and third quartiles respectively, and thus the 50% of data corresponding to a given distribution falls into the box. Therefore, above the first quartile, as well as under the third quartile, there is represented the 75% of the data related to a given data distribution. Importantly, the thresholds constructed by using the above-stated approach, are set halfway between the first and third quartiles of the particular pair of tissues being discriminated, and thus, this construction ensures a tissue condition recognition in all the cases superior to the 75% in each polarimetric channel studied. This is always true if the boxes corresponding to the different tissues to be discriminated are not overlapped in a particular polarimetric observable (as it is the case of the observables selected for the triplets, see Fig. 5 of the main manuscript). In fact, this recognition rate should be even larger if taking into account that the the remaining data (25% of data, which is represented in the whiskers) of each tissue is generally not linearly distributed throught the whiskers but especially close to the box sides (i.e., certain data between boxes above/below the threshold is still well recognized). For the sake of clarity, a representative graphic example of quartiles selection and threshold location is shown in Fig. S3.

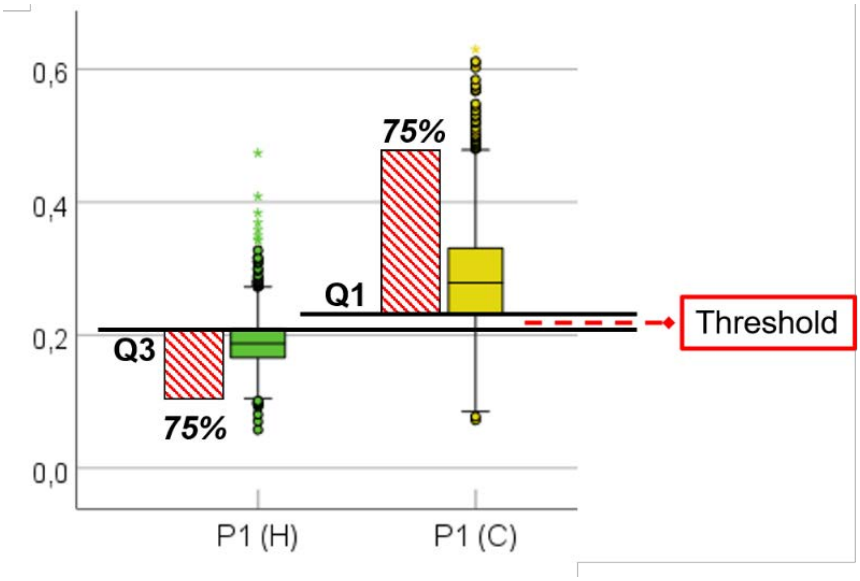


Figure S3. Representative example of quartile values location and threshold calculation for chlorotic and healthy P_1 distributions of the *Medicago sativa* inspected sample. The relative position of first and third quartiles (labels Q1 and Q3, respectively) for healthy (H) and chlorotic (C) distributions, are indicated by means of continuous black lines, whereas the location of the corresponding threshold (mean value between Q3 and Q1) is red-lined. Red-dashed boxes illustrate the percentage of data explained under the third quartile and above the first of the healthy and chlorotic polarimetric observable P_1 data distribution.

Once the thresholds were obtained as previously discussed (see values in Table S5), they were used to implement the pseudo-colored images. To do so, we associated each one of the three primary colors (R, G, B) to the different tissue conditions of the inspected samples: chlorotic regions were associated to red, healthy tissues to green and necrotic ones to blue. Afterwards, a binary colored image (color/black) is constructed for each tissue condition: values above/below the corresponding threshold are assigned to the particular tissue condition color (R-chlorotic, G-healthy, B-necrotic) or to black, respectively. Consequently, for each particular polarimetric observable P_i ($i=1, 2, 3, S$), we construct three independent images, one per primary color (R, G and B), each one carrying the information corresponding to a given tissue condition. Afterwards, the three binary colored-images corresponding to a particular polarimetric observable P_i, are added by means of a linear combination, leading to a first pseudo-colored image, as showin in Eq. (S1):

$$P_{i,approach}(x,y) = P_{i,Red-Chlorotic}(x,y) + P_{i,Green-Healthy}(x,y) + P_{i,Blue-Necrotic}(x,y). \quad \text{(S1)}$$

Therefore, according to Eq. (S1), the first-approach pseudo-colored functions for the different polarimetric observables in the case of *M. sativa* are defined as:

$$\begin{aligned}
P_{1,approach,MS}(x,y) &= P_{1,R-Chlorotic}(x,y) + P_{1,G-Healthy}(x,y),\\
P_{2,approach,MS}(x,y) &= P_{2,R-Chlorotic}(x,y) + P_{2,G-Healthy}(x,y),\\
P_{3,approach,MS}(x,y) &= P_{3,R-Chlorotic}(x,y) + P_{3,G-Healthy}(x,y),\\
P_{S,approach,MS}(x,y) &= P_{S,R-Chlorotic}(x,y) + P_{S,G-Healthy}(x,y),
\end{aligned} \quad \text{(S2)}$$

and for the case of *O. europaea*:

$$\begin{aligned}
P_{1,approach,OE}(x,y) &= P_{1,\mathrm{R}-Chlorotic}(x,y) + P_{1,G-Healthy}(x,y) + P_{1,B-Necrotic}(x,y),\\
P_{2,approach,OE}(x,y) &= P_{2,R-Chlorotic}(x,y) + P_{2,G-Healthy}(x,y) + P_{2,B-Necrotic}(x,y),\\
P_{3,approach,OE}(x,y) &= P_{3,R-Chlorotic}(x,y) + P_{3,G-Healthy}(x,y) + P_{3,B-Necrotic}(x,y),\\
P_{S,approach,OE}(x,y) &= P_{S,R-Chlorotic}(x,y) + P_{S,G-Healthy}(x,y) + P_{S,B-Necrotic}(x,y).
\end{aligned} \tag{S3}$$

where *MS* and *OE* sub-labels denote for *M. sativa* and *O. europaea*, respectively, and *R, G* and *B* for red, green and blue primary colors associated to the chlorotic, healthy and necrotic tissue regions, respectively. Note that the term regarding to necrotic tissues in *M. sativa* (Eq. (S2)) is equal to zero ($P_{i,\ Blue\text{-}Necrotic} = 0$) because there is no necrotic content within the inspected sample. Afterwards, and recalling the selected polarimetric triplets (i.e., (1) the Components of Purity: P_2, P_3, P_S and; (2) the IPPs: P_1, P_2, P_3), each of the RGB-layered images obtained from Eqs. (S2) and (S3) is put together by means of a second linear combination. In this way, the particular information of each polarimetric observable into a triplet is combined in a final RGB-layered pseudo-colored image. The final image linear combinations for the triplets (P_2, P_3, P_S) and (P_1, P_2, P_3) for *M. sativa* and *O. europaea* inspected samples are presented in the following equations (S4) and (S5), respectively:

$$\begin{aligned}
Pseudo\#1(x,y)_{MS} &= P_{2,approach,MS}(x,y) + P_{3,approach,MS}(x,y) + P_{S,approach,MS}(x,y),\\
Pseudo\#2(x,y)_{MS} &= P_{1,approach,MS}(x,y) + P_{2,approach,MS}(x,y) + P_{3,approach,MS}(x,y),
\end{aligned} \tag{S4}$$

$$\begin{aligned}
Pseudo\#1(x,y)_{OE} &= 2\times[P_{2,approach,OE}(x,y) + P_{3,approach,OE}(x,y) + P_{S,approach,OE}(x,y)],\\
Pseudo\#2(x,y)_{OE} &= 2\times[P_{1,approach,OE}(x,y) + P_{2,approach,OE}(x,y) + P_{3,approach,OE}(x,y)],
\end{aligned} \tag{S5}$$

where *MS* and *OE* sub-labels denote for *M. sativa* and *O. europaea*. For simplicity, we chose the weigths for polarimetric triplets to be the unit. Nevertheless, the overall low values on polarimetric observables on *O. europaea* lead to a dark tone final image. To improve the visualization, the weights were pondered by a factor 2.

For visual comparison, in Fig. S4 we present the regular intensity images, the images for the indices of polarimetric purity P_2 and P_1, and the obtained pseudo-colored images by means of performing the above-presented relationships in Eqs. (S4) and (S5) for *M. sativa* and *O. europaea*. The final pseudo-coloration clearly demonstrates, as initially stated, the image visual enhancement between pathology symptoms (injured regions) and healthy tissue within the inspected samples. A detailed discussion of the obtained results is provided in the discussion section of the main manuscript.

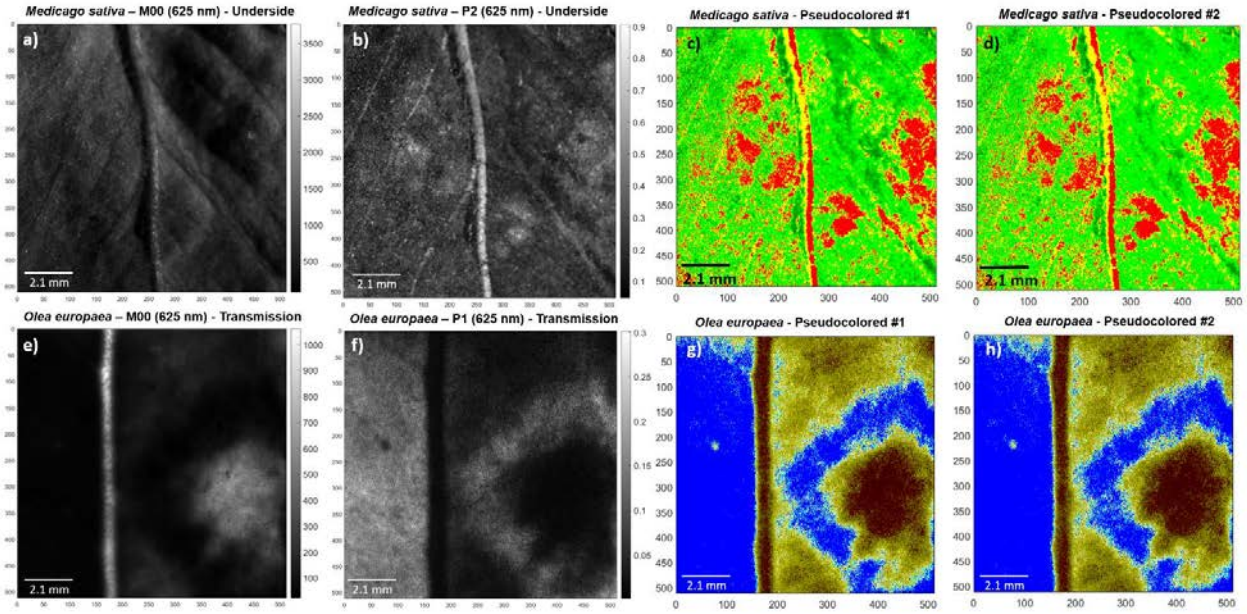


Figure S4. Visual comparison of *Medicago sativa* leaf (top row): a) 625 nm intensity image (M_{00}), b) polarimetric purity index P_2, c) processed image by means of #1 pseudo-coloration and d) processed image by means of #2 pseudo-coloration; Visual comparison of *Olea europaea* leaf (bottom row): e) 625 nm intensity image (M_{00}), f) polarimetric purity index P_1, g) processed image by means of #1 pseudo-coloration and h) processed image by means of #2 pseudo-coloration.

4. Complete image Mueller polarimeter description

The experimental Mueller matrices of the collected samples were acquired by means of a complete image Mueller polarimeter working at three different illumination wavelengths (625 nm, 530 nm and 470 nm), covering the visible range, and two measurement configurations: scattering and transmission. Importantly, the polarimeter used in this work consists of two independent optical systems gathered into two mobile compact arms, the polarization state generator (PSG) and the polarization state analyzer (PSA). The optical design for PSG consists on the arrangement of a linear polarizer oriented at 0º with respect to the laboratory vertical, followed by two Parallel Aligned Liquid Crystal (PA-LC) retarders at 45º and 0º, respectively. PSA optical set-up has the same elements as PSG but placed in the inverse order, but additionally placing a CCD camera for the capture of the sample intensity. This architecture leads to the capability of generating and analyzing, respectively for the PSG and PSA systems, any fully polarized state. Remark that scattering measurements are performed by illuminating the sample by placing the PSG at 34º with respect to the laboratory horizontal and the PSA in vertical position to avoid the direct reflections and collect scattered light. Transmission configuration is characterized by placing both PSG and PSA at 0º with respect to laboratory horizontal. Visual representation of both optical set-up configurations as well as the inner components is shown in Figs. S5 and S6, respectively.

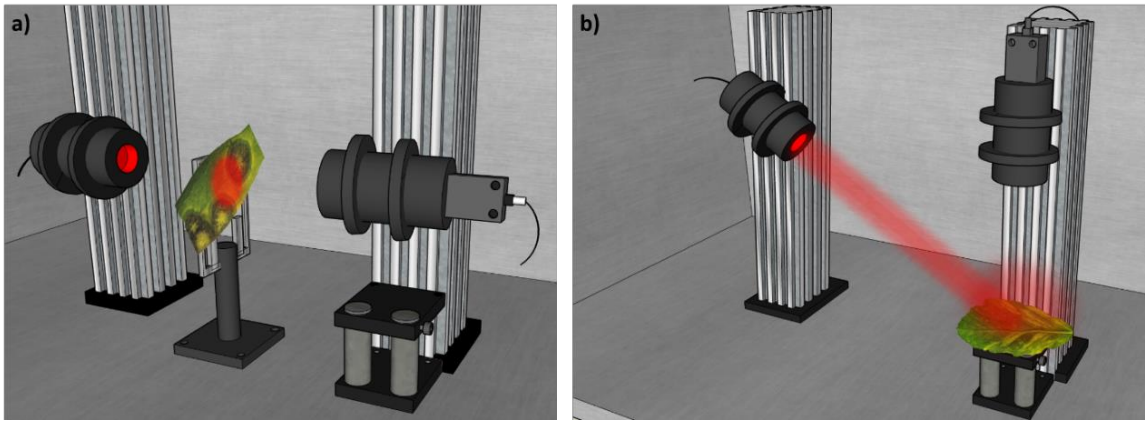


Figure S5. 3D representation of the complete image Mueller polarimeter used in this study at a) transmission configuration and b) scattering configuration.

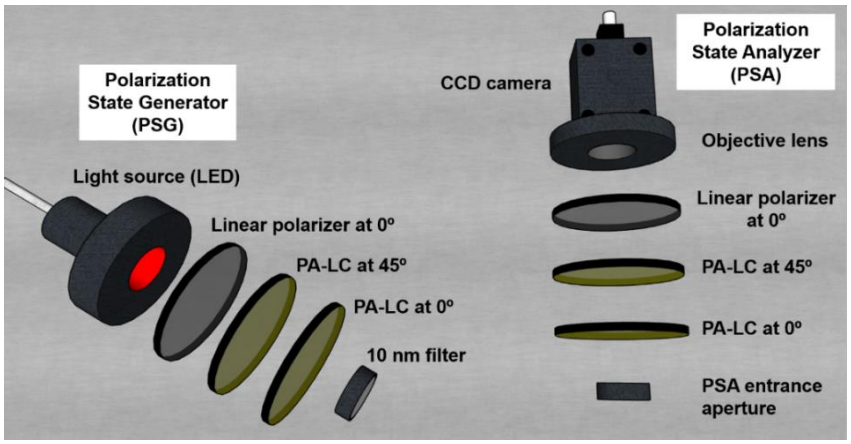


Figure S6. 3D representation of the polarization state generator (PSG) and polarization state analyzer (PSA) optical components arrangement.

In the following we provide the detailed information about the optical components within the polarimeter: the illumination is performed by means of we use a four-wavelength high-power Thorlabs LED source (LED4D211, operated by DC4104 drivers distributed by Thorlabs) complemented with 10 nm dielectric bandwidth filters distributed by Thorlabs: FB530-10 and FB470-10 for green and blue wavelengths, respectively. Imaging is performed by means of a 35 mm focal length Edmund Optics TECHSPEC® high resolution objective followed by an Allied Vision Manta G-504B CCD camera, with 5 Megapixel GigE Vision and Sony ICX655 CCD sensor, 2452(H) × 2056(V) resolution and cell size of 3.45 µm × 3.45 µm, so a spatial resolution of 22 µm is achieved. Regarding the two different linear polarizers: the one arranged on PSG is a Glan-Thompson prism-based CASIX whereas the placed in PSA is a dichroic sheet polarizer distributed by Meadowlark Optics. The four Parallel Aligned Liquid Crystals are Variable Retarders with Temperature Control (LVR-200-400-700-1LTSC distributed by Meadowlark Optics).

5. References

1. Gil, J. J. & Ossikovski, R. *Polarized Light and the Mueller Matrix Approach* (CRC Press, Boca Raton, 2016).

2. Tukey, J. W. *Exploratory Data Analysis* (Addison-Wesley, 1977).

3. McGill, R., Tuckey, J. W. & Larsen, W. A. Variations of Box Plots, *The Am. Stat*. **32**(1), 12-16; 10.2307/2683468 (1978).

PAPER B

Automatic pseudo-coloring approaches to improve visual perception and contrast in polarimetric images of biological tissues

Carla Rodríguez[1,*], Albert Van Eeckhout[1,2], Enrique Garcia-Caurel[2], Angel Lizana[1] and Juan Campos[1]

[1] *Grup d'Òptica, Physics Department, Universitat Autònoma de Barcelona, Bellaterra 08193, Spain*
[2] *LPICM, CNRS, Ecole Polytechnique, Institut Polytechnique de Paris, Palaiseau 91120, France*

This paper has been published in *Scientific Reports*, **12**, pp. 18479, 2022

scientific reports

Check for updates

OPEN

Automatic pseudo-coloring approaches to improve visual perception and contrast in polarimetric images of biological tissues

Carla Rodríguez[1]✉, Albert Van Eeckhout[1,2], Enrique Garcia-Caurel[2], Angel Lizana[1] & Juan Campos[1]

Imaging polarimetry methods have proved their suitability to enhance the image contrast between tissues and structures in organic samples, or even to reveal structures hidden in regular intensity images. These methods are nowadays used in a wide range of biological applications, as for the early diagnosis of different pathologies. To include the discriminatory potential of different polarimetric observables in a single image, a suitable strategy reported in literature consists in associating different observables to different color channels, giving rise to pseudo-colored images helping the visualization of different tissues in samples. However, previous reported polarimetric based pseudo-colored images of tissues are mostly based on simple linear combinations of polarimetric observables whose weights are set ad-hoc, and thus, far from optimal approaches. In this framework, we propose the implementation of two pseudo-colored methods. One is based on the Euclidean distances of actual values of pixels and an average value taken over a given region of interest in the considered image. The second method is based on the likelihood for each pixel to belong to a given class. Such classes being defined on the basis of a statistical model that describes the statistical distribution of values of the pixels in the considered image. The methods are experimentally validated on four different biological samples, two of animal origin and two of vegetal origin. Results provide the potential of the methods to be applied in biomedical and botanical applications.

Contrast in an image can be understood as the difference in either luminance (grey level) or color, that makes an object, or a region embedded in a portion of said image, to be distinguishable from the surrounding objects or regions. In biological samples, variations of luminosity or color of images are due to changes in the amount of light being absorbed, reflected or scattered by the probed sample. Since such fluctuations are related in a more or less involved different properties of the sample (composition, organization, order, structure….) then it can be said that absorption, reflection or scattering are sources of contrast that may allow to perceive variations in the physical properties of the imaged scene. The way in which a sample absorbs, reflects or scatters light not only depends on its intrinsic properties but also depends on the illumination conditions set to probe such sample. The simplest and maybe the most standard way of illuminating an object is to use partially coherent unpolarized, either polychromatic or monochromatic, light. Unfortunately in many cases, such a simple way of illumination may lead to poorly contrasted images which do not provide enough visual contrast. In those cases the use of more sophisticated methods of illumination, such as polarized light imaging, may be employed in order to boost the contrast in the images and therefore to retrieve information from probed samples with enhanced accuracy. The use of polarized light has been shown to be useful to reveal differences in the composition, thickness, density and structural organization as, for instance, in some tendinous structures in animal samples[1] or some organelles in plants[2–4]. This is because, when illuminating these biological structures with polarized light, the light matter interactions lead to very different polarimetric properties (in retardance, dichroism and/or depolarization), which are directly related with their structural nature[1,5]. For instance, the anisotropic behavior of the plant vascular

[1]Optics Group, Physics Department, Universitat Autònoma de Barcelona, 08193 Bellaterra, Spain. [2]LPICM, CNRS, Ecole Polytechnique, Institut Polytechnique de Paris, 91120 Palaiseau, France. ✉email: carla.rodriguez@uab.cat

structures (mainly composed by organized and well-aligned cellulose chains[6–8]), or the animal structures built-in collagen (e.g., tendons), induce retardance to the polarization of the incident light beam.

One of the main advantages of polarization is the high sensitivity of the technique to differences in composition and structural organization within the inspected biological tissues. Among others, polarimetry allows the imaging of nerve fiber bundles in the human brain[9] and the inspection of Alzheimer disease[10]. Polarization-related techniques also demonstrate high accuracy in the early detection of some cancers such as skin cancer[11], colon cancer[12], breast cancer[13,14] and brain cancer[15]. Moreover, recent studies reveal the accurate performance of some machine learning algorithms built-in the polarimetric analysis of some biological tissues[16–20]. This situation may benefit the implementation of fast in-vivo and non-invasive pathology recognition methods, even at early stages. Analogously, the use of these optical techniques in plant science has proved to be very useful, as vegetal structures present polarimetric signatures (for instance, dichroism and birefringence) which can be potentially exploited for characterizing the spatial organization of some vegetal organelles such as the thylakoid membranes[2] or the cell wall composition[8]. Recently, in addition to dichroism and birefringence, the use of depolarization observables to characterize structures in vegetal samples, is arousing growing interest[3,4].

Keeping in mind the suitability of polarimetry for organic tissues characterization, it is feasible to achieve the overall enhancement of image contrast, including the revelation of regular intensity-hidden inherent structures, by conducting polarimetric measurements of biological tissues (either animal or vegetal). Polarimetric measurements result in a diversity of polarimetric figures (depolarization, diattenuation, polarizance, linear retardance, etc.)[21–23] that may provide relevant information of tissue structures but which are commonly visualized as separated information channels. With the aim of enhancing the visualization of tissue imaging, recent works[1,24,25] suggest the construction of a pseudo-colored image whose layers contain the different polarimetric features of the sample, i.e., different polarimetric information origins are included all together in a single enhanced image. The main idea in these works is to design a pseudo-colored function based on the weighted combination of three different polarimetric observables showing the highest image contrast between the structures of interest within the sample and associate each chosen observable to a primary color, RGB (red, green and blue). So far, the methods proposed to build polarimetric based pseudo-colored images are based on the qualitative (visual) selection of the polarimetric observables and their relative weights, and so, the final pseudo-colored model is quite arbitrary and not optimal in most of cases.

In this work we present, for the first time, two pseudo-coloring models which are designed based on two different methods that allow maximizing the visual contrast of different tissues in the sample. The two applied methods for building pseudo-colored images are based on: (1) the Euclidean distance between polarimetric values of different tissues; and (2) the Normal (Gaussian) function based on polarimetric data to estimate the probability of belonging to a particular tissue. Based on previous studies showing the special suitability of some depolarizing spaces in terms of biological samples discrimination[1,3–5], we selected the Indices of Polarimetric Purity (IPPs)[26] and the Components of Purity (CPs)[21] spaces as the variables to be implemented within the pseudo-colored functions. Note that these two polarizing spaces are complementary, and fully describe the depolarizing properties of samples[27]. For completeness, we provide the comparison between pseudo-colored model results, based on the Euclidean and the Normal-based approaches, when analyzing different organic samples. Overall, images resulting from the pseudo-colored methods presented in this study overcome the regular polarimetric observables in terms of spatial location, visualization and recognition of the chosen structures within the inspected samples.

Results

This section aims to show the results obtained with the Euclidean and Normal-based pseudo-coloring methods when used to inspect diverse types of biological samples (animal and vegetal samples). In both methods, the construction of the corresponding pseudo-coloring functions is based on the selection of a triplet of polarization-based figures or observables. As previously mentioned, we selected depolarization-related observables because as reported in literature[16,17,28–30], and based in our previous experience we know that they are suitable to characterize biological samples[1,3–5]. To reinforce this argument, in section 1 of Supplemental document, we show the images of different biological structures related to a representative collection of different polarimetric observables used in the literature, as well as the images corresponding to IPPs and CPs, to show how the latter spaces give rise to greater visual contrast. In particular, two different scenarios were studied: (1) pseudo-coloring based on the Indices of Polarimetric Purity (IPPs, labeled as P_1, P_2 and P_3); and (2) pseudo-coloring based on the Components of Purity (CPs, labeled as P, D and P_Δ). These depolarization-based figures, which are discussed in detail in the "Methods" section, are selected here because they provide a complete description of the depolarization of light by the depolarization-related information of probed samples[27]. For illustrative purposes, in this section we show an application of the two image processing methods discussed here to four different types of samples, two of them were animal tissues and the other two were vegetal tissues. The two animal samples were biopsies from a lamb trachea (showing the trachea ring and the sheath), and a lamb tongue (showing the lingual papillae and epithelial tissue) respectively. The two vegetal tissues were taken from leaves of a specimen of *Quercus pubescens* presenting powdery mildew lesion caused by the fungus *Erysiphe alphitoides* (leaf powdery mildew vs leaf lamina), and, (4) leaves from a specimen of *Vitis vinifera* showing no symptoms of disease (leaf vein vs raphides vs cell cluster). Note that in the case of the *V. vinifera* leaf, we simultaneously analyzed three features (leaf vein/raphide/cell cluster) to study the ability of the proposed methods to characterize samples with more than two different features.

The pseudo-colored imaging construction procedure is summarized as follows. First, the experimental Mueller matrix (MM) image of all the studied samples was measured with a complete imaging polarimeter fully described in "Methods" section. From MM images, according to the mathematical background described in "Methods" section, we retrieved images of the depolarization-related observables corresponding to the

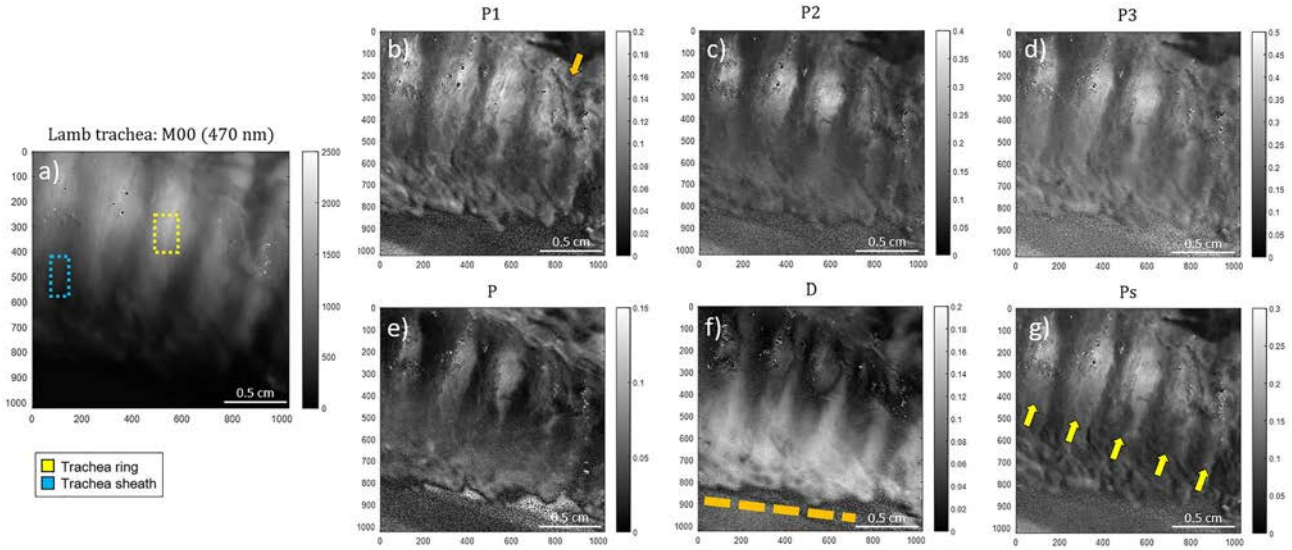


Figure 1. Images of a lamb trachea measured at 470 nm illumination wavelength: (**a**) unpolarized intensity image (M_{00}), the Indices of Polarimetric Purity (**b**) P_1, (**c**) P_2 and (**d**) P_3 and the Components of Purity (**e**) P, (**f**) D and (**g**) P_S. The yellow dotted box (Fig. 1a) and yellow arrows (Fig. 1 g), indicate the location of the cartilaginous rings. The dotted blue box (Fig. 1a) indicates the location of the trachea sheath. The orange arrow (Fig. 1b) and the orange dotted line (Fig. 1f.) show the vascular structure within the external trachea sheath and the sample border, respectively.

polarimetric triplets of IPPs (P_1, P_2 and P_3) and CPs (P, D and P_S), i.e., six polarimetric images were derived from each experimental MM of a sample. From each one of these observables, we built four pseudo-colored images per sample, according to the approaches described in "Methods" section. These four pseudo-colored images, correspond to the Euclidean-based or Normal-based designs. In a final step the pseudo-colored images of each observable obtained with the two methods are compared to each other in order to select the one which higher contrast, and therefore the one which offers the cleared description of the studied sample. In the following, we provide the results obtained for the study of the above-stated biological samples. The analysis of the results is provided in the "Discussion" section.

Polarimetric inspection of animal samples. The two animal samples discussed here correspond to a lamb trachea and a lamb tongue. The images shown are representative of the different images that we took from the same biopsy. The experimental data were taken in the scattering configuration, with a spectrally filtered blue light at 470 nm. The size of the probed samples was 2.2×2.2 cm^2 and the images were taken using a camera with 1024×1024 pixels. Figure 1 shows the intensity image of the lamb trachea obtained with unpolarized light (M_{00}, Fig. 1a; i.e., unpolarized intensity image) and the retrieved depolarization observables-based images corresponding to the IPPs (P_1, P_2, P_3 in Fig. 1b–d) and CPs (P, D, P_S in Fig. 1e–g).

As previously mentioned, the following step is to construct pseudo-colored images based on the observables of the two chosen triplets of depolarization-based observables, the IPPs (P_1, P_2 and P_3) and the CPs (P, D and P_S), and to use them to apply the pseudo-color approach based on the Euclidean and the Gaussian approaches. Therefore, this procedure leads to four different pseudo-colored images for each sample studied (i.e., Euclidean model for (1) IPPs and (2) CPs plus Normal model for (3) IPPs and (4) CPs). In this subsection we focus on the results obtained for the animal samples (lamb trachea and lamb tongue samples).

In the particular case of the lamb trachea sample, we choose the trachea cartilaginous rings (marked in yellow in Fig. 1a) and the sheath (marked in blue in Fig. 1a) as the features to be differentiated within the resulting pseudo-colored image. Under this scenario, we associate the trachea rings with the coordinates, $\overrightarrow{C^{\text{Ring}}} = [1,\ 1,\ 0]$ corresponding to the yellow color in the standard RGB color space. Analogously, the trachea sheath is associated with the RGB coordinates corresponding to the blue color, $\overrightarrow{C^{\text{Sheath}}} = [0,\ 0,\ 1]$. Figure 2 shows the pseudo-colored images of the lamb trachea in Fig. 1 resulting from the application of the two coloring approaches. The unpolarized image is also shown for visual comparison. The selected ROIs used for reference in the computations are indicated by the yellow and blue squares in Fig. 2a, respectively. Figure 2b and d correspond to the Euclidean pseudo-coloring for the polarimetric triplets IPPs and CPs, respectively. In counterpart, Fig. 2c and e illustrate the Normal pseudo-coloring for the polarimetric observables IPPs and CPs, respectively.

Likewise, we apply the same polarimetric analysis above-described, but this time to the lamb tongue. Figure 3 presents typical images of the tongue (Fig 3a corresponds to the unpolarized diffuse reflectance, M_{00}, and the corresponding polarimetric observables P_1, P_2, P_3, P, D and P_S can be seen in Fig. 3b–g respectively).

The two most well-differentiated characteristics within the lamb tongue correspond to the lingual papillae and epithelial tissue. Accordingly, we chose these features as the structures to be differentiated in the pseudo-colored

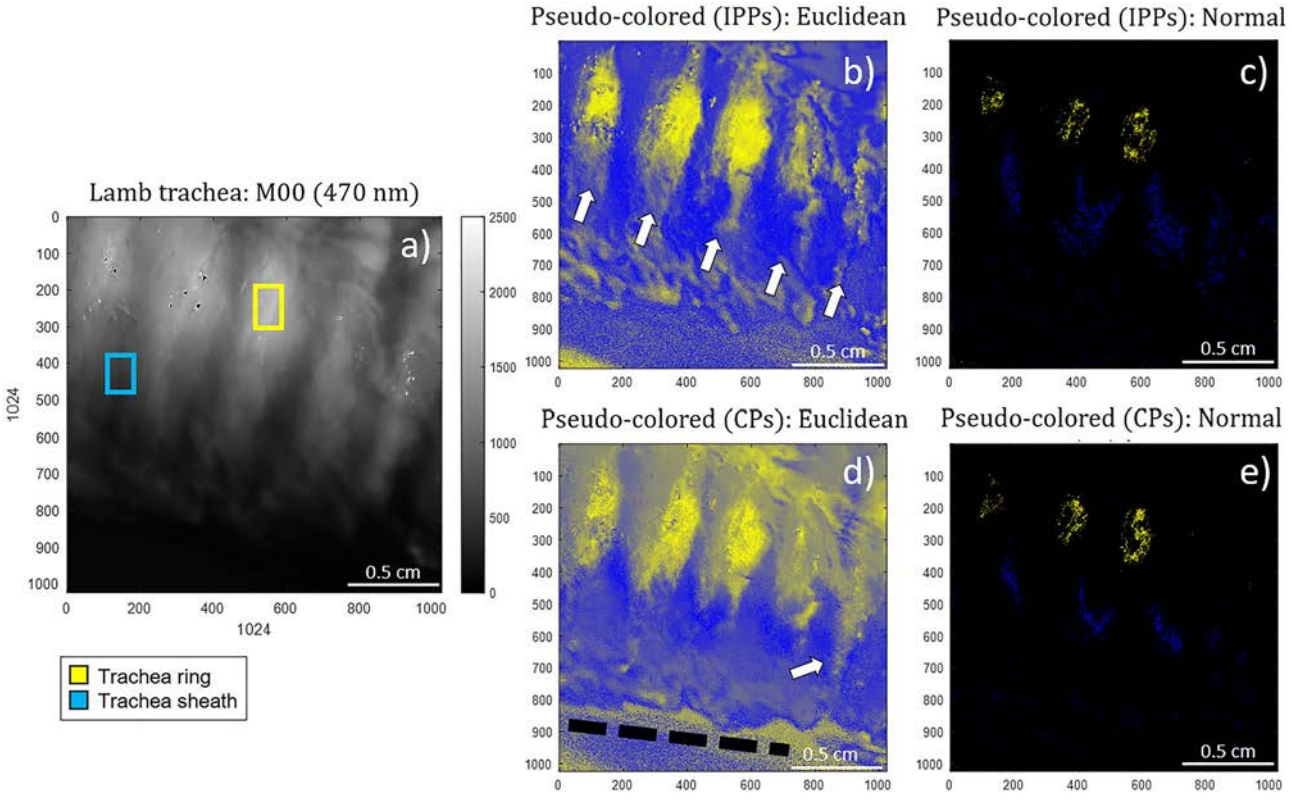


Figure 2. Raw and pseudo-colored images of the lamb trachea: (**a**) Unpolarized intensity image (M_{00}) taken at 470 nm, (**b**) Euclidean and (**c**) Normal pseudo-colored images based on the IPP triplet, (**d**) Euclidean and (**e**) Normal pseudo-coloring based on the CP triplet. Yellow and blue squares show reference areas corresponding to the trachea ring and sheath, respectively. The white arrows (Fig. 2b and d) and the black dotted line (Fig. 2d) denote for the cartilaginous rings and the trachea border, respectively.

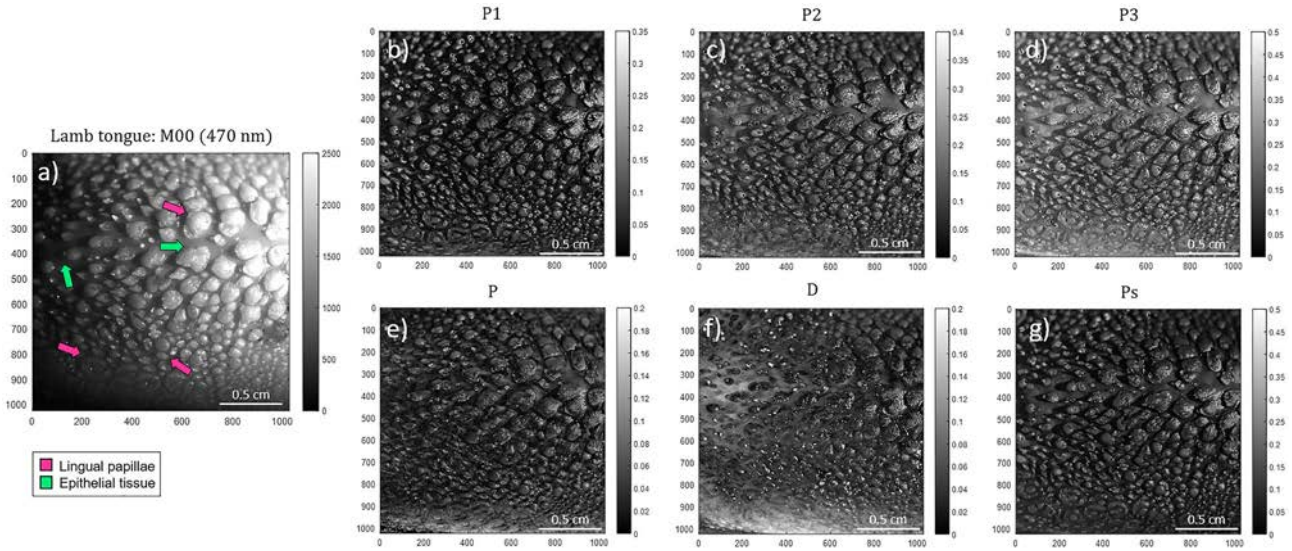


Figure 3. Polarimetric images of the lamb tongue measured at 470 nm illumination wavelength: (**a**) regular intensity image (M_{00}), the Indices of Polarimetric Purity (**b**) P_1, (**c**) P_2 and (**d**) P_3 and the Components of Purity (**e**) P, (**f**) D and (**g**) P_S. The pink and lime-green arrows in Fig. 3a indicate the location of some of the lingual papillae and epithelial tissue regions, respectively.

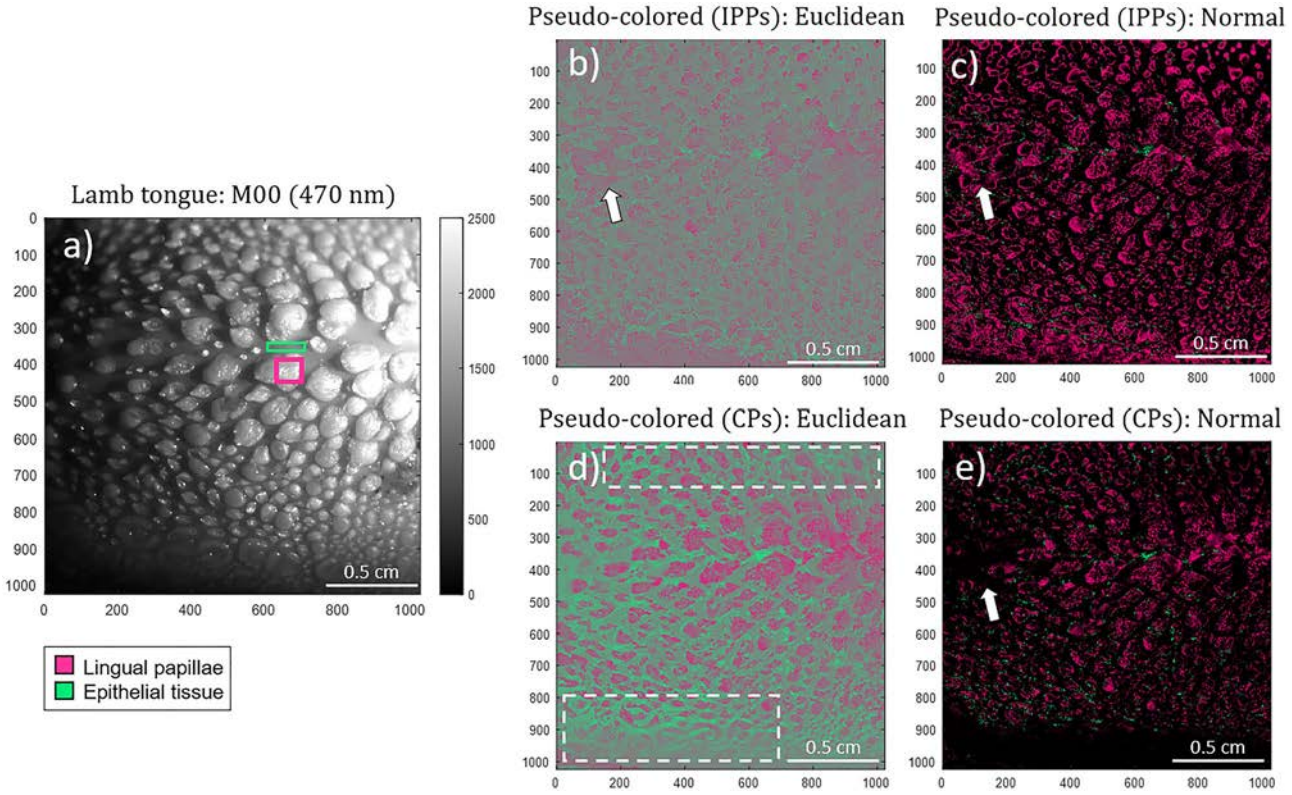


Figure 4. Intensity and pseudo-colored images for the inspected lamb tongue: (**a**) regular intensity image (M_{00}) captured at 470 nm illumination wavelength, (**b**) Euclidean and (**c**) Normal pseudo-colored images for IPPs, (**d**) Euclidean and (**e**) Normal pseudo-coloring for CPs. Pink and green squares (Fig. 4a) denote for the selected regions of interest (ROI) of lingual papillae and epithelial tissue, respectively. White arrows (Fig. 4b, c and e) indicate a particular region only containing epithelial tissue. The dotted-squares indicate the unseen / out of focus region (by means of the unpolarized image, M_{00}) of the tongue.

images. In particular, we associate the polarimetric information of the lingual papillae with pink color with the following RGB coordinates $\overrightarrow{C^{\text{Papillae}}} = [1,\ 0,\ 0.5]$, and the epithelial tissue with the RGB coordinates corresponding to lime-green color with the RGB coordinates $\overrightarrow{C^{\text{Epithelial}}} = [0,\ 1,\ 0.5]$, (both arrowed with their corresponding color in Fig. 3a). Figure 4 shows the diffuse reflectance image M_{00} (Fig. 4a) and the images resulting from the implemented pseudo-colored functions. The ROIs corresponding to lingual papillae and epithelial tissue are indicated in Fig. 4a with pink and lime-green squares, respectively. Figure 4b and d correspond to the Euclidean pseudo-coloring of the inspected sample for the depolarization spaces IPPs and CPs, respectively. Besides, the Normal pseudo-coloring for the polarimetric triplets IPPs and CPs is presented in Fig. 4c and e, respectively.

Polarimetric inspection of vegetal samples. In this section we discuss the results obtained from two different plants; one was a leaf of *Q. pubescens* and the other was a leaf of *V. vinifera*. The leaf of *Q. pubescens* show powdery mildew lesions caused by the fungus *E.rysiphe alphitoides* while the leaf of *V. vinifera* did not show any sign of infection or parasitic invasion. The leaf *Q. pubescens* was measured with the same imaging polarimeter used to obtain the images discussed in the previous section. The typical size of the field of view of the images measured with such polarimeter is1.1×1.1 cm^2 (512×512 pixels). The experimental Mueller matrices of the leaf of *V. vinifera* were measured with a polarimetric microscope in transmission at a wavelength of 533 nm. The field of view of the images taken by the microscope correspond to a circle of radius 100 µm. The non-polarized (transmission, diffuse reflectance) intensity image (M_{00}) and the retrieved depolarization observables (P_1, P_2, P_3, P, D and P_S) for the *Q. pubescens* sample are presented in Fig. 5.

In order to explore the ability of pseudo-colored images to enhance the visual contrast between healthy and infected areas of the leaf, we chose to inspect the characteristics corresponding to the powdery mildew lesions and the healthy leaf lamina. In particular, we associate the polarimetric information of the powdery mildew with the yellow color, with RGB coordinates $\overrightarrow{C^{\text{Powdery}}} = [1,\ 1,\ 0]$, and the leaf lamina with blue color with RGB coordinates, $\overrightarrow{C^{\text{Lamina}}} = [0,\ 0,\ 1]$. Note that the chosen colors, blue and yellow are complementary to each other, therefore they even the slightest shadow of them can be can be visually discriminated with ease. Figure 6 presents the visual comparison between the M_{00} image (Fig. 6a) and the four pseudo-colored images based on polarimetric observables for the inspected leaf of *Q. pubescens*. As discussed in previous section the pseudo-colored images

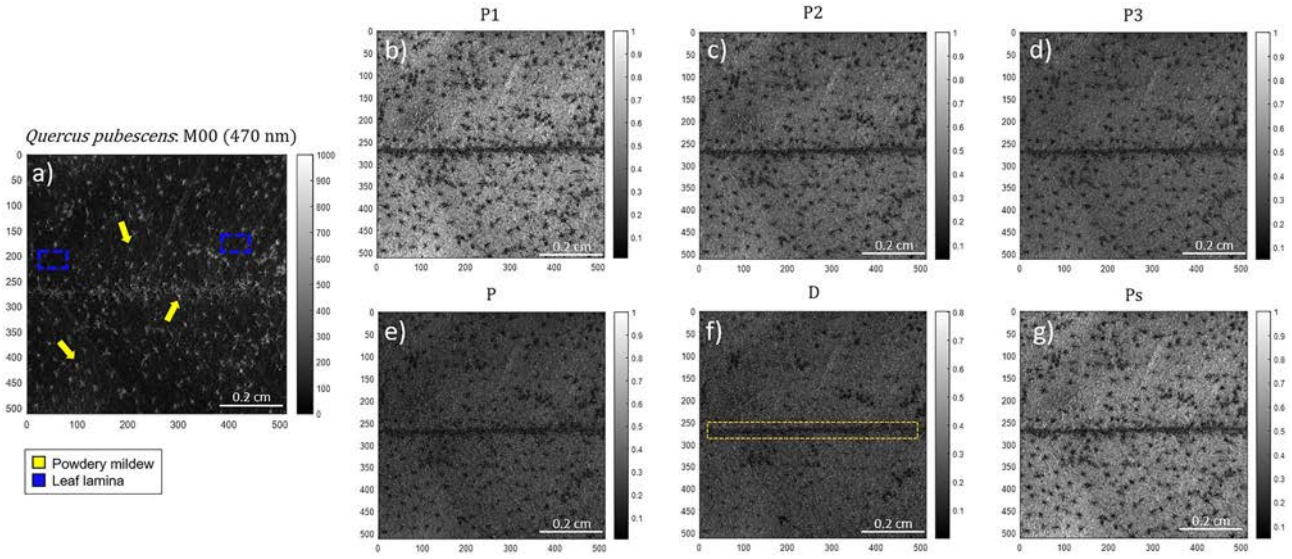


Figure 5. Polarimetric images of the Quercus pubescens leaf measured at 470 nm illumination wavelength: (**a**) regular intensity image (M_{00}), the Indices of Polarimetric Purity (**b**) P_1, (**c**) P_2 and (**d**) P_3 and the Components of Purity (**e**) P, (**f**) D and (**g**) P_S. The yellow arrows, and the blue dotted box (both in Fig. 5a) denote for some of the powdery mildew lesions and the leaf lamina location, respectively. The orange-dotted box (Fig. 5f.) indicates the location of the leaf vein.

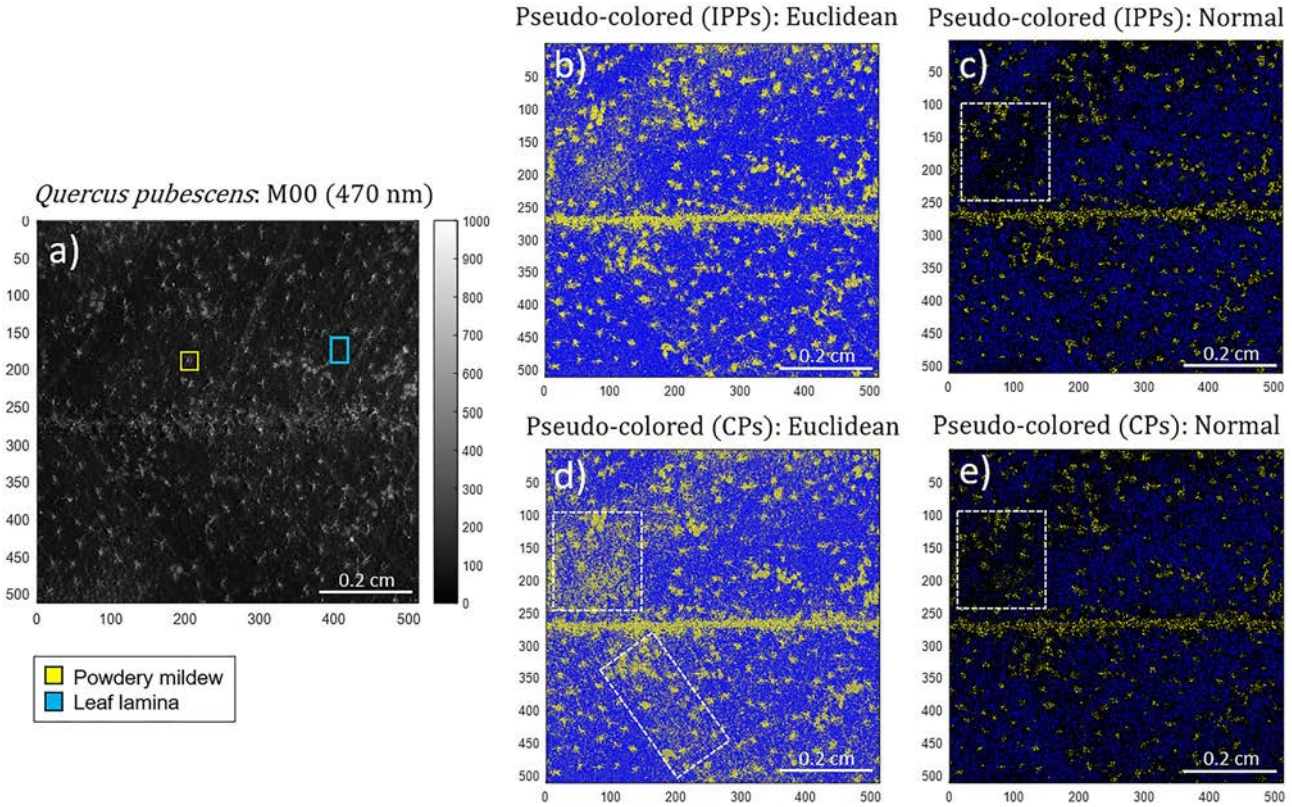


Figure 6. Intensity and pseudo-colored images for the inspected *Quercus pubescens* leaf: (**a**) non-polarized (transmission—diffuse reflection) image (M_{00}) captured at 470 nm illumination wavelength, (**b**) Euclidean and (**c**) Normal pseudo-colored images for IPPs, (**d**) Euclidean and (**e**) Normal pseudo-coloring for CPs. Yellow and blue squares (Fig. 6a) denote for the selected regions of interest (ROI) of powdery mildew lesion caused by the fungus *Erysiphe alphitoides* and the healthy leaf lamina, respectively. White-dotted squares (Fig. 6c and e) denote for the misrecognized pixel regions of both Euclidean and Normal-based methods.

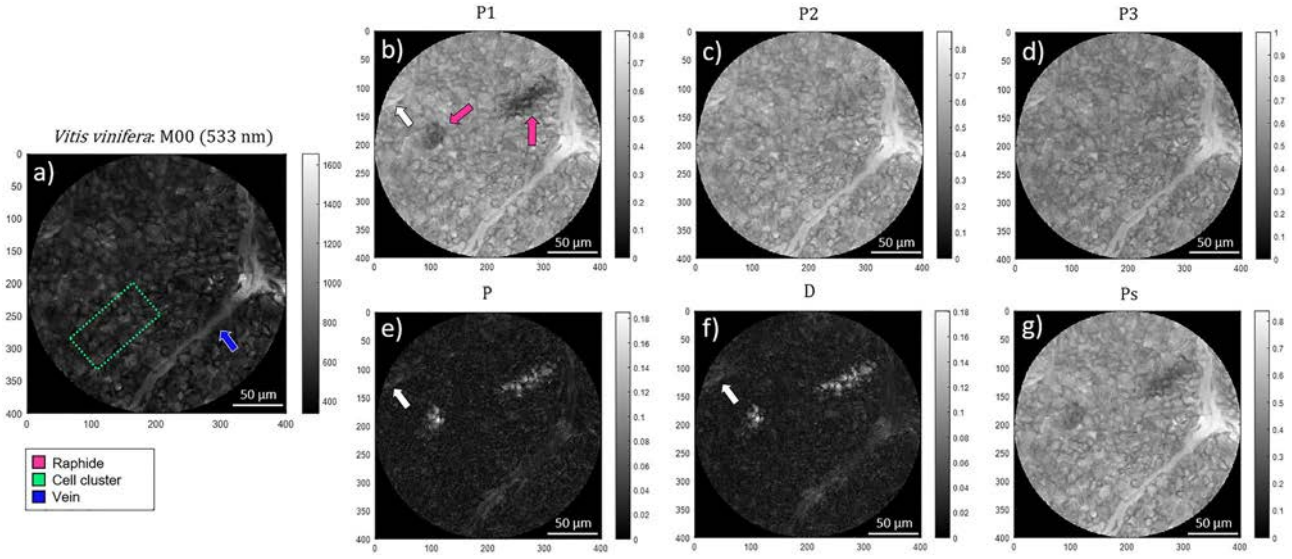


Figure 7. Polarimetric images of the *Vitis vinifera* leaf measured under the microscope for 533 nm illumination wavelength: (**a**) non-polarized transmission image (M_{00}), the Indices of Polarimetric Purity (**b**) P_1, (**c**) P_2 and (**d**) P_3 and the Components of Purity (**e**) P, (**f**) D and (**g**) P_S. The blue and pink arrows (Fig. 7a, and b) indicate the location of the vein and the raphides, respectively. The white arrows (Fig. 7b, e and f) indicate the secondary vascular structure. The lime-green dotted box indicates an illustrative region comprising a cell cluster.

were obtained using the functions described in Methods with the above-stated $\overrightarrow{C^{\text{Powdery}}}$ and $\overrightarrow{C^{\text{Lamina}}}$ values. The selected regions of interest (ROI) corresponding to powdery mildew and leaf lamina, are indicated in Fig. 6a with yellow and blue squares, respectively. The resulting Euclidean and Normal pseudo-coloring for the polarimetric spaces IPPs and CPs are shown in Fig. 6b and d and Fig. 6c and e, respectively.

Regarding to the inspection of *V. vinifera* leaf, we chose to differentiate three main structures present in the leaf, the raphides, the vegetal cells walls of the leaf lamina and the venous system, each one associated with the pink, lime-green and blue colors corresponding to the RGB coordinates $\overrightarrow{C^{\text{Raphide}}} = [1,\ 0,\ 0.5], \overrightarrow{C^{\text{Cell}}} = [0,\ 1,\ 0.5]$, and $\overrightarrow{C^{\text{Vein}}} = [0,\ 0,\ 1]$, respectively. Figure 7 presents the regular intensity image (M_{00}) and the retrieved depolarization observables (P_1, P_2, P_3, P, D and P_S) for the *V. vinifera* leaf.

Since the Euclidean distances approach is valid for just two classes, we decide to choose cell walls of the leaf lamina. In contrast, the Normal approach accepts more than two classes, so with this method we can visualize, simultaneously, the three above-mentioned features: the raphides, the leaf lamina cells and the vein. In this case, the pseudo-colored approach results in the images shown in Fig. 8. This example allows us to show the ability to recognize tissues that are invisible in non-polarimetric images thanks to the use of polarimetric images. This is because some biological structures analyzed are invisible in the intensity image (Fig. 7a) and therefore, the ROIs used to train the model are directly obtained from polarimetric channels, and in particular, from the P_1 image which was the one giving the largest visual contrast. In particular, Fig. 8 presents a visual comparison between the image corresponding to the polarimetric purity index P_1 (Fig. 8a) and the resulting images from both the Euclidean and the Normal-based pseudo-coloring methods for the *V. vinifera* inspected section. As said, note that unlike previous cases, in this case we select the polarimetric purity index P_1 (Fig. 8a) as reference instead of the intensity image (Fig. 7a). This is because some structures of interest (the raphides in this case) are visible thanks to the P_1 channel, but completely invisible in the intensity imae (Fig. 7a), and therefore, corresponding ROIs were obtained from Fig. 8a. For the sake of clarity, the ROIs of the structures corresponding to the raphide, cell cluster and leaf vein, are indicated in Fig. 8a with pink, lime-green and blue squares, respectively. Figure 8b and d illustrate the Euclidean pseudo-coloring of the sample by means of the IPPs and CPs observables, respectively. Figure 8c and e show the Normal pseudo-coloring for the IPPs and CPs implementation, respectively.

Discussion

In the following, we present a discussion regarding the pseudo-colored images of different animal tissues presented in the previous "Results" section. In this work, we choose the polarimetric channels that give larger contrast between tissues of interest, and apply them within diverse pseudo-coloration methods. Note that we do not select the intensity image as one channel for the pseudo-coloration approach. The loss of the intensity channel is a choice in our case, in favor of a more adequate space to enhance the contrast between different tissues. We use the RGB color space mainly because it is made up of three orthogonal variables that vary from 0 to 1 and the IPPs are restricted to the same value range. However, we want to note that the IPPs are not completely independent since the value of P_3 channel determines the range of P_2 and P_1[21]. For this reason, another color composition may also be useful (or more optimal) to represent the IPPs space.

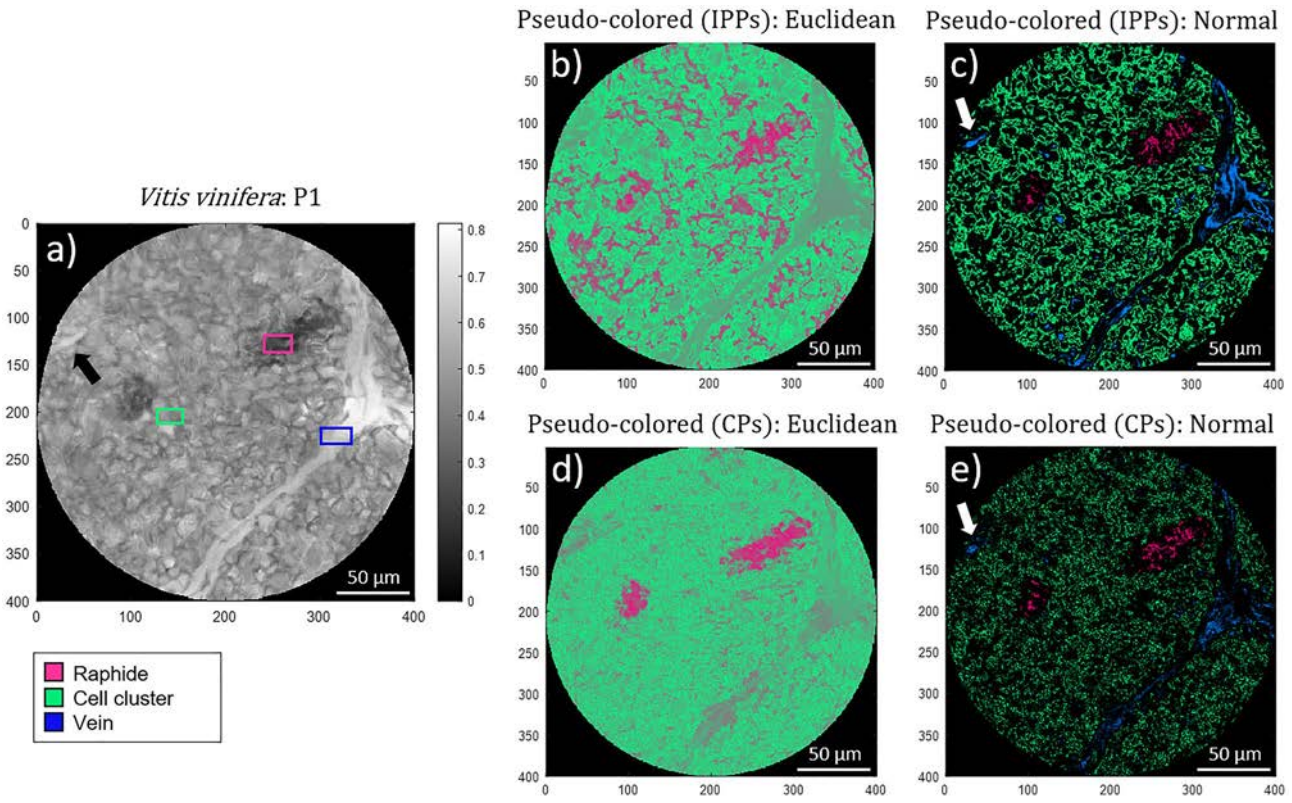


Figure 8. Non-polarized transmission (M_{00}) and pseudo-colored microscopic images from a *Vitis vinifera* leaf section captured at 533 nm illumination wavelength: (**a**) polarimetric purity index P_1, (**b**) Euclidean and (**c**) Normal pseudo-colored images for IPPs, (**d**) Euclidean and (**e**) Normal pseudo-coloring for CPs. Pink, lime-green and blue squares denote for the selected regions of interest (ROI) of raphides, leaf lamina cells and vein, respectively. The white arrows indicate the location of the secondary vascular structure.

In the case of the lamb trachea sample (Fig. 1), the two most well-differentiated structures for the lamb trachea in the non-polarized reflectance image (Fig. 1a) correspond to the cartilaginous rings, composed by hyaline cartilage, and the external sheath, also known as tunica adventitia, mainly composed by collagen (both structures are indicated in Fig. 1a with yellow and blue squares, respectively). These main differences in tissue composition give raise to the different polarimetric responses explored and discussed in this work. Figure 1 clearly shows how the images corresponding to the enpolarization metrics (P_1, P_2, P_3, P, D and P_S in Fig. 1b–e) point well the difference between the two types of tissues and also present an enhanced contrast when compared to the non-polarized diffuse reflectance (M_{00} in Fig. 1a). Regardless of the model used to assign colors to pixels, the key to ensure a vivid contrast enhancement using a pseudo-colored approach from several channels is the fact that a given channel must be sensitive to characteristics of the sample that the other channels are not. For instance, in the particular case of the trachea, P_1 provides information about the surface details of the trachea sheath (e.g. some vascular structures within the external sheath, indicated with the orange arrow in Fig. 1b), while P_2 and P_3 are sensitive to the cartilaginous rings (Fig. 1c–d). Interestingly, the cartilaginous rings demonstrate higher mean values for IPPs ($P_1 = 0.14 \pm 0.01$, $P_2 = 0.27 \pm 0.01$ and $P_3 = 0.37 \pm 0.01$) when compared with the trachea sheath ($P_1 = 0.05 \pm 0.01$, $P_2 = 0.10 \pm 0.01$ and $P_3 = 0.18 \pm 0.01$), this meaning that the cartilaginous rings are less depolarizing than the trachea sheath. Accordingly, cartilaginous rings demonstrate higher mean polarizance P (Fig. 1e) values ($P = 0.06 \pm 0.01$) than the sheath ($P = 0.02 \pm 0.01$). In contrast, the sheath induces more diattenuation ($D = 0.10 \pm 0.01$) than the rings ($D = 0.05 \pm 0.01$) (Fig. 1f). Note that in all the cases, polarizance and diattenuation channels present values lower than 0.1. This situation indicates that the trachea can be considered as a non-dichroic structure and that the observed depolarizing effects can be mostly associated with either multiple scattering or fluctuations in the value and the direction of the retardance of birefringent structures present in cartilaginous rings and the trachea sheath. In this regard, retardance is encoded in the sphericity degree P_S and, as shown in Fig. 1g, it highlights the details of the trachea rings (mean value of $P_S = 0.2 \pm 0.01$, see the yellow arrows in Fig. 1g) in addition to the surface of the sheath (mean value of $P_S = 0.07 \pm 0.01$). Interestingly, note that the IPPs provide information about the structure of the polarimetric randomness of samples, and that they are directly connected with the CPs through the relation[27]: $P^2 + D^2 + 3P_s^2 = 2P_1^2 + \frac{2}{3}P_2^2 + \frac{1}{3}P_3^2$. This fact, together with the above-discussed low values for P and D in the trachea sample, suggests that high values of P_S may be related to the fact that the cartilaginous structure induces less depolarization according to the strong alignment of collagenous fibers. In contrast, the low Ps value of the sheath demonstrates the higher depolarizing behavior of that tissue. Importantly, all the polarimetric observables unveil the trachea borders (see the orange dotted line in Fig. 1f) not seen in the intensity image (Fig. 1a).

Regarding to the pseudo-colored images, the Euclidean distance-based method for both polarimetric triplets of IPPs and CPs (Fig. 2b and , respectively) enhances the perceived contrast between cartilaginous rings and trachea sheath. Conversely, the Normal-based approach (based on IPPs and CPs in Fig. 2c and e) has not enough capability to improve visualization, as most pixels in the image are not recognized as part of any of the two classes (i.e., the probability, calculated through the normal function, of belonging to either the ring or the sheath is almost zero). Therefore, according to Eq. (10) in "Methods" section, those pixels are represented in black. Concerning the observables used to implement the pseudo-coloring model based on Euclidean distances, the use of the IPPs set of observables seem to well differentiate the five rings (indicated with white arrows in Fig. 2b) from the sheath, thus allowing a clear an accurate identification and spatial localization of the cartilaginous tissue within the sample. In turn, the same pseudo-coloring model based on the CPs set of observables (Fig. 2d) also succeeds to give a clear discrimination between the cartilaginous rings and the trachea sheath but the transition between these two classes is less accurate than in the IPPs case (i.e., the trachea sheath occupying the space between rings is misrecognized as cartilaginous tissue too). Despite this, the fifth trachea ring (white arrow in Fig. 2d) and the full trachea structure border (black dotted line in Fig. 2d) are both better spatially located in the CPs case. In summary because the poor ability of the coloring based on the Normal model to discriminate among different classes, the most adequate method to visually inspect structures within the lamb trachea sample is the Euclidean distances approach.

The second case of study is the lamb tongue. Likewise to the case of the lamb trachea above-discussed, we retrieve, from the experimental Mueller matrix measurement of the lamb tongue, the isolated polarimetric observables images (P_1, P_2, P_3, P, D and P_S). In this sample, we focus on two tissue classes: the lingual papillae and the epithelial tissue (some of those are indicated with pink and green-lime arrows in Fig. 3a, respectively). Regarding the depolarizing content, the mean values for lingual papillae and epithelial tissue are: P_1=0.09±0.01, P_2=0.14±0.01, P_3=0.21±0.01, and, P_1=0.03±0.01 (Fig. 3b), and P_2=0.10±0.01 (Fig. 3c) and P_3=0.20±0.01 (Fig. 3b,c and d), respectively. In overall, both the lingual papillae and the epithelial tissue demonstrate similar low IPPs values, which translate into the tissues having high depolarizing capability. In addition, the epithelial tissue and lingual papillae also show similar (and low) mean polarizance P values: P=0.01±0.01 and P=0.04±0.01, respectively (Fig. 3e). In turn, the diattenuation D becomes higher at sample borders (see the bottom region in Fig. 3f), reaching mean values of D=0.13±0.01 compared with the D=0.05±0.01 of the central region. The values of polarizance and diattenuation obtained for the lingual papillae and epithelial tissue demonstrate that they are mostly non-dichroic structures. Therefore, depolarization observed can be associated to either the effect of multiply scattered light or the fluctuation in the value and the direction of the birefringent structures. As in the case of the trachea sample, due to this non-dichroic behavior, the P_S values obtained for the tongue demonstrates more depolarizing behavior for the epithelial tissue (lower values of P_S=0.08±0.01) than for the papillae (P_S=0.12±0.01), this leading to a visual contrast between these two classes (Fig. 3g). Importantly, we want to highlight a significant improvement provided by some of the depolarizing channels (Fig. 3b–g) with respect to non-polarized diffuse reflectance image (Fig. 3a). In particular, note that the tongue structures placed at the bottom part of the non-polarized diffuse reflectance image are very difficult, or even impossible to be seen in some parts (check for instance the region between the Y axis pixels 800 and 1000) due to the low contrast between them. Same effect can be observed at the very top of the image. This is mostly due to intensity losses or defocusing introduced by the measure of a non-planar sample (the tongue), where only the central region is properly illuminated and in focus. In contrast to that, a clear visual enhancement of those regions is provided by depolarization observables, specially by P_1 and P_S channels, in which the lingual papillae and the epithelial tissue are clearly observed in the whole image. The latter may be explained by the fact that contrast in polarimetric images is less sensitive to focus than in non-polarized intensity images[3] and thus, lead to this image improvement. Indeed, while a fine tune of focus increases sharpness, and thus contrast in non-polarized intensity images, the relative difference between the polarimetric response of adjacent zones is what provides contrast in a polarization-based image[3]. This situation is one of the reasons that explains the interest of using polarimetric methods for tissue characterization.

Afterwards, we implement the pseudo-coloring functions for the lamb tongue sample. Accordingly, the lingual papillae and the epithelial tissue were associated with two different colors: pink and lime-green, respectively. We see how the best results, in terms of visualization, is obtained for the Euclidean pseudo-coloring method based on the CPs triplet (Fig. 4d). In agreement with the discussion related to previous Fig. 3, note how tongue structures out of focus are now well identified and discriminated in the pseudo-colored images, as they are based on polarimetric observables (see Fig. 3b–g) not in intensity gray levels such as the non-polarized diffused reflectance observable (Fig. 3a). In the case of the pseudo-coloring based on the Normal method (Fig. 4c and e), the best visualization is obtained for the implementation based on the IPPs triplet (Fig. 4c). However, the class recognition rate is lower than for the Euclidean distance case (see for instance white arrows in Fig. 4c and e, pointing the lingual papillae recognition for the Normal method, compared with correct recognition of the epithelial tissue in Fig. 4b, by the Euclidean distance method). In the case of the pseudo-coloring based on the Normal method, the mean values selected for the epithelial tissue (obtained from the selected ROIs, green rectangle in Fig. 4a) are not representative enough of the properties of the epithelial tissue, which present a large variance across the image. For this reason, an important part the epithelial could not be successfully assigned to the correct class by the method and therefore appear in black in the pseudo-colored images (Fig. 4c and e). Under this scenario, it is shown the pseudo-coloring based on the Euclidean approach performs better than the pseudo-coloration based in the Normal approach. In order to improve the performance of the pseudo-coloring based in Normal methods, it may be possible, for instance, to select multiple ROIs of each class tissue across the image and then to evaluate the corresponding mean values. Unfortunately, the latter may result in an obvious complication of the method which is to be compared to the simplicity shown by the Euclidean approach (one single ROI assignment per tissue class).

In the following, we switch to the inspection of plant tissues. Following the same protocol that we applied to animal tissues in previous cases, we measured the experimental Mueller matrix corresponding to a leaf of a specimen of *Q. pubescens* showing powdery mildew lesions due to the infection of the fungus *E. alphitoides*. Accordingly, the two classes to be differentiated within the sample correspond to the healthy tissue (leaf lamina) and the regions showing symptomatology (powdery mildew). For the sake of clarity, in Fig. 5a we present the intensity image of the leaf, and we highlight some of the locations corresponding to the lesions caused by the pathogen (yellow arrows) and the regions corresponding to healthy tissue (blue dashed squares). The evident differences in biological structure and chemical composition between a plant tissue, the leaf of *Q. pubescens*, and a fungus, the *E. alphitoides*, lead to different polarimetric responses. The polarimetric images based on the IPPs (Fig. 5b–d) demonstrate the enhancement of the overall image contrast, thus allowing a proper spatial localization of the leaf lesions. Among the IPPs, P_1 is the observable that leads to larger contrast between classes, and demonstrates a higher capability to discriminate between features (Fig. 5b), followed by P_2 and P_3 (Fig. 5c and d, respectively). Regarding the depolarization content of the inspected sample, the healthy leaf lamina shows high IPP mean values: P_1=0.82±0.01, P_2=0.89±0.01 and P_3=0.99±0.01. Conversely, the powdery mildew shows lower mean values: P_1=0.27±0.01, P_2=0.43±0.01 and P_3=0.60±0.01. Accordingly, it can be said that the leaf lamina induces less depolarization to the incident light than the powdery mildew lesions. The low depolarizing performance of the leaf lamina suggests that it presents a well-organized cell layout within the leaf, and a homogeneous polarization response through the structure. In contrast, the effect of the fungus seems to modify the cell layout structure of the leave, leading to an evident modification of the polarimetric response of the regions with lesions. This different polarimetric response between healthy and infected regions may also be observed in the polarizance P channel (Fig. 5e): the healthy lamina shows mean polarizance value of P=0.35±0.01 and the powdery mildew of P=0.12±0.01. Note that values of polarizance larger than 0.3 show a non-negligible dichroic response of the vegetal cells in the *Q. pubescens* leaf, being this dichroic capability significantly reduced after sample infection. Regarding to the diattenuation response (Fig. 5f), the sample demonstrates, overall, low mean values (D=0.13±0.01 and D=0.23±0.01 for leaf lamina and powdery mildew, respectively). Finally, among the presented CPs, the sphericity degree P_S is the observable whose performance better enhances the image contrast between the inspected features (see Fig. 5g): healthy lamina shows mean value of P_S=0.86±0.01 and powdery mildew, P_S=0.34±0.01. The differences in P_S values of the healthy and infected regions may be due to modifications in the cell layout organization (cell alignment) and/or within their polarimetric behavior (mainly birefringence or polarizance features). Furthermore, it is important to mention that the black horizontal line (indicated with the orange botted box in Fig. 5f) corresponds to an underlying principal vein of the leaf.

Concerning the pseudo-coloring of the image of the *Q. pubescens* leaf, we associate the yellow and blue colors to the powdery mildew and leaf lamina classes respectively. When using the pseudo-coloring method based on the Euclidean distance approach (Fig. 6b and d) an obvious visual enhancement of the studied classes is obtained. The infected regions by the powdery mildew clearly appear in a shiny yellow over a blue background which corresponds to the of the leaf lamina. The use of the IPPs set of observables seem to provide a better discrimination than the CPs set. The latter can be seen, because in the pseudo-colored image using the CPs set (Fig. 6d) there are some areas (highlighted with white rectangles) where some pixels are misclassified while the same pixels are correctly classified in the image pseudo-colored using the IPPs set. Analogously to the conclusions found for biological samples previously discussed, the pseudo-coloring method based on the Euclidean distance method shows a better performance than the method based on the Normal function-based approach. More in detail, for the particular case of the *Q. pubescens*, the set of observables that gives the most robust results in terms of discrimination efficiency is the set based on the IPPs because it shows larger delimitation of the infected regions and their borders.

Finally, we also studied a specimen of *V. vinifera* showing no pathologic symptomatology. As in the previous cases, we measured the experimental Mueller matrix and we retrieved the non-polarized transmission image (M_{00} in Fig. 7a) as a reference, and the polarimetric observables corresponding to the IPPs and CPs (Fig. 7b–g). In the non-polarized transmission image (Fig. 7a) we can observe two main leaf features, a vein (indicated by a blue arrow in Fig. 7a), and cell clusters (e.g., the green rectangle highlights a region of cell clusters in Fig. 7a). Importantly, polarimetric images in Fig. 8 allow the recognition of other plant structures not visible in non-polarimetric intensity images. This situation provides the importance of polarimetric channels for plant structures imaging, not only to increase visual image contrast, but also to reveal structures hidden in regular intensity images. In particular, contrarily to non-polarized intensity image (Fig. 7a), polarimetric images reveal the presence of a third structure consisting of a raphide. In vine leaves, raphides are made of calcium oxalate needle shaped crystals packed together[31,32] forming prorated clusters of typically 80 µm (long axis) × 30 µm (short axes). Raphides are completely invisible in non-polarized transmission images (Fig. 7a), but their presence and spatial location becomes clearly visible in polarization-based images, for instance, they are well visible in the P_1 image (Fig. 7b), indicated by pink arrows, and also, they are visible in the P, D and P_S images (Fig. 7e,f and g, respectively). In addition, the polarimetric analysis reveals another structure located at the upper-left part of the sample (indicated with white arrows in Fig. 7b and Fig. 7e and f, respectively) that has a polarimetric signature similar to that of the vein. This structure may correspond to a secondary vascular structure. From lower to higher, the mean values of the IPPs (Fig 7b–d) corresponding to the raphide are P_1=0.24±0.01, P_2=0.41±0.01 and P_3=0.47±0.01, followed by these corresponding to the cell clusters: P_1=0.47±0.01, P_2=0.50±0.01 and P_3=0.56±0.01, and finally the ones corresponding the leaf vein, P_1=0.68±0.01, P_2=0.72±0.01 and P_3=0.81±0.01. Accordingly it can be said that the raphide possess an individual signature, different to that of the vein and the cluster of cells and therefore it can be said that there are three different classes in the image. The reasons that may explain the elevated depolarization of the raphide are the fluctuations in the polarimetric properties, and scattering, which may be higher in the raphide that in the vein or in the cluster of cells. Concerning scattering, it is expected that the refractive index mismatch between a given region and the surrounding media, at the origin of scatting, should be higher for

the raphide, made of a solid inorganic component, than for veins or clusters of cells which are essentially made of a liquid similar to the surrounding media contained by the membranes forming the cell walls and other cell organelles. Importantly, note that P_2 and P_3 mean values in raphides and clusters of cells are quite similar to each other, but P_1 is significantly different (0.24 and 0.47 for raphides and cell cluster, respectively), which explains why the raphides are clearly seen in the P_1 image (Fig. 7b). The largest mean IPPs values correspond to the vein. This means that the vein structure induces little depolarization to the incident light, because it is a structure essentially filled with a liquid with low scattering and no polarimetric properties. Regarding the CPs observables, the highest polarizance (Fig. 7e) mean values are demonstrated for raphides (P=0.03±0.01), followed by the cell cluster (P=0.003±0.001) and the leaf vein (P=0.001±0.001). Likewise, the raphides show the highest mean diattenuation (D=0.07±0.01, Fig. 7f), to be compared to D=0.04±0.01 and D=0.02±0.01, corresponding to veins and clusters of cells respectively. Taking these values into account, the *V. vinifera* leaf shows low polarizance and diattenuation response in all its structures, and thus, it can be understood as a non-dichroic sample. Conversely, the highest mean values of the spherical purity P_S (Fig. 7g) are demonstrated for the leaf vein (P_S=0.70±0.01), this being a direct consequence of the strong alignment of the cellulose filaments within the vein. In turn, the P_S values for the cell cluster are reduced to P_S=0.49±0.01, and to P_S=0.31±0.01 in the case of the raphides. In analogy to the image of the P_1, these differences in the P_S mean values for the raphide, the cluster of cells and the vein, provide a well-contrasted image with well-differentiated regions (Fig. 7e).

The difference of the vein leaf sample with respect to the examples previously discussed is that it contains three classes instead of two of them. The latter may be a drawback to apply the pseudo-coloring based on the Euclidean distance, because it has been defined to handle only two classes. Therefore the pseudo-coloring based on Normal distribution may have an advantage in this particular situation. Accordingly and unlike in the examples previously discussed, we select three classes of structures to be simultaneously visualized (in the Normal-based approach): the raphides, a cell cluster and the leaf vein. The selected regions of interest, corresponding to the three classes, are indicated with pink, lime-green and blue squares within the purity index P_1 (Fig. 8a). In the case of the Euclidean distance method, as it is restricted to handle only two classes, the selected classes are the raphide (in pink) and the cell cluster (lime-green). Note that unlike in the previous samples, instead of using the non-polarized transmission image (M_{00}; Fig. 7a) to design the ROIs for the classes we use the P_1 image (Fig. 8a) since the raphides are not visible in the non-polarized transmission intensity channel. Implemented pseudo-colored images are provided in Figs. 8b-8e, for Euclidian distance (IPPs based in Fig. 8b and CPs based in Fig. 8d) and for the Normal-based models (IPPs based in Fig. 8c and CPs based Fig. 8e). Concerning the performance of the two methods in terms of class coloring and visual discrimination, we observe some differences. On the one hand, the IPPs are not sensitive enough to correctly identify the location of raphides when implementing the Euclidean approach (Fig. 8b). In particular, some pixels belonging to raphides are not well-colored in pink, but other pixels that do not belong to raphides, they are incorrectly painted in pink. Unlike this, the Euclidean method applied with CPs observables is much more efficient and quite accurate discriminating between raphides and cell cluster (Fig. 8d). On the other hand, the IPPs observables applied with the Normal-based approach are able to correctly identify and localize all the studied classes: raphides (pink regions in Fig. 8c), the leaf vein (blue region in Fig. 8c) and the cell cluster (lime-green pixels in Fig. 8c). Furthermore, other structure previously discussed in polarization images (Fig. 7), the vascular structure located on the upper-left part of the sample image, it is also colored in blue (indicated with a white arrow in Fig. 8c), as it is recognized as part of a vein (as previously told, vascular structure and vein presents very similar polarimetric response, and therefore they are recognized as part of the same class). Finally, when applying the CPs observables with the Normal-based approach (Fig. 8e), all the classes are correctly discriminated as well, but due to they are more affected by distances between pixel-values and mean classes-values, when applying the Gaussian probability function (Eq. (9) in "Methods" section), more pixels tends to zero probability of belonging to any class, and then, painted in black, this darkening the whole image.

Summarizing, due to the excellent recognition and visualization of the classes, as well as the capability of discriminating more than two classes simultaneously (three in this case), the best results for the *V. vinifera* sample in terms of visual structure discrimination are obtained when applying the Normal-based approach, implemented with the IPPs observables (Fig. 8c). Importantly, if one is just interested to discriminate between two classes, the Euclidean distance method based on CPs also provides excellent results (Fig. 8d).

As a final remark, we want to highlight that the present work provides the suitability of implementing two robust polarimetric image-processing methods, the Euclidean distances and the Normal (Gaussian) function (described in "Methods" section), for the visual enhancement of image contrast and the higher accuracy in spatial location of the different biological structures within the inspected samples. Both methods are based on the association of different colors with the particular tissue classes that should be highlighted, or discriminated, within the sample, leading to pseudo-colored images including information of a group of different polarimetric observables into a single image. The two presented approaches can be used to develop automatic coloring methods, and surpass proposals previously presented in the literature, which were based on very basic linear combinations of polarimetric observables whose weights were obtained with a non-optimal and mostly heuristic approach[1,24,25]. Going more specifically to the obtained results, when applying these two methods to discriminate among different structures present in organic samples (lamb trachea, lamb tongue, a *Q. pubescens* leaf and a *V. vinifera* leaf) they have revealed different strengths and drawbacks. The Euclidean method provided very good results to discriminate between two different classes in the studied samples (see for instance, Figs. 2b, 4d, 6b and 8d) but as it has been said, this method is restricted to the study of only two classes. In turn, the major strengths of the Normal-based approach are that it can provide discrimination between more than two classes (see for instance Fig. 8c) and that it provides a further interpretation, as it is based on the probability of a given pixel to belong to a given class. The Normal based approach can also be used as an automatic classifier. However, the pseudo-coloring Normal approach is very sensitive to the actual statistical distribution, the variance and the presence of outliers in the data. The Normal distribution approach works well when data is normally or

close to normally distributed without outliers. In real-life data, departures from the ideal normal distribution in data values results in either misclassification or black pixels, i.e. failure of identification to a given class (see for instance Fig. 2c and e).

In this work, we have provided optimized pseudo-colored functions based on polarimetric spaces, which enhance other approaches we previously presented in literature[1,24,25,30]. However, we want to note that the application of pseudo-colored functions is not limited to polarimetric spaces[33–36], and there exist other approaches, as image segmentation and coloring approaches[37–40], that lead to interesting results in terms of structures visualization. These approaches can be applied directly on polarimetric images, the resulting methods favoring from the inherent image enhancement of polarimetry images, but further comparison must be conducted in order to determine the best possible pseudo-colored imaging scenario. Regarding to the application of the IPPs or the CPs observables within the proposed methods, we have chosen these metrics because previous works highlight the suitability of depolarization observables for the discrimination of tissues[1,4,5,11,16,28], and because these two bases, all together, completely describe the depolarization response of a samples[27]. Therefore, IPPs and CPs represent an ideal framework to implement pseudo-colored functions for tissue discrimination. Regarding the samples discussed in this section, we have demonstrated that depending on the characteristics of the particular sample, either IPP or CP basis, provide to excellent visualization of the structures when used to apply pseudo-coloring strategies. Accordingly, for each sample to be analyzed, we recommend the use to the two bases of observables, i.e. the IPPs and CPs. Whereas IPPs are sensitive to the structure of the depolarization in samples (i.e., depolarization anisotropies), the CPs are more related to the physical properties of the constituents of the sample being at the roots of depolarization (retardance, polarizance, diattenuation).

At the end, the characteristics of each studied sample will determine which one of the two basis will provide the more vivid contrast.

Methods

Sample description. The animal samples used in this study were a section of a trachea and a section of a tongue dissected from an ex-vivo lamb sample bought from a grocery store. All experimental protocols were carried out in accordance with relevant guidelines and regulations. The two vegetal samples used in this work were (1) a leaf of *Q. pubescens* specimen infected with *E. alphitoides*, which causes powdery mildew lesions on leaf surface and (2) a leaf of *V. vinifera* specimen showing no symptoms of disease. The *Q. pubescens*, a species of white oak, belongs to the Fagaceae family and it is commonly found in central and southern Europe. It produces acorns (oak nuts) which can be consumed or extract their oil. *V. vinifera*, commonly known as grape vine, belongs to the Vitaceae family. Native from the central Europe, the land regions around the Mediterranean Sea and southwestern Asia, *V. vinifera* is cultivated worldwide for both grape (fresh or dried) consuming, and vinegar and wine production.

The *Quercus pubescens* leaf used in this study was kindly provided by Dra. Teresa Garnatje (Botanical Institute of Barcelona (IBB, CSIC-ICUB), Barcelona, Spain) and Dr. Jordi Luque (Institute of Agrifood Research and Technology (IRTA), Cabrils, 08348, Spain). The *Vitis vinifera* leaf was kindly provided by Dr. E. Garcia-Caurel. All the sample collecting methods were performed in accordance with relevant guidelines and regulation. T. Garnatje and J. Luque undertook the formal identification of the plant material used in this study. An herbarium voucher of *Q. pubescens* is deposited in the Herbarium of the Botanical Institute of Barcelona (BC-983018).

Polarization observables. Among the wide variety of mathematical approaches[21], the Mueller-Stokes (M-S) formalism is especially suitable for the description of the polarimetric properties of turbid media, as it is based on radiometric measurements and allows to deal with the depolarization content of samples[22,23]. In this approach, the state of polarization of light beams is characterized by the so-called Stokes vector (S) and the polarimetric characteristics of the sample are complexly encoded into the 4×4 real matrix called Mueller matrix (MM). The generic MM block form is defined as:

$$M = m_{00}\begin{pmatrix} 1 & D^T \\ P & m \end{pmatrix}. \tag{1}$$

From the structure of the MM (Eq. (1)) we can easily retrieve the non-polarized transmission or reflection (m_{00}) and the dichroism (diattenuation and polarizance, the 3-dimensional vectors D and P, respectively). However, the polarimetric properties related to the retardance and the depolarization are entangled in a 3×3 submatrix, m. Concerning to the polarimetric observables, whereas the diattenuation gives a measure of the transmission /reflection dependence of the sample with the input polarization state, polarizance describes the capability of the sample to polarize a fully unpolarized input light beam. In addition, the degree of spherical purity, P_S, defines the portion of depolarization which is not directly related with dichroic properties of the sample.

$$D = \frac{\sqrt{m_{01}^2 + m_{02}^2 + m_{03}^2}}{m_{00}}, \quad P = \frac{\sqrt{m_{10}^2 + m_{20}^2 + m_{30}^2}}{m_{00}}, \quad P_S = \frac{\|m\|_2}{\sqrt{3}}, \tag{2}$$

where $\|m\|_2$ is the 2-norm of the sub-matrix m. To retrieve the depolarization content, we conduct the Cloude's decomposition[21] which defines the MM as a parallel combination (i.e., convex sum) of four non-depolarizing (pure) MMs, labeled as M_{Ji}, whose statistical weights are proportional to the covariance matrix $H(M)$ eigenvalues (λ_i)[22]:

$$M = m_{00} \sum_{i=0}^{3} \hat{\lambda}_i \hat{M}_{Ji}, \quad \lambda_0 \geq \lambda_1 \geq \lambda_2 \geq \lambda_3 \geq 0 \tag{3}$$

Interestingly, the depolarizing response being synthetized within the four pure components M_{Ji} allows to retrieve the different types of depolarizers by simply looking at the weights corresponding to each pure component. The combination of the above-mentioned normalized eigenvalues results into the definition of the Indices of Polarimetric Purity (IPPs)[26]: three real, dimensionless and invariant parameters which provide information about the polarimetric randomness induced by the sample to input polarization states. The IPPs are defined as:

$$P_1 \equiv \hat{\lambda}_0 - \hat{\lambda}_1, \quad P_2 \equiv \hat{\lambda}_0 + \hat{\lambda}_1 - 2\hat{\lambda}_2, \quad P_3 \equiv \hat{\lambda}_0 + \hat{\lambda}_1 + \hat{\lambda}_2 - 3\hat{\lambda}_3, \tag{4}$$

and are restricted to $0 \leq P_1 \leq P_2 \leq P_3 \leq 1$. The IPPs define a real 3D-depolarization space whose interpretation is related to the polarimetric randomness (i.e., depolarization) induced by different mechanisms within the sample. Accordingly, the IPPs' depolarization space constitutes a suitable tool to discriminate among structures with different depolarization signatures due to their inherent components. Recalling that the depolarizing response of a sample can be synthetized as the incoherent sum of four pure components (M_{Ji} in Eq. (3)), the IPPs correspond to the statistical weight corresponding to each of these pure components[22,26]. In particular, P_1 is associated with the portion of the pure non-depolarizing component, P_2-P_1 quantifies the statistical portion of a bidimensional depolarizer, P_3-P_2 corresponds to the portion of a tridimensional depolarizer (equiprobable mixture of three pure components) and 1-P_3 quantifies the statistical weight of an ideal depolarizer. Furthermore, the depolarization index, P_Δ, estimates the overall depolarization of the MM. Importantly, this observable allows to connect the polarimetric spaces of (1) the Components of Purity (CPs; composed by the enpolarization metrics corresponding to the diattenuation D, polarizance P and the degree of spherical purity P_S) and (2) the Indices of Polarimetric Purity (IPPs; P_1, P_2 and P_3) in the following way:

$$P_\Delta = \frac{1}{\sqrt{3}}\sqrt{D^2 + P^2 + 3P_S^2} = \frac{1}{\sqrt{3}}\sqrt{2P_1^2 + \frac{2}{3}P_2^2 + \frac{1}{3}P_3^2}, \qquad 0 \leq P_\Delta \leq 1 \tag{5}$$

Importantly, when $P_1 = P_2 = P_3 = P_\Delta = 1$ is found, the indicators characterize a non-depolarizing (pure) system. Conversely, the ideal depolarizer is defined by $P_1 = P_2 = P_3 = P_\Delta = 0$.

Pseudo-coloring approaches. In the following, we define the parameters of interest involved in the pseudo-coloring models, both for the Euclidean and Normal cases. Assume, from the experimental Mueller matrix (MM) measurement of a given sample, the extraction of n MM-derived polarimetric observables, $\overrightarrow{p} = [p_1, ..., p_n]$[21]. Note that the polarimetric observables, $\overrightarrow{p}$, we use in this study are those defined in the previous section, i.e., the Components of Purity, CPs, and the Indices of Polarimetric Purity, IPPs. Additionally, consider the definition of i classes corresponding to different organic tissues. Each kind ($i = 1, \ldots, k$) of organic tissues (trachea ring, trachea sheath, tongue papillae tissue, tongue epithelial tissue, *Q. pubescens* powdery mildew, *Q. pubescens* lamina, *V. vinifera* vein, *V. vinifera* raphides and *V. vinifera* cell cluster; in our study) is characterized by the $j = 1, \ldots, n$ means, m_j^i, and standard deviations, σ_j^i, corresponding to the p_j^i observables calculated from a Region of Interest (ROIs) within the specific class i. In addition, for image coloring purposes, we also define the vector $\vec{C}^i = \left[R^i, G^i, B^i\right]$ with $i = 1, \ldots, k$ as the standard RGB color space triplet associated with a particular class i.

Euclidean distances for $k=2$ classes. In this case we assume we only need to discriminate between $k = 2$ classes (e.g. healthy/infected tissue). This pseudo-colored approach is based on the Euclidean distance from the values of $\overrightarrow{p}$ from a given image pixel to the mean values $\overrightarrow{m^i} = [m_1^i, ..., m_n^i]$ of the n polarimetric observables of a given class ($i = 1$ or $i = 2$). The normalized distance $d^{i,norm}$ is given by,

$$d^{i,norm} = \sqrt{\sum_{j=1}^{n}\left(\frac{d^i}{m_j^{i=1} - m_j^{i=2}}\right)^2} = \sqrt{\sum_{j=1}^{n}\left(\frac{m_j^i - p_j}{m_j^{i=1} - m_j^{i=2}}\right)^2}, \quad i = 1, 2, \tag{6}$$

where $m_j^{i=1} - m_j^{i=2}$ corresponds to the distance between the means of the two classes for a given polarimetric observable, p_j ($j = 1, \ldots, n$). Importantly, we associate a given color, $\overrightarrow{C_i}$, to each particular class, i. Consequently, the larger the distance from a given pixel to a given class, the lower the corresponding weight to its particular class color, $\overrightarrow{C_i}$. Accordingly, each particular class color $\overrightarrow{C_i}$ is pondered by the following subtraction,

$$\Re^i = 1 - \frac{d^{i,norm}}{d}, \quad i = 1, 2, \tag{7}$$

where $d^{i,norm}$ is normalized by the sum of the distances (i.e., $d = d^{i=1,norm} + d^{i=2,norm}$), so that the subtraction is positive-definite and ranges between 0 and 1. Thus, the larger the distance, $d^{i,norm}$, the lower the amount of i-color level included in the pixel. The final pixel color for the two classes ($i = 1, 2$), is given by,

$$\vec{C} = [R, G, B] = \Re^1\left[R^1, G^1, B^1\right] + \Re^2\left[R^2, G^2, B^2\right]. \tag{8}$$

Importantly, the method does not behave as a classifier. Note that some pixels within the image may contain more than one color, thus showing a mixed tone between the two basis (pure) colors. Moreover, there are some pixels that correspond to some parts of the sample which do not belong to any inspected class, either $i=1$ nor $i=2$. In such case, we can equally interpret both distances (i.e., $d^{i=1,norm} = d^{i=2,norm}$), and thus, the $d^{i,norm}/d$ term in Eq. (7) tends to 1/2 so that $R^i \approx 1/2$. Under this approach, as both R^i tend to be 1/2, the resulting color for these pixels is an equal mixture of the two selected colors. Therefore, the resulting color encoding cannot be used for classificatory proposes, but just for visual enhancement of the images as shown in "Results" and "Discussion" sections. Pseudo-coloring based on the Euclidean-distances method is very efficient in terms of tissues visualization and discrimination.

Normal distribution for *k* classes. The previous method is difficult to generalize to more than two classes ($k>2$). For this reason, in this subsection we propose a second method to construct pseudo-colored functions, in this case, based on the Normal (Gaussian) probability distribution of the n polarimetric observables. Contrary to the above-presented Euclidean method, the Normal approach allows to inspect an unlimited number of features (i.e., classes), k. The probability P^i of a pixel corresponding to the jth observable ($j=1,..,n$) of being part of a given class i ($i=1,..,k$) is defined as,

$$P^i = \prod_j \exp\left[-\left(\frac{p_j - m_{i,j}}{\sigma_{i,j}}\right)^2\right], \tag{9}$$

which is limited to range between 0 and 1. Therefore, for each pixel we get as many probability functions, P^i, as classes, i, we deal with. Afterwards, the pseudo-colored image is constructed by associating each standard RGB color space triplet to their corresponding probability function (P^i) in the following way:

$$\vec{C} = [R, G, B] = \sum_i P^i \cdot C^i = \sum_i P^i \left[R^i, G^i, B^i\right]. \tag{10}$$

This approach outputs a colored polarimetric image, $\vec{C} = [R, G, B]$, which is based on the linear combination of the $P^i \cdot C^i$ terms of the i classes involved. In other words, the amount of i-color level within a pixel is pondered by the probability of the particular pixel to be recognized as belonging to the class i. Therefore, unlike the Euclidean distance approach, the pixel color coding associated to this Normal-based model can be used for classificatory proposes.

Complete image Mueller polarimeter. The experimental Mueller matrices images of tissues are acquired by means of a complete image Mueller polarimeter consisting of two independent and mobile arms, the Polarization State Generator (PSG) and the Polarization State Analyzer (PSA), which are capable to generate and analyze, respectively, any fully polarized state. The optical systems comprising both PSG and PSA are based on Parallel Aligned Liquid Crystals (PA-LC, Variable Retarders with Temperature Control) retarders (LVR-200-400-700-1LTSC distributed by Meadowlark Optics). In particular, the PSG is composed by a linear polarizer (Glan-Thompson prism-based CASIX) oriented at 0°, followed by two PA-LC oriented at 45° and 0°. Likewise, the PSA optical set-up is composed by the same optical elements as the PSG but arranged in reverse order. The polarizer within the PSA corresponds to a dichroic sheet polarizer distributed by Meadowlark Optics. All orientations are with respect to the laboratory vertical. The light source device is a four-wavelength high-power Thorlabs LED source (LED4D211, operated by DC4104 drivers distributed by Thorlabs). It is placed on the PSG and allows to illuminate with different wavelengths covering the visible spectrum (from 400 to 700 nm, approx.). Complementary, we use 10 nm dielectric bandwidth filters distributed by Thorlabs: FB530-10 and FB470-10 for green and blue wavelengths, respectively. The image acquisition (i.e., sample intensity) is conducted by means of the set comprising a 35 mm focal length Edmund Optics TECHSPEC® high resolution objective followed by an Allied Vision Manta G-504B CCD camera, with 5 Megapixel GigE Vision and Sony ICX655 CCD sensor (of 2452(H) × 2056(V) resolution and cell size of 3.45 × 3.45 μm). Both optical components are placed on the PSA system, achieving a spatial resolution of 22 μm.

The Mueller matrix measurements of the lamb trachea, the lamb tongue and the *Q. pubescens* are conducted at 470 nm illumination wavelength by tilting by 34° the PSG with respect to the laboratory horizontal reference and holding the PSA at 0°. This configuration allows us to avoid the ballistic reflection. From the whole sample, we select a region of interest (ROI) of 512 × 512 pixels, thus corresponding to an area of 1.1 × 1.1 cm^2. With regards to the experimental Mueller matrix of the *V. vinifera* leaf, it is acquired by means of a multimodal microscope polarimeter[3] coupled to a light source of white light LED with a narrow-band spectral filter of 533 nm and a spectral width of 15 nm. The microscope is set at transmission configuration, where the sample leaf is placed between two identical microscope objectives (for imaging and illumination, respectively). The achieved magnifications are of 50X, 20X, or 5X.

Data availability

The datasets generated during and/or analysed during the current study are not publicly available due to the conduction of different research studies but are available from the corresponding author on reasonable request.

Received: 16 June 2022; Accepted: 29 October 2022
Published online: 02 November 2022

References

1. Van Eeckhout, A. *et al.* Polarimetric imaging of biological tissues based on the indices of polarimetric purity. *J. Biophotonics* **11**, e201700189. https://doi.org/10.1002/jbio.201700189 (2017).
2. Tóth, T. N. *et al.* Fingerprinting the macro-organisation of pigment-protein complexes in plant thylakoid membranes in vivo by circular-dichroism spectroscopy. *Biochim. et Biophys. Acta* **1857**(9), 1479–1489. https://doi.org/10.1016/j.bbabio.2016.04.287 (2016).
3. Van Eeckhout, A. *et al.* Polarimetric imaging microscopy for advanced inspection of vegetal tissues. *Sci. Rep.* **11**(1), 3913. https://doi.org/10.1038/s41598-021-83421-8 (2021).
4. Van Eeckhout, A. *et al.* Depolarizing metrics for plant samples imaging. *PLoS ONE* **14**(3), e0213909. https://doi.org/10.1371/journal.pone.0213909 (2019).
5. Van Eeckhout, A. *et al.* Synthesis and characterization of depolarizing samples based on the indices of polarimetric purity. *Opt. Lett.* **42**(20), 4155–4158. https://doi.org/10.1364/OL.42.004155 (2017).
6. Bischoff, V. *et al.* Trichome birefringence and its homolog AT5G01360 encode plant-specific DUF231 proteins required for cellulose biosynthesis in arabidopsis. *Plant Phys.* **153**(2), 590–602. https://doi.org/10.1104/pp.110.153320 (2010).
7. Gao, C. J., Xia, X. J., Shi, K., Zhou, Y. H. & Yu, J. Q. Response of stomata to global climate changes and the underlying regulation mechanism of stress responses. *Plant Physiol. J.* **48**(1), 19–28 (2012).
8. Shtein, I. S. *et al.* Stomatal cell wall composition: Distinctive structural patterns associated with different phylogenetic groups. *Ann. Bot.* **119**(6), 1021–1033. https://doi.org/10.1093/aob/mcw275 (2017).
9. Jain, A. *et al.* Backscattering polarimetric imaging of the human brain to determine the orientation and degree of alignment of nerve fiber bundles. *Biomed. Opt. Express* **12**, 4452–4466. https://doi.org/10.1364/BOE.426491 (2021).
10. Borovkova, M. *et al.* Evaluating β-amyloidosis progression in Alzheimer's disease with Mueller polarimetry. *Biomed. Opt. Express* **11**, 4509–4519. https://doi.org/10.1364/BOE.396294 (2020).
11. Le, D. L., Nguyen, D. T., Le, T. H., Phan, Q. H. & Pham, T. T. H. Characterization of healthy and cancerous human skin tissue utilizing Stokes-Mueller polarimetry technique. *Opt. Commun.* **480**, 126460. https://doi.org/10.1016/j.optcom.2020.126460 (2021).
12. Ivanov, D. *et al.* Colon cancer detection by using Poincaré sphere and 2D polarimetric mapping of ex vivo colon samples. *J. Biophotonics* **13**(8), e202000082. https://doi.org/10.1002/jbio.202000082 (2020).
13. Patel, A., Khan, A., Quinlan, R. & Yaroslavsky, A. N. Polarization-sensitive multimodal imaging for detecting breast cancer. *Cancer Res.* **74**(17), 4685–4693. https://doi.org/10.1158/0008-5472.CAN-13-2411 (2014).
14. Yaroslavsky, A. N., Feng, X., Muzikansky, A. & Hamblin, M. R. Fluorescence polarization of methylene blue as a quantitative marker of breast cancer at the cellular level. *Sci. Rep.* **9**(1), 940. https://doi.org/10.1038/s41598-018-38265-0 (2019).
15. Schucht, P. *et al.* Visualization of white matter fiber tracts of brain tissue sections with wide-field imaging Mueller polarimetry. *IEEE Trans. Med. Imaging* **39**(12), 4376–4382. https://doi.org/10.1109/TMI.2020.3018439 (2020).
16. Rodríguez, C. *et al.* Polarimetric data-based model for tissue recognition. *Biomed. Opt. Express* **12**, 4852–4872. https://doi.org/10.1364/BOE.426387 (2021).
17. Ivanov, D. et al. Polarization-Based Histopathology Classification of Ex Vivo Colon Samples Supported by Machine Learning. *Front. in Phys.* **9**; https://doi.org/10.3389/fphy.2021.814787 (2022).
18. Li, P. *et al.* Analysis of tissue microstructure with Mueller microscopy: logarithmic decomposition and Monte Carlo modeling. *J. Biomed. Opt.* **25**(1), 1–11. https://doi.org/10.1117/1.JBO.25.1.015002 (2020).
19. Luu, N., Le, T. H., Phan, Q. H. & Pham, T. T. H. Characterization of Mueller matrix elements for classifying human skin cancer utilizing random forest algorithm. *J. Biomed. Opt.* **26**, 075001. https://doi.org/10.1117/1.JBO.26.7.075001 (2021).
20. Ahmad, I., Ahmad, M., Khan, K. & Ikram, M. Polarimetry based partial least square classification of ex vivo healthy and basal cell carcinoma human skin tissues. *Photodiag. Photody. Ther.* **14**, 134–141. https://doi.org/10.1016/j.pdpdt.2016.04.004 (2016).
21. Gil, J. J. & Ossikovski, R. *Polarized light and the Mueller matrix approach* (CRC Press, 2016).
22. Goldstein, D. *Polarized light* 2nd edn. (Marcel Dekker, 2003).
23. Chipman, R. A. *Polarimetry: Handbook of optics* 2nd edn. (McGraw Hill, 1995).
24. Rodríguez, C. et al. Indices of polarimetric purity: Application in biological tissues. *Proc. SPIE*, 116460P. https://doi.org/10.1117/12.2578254 (2021).
25. Van Eeckhout, A. et al. Indices of polarimetric purity for biological tissues inspection. *Proc. SPIE*, 104971V. https://doi.org/10.1117/12.2290283 (2018).
26. San José, I. & Gil, J. J. Invariant indices of polarimetric purity: Generalized indices of purity for NXN covariance matrices. *Opt. Commun.* **284**(1), 38–47. https://doi.org/10.1016/j.optcom.2010.08077 (2011).
27. Van Eeckhout, A. *et al.* Unravelling physical information of depolarizers. *Opt. Exp.* **29**(23), 38811–38823. https://doi.org/10.1364/OE.438673 (2021).
28. Wan, J. *et al.* Polarization-based probabilistic discriminative model for quantitative characterization of cancer cells. *Biomed. Opt. Express* **13**, 3339–3354. https://doi.org/10.1364/BOE.456649 (2022).
29. Ivanov, D. *et al.* Polarization and depolarization metrics as optical markers in support to histopathology of ex vivo colon tissue. *Biomed. Opt. Express.* **12**(7), 4560–4572. https://doi.org/10.1364/BOE.426713 (2021).
30. Rodríguez, C. *et al.* Polarimetric observables for the enhanced visualization of plant diseases. *Sci. Rep.* **12**, 14743. https://doi.org/10.1038/s41598-022-19088-6 (2022).
31. Fabbri, A., Benelli, C. & Di Collalto, G. Calcium oxalate crystals in vegetative and reproductive organs of the grapevine. *Acta Hortic.* **292**, 107–12 (1992).
32. Jáuregui-Zúñiga, D., Reyes-Grajeda, J. P., Sepúlveda-Sánchez, J. D., Whitaker, J. R. & Moreno, A. Crystallochemical characterization of calcium oxalate crystals isolated from seed coats of *Phaseolus vulgaris* and leaves of *Vitis vinifera*. *J. Plant. Physiol.* **160**, 239–245. https://doi.org/10.1078/0176-1617-00947 (2003).
33. Aïnouz, S., Zallat, J., de Martino, A. & Collet, C. Physical interpretation of polarization-encoded images by color preview. *Opt. Express* **14**(13), 5916–5927. https://doi.org/10.1364/OE.14.005916 (2006).
34. Fu, J. & Caulfield, H. J. Applying color discrimination to polarization discrimination in images. *Opt. Comm.* **272**(2), 362–366. https://doi.org/10.1016/j.optcom.2006.11.058 (2007).
35. Li, N., Zhao, Y., Pan, Q., Kong, S. G. & Chan, J. C. W. Illumination-invariant road detection and tracking using LWIR polarization characteristics. *ISPRS J. Photo. Rem. Sens.* **180**, 357–369. https://doi.org/10.1016/j.isprsjprs.2021.08.022 (2021).
36. Li, N., Zhao, Y., Wu, R. & Pan, Q. Polarization-guided road detection network for LWIR division-of-focal-plane camera. *Opt. Lett.* **46**(22), 5679–5682. https://doi.org/10.1364/OL.441817 (2021).
37. Jones, G. R. & Spencer, J. W. *Advanced chromatic monitoring* (CRC Press, 2020).
38. Muñoz-Aguirre, M., Ntasis, V. F., Rojas, S. & Guigó, R. PyHIST: a histological image segmentation tool. *PLoS Comput. Biol.* **16**(10), e1008349. https://doi.org/10.1371/journal.pcbi.1008349 (2020).
39. Dang, Vu. *et al.* Methods for segmentation and classification of digital microscopy tissue images. *Front. Bioeng. Biotech.* **7**, 53. https://doi.org/10.3389/fbioe.2019.00053 (2019).
40. Kang, H. K. *et al.* A simple segmentation and quantification method for numerical quantitative analysis of cells and tissues. *Technol Health Care* **28**(S1), 401–410. https://doi.org/10.3233/THC-209041 (2020).

Acknowledgements
We acknowledge the financial support of Spanish MINECO (PID2021-126509OB-C21, and Fondos FEDER); Catalan Government (2017-SGR-001500).

Author contributions
C.R., A.L. and J.C. were involved in the study design. C.R., A.V.E. collected samples. C.R. and A.V.E. conducted measurements. C.R., A.L. and J.C. developed software for data analysis. C.R. analyzed data. C.R., A.L., A.V.E., E.G.C., and J.C. drafted the manuscript. C. R. prepared Fig. 1, 2, 3, 4, 5, 6, 7, 8. A.L. and J.C. supervised the project. All authors reviewed the manuscript.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-022-23330-6.

Correspondence and requests for materials should be addressed to C.R.

Reprints and permissions information is available at www.nature.com/reprints.

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

Automatic pseudo-coloring approaches to improve visual perception and contrast in polarimetric images of biological tissues

Carla Rodríguez[1,*], Albert Van Eeckhout[1,2], Enrique Garcia-Caurel[2], Angel Lizana[1] and Juan Campos[1]

[1]Optics group, Physics Department, Universitat Autònoma de Barcelona, Bellaterra, 08193, Spain
[2]LPICM, CNRS, Ecole Polytechnique, Institut Polytechnique de Paris, Palaiseau, 91120, France
[*]carla.rodriguez@uab.cat

1. Polarimetric observables for biological samples inspection

This section is devoted to show the suitability of using polarimetric observables for biological tissue inspection. In the following we show the polarimetric images corresponding to the non-polarized intensity (M_{00}), the indices of polarimetric purity (IPPs: P_1, P_2 and P_3), the depolarization index P_Δ, the components of purity (CPs: P, D and Ps) and the global retardance R, from a set of four different biological samples. In particular, we inspect the lamb trachea section and the *Vitis vinifera* plant sample, already shown in the manuscript. In addition, we provide the polarimetric images taken from a leaf of *Prunus dulcis* plant specimen showing disease symptoms from fungal infection and a section of a lamb heart.

Figure S1 shows the polarimetric images from the lamb trachea section. The cartilaginous rings and the trachea sheath is clearly distinguished by the enpolarization metrics of P_2, P_3, P, and P_S (Figs. S1c, S1d, S1f and S1h), as they point well the difference between the two types of tissues and also present an enhanced contrast when compared to the non-polarized diffuse reflectance (M_{00} in Fig. S1a).

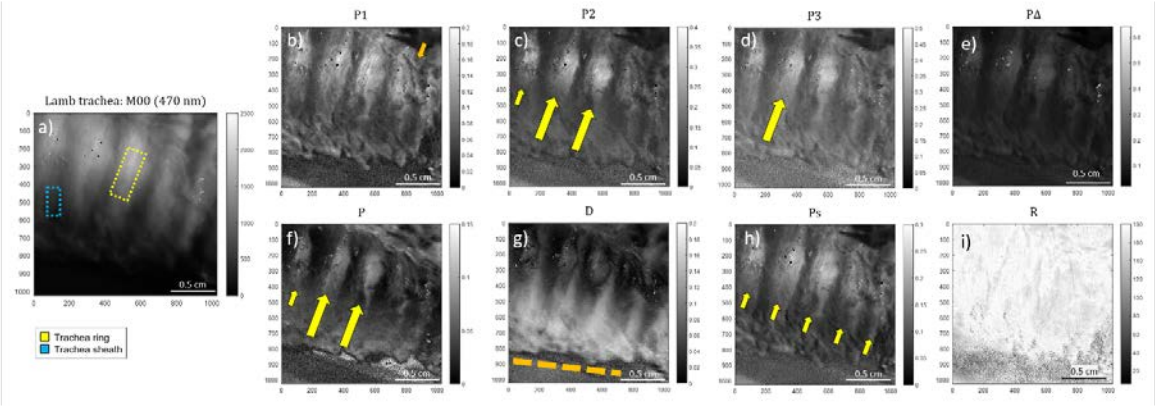


Figure S1. Polarimetric images of a lamb trachea measured at 470 nm illumination wavelength: a) unpolarized intensity image (M_{00}), the Indices of Polarimetric Purity b) P_1, c) P_2 and d) P_3, the depolarization index e) P_Δ, the Components of Purity f) P, g) D and h) P_S and i) the global retardance R. The yellow dotted box (Fig. S1a) and yellow arrows (Figs. S1c, S1d, S1f and S1h), indicate the location of the cartilaginous rings. The dotted blue box (Fig. S1a) indicates the location of the trachea sheath. The orange arrow (Fig. S1b) and the orange dotted line (Fig. S1g) show the vascular structure within the external trachea sheath and the sample border, respectively.

To highlight the behavior of the polarimetric observables, we computed the visibility, $V = \left[I_{\max} - I_{\min}\right] / \left[I_{\max} + I_{\min}\right]$ (see Table S1) of the sample regions corresponding to the trachea rings and the trachea sheath for each of the metrics here discussed. Additionally, we present the subtraction of the visibility value between both structures. The largest visibility difference between

the lamb trachea rings and sheath is demonstrated by the diattenuation D (0.4753 in Table S1), followed by the index of polarimetric purtiy P_3 and the depolarization index P_Δ (0.2621 and 0.2363 in Table S1, respectively).

Table S1. Visibility of different polarimetric observables corresponding to the trachea ring and trachea sheath in lamb trachea sample. The largest differences in visibility values between the inspected structures is highlighted in gray.

		M_{00}	P_Δ	P_1	P_2	P_3	P	R	D	P_S
Lamb trachea	Ring	0.2596	0.4639	0.4258	0.5644	0.5130	0.6974	0.1030	0.9434	0.4542
	Sheath	0.3145	0.2276	0.4599	0.3859	0.2509	0.5782	0.1348	0.4681	0.4168
Subtraction: Ring *vs* Sheath		0.0549	0.2363	0.0341	0.1785	0.2621	0.1192	0.0318	0.4753	0.0374

Similar analysis is conducted for the *V. vinifera* plant sample. Figure S2 shows the polarimetric images from *V. vinifera* specimen. Importantly, raphides are completely invisible in non-polarized transmission images (i.e., M_{00} in Fig. S2a), but their presence and spatial location becomes clearly visible in polarization-based images, for instance, they are clearly visible in the index of polarimetric purtiy P_1 image (Fig. S2b), indicated by pink arrows, and also, they are visible in the P_Δ, P, D and P_S images (Figs. S2e-S2h, respectively).

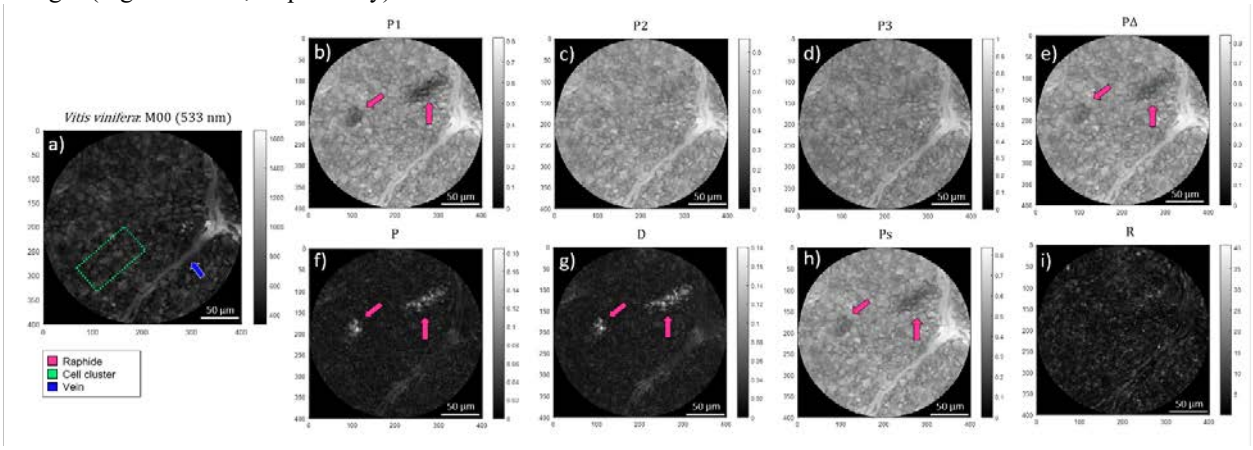


Figure S2. Polarimetric images of the *Vitis vinifera* leaf measured under the microscope for 533 nm illumination wavelength: a) non-polarized transmission intensity image (M_{00}), the Indices of Polarimetric Purity b) P_1, c) P_2 and d) P_3, the depolarization index e) P_Δ, the Components of Purity f) P, g) D and h) P_S and i) the global retardance R. The blue arrow (Figs. S2a) indicates the location of the vein. The pink arrows (Figs. S2b, S2e-S2h) indicate the location of the raphides. The lime-green dotted box indicates an illustrative region comprising a cell cluster.

Likewise, the visibility values corresponding to the raphides and the cell cluster are presented in Table S2 as well as the subtraction between their values. The largest visibility difference between the raphide and the cell cluster is achieved by the index of polarimetric purity P_1 (the 0.3154 In Table S2), followed by polarizance, P, and diattenuation, D (0.1269 and 0.1234 in Table S2, respectively).

Table S2. Visibility of different polarimetric observables corresponding to the raphide and the cell cluster in *Vitis vinifera* plant sample. The largest difference in visibility values between the inspected structures is highlighted in gray.

		M_{00}	P_Δ	P_1	P_2	P_3	P	R	D	P_S
V. vinifera	Raphide	0.2319	0.2829	0.5410	0.2323	0.2305	0.9059	0.9698	0.8727	0.2963
	Cell cluster	0.2312	0.2192	0.2256	0.2029	0.2484	0.7790	0.9763	0.7493	0.2195
Subraction: Raphide vs Cell cluster		0.0007	0.0637	0.3154	0.0294	0.0179	0.1269	0.0065	0.1234	0.0768

Additionally, in Fig. S3 we show the polarimetric images corresponding to a 2.2 x 2.2 cm^2 section of a lamb heart. In particular,

the different polarimetric properties of the sample allow the enpolarizing metrics to reveal different structures. That is the case of small vascular structures within the lamb heart revealed by the Index of Polarimetric Purity, P_1, and the polarizance, P (yellow arrows in Figs. S3b and S3f, respectively). Moreover, the roughness pattern of the tissue is enhanced by Index of Polarimetric Purity, P_2 and the degree of spherical purity P_S (green arrows in Figs. S3c and S3h, respectively). Retardance R (Fig. S3i) also reveals some vascular structures and the striated tissue region (yellow and green arrow in Fig. S3i, respectively). However, this polarimetric observable does not enhance the overall image contrast as much as the depolarizing metrics of the Indices of Polarimetric Purity (P_1, P_2 and P_3) and the degree of spherical purity P_S do (Figs. S3b-S3d and S3h, respectively).

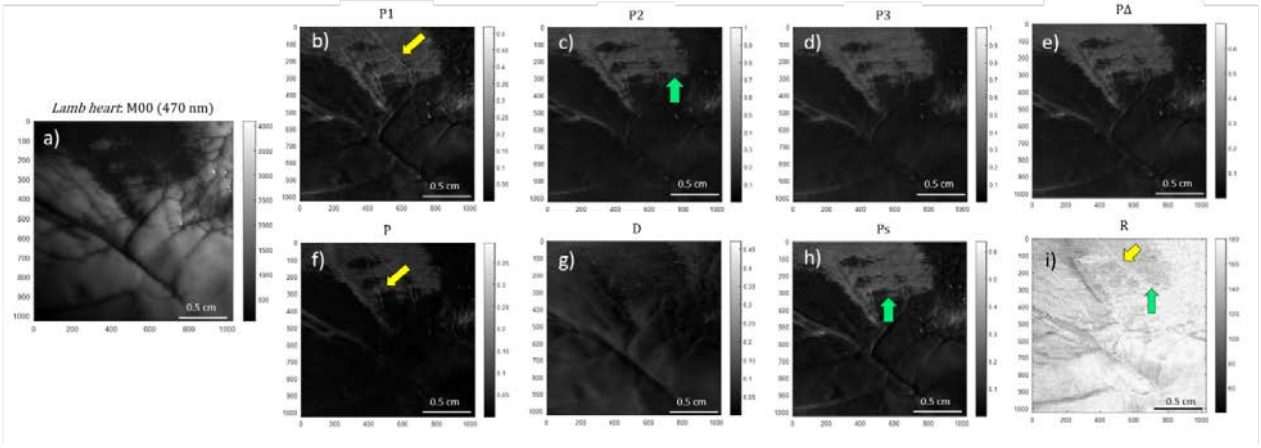


Figure S3. Polarimetric images of the lamb heart measured at 470 nm illumination wavelength: a) non-polarized transmission intensity image (M_{00}), the Indices of Polarimetric Purity b) P_1, c) P_2 and d) P_3, the depolarization index e) P_Δ, the Components of Purity f) P, g) D and h) P_S and i) the global retardance R. The yellow arrows (Fig. S3b, S3f and S3i) indicate the location of a particular vascular structure. The green arrows (Fig. S3c, S3h and S3i) indicate the region showing different roughness.

Finally, in Fig. S4 we show the polarimetric images corresponding to a leaf of *Prunus dulcis* plant specimen showing disease symptoms from fungal infection. The disease is mostly located within the dark spot located at the down-left part of the intensity image M_{00} (indicated with the pink arrow in Fig. S4a). The depolarization-related observables (i.e., P1, P2 and P3 and P_Δ in Figs. S4b-S4e, respectively) as well as the Components of Purity (i.e., P, D and Ps in Figs. S4f-S4h) reveal the small vascular structures contained inside the spot (pink arrow in Fig. S4b), which are unable to be detected by means of regular intensity image (Fig. S4a). On the contrary, retardance R focuses on the principal vein (Fig. S4i).

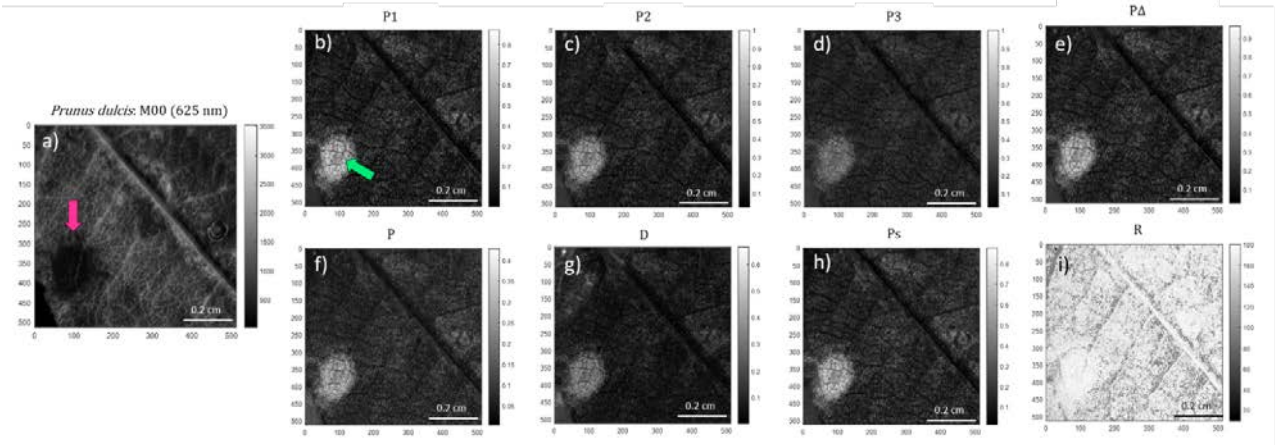


Figure S4. Polarimetric images of the *Prunus dulcis* leaf measured at 625 nm illumination wavelength: a) non-polarized transmission intensity image (M_{00}), the Indices of Polarimetric Purity b) P_1, c) P_2 and d) P_3, the depolarization index e) P_Δ, the Components of Purity f) P, g) D and h) P_S and i) the global retardance R. The pink and green arrows (Figs. S4a and S4b, respectively) indicate the location of the fungal disease and the small vascular structures contained.

PAPER C

Polarimetric data-based model for tissue recognition

Carla Rodríguez[1,*], Albert Van Eeckhout[1], Laia Ferrer [1], Enrique Garcia-Caurel[2], Emilio González-Arnay[3,4], Juan Campos[1] and Angel Lizana[1]

[1] *Grup d'Òptica, Physics Department, Universitat Autònoma de Barcelona Bellaterra 08193, Spain*
[2] *LPICM, CNRS, Ecole Polytechnique, Institut Polytechnique de Paris, Palaiseau 91120, France*
[3] *Depto. de Anatomía, Histología y Neurociencia, Universidad Autónoma de Madrid, Madrid 28049, Spain*
[4] *Servicio de Anatomía Patológica, Hospital Universitario de Canarias, Santa Cruz de Tenerife 38320, Spain*

This paper has been published in *Biomedical Optics Express*, **12**(8), pp. 4852–4872, 2021.

Research Article Vol. 12, No. 8 / 1 August 2021 / Biomedical Optics Express 4852
Biomedical Optics EXPRESS

Polarimetric data-based model for tissue recognition

CARLA RODRÍGUEZ,[1,5] ALBERT VAN EECKHOUT,[1] LAIA FERRER,[1] ENRIQUE GARCIA-CAUREL,[2] EMILIO GONZÁLEZ-ARNAY,[3,4] JUAN CAMPOS,[1] AND ANGEL LIZANA[1,6]

[1]*Grup d'Òptica, Physics Department, Universitat Autònoma de Barcelona, Bellaterra 08193, Spain*
[2]*LPICM, CNRS, Ecole Polytechnique, Institut Polytechnique de Paris, Palaiseau 91120, France*
[3]*Departamento de Anatomía, Histología y Neurociencia, Universidad Autónoma de Madrid, Madrid 28049, Spain*
[4]*Servicio de Anatomía Patológica, Hospital Universitario de Canarias, Santa Cruz de Tenerife 38320, Spain*
[5]*carla.rodriguez@uab.cat*
[6]*angel.lizana@uab.cat*

Abstract: We highlight the potential of a predictive optical model method for tissue recognition, based on the statistical analysis of different polarimetric indicators that retrieve complete polarimetric information (selective absorption, retardance and depolarization) of samples. The study is conducted on the experimental Mueller matrices of four biological tissues (bone, tendon, muscle and myotendinous junction) measured from a collection of 157 ex-vivo chicken samples. Moreover, we perform several non-parametric data distribution analyses to build a logistic regression-based algorithm capable to recognize, in a single and dynamic measurement, whether a sample corresponds (or not) to one of the four different tissue categories.

1. Introduction

Polarimetry comprises a set of non-invasive and non-contact testing techniques that provide information about the optical properties of a sample [1]. Those techniques appear as very useful tools to be applied in different research fields, such as material classification [2], gas sensing [3], astronomy [4], biomedicine [5–8], and remote sensing [9] but among all, for the purpose of this work, they are an especially useful for analyzing and characterizing biological samples [10–12]. Biological tissues are turbid media that strongly scatter light and present certain characteristic degree of anisotropy due to their ultrastructure components (e.g., the dense collagen and elastin matrix with oriented fibers, etc.) [13] so these properties lead to certain polarimetric response (depolarization, retardance, etc.). In several pathologies, morphological changes occur that alter the polarization properties of tissues [14]. Therefore, polarimetry constitutes a powerful and promising optical tool for the study and classification of biological samples [11,15]. Moreover, polarimetry can be combined with other optical techniques such as multispectral imaging [15–21] or microscopy [22–24].

The Mueller-Stokes (M-S) is a widespread mathematical formalism which allows to characterize the state of polarization of light beams (Stokes vector) and the polarimetric properties of samples (Mueller matrices, **M**). The Mueller matrix, MM, describes the effect of a material media, such as an optical element, or a sample, upon the polarization state of light by generating an output Stokes vector, S_{out}, from an input light state of polarization, S_{in}, for linear processes [25,26]. One interesting property of the (M-S) formalism is that it is based on radiometric measurements and therefore is easy to be implemented to interpret experimental data. In addition, unlike other formalism used to describe polarization, as for instance, the Jones formalism, the M-S is valid to deal with fully polarized light, but also with depolarization content. In

#426387 https://doi.org/10.1364/BOE.426387
Journal © 2021 Received 31 Mar 2021; revised 18 Jun 2021; accepted 25 Jun 2021; published 15 Jul 2021

Biomedical Optics EXPRESS

addition, the Mueller matrix provides the intensity and the complete set of polarization properties, diattenuation and polarizance (3-dimensional vectors **D** and **P**, respectively, in Eq. (1)), retardance and depolarization (both mixed in the 3×3 **m** submatrix) [27–30], which are sensitive to the microarchitecture of the tissue [16].

$$\mathbf{M} = m_{00}\begin{pmatrix} 1 & \mathbf{D}^T \\ \mathbf{P} & m \end{pmatrix}. \tag{1}$$

The polarimetric characteristics are not related to a particular element of the matrix **M** but they appear spread (or entangled) in multiple **M** elements in a complex way. When a physical model is not applicable, the polarimetric properties can be evaluated after separating the original **M** into a set of simpler matrices. The process of separation is known as matrix decomposition. In practice, there exists different strategies to perform **M** decomposition which can divided in three classes; product, sum and differential decompositions. The choice of the decomposition method is not trivial, and in general, it is specific to the sample being probed, and sometimes different decompositions may lead to comparable results. The best practice is to apply different decompositions and then to chose the one which performs the best in terms of computation time and reliability of the obtained results. For instance for the present work, after testing the symmetric decomposition [31] and the Lu-Chipman decomposition [32] we chose the later because it was more numerically efficient. Depolarization has been considered as a deletary effect which tends to screen or to erase the polarimetric properties (traditionally considered as the desired observables). Recently the introduction of different depolarization metrics [33–35], and in particular the depolarization spaces [15], show that there may be an interest of considering depolarization at the same level than the polarization properties. The different depolarization metrics are related to the way in which the purity of polarization of the illuminating beam is lost after interaction with the sample. In biological tissues, light scattering and the non uniform distribution of polarization properties along the probed samples, are the main causes of depolarization. Therefore depolarization metrics may unveil differences in the way light scattered or the polarization modified. Since scattering and polarization properties are linked to the ultrastructures present in the tissues, the depolarization metrics can serve to unveil differences or characteristics in the tissues, overlooked when probed with non polarized light, and which can present a true interest for biological / medical purposes or characterization. Although we focused the attention to the physical properties of the medium, other metrics such as the higher order moments of polarimetric observables recently discussed in the literature [36–39] are of potential interest and could be used in analogous procedures. Among the known depolarization metrics we chose to work with the indices of polarimetric purity (IPPs) [11,12,20]. The IPPs can be calculated from the combination of the eigenvalues (λ_i) of the covariance matrix **H(M)** associated to **M** [40]. These IPPs provide well-defined polarimetric magnitudes with easier physical interpretation and further synthetization of sample's depolarizing mechanisms.

In the present study, we develop a predictive optical model for tissue classification based on the statistical analysis of different polarimetric metrics, directly derived from the experimental Mueller matrix of a collection of several ex-vivo chicken samples. The efficiency of the developed models, when classifying different biological tissues, is given in terms of sensitivity and specificity. We also provide the potential of the developed models for tissue characterization imaging techniques. This study could be of interest for biomedical applications such as for guided systems in surgical procedures and for early detection of some diseases.

The outline of the manuscript is as follows. In section 2, we briefly review the mathematical background of the polarimetric metrics selected for the study. In section 3, we provide the experimental methodology and the pathological description of the four different studied tissues: tendon, muscle, myotendinous junction and bones (subsection 3.1). We also detail the description

Biomedical Optics EXPRESS

of the statistical analysis conducted on polarimetric data (subsection 3.2). In section 4, we describe the construction of an optical probabilistic model suitable for the classification of biological samples, which is based on a logistical regression applied to polarimetric data, and we discuss the validity of developed models in terms of sensitivity and specificity, as well as, their suitability for imaging classification of tissues. Finally, the main conclusions of the study are provided in section 5.

2. Methods

In this section, we present the mathematical background necessary to build the proposed tissue recognition model. The model is based on a set of metrics derived from experimental Mueller matrices of biological samples. Those observables allow us to obtain a clear interpretation of some physical properties of the samples. We also provide the experimental methodology followed for the completion of the study.

2.1. Polarimetric metrics

Recalling the **M** expression in Eq. (1) we define the diattenuation vector **D** as the metric that describes the property of an optical element by which the intensity of the exiting beam depends on the polarization state of the incident beam. Analogously, the polarizance vector **P** describes the polarization capability of a sample, when illuminated by an unpolarized input state [41]. Such magnitudes can be computed and described directly from the combination of **M** coefficients by using:

$$a)\ D = \frac{1}{m_{00}}\sqrt{m_{01}^2 + m_{02}^2 + m_{03}^2}, \qquad b)P = \frac{1}{m_{00}}\sqrt{m_{10}^2 + m_{20}^2 + m_{30}^2}, \tag{2}$$

where m_{ab} (a,b = 0, 1, 2, 3) represents an element of **M**.

Other polarimetric characteristics of a sample are also encoded in its Mueller matrix in a complex way. This information can be synthesized in a product of three pure Mueller matrices, which are functions of well-defined polarimetric magnitudes for an easier physical interpretation, by employing the so-called Lu-Chipman decomposition [32]:

$$M \equiv m_{00} \cdot M_{\Delta P} \cdot M_R \cdot M_D, \tag{3}$$

where m_{00} is related to the intensity transmittance or reflectance of the sample and those pure matrices are defined as depolarizers ($M_{\Delta P}$), pure retarders (M_R) and pure diattenuators (M_D).

Therefore, the information related to the depolarization, retardance and diattenuation properties of samples is encoded in different scalar polarimetric observables derived from those pure matrices. In particular, the total retardance R, the linear retardance δ and the optical rotation Ψ are obtained from the pure retarder matrix M_R in this way [28]:

$$R = \cos^{-1}\left|\frac{Tr(M_R)}{2} - 1\right|, \tag{4}$$

$$\delta = \cos^{-1}\left(\sqrt{[M_{R11} + M_{R22}]^2 + [M_{R21} - M_{R12}]^2} - 1\right), \tag{5}$$

$$\psi = \tan^{-1}\left(\frac{M_{R21} - M_{R12}}{M_{R11} + M_{R22}}\right), \tag{6}$$

where Tr denotes for the trace of a matrix, and $M_{Ri,j}$ are the coefficients of the pure retarder matrix.

The total retardance (R) allows the description of the global behavior of a general retarder (i.e., an elliptical retarder), as it combines the linear retardance and the optical rotation [19].

Biomedical Optics EXPRESS

However, in the context of tissue discrimination, to include individual characteristics of the retarder constituents, as the linear retardance (δ) or the rotation applied to the linear retarder (Ψ), may be of interest. Thereby, in the present work, all them will be analyzed individually in order to consider all the contributions.

Further description of a sample can be achieved by studying its depolarizing properties. To this aim, it is worth using the indices of polarimetric purity (IPP) [11,12,20]: three invariant depolarizing indicators that contain complete and detailed information of the capability of samples to induce randomness to input polarization states [29]. The nature of the IPPs lies in the statement that any depolarizer response can be synthesized as an incoherent sum of four non-depolarizing components whose relative statistical weights are performed by combinations of IPPs. In this way, we can discriminate different types of depolarizers by only looking at the statistical weight of each pure component.

The IPP observables are obtained from the four eigenvalues λ_i (which fulfill the condition $\lambda_0 \geq \lambda_1 \geq \lambda_2 \geq \lambda_3$) of the covariance matrix $\mathbf{H(M)}$, a positive semidefinite Hermitian matrix whose elements are defined through linear combinations of the Mueller matrix [20]. Particularly, the IPPs (real magnitudes P_1, P_2 and P_3) are defined as:

$$P_1 \equiv \frac{\lambda_0 - \lambda_1}{Tr(H)}, \qquad P_2 \equiv \frac{\lambda_0 + \lambda_1 - 2\lambda_2}{Tr(H)}, \qquad P_3 \equiv \frac{\lambda_0 + \lambda_1 + \lambda_2 - 3\lambda_3}{Tr(H)}. \tag{7}$$

In addition, the depolarization index P_Δ, which estimates the overall depolarization of a $\mathbf{M}$, is built from IPPs as follows:

$$P_\Delta = \frac{1}{\sqrt{3}}\sqrt{2P_1^2 + \frac{2}{3}P_2^2 + \frac{1}{3}P_3^2}, \qquad 0 \leq P_\Delta \leq 1. \tag{8}$$

These indicators characterize a pure (non-depolarizing) system when $P_1 = P_2 = P_3 = P_\Delta = 1$ is found [42]. Due to eigenvalue restrictions, IPPs are limited by the $0 \leq P_1 \leq P_2 \leq P_3 \leq 1$ inequalities. The ideal depolarizer is described for $P_1 = P_2 = P_3 = P_\Delta = 0$. Moreover, IPPs allow synthetizing the global information provided by P_Δ, allowing to transform a 1-dimensional space into a 3-dimensional information space.

3. Experimental methodology and statistical analysis

In this section, we provide the experimental methodology and the description of the inspected tissue samples (subsection 3.1) as well as the description of the data analysis conducted on Mueller matrix-based polarimetric data (subsection 3.2). The study provided in this section lays the foundation for the subsequent construction of an optical model suitable for the classification of biological samples, that will be discussed in section 4.

3.1. Experimental methodology and sample description

In the current study, the experimental Mueller matrices of 157 biological samples are measured at three different wavelengths covering the visible range (625 nm, 530 nm and 470 nm), since different wavelengths are associated with different light penetration capability in tissues [43]. Such measurements are performed in a scattering configuration (capturing diffuse light) by using a complete Mueller imaging polarimeter (described in section 1 of Supplement 1) based on parallel-aligned liquid crystal retarders [11,12,15,25].

The inspected tissues were obtained from 25 different *ex-vivo* specimens of chicken thighs provided by the same commercial brand. Tissues were taken from unambiguous anatomic locations under the supervision of a pathologist: we disinserted the muscles in order to dissect skeletal muscle, tendon and myotendinous junction as well as bone and worked with the same decomposition conditions. To do so, two samples per each tissue type were obtained (a total of

Biomedical Optics EXPRESS

23 bones, 50 muscles, 50 myotendinous junctions and 34 tendons) and subsequently frozen at $-16°C$. Three hours before measuring, the corresponding samples were systematically defrosted in order to equalize the decomposition stage. Also, histological sections from the same regions that were studied were collected and stained using a Masson Trichromic technique, which allows a clear-cut differentiation of all connective tissue (and otherwise mesodermal as bone and muscle). The detailed process and the histological images (Figs. S2 and S3) are provided in the section 2 of Supplement 1.

An example of an image of each one of the four studied tissues is shown in Fig. 1. Soft tissues (muscle, tendon and myotendinous junction) and bone present different properties owing to their physiologic functionality and structure. In the following, we describe a physiological analysis of each soft tissue in order to well-differentiate its structure and biological components. The skeletal muscle (Fig. 1(a)) is composed by contractile myofibril chains bundled into fascicles sheathed by perimysium (connective tissue which contains type-I collagen fibers [44]) and packed within a collagen-rich epimysium layer, which protects muscle fibers from friction. By comparison, tendons (Fig. 1(b)) are composed by dense collagen (60-80% type-I collagen) fibers arranged into parallel fascicles by following the same orientation as muscle fibers [45,46] which are completely enclosed by paratenon (a thick fibroadipose layer) and fascia (collagen-rich tissue that covers both muscle and tendon) [47,48]. In turn, myotendinous junction (Fig. 1(c)) is a variable combination of both previously mentioned soft tissues, which are progressively mixed along the transition muscle-tendon: contractile myofibrils and collagen fibers are bundled into separated fascicles [49] covered by fasciae. Each fascicle contains fibers of single type (either contractile or collagenous) but fascicles of each type are intermingled in varying proportions. Finally, bones (Fig. 1(d)) are composed by a dense matrix of collagen fibers with a varying arrangement that depends on the type of bony tissue examined. The cortical bony tissue of a long bone like the one here examined is formed by parallel collagen fibers arranged following the main axis of the bone and interrupted by *lacunae* (spaces containing cells) at more or less regular distances. Around *lacunae* collagen fibers are distributed concentrically. Mineral deposits are located in the collagen fibers and, particularly, in the spaces between adjacent collagen fibers.

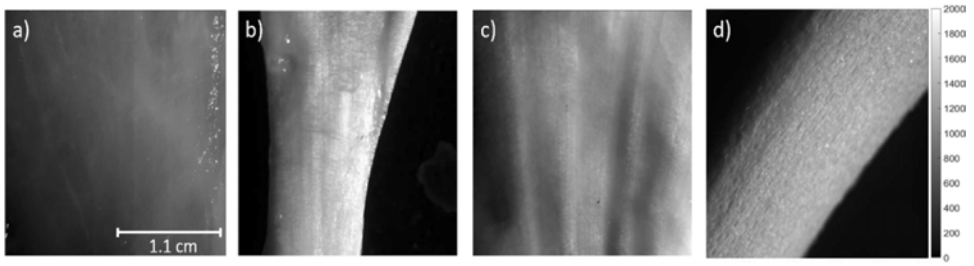


Fig. 1. Intensity image measured at 625 nm of a) muscle, b) tendon, c) myotendinous junction and d) bone tissues of a given chicken thigh. Images correspond to an area of 1.1×1.1 cm^2.

Obtained experimental Mueller images correspond to a region of interest (ROI) of 512×512 pixels which correlates to an area of 1.1×1.1 cm^2 (the images of the experimental Mueller matrices, MMs, as well as MMs of the decomposed parameters according to section 2 are presented on Figs. S4 to S11 on Section 3 of Supplement 1). The polarimetric analysis is performed for a pure tissue region, which exclusively contains the pixels related to the desired type of tissue and from which the background has been mostly removed. By using the measured experimental Mueller images, polarimetric images corresponding to different polarimetric metrics were calculated, according to the equations provided in section 2. To suppress the possible specular reflection effects on camera and to perfectly isolate the desired type of tissue, we selected,

Biomedical Optics EXPRESS

from the original 512 × 512 pixels polarimetric images, a specific sub-image with a sub-ROI of 150 × 150 pixels (which does not include the saturated -specular reflection originated- regions) from which we compute the mean value and standard error for the different polarimetric metric sub-images. The obtained standard errors of the mean for all the metrics calculated are within a value ranging between 0.006% and 0.02%, except for the optical rotation Ψ, whose highest standard errors are ranging from 0.3% to 3.4%: those optical rotation errors could be related with the heterogeneity of the tissue fibers orientations through the whole sample. The general low values of the standard errors suggest that the computed mean values are reliable and consistent, probably due to the large number of pixels (150 × 150) measured per sample.

3.2. Statistical treatment

The object of the study is a data matrix composed of 157 tissues cases, analysed through 27 predictor variables (9 polarimetric metrics measured at three wavelength channels; 625 nm, 530 nm, and 470 nm) and one classifier variable with four categories (the type of biological tissues: muscle, tendon, myotendinous junction and bone). The 9 selected polarimetric metrics, calculated from the experimental Mueller matrix images according to section 2, are: P_Δ, P_1, P_2, P_3, P, R, D, δ and Ψ. Under this scenario, the whole data matrix corresponds to 157 x (27 + 1) items. It is important to note that there are not conditional relations when experimentally measuring the 157 cases, so they can be considered as independent observations.

In the following, two statistical studies we applied on measured polarimetric data are described: (1) an exploratory univariant statistical analysis of the polarimetric data distributions profile; and (2) a factor analysis on data matrix with principal components extraction, based on the results obtained with the exploratory analysis, to obtain a more adequate basis to construct a logistic model capable to discriminate between the four studied tissue categories. These two studies are provided in subsections 3.2.1 and 3.2.2, respectively. Note that the different statistical tests, following discussed, applied on experimental data were conducted by using the SPSS software.

3.2.1. Exploratory univariant statistical analysis on polarimetric data distributions

The exploratory univariant statistical analysis applied on the data matrix above-mentioned was performed at a global level and according to classificatory variable groups (the type of tissues). Particularly, data distribution is analyzed by studying the central tendency, the dispersion and asymmetry behavior, outlier's identification and graphical differences between samples sorted by type of tissue characterization from Boxplot [50,51], as well as by checking normal distribution fit assumptions (assess of normality charts and contrast tests Kolmogorov-Smirnov and Shapiro-Wilk [52–55]). The conducted analysis proved that, in general, neither these data distributions nor their transformations (according to Tuckey's ladder of powers [56]) fit the normal distribution behavior. Particularly, primary statistics show that an asymmetric data cloud cannot be fitted into a normal distribution. In this context, we observe that the 27 metrics present asymmetry: skewness values different from 0 (associated with normal distribution) indicate the mean position with respect to the median. The closest skewness value to zero is -0.111 (linear retardance, δ, distribution at 625 nm) meanwhile the highest one, 2.257, is achieved by polarizance, P, measured at 530 nm. Only 7 from the 27 measured polarimetric distributions can be associated to normal distributions, with skewness values close to zero (they present values lower than |0.840|), and they are mostly found at 625 nm. The remaining 20 polarimetric distributions present higher skewness values, being the polarimetric distributions based on the 470 nm channel those with the clearest asymmetric tendency.

Additionally, regarding the Boxplot analysis [50,51], we evaluate the total percentage of outliers (mild and extreme) to show that there is no relevant measurement bias: among the different polarimetric distributions analyzed, the outlier values range from 0% to 5.09% of the total distribution data. Because we have not done any normality assumption on measured data and

Biomedical Optics EXPRESS

because the subsequent statistical analysis is based on robust statistical techniques, we conduct a conservative approach and those detected outlier values are not removed from the study.

In Fig. 2, we present, as a representative example to illustrate the small number of outliers associated to the observed data distributions, the Boxplot for the Index of Purity P_2 distributions measured at 625 nm illumination channel and corresponding to the four studied tissues. The few observed outliers are represented in Fig. 2 by circles (mild) and stars (extremes). What is more, the asymmetric tendency of the distributions is suggested by the relative position of the median, which is not equidistant from the first and third quartiles (they are not placed at the middle of boxes). In addition, the tissue discrimination capability of the metrics is also suggested by the boxplot analysis as medians do not fit within the boxes of other tissues. For instance, Fig. 2 illustrates the potential of P_2 to discriminate muscle among the remaining tissues (see dashed red line). Similar tendency is obtained when analyzing 530 nm and 470 nm illumination wavelength (represented in Supplement 1, Figs. S12 and S13, respectively). These results can be extrapolated to the remaining **M**-metrics studied in this work.

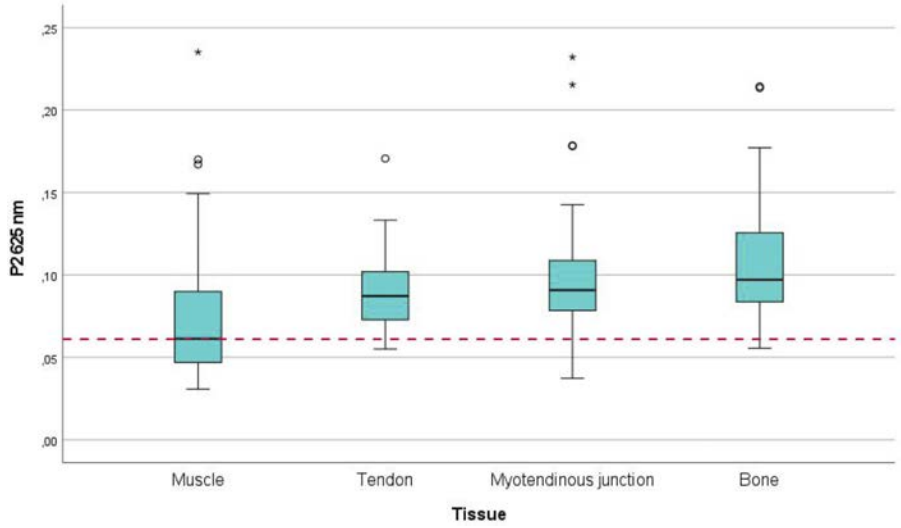


Fig. 2. Boxplot of P_2 index for all tissues measured at 625 nm illumination channel. Red-dashed line visually represents the potential of the metric to discriminate muscle among remaining tissue types: the median of the muscle box does not fit within the other tissues' boxes. Boxplot points out the low quantity of outliers on data distributions (which can be extrapolated to the remaining metrics): mild and extreme values are represented by circles and stars, respectively.

The normality analysis above conducted on polarimetric distributions proved that subsequent data analysis must be done by following non-parametric statistics. Under this scenario, the different polarimetric distributions, sorted by type of tissue, were studied by performing a non-parametric homogeneity multiple contrast analysis: the Kruskal-Wallis test [57,58]. This test allows us to determine if there are significant statistical differences between two or more data distributions. Therefore, in our case, the outputs are connected with the tissue discriminatory potential of the studied metrics: p-values lower than 0.05 confirm, with a 95% of confidence, the capability of the polarimetric indicator to discriminate the origin of the data between different pair of tissues.

Concerning the Kruskal-Wallis test results (whole data is presented in Table S1 of Supplement 1) we confirm that P_Δ, IPPs (P_1, P_2 and P_3) and polarizance P, are the most sensitive and

Biomedical Optics EXPRESS

interesting polarimetric indicators for the studied tissue characterization, while the optical rotation Ψ is not capable to differentiate between any pair of tissues. Particularly, polarizance P seems to be the only metric capable to discriminate between myotendinous junction and bone pair. With regards to retardance indicators, they also provide some predictive potential. From this, we conclude that the main polarimetric characteristics of samples (selective absorption, retardance and depolarization) present certain signatures in the studied tissues, and therefore, the ensemble of polarimetric indicators selected seems to be adequate to be used in optical models to discriminate between the studied tissues.

The univariant analysis results above-discussed, particularly the non-normal tendency of studied polarimetric distribution, manifest the impossibility to implement our solution based on techniques such as discriminant analysis: this requires, in addition to the assumption of equality of original metrics, sorted by type of tissue, variance-covariance matrices, that these metrics follow a multivariant normal distribution behavior. As a consequence, our proposal considers an elaborated model based on a more robust technique, the logistic regression, which does not require these assumptions. At the same time, a dimension reduction of the original matrix in pursuit of non-correlated predictors (that minimize or delete the multicollinearity that generates instability on models) is aimed. Particularly, this solution considers:

(1) A factor analysis with principal components extraction [59–73] that provides a set of independent predictor variables (defined as principal components) that maintains the information contained in the original variables and describes an acceptable proportion of the variance or inertia of point clouds.

(2) The implementation of multivariant predictive models, based on binary logistic regression [74] techniques with the extracted principal components as parameters, for the type of biological tissue classification, as well as the independent predictor and implemented model associated to ROC curves [75–78] draft for their predictive capacity evaluation, comparison and optimization.

The factor analysis is described in subsection 3.2.2, whereas the implementation of the statistical models and their evaluation are described in section 4.

3.2.2. Factor analysis: principal components extraction

The goal of this work is to implement multivariant probabilistic models, particularly a binary logistic regression, with the capability of determining whether a certain sample corresponds to any of the chicken tissue categories studied in this work: muscle, bone, tendon and connective tissue. Because the application of logistic regression on data does not necessarily converge when using a non-orthogonal basis, it is convenient to obtain a set of uncorrelated (orthogonal) predictors for the regression formulation which, in turn, define a reduced-dimension space. To do so, when performing factor analysis, we choose the principal components extraction as it always provides a solution in which the output factors compound a set of independent (uncorrelated) variables obtained by linear combinations of the original indicators. The extracted principal components maintain the same information of the original variables while keeping the maximum possible variance: they constitute the ideal option to implement on predictive models. It is important to remark that, because of the uncorrelation of these new components, the multicollinearity problem in a regression model is avoided.

Before conducting principal components analysis (PCA) on our particular polarimetric data matrix, we have previously computed the Bartlett's test of Sphericity [79] and the Kaiser-Meyer-Olkin Measure of Sampling Adequacy (KMO) [80] to ensure the suitability of factor analysis for the particular studied polarimetric data. Particularly, Bartlett's test of Sphericity is based on the null hypothesis according to which the population correlation matrix is equal to the identity matrix: in such case, the factor model is not appropriate to treat the input data [79]. On the other

Biomedical Optics EXPRESS

hand, KMO allows evaluating the degree in which each variable is predictable from the remaining ones: obtained values larger than 0.5 indicate that at least one common factor underlying the observed variables exists [80]. The output of both tests shows the suitability of applying factor analysis (PCA) to our studied data matrix. Bartlett's test clearly rejects the null hypothesis (p-value 0.000). Likewise, Kaiser-Meyer-Olkin outputs a value of 0.650, larger than the typical cut-off, meaning that there's a common subjacent factor for the studied variables. Because a correlation between the variables exists, the factor model is an appropriate methodology for treating the data. Hence, considering the output of both Bartlett's and KMO tests, they all suggest that PCA may suit our particular data set when being applied.

Once the Principal Components have been calculated by using the SPSS software on the data matrix, we have computed the scree plot [71] (Fig. 3) which allows us to select a reasonable number of components to be used for the predictive models. In particular, we decided to hold all components preceding the sedimentation zone: 10 components explain more than 90% of the original metrics variance.

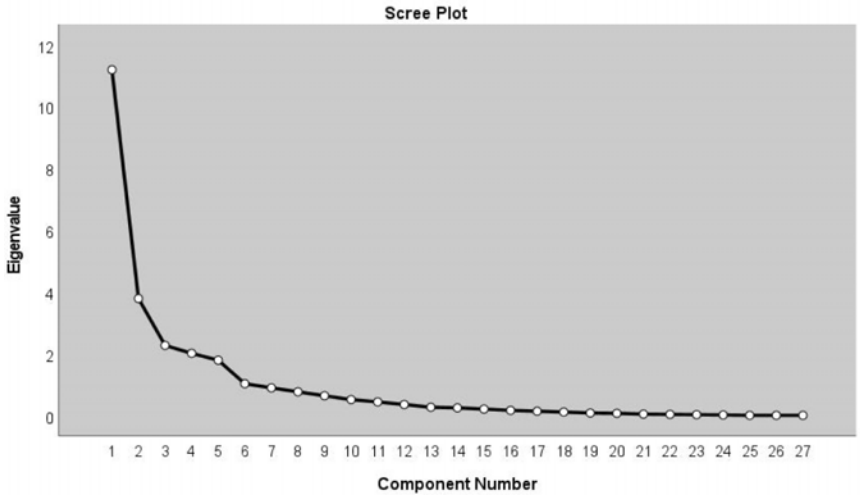


Fig. 3. Scree plot of the principal component analysis.

The eigenvalues of those components, together with the data variance (in %) and the cumulative variance (in %) are shown in Table S2 of the Supplement 1. In this way, PCA reduces the 27-dimension space of **M**-metrics into a new 10-dimension space by identifying the original variables set underlying dimensions. The extracted factors cannot be directly measured but they enable the data structuration around a reduced number of variability axis. As the new factors are mathematical constructions, they do not necessarily have to be connected to physical properties, but they can be further interpreted by analyzing their dependence with polarimetric metrics. From a geometrical point of view, the principal components correspond to axes perpendicular to each other that better fit with the point cloud that sets the data matrix.

4. Predictive model construction: results and discussion

The selected principal components encode the polarimetric information so we can write each component as the linear combination of the 27 studied polarimetric metrics weighed by some constants provided by the so-called component score coefficient matrix, shown in Table S3 of Supplement 1. The fact that the component score coefficient matrix allows expressing the different polarimetric metrics as a linear combination of the principal components, or vice versa,

Biomedical Optics EXPRESS

is determinant because this situation defines the basis for building a probabilistic model. In a general way, the principal components (written as $\widehat{C}$) are obtained by multiplying the **M**-metrics 27×1-dimension vector, $\widehat{P}$, by the transposed component score matrix, $\widehat{CS}$ (a 27×10-dimension data matrix in Table S3):

$$\widehat{C} = \widehat{CS}^T \times \widehat{P}, \tag{9}$$

where $\widehat{C}$ elements correspond to C_i ($i = 1, \ldots, 10$) and $\widehat{P}$ is written as follows,

$$\widehat{P} = (P_{\Delta R}, P_{1R}, P_{2R}, P_{3R}, P_R, R_R, D_R, \delta_R, \Psi_R, \ldots, \Psi_G, \ldots, \Psi_B)^T, \tag{10}$$

where R, G and B indicate the wavelength channel (625 nm, 530 nm and 470 nm, respectively) of the polarimetric indicator and the order of the metric is conserved along the vector.

To illustrate the correspondence between the principal components and the physical variables, we represent a two-dimensional space plot where the 27 indicators are scattered as a function of the two main principal components, C_1 and C_2. The obtained results are shown in Fig. 4.

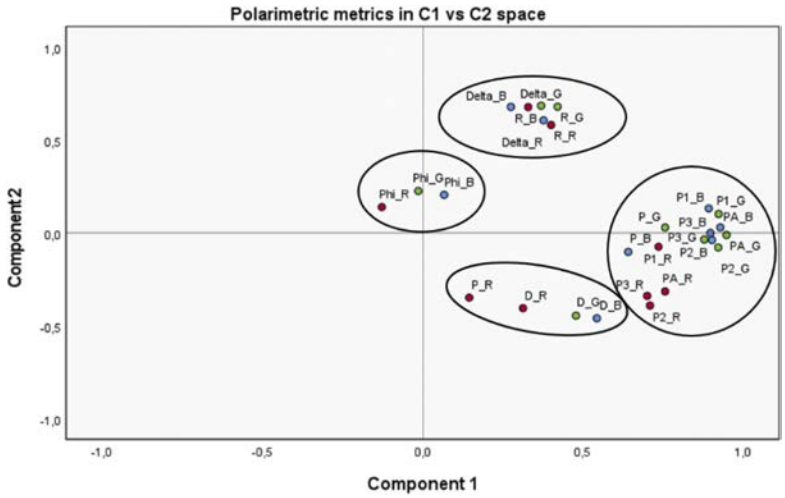


Fig. 4. The plot showing principal components C_1 against C_2 represents the correlation coefficients between the 27 polarimetric indicators and the two first extracted principal components. The notation of the polarimetric indicators names is composed of the word that represents the measured parameter: P1, P2, P3 and PA (IPPs and P_Δ, respectively), D and P (diattenuation and polarizance), R and Delta (global and linear δ retardance) and Phi (optical rotation Ψ), followed by R, G or B (corresponding to red, green and blue measured wavelength, respectively).

Note that the metrics are clustered in four main areas (highlighted by black circles) and more interestingly, those clusters mainly coincide with their physical characteristics regardless its measured wavelength: retardance related indicators (total and linear retardance, R and δ respectively), depolarization (IPPs and P_Δ), dichroism (diattenuation and polarizance, D and P, respectively) and optical rotation Ψ. By comparing the weight or influence of the two first principal components when analyzing polarimetric indicators we find that IPPs and P_Δ (scattering, depolarization) are mostly described by C_1 as they are clustered with weights around 0.5 and 1.0 meanwhile the weights of C_2 are between 0 and 0.5. When looking at how retardance indicators are grouped, we find that the presence of the C_2 principal component (between 0.5 and 1.0) is higher than the influence on the C_1 component (for the last one, its weight value is close to 0). For diattenuation and polarizance, they are both influenced by the two main principal components

Biomedical Optics EXPRESS

but they are especially represented by the C_1 component as most dispersion is observed in such a component. Finally, optical rotation principal components C_1 and C_2 values are both near 0 so they do not provide enough information to describe Ψ (the two first principal components are not able to explain optical rotation so we expect to find weights different from zero when analyzing the remaining components).

Pointing out some interesting facts: plotting principal components against each other provides us a very intuitive way to understand in which way those components are connected with actual physical information of samples (the 27 polarimetric indicators), and which one of those polarimetric indicators (or collection of indicators) carry more data variability in the original raw data matrix. Interestingly, the fact that physical variables of the same type (for instance retardance) are grouped in well-defined areas of the C_1-C_2 space (or of other combinations of 2-dimensional C_i spaces) indicates that different origin in data variability (orthogonal dimensions) is well connected with actual different physical structures present in the studied tissue samples.

The discussion of the principal component's space characteristics (such as its discriminatory potential, sensitivity and specificity) and the predictive model construction of each type of tissue and its predictive features are described in subsections 4.1 and 4.2, respectively.

4.1. Discriminatory potential of the principal components space

In this section, we are interested in studying the the performance of 10-principal components space as tissue classifiers. With this aim, the Receiver Operating Characteristic (ROC) curve analysis [75–78] is carried out: it describes the performance of a classifier (an algorithm or a particular variable) when classifying measures into two categories by plotting the true positive rate (TPR), or sensitivity, against the false positive rate (FPR), or 1-specificity, for multiple threshold values of the classifier. For each classifier and measure, the combination of what the sample actually is (real value, positive or negative) and how it is classified by the model (prediction; positive or negative) results in four possible outcomes [76,77]. The mathematical relations connecting these four criteria define the sensitivity and specificity, which are the base of ROC curves, and are given by:

$$Sensitivity = \frac{TP}{P}, \tag{11}$$

$$Specificity = \frac{TN}{N}, \tag{12}$$

where TP, P, TN and N denote for true positives, total positives, true negatives and total negatives, respectively. Another way to define sensitivity is the ability to correctly identify positive samples while specificity results in the opposite case, correctly identifying negative samples [78].

By using the SPSS software, we computed the ROC curve for all the 10-principal components and for the four different studied tissues (a specific dichotomic predictive model is constructed for each studied tissue, so four models are constructed). As an example, the ROC curve of the principal component C_1 for each biological tissue is shown in Fig. 5. Note that the different threshold are all the possible values in C_1. However, all principal components present statistical significance in ROC curves for one or more tissues, which bolsters the interest in keeping them in the process.

Usually, to compare the performance of classifiers (in our case, to estimate which component provides better sensitivity-specificity values for a particular tissue), the area under the curve (AUC) is calculated, with values ranging from 0, when the variable has no predictive capability, to 1, with 100% both sensitivity and specificity. The AUC values of all principal components when discriminating each biological tissue against the remaining three are summarized in Table 1.

With regards to the exposed information, we see how most of the principal components provide certain discriminative potential, with values larger than 0.5. The components with the largest AUC for a particular tissue are highlighted in Table 1. Notice that C_1 present significant

Biomedical Optics EXPRESS

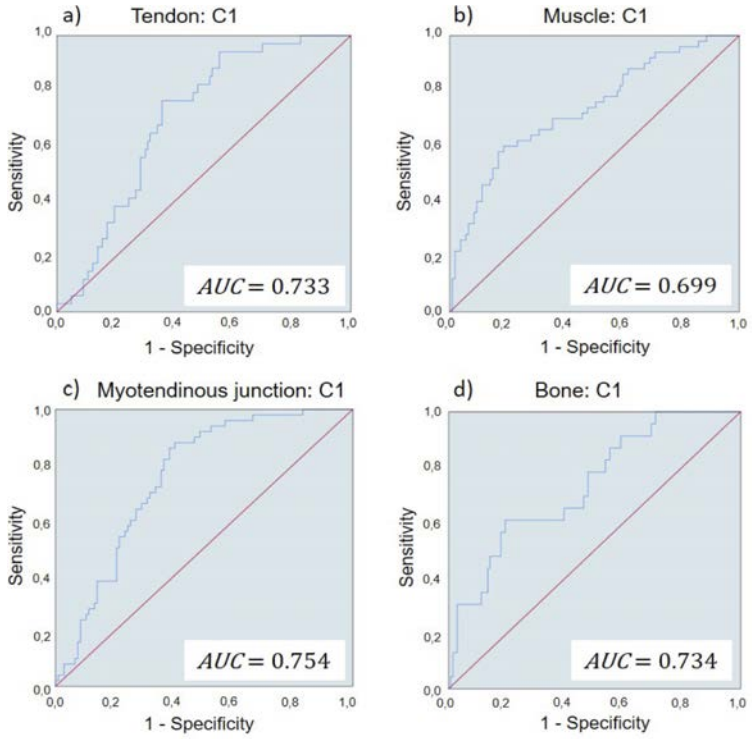


Fig. 5. ROC curve of the principal component C_1 for (a) muscle, (b) tendon, (c) myotendinous junction and (d) bone.

Table 1. AUC of 10 first principal components for all tissues.

	Muscle	Tendon	Myotendinous junction	Bone
C_1	0.733	0.699	0.754	0.734
C_2	0.559	0.717	0.567	0.577
C_3	0.528	0.651	0.583	0.512
C_4	0.731	0.745	0.563	0.542
C_5	0.667	0.733	0.500	0.527
C_6	0.516	0.714	0.521	0.725
C_7	0.567	0.567	0.508	0.722
C_8	0.703	0.706	0.544	0.649
C_9	0.581	0.540	0.616	0.507
C_{10}	0.569	0.495	0.623	0.601

discrimination values in all tissues, C_2 and C_3 mostly provide discriminatory information for tendon, C_4 and C_5 focus on muscle and tendon discrimination, C_6 on tendon and bone, C_7 on bone, C_8 keeps the information about muscle, tendon and bone and, finally, C_9 and C_{10} provide discrimination for the myotendinous junction.

Interestingly, it has been noticed that the principal components predictive potential (given by the large area under the curve, AUC, values exposed in Table 1) is different for distinct tissues. Such behavior was expected as every component is written as a linear combination of

Biomedical Optics EXPRESS

the polarimetric indicators (rows in Table S3 in Supplement 1), and the metric values strongly depend on the composition characteristics of each biological tissue. Therefore, considering the structural differences between them, some polarimetric indicators and, consequently, some principal components, are expected to have a higher capacity to discriminate a particular tissue between the rest.

4.2. Predictive model construction

In this section, the determination of the predictive model for tissue recognition is described. The constructed model is based on a probabilistic function achieved by fitting the experimental data (array of principal components values for all chicken samples) to a sigmoid function, more specifically, to a logistic function. In other words, we perform a non-linear data model transformation into a linear one by means of a logistic regression fit of data. The logistic function with the principal components as variables and with a curve maximum value equal to 1 is written as follows,

$$p = \frac{1}{1 + e^{-\left(\beta_0 + \sum_i^n \beta_i C_i\right)}}, \tag{13}$$

where C_i are the i principal components, β_i the weights of the components and β_0 a constant value. The function in Eq. (13) is bounded between 0 and 1 and can be interpreted as the probability of a given outcome.

We have chosen a logistic function because it is valid for non-parametric data and it does not require the relation between the predictors and the probability of a target outcome to be linear. Moreover, despite other fitting methods, as it is the case of those based on the ordinary least square regression (OLS), the logistic function neither requires the residuals to be normally distributed and exhibit constant variance. Particularly, we have previously verified the non-dependency of the model predictors (principal components) and their connection with the dichotomic dependent variable. Moreover, by performing the PCA, the multicollinearity between predictors have been removed so stable predictive models and convergent iteration processes are achieved. Under this scenario, four logistic regression functions are designed: one model for each studied biological type of tissue. As seen in Eq. (13) the obtained probability function depends on some of the principal components factors which, in turn, depend on the polarimetric indicator values measured from the image sample. The logistic regression fit on the experimental principal components (obtained from the experimental polarimetric indicators according to the relations provided in the component score coefficient matrix in Table S3 in Supplement 1) is conducted by using the SPSS software. By applying a stepwise regression approach (backward elimination) and using the Wald estimator throughout multiple steps, the principal components are removed until only the most significant ones remain [74].

This routine gave us the values for the weights in Eq. (13) and, accordingly, the obtained probabilistic functions for the four studied tissues are exhibit below:

$$P_{Muscle} = \frac{1}{1 + e^{-(-2.183C_1 + 1.535C_4 - 1.280C_5 - 1.077C_8 + 0.629C_9 - 1.879)}}, \tag{14}$$

$$P_{Tendon} = \frac{1}{1 + e^{-(-2.213C_1 - 2.358C_4 + 1.625C_5 + 1.257C_6 + 1.059C_8 - 3.128)}}, \tag{15}$$

$$P_{Myotendinous\ junction} = \frac{1}{1 + e^{-(0.900C_1 - 0.573C_9 + 0.589C_{10} - 0.934)}}, \tag{16}$$

$$P_{Bone} = \frac{1}{1 + e^{-(1.018C_1 + 0.715C_2 + 0.621C_4 - 1.452C_6 - 1.437C_7 - 0.688C_{10} - 3.132)}}. \tag{17}$$

To illustrate the goodness-of-fit for the four regressions, the Hosmer-Lemeshow significance [79,80] and the R^2 of Nagelkerke [81,82] indicators are provided in Table S4 of Supplement 1.

Biomedical Optics EXPRESS

Afterwards, to study the efficiency of the probabilistic models in terms of sensitivity and specificity, the associated ROC curves were also computed for each one of the four models. They are provided in Fig. 6, in which their corresponding AUC are also indicated. We see how the AUC are significantly larger than those we obtained when representing the principal components (see Fig. 5 as an example), so we see how the discriminatory potential of the probabilistic functions clearly overcome the obtained with the principal components by themselves. We want to highlight that the AUC values obtained for muscle, tendon and bone tissues models, are significantly high (0.92, 0.95 and 0.89 respectively), providing the efficiency of these models for tissue discrimination. In turn, the model providing lower discriminatory potential is that associated to the myotendinous junction (AUC=0.79). This situation agrees with the fact that only C_1, C_9 and C_{10} components present discriminatory potential for the myotendinous junction tissue (see Table 1), and the two latter, represent a very small data variation of the whole data matrix (2.350% and 1.885% of the variance (Table S2), respectively).

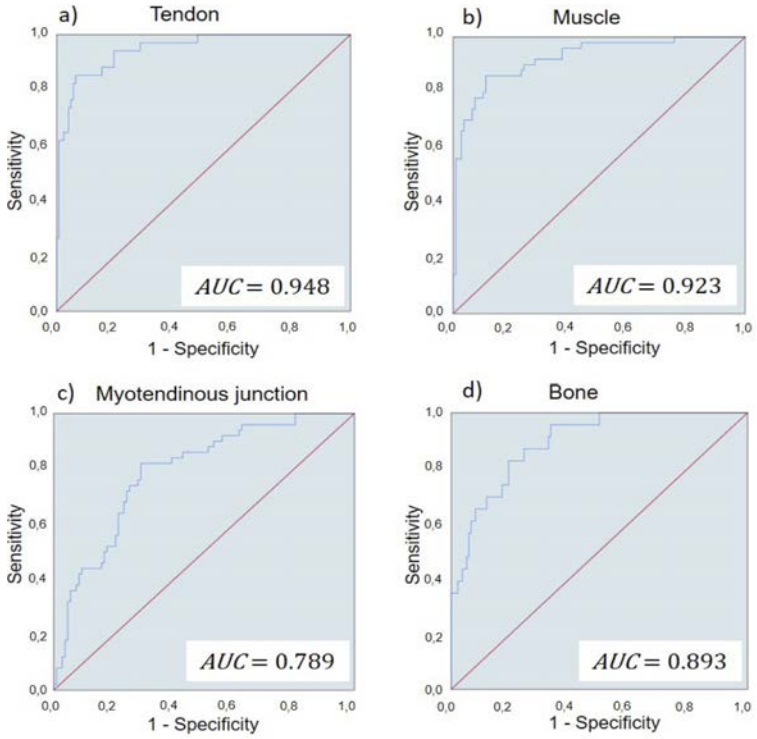


Fig. 6. ROC curve of the probabilistic model for (a) tendon, (b) muscle, (c) myotendinous junction and (d) bone.

Note that models in Eqs. (14–17) provide as output a real number between 0 and 1, associated to the probability of a given tissue to belong to a particular tissue category. To build a dichotomic model determining if a tissue belongs or not to a particular category, a threshold (specific probability value) must be set in a way that values above/below such threshold are associated with Yes/No answers (i.e., is the unknown sample a bone? Yes-No; is it a tendon? Yes-No, etc.). However, each possible threshold (probability), as can be seen from ROC curves in Fig. 6, will lead to a particular pair of sensitivity and specificity values for the constructed models. In our work, the criterion used to set such threshold is by using the Youden's Index [83] of the ROC curves in Fig. 6. According to this criterion, the optimal cut-off for our models corresponds to

Biomedical Optics EXPRESS

the farthest point of the probabilistic function in the ROC curves (blue curve in Fig. 6), from the ROC diagonal (red line in Fig. 6) along Y axis, which is calculated as the maximum value, d_{max}, of d = sensitivity + specificity – 1.

The computed Youden's index [83] for each model, and its corresponding sensitivity and specificity values, are shown in Table 2. Note that the models for the muscle, tendon and bone present high values of specificity and sensitivity (values higher than 80% in all the cases). In turn, the worst obtained result is that for the myotendinous junction tissue specificity, which descends to a value of 71% (in agreement with the limited potential of the myotendinous model discussed in the previous analysis).

Table 2. AUC, Youden's index, sensitivity and specificity of each predictive model.

	Muscle	Tendon	Myotendinous junction	Bone
AUC	0.923	0.948	0.789	0.893
Threshold (Youden's Index)	0.392	0.351	0.285	0.162
Sensitivity (%)	86.0	85.3	82.0	82.6
Specificity (%)	88.8	93.5	71.0	80.6

Therefore, the sensitivity and specificity obtained results for the constructed predictive models (data in Table 2) quantify the efficiency of those models to categorize the studied organic tissues. What is more, the current study highlights the potential of predictive models based on polarimetric data to discriminate between animal tissues, and the methods provided could be applied in the future, for instance, for medical applications.

Following this same idea, to show the potential of the method to be applied to imaging techniques to discriminate between tissues, we encoded an algorithm that computes the probabilistic function for each pixel of the measured sample. In particular, for an arbitrary chicken tissue, the Mueller matrix image is measured at the three studied wavelengths (625 nm, 530 nm, 470 nm), from which the 27 polarimetric images are calculated. From this polarimetric image database, the probability image of the arbitrary tissue to be categorized as a particular category (muscle, bone, tendon and myotendinous junction) is calculated according to the logistic probability functions in Eqs. (14–17). The algorithm outputs which is the probability of the analyzed pixel to be recognized as a particular tissue, so the probability image is constructed, with values ranging between 0 and 1.

As an example, in Fig. 7 we provide the probability function images, corresponding to the four predictive models when analyzing an arbitrary sample of a chicken tendon. Figure 7 presents the intensity image of the tendon (M00, Fig. 7(a)), and the output images of the probabilistic models for the recognition of muscles (Fig. 7(b)), tendons (Fig. 7(c)), myotendinous junctions (Fig. 7(d)) and bones (Fig. 7(e)). Probability images are given in grayscale, in which white is associated to the maximum probability value (1) and black to the minimum (0). By comparison, the highest probabilistic values (closer to 1) are clearly obtained by the tendon predictive model (Fig. 7(c)), recognizing almost all the pixels along the tendon tissue image, and thus, perfectly classifying the sample as a tendon. In turn, the probability images for muscle (Fig. 7(b)), myotendinous junction (Fig. 7(d)), and bone (Fig. 7(e)) are based, despite some clustered high output pixels, on significantly smaller probabilistic values than those present in tendon recognition image (Fig. 7(c)). Therefore, those models do not assign the tendon sample as muscle, myotendinous junction or bone, respectively. Among these three probability images, the larger values are obtained when using the myotendinous junction model, providing that this model prediction is the less sensitive and specific one, in agreement with previous discussions on model efficiency.

Analogously, we provide a second example, for the study of an arbitrary muscle sample. Results are provided in Fig. 8, for the intensity image (M00, Fig. 8(a)), and the four predictive model output images (muscle, tendon, myotendinous junction and bone recognition, Figs. 8(b-e),

Biomedical Optics EXPRESS

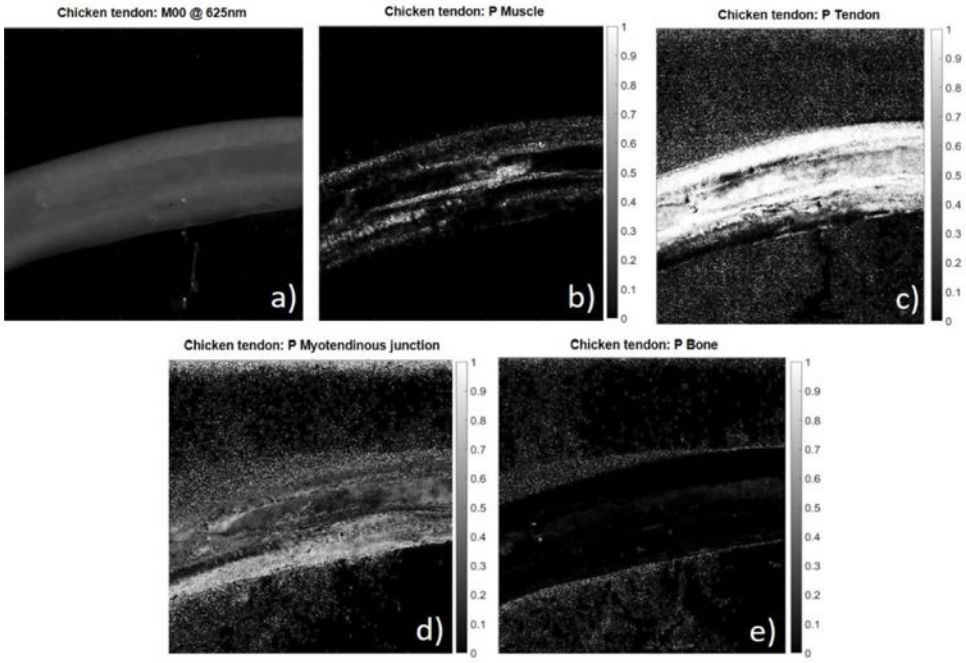


Fig. 7. Intensity image M00 (a) and probability image outcome when applying the Muscle-model (b), Tendon-model (c), Myotendinous junction-model (d) and Bone-model (e) on an arbitrary chicken tendon sample. The gray level bars, placed to the right of the corresponding probability function images, defines the probability of the pixel to be recognized as a particular tissue, in a range between one (white) or zero (black).

respectively). The obtained predictive results demonstrate the high capability of the muscle-recognition model: the highest overall values for the output probability image correspond to the muscle recognition model (Fig. 8(b)), perfectly classifying the muscle tissue. When analyzing the probability distribution for the three remaining models, a good discriminating potential is also demonstrated: the sample's output shows low probability values for classifying the muscle as a tendon (Fig. 8(c)), a myotendinous junction (Fig. 8(d)) or a bone (Fig. 8(e)). Once again, and as expected by the previous analysis (sensitivity and specificity model values in Table 2), the worst predictive efficiency is related to the myotendinous junction model, leading to larger probability values than the tendon and bone models.

Note that the tendon and muscle study cases (results in Figs. 7 and 8) highlight the strong qualitative and quantitative discriminative potential of the models to be applied for tissue recognition imaging techniques. For the sake of completeness, the remaining tissue study cases, bone and myotendinous junction recognition test, can be found in section 6 of Supplement 1. These examples (Fig. 7, Fig. 8 and Fig. S14 and Fig. S15 of Supplement 1) point in the direction that the proposed models are suitable for tissue classification and recognition, and the corresponding sensitivity and specificity outcomes will be limited by data provided in Table 2.

Biomedical Optics EXPRESS

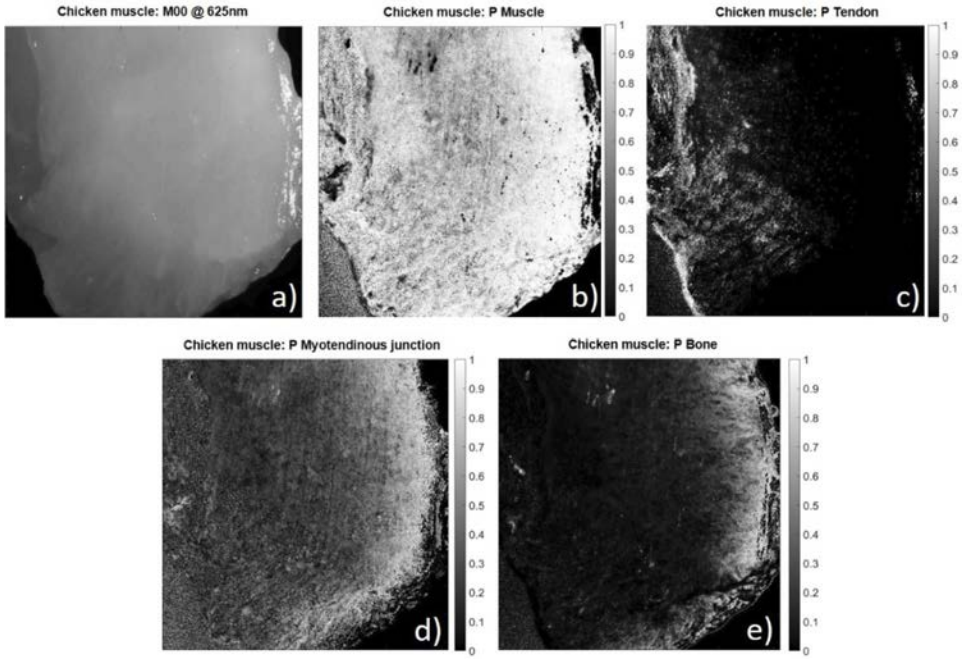


Fig. 8. Intensity image M00 (a) and probability function of muscle (b), tendon (c), myotendinous junction (d) and bone (e) for chicken muscle measurements. The gray level bars, placed to the right of the corresponding probability function images, defines the probability of the pixel to be recognized as a particular tissue, in a range between one (white) or zero (black).

5. Conclusions

Four predictive models based on the measure of diverse polarimetric metrics derived from the experimental MM of four ex-vivo chicken tissues (bone, tendon, muscle and myotendinous junction tissue) have been designed for tissue recognition and provided in this manuscript. The predictive models start by studying the distribution of the experimental data and their different statistical origins. Obtained results proved that we deal with non-parametric data with significant potential to discriminate between studied tissues. Afterwards, with the idea of constructing a robust non-parametric predictive probabilistic model, we applied a Principal Component Analysis on data, which allowed us to reduce the dimension of the data space (from 27-dimension to 10-dimensions) but dealing with maxima information as well as to obtain an orthogonal basis of parameters. By analyzing the connection between the principal components basis and the original polarimetric metrics, we realized that different variability directions of data in the principal components space can be roughly associated by different physical origins of the tissue structures (dichroism, retardance and depolarization).

Subsequently, based on the computed principal components basis, we constructed a logistic regression on data, leading to four probabilistic models (a model for each one of the four studied tissues) able to categorize if an arbitrary tissue belongs to a certain tissue category. The sensitivity and specificity values reached for each one of the four models are respectively: 85.3% and 93.5% (Tendon-model), 86.0% and 88.8% (Muscle-model), 82.6% and 80.6% (Bone-model) and 82.0% and 71.0% (Myotendinous junction-model), proving the potential of the method.

Finally, to highlight the suitability of the provided methods to be applied in biological samples imaging recognition techniques, we computed an algorithm, based on the constructed probabilistic

Biomedical Optics EXPRESS

models, that provides the probability, between 0 to 1, of a given pixel to be categorized as a particular tissue, and this is done for all the pixels of a given tissue image. In short, the procedure outputs four different probability images, corresponding to the probability of an arbitrary tissue to be muscle, bone, tendon or myotendinous junction tissue. The method provided complementary visual interpretation for tissue recognition, and a satisfactory categorization of the analyzed tissue image, according to the sensitivities and specificities stated above.

The proposed non-invasive methods discussed in this work could be applied, by conducting the required statistical data feeding, in multiple biomedical scenarios, for example, for the early diagnosis of some pathologies.

Funding. Generalitat de Catalunya (2017-SGR-001500); Ministerio de Economía, Industria y Competitividad, Gobierno de España (Fondos FEDER, RTI2018-097107-B-C31).

Acknowledgements. We acknowledge B. Jauregui-Ríos for the fruitful discussions about the statistical treatment on data.

Disclosures. The authors declare no conflicts of interest.

Data availability. Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

Supplemental document. See Supplement 1 for supporting content.

References

1. J. C. Suárez-Bermejo, J. C. G. D. Sande, M. Santarsiero, and G. Piquero, "Mueller matrix polarimetry using full Poincaré beams," Optics and Lasers in Engineering **122**, 134–141 (2019).
2. S. A. Hall, M. A. Hoyle, J. S. Post, and D. K. Hore, "Combined Stokes vector and Mueller matrix polarimetry for materials characterization," Anal. Chem. **85**(15), 7613–7619 (2013).
3. Q. H. Phan, P. M. Yang, and Y. L. Lo, "Surface plasmon resonance prism coupler for gas sensing based on Stokes polarimetry," Sensors and Actuators B: Chemical **216**, 247–254 (2015).
4. J. Hough, "New opportunities for astronomical polarimetry," Journal of Quantitative Spectroscopy and Radiative Transfer **106**(1-3), 122–132 (2007).
5. S. H. Yoo, R. Ossikovski, and E. Garcia-Caurel, "Experimental study of thickness dependence of polarization and depolarization properties of anisotropic turbid media using Mueller matrix polarimetry and differential decomposition," Appl. Surf. Sci. **421**, 870–877 (2017).
6. V. V. Tuchin, "Polarized light interaction with tissues," J. Biomed. Opt. **21**(7), 071114 (2016).
7. S. Anwar, S. Firdous, A. Rehman, and M. Nawaz, "Optical diagnostic of breast cancer using Raman, polarimetric and fluorescence spectroscopy," Laser Phys. Lett. **12**(4), 045601 (2015).
8. D. Ivanov, V. Dremin, A. Bykov, E. Borisova, T. Genova, A. Popov, R. Ossikovski, T. Novikova, and I. Meglinski, "Colon cancer detection by using Poincaré sphere and 2D polarimetric mapping of ex vivo colon samples," J. Biophotonics **13**(8), e202000082 (2020).
9. C. H. L. Patty, D. A. Luo, F. Snik, F. Ariese, W. J. Buma, I. L. ten Kate, R. J. M. van Spanning, W. B. Sparks, T. A. Germer, G. Garab, and M. W. Kudenov, "Imaging linear and circular polarization features in leaves with complete Mueller matrix polarimetry," Biochimica et Biophysica Acta (BBA) - General Subjects **1862**(6), 1350–1363 (2018).
10. I. Ahmad, A. Khaliq, M. Iqbal, and S. Khan, "Mueller matrix polarimetry for characterization of skin tissue samples: a review," Photodiagn. Photodyn. Ther. **30**, 101708 (2020).
11. A. Van Eeckhout, A. Lizana, E. Garcia-Caurel, J.J. Gil, A. Sansa, C. Rodríguez, I. Estévez, E. González, J.C. Escalera, I. Moreno, and J. Campos, "Polarimetric imaging of biological tissues based on the indices of polarimetric purity," J. Biophotonics **11**, e201700189 (2017).
12. A. Van Eeckhout, A. Lizana, E. Garcia-Caurel, J.J. Gil, R. Ossikovski, and J. Campos, "Synthesis and characterization of depolarizing samples based on the indices of polarimetric purity," Opt. Lett. **42**(20), 4155–4158 (2017).
13. F. Fanjul-Vélez, N. Ortega-Quijano, and J. L. Arce-Diego, "Polarimetry group theory analysis in biological tissue phantoms by Mueller coherency matrix," Opt. Commun. **283**(22), 4525–4530 (2010).
14. S. Badieyan, A. Ameri, M. R. Razzaghi, H. Rafii-Tabar, and P. Sasanpour, "Mueller matrix imaging of prostate bulk tissues, polarization parameters as a discriminating benchmark," Photodiagn. Photodyn. Ther. **26**, 90–96 (2019).
15. A. Van Eeckhout, E. Garcia-Caurel, R. Ossikovski, A. Lizana, C. Rodríguez, E. González-Arnay, and J. Campos, "Depolarization metric spaces for biological tissues classification," J. Biophotonics **13**, e202000083 (2020).
16. I. Ahmad, K. Khan, and M. Ikram, "Polarimetry based Partial Least Square classification of ex vivo healthy and basal cell carcinoma human skin tissues," Photodiagn. Photodyn. Ther. **14**, 134–141 (2016).
17. R. A. Chipman, W. S. T. Lam, and G. Young, *Polarized light and optical systems* (CRC Taylor & Francis Group, 2018).
18. A. S. Zachary, S. H. Derek, K. C. Kengyeh, and A. Wax, "Light-scattering methods for tissue diagnosis," Optica **6**(4), 479–489 (2019).

Biomedical Optics EXPRESS

19. N. Ghosh, M. F. Wood, and I. A. Vitkin, "Influence of the order of the constituent basis matrices of the Mueller matrix decomposition-derived polarization parameters in complex turbid media such as biological tissues," Opt. Commun. **283**(6), 1200–1208 (2010).
20. I. San José and J. J. Gil, "Invariant indices of polarimetric purity: generalized indices of purity for NXN covariance matrices," Opt. Commun. **284**(1), 38–47 (2011).
21. A. Lizana, A. Van Eeckhout, K. Adamczyk, C. Rodríguez, J. C. Escalera, E. Garcia-Caurel, I. Moreno, and J. Campos, "Polarization gating based on Mueller matrices," J. Biomed. Opt. **22**(05), 1 (2017).
22. A. Van Eeckhout, E. Garcia-Caurel, T. Garnatje, J. C. Escalera, M. Durfort, J. Vidal, J. J. Gil, J. Campos, and A. Lizana, "Polarimetric imaging microscopy for advanced inspection of vegetal tissues," Sci. Rep. **11**(1), 3913 (2021).
23. A. Le Gratiet, M. d'Amora, M. Duocastella, R. Marongiu, A. Bendandi, S. Giordani, P. Bianchini, and A. Diaspro, "Zebrafish structural development in Mueller-matrix scanning microscopy," Sci. Rep. **9**(1), 19974 (2019).
24. R. L. Long, M. P. Bange, S. G. Gordon, and G. A. Constable, "Measuring the maturity of developing cotton fibers using an automated polarized light microscopy technique," Text. Res. J. **80**(5), 463–471 (2010).
25. A. Van Eeckhout, E. Garcia-Caurel, T. Garnatje, M. Dufort, J. C. Escalera, J. Vidal, J. J. Gil, J. Campos, and A. Lizana, "Depolarizing metrics for plant simples imaging," PLoS One **14**(3), e0213909 (2019).
26. S. Y. Lu and R. A. Chipman, "Mueller matrices and the degree of polarization," Opt. Commun. **46**(1-6), 11–14 (1998).
27. F. Boulvert, G. L. Brun, B. L. Jeune, J. Cariou, and L. Martin, "Decomposition algorithm of an experimental Mueller matrix," Opt. Commun. **282**(5), 692–704 (2009).
28. P. Babilotte, K. Sathaye, and V. N. Silva, "Retardance properties of twisted ferro-electric liquid crystal," Optik **125**(14), 3646–3652 (2014).
29. J. J. Gil and I. S. José, "3D polarimetric purity," Opt. Commun. **283**(22), 4430–4434 (2010).
30. J. J. Gil, "Structure of polarimetric purity of a Mueller matrix and sources of depolarization," Opt. Commun. **368**, 165–173 (2016).
31. R. Ossikovski, "Analysis of depolarizing Mueller matrices through a symmetric decomposition," J. Opt. Soc. Am. A **26**(5), 1109–1118 (2009).
32. S. Y. Lu and R. A. Chipman, "Interpretation of Mueller matrices based on polar decomposition," J. Opt. Soc. Am. A **13**(5), 1106–1113 (1996).
33. A. Tariq, H. He, P. Li, and H. Ma, "Purity-depolarization relations and the components of purity of a Mueller matrix," Opt. Express **27**(16), 22645–22662 (2019).
34. R. Ossikovski and J. Vizet, "Eigenvalue-based depolarization metric spaces for Mueller matrices," J. Opt. Soc. Am. A **36**(7), 1173–1186 (2019).
35. C. J. R. Sheppard, A. Bendandi, A. le Gratiet, and A. Diaspro, "Eigenvectors of polarization coherency matrices," J. Opt. Soc. Am. A **37**(7), 1143–1154 (2020).
36. M. Borovkova, M. Peyvasteh, O. Dubolazov, Y. Ushenko, A. Bykov, S. Deby, J. Rehbinder, T. Novikova, and I. Meglinski, "Complementary analysis of Mueller-matrix images of optically anisotropic highly scattering biological tissues," J. Eur. Opt. Soc.-Rapid Publ. **14**(20), s41476-018-0085-9 (2018).
37. M. Borovkova, A. Bykov, A. Popov, A. Pierangelo, T. Novikova, J. Pahnke, and I. Meglinski, "Evaluating β-amyloidosis progression in Alzheimer's disease with Mueller polarimetry," Biomed. Opt. Express **11**(8), 4509–4519 (2020).
38. V. A. Ushenko, B. T. Hogan, A. Dubolazov, A. V. Grechina, T. V. Boronikhina, M. Gorsky, A. G. Ushenko, Y. O. Ushenko, A. Bykov, and I. Meglinski, "Embossed topographic depolarization maps of biological tissues with different morphological structures," Sci. Rep. **11**(3871), s41598-021-83017-2 (2021).
39. V. A. Ushenko, B. T. Hogan, A. Dubolazov, G. Piavchenko, S. L. Kuznetsov, A. G. Ushenko, Y. O. Ushenko, M. Gorsky, A. Bykov, and I. Meglinski, "3D Mueller matrix mapping of layered distributions of depolarization degree for analysis of prostate adenoma and carcinoma diffuse tissues," Sci. Rep. **11**(5162), s41598-021-83986-4 (2021).
40. J. J. Gil, "Invariant quantities of a Mueller matrix under rotation and retarder transformations," J. Opt. Soc. Am. A **33**(1), 52–58 (2016).
41. K. Dev and A. Asundi, "Mueller-Stokes polarimetric characterization of transmissive liquid crystal spatial light modulator," Optics and Lasers in Engineering **50**(4), 599–607 (2012).
42. A. Tariq, P. Li, D. Chen, D. Lv, and H. Ma, "Physically realizable space for the purity-depolarization plane for polarized light scattering media," Phys. Rev. Lett. **119**(3), 033202 (2017).
43. F. H. Mustafa and M. S. Jaafar, "Comparison of wavelength-dependent penetration depths of lasers in different types of skin in photodynamic therapy," Indian J. Phys. **87**(3), 203–209 (2013).
44. B. Charvet, F. Ruggiero, and D. le Guellec, "The development of the myotendinous junction: a review," Musc. Liga. Tend. J. **2**(2), 53–63 (2012).
45. D. L. Butler, E. S. Grood, F. R. Noyes, and R. F. Zernucke, "Biomechanics of ligaments and tendons," Exer. Sports Sci. Rev. **6**, 125–181 (1978).
46. R. James, G. Kesturu, G. Balian, and A. B. Chhabra, "Tendon: biology, biomechanics, repair, growth factors, and evolving treatment options," The Journal of Hand Surgery **33**(1), 102–112 (2008).
47. D. J. Hulmes, "Building collagen molecules, fibrils and suprafibrillar structures," J. Struct. Biol. **137**(1-2), 2–10 (2002).
48. S. Schiaffino and C. Reggiani, "Fiber types in mammalian skeletal muscles," Physiol. Rev. **91**(4), 1447–1531 (2011).

Biomedical Optics EXPRESS

49. H. Asahara, M. Inui, and M. K. Lotz, "Tendons and ligaments: connecting developmental biology to musculoskeletal disease pathogenesis," J. Bone Miner. Res. **32**(9), 1773–1782 (2017).
50. J. E. V. Ferreira, M. T. S. Pinheiro, W. R. S. D. Santos, and R. D. S. Maia, "Graphical representation of chemical periodicity of main elements through Boxplot," Educación Química **27**(3), 209–216 (2016).
51. R. McGill, J. W. Tuckey, and W. A. Larsen, "Variations of Box Plots," The American Statistician **32**(1), 12–16 (1978).
52. Y. L. Katchanov, Y. V. Markova, and N. A. Shmatko, "The distinction machine: physics journals from the perspective of the Kolmogorov–Smirnov statistic," Journal of Informetrics **13**(4), 100982 (2019).
53. Z. Drezner and O. Turel, "Normalizing variables with too-frequent values using a Kolmogorov–Smirnov test: a practical approach," Computers and Industrial Engineering **61**(4), 1240–1244 (2011).
54. N. Razali and Y. B. Wah, "Power comparisons of Shapiro–Wilk, Kolmogorov–Smirnov, Lilliefors and Anderson–Darling tests," Journal of Statistical Modeling and Analytics **2**(1), 21–33 (2011).
55. M. B. Wilk and R. Gnanadesikan, "Probability plotting methods for the analysis of data," Biometrika **55**(1), 1–17 (1968).
56. J. W. Tukey, *Exploratory Data Analysis* (Addison-Wesley, 1977).
57. S. Guo, S. Zhong, and A. Zhang, "Privacy-preserving Kruskal–Wallis test," Computer Methods and Programs in Biomedicine **112**(1), 135–145 (2013).
58. G. D. Ruxton and G. Beauchamp, "Some suggestions about appropriate use of the Kruskal–Wallis test," Animal Behaviour **76**(3), 1083–1087 (2008).
59. V. Srinivasarao and U. Ghanekar, "Speech enhancement - an enhanced principal component analysis (EPCA) filter approach," Computers and Electrical Engineering **85**, 106657 (2020).
60. P. Cai and X. Deng, "Incipient fault detection for nonlinear processes based on dynamic multi-block probability related kernel principal component analysis," ISA Trans. **105**, 210–220 (2020).
61. S. A. Ghorashi, F. Honarvar, and M. Tabatabaeipour, "Automated extraction of local defect resonance using the principal component analysis in lock-in ultrasonic vibrothermography," Infrared Physics and Technology **105**, 103204 (2020).
62. F. Reverter, E. Vegas, and P. Sánchez, "Mining gene expression profiles: an integrated implementation of kernel principal component analysis and singular value decomposition," Genomics, Proteomics and Bioinformatics **8**(3), 200–210 (2010).
63. R. V. Ramirez-Velarde, M. Roderus, C. Barba-Jimenez, and R. Perez-Casares, "A parallel implementation of singular value decomposition for video-on-demand services using principal components analysis," Procedia Computer Science **29**, 1876–1887 (2014).
64. W. Zhou, H. Liu, Q. Xu, P. Li, L. Zhao, and H. Gao, "Glycerols generalized two-dimensional correlation IR/NIR spectroscopy and its principal component analysis," Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy **228**, 117824 (2020).
65. B. Olasege, S. Zhang, Q. Zhao, D. Liu, H. Sun, Q. Wang, P. Ma, and Y. Pan, "Genetic parameter estimates for body conformation traits using composite index, principal component, and factor analysis," J. Dairy Sci. **102**(6), 5219–5229 (2019).
66. H. H. Jolliff, "Principal component factor analysis," *All Graduate Plan B and other Reports*, 1117 (1968).
67. C. C. Clark, C. M. Barnes, M. J. Duncan, H. D. Summers, and G. Stratton, "Physical activity, motor competence and movement and gait quality: a principal component analysis," Human Movement Science **68**, 102523 (2019).
68. B. Mavule, V. Muchenje, C. Bezuidenhout, and N. Kunene, "Morphological structure of Zulu sheep based on principal component analysis of body measurements," Small Ruminant Research **111**(1-3), 23–30 (2013).
69. S. Gaijar, M. Kulahci, and A. Palazoglu, "Selection of non-zero loadings in sparse principal components analysis," Chemom. Intell. Lab. Syst. **162**, 160–171 (2017).
70. Principal Components Analysis. SPSS Annotated Output, UCLA: Statistical Consulting Group. From https://stats.idre.ucla.edu/spss/output/principal_components/ (accessed Jun 1, 2020).
71. M. Zhu and A. Ghodsi, "Automatic dimensionality selection from the Scree plot via the use of profile likekihood," Computational Statistics and Data Analysis **51**(2), 918–930 (2006).
72. Factor Analysis Extraction. IBM Knowledge Center, From https://www.ibm.com/support/knowledgecenter/en/SSLVMB_sub/statistics_mainhelp_ddita/spss/base/idh_fact_ext.html (accessed May 31, 2020).
73. S. de la and F. Fernandez, "Componentes Principales," (Facultad de Ciencias Económicas y Empresariales, Universidad Autónoma de Madrid, 2011).
74. S. de la and F. Fernandez, "Regresión Logística," (Facultad de Ciencias Económicas y Empresariales, Universidad Autónoma de Madrid, 2011).
75. A. Jokiel-Rokita and R. Topolnicki, "Estimation of the ROC curve from the Lehmann Family," Computational Statistics and Data Analysis **142**, 106820 (2020).
76. T. Fawcett, "An introduction to ROC analysis," Pattern Recognition Letters **27**(8), 861–874 (2006).
77. J. A. Cook, "ROC curves and nonrandom data," Pattern Recognition Letters **85**, 35–41 (2017).
78. M. D. Figueiredo, C. B. Cordella, D. J. R. Bouveresse, X. Archer, J. M. Bégué, and D. N. Rutledge, "A variable selection method for multiclass classification problems using two-class ROC analysis," Chemom. Intell. Lab. Syst. **177**, 35–46 (2018).
79. Y. Tahtali, "Use of Factor Scores in Multiple Regression analysis for estimation of body weight by certain body measurements in Romanov lambs," PeerJ **7**, e7434 (2019).

Biomedical Optics EXPRESS

80. B. D. Hill, "The sequential Kaiser-Meyer-Olkin procedure as an alternative for determining the number of factors in common-factor analysis: a Monte Carlo simulation," Dissertation (Oklahoma State University, 2011).
81. K. Charry, K. Coussement, N. Demoulin, and N. Heuvinck, "Getting started with IBM SPSS statistics," Marketing Research with IBM® SPSS Statistics **1**, 1–30 (2016).
82. N. J. D. Nagelkerke, "A note on the general definition of the coefficient of determination," Biometrika **78**(3), 691–692 (1991).
83. W. J. Youden, "Index for rating diagnostic tests," Cancer **3**(1), 32–35 (1950).

Polarimetric data-based model for tissue recognition: supplemental document

1. Complete image Mueller polarimeter description

The current study is performed by measuring the experimental Mueller matrices from different chicken samples at three different illumination wavelengths covering the visible range: 625 nm, 530 nm and 470 nm. To do so, a complete image Mueller polarimeter [1-4] is used (Fig. S1) Particularly, it is based on Parallel Aligned Liquid Crystals (PA-LC) retarders and consists of two compact and mobile arms. The first one, the Polarization State Generator (PSG), allows us to generate any fully polarized state so we can illuminate the sample with controlled polarized light. It is composed by a linear polarizer oriented at 0º followed by two PA-LC at 45º and 0º, respectively, with respect to the laboratory vertical. We illuminate the sample at 56º with respect to the vertical so the second arm, called Polarization State Analyzer (PSA) and placed in the vertical position avoids direct reflections and collects the scattered light. PSA is composed of the same elements as PSG but placed in the inverse order to finally capture, by means of a CCD camera, the sample's intensity image. With this architecture, the polarimeter is capable to analyze and determine the polarimetric response of the sample.

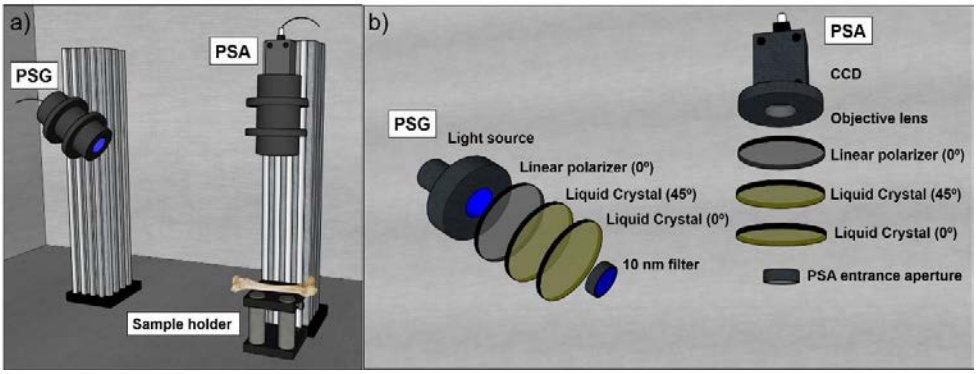


Fig. S1. a) 3D representation of the complete image Mueller polarimeter used in this study, b) 3D representation of the arrangement of the optical components in the PSG and the PSA.

In the employed set-up, the illumination is performed by using a four-wavelength high-power Thorlabs LED source (LED4D211, operated by DC4104 drivers distributed by Thorlabs), complemented with 10 nm dielectric bandwidth filters for green (530 nm) and blue (470 nm) wavelengths (Thorlabs FB530-10 and FB470-10, respectively). In counterpart, imaging is performed by means of a TECHSPEC® high-resolution objective (distributed by Edmund Optics) with a 35 mm focal length so a resolution of 22 μm is achieved followed by an Allied Vision Manta G-504B CCD camera.

To build the experimental Mueller matrix, at least 16 independent measurements are required (related with different Stokes generation, PSG, and detection, PSA, configurations). In the current study, we use 6 illumination states of polarization (generators) and 6 detection analyzers instead, so taking 36 images, to minimize the measurement noise of each sample's Mueller matrix. A complete Mueller measurement waiting time lasts ~ 3.6 seconds: liquid crystal molecules rearrange the orientation process from one polarization state to the following one lasts approx. 100 ms.

2. Histological process of samples

For histological countercheck of the images, samples were obtained from regions analogous to the ones submitted to polarimetric analysis (i.e., striated muscle, tendon, myotendinous junction and diaphyseal bone). After extraction, they were immersion-fixed in 4% formaldehyde overnight and dehydrated through a protocol including immersion in: 70% Ethanol (30'), 80% Ethanol (30'), 90% ethanol (30'), 95% ethanol (30'), two successive immersions in 100% Ethanol (60' each) and xylene (60' each). After dehydration, they were embedded and in a paraffin wax. After cooling they were sliced in 5µm-thick sections through microtomy. Sections were mounted in crystal slides which were stained after deparaffination. Deparaffination protocol includes a 24-hour immersion in xylene and brief, successive rinses in 96% ethanol, 90% ethanol, 80% ethanol, 70% ethanol and distilled water. Two different histological staining-techniques were used.

a) Hematoxylin-eosin is the standard histological technique for diagnostic microscopy. It involves the combined use of an oxidized cationic colorant (hematoxylin) that highlights nucleic acid-rich regions of the tissue in purple and an anionic colorant (eosin) that stains in a characteristic pink color protein-rich area of the tissue, e.g. collagen in pale pink and muscle in dark pink. Sections were immersed in a bath containing Harris hematoxylin solution for 10' (Merck ® H9627), differentiated in 1% acid alcohol solution (5 ml of 37% HCl in 495 ml of 70% ethanol) for around 20", blued in a bluing solution for 20" (Sigma-Aldric ® S5134) and counterstained with Eosin Y for 1' (Sigma-Aldric ® E4009). Brief rinses in tap water were performed between each step.

b) Masson technique allows the differential staining of connective tissue. It includes Weigert's hematoxylin (a variant of hematoxylin) for nuclei acid-rich regions, acid fuchsin as an acid dye for protein-rich cytoplasm, a combination of phosphotungstic and phosphomolybdic acids to remove the excess fuchsin from fibers and aniline blue to counterstain the fibers. Overall, it provides a blueish color to collagen-rich tissue and deep red color to muscle fibers, also allowing the identification of other kinds of tissue. Staining protocol includes a 5' rinse in Weigert Hematoxylin (Merck® 115973), a 5' incubation in acid fuchsin solution (Sigma-Aldrich ® F8219), successive incubations (no longer than 3') in a 20% solution of phosphotungstic-phosphomolybdic acid in ethanol (Sigma-Aldrich ® 319279) and a 5' incubation in a 2.5% solution of aniline blue prepared in 2% acetic acid (Sigma-Aldric ® B8563). If necessary, a rapid rinse in 1% acetic was performed to remove the excess dye. Brief rinses in distilled water were performed between steps.

After staining, samples were dehydrated in ethanol (5' in ascending concentrations of 70%, 80%, 90%, 96%, 100%) and xylene before being covered using Eukitt® as mounting media. After drying for 24 hours, sections were scanned (Zeiss AxioScan ®) and analyzed using a QuPath free software. Obtained images are presented in Figs. S2 and S3.

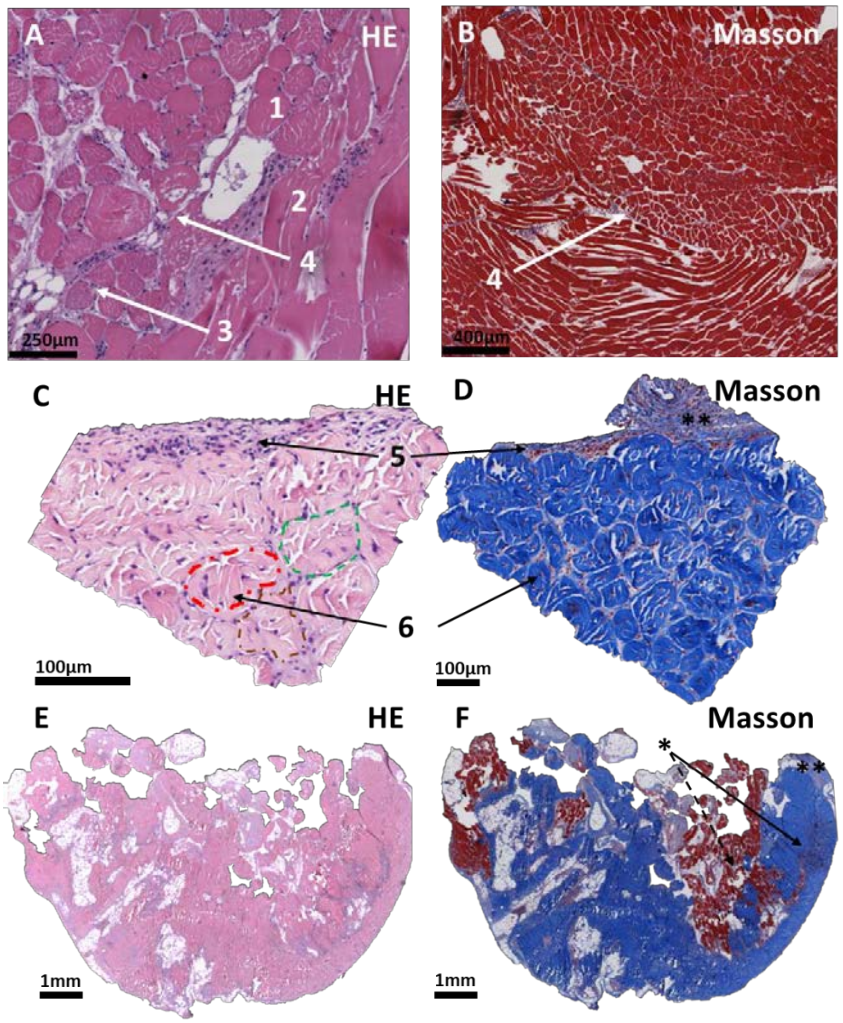


Fig. S2: Histochemical stainings of muscle (A, B), tendon (C, D) and myotendinous junction (E, F). Bundles (1, 2, 6) of either contractile (A, B, E, F) or dense fibrous (collagen) tissue (C-F) are surrounded by sheets of lax connective tissue (collagen) concentrically organized as epimysium (**)/epitenon (**), perimysium/peritenon (5) and endomysium/endotenon (6). In E, F a transverse section of the myotendinous junction is shown, where the Masson staining reveals the intermixed (*) fascicles of contractile fibers (discontinuous arrow in (F) and collagenous fibers (arrow in F)). Different fascicles of tendinous collagen are highlighted in red, green and gold (C).

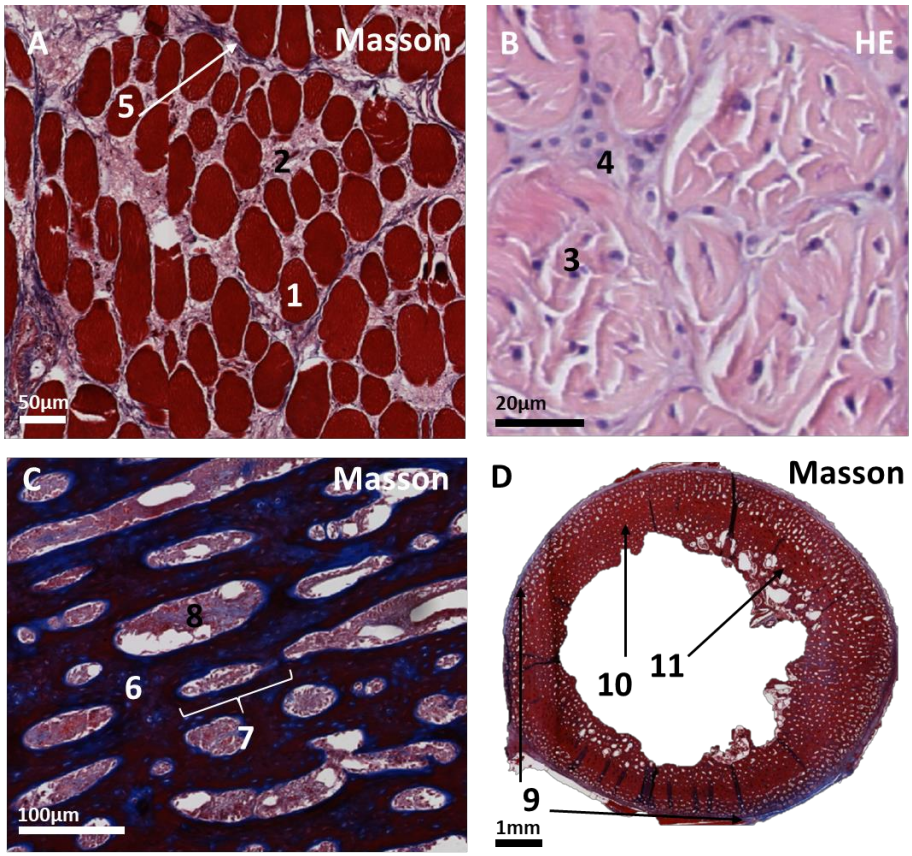


Fig. S3: Transversal section of striated muscle (A) showing bundles of contractile fibers (red colored, 1) surrounded by endomysium (2, compare to 3, 4) and included in a single fascicle surrounded by collagen (perimysium, 5). This structure is analogous to the tendinous structure (B), where the collagen fibers (3) are densely packed into a sheet of peritenon (4). Longitudinal section of cortical bone (C), showing bony matrix (6) and lacunae (7) containing osteocytes, vessels and collagen that are heavily artifacted due to pre-treatment (8); Transverse section through the diaphysis of a long bone (D), where collagenic periosteum (9), cortical bone (10) and rests of trabecular bone (11) are arranged in a concentrically layers. Note the regular arrangement of lacunae, both in longitudinal and transverse sections. Color (from reddish to purple) depends on the amount of mineral deposits in any given region of the bone.

3. Experimental Mueller matrices

The experimental typical Mueller matrices (MMs) obtained for each type of tissue (muscle, tendon, myotendinous junction and bone) measured at 625 nm illumination wavelength are shown below followed by the respective parameters (P_Δ, P_1, P_2, P_3, P, R, D, δ and Ψ) resulting from matrix decomposition. The obtained images correspond to a region of interest (ROI) of 512 x 512 pixels which corresponds to an area of 1.1 x 1.1 cm^2.

Particularly, typical MM of muscle tissue and the respective polarimetric parameters are shown in Figs. S4 and S5, respectively. MMs of tendon, myotendinous junction and bone tissue are presented in Figs. S6 and S7, Figs. S8 and S9, and Figs. S10 and S11, respectively.

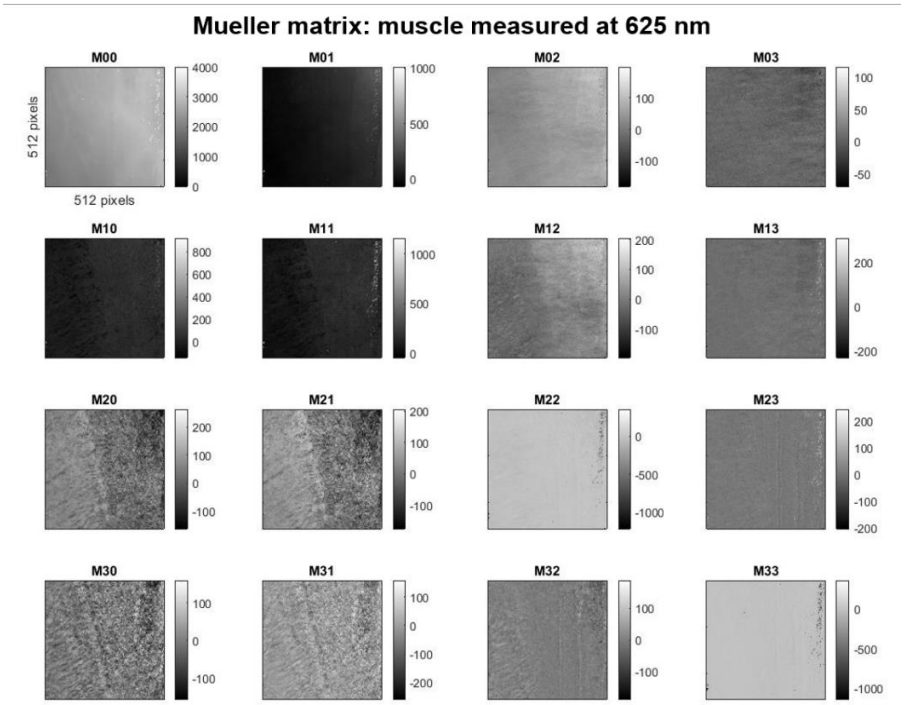


Fig. S4. Experimental Mueller matrix image of a sample of muscle tissue.

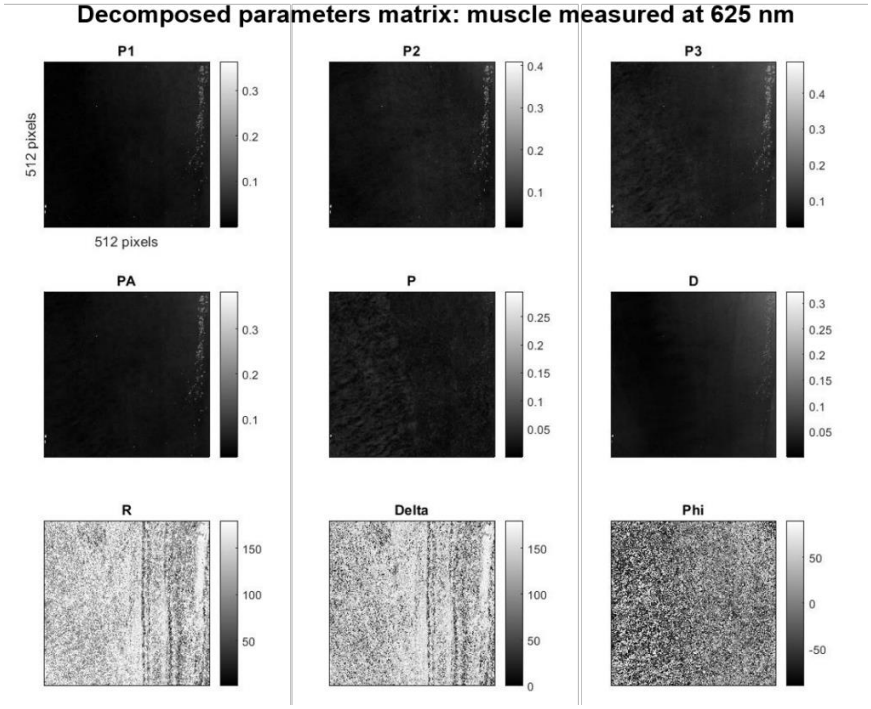


Fig. S5. Images of the polarimetric parameters P_Δ, P_1, P_2, P_3, P, R, D, δ and Ψ from a sample of muscle tissue.

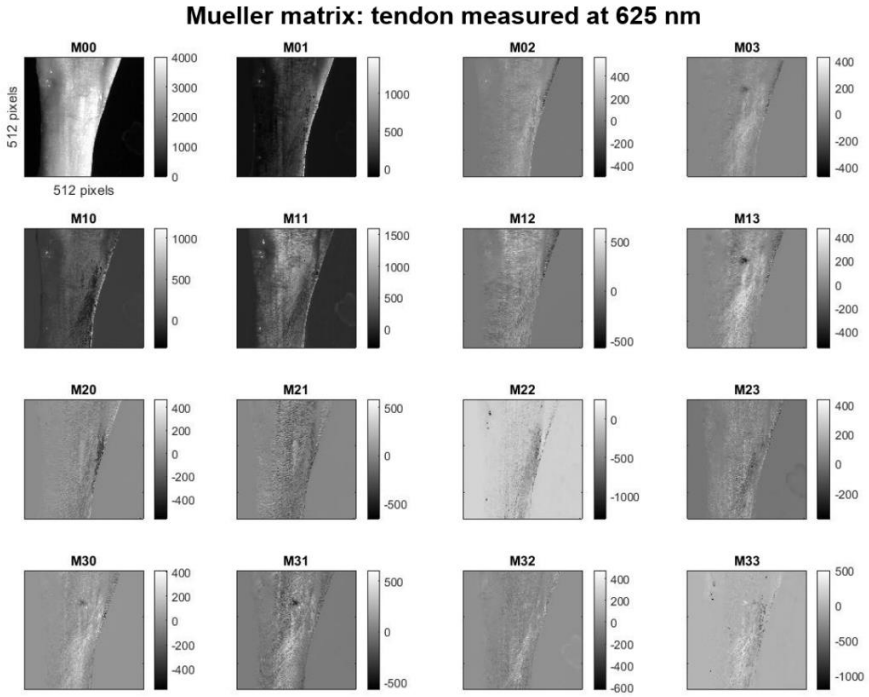


Fig. S6. Experimental Mueller matrix image of a sample of tendon tissue.

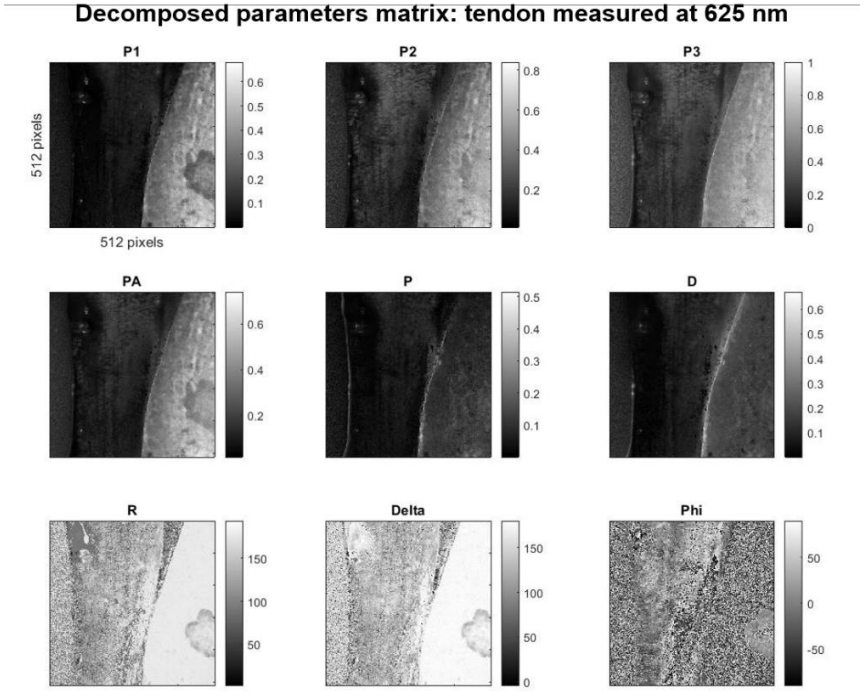


Fig. S7. Images of the polarimetric parameters P_Δ, P_1, P_2, P_3, P, R, D, δ and Ψ from a sample of tendon tissue.

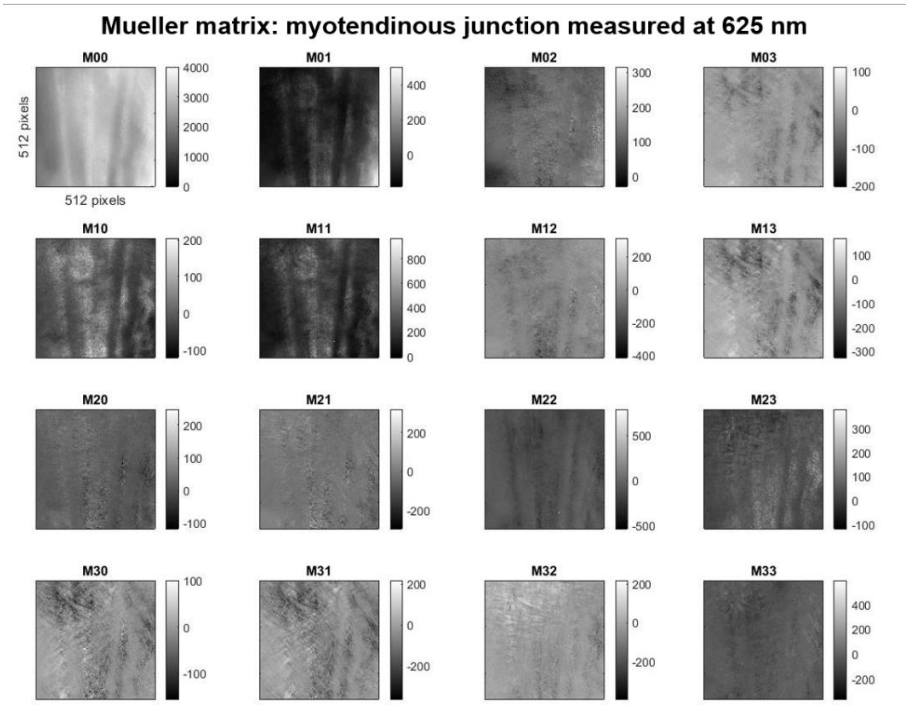


Fig. S8. Experimental Mueller matrix image of a sample of myotendinous tissue.

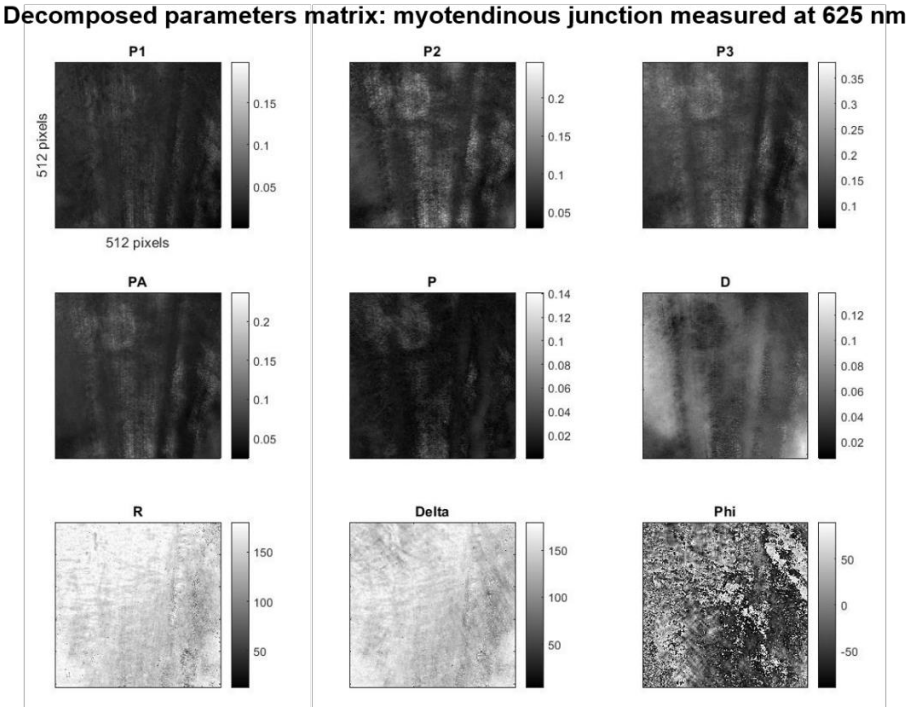


Fig. S9. Images of the polarimetric parameters P_Δ, P_1, P_2, P_3, P, R, D, δ and Ψ from a sample of myotendinous tissue.

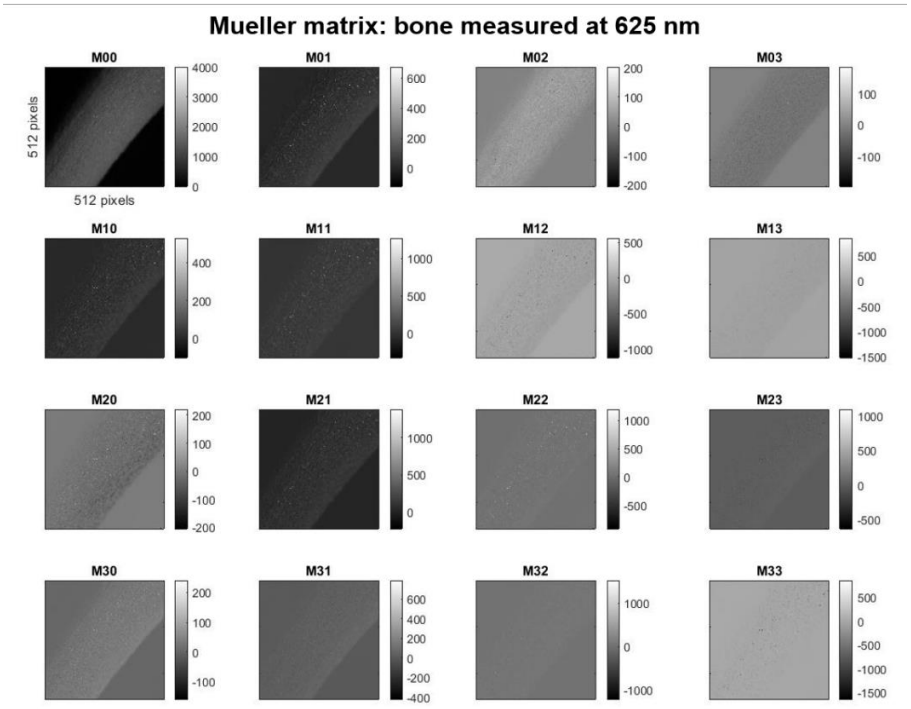


Fig. S10. Experimental Mueller matrix image of a sample of bone tissue.

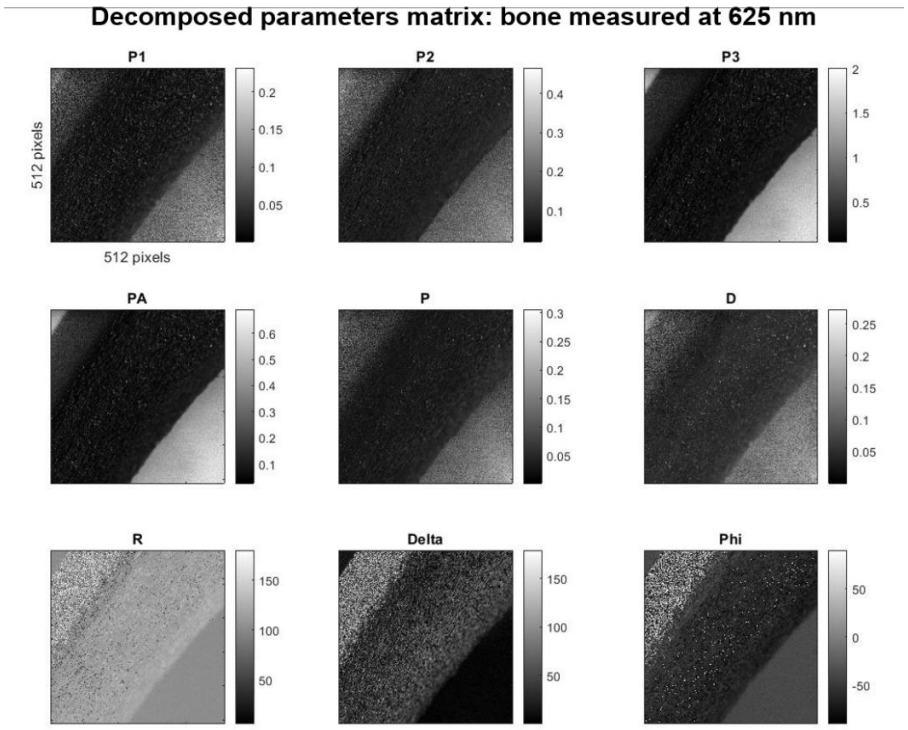


Fig. S11. Images of the polarimetric parameters P_Δ, P_1, P_2, P_3, P, R, D, δ and Ψ from a sample of bone tissue.

4. Statistical analysis

For the correct manipulation of data to construct a predictive model it is mandatory to know which are the distributions profile (parametric or not) and, depending on the results, apply the corresponding statistical tests to formulate the tissue-discrimination functions. In this way, we perform an exploratory univariant analysis (to analyze the distribution profile) and a posterior factor analysis with principal components extraction for the dichotomic predictive models' construction. The additional material for the above-mentioned analysis is described in the following sections.

4.1 Exploratory univariant analysis

To ensure the correct statistical manipulation of polarimetric data distributions classified by type-of-tissue variable groups, it is necessary to previously perform an exploratory univariant analysis based on the study of the graphical differences between sample's distributions, as well as the central tendency, data dispersion, asymmetry and outlier's identification by means of Boxplot [5,6].

Recalling the P_2 purity index (measured at 625 nm illumination) presented in Fig. 2, in this section we provide, in complement, the results in the format of Boxplots corresponding to the data acquired at 530 nm and 470 nm in Figs. S12 and S13, respectively.

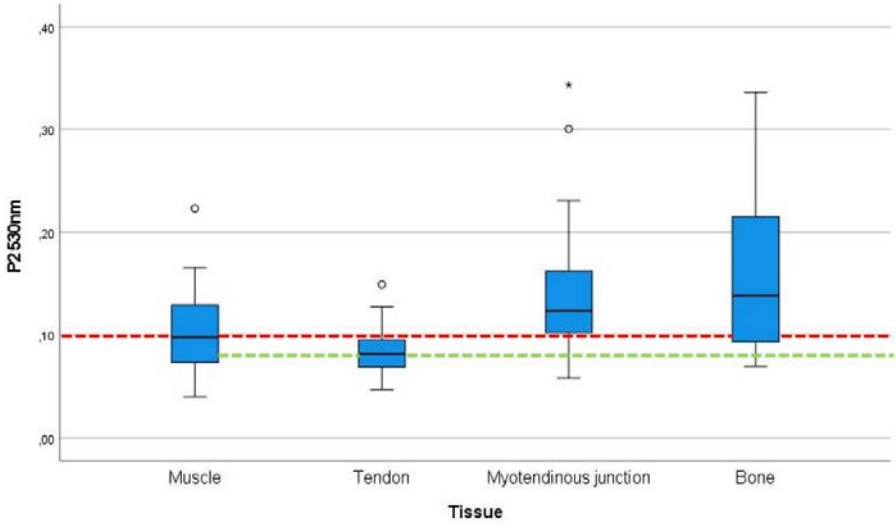


Fig. S12. Boxplot of P_2 index for all tissues measured at 470 nm illumination channel. Red-dashed line visually represents the potential of the metric to discriminate muscle among remaining tissue types: the median of the muscle box does not fit within the other tissues' boxes. Green-dashed line visually represents P_2 tendon discrimination capability. Boxplot points out the low quantity of outliers on data distributions (which can be extrapolated to the remaining metrics): mild and extreme values are represented by circles and stars, respectively.

The exposed P_2 data distributions for all type of tissues measured at 530 nm wavelength (Fig. S12) clearly shows the same small number of outliers' behavior as 625 nm (manuscript Fig. 2) and 470 nm (Fig. S13). What is more, asymmetric tendency (position of the median is not in the middle of boxes) is common for all wavelengths. Moreover, tissue discrimination capability of P_2 is pointed out (median line does not fit within the remaining boxes): at 530 nm illumination (Fig. S12), the metric shows muscle discriminative potential (see dashed red line) but also tendon differentiation from myotendinous junction and bone (see dashed green line). At 470 nm

illumination conditions (Fig. S13), P_2 keeps differentiating muscle (see dashed red line) but also works well as a myotendinous junction discriminator (see dashed blue line).

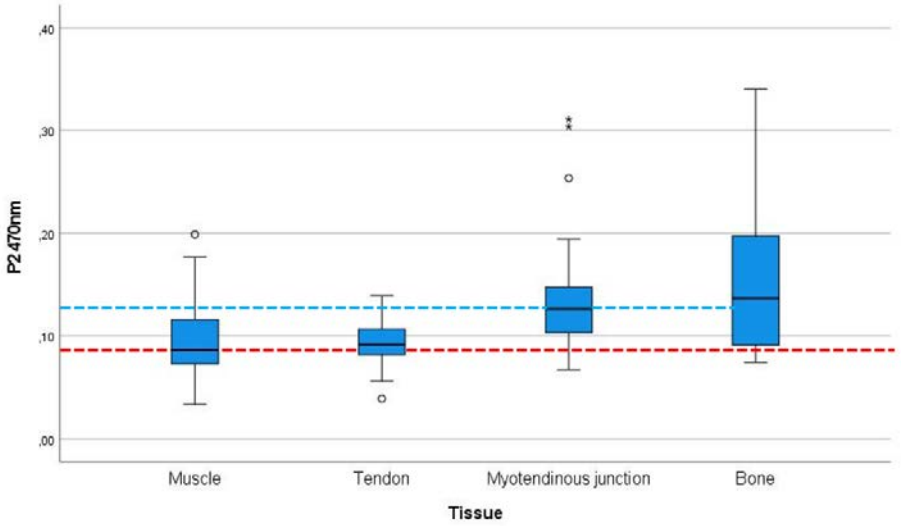


Fig. S13. Boxplot of P_2 index for all tissues measured at 470 nm illumination channel. Red-dashed line visually represents the potential of the metric to discriminate muscle among remaining tissue types: the median of the muscle box does not fit within the other tissues' boxes. Blue-dashed line visually represents P_2 Myotendinous junction discrimination capability. Boxplot points out the low quantity of outliers on data distributions (which can be extrapolated to the remaining metrics): mild and extreme values are represented by circles and stars, respectively.

For the Kruskal-Wallis [7,8] homogeneity multiple contrast analysis, the output significance of the 27 measured polarimetric indicators for each pair of tissue discriminative power is presented in Table S1. Particularly, we assume, as null hypothesis, that the data comes from samples with the same statistical distribution. By setting a significance level, α, of 0.05, Kruskal-Wallis outcomes reject, in most of the cases, the mentioned hypothesis: p-values lower than the significance level show the discriminatory potential of the polarimetric indicator between the corresponding pair of tissues. Note that, because the SPSS software found no significant difference when sorting by pair of tissues, only one value is output for optical rotation, Ψ.

Table S1. Significance (p-value, rounded to three digits) of the M-metrics pair-of-tissue discrimination.

		Muscle – Tendon	***Muscle – Myo.***	***Muscle – Bone***	***Tendon – Myo.***	***Tendon – Bone***	***Myo. – Bone***
625 nm	P_Δ	0.000	0.000	0.000	0.559	0.331	0.599
	P_1	0.000	0.000	0.000	0.210	0.274	0.948
	P_2	0.003	0.000	0.000	0.741	0.181	0.254
	P_3	0.000	0.000	0.000	0.723	0.714	0.936
	P	0.427	0.000	0.000	0.036	0.000	0.004
	R	0.378	0.000	0.000	0.002	0.000	0.223

	D	0.070	0.002	0.000	0.149	0.008	0.135
	δ	0.374	0.005	0.000	0.062	0.004	0.614
	Ψ	0.110					
530 nm	P_Δ	0.351	0.000	0.000	0.000	0.000	0.730
	P_1	0.594	0.000	0.000	0.000	0.000	0.389
	P_2	0.044	0.000	0.000	0.000	0.003	0.982
	P_3	0.995	0.000	0.000	0.000	0.000	0.268
	P	0.054	0.001	0.000	0.062	0.001	0.436
	R	0.001	0.000	0.000	0.358	0.210	0.615
	D	0.695	0.017	0.010	0.060	0.031	0.534
	δ	0.008	0.000	0.000	0.283	0.042	0.247
	Ψ	0.101					
470 nm	P_Δ	0.480	0.000	0.000	0.000	0.000	0.877
	P_1	0.253	0.000	0.000	0.000	0.000	0.436
	P_2	0.930	0.000	0.000	0.000	0.000	0.885
	P_3	0.340	0.000	0.000	0.000	0.000	0.508
	P	0.010	0.020	0.000	0.835	0.001	0.015
	R	0.024	0.002	0.002	0.358	0.211	0.616
	D	0.383	0.021	0.000	0.181	0.005	0.079
	δ	0.009	0.004	0.009	0.749	0.634	0.830
	Ψ	0.299					

4.2 Principal Components extraction and component score coefficient matrix

Linked with scree plot [9], Table S2 exposes the eigenvalues of the chosen 10 principal components [10,11], together with the data variance and the cumulative variance (both in %). Note that C_1 carries the highest amount of explained variance. This behavior will be reflected in predictive model construction as this component is going to have a strong influence. Lower amounts of variance are taken by the remaining components. However, it is important to remark that when considering the 10 components as a whole, we achieve a 92% of explained variance.

Table S2. Percentage of variance explained.

	Extraction Sums of Squared Loadings		
	Total	***% of Variance***	***Cumulative***
C_1	11.195	41.462	41.462
C_2	3.783	14.012	55.473
C_3	2.265	8.391	63.864
C_4	2.008	7.436	71.300

C_5	1.786	6.616	77.915
C_6	1.026	3.799	81.714
C_7	0.890	3.298	85.012
C_8	0.762	2.824	87.836
C_9	0.634	2.350	90.186
C_{10}	0.509	1.885	92.071

The 27-dimension metrics space is reduced to a new 10-dimension principal components space. In this way, the original polarimetric information is encoded on the principal components: we can write each component as a linear combination of the 27 different polarimetric indicators weighed by some constants provided by columns of the so-called component score coefficient matrix, shown in Table S3.

Table S3. Component score coefficient matrix of the 10 first principal components.

		Principal component									
		C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8	C_9	C_{10}
625 nm	P_Δ	0.068	-0.083	0.219	-0.061	0.113	0.005	-0.079	-0.050	-0.066	-0.204
	P_1	0.066	-0.020	0.154	-0.126	0.205	-0.024	-0.128	0.090	-0.129	-0.178
	P_2	0.063	-0.103	0.214	-0.017	0.053	0.013	-0.015	-0.162	-0.035	-0.241
	P_3	0.063	-0.090	0.230	-0.060	0.102	0.022	-0.088	-0.026	-0.053	-0.146
	P	0.013	-0.092	0.097	-0.205	0.168	0.441	0.442	0.329	0.252	0.217
	R	0.036	0.154	0.138	-0.032	0.117	-0.329	-0.074	-0.018	0.484	0.636
	D	0.028	-0.107	0.218	0.232	-0.213	-0.082	0.032	-0.252	0.099	-0.077
	δ	0.034	0.160	0.139	-0.098	0.018	-0.114	-0.014	0.402	-0.505	0.532
	Ψ	-0.011	0.037	0.008	0.280	0.196	-0.313	0.434	0.472	0.120	-0.622
530 nm	P_Δ	0.085	-0.003	-0.105	0.008	0.017	-0.096	-0.044	-0.076	-0.058	0.018
	P_1	0.082	0.027	-0.069	0.005	0.045	-0.127	0.020	-0.029	-0.011	-0.044
	P_2	0.082	-0.021	-0.095	0.019	-0.023	-0.077	-0.072	-0.121	0.066	-0.002
	P_3	0.078	-0.009	-0.132	-0.010	0.035	-0.085	-0.062	-0.094	-0.201	0.083
	P	0.068	0.008	-0.094	-0.016	0.066	-0.004	0.434	-0.315	-0.235	0.111
	R	0.038	0.180	0.105	0.037	-0.038	-0.154	0.250	-0.400	0.294	-0.022

	D	0.043	-0.118	0.089	0.196	-0.245	0.034	0.036	0.239	0.073	0.484
	δ	0.029	0.179	0.111	0.038	-0.156	0.038	0.294	-0.065	-0.553	-0.164
	Ψ	-0.001	0.060	-0.005	0.346	0.298	0.145	-0.064	0.098	0.008	0.251
470 nm	P_4	0.083	0.008	-0.127	0.005	-0.011	0.027	-0.121	0.169	0.079	-0.117
	P_1	0.080	0.035	-0.101	0.005	0.005	0.007	-0.143	0.287	0.111	-0.099
	P_2	0.081	-0.010	-0.136	0.007	-0.033	0.051	-0.106	0.082	0.097	-0.152
	P_3	0.080	0.000	-0.126	-0.002	-0.003	0.017	-0.113	0.124	0.030	-0.065
	P	0.057	-0.027	-0.143	-0.026	0.000	0.244	0.459	-0.190	0.108	0.216
	R	0.033	0.181	0.084	-0.034	-0.119	0.283	-0.083	0.023	0.616	-0.241
	D	0.049	-0.122	0.036	0.184	-0.239	0.002	0.095	0.258	0.005	0.254
	δ	0.025	0.179	0.062	0.012	-0.233	0.356	-0.087	0.196	-0.149	-0.308
	Ψ	0.006	0.054	0.017	0.303	0.236	0.444	-0.235	-0.285	-0.176	0.151

5. Predictive model construction

To test the goodness-of-fit of the four constructed predictive models (muscle, tendon, myotendinous junction and bone), we rely on the Hosmer-Lemeshow [12] significance and the R^2 of Nagelkerke [13] measure. Outputs are presented in Table S4.

Table S4. Hosmer-Lemeshow significance (p-value) and R^2 of Nagelkerke, of each predictive model.

	Muscle	**Tendon**	**Myotendinous junction**	**Bone**
Hosmer-Lemeshow Sig.	0.192	0.980	0.530	0.956
R^2 of Nagelkerke	0.640	0.690	0.282	0.481

Because the obtained Hosmer-Lemeshow [12] significance values are larger than the threshold (p-value > 0.05), we accept the null hypothesis: all the four predictive models fit the data. Regarding to the of R^2 of Nagelkerke [13] values, the larger the value the better the fit, being 1 the maxim value (optimal fit). However, generally speaking, R^2 of Nagelkerke values larger than 0.2 are acceptable, and larger than 0.6 can be associated to excellent fits. In this vein, obtained results show a nice data fit for the muscle (0.64), tendon (0.69) and bone (0.48) regressions. In turn, the myotendinous junction data regression lead to an acceptable result (0.28), but clearly worst than those obtained for the other tissues. This result is in agreement with the models analysis provided in the main manuscript, where the worst results in terms of sensitivity and specificity are obtained for the myotendinous junction model.

6. Tissue recognition through binary logistic model

For a complementary visual analysis, the output of the probabilistic model for the particular cases of an arbitrary bone and an arbitrary myotendinous junction tissue are provided in Figs.

S14 and S15, respectively. Particularly, Fig. S14 shows the intensity image M00 (Fig. S14(a)) and the four probability function images (muscle, tendon, myotendinous junction and bone recognition (Figs. S14(b-e), respectively) for a bone sample. Recall that sample holder polarimetric information is not characterized: pixel values for background acquired data have no physical meaning. The obtained probability function images show how the bone model successfully recognize the bone tissue pixels (Fig. S14(e)), clearly showing, in addition, the two longitudinal bone edges (marked with yellow and red dotted lines) not visible in intensity image in Fig. S14(a). When analyzing the probability distribution of the three remaining functions, a good discriminating potential is demonstrated, specially for the tendon model in Fig. S14(c), which clearly does not recognize the tissue as tendon. On the other hand, muscle and myotendinous junction probability images (Figs. S14(b) and S14(d)) show, with low probability values, an overall discard of analyzed tissue to be muscle or myotendinous junction tissues, respectively. As in previous discussion, worst predictive results are obtained for the myotendinous junction model (Fig. S14(d)).

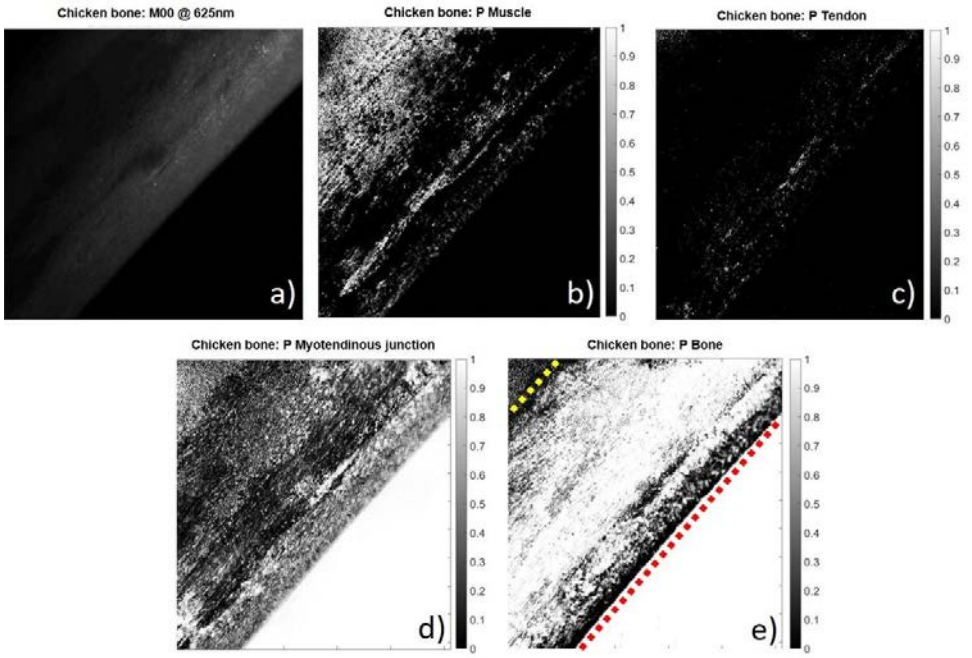


Fig. S14. Intensity image M00 (a) and probability function of muscle (b), tendon (c), myotendinous junction (d) and bone (e) for chicken bone measurements. The gray level bars, placed to the right of the corresponding probability function images, defines whether the probability of the pixel to be recognized as a particular tissue is one (white) or zero (black). Yellow and red dotted lines highlight the position of the edges of the bone.

Analogously, Fig. S15 shows the intensity image M_{00} (Fig. S15(a)) and the four probability function images (muscle, tendon, myotendinous junction and bone recognition (Figs. S15(b-e), respectively) for a particular myotendinous junction sample. The obtained probability images when measuring the myotendinous junction sample demonstrate how all the models discard the analyzed myotendinous junction tissue to be muscle (Fig. S15(b)), tendon (Fig. S15(c)) or bone (Fig. S15(e)), except the actual myotendinous junction model (Fig. S15(d)), who succefully identify the tissue as myotendinous junction one. Therefore, the overall recognition of the tissue is correct, as the above-stated results are consistent in most image pixels. Only few individual pixels throughout the images show deviation from the general tendency, but still they could be properly categorized depending on the selected categorization probability threshold selected.

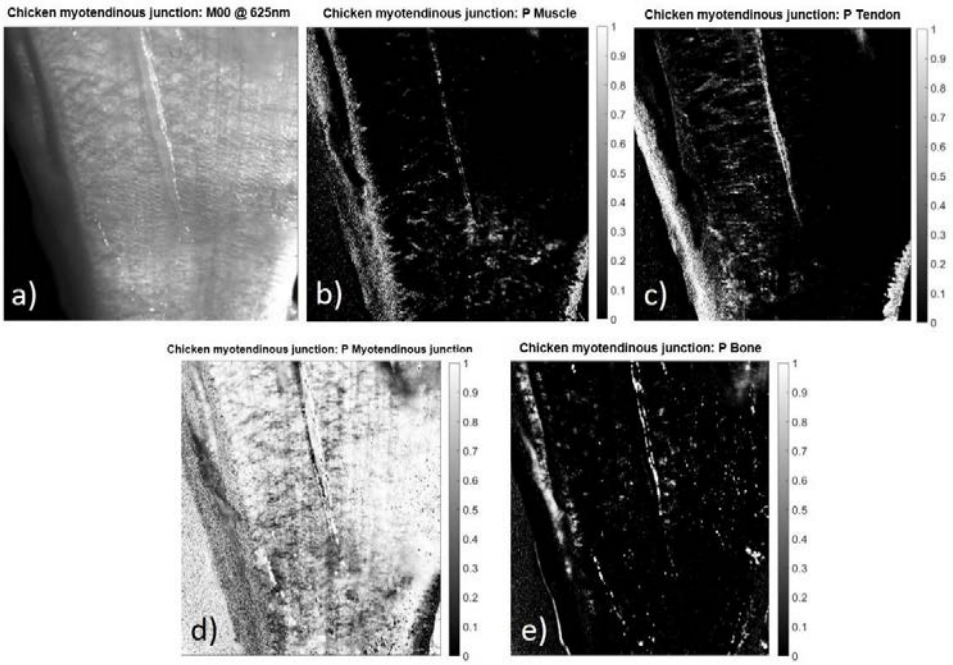


Figure S15. Intensity image M00 (a) and probability function of muscle (b), tendon (c), myotendinous junction (d) and bone (e) for chicken myotendinous junction measurements. The gray level bars, placed to the right of the corresponding probability function images, defines whether the probability of the pixel to be recognized as a particular tissue is one (white) or zero (black).

7. References

1. A. Van Eeckhout, A. Lizana, E. Garcia-Caurel, J.J. Gil, A. Sansa, C. Rodríguez, I. Estévez, E. González, J.C. Escalera, I. Moreno, and J. Campos. "Polarimetric imaging of biological tissues based on the indices of polarimetric purity," Journal of Biophotonics, e201700189 (2017).
2. A. Van Eeckhout, A. Lizana, E. Garcia-Caurel, J.J. Gil, R. Ossikovski, and J. Campos, "Synthesis and characterization of depolarizing samples based on the indices of polarimetric purity," Opt. Lett. **42**(20), 4155-4158 (2017).
3. A. Van Eeckhout, E. Garcia-Caurel, R. Ossikovski, A. Lizana, C. Rodríguez, E. González-Arnay, and J. Campos, "Depolarization metric spaces for biological tissues classification," Journal of Biophotonics, e202000083 (2020).
4. A. Van Eeckhout, E. Garcia-Caurel, T. Garnatje, M. Dufort, J. C. Escalera, J. Vidal, J. J. Gil, J. Campos, and A. Lizana, "Depolarizing metrics for plant simples imaging," PLoS ONE **14**(3), e0213909 (2019).
5. J. E. V. Ferreira, M. T. S. Pinheiro, W. R. S. D. Santos, and R. D. S. Maia, "Graphical representation of chemical periodicity of main elements through Boxplot," Educación Química **27**(3), 209–216 (2016).
6. R. McGill, J. W. Tuckey, and W. A. Larsen, "Variations of Box Plots," The American Statistician **32**(1), 12–16 (1978).
7. S. Guo, S. Zhong, and A. Zhang, "Privacy-preserving Kruskal–Wallis test," Computer Methods and Programs in Biomedicine **112**(1), 135–145 (2013).
8. G. D. Ruxton, and G. Beauchamp, "Some suggestions about appropriate use of the Kruskal–Wallis test," Animal Behaviour **76**(3), 1083–1087 (2008).
9. M. Zhu, and A. Ghodsi, "Automatic dimensionality selection from the Scree plot via the use of profile likekihood," Computational Statistics and Data Analysis **51**(2), 918–930 (2006).
10. H. H. Jolliff, "Principal Component Factor Analysis," All Graduate Plan B and other Reports, 1117 (1968).

11. C. C. Clark, C. M. Barnes, M. J. Duncan, H. D. Summers, and G. Stratton, “Physical activity, motor competence and movement and gait quality: A Principal Component Analysis,” Human Movement Science **68**, 102523 (2019).
12. Y. Tahtali, “Use of Factor Scores in Multiple Regression analysis for estimation of body weight by certain body measurements in Romanov Lambs,” PeerJ, **7**, e7434 (2019).
13. N. J. D. Nagelkerke, “A note on the general definition of the coefficient of determination,” Biometrika, **78**(3), 691–692 (1991).

PAPER D

Optimizing the classification of biological tissues using Machine Learning models based on polarized data

Carla Rodríguez[1,*], Irene Estévez[1,2], Emilio González-Arnay[3,4], Angel Lizana[1] and Juan Campos[1]

[1] *Grup d'Òptica, Physics Department, Universitat Autònoma de Barcelona, Bellaterra, 08193, Spain*
[2] *Centre of Physics, Deparment of Physics, University of Minho, Azurém Campus, 4800-058, Guimaraes, Portugal*
[3] *Servicio de Anatomía Patológica, Hospital Universitario de Canarias, Santa Cruz de Tenerife, 38320, Spain*
[4] *Departamento de Anatomía, Histología y Neurociencia, Universidad Autónoma de Madrid, Madrid, 28049, Spain*

This paper has been published in *Journal of Biophotonics*, e202200308, 2022.