Mueller polarimetry for the assessment of tissue microstructure in digital histology and cryopreservation

Deyan Ivanov

INSTITUT POLYTECHNIQUE DE PARIS

NNT : 2023IPPAX069

Thèse de doctorat

ÉCOLE POLYTECHNIQUE

IP PARIS

Mueller polarimetry for the assessment of tissue microstructure in digital histology and cryopreservation

Thèse de doctorat de l'Institut Polytechnique de Paris
préparée à l' École Polytechnique

École doctorale n°626 (EDIPP)
Spécialité de doctorat: Optique, Lasers et Plasmas

Thèse présentée et soutenue à Palaiseau, le 07.09.2023, par

DEYAN IVANOV

Composition du Jury :

Adolf Canillas Professeur, University of Barcelona, Spain	Rapporteur
Anabela Da Silva DR, CNRS, Institut Fresnel, Marseille, France	Rapporteur, Présidente
Walter Blondel Professeur, Université de Lorraine, Nancy, France	Examinateur
Dominique Pagnoux CR, CNRS, Université de Limoges, Limoges, France	Examinateur
Razvigor Ossikovski Professeur, École Polytechnique (LPICM)	Directeur de thèse
Tatiana Novikova IR, École Polytechnique (LPICM)	Co-directrice de thèse

With the completion of the manuscript and the thesis defense,
I would like to honor my Family, Friends and Relatives
for both the support and the patience provided throughout the years so far!

Also, I would like to dedicate the current manuscript and pay special tribute
in memory of Assoc. Prof. Ekaterina Borisova, PhD,
who believed in me and was always ready to support me!

Acknowledgements

Firstly, I would like to acknowledge all my former *Teachers* and *Mentors*, starting from *Elementary School* and continuing through *High School* and also the *University* for all the knowledge and lessons, which I learned.

Next, I want to thank all my former *Supervisors* and *Colleagues* from the *Bulgarian Academy of Sciences* for giving me the opportunity and teaching me how to work in Academic environment and how to be a good Researcher.

Furthermore, I would like to appreciate the PhD funding of *Institut Polytechnique de Paris* and also the opportunity provided from *LPICM, École Polytechnique* to work in World-leading laboratory in Polarimetry - it was an Honor for me.

Without any doubt, the current work would not have been possible to be finished without the collaboration of our partners and all Colleagues, Researchers and Scientist from: *Institute of Electronics, Bulgarian Academy of Sciences, Sofia, Bulgaria; University Hospital Queen Joanna, Sofia, Bulgaria; Oulu University, Oulu, Finland; Aston University, Birmingham, England; Fraunhofer Institute for Silicate Research, Wuerzburg, Germany; University Hospital, Wuerzburg, Germany; University Hospital, Bern, Switzerland*; *University of Bern, Bern, Switzerland* and *Tsinghua University, Beijing, China*.

Then, I can't miss the opportunity and the chance to acknowledge the contribution to all anonymous *Patients*, who volunteered for the studies.

Acknowledgements to my *Academic Supervisors - Prof. Razvigor Ossikovski and Dr. Tatiana Novikova* for their dedicated supervision during the years. Specifically, I want to thank for all the efforts, lessons, help and the devotedly time spent from *Prof. Razvigor Ossikovski* and *Dr. Enrique Garcia-Caurel* to teach and mentor me, thus aiding me to learn and further progress in the field of Polarimetry. Sincerely, I would like to thank all *Colleagues* and *Lab members* from *LPICM*, especially to the *administrative staff*, for their support and help.

Last but not least, I want to hugely express my gratitude to *Madame Sylvie Coussot, Prof. Yvan Bonnassieux, Prof. Jean-Charles Vanel* and *Prof. Razvigor Ossikovski* by helping me to overcome an extremely difficult situation and by this way teaching me to improve myself as a person, to have better work-life balance and better working ethics.

”Le savant n’étudie pas la nature parce que cela est utile ; il l’étudie parce qu’il y prend plaisir et il y prend plaisir parce qu’elle est belle .” - Henri Poincaré

Contents

1 **Theory of polarimetry** 22
- 1.1 Polarization of light . . . 22
- 1.2 Stokes-Mueller formalism . . . 24
- 1.3 Poincaré sphere representation . . . 25
- 1.4 Coherency matrix and Pauli spin matrices . . . 28
- 1.5 Physical realizability . . . 29
- 1.6 Decomposition algorithms . . . 30
  - 1.6.1 Polarization purity and polarization entropy . . . 30
  - 1.6.2 Differential decomposition . . . 31
  - 1.6.3 Lu-Chipman decomposition . . . 34
  - 1.6.4 Symmetric decomposition . . . 36
- 1.7 Depolarization spaces . . . 38

2 **Polarimetric instrumentation** 40
- 2.1 Eigenvalue calibration method . . . 41
- 2.2 Transmission geometry . . . 44
- 2.3 Reflection geometry . . . 49

3 **Image processing algorithms** 53
- 3.1 Substrate contribution removal . . . 53
- 3.2 Image segmentation . . . 54
- 3.3 Histogram normalization . . . 55
- 3.4 Entropy and Standard deviation . . . 56

4 **Experimental results and data post-processing** 59
- 4.1 Multi-layered stacks of thin sections of anisotropic scattering white matter of brain . . . 61
- 4.2 Assessment of different thawing mechanisms of tissue models . . . 69
  - 4.2.1 Polarimetric results . . . 69
  - 4.2.2 Statistical analysis . . . 78
- 4.3 Digital histology of thin sections of human skin . . . 82
  - 4.3.1 Polarimetric results . . . 82
  - 4.3.2 Deep learning . . . 89

4.4 Colon cancer detection: ex-vivo studies . . . 95
4.4.1 Polarimetric results . . . 95
4.4.2 Statistical analysis . . . 104
4.4.3 Principal component analysis . . . 108
4.4.4 Machine learning . . . 111

5 Future perspectives **118**

6 Appendix **120**
6.1 Principal component analysis . . . 120
6.2 Logistic regression . . . 121
6.3 Random forest . . . 123
6.4 Support vector machines . . . 124
6.5 Deep learning . . . 127

7 Bibliography **132**

8 Educational and scientific part **140**
8.1 Courses taken: . . . 140
8.2 Conferences: . . . 141
8.3 Project participation: . . . 142
8.4 Publications: . . . 143
8.4.1 Published in peer-review journals: . . . 143
8.4.2 Published in conference proceedings: . . . 143
8.5 Award: . . . 145
8.6 Mentorship: . . . 145

List of Figures

2 Poincaré sphere representation. 26

3 Graphical representation of different polarization states with ρ = 1: (a) horizontal, (b) vertical, (c) +45 and (d) left circular. All red lines represent the vertical component of the electric field vector, all black lines - the horizontal component, while the blue line is the resultant wave. `https://fr.mathworks.com/help/phased/ug/modeling-and-analyzing-polarization.html` 27

4 Schematic representation of polarized light propagation in an anisotropic medium along the *z*-axis of a Cartesian coordinate system. The medium is transversally homogeneous (along *x,y*) and longitudinally inhomogeneous (along *z*). 32

5 Schematic representation of polarized light propagation in anisitropic medium along the *z*-axis of a Cartesian coordinate system, following sequence of diattenuator, retarder and depolarizer. 36

6 3D representation of: (a) Canonical, (b) Natural, (c) Indices of polarimetric purity depolarization spaces, where the symbols represent • – pure polarizer, • – pure depolarizer. 38

7 Schematic representation of the calibrating samples: a) horizontal polarizer, b) vertical polarizer (the red line shows the transmission axes), c) quarter-wave plate with fast axis (blue line), oriented at 30^o with respect to the x - axis, d) laboratory co-ordinate frame, light propagation assumed along z - axis. 43

8 Experimental set-up in transmission geometry: (a) Schematic representation and (b) Side view. Figure available from the published material in [59]. 44

9 Schematic representation of the components of the PSG and the PSA: a) HP: horizontal, polarizer; FLC1/2: ferroelectric liquid crystal retarders; WP, waveplate, and b) laboratory co-ordinate frame, light propagation assumed along z - axis. . . 46

10 Schematic representation of the experimental setup. Figure available from the published material in [63]. Special acknowledgments for the figure preparation to Dr. Viktor Dremin. 50

11 Schematic representation of the image segmentation. 55

12 Flowchart of the data post-processing sequence and algorithms. 58

13 example 62

14 Net scalar retardance in degrees, obtained with either differential (left column) or Lu-Chipman (right column) decomposition, corresponding to different spatial arrangements of the brain corpus callosum stripes: (a) and (b) single layer (10 μm); (c) and (d) parallel overlap (10 μm + 5 μm); (e) and (f) 45^o overlap (10 μm + 5 μm); (g) and (h) 90^o overlap (10 μm + 5 μm). 63

15 Box-whisker plots of the net scalar retardance from: a) differential and b) Lu-Chipman decomposition, corresponding to the different spatial configurations of the *corpus callosum*, brain samples. 64

16 Quiver plots of the net scalar retardance, obtained with differential (left column) or Lu-Chipman (right column) decomposition, corresponding to different spatial arrangements of the brain corpus callosum layers: (a) and (b) single layer (10 μm); (c) and (d) parallel overlap (10 μm + 5 μm); (e) and (f) 45° overlap (10 μm + 5 μm); (g) and (h) 90° overlap (10 μm + 5 μm). 65
17 Total depolarization, obtained with either differential (left column) or Lu-Chipman (right column) decomposition, corresponding to different spatial arrangements of the brain corpus callosum stripes: (a) and (b) single layer (10 μm); (c) and (d) parallel overlap (10 μm + 5 μm); (e) and (f) 45° overlap (10 μm + 5 μm); (g) and (h) 90° overlap (10 μm + 5 μm). 67
18 Box-whisker plots of the total depolarization from: a) differential and b) Lu-Chipman decomposition, corresponding to the different spatial arrangement of the corpus callosum brain stripes. 68
19 Photo of the microscopy glass slides with the samples under examination. From left to right: Control/Reference, water thawed and RF thawed with MNPs. 71
20 Microscopy image with the 3D artificial skin model (the darker zone) and the paraffin host medium around it in which the skin stripe was embedded. 71
21 The flowchart of the study design including the steps of samples' preparation, rewarming with MNPs, tissue models sectioning and polarimetric measurements. Special acknowledgements to Anika Hoeppel for the figure preparation. 72
22 Photo of the setup for the inductive thawing of the fibroblast-containing 3D tissue models. Special acknowledgements to Anika Hoeppel for the figure preparation. . 72
23 Intensity images of m_{11} (scales in bit depth): (a) reference, (b) water thawed and (c) RF thawed with MNPs. 73
24 Maps of the azimuth θ of the optic axis (in degrees): (a) reference, (b) rater thawed and (c) RF thawed with MNPs. 74
25 Normalized histograms of the θ angle for the three groups. 75
26 Entropy H of the azimuth of the optic axis θ for each patch with size 100$\times$100 pixels: (a) reference, (b) water thawed and (c) RF thawed with MNPs. 76
27 Standard deviation σ in bit depth of the azimuth of the optic axis θ (after converting in gray-scale) for each patch with size 100 x 100 pixels: (a) reference, (b) water thawed and (c) RF thawed with MNPs. 77
28 Box-Whisker plot of the intensity from all three groups. 78
29 Box-Whisker plot of the net dichroism Δd from all three groups. 79
30 Box-Whisker plot of the net depolarization α_t from all three groups. 79
31 Box-Whisker plot of the net, scalar retardance ϕ from all three groups. 79
32 Box-Whisker plot of the orientation angle θ from all three groups. 79
33 Graphical representation of the dependence between the error rate $\mathcal{E}$ and the number of comparisons γ for significance level α = 0.05. 80
34 Intensity images of m_{11} in bit depth units: (a) Healthy, (b) Scleroderma, (c) Lupus, (d) Psoriasis, (e) Syndrome of Raynaud, (f) Squamous-cell carcinoma, (g) Basal-cell carcinoma and (h) Melanoma. 83

35 Images of dichroism Δd: (a) Healthy, (b) Scleroderma, (c) Lupus, (d) Psoriasis, (e) Syndrome of Raynaud, (f) Squamous-cell carcinoma, (g) Basal-cell carcinoma and (h) Melanoma. 84
36 Images of depolarization α_t: (a) Healthy, (b) Scleroderma, (c) Lupus, (d) Psoriasis, (e) Syndrome of Raynaud, (f) Squamous-cell carcinoma, (g) Basal-cell carcinoma and (h) Melanoma. 85
37 Images of retardance ϕ: (a) Healthy, (b) Scleroderma, (c) Lupus, (d) Psoriasis, (e) Syndrome of Raynaud, (f) Squamous-cell carcinoma, (g) Basal-cell carcinoma and (h) Melanoma. 86
38 Images of the orientation angle θ of the optic axis: (a) Healthy, (b) Scleroderma, (c) Lupus, (d) Psoriasis, (e) Syndrome of Raynaud, (f) Squamous-cell carcinoma, (g) Basal-cell carcinoma and (h) Melanoma. 87
39 Normalized histograms: (a) dichroism – Δd, (b) depolarization – α_t, (c) retardance – ϕ and (d) orientation angle of the optic axis – θ. 88
40 Architecture of the CNN as a flowchart. 91
41 Architecture of the CNN as an illustration. 92
42 CNN evaluation during the training and validation phases: **(a)** model loss and **(b)** model accuracy. 92
43 Confusion matrix (CM) of the Skin-HDT700 model. The notations stand for 0 - healthy, 1 - degenerative and 2 - tumor classes. 93
44 ROC curves for the CNN and all classes of the Skin-HDT700 model: **(a)** whole ranges of the false positive and true negative rates, dashed line represents the ROC curve for the completely random classifier and **(b)** reduced ranges for better clarity. 93
45 Simulated and measured Mueller matrices: (a) Monte Carlo (MC) (b) Tumor (T), (c) Healthy (H) and (d) Phantom (Ph). Results available from the published material in [63]. 96
46 Mueller matrices of the Diattenuator D_2 obtained from the symmetric decomposition for: (a) Monte Carlo (MC) (b) Tumor (T), (c) Healthy (H) and (d) Phantom (Ph). Results available from the published material in [63]. 97
47 Mueller matrices of the Retarder R_2 obtained from the symmetrical decomposition for: (a) Monte Carlo (MC) (b) Tumor (T), (c) Healthy (H) and (d) Phantom (Ph). Results available from the published material in [63]. 97
48 Mueller matrices of the canonical depolarizer M_Δ obtained from the symmetrical decomposition for: (a) Monte Carlo (MC) (b) Tumor (T), (c) Healthy (H) and (d) Phantom (Ph). Results available from the published material in [63]. 98
49 Mueller matrices of the Retarder R_1 obtained from the symmetric decomposition for: (a) Monte Carlo (MC) (b) Tumor (T), (c) Healthy (H) and (d) Phantom (Ph). Results available from the published material in [63]. 98
50 Mueller matrices of the Diattenuator D_1 obtained from the symmetric decomposition for: (a) Monte Carlo (MC) (b) Tumor (T), (c) Healthy (H) and (d) Phantom (Ph). Results available from the published material in [63]. 99

51 Symmetric decomposition products: a), b), c) depolarization indices – d_i; d) net depolarization index – Δ; e), f) diattenuadtion D_i and g), h) retardance – ϕ_i. Results available from the published material in [63]. 100
52 3D representation of depolarization: (a) Canonical, (b) Natural, (c) Indices of polarimetric purity depolarization spaces, where the symbols represent ● – Healthy, ◆ – Tumor data points. Results available from the published material in [97]. 101
53 Scatter plots of: (a) D vs P, (b) D_2 vs D_1 and (c) R_2 vs R_1, for both healthy green points) and tumor (red points) zones, where $\phi_{1,2} \equiv R_{1,2}$. Results available from the published material in [97]. 101
54 Spatial distribution of the values of polarizance and diattenuation within both tissue zones: (a) *P* – Healthy, (b) *P* – Tumor, (c) D_2 – Healthy, (d) D_2 – Tumor. Results available from the published material in [97]. 102
55 Spatial distribution of the values of depolarization Δ and polarization entropy *S* within both tissue zones: (a) Δ – Healthy, (b) Δ – Tumor, (c) *S* – Healthy, (d) *S* – Tumor. Results available from the published material in [97]. 103
56 Spatial distribution of the values of the polarimetric purity – *PI* and the Gil-Bernabeu's depolarization index – P_Δ within both tissue zones: (a) *PI* – Healthy, (b) *PI* – Tumor, (c) P_Δ – Healthy, (d) P_Δ – Tumor. Results available from the published material in [97]. 104
57 Visualization of polarimetric data sets in different parametric spaces at all spatial locations of the measurements of both healthy and tumor zones of colon tissue ● – healthy and ◆ – tumor via: **(a)** Poincaré sphere for probing (or incident) circular polarization, **(b)** natural, **(c)** IPP and **(d)** canonical depolarization spaces. Results available from the published material in [99]. 105
58 Flowchart of the feature selection. 107
59 Correlation matrix of the eigenvalue model ($\phi_1 \equiv R_1$ and $\lambda_i \equiv l_i$). 108
60 Correlation matrix of the IPP model $\phi_1 \equiv R_1$). 109
61 PCA for N number of components, explaining the corresponding percentage of variance σ^2 for both submodels. Results available from the published material in [99]. 110
62 **(a)** Correlation matrices for: **(a)** 7 PCs and their scores – eigenvalue model and **(b)** 6 PCs and their scores – IPP model. Results available from the published material in [99]. 110
63 Probability for tumor detection, calculated from LR algorithm: **(a)** d_1, **(b)** R_1, **(c)** λ_2, **(d)** P, **(e)** D_2 and **(f)** D_2. For the panels **a,b,d** and **e** the results are comparable for both sub-models, whereas the panels **c** and **f** show the results computed from the eigenvalue and the IPP sub-models, respectively ($\phi_1 \equiv R_1$ and $\lambda_2 \equiv l_2$). Results available from the published material in [99]. 111
64 ROC curves for: **(a)** eigenvalue sub-model (trained with 7 PCs and their scores) and **(b)** IPP sub-model (trained with 6 PCs and their scores). Results available from the published material in [99]. 113
65 Flowchart of the supervised ML process. 114

66 Sigmoid function $\sigma(x)$ of real variable x. 122
67 Arbitrary data and provided code solutions available from the course *Machine and Deep Learning, École Polytechnique, INF 554*: **(a)** Linearly separable and **(b)** Non-linearly separable. 124
68 Arbitrary data and provided code solutions available from the course *Machine and Deep Learning, École Polytechnique, INF 554*: **(a)** SVM model with linear kernel, C = 100 and **(b)** SVM model with linear kernel, C = 1000. 125
69 Arbitrary data and provided code solutions available from the course *Machine and Deep Learning, École Polytechnique, INF 554*: **(a)** Linearly separable data and **(b)** Non-linearly separable data. In both cases Gaussian kernel was applied with C = 1 and σ = 0.1. 127

List of Tables

1 Retardance and orientation angle values (in degrees) of the optical elements in the optimized PSG and PSA. For the ferro-electric liquid crystals the retardance values are at the wavelength of 510 nm, while for the two wave plates, at 633 nm. . 48
2 Supplementary table associated with decomposition data. Data available from the published material in [63]. 99
3 Supplementary table associated with all classification MLAs performances, where all numerical values are in %. All MLAs were trained with 7 PCs and their scores for the eigenvalue sub-model and 6 PCs and their scores for the IPP sub-model. Data available from the published material in [99]. 113

List of Polarimetric parameters and statistical quantities, used as optical markers and/or predictors for modeling

- S_i, $i \in$ [0,3] or **S** - Stokes parameters and Stokes vector for a given polarization state;
- **M** and **H** - Mueller matrix and its covariance matrix;
- m_{ij} - matrix elements of **M**;
- λ_i - eigenvalues of **H**;
- P_i, $i \in$ [1,3] - indices of polarimetric purity, obtained from λ_i;
- PI - overal polarimetric purity, obtained from P_i;
- P_Δ - overal depolarization ability, obtained from P_i;
- S - Cloude's entropy, obtained from λ_i;
- P and **P** - Polarizance and its vector, obtained from the first row of **M**;
- D and **D** - Diattenuation and its vector - obtained from the first column of **M**;
- **m**, $\mathbf{m_m}$ and $\mathbf{m_u}$ - differential Mueller matrix, polarization and depolarization tensors;
- ϕ_t - net, scalar retardance, obtained from $\mathbf{m_m}$;
- θ - orientation angle of the optic axis, used for all decomposition algorithms;
- Δd - net dichroism, obtained from $\mathbf{m_m}$;
- α_t - total depolarization, obtained from $\mathbf{m_u}$;
- d_i, $i \in$ [1,3] - principal depolarization factors from the symmetric decomposition;
- Δ - net depolarization index - used for both Lu-Chipman and symmetric decompositions;
- $D_{1,2}$ - diattenuation parameters from the symmetric decomposition;
- $\phi_{1,2}$ - retardance values from the symmetric decomposition;
- H - entropy of a given statistical distribution;
- σ - standard deviation;
- PC - principal component;
- PCS - principal component scores.

Partner organizations

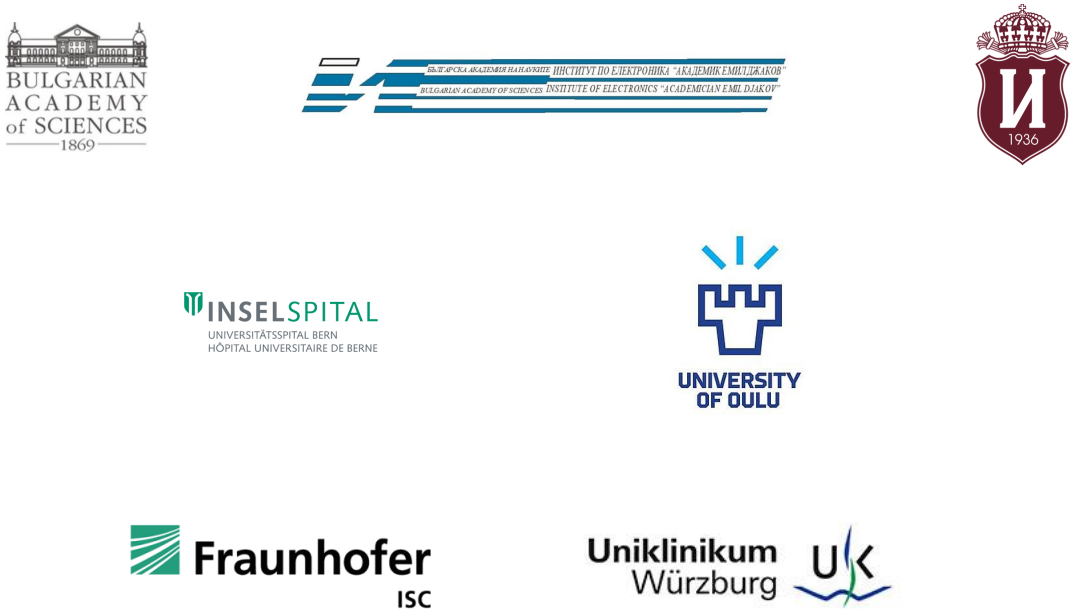


Figure 1: List of all partner organizations and collaborations.

During the PhD period, a multi-partner collaborative framework was established within which this work was carried out, thus defining the scope of the dissertation. Firstly, for the brain polarimetry studies with multi-layered brain phantoms, the partner organization is University hospital, Bern, Switzerland. Then, for the polarimetric assessment of different thawing methods of artificial skin complemented by statistical analysis, the collaboration continues with Fraunhofer Institute for Silicate Research, Wuerzburg, Germany and University Hospital, Wuerzburg, Germany. Then, the project related about digital histology of skin tissues with polarized light complemented by AI, became possible after collaboration between Institute of Electronics, "Acad. E. Djakov", Bulgarian Academy of Sciences, Sofia, Bulgaria and University Hospital, Queen Joanna, Sofia, Bulgaria. Finally, the project related about digital histology of gastric tissues with polarized light complemented by AI, became possible after collaboration between Institute of Electronics, "Acad. E. Djakov", Bulgarian Academy of Sciences, Sofia, Bulgaria, University Hospital, Queen Joanna, Sofia, Bulgaria and Oulu University, Oulu, Finland.

Abstract

Aim and Significance: Polarization of light can be used as an optical modality sensitive to tissue structures and biomedical samples in general. Thus, this vectorial property of light can be applied to diagnose non-invasively healthy, benign and malignant formations. Likewise, when biomedical samples are frozen and then defrosted, the internal damages occurring in the tissue structures due to defrosting can be also detected with polarized light.

Approach: An imaging Mueller polarimeter operating in the visible range was used to measure the full polarimetric response of the tissue specimens under study. The optical system was configured as a Mueller microscope in transmission geometry. Thus, medical doctors could be assisted in their diagnosis by performing polarization-sensitive digital histology. For thick tissue specimens, reflection rather than transmission, geometry and scanning were chosen. In that case, the backscattered photons convey diagnosis information by experiencing more scattering events and, thus, being more depolarized, however at the price of lower light intensity and reduced signal-to-noise ratio. Regardless of the experimental geometry, precise calibration (the eigen-value calibration method), as well as a decomposition algorithm to filter data noise and/or experimental errors (physical realizability filtering), are required. Depending on the experimental configuration, transmission or reflection, the differential and Lu-Chipman decompositions or the symmetric decomposition were respectively used. All of these algorithms make it possible to interpret phenomenologically an experimental Mueller matrix, without the need for an explicit physical model of the given tissue sample. Nevertheless, efforts were also spent on relating the decomposition results to the available structural and biomedical information.

Results: In transmission, corpus callosum ex vivo brain slides with different spatial orientations were used to study the behavior of the polarimetric parameters provided by the differential and the Lu-Chipman decomposition algorithms at 532 nm and to decide on the most suitable decomposition. In this study, the Mueller microscope was used to visualize the orientation of brain fiber tracts as well, which can be helpful to guide neurosurgeons during operations. Next, non-frozen (initially frosted at -80^o C and then defrosted either with conventional heating or with magnetic-nanoparticle-assisted thawing) tissue models were measured with the Mueller microscope at 700 nm. The experimental results were processed with the differential decomposition algorithm, with the help of which small differences in the tissue structures were detected and confirmed by statistical analysis. This approach has the potential to be used in the field of regenerative medicine to assess any tissue alterations due to cryopreservation and defrosting. In another study, skin samples with degenerative and malignant lesions were measured with the same instrument, geometry and wavelength (700 nm), and a deep learning model was elaborated for image classification. In this way, a step closer to digital histology assessment was made to assist medical doctors. Finally, a non-imaging polarimetric set-up in reflection was used to scan and discriminate the polarization responses of ex vivo colon samples with healthy and tumorous sections at 635 nm. With the increasing amount of experimental data, the problem of the inter-patient variability had to be handled inevitably. Additionally, when classifying the

polarization measurements in relation to the medical doctors' ground truth histological analysis, false positive and false negative data points had to be taken into account. By using statistical analysis, together with supervised and unsupervised machine-learning algorithms, it became possible to find adequate solutions to the aforementioned issues.

Conclusions: The scope of the PhD thesis covered various scientific domains including *Mueller polarimetry*, *Image processing & segmentation*, *Statistics* and *Artificial intelligence*, all of which in combination were found able to provide medical doctors and biologists with relevant supporting information on the internal structure and organization of the samples, either modified by the presence of malignancy or caused by the thawing mechanisms. More generally, the work carried throughout the thesis has been a part of a valuable partnership between LPICM and various, local and international, collaborative research groups. In the field of biomedical optics and photonics, a wide variety of aspects have been approached profoundly in order to find a better interpretation of complex data and provide complementary diagnosis techniques. In the end, a summary and an evaluation of the articles published in peer-review journals, the reports presented in scientific conferences and the list of courses undertaken in the framework of the current thesis are given.

Résumé

Objectifs et intérêt: La polarisation de la lumière peut être utilisée comme modalité optique sensible à la structure des tissus et aux échantillons biologiques en général. Concrètement, cette propriété vectorielle de la lumière peut être appliquée au diagnostic non invasif de formations saines, bénignes ou malignes. De façon similaire, suite à la congélation et à la décongélation d'échantillons biologiques, les dégâts internes apparaissant dans la structure des tissus à cause de la décongélation peuvent être aussi détectés par le biais de la lumière polarisée.

Approche: Un polarimètre-imageur de Mueller, fonctionnant dans le visible, a été utilisé afin de mesurer la réponse polarimétrique complète des échantillons de tissu étudiés. Le dispositif optique a été configuré comme microscope de Mueller en configuration de transmission. De cette façon, les médecins peuvent être assistés dans leur diagnostic par l'histologie numérique sensible à la polarisation. Dans le cas d'échantillons épais, la configuration de réflexion plutôt que de transmission a été choisie. Dans ce cas, les photons rétrodiffusés contiennent des renseignements utiles au diagnostic du fait d'éprouver de plus nombreux actes de diffusion et donc, d'être plus dépolarisés, quoique au prix d'une intensité lumineuse réduite et un rapport signal au bruit dégradé. Indépendamment de la configuration expérimentale, une étalonnage précis (par la méthode des valeurs propres), de même qu'un algorithme de décomposition pour filtrer le bruit des données et/ou les erreurs expérimentales sont nécessaires. En fonction de la configuration expérimentale – transmission ou réflexion – les décompositions différentielle, de Lu-Chipman ou la symétrique ont été respectivement utilisées. Tous ces algorithmes rendent possible l'interprétation phénoménologique d'une matrice de Mueller expérimentale, sans nécessiter un modèle physique explicite de l'échantillon de tissu. Cependant des efforts ont été également dirigés vers l'établissement de liens entre les résultats des décompositions et les informations structurelles et biomédicales disponibles.

Résultats: En transmission, des tranches ex vivo de cerveau (corpus callosum) avec des orientations spatiales différentes ont été utilisées pour étudier le comportement des paramètres fournis par les algorithmes de décomposition différentielle et de Lu-Chipman à 532 nm afin de décider de la décomposition la plus adaptée. Dans cette étude, le microscope de Mueller a été également employé pour visualiser l'orientation des tracts de fibres du cerveau ce qui pourrait être utile comme guide pour les chirurgiens lors des opérations. Ensuite, des modèles de tissus non congelés (initialement congelés à –80 ℃ et ensuite décongelés soit par échauffement classique, soit par fonte assistée par des nanoparticules métalliques) ont été mesurés avec le microscope de Mueller à 700 nm. Les résultats expérimentaux ont été traités avec l'algorithme de la décomposition différentielle à l'aide duquel de petites différences dans la structure des tissus ont été détectées et confirmées par une analyse statistique. Cette approche peut potentiellement être employée dans le domaine de la médecine régénérative afin d'évaluer toute altération du tissu causée par la cryoconservation et la décongélation. Dans une autre étude, des échantillons de peau avec des lésions dégénératives et malignes ont été mesurés avec le même instrument, même configuration et longueur d'onde (700 nm) et un modèle d'apprentissage profond a été

construit pour la classification des images. De cette façon, un pas de plus a été fait dans la direction de l'évaluation par histologie numérique afin d'assister les médecins. Finalement, un dispositif non imageur en réflexion a été utilisé pour topographier et discriminer les réponses en polarisation d'échantillons ex vivo de tranches de colon saines et tumorales à 635 nm. Le volume de données croissant, le problème de la variabilité inter-patients a dû inévitablement être traité. De plus, les faux positifs et les faux négatifs dans les données ont dû être pris en compte lors de la classification des mesures polarimétriques par rapport à la vérité de base issue de l'analyse histologique des médecins. En employant l'analyse statistique, accompagnée d'algorithmes d'apprentissage automatique supervisés et non supervisés, il a été possible de trouver des solutions adéquates aux problèmes cités.

Conclusions: La portée de cette thèse de doctorat a couvert des domaines scientifiques divers tels que la polarimétrie de Mueller, le traitement des images et la segmentation, les statistiques et l'intelligence artificielle. Ils ont tous été démontrés capables de fournir aux médecins des renseignements supplémentaires pertinents sur la structure interne et l'organisation des échantillons, modifiées soit par la présence de formations malignes, soit par les mécanismes de décongélation. Dans un contexte plus général, ce travail de thèse faisait partie de partenariats étroits entre LPICM et différents groupes de recherche nationaux et étrangers. Une grande diversité d'aspects propres aux domaines de l'optique biomédicale et de la photonique ont été traités en profondeur afin d'établir une meilleure interprétation de données complexes et de fournir des techniques de diagnostic complémentaires. En annexe, on trouve un résumé ainsi que les listes des articles publiés dans des journaux à comité de lecture, des rapports présentés à des congrès scientifiques et des cours suivis lors de la thèse.

Introduction

Polarimetry has already been recognized as an efficient optical technique for material characterization [1–4]. From an experimental point of view, with the currently available instrumentation, one should be able to measure and characterize a wide range of materials, preserving or altering the initial state of light polarization. The full polarization fingerprint is contained in the Mueller matrix (**M**) of the sample under study [5]. Within the expanding field of biophotonics, tissue polarimetry can be likewise utilized as label-free modality for supplementary diagnostic information provided to physicians [6–13]. Thus, polarization of light is sensitive to tissue morphology and can reveal relevant information about the tissue internal structure and reorganization. In such cases, the availability of a physical model for the corresponding turbid medium is highly desirable. Unfortunately, this is not always the case, while it remains of huge importance to interpret physically an experimentally measured Mueller matrix and consequently, to retrieve all polarization and depolarization properties. If one knows the physically realizable form of **M** for the corresponding medium under interest, all polarization and depolarization properties can be recovered via pertinent decomposition algorithms [5].

Prior to probing biological specimens, their basic optical properties should be ideally known. In general, biological specimens are considered as turbid optical media [14] so that the relevant properties are the scattering and the absorption coefficients μ_s, μ_a, and respectively the anisotropy factor *g*, as well as the refractive index *n* [14]. Unfortunately, these are usually not known *a priori* and the measurements must be conducted without explicit knowledge of the optical parameters of the samples. Needless to say that this adds more complexity to the final interpretation of the experimental results. Whenever tissue damage, alteration or modification occur the above-mentioned optical properties are modified. As a result, the light scattering also changes and affects the (de)polarization properties of the samples under study.

One possible reason for tissue damage is the so called cryopreservation. Briefly, this is a method for long-term storage of biological tissues at low temperatures [15, 16]. It remains one of the main actual problems of regenerative medicine and tissue engineering. Delivery of transplants and organs requires maintenance of the tissues by means of cryopreservation, i.e. the long-term storage at very low temperatures, e.g. -140^o C. Standard cryopreservation often fails due to the growth of large ice crystals during tissue freezing and thawing. It causes damage of the cell membranes and the extracellular collagen matrix [17–19]. Fast and more uniform warming of a large volume of cryopreserved tissue may be achieved by introducing magnetic nano-particles (MNPs) into the tissues before freezing. Generating heat internally by applying an external magnetic field causing periodic oscillations of the MNPs favors rapid and homogeneous heat distribution. This is known to help avoiding the recrystallization (growth of large ice crystals at the expense of small ones) and associated tissue damage [20–22]. The latter affects both the extra-cellular matrix (ECM) and the cellular components of the biological tissues. As a consequence, a break in the membrane integrity is followed by a change in the tissue scattering properties and modification of the degree of polarization of transmitted and/or

reflected light. Damage to the ECM manifests in both loss of tissue anisotropy and a decrease in tissue retardance.

Another reason for tissue-structure damages is the presence of malignant formations. In short, cancer development can be summarized as: *i)* morphological and biochemical alterations in the cells in which the malignancy has occurred, followed by cellular density growth; *ii)* invasion by the tumor cells of the adjacent tissue, which leads to the destruction of the healthy tissue; *iii)* metastasis as a result of the tumor spread in other locations of the body via the lymph or the blood [23]. Pathologists are in charge of conducting macroscopic and microscopic examinations of the excised tissues from surgery. These types of examinations aim at making accurate diagnosis and at localizing any zones of the tissues with residual cancer [24]. However, this is not a trivial task, and requires extensive knowledge and experience. It is time consuming, since involving sample cutting and fixing, with possible tissue staining with hematoxylin and eosin, as well as preparation and examination of many histological slides [25]. To aid and supplement medical doctors, polarization-sensitive digital histology holds great promise as a fast, non-invasive and relatively cheap optical method in cancer detection.

Additionally, with the increasing amount of experimental data, apt post-processing algorithms are required, alongside the inclusion of statistical analyses and implementation of an artificial intelligence (AI) framework. The latter could be utilized to mimic human-like intellect when handling large and complex data sets, images, etc. Being a part of AI, the vastly expanding field of machine learning (ML) covers a wide spectrum of applications for solving multiple scientific problems [26–30], as well as for cancer classification [31–35]. Since conventional programming processes input data by means of particular syntax and semantics to produce a desired output, it is a method prone to multiple error repetition. To overcome this issue, ML uses both the input and output data to train an algorithm for an *a priori* defined purpose. Depending on their purpose, ML algorithms can be grouped into three distinct classes [36–38], namely supervised, unsupervised and reinforcement. The performance of each ML algorithm is evaluated by means of computing the corresponding confusion matrix, areas under the curves (AUC), as well as by loss and relative risk calculations related to misclassifications. Without any doubt, the data collected from the measurements during the PhD project are not enough to form a database and a model for clinical applications, but are clearly sufficient for pilot study. The absence of general physical model describing the light tissue interactions with polarized light could potentially be compensated by sufficiently large data sets of ground truth polarization-sensitive measurements, initially discussed with the medical doctors. By means of classification of newly obtained data, the work of the medical staff may be eased, while the classification time may be significantly reduced.

In order to cover all of the aforementioned aspects, the outline of the dissertation is organized as follows. In Chapter 1 the theoretical description of light polarization is available, including also the mathematical foundation of all decomposition algorithms used in the dissertation. In Chapter 2 the experimental configurations and calibration methods are described accordingly. Chapter 3 deals with the algorithms used for image processing, before describing the experimental results in Chapter 4, including various data post-processing approaches when analysing larger data sets, by applying the concepts of artificial intelligence, machine-learning, as well as statistical

analysis. Additionally, the mathematical background behind the artificial intelligence algorithms is reported in the Appendix 6. Finally, the bibliography together with the list of all conferences, projects and courses taken are listed in Chapters 7 and 8, respectively.

To sum up in due detail, the experimental aspect of the dissertation is focused on the following key points:

1. Validation of the proposed decomposition algorithms and their comparison whenever possible;
2. Detailed study and application of various decomposition algorithms to extract the underlying polarization and depolarization information of tissue samples and tissue-like micro-structures;
3. Application of the aforementioned image processing algorithms on the available data;
4. Study of the tissue thawing mechanisms after heating at room temperature as well as with external magnetic field applied to bio-compatible FeO magnetic nano-particles, incorporated in the tissue prior to cryopreservation;
5. Comparison of the results for reference (i.e. non frosted and defrosted) samples, conventionally defrosted tissue samples and samples defrosted with external magnetic field and magnetic nano-particles;
6. Analysis and evaluation of the behaviour of the polarimetric parameters in multilayered structures of brain samples;
7. Comparison of the results for healthy, degenerative and tumorous skin samples initially diagnosed from histopathologists;
8. Comparison of the results for healthy and tumorous colon samples initially diagnosed from histopathologists;
9. Implementation of data post-processing algorithms for more advanced and deeper analysis of the experimental results such as statistical analysis, image processing and artificial intelligence.

1 Theory of polarimetry

In order to define the polarization of light, one can simply consider its wave nature and disregard the particle-wave duality. In this essence, electromagnetic radiation is a transverse wave in which oscillations of the electric and magnetic fields take place, the latter being represented by corresponding vectors along directions orthogonal to the direction of propagation described by the wave vector. [39–41] It is to be mentioned that the scope of the current dissertation is focused only on the visible spectrum of light. Nevertheless, the theoretical part can be as equally applied to the UV and the IR spectral ranges.

When we define and refer to light polarization, we mean the manner in which the vibration of the electric field vector takes place in space. The most general polarization state is the elliptical one, which typically results from reflection and scattering of linearly polarized light. [40, 41]

When biological specimens are probed with polarized light, the light-tissue interactions can be mathematically described via the Stokes-Mueller formalism, valid for light with arbitrary degree of polarization. [42–44] The knowledge of how optical polarization evolves in biological specimens can reveal additional information which can be related to the health condition of the specimen under study. [44–48] This remarkable opportunity makes it possible to use polarized light for biomedical diagnostics and imaging [49–52] and, in this way, to complement medical doctors in more sensitive diagnostics at an earlier stage.

1.1 Polarization of light

For a plane, polarized, monochromatic electromagnetic wave propagating in the *z*-direction, the electric field can be described at time *t* and point with coordinates *x,y* as [39–41]:

$$E_x(z,t) = \Re[E_{0x}e^{i(\omega t - kz + \delta_x)}] = E_{0x}cos(\omega t - kz + \delta_x) \tag{1a}$$

$$E_y(z,t) = \Re[E_{0y}e^{i(\omega t - kz + \delta_y)}] = E_{0y}cos(\omega t - kz + \delta_y), \tag{1b}$$

where E_0 denotes amplitude value, $\omega = d\varphi/dt$ is the angular frequency, $k = 2\pi n/\lambda$ is the wave-number for a given refractive index *n*, while δ is an arbitrary phase. By dividing both parts of Eq.1a and b by the factor of E_0 and using the trigonometric properties, one gets:

$$\frac{E_x}{E_{0x}} = cos(\omega t - kz)cos(\delta_x) - sin(\omega t - kz)sin(\delta_x) \tag{2a}$$

$$\frac{E_y}{E_{0y}} = cos(\omega t - kz)cos(\delta_y) - sin(\omega t - kz)sin(\delta_y). \tag{2b}$$

After rearranging, the following equations can be derived:

$$\frac{E_x}{E_{0x}}sin(\delta_y) - \frac{E_y}{E_{0y}}sin(\delta_x) = cos(\omega t - kz)sin(\delta_y - \delta_x) \tag{3a}$$

$$\frac{E_x}{E_{0x}}cos(\delta_y) - \frac{E_y}{E_{0y}}cos(\delta_x) = sin(\omega t - kz)sin(\delta_y - \delta_x). \tag{3b}$$

Finally, in order to derive the equation for the polarization ellipse, Eqs.3a and b must be squared and added to give:

$$\frac{E_x^2}{E_{0x}^2} + \frac{E_y^2}{E_{0y}^2} - 2\frac{E_x}{E_{0x}}\frac{E_y}{E_{0y}}cos(\delta_y - \delta_x) = sin^2(\delta_y - \delta_x). \tag{4}$$

At this point, the Stokes parameters can be defined. First, several assumptions must be made. For example, considering no loss of coherence would factor out the term ωt, while considering a pair of orthogonal plane waves with their corresponding electric field vectors at a space point *z* = 0 would lead to dropping the term *kz*. Further, one can set δ = δ_y - δ_x so that the evolution with respect to time of Eq.4 can be expressed likewise [41]:

$$\frac{\langle E_x^2(t)\rangle}{E_{0x}^2} + \frac{\langle E_y^2(t)\rangle}{E_{0y}^2} - 2\frac{\langle E_x(t)\rangle}{E_{0x}}\frac{\langle E_y(t)\rangle}{E_{0y}}cos(\delta) = sin^2(\delta), \tag{5}$$

where, for monochromatic electro-magnetic waves, both the phase factors and the amplitudes are constant and do not depend on time. To remove completely the time dependence, the operator $\langle$ ... $\rangle$ denoting time averaging over a single period *T* of the oscillations has been applied. Next, by multiplying both sides of the above equation by the factor $4E_{0x}^2E_{0y}^2$ one gets:

$$4E_{0y}^2\langle E_x^2(t)\rangle + 4E_{0x}^2\langle E_y^2(t)\rangle - 8E_{0x}E_{0y}\langle E_x(t)E_y(t)\rangle cos(\delta) = [2E_{0x}E_{0y}sin(\delta)]^2 \tag{6}$$

In what follows, we use the following expressions for the time averages:

$$\langle E^2(t)\rangle = \frac{1}{2}E_0^2 \tag{7a}$$

$$\langle E_x(t)E_y(t)\rangle = \frac{1}{2}E_{0x}E_{0y}cos(\delta). \tag{7b}$$

Substituting Eq.7a and b into Eq.6 with the subsequent addition and subtraction of E_{0x}^4+E_{0y}^4 from the left-hand side yields and completes the derivation of the Stokes parameters:

$$(E_{0x}^2 + E_{0y}^2)^2 - (E_{0x}^2 - E_{0y}^2)^2 - [2E_{0x}E_{0y}cos(\delta)]^2 = [2E_{0x}E_{0y}sin(\delta)]^2 \tag{8a}$$

$$S_0^2 - S_1^2 - S_2^2 = S_3^2. \tag{8b}$$

Finally, we can relate the light intensity to the electric field amplitude:

$$I = \langle E_0^2\rangle. \tag{9}$$

In this way, it becomes possible to connect the Stokes parameters to real and measurable quantities, namely the light intensities of given polarization states. [40, 41]

$$\begin{bmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \end{bmatrix} = \begin{bmatrix} I_x + I_y \\ I_x - I_y \\ I_{+45} - I_{-45} \\ I_{RC} - I_{LC} \end{bmatrix} \tag{10}$$

Generally, light polarization is defined via the two orthogonal projections of the electric field vector. Following this convention, the term S_0 is the total intensity of light, S_1 indicates the difference between horizontal or vertical polarization, S_2 indicates the difference between $\pm 45^o$ polarization, while S_3 is that between right and left circular polarization. The equality in Eq.8b is valid only for total degree of polarization ρ = 1. In reality, especially when dealing with turbid media and tissue specimens, ρ is always less than unity. In that case the equality becomes inequality obeying the relation [40, 41]:

$$S_0^2 > S_1^2 + S_2^2 + S_3^2, \tag{11}$$

while ρ is defined from [40, 41]:

$$\rho = \frac{(S_1^2 + S_2^2 + S_3^2)^{1/2}}{S_0}. \tag{12}$$

The above-mentioned considerations justify completely the use of the Stokes parameters to describe and trace the light polarization when interacting with turbid media and biological samples, as it will be shown later.

1.2 Stokes-Mueller formalism

Instead of having dimensions of the Stokes parameters of intensity in *W/cm²*, it is often more convenient to work with normalized and dimensionless quantities. Generally, this is achieved by normalizing all Stokes parameters to S_0. In this way, a single column vector can be formed for a given polarization state. For instance, $\mathbf{S_H}$=(1,1,0,0)^T, $\mathbf{S_V}$=(1,-1,0,0)^T, $\mathbf{S_P}$=(1,0,1,0)^T and $\mathbf{S_{RC}}$=(1,0,0,1)^T, where the subscripts denote horizontal linear (H), vertical linear (V), +45 linear (P), and right-circular (RC) polarization states. On the other hand, each medium has a given polarization response **M**, which could be considered also as a transfer function. The input Stokes vector ($\mathbf{S_i}$) will be modified by the transfer function of the medium (**M**), producing an output Stokes vector ($\mathbf{S_0}$). This can be written as follows [40]:

$$\mathbf{S_0} = \mathbf{M} \cdot \mathbf{S_i}, \tag{13}$$

where **M** is a real 4x4 matrix containing all polarization and depolarization properties of the specimen under study. Knowledge of the Mueller matrix would be sufficient to reconstruct the output Stokes vectors for the aforementioned polarization states. To do so, one needs to solve the following system of four vector linear equations [42]:

$$\begin{cases} \mathbf{S^0_{H/V}} = \mathbf{M_{i1}} \pm \mathbf{M_{i1/2}} \\ \\ \mathbf{S^0_{P/RC}} = \mathbf{M_{i1}} + \mathbf{M_{i3/4}}, \end{cases} \tag{14}$$

From experimental point of view, it is of growing importance to interpret and link the Mueller matrix elements to real, physical properties of the sample under study. Such task generally requires knowledge of a physical model *a priori*, which is far from straightforward and may be even impossible. For instance, it is of high importance to extract and interpret the polarization fingerprint of an anisotropic turbid media, like biological specimens. This task will be addressed in detail in the following chapters of the dissertation. At this point, it is important to recall several main properties of the Mueller matrix. First, it is usually normalized to the total unpolarized light intensity or m_{11}, so that all elements are dimensionless. Second, the first row is mainly responsible for diattenuation, while the first column for polarizance, both of which being characterized by their corresponding vectors **D** and **P**, such as [41]:

$$\mathbf{D} = (m_{12}, m_{13}, m_{14}) \tag{15a}$$

$$\mathbf{P} = (m_{21}, m_{31}, m_{41})^T, \tag{15b}$$

and their scalar values can be obtained as [41]:

$$D = \frac{1}{m_{11}} \sqrt{\sum_j m_{1j}^2}, \quad j = 2,3,4, \quad 0 \leq D \leq 1 \tag{16a}$$

$$P = \frac{1}{m_{11}} \sqrt{\sum_i m_{i1}^2}, \quad i = 2,3,4, \quad 0 \leq P \leq 1 \tag{16b}$$

The remaining 3x3 submatrix **m**$'$ is accountable for both retardance and depolarization and a shortened notation can be utilized:

$$\mathbf{M} = \begin{bmatrix} 1 & \mathbf{D} \\ \mathbf{P} & \mathbf{m}' \end{bmatrix} \tag{17}$$

1.3 Poincaré sphere representation

The generic concept of the polarizaton ellipse is very useful to describe various polarization states for light with ρ = 1. Unfortunately, this representation would not be valid in the case of turbid media, where $\rho < 1$. For this particular reason, it is more suitable to adopt the Poincaré sphere representation [40, 41]. To do so, firstly the Cartesian coordinates of the Stokes parameters must be transformed into spherical coordinates, like [5]:

$$\mathbf{S} = I(1, \mathbf{p})^T \tag{18a}$$

$$\mathbf{p} = \rho \mathbf{u} \tag{18b}$$

$$\mathbf{u} = [cos(2\theta)cos(2\epsilon), sin(2\theta)cos(2\epsilon), sin(2\epsilon)], \tag{18c}$$

where in the equation above **p** and **u** are the polarization and the Poincaré vectors, while $\theta \in$ [-π/2; π/2] and $\epsilon \in$ [-π/4; π/4] denote respectively the azimuth and ellipticity angles. Expressed through the Stokes parameters, both angles are given by [40]:

$$\theta = \frac{1}{2} atan\left(\frac{S_2}{S_1}\right) \tag{19a}$$

$$\epsilon = \frac{1}{2} asin\left(\frac{S_3}{S_0}\right) \tag{19b}$$

Every single point on the surface of the Poincaré sphere represents a polarization states with ρ = 1. Inside the sphere, $\rho < 1$ and is 0 at the center. Briefly, all linear polarization states are located on the equator, whereas the two circular polarization states lie on the poles; elsewhere, the polarizations are elliptical. The upper hemisphere represents all states with ellipticity > 0, while the lower hemisphere represents all states with ellipticity < 0. In addition, antipodal points on the surface of the sphere have mutually orthogonal polarization states.

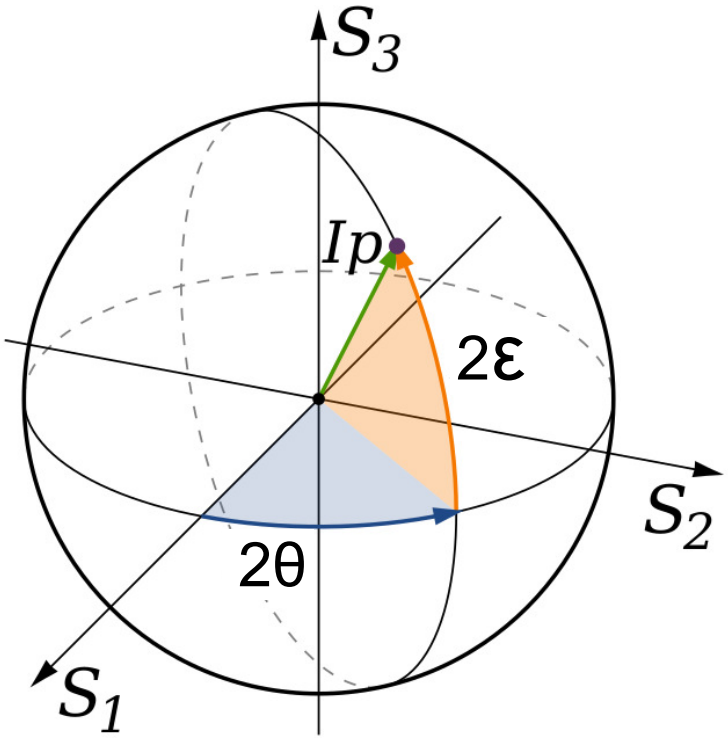


Figure 2: Poincaré sphere representation.
https://fr.m.wikipedia.org/wiki/Fichier:Poincar%C3%A9_sphere.svg

At this point it would be practical to evaluate and graphically represent the aforementioned polarization states, alongside their visualization on the surface of the Poincaré sphere. For simplicity, no loss of polarization degree was assumed and the following notations were used interchangeably: $E_x \equiv E_h$ - horizontal, $E_y \equiv E_v$ - vertical and *k* $\equiv$ *r* - for direction of propagation.

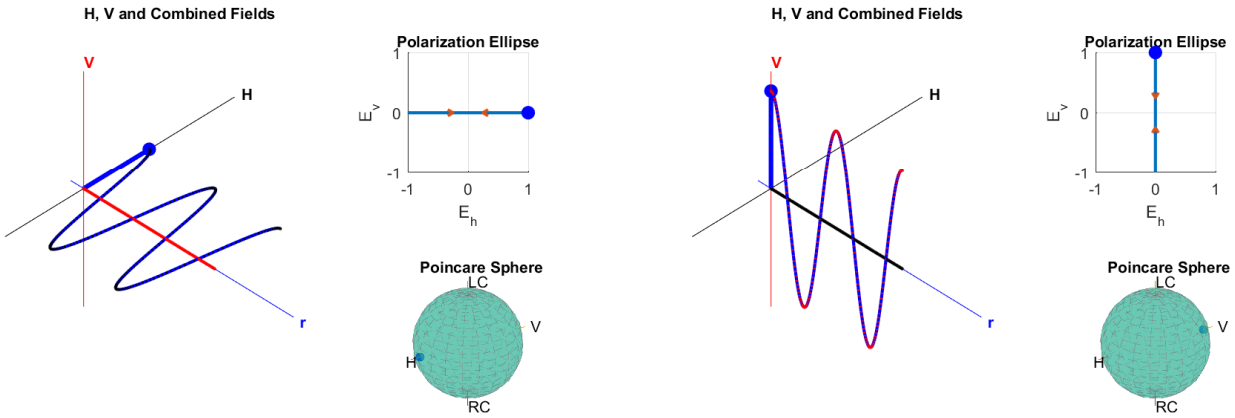


Figure 3: Graphical representation of different polarization states with $\rho = 1$: (a) horizontal, (b) vertical, (c) +45 and (d) left circular. All red lines represent the vertical component of the electric field vector, all black lines - the horizontal component, while the blue line is the resultant wave.
`https://fr.mathworks.com/help/phased/ug/modeling-and-analyzing-polarization.html`

1.4 Coherency matrix and Pauli spin matrices

Apart from the Stokes parameters representation of a given polarization state, there is another way to do so, namely, by using a 2×2 matrix, also known as the coherency matrix - **J**. There is a direct relationship between the elements of **J** and the Stokes parameters. First, one needs to define the following quantities forming the elements of the coherence matrix [5, 41]:

$$J_{ij} = \langle E_i E_j^* \rangle = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} E_i E_j^* \, dt \qquad (i, j = x, y) \tag{20}$$

The coherency matrix is Hermitian, so that $\mathbf{J^T} = \mathbf{J^*}$ and $J_{xy} = J_{yx}^*$. In this way, **J** can be defined by the equation below and its trace will represent the total light intensity [5, 41]:

$$\mathbf{J} = \begin{bmatrix} \langle E_x E_x^* \rangle & \langle E_x E_y^* \rangle \\ \langle E_y E_x^* \rangle & \langle E_y E_y^* \rangle \end{bmatrix} \tag{21}$$

The determinant of the coherency matrix can define the cases of complete and partial polarization if det**J** = 0 or det**J** > 0, while ρ reads [5, 41]:

$$\rho = \sqrt{1 - \frac{4 det \mathbf{J}}{(Tr\mathbf{J})^2}} \tag{22}$$

Finally, one can connect the Stokes vector elements to the elements of the coherency matrix as follows [5, 41]:

$$\begin{bmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \end{bmatrix} = \begin{bmatrix} J_{xx} + J_{yy} \\ J_{xx} - J_{yy} \\ J_{xy} + J_{yx} \\ i(J_{xy} - J_{yx}) \end{bmatrix} \tag{23}$$

Analogously, to provide a measure of how close a given polarization state is to the surface of the Poincaré sphere, both the polarization and the Poincaré vectors can be connected to the coherency matrix as [5, 41]:

$$\mathbf{J} = \begin{bmatrix} 1 + \rho u_1 & \rho(u_2 - i u_3) \\ \rho(u_2 + i u_3) & 1 - \rho u_1 \end{bmatrix} \tag{24}$$

Due to the fact that electromagnetic radiation carries spin-orbital momentum, there is a need to make the connection of that property of light with pure states of polarization by using the Pauli spin matrices in the following manner [5, 41]:

$$\mathbf{J} = \frac{1}{2} \sum_{i=1}^{4} \sigma_i S_i, \tag{25}$$

recalling, that:

$$\sigma_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{26a}$$

$$\sigma_2 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \tag{26b}$$

$$\sigma_3 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \tag{26c}$$

$$\sigma_4 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \tag{26d}$$

1.5 Physical realizability

In the previous sections it was assumed that the algebraic quantity **M** is indeed a Mueller matrix. In reality, not every real-valued 4x4 matrix can be considered as a Mueller matrix, since the input Stokes vector of the incident light that is modified by the transfer function should be transformed into an output Stokes vector with a degree of polarization less or equal to 1. Furthermore, physically realizable, depolarizing Mueller matrices must be representable as weighted averages of non-depolarizing Mueller matrices [5]:

$$\mathbf{M} = \sum_{i=1}^{4} \lambda_i \mathbf{M}_i \tag{27}$$

Due to measurement errors and noise, this condition may not be met. Such matrices need to undergo a physical realizability filtering, according to the method proposed by Cloude [53]. To do so, each experimental 4x4-matrix candidate for a Mueller matrix can be cast into the form of the so called covariance matrix **H**, a Hermitian matrix defined as [5]:

$$\mathbf{H} = \frac{1}{4} \sum_{i,j=1}^{4} m_{ij} (\sigma_i \otimes \sigma_j), \tag{28}$$

where $\sigma_\mathbf{i}$ are the four Pauli spin matrices and the symbol $\otimes$ denotes the Kronecker product. The eigenvalues of the matrix **H** are usually arranged in decreasing order ($\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \lambda_4$). If there is at least one negative eigenvalue, in that case the experimental matrix is not a Mueller matrix and needs to be filtered by adopting the physical realizability concept. To do so, all negative eigenvalues are set to 0, while their corresponding eigenvectors remain unchanged. Then, the matrices Λ and **V** are comprising the filtered eigenvalues and the eigenvectors $\mathbf{V}_\mathrm{i}$ of **H**, are constructed as:

$$\Lambda = diag(\lambda_i) \tag{29a}$$

$$\mathbf{V} = [\mathbf{V_1}, \mathbf{V_2}, \mathbf{V_3}, \mathbf{V_4}] \tag{29b}$$

The closest possible covariance matrix $\mathbf{H_f}$, originating from a physically realizable Mueller matrix is obtained from the relation [5]:

$$\mathbf{H_f} = \mathbf{V}\Lambda\mathbf{V}^{-1}. \tag{30}$$

To complete the physical realizability, every single element of the filtered Mueller matrix can be reconstructed from [5]:

$$m_{ij}^{f} = Tr[(\sigma_i \otimes \sigma_j)\mathbf{H_f}]. \tag{31}$$

From here on, the physical interpretation and the extraction of the polarimetric properties of a given sample will be directly performed on the filtered Mueller matrix $\mathbf{M_f}$ and its corresponding $\mathbf{H_f}$, both of which for brevity sake will be denoted as **M** and **H**, taking into account the physical realizability filtering as described above.

1.6 Decomposition algorithms

After introducing the concept and the importance of the physical realizability, one can proceed towards the extraction of the polarimetric fingerprint of the samples under examination. Once done, the polarimetric properties can be related to the sample inner structure and morphology. To do so, a distinction has to be made between changing the polarization state and changing the polarization degree of the probing light beam by the sample. For biomedical samples, the former is a result of surface and volume effects, while the latter is due to spatial or temporal coherence loss. Most often, it is more practical to focus only on the polarimetric properties, by setting tr(**H**) = 1 or simply normalizing the eigenvalues sum to unit. In this way, both values of transmissivity and reflectivity can be disregarded. Then, depending on the experimental configuration and the physical model, if any, the most suitable decomposition algorithm can be chosen. In this subsection, all decomposition algorithms used in the scope of the dissertation are presented in detail.

1.6.1 Polarization purity and polarization entropy

The first decomposition algorithm to discuss is the parallel or the sum decomposition. Eq.27 compactly summarizes its essence of a weighted average representation of the covariance matrix, the eigenvalues λ_i serving as weighting coefficients. Depolarization information can be extracted from **H**, which can be related to the depolarization properties of the underlying medium and its Mueller matrix **M**. Three depolarization indicators, called indices of polarization purity, can be simply derived from λ_i(**H**), assuming that tr(**H**) = 1 and $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \lambda_4$ [5]:

$$P_1 = \lambda_1 - \lambda_2, \tag{32a}$$

$$P_2 = \lambda_1 + \lambda_2 - 2\lambda_3, \tag{32b}$$

$$P_3 = \lambda_1 + \lambda_2 + \lambda_3 - 3\lambda_4. \tag{32c}$$

The overall depolarization ability, called depolarization index P_Δ and the polarization purity PI are defined as [5]:

$$P_\Delta = \sqrt{\frac{1}{3}(2P_1^2 + \frac{2}{3}P_2^2 + \frac{1}{3}P_3^2)}, \tag{33a}$$

$$PI = \sqrt{\frac{1}{3}(P_1^2 + P_2^2 + P_3^2)}. \tag{33b}$$

From Eq. 32 and Eq. 33 two limiting cases could be identified: pure non-depolarizing media, when P_i = P_Δ = PI = 1 and pure depolarizing media, when P_i = P_Δ = PI = 0. Yet, even more information is encoded in **H** via the Cloude's entropy S, which is related to the spatial heterogeneity of a given sample of interest [5]:

$$S = -\sum_{i=1}^{4} \lambda_i log_4(\lambda_i). \tag{34}$$

Unlike P_i, P_Δ and PI, S = 1 would lead to an assumption of heterogeneous inner structure, responsible for a complete randomization of the input light polarization state(s). This corresponds to an ideal depolarizer (i.e. λ_1=λ_2=λ_3=λ_4) and the respective turbid medium exhibits polarimetric properties that fluctuate completely at random. On the contrary, S = 0 presumes homogeneous inner structure, indicative for a complete preservation of ρ for fully polarized light.

1.6.2 Differential decomposition

Let us now consider a propagation of a polarized light beam along the z-axis of a Cartesian coordinate system. The medium is highly anisotropic and introduces both changes in the polarization state and degree. Also, the medium is assumed to be transversally homogeneous (along x,y) and longitudinally inhomogeneous (along z [54]). Following these concepts, the differential matrix **m** relates the Mueller matrix **M** of the anisotropic medium as follows:

$$\frac{d\mathbf{M}}{dz} = \mathbf{mM}. \tag{35}$$

To schematically represent the initial assumptions, the following figure is shown:

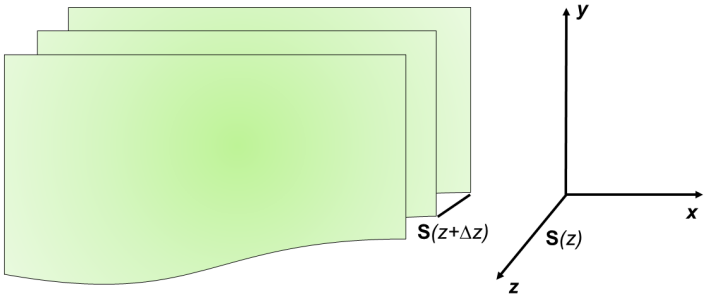


Figure 4: Schematic representation of polarized light propagation in an anisotropic medium along the *z*-axis of a Cartesian coordinate system. The medium is transversally homogeneous (along *x,y*) and longitudinally inhomogeneous (along *z*).

According to the differential equation in Eq.35, the initial Mueller matrix will be modified by factor of **m**. This factor is another 4x4 matrix, containing the mean values and the mean squared deviations of the polarization properties [55]:

$$\mathbf{m} = \mathbf{m_m} + \mathbf{m_u} = \langle \mathbf{m} \rangle + \langle \Delta \mathbf{m^2} \rangle z, \tag{36}$$

where the spatial averaging $\langle\rangle$ is in the transverse plane. After rearranging, substituting **m** and integrating with respect to the boundary conditions, Eq.35 is found to admit the solution:

$$ln\mathbf{M} = \langle \mathbf{m} \rangle z + \frac{1}{2} \langle \Delta \mathbf{m}^2 \rangle z^2 \tag{37}$$

The matrix logarithm can be obtained by first solving the eigenvalue-eigenvector problem for **M** and then forming **D** = diag[ln(d_i)], where d_i are the eigenvalues of **M**. Finally, using the matrix **V** comprising the eigenvectors of **M**, the matrix logarithm and the two counterparts of **m** can be calculated as [55]:

$$\mathbf{L} \equiv ln\mathbf{M} = \mathbf{V D V^{-1}}, \tag{38a}$$

$$\mathbf{m_m} = \frac{1}{2}(\mathbf{L - G L^T G}), \tag{38b}$$

$$\mathbf{m_u} = \frac{1}{2}(\mathbf{L + G L^T G}), \tag{38c}$$

where **G** = diag(1,-1,-1,-1) is the Minkowski metric tensor. For better clarity, both the polarizing and the depolarizing parts of the differential matrix can be summarized as [55]:

$$\mathbf{m_m} = \begin{bmatrix} 0 & p_1 & p_2 & p_3 \\ p_1 & 0 & p_6 & p_5 \\ p_2 & -p_6 & 0 & p_4 \\ p_3 & -p_5 & -p_4 & 0 \end{bmatrix}, \tag{39a}$$

$$\mathbf{m_u} = \begin{bmatrix} \alpha & d_1 & d_2 & d_3 \\ -d_1 & \alpha_1 & d_6 & d_5 \\ -d_2 & d_6 & \alpha_2 & d_4 \\ -d_3 & d_5 & d_4 & \alpha_3 \end{bmatrix}, \tag{39b}$$

where the notations used denote the elementary polarization properties as follows: p_1 - linear dichroism along *x-y* axis (LD), p_2 - linear dichroism along $\pm 45^o$ axis (LD'), p_3 - circular dichroism (CD), p_4 - linear birefringence along *x-y* axis (LB), p_5 - linear birefringence along $\pm 45^o$ axis (LB') and p_6 - circular birefrigence (CB'). For a depolarizing medium, the off-diagonal elements of $\mathbf{m_m}$ represent the mean values of the six elementary properties. For the depolarizing part $\mathbf{m_u}$, the off-diagonal elements d_i are the uncertainties of the six elementary properties, while the main diagonal contains the depolarizing factors α_i (provided that α = 0). Again, α_1 is along the *x-y* axis, α_2 is along the $\pm 45^o$ axis and α_3 is the circular component. Often, α_1, α_2 and α_3 are referred as anisotropic absorptions and are consequently denoted as LA, LA' and CA, while α_0 is the isotropic component [55]. By using Eq.40, one can obtain the net scalar retardance ϕ_t, the total depolarization α_t (provided that α = 0), the net dichroism Δd and the orientation θ [55] of the optic axis:

$$\phi_t = \sqrt{p_6^2 + p_5^2 + p_4^2}, \tag{40a}$$

$$\alpha_t = \frac{1}{3}|\alpha_1 + \alpha_2 + \alpha_3|, \tag{40b}$$

$$\Delta d = \sqrt{p_1^2 + p_2^2 + p_3^2} \tag{40c}$$

$$\theta = \frac{1}{2} tan^{-1}\left[\frac{p_5}{p_6}\right]. \tag{40d}$$

In the above equation, the depolarization factors are bound within the interval (–∞, 0] for physically realizable differential Mueller matrices, while the factor of 1/2 in the orientation definition is relating to a physical azimuth in the absence of circular component. If one wants to relate to an orientation angle for the Poincaré sphere representation, then the factor of 1/2 should be omitted.

The proposed decomposition algorithm can be adopted for use when modelling the propagation of polarized light into an ansiotropic medium. Nevertheless, it is very important to compare results with samples with similar thicknesses, since Eq.37 shows that the polarization properties scale linearly with sample thickness, while the depolarization ones evolve quadratically. In reality, especially when dealing with biological samples, the thickness is always varying from one sample to another and the aforementioned considerations are not met. In order to avoid this obstacle, thickness invariant parameters must be implemented. Analogously to Eq.35, the Beer-Lambert law has also an exponential solution [39]:

$$dI = -I\mu_t z, \tag{41a}$$

$$I(z) = I_0 e^{-\mu_t z}, \tag{41b}$$

where *I* denotes the light intensity and μ_t is the total attenuation coefficient which includes both scattering and absorption. In practice, the latter is again in most cases unknown, but can be evaluated approximately from:

$$ln\left[\frac{I(z)}{I_0}\right] = -\mu_t z \approx ln(m_{11}). \tag{42}$$

In this way, m_{11} can be obtained from the non-normalized **M** and each polarimetric quantity from Eq.40a through Eq.40c can be normalized, thus working with and considering only thickness-invariant parameters. All polarization quantities must be normalized to $\ln(m_{11})$, while all depolarizing ones, to $\ln^2(m_{11})$, as follows by referring back to Eq.37.

1.6.3 Lu-Chipman decomposition

A widespread in the polarimetric community decomposition algorithm, first proposed by Lu and Chipmanin, can also be used accordingly both in transmission and in reflection geometry, after having complied with the physical realizability. Its representation is based on a matrix factorization into a sequence of the Mueller matrices of a diattenuator $\mathbf{M_D}$, a retarder $\mathbf{M_R}$ and a depolarizer $\mathbf{M}_\Delta$ (a.k.a. forward decomposition; reversed decomposition also exists, but the present manuscript deals only with the former) [56]:

$$\mathbf{M = M}_\Delta\mathbf{M_R M_D}. \tag{43}$$

For better clarity, it is convenient to adopt a partitioned form for all product matrices in Eq.43. i.e. their general form reads:

$$\mathbf{M_D} = \begin{bmatrix} 1 & \mathbf{D}^\mathsf{T} \\ \mathbf{D} & \mathbf{m_D} \end{bmatrix}, \tag{44a}$$

$$\mathbf{M_R} = \begin{bmatrix} 1 & \vec{\mathbf{0}}^\mathbf{T} \\ \vec{\mathbf{0}} & \mathbf{m_R} \end{bmatrix}, \tag{44b}$$

$$\mathbf{M}_\Delta = \begin{bmatrix} 1 & \vec{\mathbf{0}}^\mathbf{T} \\ \boldsymbol{P} & \mathbf{m}_\Delta \end{bmatrix}. \tag{44c}$$

The first step is to find the diattenuation matrix from the diattenuation vector:

$$\mathbf{m_D} = \sqrt{1-\mathbf{D}^2}\mathbf{I} + (1-\sqrt{1-\mathbf{D}^2})\hat{\mathbf{D}}\hat{\mathbf{D}}^\mathbf{T}, \tag{45}$$

where **I** is 3x3 identity matrix and $\hat{\mathbf{D}}$ is the unit vector along **D**. Knowledge of the diattenuation matrix can facilitate the completion of the decomposition method by calculating:

$$\mathbf{M M_D^{-1}} = \mathbf{M}_\Delta\mathbf{M_R} = \begin{bmatrix} 1 & \vec{\mathbf{0}}^\mathbf{T} \\ \boldsymbol{P} & \mathbf{m}_\Delta\mathbf{m_R} \end{bmatrix} = \mathbf{M}^{'}. \tag{46}$$

Next, the 4x4 $\mathbf{M}'$ matrix decomposition reduces to a 3x3 matrix decomposition, as the polarizance vector is already known from Eq.15b. What is left is $\mathbf{m}' = \mathbf{m}_\Delta \mathbf{m}_\mathbf{R}$. Let λ_i be the eigenvalues of $\mathbf{m}'(\mathbf{m}')^\mathsf{T}$. Then, $\mathbf{m}_\Delta$ is obtained from [56]:

$$\mathbf{m}_\Delta = \pm[\mathbf{m}'(\mathbf{m}')^\mathsf{T} + (\sqrt{\lambda_1\lambda_2} + \sqrt{\lambda_2\lambda_3} + \sqrt{\lambda_3\lambda_1})\mathbf{I}]^{-1} \times [(\sqrt{\lambda_1} + \sqrt{\lambda_2} + \sqrt{\lambda_3})\mathbf{m}'(\mathbf{m}')^\mathsf{T} + \sqrt{\lambda_1\lambda_2\lambda_3}\mathbf{I}]. \tag{47}$$

The sign is given by the sign of the determinant of $\mathbf{m}'$. Knowledge of the polarizance vector and $\mathbf{m}_\Delta$ is sufficient to reconstruct $\mathbf{M}_\Delta$. Finally, $\mathbf{M_R}$ is obtained from:

$$\mathbf{M_R} = \mathbf{M}_\Delta^{-1}\mathbf{M}' \tag{48}$$

After the completion of this decomposition algorithm, the net scalar retardance, the net depolarization and the orientation of the optic axis can be found from [56]:

$$\phi = cos^{-1}\left[\frac{tr(\mathbf{M_R})}{2} - 1\right], \tag{49a}$$

$$\Delta = 1 - \frac{|m_\Delta(1,1)| + |m_\Delta(2,2)| + |m_\Delta(3,3)|}{3}, \tag{49b}$$

$$\theta = \frac{1}{2}tan^{-1}\left[\frac{m_R(2,4)}{m_R(4,3)}\right]. \tag{49c}$$

In this particular decomposition algorithm, the depolarization Δ is in the interval [0, 1] and, analogously, the physical azimuth interpretation of the orientation is meaningful in the absence of circular retardance only. Attention should be paid when comparing the depolarization parameters from the differential and the Lu-Chipman decompositions. For instance, one can simply calculate the following expression of α_t:

$$\Delta_t = 1 - \frac{1}{3}(e^{\alpha_1} + e^{\alpha_2} + e^{\alpha_3}), \tag{50}$$

in order to make it comparable to Lu-Chipman's Δ. To illustrate graphically the assumption on the discrete sequence of polarization elements characterized by their corresponding Mueller matrices, the following figure is shown:

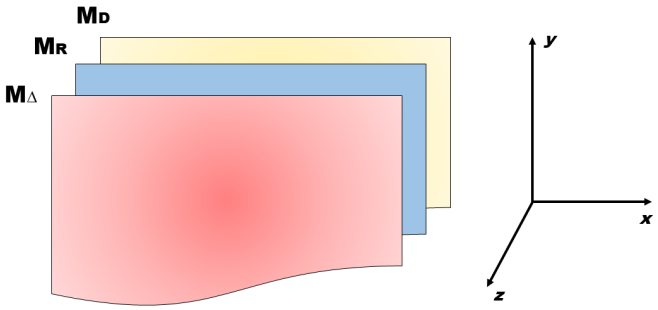


Figure 5: Schematic representation of polarized light propagation in anisitropic medium along the *z*-axis of a Cartesian coordinate system, following sequence of diattenuator, retarder and depolarizer.

The convenience of this decomposition algorithm is that it can be applied for both transmission and reflection geometries. However, it assumes discrete order or occurance of the polarimetric responses, characterized by the matrix factorization in Eq.43. Also, it is important whether the forward or reversed decomposition is chosen, as this will determine if the depolarizer will include polarizance or diattenuation respectively.

1.6.4 Symmetric decomposition

The interpretation of depolarizing systems and samples has been extensively studied either with Lu-Chipman or differential decompositions. The former may exhibit forward and reverse forms, thus yielding two asymmetric depolarizers containing either polarizance or diattenuation. On the other hand, the latter assumes transversally homogeneous and longitudinally inhomogeneous anisotropic medium with continuous distribution of all optical features throughout the sample volume. Such a condition might not be met, due to macroscopic variations of the refractive index, to the highly anisotropic structure of bio-tissues or, simply, to the use of backscattering configuration whereby the probing light follows two, forward and backward, paths. Furthermore, a large variety of samples requires the implementation of angular-resolved measurements which typically involve pure depolarizers without polarizance and diattenuation. Hence, an arbitrary **M** can be decomposed into the so called symmetric factorization in such a way so that the canonical diagonal depolarizer is placed between pairs of diattenuators and retarders [57]:

$$\mathbf{M} = \mathbf{M}_{\mathbf{D}_2}\mathbf{M}_{\mathbf{R}_2}\mathbf{M}_{\triangle}\mathbf{M}_{\mathbf{R}_1}^{\mathbf{T}}\mathbf{M}_{\mathbf{D}_1}. \tag{51}$$

For better clarity, it is convenient to adopt a partitioned form for all product matrices in Eq.52, i.e. their general form reads:

$$\mathbf{M_D} = \begin{bmatrix} 1 & \mathbf{D}^{\mathsf{T}} \\ \mathbf{D} & \mathbf{m_D} \end{bmatrix}, \tag{52a}$$

$$\mathbf{M_R} = \begin{bmatrix} 1 & \vec{\mathbf{0}}^{\mathbf{T}} \\ \vec{\mathbf{0}} & \mathbf{m_R} \end{bmatrix}, \tag{52b}$$

$$\mathbf{M}_\Delta = diag(1, d_1, d_2, d_3), \tag{52c}$$

where $\mathbf{m_D}$ and $\mathbf{m_R}$ are 3x3 submatrices. The former is constructed from the diattinuation vector in 45 and the latter is a rotatory matrix. On the other hand, all d_i are termed as the principal depolarization factors for 1 - linear, horizontal or vertical, 2 - linear $\pm\ 45^o$ and 3 - right/left circular polarization states. The first step to the symmetric decomposition would be to find the diattenuation matrices by solving the eigenvector-eigenvalue problem of [57]:

$$\mathbf{(M^T GMG)}\xi_{\mathbf{1}} = \beta^2 \xi_{\mathbf{1}} \tag{53a}$$

$$\mathbf{(MGM^T G)}\xi_{\mathbf{2}} = \beta^2 \xi_{\mathbf{2}}, \tag{53b}$$

where $\mathbf{G}$=*diag*(1,-1,-1,-1) is the Minkowski metric tensor and β^2 is a common eigenvalue. When the eigenvectors $\xi_{\mathbf{1,2}} = \mathbf{(1, D_{1,2})^T}$ are found, then the diattenuation vectors can be used to obtain $\mathbf{m_{D_i}}$ and $\mathbf{M_{D_i}}$:

$$\mathbf{m_D} = \sqrt{1 - \mathbf{D}^2}\mathbf{I} + (1 - \sqrt{1 - \mathbf{D}^2})\hat{\mathbf{D}}\hat{\mathbf{D}}^{\mathbf{T}}, \tag{54}$$

where $\mathbf{I}$ is 3x3 identity matrix and $\hat{\mathbf{D}}$ is the unit vector along $\vec{\mathbf{D}}$. Once the diattenuation matrices are determined, one calculates:

$$\mathbf{M_{D_2}^{-1} M M_{D_1}^{-1}} = \mathbf{M_{R_2} M_\Delta M_{R_1}^T} = \mathbf{M}' = \begin{bmatrix} \beta & \vec{\mathbf{0}}^{\mathbf{T}} \\ \vec{\mathbf{0}} & \mathbf{m}' \end{bmatrix}. \tag{55}$$

Since $\mathbf{M}'$ and $\mathbf{M}_\Delta$ contain no diattenuation and polarizance, by virtue of the singular value decomposition (SVD) the 3x3 sub-matrix $\mathbf{m}'$ can be obtained, which will be sufficient to construct the retarder matrices $\mathbf{M_{R_{1,2}}}$ and the canonical depolarizer matrix $\mathbf{M}_\Delta$, thus completing the symmetric decomposition algorithm. After this step, it becomes possible to calculate the retardance and the net depolarization values from [57]:

$$\phi = \cos^{-1}\left[\frac{tr(\mathbf{m_R}) - 1}{2}\right], \tag{56a}$$

$$\Delta = 1 - \frac{|d_1| + |d_2| + |d_3|}{3}, \quad 0 \leq \Delta \leq 1. \tag{56b}$$

Up to this point, it was shown the sequence to obtain all products of the symmetric decomposition.

As for the physical interpretation of the symmetric decomposition applied to backscattering configuration measurements, it is to be noted that the forward propagating photons and the backscattered ones will have different polarization responses, due to the light interaction with the turbid media under study. Their separate contributions can be evaluated from the corresponding

matrix products $\mathbf{M_{R_1}M_{D_1}}$ and $\mathbf{M_{R_2}M_{D_2}}$ respectively. Thus, the diattenuation and the retardance are responsible for the changes in the initial probing light polarization. In both cases, due to multiple scattering events and a coherence loss, a depolarization will be introduced as well, characterized by the canonical depolarizer matrix $\mathbf{M}_\Delta$, which will not be "contaminated" by diattenuation or polarizance (wherefrom termed "pure"). In this way, the symmetric decomposition appears as a good choice for interpreting backscattering and, more generally, angular-resolved, measurements.

1.7 Depolarization spaces

Often, it is useful to graphically represent the depolarization properties of a given sample. To do so, one needs to implement the so called depolarization space. Based on the already presented Mueller calculus, it is practical to introduce extrinsic and intrinsic depolarization parameters of a given **M** both of which define corresponding depolarization spaces. The extrinsic parameters and spaces are described by the eigenvalues of the covariance matrix **H** associated with **M**, whereas the intrinsic parameters and spaces are defined by the eigenvalues of the covariance matrix $\mathbf{H}_\Delta$ of the canonical depolarizer $\mathbf{M}_\Delta$. In the absence of diattenuation, the matrices of both diattenuators in the symmetric decomposition are equal to the identity matrix and the extrinsic and intrinsic depolarization parameters are equal, and therefore the extrinsic and intrinsic spaces converge. The diattenuators have the capability of altering the degree of polarization of partially polarized light by introducing repolarizing effects and thus, making the difference between the two sets of depolarization parameters and spaces. To sum up, the depolarization spaces can be grouped into three categories: *canonical*, *natural* and *IPP* [5].

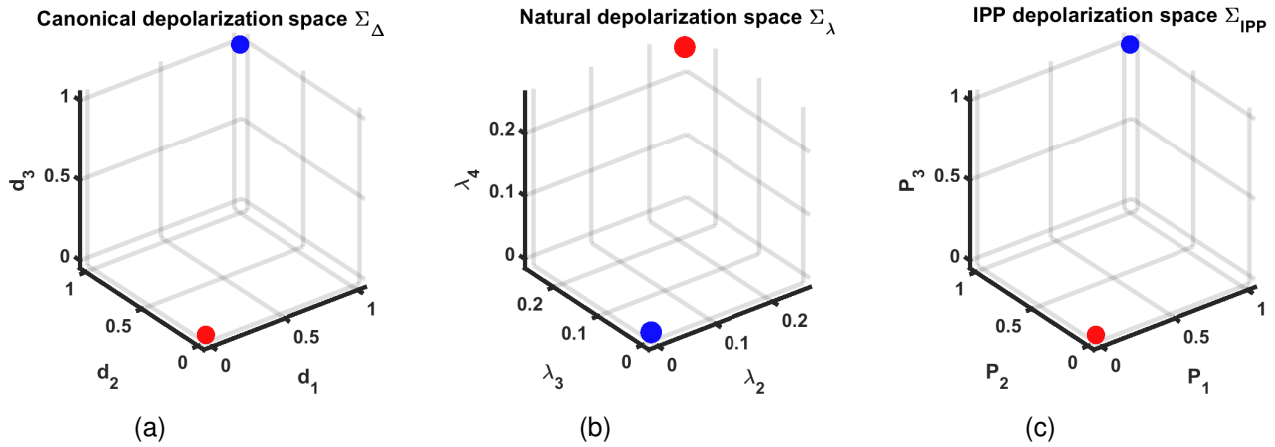


Figure 6: 3D representation of: (a) Canonical, (b) Natural, (c) Indices of polarimetric purity depolarization spaces, where the symbols represent ● – pure polarizer, ● – pure depolarizer.

As can be seen from the figure above, the three depolarization spaces consist of the following components: the *canonical* one includes the three depolarization factors from the symmetric decomposition d_1, d_2 and d_3, which can be regarded as weight coefficients to the Stokes parameters S_i; the *natural* space includes the last three eigenvalues from the covariance matrix **H** λ_2,

λ_3 and λ_4; the *IPP* space includes the three indices of polarimetric purity P_1, P_2 and P_3. Based on the values of all these parameters, two extreme cases can be illustrated, depending on the depolarization capabilities of the corresponding Mueller matrix: a nondepolarizing matrix, which preserves ρ at 1 and pure depolarizer in which case ρ = 0. The former case holds, when d_i and also P_i = 1, while λ_i = 0. On the other hand, the latter case holds when d_i and also P_i = 0, while λ_i = 0.25. (This specific value of λ_i comes from the normalizing factor of 1/4 from Eq.28).

General conclusions

Up to this point, we have presented and discussed the mathematical foundations and the most important theoretical aspects for describing polarized light and its interaction with biomedical samples. Hence, significant attention is to be paid on the algebraic model, which depends on the experimental geometry and the type of samples to be measured. Consequently, the model choice leads to the selection of the most suitable decomposition algorithm to extract the polarimetric fingerprint of the measured samples. To draw a parallel between all decomposition algorithms, one typically needs first to assess whether there are significant diattenuation and retardance as polarimetric effects from the samples. If no, then using only the parallel docomposition would be sufficient, thus focusing only on the depolarization properties. In this way, two depolarization spaces (natural and IPP) and net parameters such as PI, P_Δ and S would be enough for the analysis. On the contrary, if at least the retardance introduced by the samples is non-negligible, then another decomposition is needed, depending on the experimental geometry and the initial assumptions. Following the same logic, in case the diattenuation cannot be neglected, then another depolarization space could be adopted (i.e. canonical), thus establishing a difference between extrinsic and intrinsic depolarization properties.

Historically, the Lu-Chipman decomposition was the first one to be proposed and used, and is considered as a state-of-the-art algorithm for both transmission and reflection measurements. Later, it was partially superseded by the extension of the Azzam's differential calculus to the depolarizing case by the differential decomposition, as well as by the symmetric decomposition. The former is better suited for transmission measurements, while the latter, for angular-resolved ones. In both the differential and symmetric decompositions, treatment of the depolarization is performed without inclusion of diattenuation or polarizance, while additional parameters can be extracted as well to enrich the polarimetric analysis, such as L/CB, L/CD, α_i, φ_i, d_i, Δ and D_i, as we shall see later.

2 Polarimetric instrumentation

Polarimetric measurements may have different applications, depending on the samples used and the information desired to be retrieved from them. From here on, one needs to make several important choices concerning the polarimetric prerequisites, such as: proper wavelength selection, choice of light source (i.e. laser polarimetry, spectropolarimetry or via using low coherent light sources), single shot or scanning measurements, Stokes vector or Mueller matrix domain, transmission or reflection configuration, imaging or non-imaging methods and etc. As there is no perfect experimental system and calibration method, each specific choice for the above-mentioned criteria would have its own advantages and limitations.

In the scope of the PhD thesis, all measurements were conducted with low coherent light sources, in order to avoid the formation of laser speckle, with discrete wavelength selection. Two different configurations were adopted: in transmission and in reflection. For the former, imaging of the corresponding regions of interest (ROIs) was sought, while for the latter the main priority was to scan two-dimensionally the ROIs, the detector being a single photodiode without imaging capabilty. While the first case has at avail each Mueller matrix as an image or array with dimensions corresponding to the pixels count of the camera, the second option provides each Mueller matrix element as a scalar. Nevertheless, the basic principle is the same in both instances - shine the sample with known input polarization states; project each of the transmitted/reflected signal independently on the same set of polarization states and reconstruct the full Mueller matrix. Wherever imaging is used, medical doctors could superimpose their images from conventional microscopy in transmission to the polarimetric ones (also obtained in transmission) and, in this way, have another optical marker related to the health condition. On the other hand, this could be at the price of larger data sets to process and analyze, increased computational time and additional inclusion of image processing algorithms. With single-shot measurements, these drawbacks can be overcome. It makes possible scanning and measuring faster the desired ROIs, while non-conventional imaging can be performed by interpolating between the data points from the scanned zones. Such an approach might be more useful in reflection geometry where the backscattered photons are considered as diagnostically relevant.

Additionally, two different techniques for Mueller matrix measurements were used. The first one measures the Mueller matrix directly whereas the second one obtains it indirectly from measured Stokes vectors. Generally, only complete Stokes-Mueller polarimeters were considered, i.e. polarimeters capable of generating and analyzing at least four independent polarization states of light. In the next sections, all of the aforementioned specifications are described in detail, including information on the calibration methods used.

2.1 Eigenvalue calibration method

The Mueller matrix formalism becomes very important and necessary to implement, whenever the samples are partially or totally depolarizing and/or they possess rough surfaces, with significantly increased scattering than absorption. It is of great importance to accurately calibrate the optical system, in order to minimize experimental errors and to extract only the Mueller matrix of the sample under study, by factoring out the Mueller matrix of the instrument. Usually, the calibration of a polarimetric device is a two-step process. First, a precise orientation and alignment of the optical elements comprising the optical set-up is required. Second, the use of reference or calibrating samples is obligatory. A lot of difficulties and uncertainties may arise at those two stages, due to imperfections in the polarization optical elements, as well as in the reference ones, introduction of alterations in the polarization state from the image forming optical elements, presence of dust, vibrations, etc. The eigenvalue calibration method, shortly denoted as ECM, has three main advantages: $i)$ the polarization system is complete and no other assumptions are necessary; $ii)$ the orientations and positions of the optical elements of the instrument do not have to be precisely adjusted, thus making this calibration method only dependent on the reference samples used; $iii)$ after performing the calibration, the matrices of the polarization state generator (PSG) **W** and the polarization state analyser (PSA) **A**, along with the Mueller matrix of the sample **M** can be evaluated, and the ECM can be performed for each desired wavelength. The algorithm has been reported earlier in [58] and its description is provided below in detail.

Let us now consider a sequence of four input polarization states of light, characterized by their normalized (S_0=1) Stokes vectors : $\mathbf{S_{Hi}}$=(1,1,0,0)^T, $\mathbf{S_{Vi}}$=(1,-1,0,0)^T, $\mathbf{S_{Pi}}$=(1,0,1,0)^T and $\mathbf{S_{RCi}}$=(1,0,0,1)^T, where the subscript i denotes input and the superscript T denotes transposition. Then the **W** and the **A** matrices, describing the PSG and the PSA responses respectively, are given by:

$$\mathbf{W} = [\mathbf{S_{Hi}}, \mathbf{S_{Vi}}, \mathbf{S_{Pi}}, \mathbf{S_{RCi}}], \tag{57a}$$

$$\mathbf{A} = \mathbf{W}^{\mathsf{T}}. \tag{57b}$$

The first step is to measure the experimental matrix of air $\mathbf{B}_0$, which has an identity Mueller matrix $\mathbf{M}_{\text{air}} = \mathbf{I}$. This could be mathematically represented as:

$$\mathbf{B}_0 = \mathbf{A}\ \mathbf{M_{air}}\ \mathbf{W}, \tag{58a}$$

$$\mathbf{A} = \mathbf{B}_0\mathbf{W}^{-1}. \tag{58b}$$

The calibration can be performed with the use of three reference samples (i = [1,3]), such as horizontal/vertical polarizer and a quarter-wave plate. Similarly, Eq.58a can be used for each i:

$$\mathbf{B}_{\text{i}} = \mathbf{A}\ \mathbf{M_i}\ \mathbf{W} \tag{59}$$

The Mueller matrices of the calibrating samples are given by:

$$\mathbf{M_i}(\tau_i, \psi_i, \phi_i, \theta_i) = \tau_i \mathbf{R}(\theta_i) \begin{bmatrix} 1 & -cos(2\psi_i) & 0 & 0 \\ -cos(2\psi_i) & 1 & 0 & 0 \\ 0 & 0 & sin(2\psi_i)cos(\phi_i) & sin(2\psi_i)sin(\phi_i) \\ 0 & 0 & -sin(2\psi_i)sin(\phi_i) & sin(2\psi_i)cos(\phi_i) \end{bmatrix} \mathbf{R}(-\theta_i), \quad (60)$$

where τ_i is the sample transmission coefficient for the totally depolarized fraction of light, ψ_i is an angle describing the linear dichroism introduced by the sample, ϕ_i is the retardance of the sample and θ_i is the rotation angle of the optic axis. For each $\mathbf{M_i}$, the respective eigenvalues are rotationally invariant, i.e. they do not depend on θ_i:

$$\lambda_{1i} = 2\tau_i cos^2(\psi_i), \quad (61a)$$

$$\lambda_{2i} = 2\tau_i sin^2(\psi_i), \quad (61b)$$

$$\lambda_{3i} = \tau_i cos(2\psi_i) e^{i\phi_i}, \quad (61c)$$

$$\lambda_{4i} = \tau_i cos(2\psi_i) e^{-i\phi_i}, \quad (61d)$$

$$\frac{\lambda_{1i}\lambda_{2i}}{\lambda_{3i}\lambda_{4i}} = 1. \quad (61e)$$

Afterwards, the elements of $\mathbf{M_i}$ can be calculated:

$$\tau_i = 0.5(\lambda_{1i} + \lambda_{2i}), \quad (62a)$$

$$cos(2\psi_i) = \frac{\lambda_{1i} - \lambda_{2i}}{\lambda_{1i} + \lambda_{2i}}, \quad (62b)$$

$$sin(2\psi_i)cos(\phi_i) = \frac{\lambda_{3i} + \lambda_{4i}}{\lambda_{1i} + \lambda_{2i}}, \quad (62c)$$

$$sin(2\psi_i)sin(\phi_i) = -\frac{\lambda_{3i} - \lambda_{4i}}{\lambda_{1i} + \lambda_{2i}}. \quad (62d)$$

Then, another matrix $\mathbf{C_i}$ is defined in accordance with:

$$\mathbf{C_i} = \mathbf{B_0}^{-1}\mathbf{B_i} = \mathbf{W}^{-1}\mathbf{M}\,\mathbf{W}, \quad (63)$$

$\mathbf{C_i}$ has the same eigenvalues as $\mathbf{M_i}$. The latter can be thus calculated, except for the orientation angle θ_i. Additionally, one can write:

$$\mathbf{M_i W I} - \mathbf{I\,W C_i} = 0. \quad (64)$$

At this point, it becomes necessary to construct the error matrix $\mathbf{E}$, which is a real, 16 x 16 matrix containing all experimental errors. To do so, first needs to vectorize $\mathbf{W}$ to $\mathbf{W_v}$, which is now of dimensions 16 x 1:

$$(\mathbf{I} \otimes \mathbf{M_i})\mathbf{W_v} - (\mathbf{C_i}^{\mathsf{T}} \otimes \mathbf{I})\mathbf{W_v} = 0, \quad (65a)$$

$$(\mathbf{I} \otimes \mathbf{M}_i) - (\mathbf{C}_i^T \otimes \mathbf{I}) = \mathbf{E}_i, \tag{65b}$$

$$\mathbf{E}_i \mathbf{W}_v = \mathbf{0}. \tag{65c}$$

Before completing the calibration algorithm, the symmetric non-negative-definite **K** matrix is calculated:

$$\mathbf{K} = \sum_{i=1}^{3} \mathbf{E}_i^T \mathbf{E}_i. \tag{66}$$

and Eq.65c is rewritten as:

$$\mathbf{K}\ \mathbf{W_v} = \mathbf{0}, \tag{67}$$

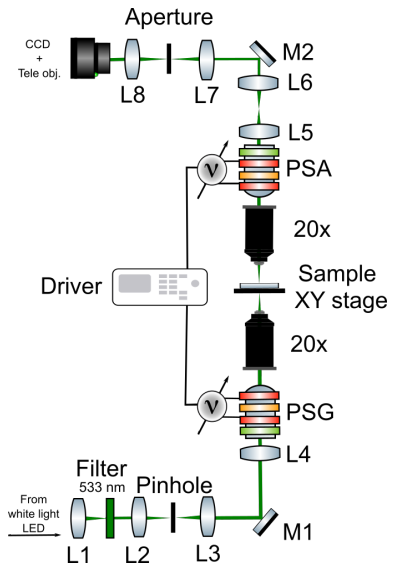
where the eigenvalues of **K** have the following property: $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq ... \geq \lambda_{16} \approx 0$. Hence, the last eigenvalue, in the ideal case equal to 0, has a corresponding eigenvector $\mathbf{W_v}$, with no experimental errors. Then, from $\mathbf{W_v}$ with dimension 16 x 1 **W** is reconstructed, now with dimension 4 x 4. Finally, by recalling Eq.58b one can find also **A**. The optimal solution to find **W** is to minimize $\lambda_{16}/\lambda_{15} << 1$ by changing the rotation angle θ_i. Experimentally, three types of calibrating samples were used: horizontal/vertical polarizers and a quarter-wave plate with the optic axis oriented at 30^o. Their schematic representation is shown in Fig.7.

(a) (b) (c) (d)

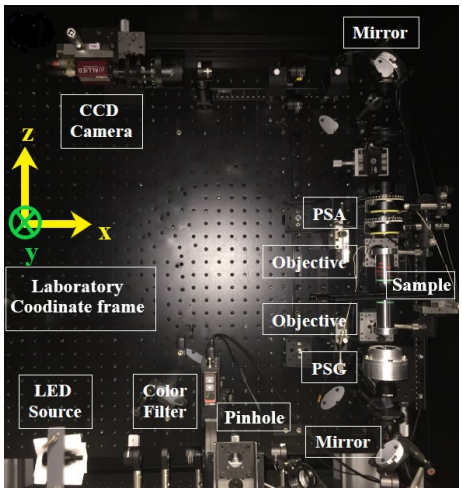
Figure 7: Schematic representation of the calibrating samples: a) horizontal polarizer, b) vertical polarizer (the red line shows the transmission axes), c) quarter-wave plate with fast axis (blue line), oriented at 30^o with respect to the x - axis, d) laboratory co-ordinate frame, light propagation assumed along z - axis.

2.2 Transmission geometry

In this experimental configuration, the method to measure the full Mueller matrix is based on using four input and output polarization states, enabling one to obtain 16 polarization combinations. For each combination the light intensity will be different and these intensity values can be used to solve 16 linear equations, in order to reconstruct all elements of the experimental Mueller matrix (i.e. see [43]). Based on this principle, a custom-built Mueller polarimeter was used to operate in transmission. The layout of the instrument is shown in the side view in Fig.8.



(a)



(b)

Figure 8: Experimental set-up in transmission geometry: (a) Schematic representation and (b) Side view. Figure available from the published material in [59].

This experimental set-up has several advantages. For instance, the CCD camera used as a detector allows to image the same region of interest as the one from the histology analysis,

provided by the medical doctors from digital microscopy. By merging both images into a single array, both the polarization response and the ground truth health condition can be mapped at each pixel location. Also, more measurement counts are available when having an array of pixels instead of having single measurements with a single photodiode. Additionally, there are no moving parts (i.e. wave plates, color filters, etc.), only liquid crystal-based polarizing elements, as well as there are no optical elements displacing the optical beam (i.e. Fresnel rhombs). Nevertheless, a potential drawback is the spectral region of operation, as liquid crystal-based polarizing elements are efficient in the visible range, but under-perform in the UV range. Consequently, all measurements were performed in the visible spectrum only.

The experimental set-up description is as follows: A white-light LED was chosen as a light source, whereas a color filter at the wavelength 533 nm (FWHM 15 nm) was selected to filter spectrally the probing light beam. A μm in diameter spatial filters (pinholes) can be inserted in the illumination arm to simultaneously control the direction and the angular aperture of the light beam. The oblique angle of incidence θ_{AOI} at which the light illuminates the sample depends both on the off-axis distance between the center of the pinhole and the optical axis of the microscope d_{OA}, as well as on the focal length of the objective lense f_{OL}:

$$\theta_{AOI} = asin\left(\frac{d_{OA}}{f_{OL}}\right). \tag{68}$$

If the pinhole is precisely aligned to the optical axis of the objective, then the sample will be illuminated at normal incidence. Additionally, the illuminating beam divergence θ_{div} can be expressed as:

$$\theta_{div} = \frac{d_{PH}}{cos(\theta_{AOI}) f_{OL}}, \tag{69}$$

where d_{PH} is the pinhole diameter. In accordance with the Koehler configuration, L3 serves as a condenser lens in order to illuminate the sample with uniform intensity and polarization. Both the PSG and the PSA contain identical optical elements however, arranged in reverse order. In short, they include a linear polarizer, two ferroelectric liquid crystal retarders (Meadowlark FPR-200-1550) and a quarter-wave retarder placed between the two retarders as shown in Fig. 9.
The transmission axis of the polarizer coincides with the x-axis of the laboratory co-ordinate frame. FLC1 serves as quarter-wave plate, while FLC2 serves as half-wave plate with fast axis oriented at an angle with respect to the x - axis of the laboratory co-ordinate frame. The orientation of the fast optic axis of either of the FLC in the plane orthogonal to the direction of light propagation (assumed along z - axis) can be modified by applying a given voltage V. Let θ_1, θ_2 and θ_{WPG} be the orientation angles of the two FLCs and the wave plate from the PSG, with respect to the x - axis. Analogously, similar notation can be applied for the PSA and its elements, θ_3, θ_4 and θ_{WPA}, where, to avoid confusion with the PSG, FLC2 is denoted as FLC3, while FLC1 as FLC4. Taking as a reference the laboratory coordinate frame, the angles of the optical axes of the PSA elements are related to the angles of the optical axes of the PSG one as follows:

$$\theta_1 = 180 - \theta_4, \tag{70a}$$

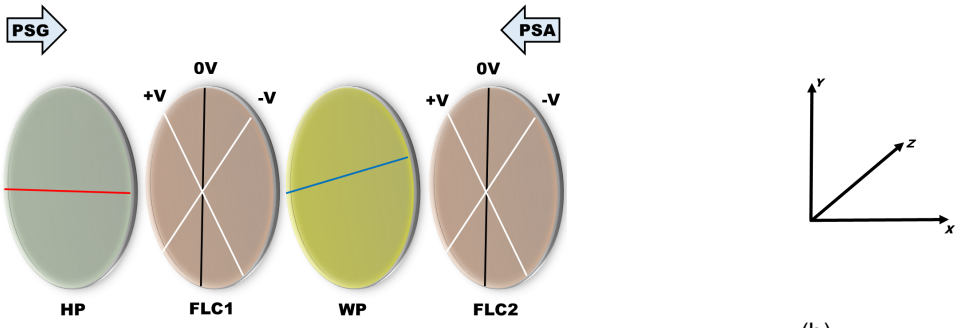


Figure 9: Schematic representation of the components of the PSG and the PSA: a) HP: horizontal, polarizer; FLC1/2: ferroelectric liquid crystal retarders; WP, waveplate, and b) laboratory co-ordinate frame, light propagation assumed along z - axis.

$$\theta_2 = 180 - \theta_3, \tag{70b}$$

$$\theta_{QWG} = 180 - \theta_{QWA}. \tag{70c}$$

It is important to note the ratios between the operating times τ_i of the voltage-modulated FLCs:

$$\tau_1 = \frac{\tau_{FLC1}}{\tau_{FLC4}} = \frac{4}{1}, \tag{71a}$$

$$\tau_2 = \frac{\tau_{FLC2}}{\tau_{FLC3}} = \frac{4}{1}, \tag{71b}$$

$$\tau_3 = \frac{\tau_{FLC1}}{\tau_{FLC2}} = \frac{1}{2}, \tag{71c}$$

$$\tau_4 = \frac{\tau_{FLC3}}{\tau_{FLC4}} = \frac{2}{1} \tag{71d}$$

After the probing light interacts with the sample, the scattered light is collected by another objective lens and with the help of L4, and the microscopic image can be formed in the real plane. The microscope objective lenses can be selected so as to provide different magnifications (i.e., 50x, 20x, 10x and 5x), depending on the sample specifications, field of view, spatial resolution and the numerical aperture which contributes to light depolarization. For that reason, the condenser and the imaging optics have always been kept identical in order to match their respective numerical apertures. On the other hand, the projection of the light beam to the Bertrand lens L6 allows to automatically switch reversibly between the real and the Fourier plane. In the real plane imaging mode, the instrument provides spatial images of the studied samples and their structures, while by switching into the Fourier imaging mode, the instrument provides images of the angular distribution of the light transmitted and/or scattered by the sample. Also, in this way, more precise optical alignment was achieved during the calibration. To monitor the image preview and to capture the images of interest, a telephoto lens was coupled and set to infinity to matrix photodetector (16-bit, single-channel, CCD camera, 600 x 800 pixels, AV Stingray F-080B). In order to calibrate or obtain the Mueller matrix of the samples under study, 4 input and 4 output polarization states are required. The voltage modulation of the optical

elements in the PSG and the PSA ensures this function. All 16 combinations of polarization states result in 16 intensities to be measured, hence capturing 16 intensity images. Afterwards, by virtue of the method proposed in reference [43] the Mueller matrix elements can be obtained. Before conducting the measurements, the Mueller microscope was calibrated adopting the eigenvalue calibration method, as described in detail in [58]. During the measurements, the use of various samples with different thicknesses caused variations in the intensity detected by the CCD camera. To avoid over- or underexposed intensity measurements, the experiments were conducted in the linear range of the detector, by varying the exposure time for each sample while the gain was always kept at minimum. Additionally, the same ambient temperature in the laboratory was kept during calibration and the measurements in order to reduce the electronic noise and keep it at a constant level.

In order to increase the accuracy and the precision of the experimental data, the condition numbers of both PSG, PSA (described correspondingly by **W, A**) must be optimized. In this way, the signal-to-noise ratio is increased [60]. The condition number of an arbitrary matrix **A** is given by:

$$k(\mathbf{A}) = ||\mathbf{A}^{-1}||||\mathbf{A}||, \tag{72}$$

where $||\quad||$ represents a matrix norm and a_{ij} are the elements of **A**. Next, assume both **W** and **A** instrument matrices are known, and $\mathbf{I_{int}}$ is the experimental intensity matrix of an arbitrary sample. Then, the following equation yields the Mueller matrix of the sample:

$$\mathbf{M} = \mathbf{A}^{-1}\mathbf{I_{int}}\mathbf{W}^{-1} \tag{73}$$

The latter equation can be vectorized like:

$$\mathbf{M_v} = (\mathbf{W}^{T} \otimes \mathbf{A})^{-1}\mathbf{I_{int_v}} = \mathbf{Q}^{-1}\mathbf{I_{int_v}} \tag{74}$$

From the equation above follows that if one wants to minimise the errors in **M** then the condition number of **Q** needs to be kept at minimum as well. This can be also expressed as [61]:

$$k_2(\mathbf{Q}) = k_2(\mathbf{W})k_2(\mathbf{A}), \tag{75}$$

where k_2 indicates the Euclidean norm. Hence, there is a need to optimize the condition numbers of both **W** and **A** by a proper selection of the optical elements comprising the PSG and the PSA parts of the optical system [62]. The Euclidean norms of the PSG and the PSA are lower bounded by $\sqrt{3}$ as the columns and rows of **W** and **A** respectively are Stokes vectors [60]. To obtain the minimum values of the condition numbers or, equivalently, the maximum values of $1/k_2$, the following optimized values for the orientation angle of the optic axis and the retardance values of the optical elements of the PSG and the PSA are summarized in Table1 below:

ϕ_{FLC1}	ϕ_{WPG}	ϕ_{FLC2}	ϕ_{FLC3}	ϕ_{WPA}	ϕ_{FLC4}
90	90	180	180	90	90
θ_{FLC1}	θ_{WPG}	θ_{FLC2}	θ_{FLC3}	θ_{WPA}	θ_{FLC4}
51.25	154.70	82.52	128.75	25.30	97.48

Table 1: Retardance and orientation angle values (in degrees) of the optical elements in the optimized PSG and PSA. For the ferro-electric liquid crystals the retardance values are at the wavelength of 510 nm, while for the two wave plates, at 633 nm.

With these sets of retardance and orientation angle values, the $1/k_2$ values for the PSG and the PSA are 0.55, 0.52 at 500 nm and 0.38, 0.35 at 700 nm. Later on, as it will be described in the experimental part, these two wavelenghts will be selected for the measurements in transmission.

2.3 Reflection geometry

In this configuration, it is feasible to measure the backscattered photons. In our case, the detector being a single photodiode, two-dimensional scans were performed whenever necessary, in order to study the (de)polarization response of the samples within their corresponding regions of interest. To retrieve the Mueller matrix, four input polarization states were used, described by their normalized (S_0=1) Stokes vectors: $\mathbf{S_{Hi}}$=(1,1,0,0)^T, $\mathbf{S_{Vi}}$=(1,-1,0,0)^T, $\mathbf{S_{Pi}}$=(1,0,1,0)^T and $\mathbf{S_{RCi}}$=(1,0,0,1)^T. The measured output Stokes vectors (also normalized) are: $\mathbf{S_{Ho}}$, $\mathbf{S_{Vo}}$, $\mathbf{S_{Po}}$ and $\mathbf{S_{RCo}}$. The subscripts denote horizontal linear (H), vertical linear (V), +45 linear (P), and right-circular (RC) polarization states; 'i' stands for the input states, 'o' for the output states and the superscript T is for transposition. The input and output Stokes vectors are connected by the relation $\mathbf{S_o}=\mathbf{M}\cdot\mathbf{S_i}$, where **M** is the Mueller matrix of the sample under measurement. An important relation between Stokes vector elements accounting for the presence of depolarization is given by:

$$S_0 > (S_1^2 + S_2^2 + S_3^2)^{1/2} = \rho \cdot S_0. \tag{76}$$

By knowing the total degree of polarization ρ and the detected signal power $\wp$, the normalized output Stokes vectors reads:

$$\mathbf{S}_{\text{dep}} = \wp \cdot \rho \cdot (S_0 \cdot \rho^{-1}, S_1, S_2, S_3)^T. \tag{77}$$

The four output Stokes vectors $\mathbf{S}_{\text{dep}}$ can be represented as algebraic sums of the columns $\mathbf{C}_l$ (l=1, ..., 4) of the sample Mueller matrix as follows [42]:

$$\mathbf{S^H_{dep}} = \mathbf{C}_1 + \mathbf{C}_2, \tag{78a}$$

$$\mathbf{S^V_{dep}} = \mathbf{C}_1 - \mathbf{C}_2, \tag{78b}$$

$$\mathbf{S^P_{dep}} = \mathbf{C}_1 + \mathbf{C}_3, \tag{78c}$$

$$\mathbf{S^{RC}_{dep}} = \mathbf{C}_1 + \mathbf{C}_4, \tag{78d}$$

where all superscripts in Eq.(78) denote the respective input polarization states. By measuring the output Stokes vector for each one of the four different input polarization states, all columns of the Mueller matrix $\mathbf{C}_l$ and therefore, the complete Mueller matrix, can be derived [42]:

$$\mathbf{C}_1 = \frac{\mathbf{S^H_{dep}} + \mathbf{S^V_{dep}}}{2}, \tag{79a}$$

$$\mathbf{C_2} = \frac{\mathbf{S^H_{dep}} - \mathbf{S^V_{dep}}}{2}, \tag{79b}$$

$$\mathbf{C_3} = \mathbf{S^P_{dep}} - \mathbf{C}_1, \tag{79c}$$

$$\mathbf{C_4} = \mathbf{S^{RC}_{dep}} - \mathbf{C}_1. \tag{79d}$$

As previously discussed, to measure the full Mueller matrix, at least sixteen independent measurements have to be conducted with different polarization states generated by both the polarization state generator (PSG) and analyzed by the polarization state analyzer (PSD) channels of the polarimetric equipment. Instead of using discrete polarization states, one may use continuous polarization modulation, typically achieved either with liquid crystal retarders or with rotating wave-plates. In this configuration, the PSG polarization states were set discretely, while those of the PSA were modulated continuously with a rotating quarter-wave plate, thus reducing the number of discrete measurements to only four. In Fig.10 the PSG and PSA channels are presented in a side view for better clarity. All optical elements were inserted in tube systems in order to block stray light that could interfere with the probing or detected beams. The field of view (FoV) of the system was found to be 100×100 μm. The angles of incidence and detection were respectively set to $55°$ and $30°$.

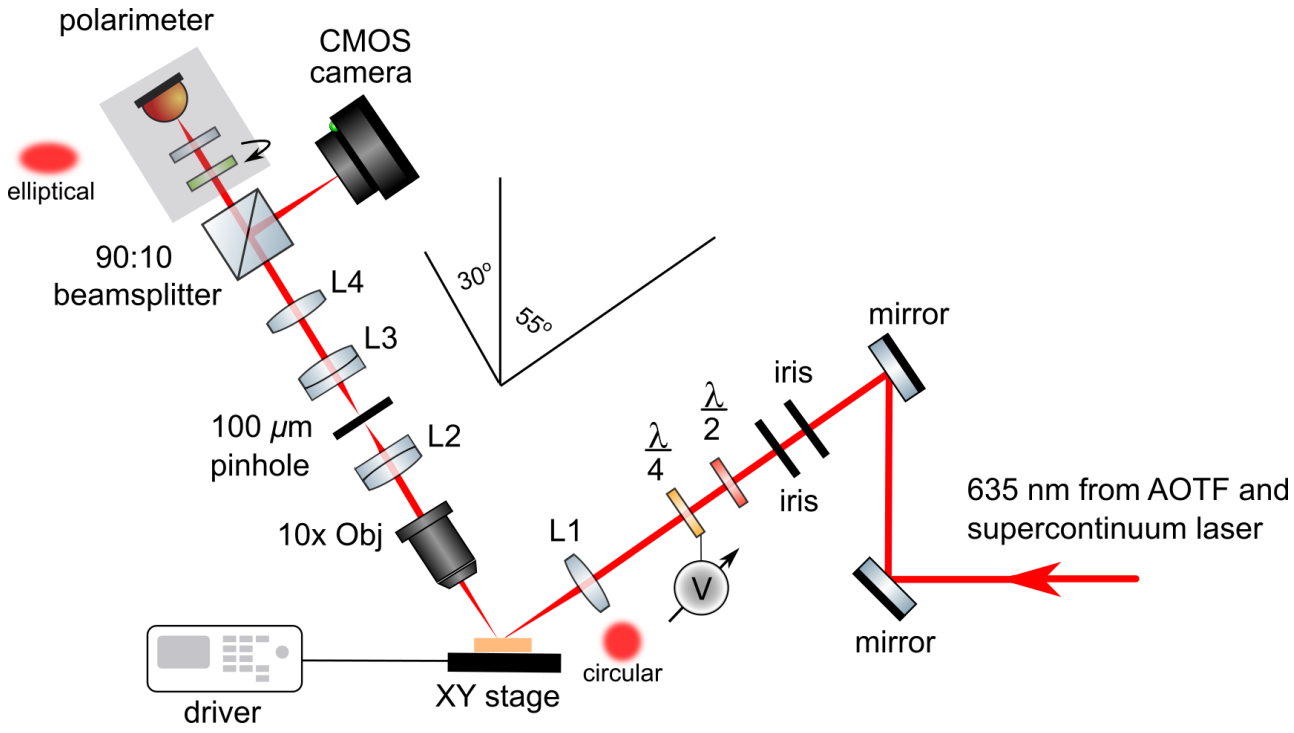


Figure 10: Schematic representation of the experimental setup. Figure available from the published material in [63]. Special acknowledgments for the figure preparation to Dr. Viktor Dremin.

Such oblique reflection configuration was found to enrich the polarimetric information obtained from the measurements, as the off-diagonal Mueller matrix (MM) elements are non-zero and therefore, the final form of the MM is different from diagonal depolarizer [63]. For instance, in [64] Pierangelo et al. used normal detection for the backscattered photons and the experimental MM at 600 nm bares strong resemblance to diagonal depolarizer. Additionally, Nishizawa et al. have also reported valuable angular optimizations for tissue polarimetric measurements in [65]. The light source, a supercontinuum fiber laser SC (Leukos Ltd., France), connected to an acousto-optic tunable filter AOTF (Leukos Ltd., France), generated the selected measurement wavelength of 635 nm with a spectral width of 8 nm and output power of 2 mW. Two irises were used to

obtain a collimated beam. A half-wave plate was inserted to vary the azimuth of the linearly polarized laser beam, while an electrically-driven liquid crystal variable quarter-wave plate was used to generate circular polarization. By means of the lens L_1 the light beam was focused to a particular depth of the sample volume. Objective lenses with $10\times$ magnification collected the diffusely scattered light, whereas the lens L_2, the 100 μm pinhole and the lens L_3 were employed to eliminate any out-of-focus photons. By using another lens, L_4, the light was focused onto a 90-10 beam splitter; the reflected beam was detected by a CMOS camera for accurate focus adjustment while the transmitted beam was analyzed by the polarimetric device (Thorlabs Ltd., USA). A rotating quarter-wave plate and a fixed linear polarizer inside the polarimetric device continuously modulate the light beam before it reaches the photo-detector *Si* photodiode. The commercially available polarimetric device was initially calibrated by the manufacturer. Prior to conducting the experiments with the biological specimen, the optical set-up was tested by measuring a mirror MM ($\mathbf{M}_{\text{exp}}$), which was compared to the theoretical MM in reflection ($\mathbf{M}_{\text{th}}$) without any depolarization properties, i.e. diag(1,1,1,1). Afterwards, a calculation of the rooot mean square error (*RMSE*) for each matrix element, except for m_{11} was performed.

$$\mathbf{M_{th}} = diag(1,1,1,1), \qquad \mathbf{M_{exp}} = \begin{bmatrix} 1 & 0.024 & 0.039 & -0.020 \\ 0.010 & 0.938 & -0.002 & 0.004 \\ -0.006 & 0.002 & 0.978 & 0.002 \\ -0.003 & 0.002 & 0.006 & 1.013 \end{bmatrix}, \qquad \text{RMSE} = 0.022. \tag{80}$$

By using a motorized translation stage, all samples were scanned independently over a region of interest (ROI) of 1 mm^2 and a step size of 0.2 mm in x-y directions. For each one of the four input polarization state measurements, the degree of polarization ρ of the measured output Stokes vector was calculated and a normalization procedure was applied (see Eq. (77)).

General conclusions

It is of great importance to possess the ability to measure different biological specimens, varying in type, thickness and other specifications. In the case of thin histological slides deposited on glass substrates, the transmitted amount of light is usually sufficient enough to be measured by the photodetector. In that case, the polarization response of the ROIs evaluated by the medical doctors can be extracted with meticulous calibration and substrate contribution removal. In this way, only the Mueller matrix of the sample is obtained, and each matrix element or polarimetric parameter from the selected decomposition can be represented as an image. Next, the final images must be post-processed to select, wherever applicable, only particular ROIs and thus, to avoid redundant features. Otherwise, the amount of data will increase with the number of measurements, and its processing and interpretation will become computationally ineffective.

Conversely, when thick tissue samples are to be measured, reflection geometry in which the focus lies on the detection of the backscattered photons, is better suited. Immediately, another difficulty arises, namely to obtain an image in reflection geometry from a particular depth in the tissue. To compensate this, spatial filtering is applied and, further, if a single photodiode is used as detector, 2D scanning is performed. Eventually, the data points from the scanned ROIs can be used to interpolate between them and to construct non-conventional images of the 2D spatial distribution of a particular polarimetric parameter.

Finally, the measurement of a complete Mueller matrix with an accurately calibrated Mueller matrix microscope provides all the necessary information in terms of polarization and depolarization from the measured samples. Moreover, the Stokes parameters for horizontal, vertical, $\pm 45^o$ and R/L circular polarizations can be obtained from a physically realizable Mueller matrix. They can be used for the purpose of complementary tissue diagnosis. Conversely, it is not always possible to have a Mueller polarimeter at the laboratory. In that case, the use of Stokes polarimeter, supplemented with a polarization state generator (PSG), is enough to measure indirectly Mueller matrices and to apply subsequently the required algorithms for physical realizability and decompositions.

3 Image processing algorithms

Throughout this work, only gray-scale images were processed and only monoscpectral polarimetric response was considered. The native spatial resolution of the imaging sensor was 600 x 800 pixels. The temporal resolution was disregarded as the work was concentrated on discrete measurements and images, rather than continuous capturing and imaging. The bit resolution was 2^{16} = 65 536 discrete gray-level values. Depending on the field of view and the samples choice, different magnifications were achieved (in units μm per pixel).

In short, the image formation can be summarized mathematically by a model, like [66]:

$$I = PSF * OF + N, \tag{81}$$

where *I* is the captured image, *PSF* is the point spread function, *OF* is the object function and *N* is additive noise, while * denotes convolution. Basically, *PSF* describes the way in which the *OF* is recorded from the imaging system to form the image *I*. *PSF* is a deterministic function that depends on the imaging system and is also dependent on the noise *N*. So, in practice it is "impossible" to image the true scene, characterized by its *OF*, but instead the closest possible representation of it is sought. In fact, the *OF* contains the full information of how the light is reflected, transmitted or even depolarized from the imaged sample towards the imaging sensor. Unlike the *PSF*, the noise function *N* is stochastic by nature. It can be described in terms of statistical noise distribution. Most commonly, the random variation of the image signal is approximated by a Gaussian function, considered as the most commonly used noise model. There are many sources of noise which cannot be avoided, but in the best case scenario these can be disregarded if there is a sufficient signal-to-noise ratio. For example, variations in the light source intensity, the image sensor temperature, presence of dust, vibrations, lens distortions, etc. can be potential reasons for noise presence in the final image. In short, the FWHM of a given image is representative of the signal noise (known as the electron noise), while $\sqrt{N_{pix}}$ (N_{pix} being the number of pixels) is representative of the pixel noise (known as the photon noise).

3.1 Substrate contribution removal

Most often, the biomedical samples are deposited on different types of substrates (glass, quartz, etc.). Ideally, the substrate would have no polarization response, but in practice there might be some residual one. When thin films are analyzed, their polarization response can be comparable to the unwanted polarization response of the substrate or, at least, the substrate may slightly shift the polarimetric values which is undesired for accurate diagnostics and must be avoided. To overcome this obstacle, the following algorithm was proposed. After the calibration, the bare substrate was measured. Then any measurements with a deposited biological sample and a substrate would yield a cumulative polarization response. To filter to substrate contribution,

the algorithm operates prior to normalizing the experimental 4x4 matrix and proceeding to the physical realizability. The filtration itself is performed by using the following metric:

$$RMSE(x,y) = \sqrt{\sum_{i=1}^{n=16} \frac{[\hat{I}_i(x,y) - I_i(x,y)]^2}{n}}, \tag{82}$$

where *RMSE* stands for root-mean-squared error, the index i runs over the experimental matrix elements, $\hat{I}_i$ represents the sample-with-substrate measurement, while I_i is the bare substrate measurement. As each matrix element is an image with size 600x800 pixels, the above-mentioned metric is calculated for each pixel of the imaging sensor. An *a priori* threshold value t is assigned, depending on how soft or hard the filtration is desired to be. Then, if *RMSE(x,y)* $> t$, the pixel value is preserved since the difference in the intensity images between the sample and the substrate is significant. On the contrary, if *RMSE(x,y)* $<= t$, then the pixel value is replaced by not-a-number value (NaN) because the difference in the intensity images between the sample and the substrate is not significant enough, hence assuming a non-removable substrate contribution. In this way, although at the price of increased computational time, the polarization response of the substrate can be filtered with a relatively simple metric. From here on, the intensity matrix can be normalized and fed towards the physical realizability, before decomposing the resulting Mueller matrix.

3.2 Image segmentation

Very often, the information in some zones of the whole image may be redundant or superfluous. Thus, a smaller region of interest (ROI) may be of greater interest. It may happen that, in a single image, different ROIs may have different polarimetric properties, e.g. corresponding to two different health conditions, or one may have a ROI with only NaNs (Not-a-Number) after filtration. Hence, it is important, whenever necessary, to split the initial image into multiple ROIs with similar polarimetric properties, and to omit the non-informative ones containing only NaNs. In the present work, the image segmentation was achieved by splitting the images into patches with smaller numbers of pixels in both directions. An example of such kind of splitting is shown in Fig.11, where each patch is of size 100 x 100 pixels, thus splitting the initial image into 48 patches, while keeping their corresponding number labelling column-wise. As discussed in Chapter2, the optical system has been modified to implement the Koehler illumination and thus, to ensure as even as possible light distribution throughout the whole image. However, this may not be always possible, because of the different numerical apertures of the objective lenses and the different fields of view chosen. Hence, a circular mask for removing the zones with low light intensity (according to a prescribed threshold value) at the side zones of the image may have to be applied. Again, as in the case in the previous subsection, the pixels values from the filtered zones are set to NaNs.

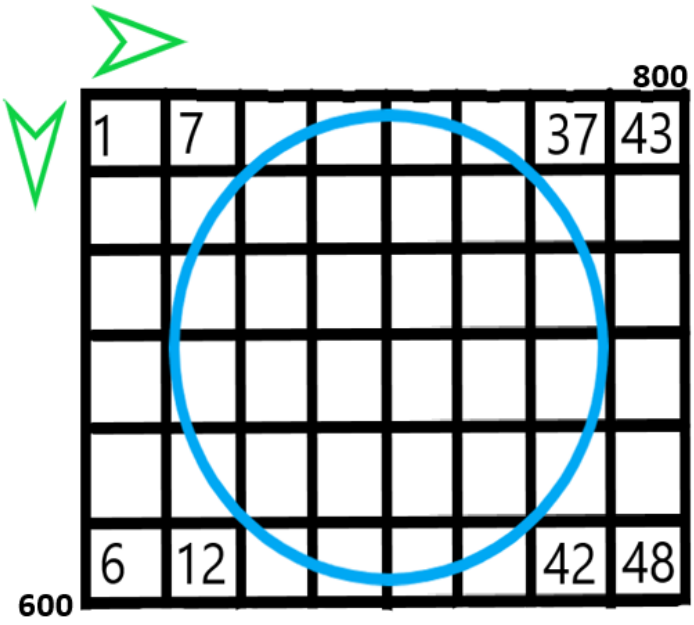


Figure 11: Schematic representation of the image segmentation.

3.3 Histogram normalization

Typically, after all filtration and segmentation procedures on the image have been completed, a histogram can be constructed so as to represent the distribution of the parameter of interest. However, when comparing polarimetric parameters from different images, the statistical moments characterizing the distributions will be different. Without appropriate normalization, the comparison between the distributions may bias the final diagnosis conclusions. For this particular reason, all histograms in the experimental part were normalized to a probability density function. Thus, the *x*-axis represents the value of the polarimetric quantity, while the *y*-axis describes its density D_y:

$$D_{y_i} = \frac{C_i}{N_{pix} \cdot w}, \tag{83}$$

where C_i is the number of counts for the i-th bin, N_{pix} is the total number of pixel values, omitting all NaNs, while w is the bin width, chosen *a priori* . Next, the i-th bin and the sum of all probabilities for the distribution are given by:

$$p_i = D_{y_i} \cdot w, \quad p_t = \sum_{i=1}^{N_{pix}} p_i = 1. \tag{84}$$

If one plots $D_y = f(X)$, where X is a given polarimetric parameter, this would represent the probability density function $\mathscr{P}$ of X. Then, the sum over all probabilities in the continuous case

will be given by the (normalized) area under the curve $\mathscr{A}$:

$$\mathscr{A} = \int_{X_{min}}^{X_{max}} \mathscr{P}(X)dX = 1. \tag{85}$$

3.4 Entropy and Standard deviation

To further develop the data analysis and the image segmentation techniques, various metrics for quantitative evaluation of the different patches can be adopted. Firstly, the entropy H and the standard deviation σ per each patch j can be adopted and computed, respectively. These metrics are calculated as follows:

$$H_j(\theta) = -\sum_{x,y} D_j^{x,y} log_2(D_j^{x,y}), \tag{86a}$$

$$\sigma_j(\theta) = \sqrt{\frac{\sum_{x,y}(\theta_j^{x,y} - \bar{\theta}_j^{x,y})^2}{N_{pix} - 1}}, \tag{86b}$$

where D contains the normalized histogram counts, while θ is an arbitrary parameter in this case. These metrics can be applied for any polarimetric quantity from the selected decomposition. Special attention should be paid to the entropy, as a built-in MATLAB function was used to calculate it. This function takes as an input a gray scale image. Hence, apart from the intensity images, Eq.86a can be applied to any other polarimetric quantity if its image has been converted to gray scale. In order to have an identical range, a bit depth of 16 was used for all images. In this way, the entropy can be considered as a statistical measure of randomness that can be used to characterize the texture of the input image. Minimum value of the entropy would be expected when the pixel value of each patch is constant, thus indicating a fully homogeneous medium. On the other hand, maximum value of the entropy, for a given image patch, would be attained when the image pixel values are uniformly distributed in the closed interval [0, 2^{16}], representing random fluctuations. In reality, all the polarimetric parameters from the decompositions have different ranges of variation. If we want to compare the entropy of different parameters, with different ranges, and from different samples, one has to work with quantities having the same range. Hence, to perform an entropy calculation on any polarimetric quantity retrieved from the decompositions, the quantity has to be converted to gray scale image ranging within the interval [0, 2^{16}].

General conclusions

It becomes very important, in the field of biomedical optics, to provide medical doctors with excellent quality polarimetric images so that their diagnosis could be adequately aided. As discussed in the current Chapter, this is not a trivial task in general, since requiring a lot of post-processing and image filtering before getting the final representation. To start with, all measurements in the present work were performed in the same laboratory conditions and, in particular at the same ambient temperature, since the latter can be a significant factor for noise presence in the images. As for Mueller polarimetry, by normalizing to m_{11} and filtering the substrate contribution, any imperfections in the optical system, fluctuations in the light intensity or any polarimetric response from the substrate can be compensated. Otherwise, each of these factors would inevitably deteriorate the final image quality. Further, by applying the physical realizability filtering, one eliminates the experimental mistakes (if any). Next, with the help of image segmentation, the desired ROI is selectively analysed, thus, focusing the attention on particular features of the samples, important for the medical doctors. Finally, to avoid potential bias in the interpretation of the results, the distributions with different number of counts can be compared, by normalizing their histograms to their probability density functions. For the scope of the current dissertation, Fig. 12 shows a flowchart that summarizes the generalized data post-processing steps, which can be applied both for images and single-shot measurements. It could be well observed, that there is no universal solution and sequence, while the complexity behind the theory of all artificial intelligence and statistical methods is significant.

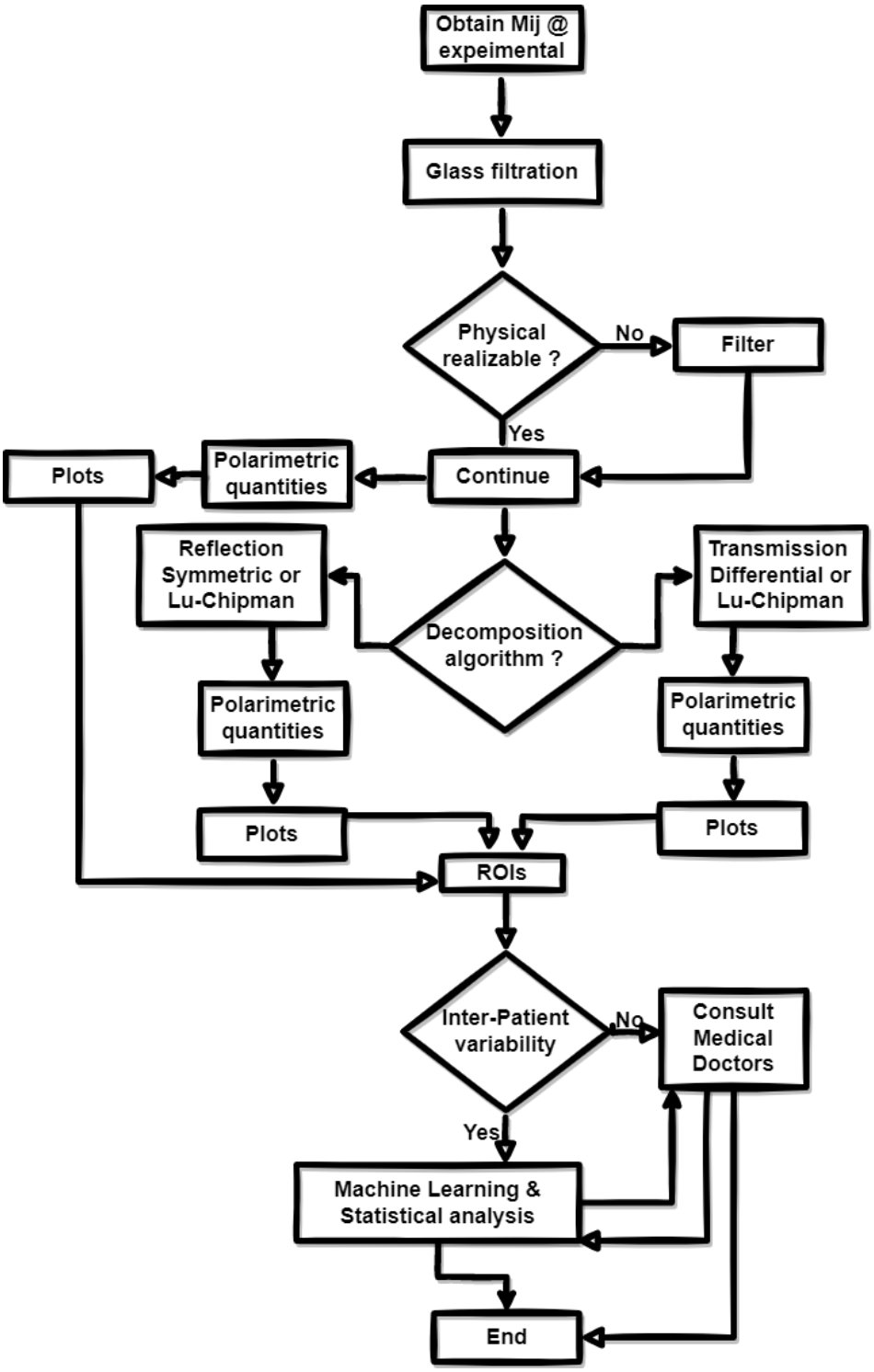


Figure 12: Flowchart of the data post-processing sequence and algorithms.

4 Experimental results and data post-processing

In this section, the experimental results are shown and discussed in detail, using the already introduced formalism and algorithms. The general purpose is to present, analyze and interpret the polarimetric measurements of various biological specimens and samples. The workflow approach included, in all instances, an initial discussion and guidance by the medical doctors and/or biologists and, after that, the reference samples were measured. Then, all other samples were measured under the same experimental conditions as the reference ones, and the results were compared with those obtained from the reference samples. Afterwards, all polarimetric results were sent back to the medical doctors for further analysis, as well as for ground truth validation.

With the increasing amount and size of the available data sets, its processing should be optimized. More precisely, for larger data sets instead of applying all decomposition algorithms described in Chapter 1 sample by sample, it would be time saving and computationally more effective to process all data at once. Then more general conclusions could be reached, in terms of the polarimetric response of particular disease presented in the tissue sample or the structural modifications upon different thawing mechanisms. Furthermore, this gives the opportunities to form polarimetric model(s) and train various AI algorithms to mimic human-like intellect, especially when handling large and complex data sets, images and etc. Since conventional programming processes the input data by means of particular syntax and semantics to produce a desired output, such method is prone to multiple errors repetition. On the other hand, ML uses both the input data and the output to train an algorithm for an *a priori* defined purpose. Depending on the purpose desired, ML algorithms can be grouped into three distinct classes [36–38]: supervised, unsupervised and reinforcement learning. The scope of the dissertation is focused on the application of both supervised and unsupervised ML algorithms. Hence, digital histology analysis could be supplemented by the appropriate ML algorithms and, potentially, by this way the medical doctors' diagnosis could be supported as well.

Another way for additional extraction of the important features from the experimental data consists in adopting the statistical analysis approach. It has already made a substantial contribution in the biomedical research [67]. By this way key patterns or features could be extracted from the available data. Generally, the statistical approach can be divided into two groups: parametric and non-parametric. In the former, the the data distribution, as well as its statistical moments are assumed to be known, whereas in the latter this assumption does not hold. Generally, the parametric tests are more powerful compared to the non-parametric tests when the above-mentioned assumptions are met. On the contrary parametric tests are less robust to violations of these assumptions, which can consequently lead to incorrect conclusions [67].

First, the results reported in subsection 4.1 were obtained in collaboration with the Bern University Hospital to validate the most suitable decomposition algorithm in transmission, as well as to study the behaviour of multi-layered stacks of scattering, anisotropic brain tissue sections. Second, subsection 4.2 presents the experimental results obtained in collaboration with the University Hospital Wuerzburg and Fraunhofer Institute for Silicate Research in Wuerzburg

with the purpose to study and monitor the impact of tissue thawing and viability of cells. Next, the results presented in subsection 4.3 originate from a collaboration with the Institute of Electronics, Bulgarian Academy of Sciences and University Hospital Queen Joanna - Sofia under Campus France PHC RILA project AURORA - Digit**A**l histology of tiss**U**e with Muelle**R** micr**O**scopy and machine lea**R**ning **A**pproach. The aim was to monitor and compare the behaviour of the polarimetric response of *ex vivo* skin samples with different histological conditions. Finally, the results in subsection 4.4 are obtained in the framework of a joint collaboration with the Institute of Electronics, the University Hospital Queen Joanna, and the Oulu University, Finland. They include polarimetric measurements of the thick colon specimens in reflection geometry for colon cancer detection.

4.1 Multi-layered stacks of thin sections of anisotropic scattering white matter of brain

The rationale for this study is the visualization of the brain fiber bundles with polarized light by exploring the anisotropy of the refractive index of healthy brain white matter and the assessment of the impact of brain fiber crossing on the experimental polarimetric images of brain.

It is known that the healthy white matter of the mammals' brain is comprised of the nerve fiber bundles, which results in an optical and structural anisotropy [68, 69]. Upon malignancy, this anisotropy would be altered and by this way, polarimetric imaging could be used as contrast enhancing method for tumor delineation during neurosurgery [70–73]. This is of paramount importance for brain tumor surgery, where the precise localisation of a tumor zone border is needed for the most complete and safe resection of brain tumor. In clinical practice, the conventional methods to localise brain tumors during neurosurgery include intra-operative magnetic resonance imaging (iMRI) [74] and 5-aminolevulinic acid (5-ALA) intra-operative fluorescence imaging assisted surgery [75]. However, both imaging techniques have limitations in terms of the acquisition time and detection of the border of low grade tumors, respectively. The prior studies showed the potential of imaging Mueller polarimetry operating in reflection to detect the optical anisotropy by visualizing the in-plane orientation of the nerve fiber bundles [76–79] by applying Lu-Chipman decomposition [56] of Mueller matrix images and suggested to explore the absence of such anisotropy as an optical marker for brain tumor.

An imaging Mueller polarimeter should be able to detect and visualise the underlying crossing nerve fibers with the polarimetric maps of linear retardance, which may provide a valuable information for neurosurgeons. Additionally, the polarimetric maps of depolarization and the quiver plots, corresponding to the orientation angle of the optic axis could also be used as sensitive optical markers. It could be also important to pay attention to the most suitable Mueller matrix decomposition algorithm due to the different initial assumptions and mathematical foundations.

For this set of measurements, two thin paraffin-embedded sections of corpus callosum of human brain obtained from the autopsy of anonymous donor were chosen (see Fig. 13). This anatomical part of brain serves as a connection between the two brain hemispheres and has a well-defined orientation of fiber bundles [80]. A waiver for ethical approval was obtained from the Ethics Committee of the Canton of Bern (BASEC-Nr: 84 Req-2021-01173). Both differential and Lu-Chipman decomposition algorithms were used for the post-processing of the recorded Mueller matrices.

The excision of a 2.5 cm × 1.5 cm × 0.5 cm coronal section of the corpus callosum was performed first in the central area of brain at the Neuropathology Department of the Bern University Hospital, Switzerland. Subsequently, the formalin-fixed specimen was embedded in paraffin. This block was then cut in sections of different thicknesses (e.g., 5 μm, 10 μm) for polarimetric measurements. The photo and schematic layout of the standard microscopy glass slides with thin (5 μm and 10 μm) histological sections of corpus callosum embedded in paraffin are shown

in Fig.13. First, we measured the single layer of 10 μm nominal thickness, then the stack of two superimposed layers (10 μm + 5 μm) aligned parallel, then crossed at 45° and, finally, at 90°. By this way, various multi-layered structures were created to monitor the behaviour of the (de)polarization parameters from the decomposition algorithms. All measurements from this set were performed in transmission. The selected FOV of the Mueller microscope was approximately 600 μm with 0.75 μm/pixel spatial resolution.

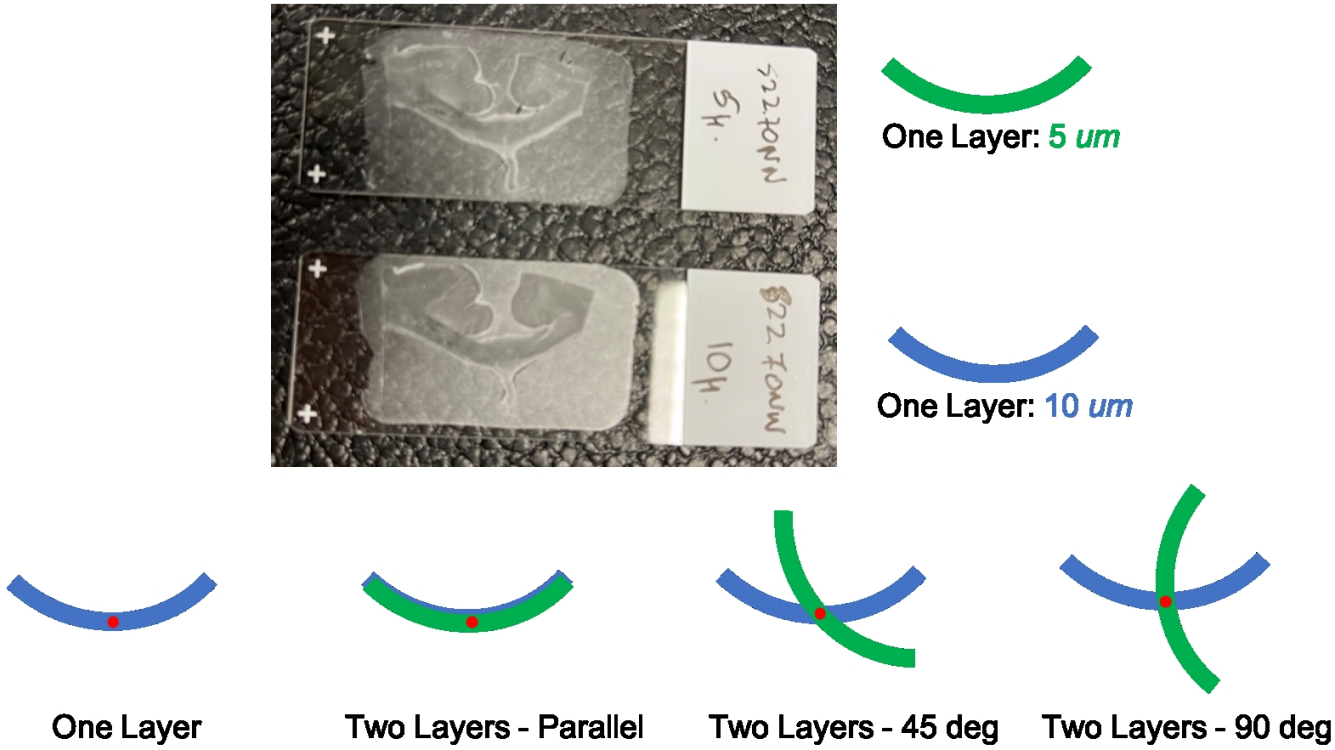


Figure 13: Thin histological sections of corpus callosum: photo of the glass slides with the coronal brain tissue sections of 5 μm (green) and 10 μm (blue) nominal thicknesses; schematic layout of the different spatial arrangement of the thin sections of corpus callosum used for the measurements. Red spots indicate the locations of the measurement site. Special acknowledgements to Lu Si (Tsinghua-Berkeley Shenzhen Institute, Tsinghua University, Shenzhen, China) for his contribution to figure preparation. Figure available from the published material in [59].

We begin with the discussion of the results. Firstly, we calculated and plot the maps of the net scalar retardance for all samples' spatial configurations, with both decomposition algorithms, respectively. The corresponding maps are shown in Fig.14. All stacks comprised of two corpus callosum tissue stripes with different thicknesses represent an anisotropic medium and their polarimetric responses vary with the relative orientation of the samples. As can be seen in Fig.14 both decomposition algorithms produce the same results. This is also confirmed by the box-whisker plots of the corresponding images as illustrated Fig.15. The addition of another corpus callosum stripe in parallel arrangement increases the optical path length of the transmitted light and consequently of the detected signal compared to a single-layer configuration. As a result, this leads to an increase of the net scalar retardance.

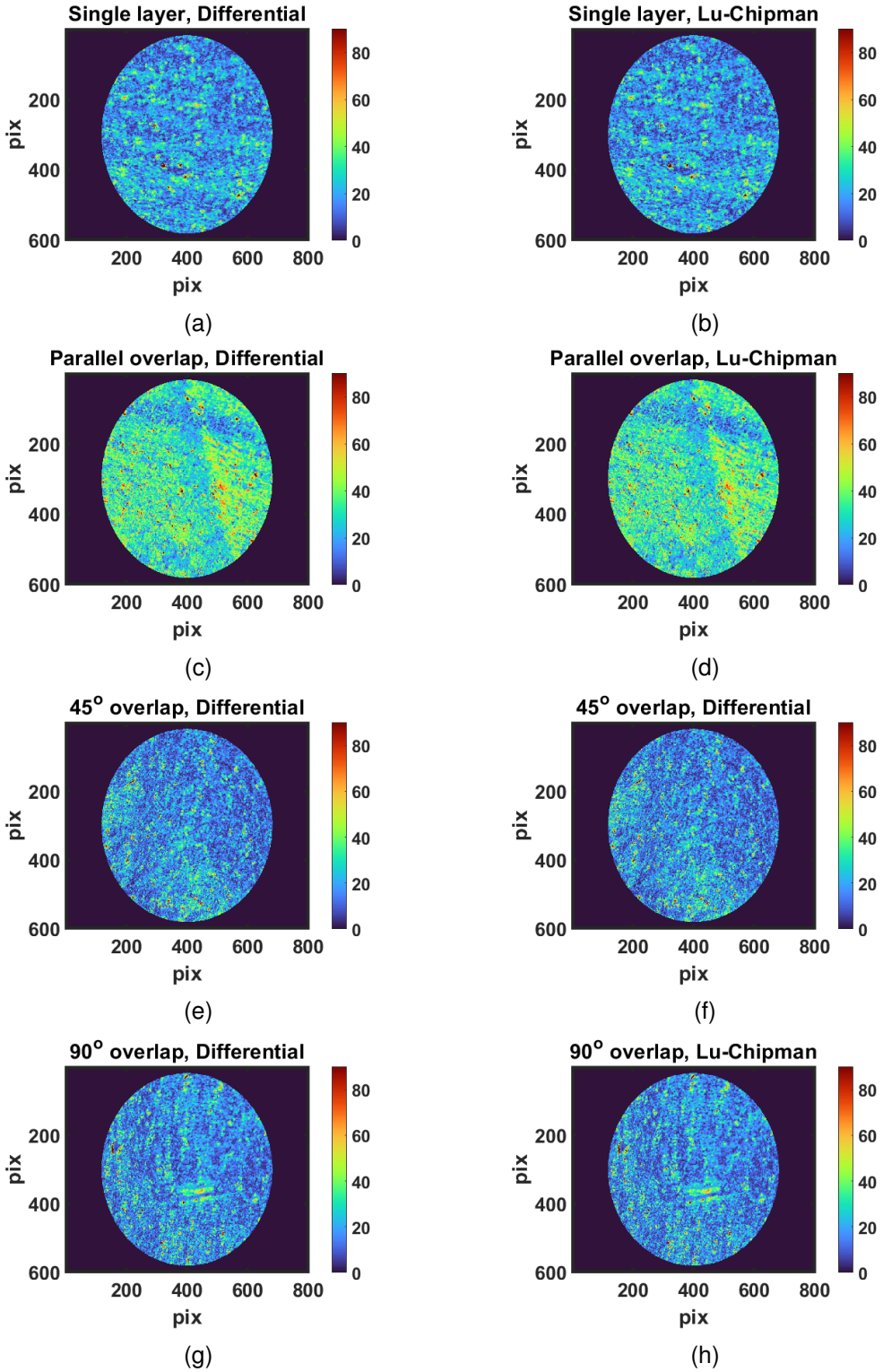


Figure 14: Net scalar retardance in degrees, obtained with either differential (left column) or Lu-Chipman (right column) decomposition, corresponding to different spatial arrangements of the brain corpus callosum stripes: (a) and (b) single layer (10 μm); (c) and (d) parallel overlap (10 μm + 5 μm); (e) and (f) 45^o overlap (10 μm + 5 μm); (g) and (h) 90^o overlap (10 μm + 5 μm).

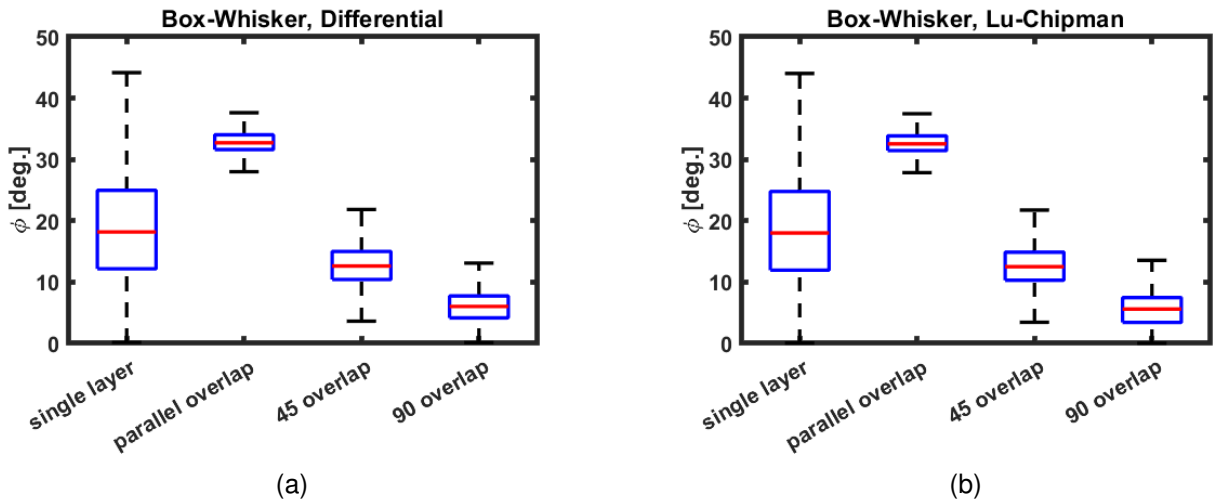


Figure 15: Box-whisker plots of the net scalar retardance from: a) differential and b) Lu-Chipman decomposition, corresponding to the different spatial configurations of the *corpus callosum*, brain samples.

Upon rotation of the top tissue stripe, the net scalar retardance decreases, reaching its lowest value for the mutually orthogonal arrangement of the two stripes. This is attributed to partial phase retardance compensation. Moreover, a preferential orientation of a densely packed nerve fiber bundles within the imaging plane is also detectable in Fig.14a-d, whereas it is lost in Fig.14e-h.

Next, the quiver plots in Fig.16 show the orientation of the slow optic axis that is aligned with the preferential orientation of the corpus callosum fibers. The length of each line is proportional to the corresponding value of the net scalar retardance, normalized to the maximum value of it over all pixels of the image. As can be seen in Fig.16, both decomposition algorithms produce similar results. As confirmed from the retardance maps and the corresponding box-whisker plots, the parallel arrangement of two super-imposed brain tissue stripes results in the highest scalar retardance values and also the largest orientational lines and the sub-domains of preferentially oriented fibers are clearly distinguishable in Fig.16(c),(d). Also, a randomization of the azimuth of the optic axis and the loss of the preferential orientation is observed, when comparing the corresponding quiver plots for the spatial arrangement of brain corpus callosum stripes at 45° and 90°, respectively.

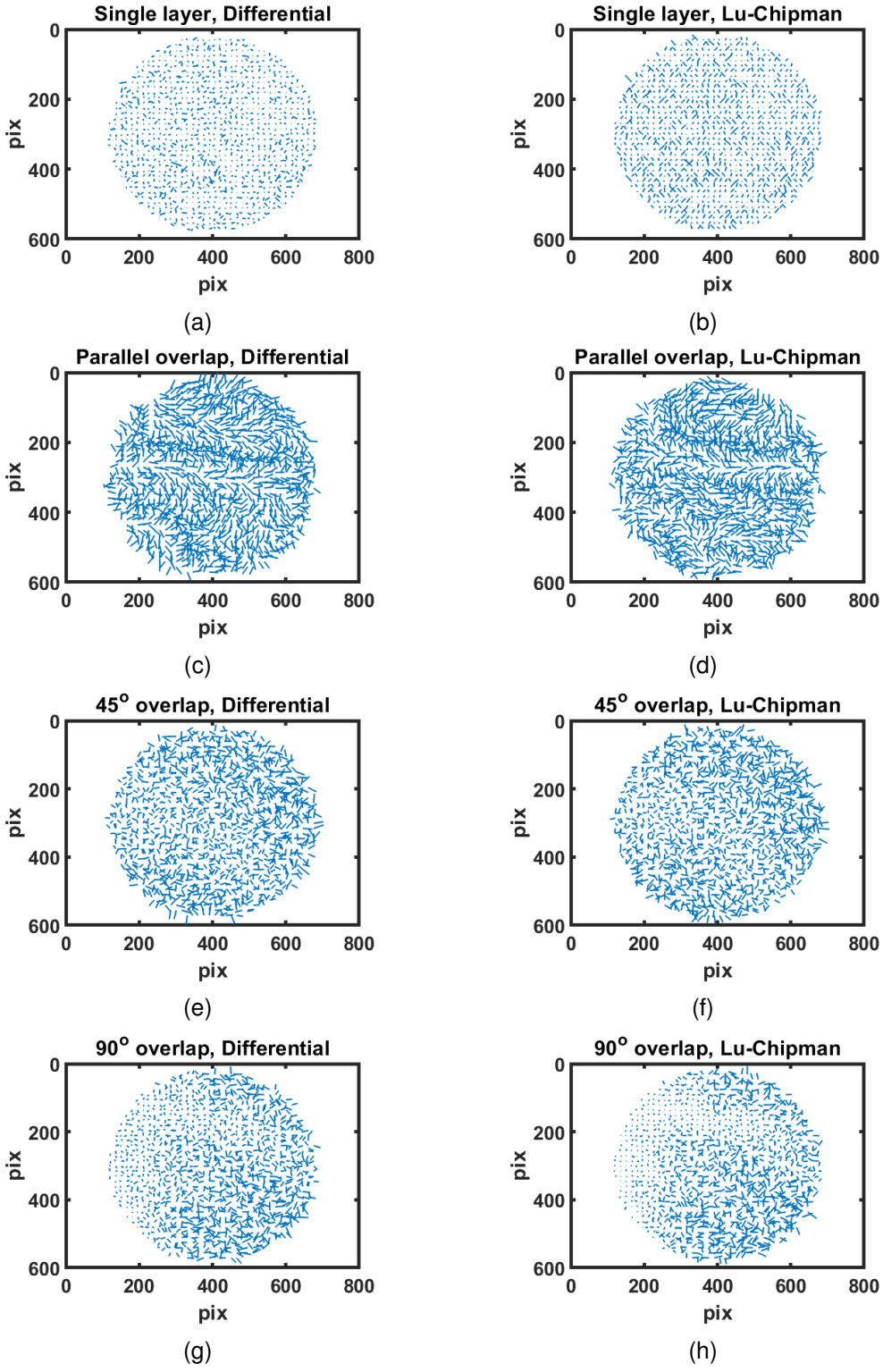


Figure 16: Quiver plots of the net scalar retardance, obtained with differential (left column) or Lu-Chipman (right column) decomposition, corresponding to different spatial arrangements of the brain corpus callosum layers: (a) and (b) single layer (10 μm); (c) and (d) parallel overlap (10 μm + 5 μm); (e) and (f) 45° overlap (10 μm + 5 μm); (g) and (h) 90° overlap (10 μm + 5 μm).

Finally, the polarimetric analysis involves the values of total depolarization obtained from both decomposition algorithms. The corresponding images are presented in Fig.17 from which it can be seen that the lowest depolarization values account for a single layer of corpus callosum, whereas the depolarization values increase for configurations of the two superimposed stripes, regardless of their mutual spatial orientation. Also, a physical thickness increase of the measured samples can be attributed to the addition of the second tissue stripe, leading to an increase of the number of scattering events. This, in turn, causes the loss of spatial coherence and an increase of depolarization of the detected light. Contrary to the variation of the retardance with the spatial arrangement of two stripes, the above-mentioned process does not depend on the orientation of the fiber bundles, but rather depends on the density of the scatterers and the optical thickness of the medium. By rotating the top slide manually, we detect a slight variation of the depolarization values when changing the spatial arrangement of two overlapped brain corpus callosum stripes. We attribute it to the local variations in the top tissue stripe thickness. Contrary to his trend, the net scalar retardance values decrease significantly for 45° and 90° spatial arrangements of two overlapped tissue stripes.

From the box-whisker plots of the total depolarization (Fig.18) one can observe that slightly higher values are obtained from the Lu-Chipman decomposition algorithm in comparison to the differential one. This behaviour could be explained by the fact that in the case of the differential decomposition there is neither diattenuation, nor polarizance present in the asymmetric tensor $\mathbf{L_u}$. On the other hand, in the forward Lu-Chipman decomposition the depolarizer Mueller matrix $\mathbf{M}_{\triangle}$ is "contaminated" with the polarizance. Moreover, to make a direct comparison between the depolarization values calculated with different decompositions, one needs to use Eq.50. Otherwise, the depolarization values calculated with the Lu-Chipman decomposition varies between 0 for non-depolarizing sample and 1 for completely depolarizing sample, whereas the depolarization values from the differential decomposition cover larger interval from 0 for a non-depolarizing sample to $+\infty$ for a completely depolarizing sample. Last but not least, the differential decomposition assumes continuous variation of the polarization and depolarization properties for a given medium, while the Lu-Chipman decomposition makes use of the discrete assumption due to the matrix multiplication for the comprising Mueller matrices, being part of this algorithm. Hence, the choice of decomposition algorithm is of non-negligible importance for the final results, as shown from the results so far. The continuous variation of the polarization and depolarization properties holds for the majority of the biological tissues, so for the rest of the polarimetric experiments in transmission geometry, we decided to use the differential decomposition only for the Mueller matrix data post-processing.

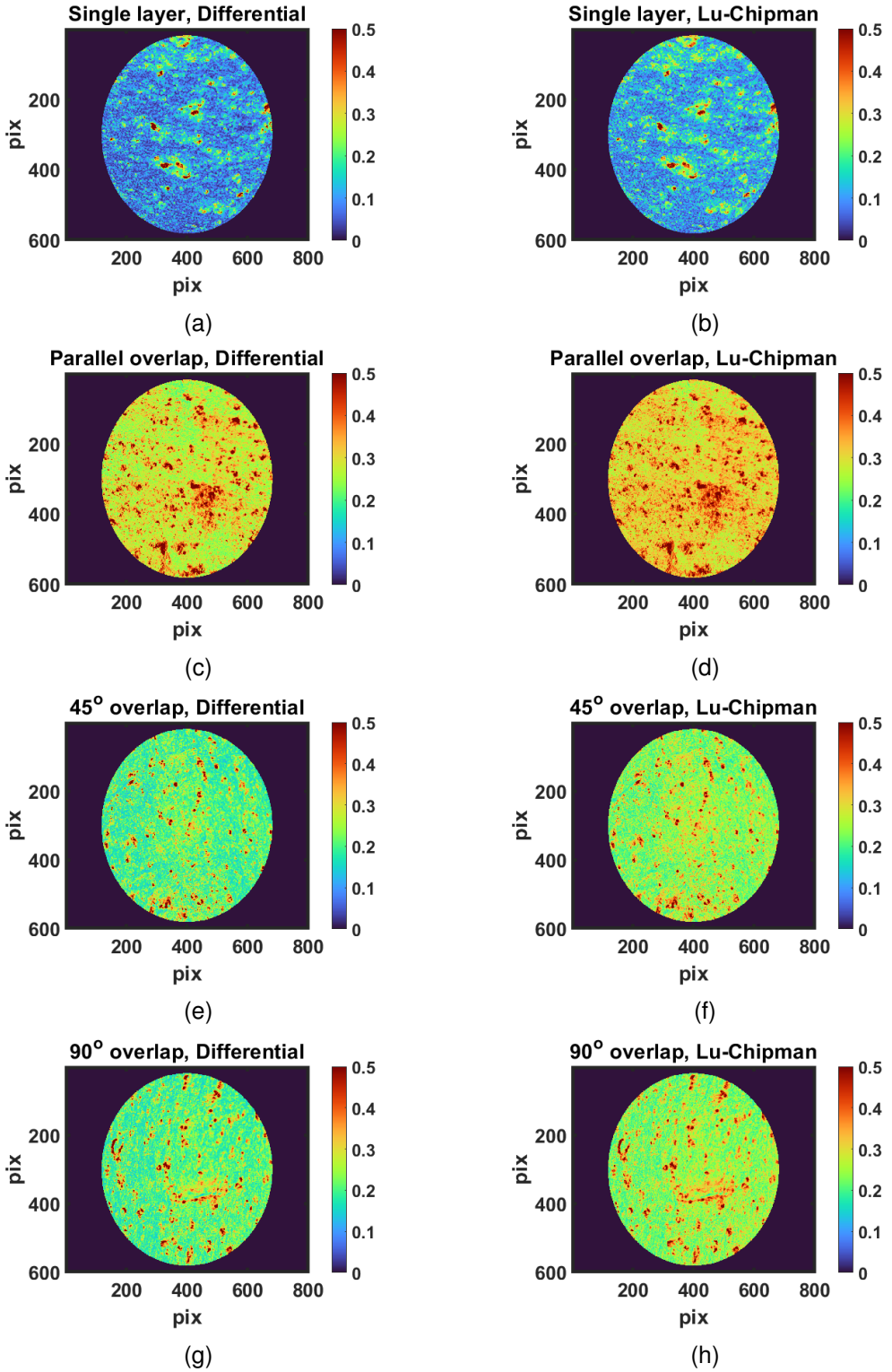
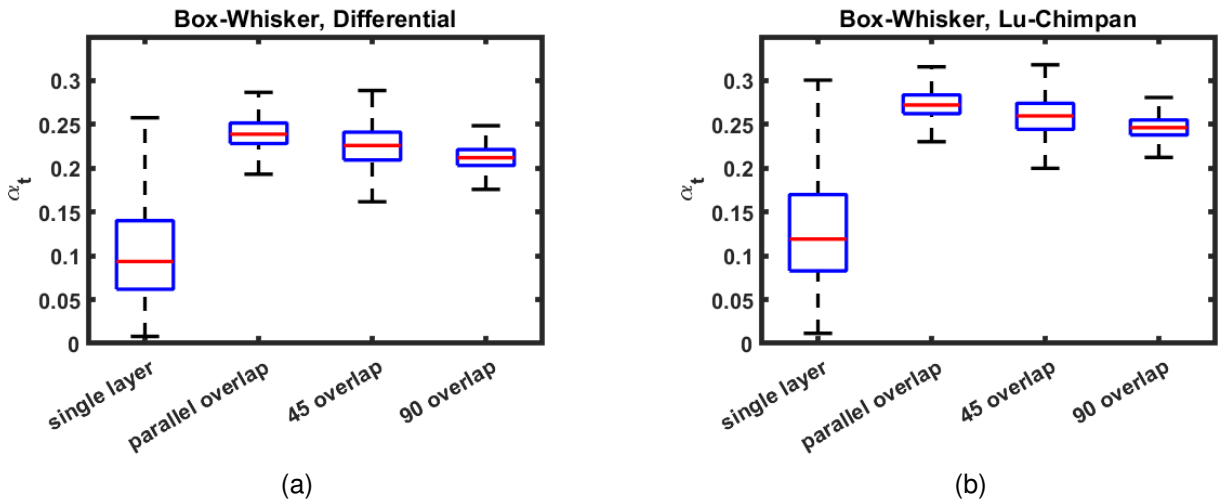


Figure 17: Total depolarization, obtained with either differential (left column) or Lu-Chipman (right column) decomposition, corresponding to different spatial arrangements of the brain corpus callosum stripes: (a) and (b) single layer (10 μm); (c) and (d) parallel overlap (10 μm + 5 μm); (e) and (f) 45° overlap (10 μm + 5 μm); (g) and (h) 90° overlap (10 μm + 5 μm).



(a) (b)

Figure 18: Box-whisker plots of the total depolarization from: a) differential and b) Lu-Chipman decomposition, corresponding to the different spatial arrangement of the corpus callosum brain stripes.

4.2 Assessment of different thawing mechanisms of tissue models

4.2.1 Polarimetric results

Freezing biological specimens at cryogenic temperatures and then defrosting them at room temperatures is often followed by a growth of large ice crystals [15, 16]. Such a process causes damage to the cellular membranes and the extracellular collagen matrix [17–19]. To minimize ice formation, so-called cryoprotectants (CPAs) are added before freezing. CPAs can either enter the cell (penetrating CPAs) or act outside of the cell (non-penetrating CPAs) [81]. However, CPAs are challenged by several requirements, such as low toxicity and biocompatibility [81]. Traditionally, the cryopreserved biological sample is rewarmed by immersion in a waterbath at temperatures around 37°C. However, this approach was found to be successful for small samples only (with a volume below 3 mL [82, 83]). Larger biological specimens such as organs or tissues require faster and more uniform rewarming to prevent cracking and damage to the biological specimen. A promising approach is the process of nanowarming, where magnetic nanoparticles (MNPs) are inductively heated by using a high frequency coil [83–85].

Initially, the membrane integrity break, caused by cryoinjury, is anticipated to be followed by changes in the tissue structure and its morphology [20–22]. As a consequence, the tissue scattering properties will be altered. These alterations can be detected with a polarized light probing beam, because the changes of tissue scattering properties will result in changes of the polarization degree of the reflected/refracted light. Some of the pioneer studies in the tissue cryopreservation domain suggested the exploitation of MNPs and radio-frequency (RF) inductive heating as a pertinent replacement method [20] for tissue thawing in a water bath. Nevertheless, the core mechanisms of MNP-tissue interactions and MNP-aided tissue defrosting are to be better understood and comprehended, as this approach holds great potential for the cryopreservation of cell suspensions and bio-tissues [21, 22].

The flowchart of our study design that includes the steps of 3D tissue model preparation, MNP-assisted nanowarming, sectioning of the tissue models and polarimetric Mueller microscope measuremants is illustrated in Fig.21. Fibroblast cells are the most common type of cells comprising connective tissues and are the primary source of the extra-cellular matrix (ECM) [86]. For the scope of this study, the fibroblasts containing 3D models of tissue were prepared in the Institute of the Regenerative Medicine, Fraunhofer Institute for Silicate Research, Wuerzburg, Germany and University Hospital, Wuerzburg, Germany. Electrospun fiber fleeces were used as the scaffold of the 3D in vitro models to mimic the extracellular matrix. The fleeces were made of the polymer polyamide 6, in which magnetite NPs were embedded. During the spinning process, NaCl particles were added into the fleece, resulting in an open, highly porous 3D structure and facilitating fibroblast cell migration and growth. The resulting MNPs-modified 3D fiber fleeces were cultivated with fibroblast cells for 4 weeks.

Then tissue model samples were split into three groups. Firstly, a control sample was selected and used as a reference without being frozen and thawed. Two other groups were frozen at -80°C. One group was defrosted within a water bath, whereas the remaining one - with radio frequency (RF), inductive heating of the magnetic nanoparticles (MNPs), initially incorporated into the tissue model. For histological observations of the fibroblast-containing 3D models with Mueller microscope the tissue sections were performed. For this purpose, the tissues were first fixed and embedded in paraffin and then 10 μm thick tissue sections were prepared. A detailed description of the protocol for the samples preparation can be found in [87], while the photo of the setup for RF rewarming with the incorporated MNPs is illustrated in Fig.22 .

In order to assist medical doctors to detect the changes in the internal structure of the tissue samples, caused by the freezing and defrosting processes, Mueller polarimetry is a suitable candidate, as polarization of light is sensitive to the microsctructural alterations of tissue. In total, 20 measurements per group were made. As most parts of the images were occupied with paraffin (see Fig.19 and 20), only the regions with the stripes of tissue models were selected by using the image segmentation technique with the patches, described in details in Chapter3. The size of each patch was set to be 100 x 100 pixels. Again, all measurements were conducted in transmission, using 700 nm wavelength of the probing beam. The workflow in data post-processing includes also the physical realizability filtering, differential decomposition [88] and normalization to $\ln(m_{11})$ or $\ln^2(m_{11})$, where the FOV was set to 400 μm (or 0.5 μm/pixel). After the post-processing of all data from all measurements, it was found that the azimuth angle θ of the optic axis is the most sensitive parameter to discriminate between the stripes of tissue models and paraffin. Therefore, we focused on the analysis of the maps of this parameter only, whereas, for better clarity, the intensity images are also included to illustrate the contrast between the zones of paraffin and the stripes of tissue models before applying the decomposition.

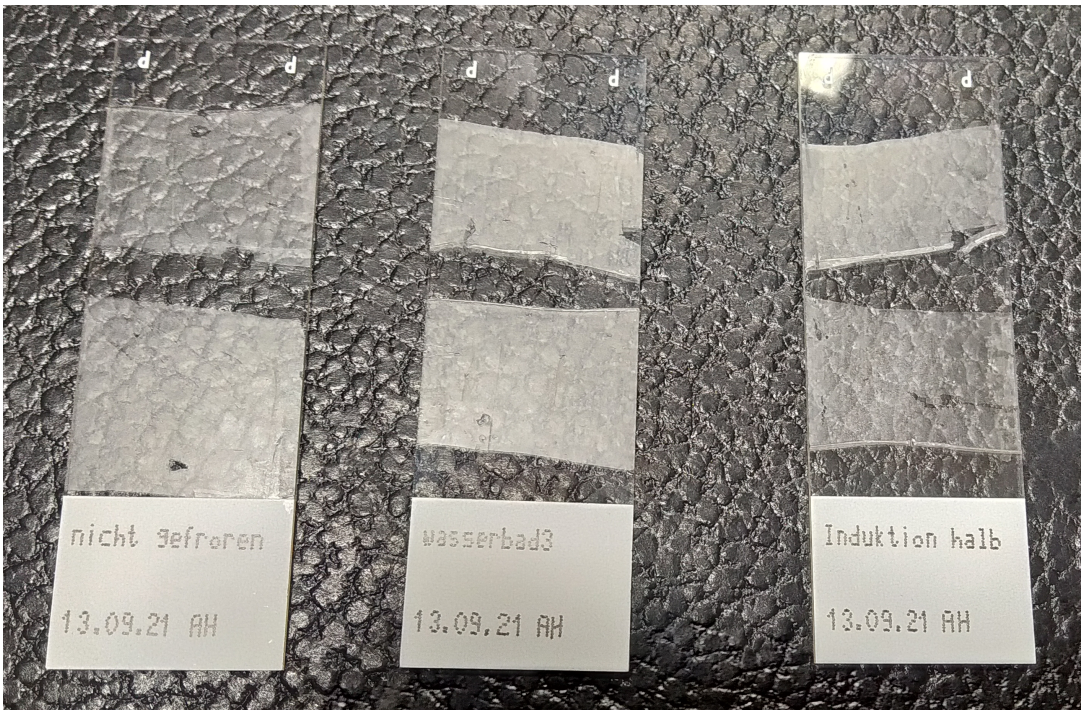


Figure 19: Photo of the microscopy glass slides with the samples under examination. From left to right: Control/Reference, water thawed and RF thawed with MNPs.

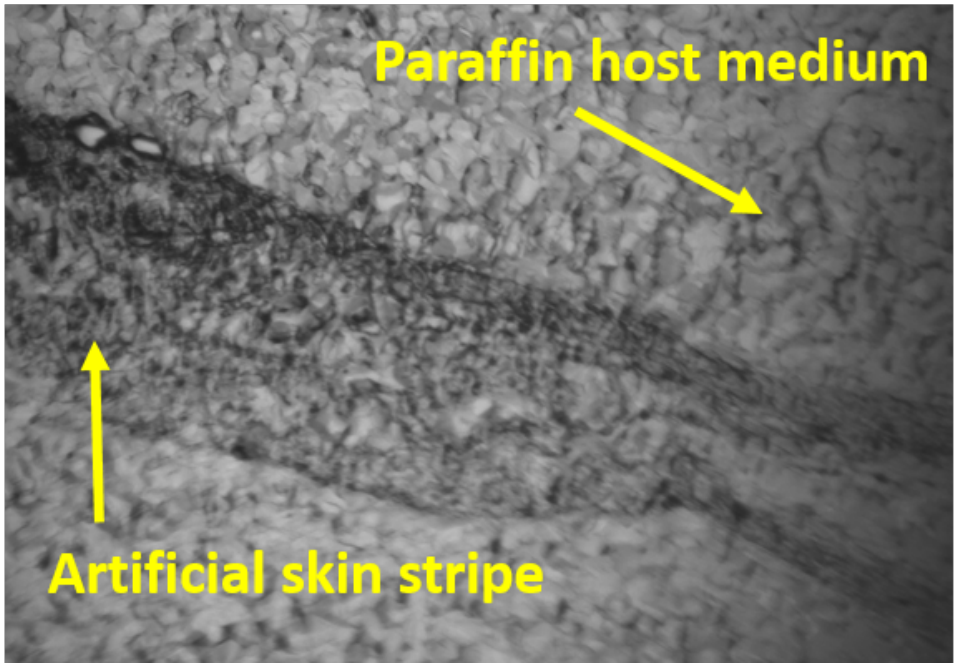


Figure 20: Microscopy image with the 3D artificial skin model (the darker zone) and the paraffin host medium around it in which the skin stripe was embedded.

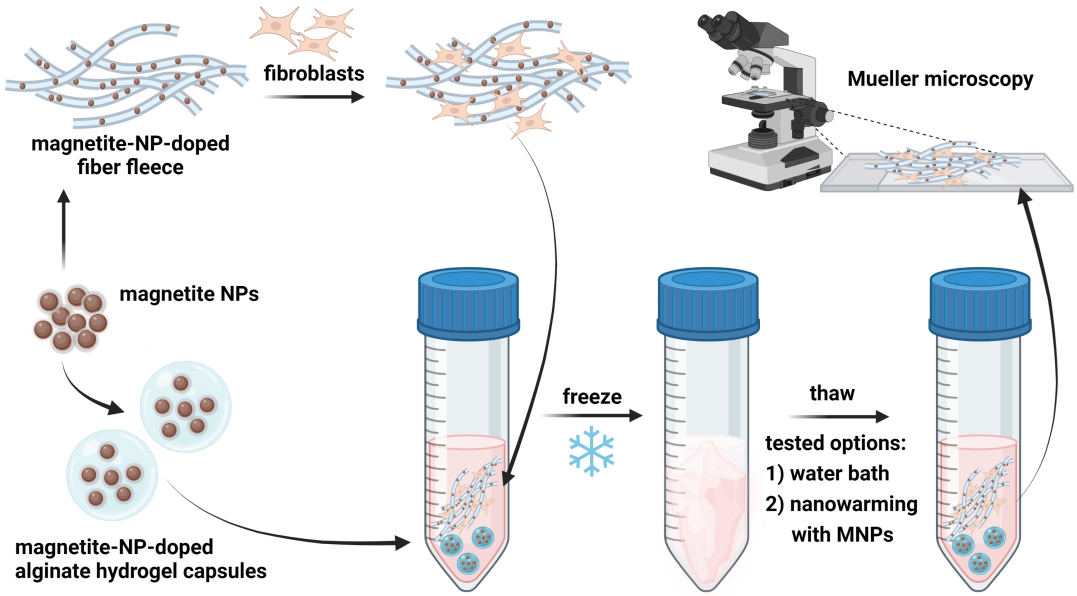


Figure 21: The flowchart of the study design including the steps of samples' preparation, rewarming with MNPs, tissue models sectioning and polarimetric measurements. Special acknowledgements to Anika Hoeppel for the figure preparation.

Figure 22: Photo of the setup for the inductive thawing of the fibroblast-containing 3D tissue models. Special acknowledgements to Anika Hoeppel for the figure preparation.

The intensity images of the samples from each class are shown in Fig. 23.

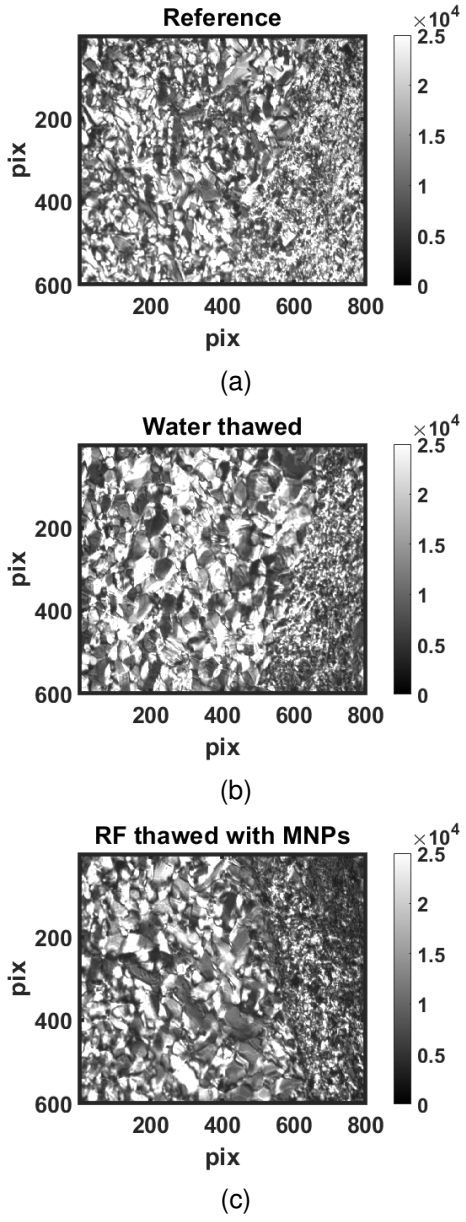


Figure 23: Intensity images of m_{11} (scales in bit depth): (a) reference, (b) water thawed and (c) RF thawed with MNPs.

As can be seen, the right-hand side of all three images contains a region with different intensity values that corresponds to the stripes of tissue models, while the remaining parts of the images show the paraffin host medium. Next, we look at the same regions of interest in the maps of the azimuth θ of the optic axis for each class (see Fig. 24).

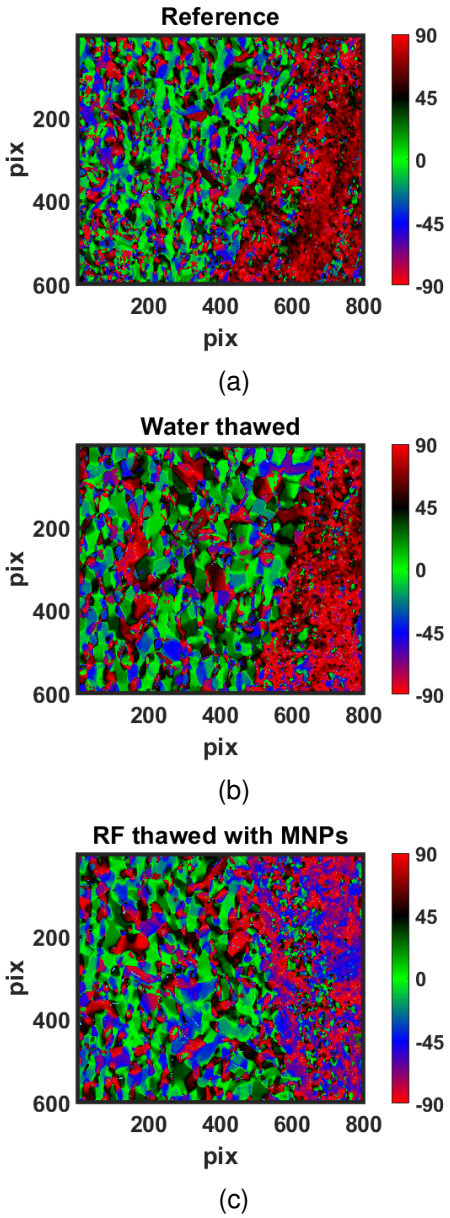


Figure 24: Maps of the azimuth θ of the optic axis (in degrees): (a) reference, (b) rater thawed and (c) RF thawed with MNPs.

Again, but this time more clearly, The right-hand side of the images, contains the regions with almost constant azimuth values (about $\pm$ 90°) that indicates the preferential vertical orientation of the optic axis. On the other hand, in the zones of paraffin, there are multiple small clusters with varying orientations.

Then, the algorithm for image segmentation was applied (see subsection 3.2). After extracting the data from the patches corresponding to the cell stripes only from all samples and all measurements, we plot the normalized histograms of the azimuth θ (see Fig. 25). This would allows us to compare the distributions of this polarimetric quantity, depending on the thawing mechanism used and compare it with the reference (or control) sample.

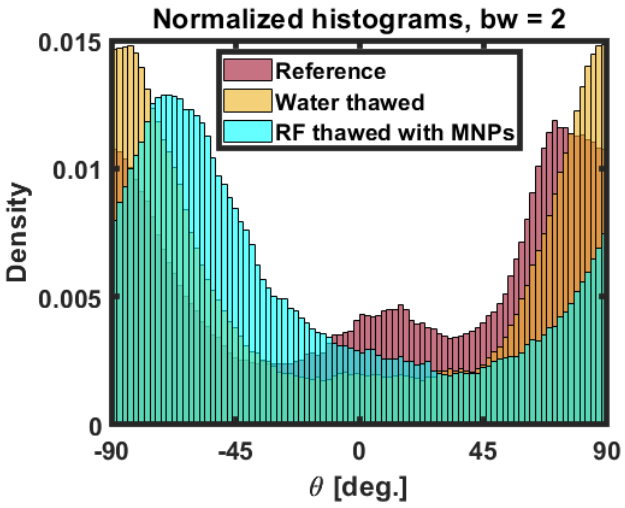
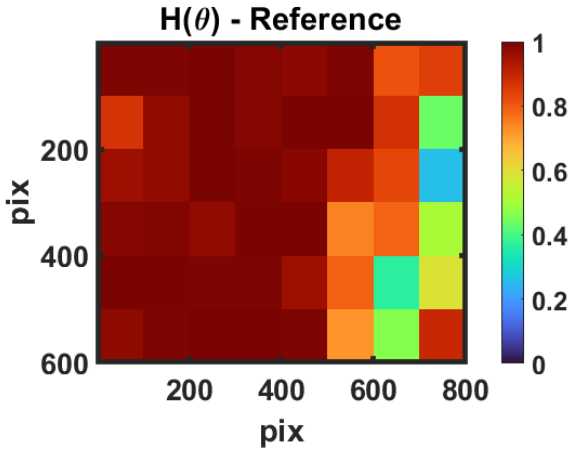


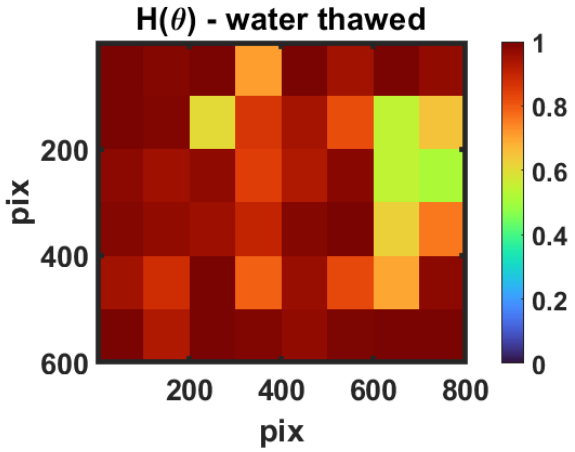
Figure 25: Normalized histograms of the θ angle for the three groups.

After plotting the normalized histograms, it becomes evident, that there is a difference in the orientation angle, depending on the defrosting methods used, with respect to the reference samples. For better clarity, a circular colormap is shown as well to better visualize and illustrate the angle distribution of the θ values. In Fig.25 one can observe, that the predominant orientation of the azimuth of the optic axis for the reference sample is close to vertical or $\pm$ 90°. However, there is also another peak, that correspond to the horizontal orientation or 0°, although with the decreased probability. Analysing the distribution for the the water-thawed sample, we again observe the peak at $\pm$ 90°, whereas the second peak at 0° disappeared. Finally, the predominant orientation for the RF thawed sample with MNPs tends towards the negative angle values about -70°. Nevertheless, the three distributions are still overlapping for most of the angle values. In order to analyse in details the corresponding distributions, whether they differ significantly or not, a detailed statistical analysis is needed, which will be presented in the next subsection.

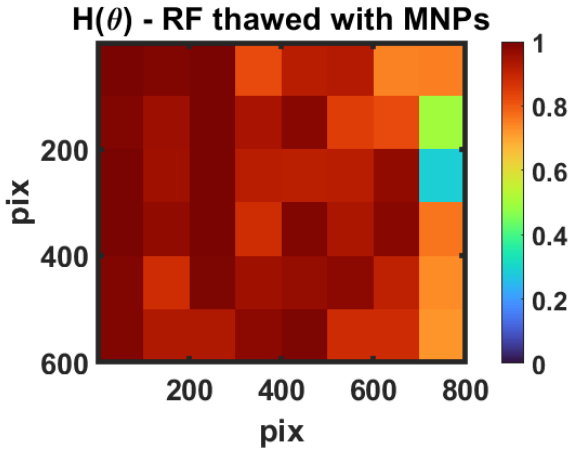
Next, we would like to continue the analysis by plotting the entropy H and the standard deviation σ for each 100$\times$100 patch of the images of the azimuth θ. To do so, all azimuth images have been converted to 16-bit, gray scale, in order to comply with the requirements for the built-in MATLAB function to calculate the entropy H. After the calculation of these metrics, for each patch we will have a scalar value, which we assign to each pixel of the corresponding patch. By this way, we create the super-pixels that illustrate the spatial heterogeneity of the corresponding tissue zones and, thus, reveal information about the changes within the samples related to the defrosting mechanisms used.



(a)

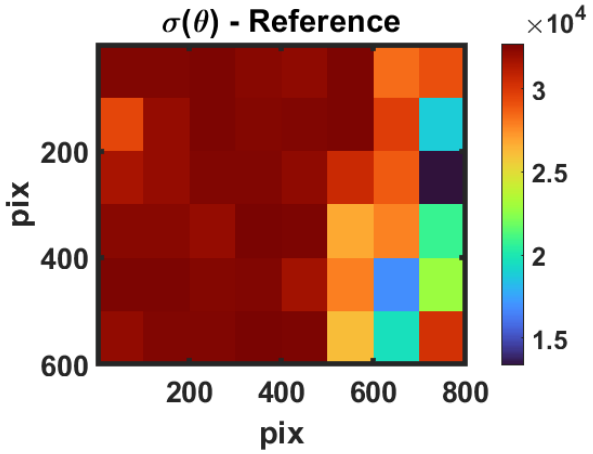


(b)

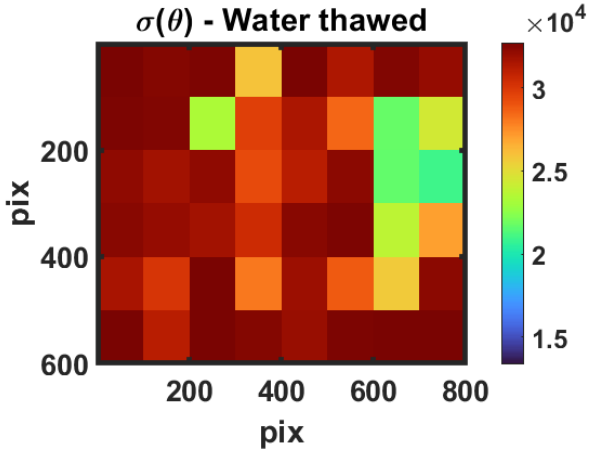


(c)

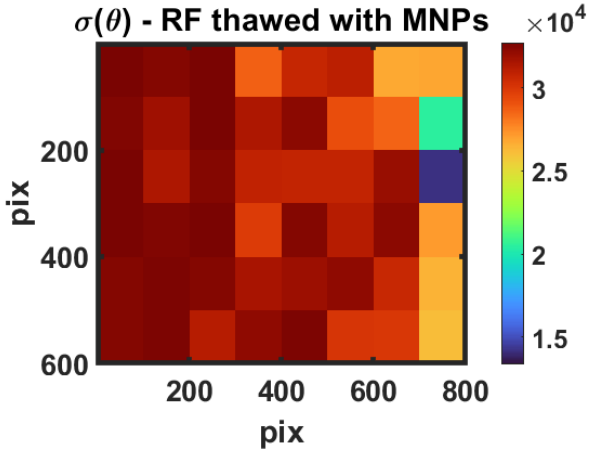
Figure 26: Entropy H of the azimuth of the optic axis θ for each patch with size 100×100 pixels: (a) reference, (b) water thawed and (c) RF thawed with MNPs.



(a)



(b)



(c)

Figure 27: Standard deviation σ in bit depth of the azimuth of the optic axis θ (after converting in gray-scale) for each patch with size 100 x 100 pixels: (a) reference, (b) water thawed and (c) RF thawed with MNPs.

From Fig.26 and Fig.27 several common trends can be observed. For instance, both metrics have lower values for the zones of tissue model stripes compared to the zones of the paraffin. This observation is valid for the images of all three classes. By this way, we can assume that the spatial heterogeneity of the paraffin zones is higher than the spatial heterogeneity of the zones with the tissue model stripes. On the other hand, both metrics for the tissue model stripes from the reference sample produce the lowest values. After the water thawing the values of the entropy and the standard deviation increase for the tissue model zones, suggesting the destruction or alteration of the initial preferential orientation of tissue model stripes. On the other hand, with the RF-assisted thawing with MNPs, the lower values of the entropy and the standard deviation are observed for the tissue model stripes, indicating a better preservation of the orientation angle θ with this defrosting mechanism.

4.2.2 Statistical analysis

Due to the fact, that the images presented contain the zones of paraffin host medium, it would be useful to separate them and exclude from the analysis of the experimental results. As the thin sections of 3D tissue models are embedded into the paraffin, it would not be feasible to use supervised or unsupervised machine learning algorithms for this task. Rather, a statistical approach could be implemented. To do so, after segmenting all images into patches with the size of 100 x 100 pixels, only the patches from zones that contain tissue models were used to form a data base, thus minimizing as much as possible the paraffin contribution to the data. Afterwards, all thickness normalized parameters obtained from the differential decomposition were split into three distinct groups: 1) control or reference, where no freezing or defrosting was applied; 2) water-bath thawed, where the samples were frosted at -80°C and later defrosted within the water bath at 37°C; 3) magnetic nano-particles (MNP) thawed, where the samples were frosted at -80°C, and thawed by applying the radio-freguency (RF) electromagnetic field. In the figures below the box-whisker plots are shown for all polarimetric parameters from all three groups, including the intensity images.

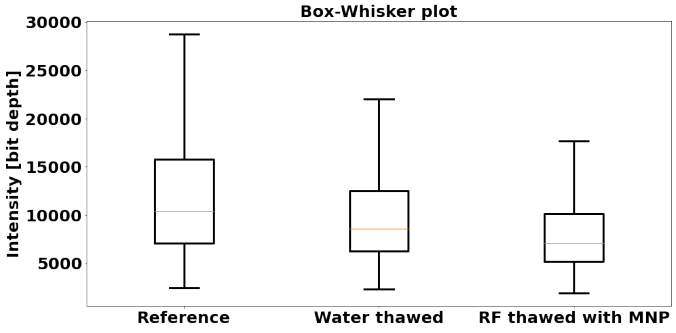


Figure 28: Box-Whisker plot of the intensity from all three groups.

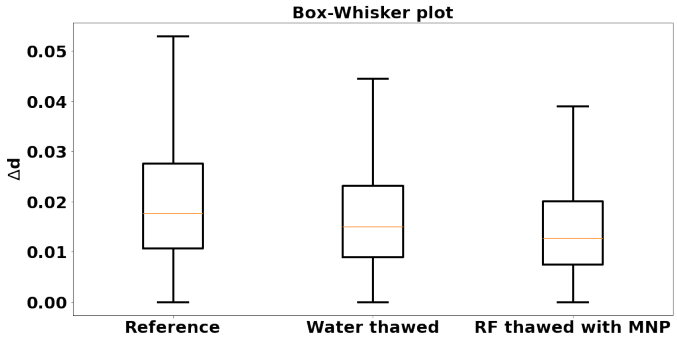


Figure 29: Box-Whisker plot of the net dichroism Δd from all three groups.

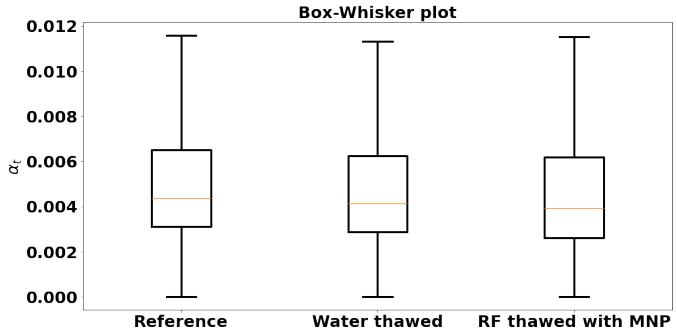


Figure 30: Box-Whisker plot of the net depolarization α_t from all three groups.

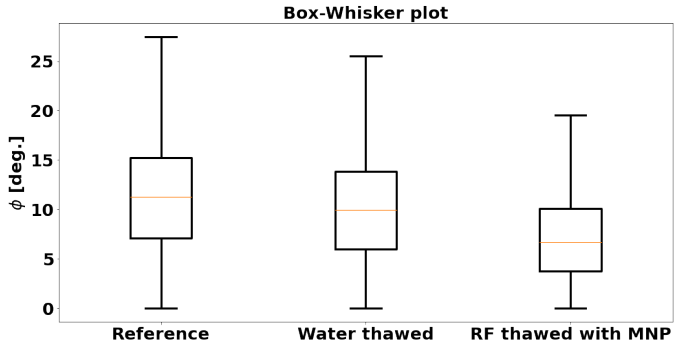


Figure 31: Box-Whisker plot of the net, scalar retardance ϕ from all three groups.

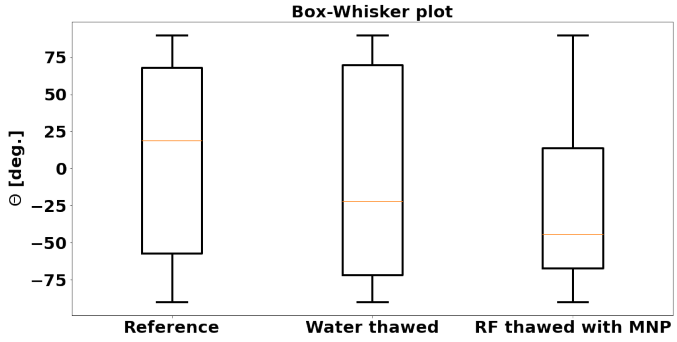


Figure 32: Box-Whisker plot of the orientation angle θ from all three groups.

After this graphical representation, one could observe the overlap between the three distributions, whereas their median values differ from each other. For this particular reason, it could be practical to use the ANOVA test for unequal variances and compare the mean values [67] of the above mentioned distributions. The initial significance level α was set to 0.05. This value was compared to the computed p - value. The zeroth hypothesis for this test H_0 is that the data from all three groups are drawn from the same distribution. This would indicate no statistical difference between the control group and the groups thawed differently. It will be valid if the mean values from the three distributions are equal and $p > \alpha$. On the contrary, the alternative hypothesis H_0^* would be that the data from all three groups are drawn from the different distributions. This would indicate a statistical difference between the control group and the two groups thawed differently. It would be valid if at least one mean value from the three groups compared is not equal to the remaining two and $p < \alpha$. After the application of this statistical test for all polarimetric parameters and including the intensity, the computed p - value is less than α. Thus, on a significance level of 0.05, we can assume, that the data are drawn from the different distributions and there is a statistical difference between the control group and the groups that were thawed using two different defrosting mechanisms.

Next, we want to make similar tests, but between all groups of all parameters. At this point, another approach should be implemented by using *post-hoc* analysis [67]. This is needed, as the error rate $\mathcal{E}$ would scale with the number of comparisons γ and groups respectively (see Fig.33).

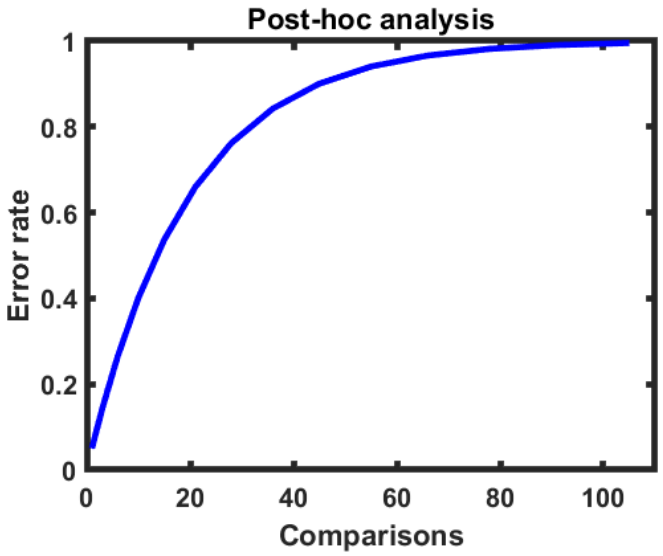


Figure 33: Graphical representation of the dependence between the error rate $\mathcal{E}$ and the number of comparisons γ for significance level α = 0.05.

The latter two parameters discussed above can be calculated as:

$$\mathcal{E} = 1 - (1 - \alpha)^{\gamma}, \tag{87a}$$

$$\gamma = \frac{n(n-1)}{2}, \tag{87b}$$

where n is the number of groups. For instance, for the current case with n = 3, the value of γ will be also 3 and the value of $\mathcal{E}$ is around 0.14. This is almost three times higher than the value of α = 0.05. When n = 15, the value of γ will be 105 and the value of $\mathcal{E}$ is close to unity. By this way, with the increasing number of groups and comparisons, a false positive result will be guaranteed. For this particular reason, we initially set a constant error rate, when performing *post-hoc* analysis, on the price of adjusted p - values. The adjustment is being done by the built-in Python libraries for this test. While we decrease the error rate by this way, if the adjusted p - values are reduced, this would make the test less sensitive to smaller changes into the data. Again as for the ANOVA test, the zeroth hypothesis for each multi-group comparison H_0 is that the data from the two groups are drawn from the same distributions. This would indicate no statistical difference between them. It will be valid if their mean values are equal and the adjusted $p > \alpha$. On the contrary, the alternative hypothesis H_0^* would be that the data from each muti-group comparisons are drawn from the different distributions. This would indicate a statistical difference between them. It would be valid if their mean values are different and the adjusted $p < \alpha$. For this purpose, Tukey HSD post-hoc test was used [67]. After the application of this statistical test for all multi-group comparisons per each polarimetric parameters, including the intensity data, the adjusted p - values are less than α. Thus, on a significance level 0.05, we can assume, that there is a statistical difference between all three groups or, that by means of polarimetry, the small changes in the samples' structures, resulting from the difference in the defrosting mechanisms could be detected. Nevertheless, this statement is to be confirmed or rejected with another set of tissue model samples without the paraffin host medium.

4.3 Digital histology of thin sections of human skin

4.3.1 Polarimetric results

In this subsection, all results from *ex vivo* human skin thin sections are presented. Thin histological sections with nominal thicknesses varying from 4 μm to 12 μm were placed on the glass microscopy slides. The samples were measured with the Mueller microscope in transmission geometry, while the corresponding Mueller matrices were decomposed with the differential decomposition. The results of the histological analysis by a pathologist were used as a ground truth for polarimetry. In total we had eight different tissue types (or classes) that were labelled as: H – *Healthy*; SC – *Scleroderma*; LU – *Lupus erythematosus*; PS – *Psoriasis*; RE – *Syndrome of Raynaud*; SCC – *Squamous-cell carcinoma*; BCC – *Basal-cell carcinoma* and MM – *malignant Melanoma*.

The *Scleroderma* is an autoimmune diseases, which is usually localized on the skin, where some of the most common symptoms include thickened skin and abnormal growth of connective tissue [89–92]. Another autoimmune disease is the *Lupus erythematosus*, causing widespread inflammation and tissue damage in the affected organs, including the skin [89–92]. Also, *Psoriasis* is an immune, inflammatory disease, mainly affecting the skin. The inflammations concern the dermis and the epidermis, with excessive renewal of skin cells [89–92]. On the other hand, the *Raynaud syndrome* may be associated with skin lesions or with connective-tissue disorder [89–92]. Usually, the aforementioned skin conditions are termed as degenerative.

The group of malignant skin lesions includes Squamous-cell carcinoma that starts from the squamous cells, which are located on the surface of skin, on the lining of hollow organs in the body and on the lining of the respiratory and digestive tracts [89–92]. Another type of malignant lesion is the *Basal-cell carcinoma* that represents the most common type of skin cancer and is named after the basal cells that form the lowest layer of the epidermis [89–92]. Finally, *malignant Melanoma* is a type of skin cancer that develops from the pigment-producing cells, known as melanocytes and it is the most dangerous type of skin cancer [89–92]. The early detection of the malignant pathologies of skin will increase the life expectancy and the quality of life of the patients and tissue polarimetry might be used as a suitable optical technique to address this problem.

In total, the measurements of thin sections (26 healthy, 31 degenerative and 33 tumorous) were conducted in transmission, with 700 nm wavelength, as in this spectral region, the scattering dominates over the absorption in biological tissues [14]. The FOV was set to 400 μm or 0.5 μm/pix. The data post-processing workflow includes in succession the physical realizability filtering, differential decomposition and thickness invariant normalization. Data post-processing includes a glass filtration step before applying the decomposition of the recorded Mueller matrices. The results are presented and discussed bellow.

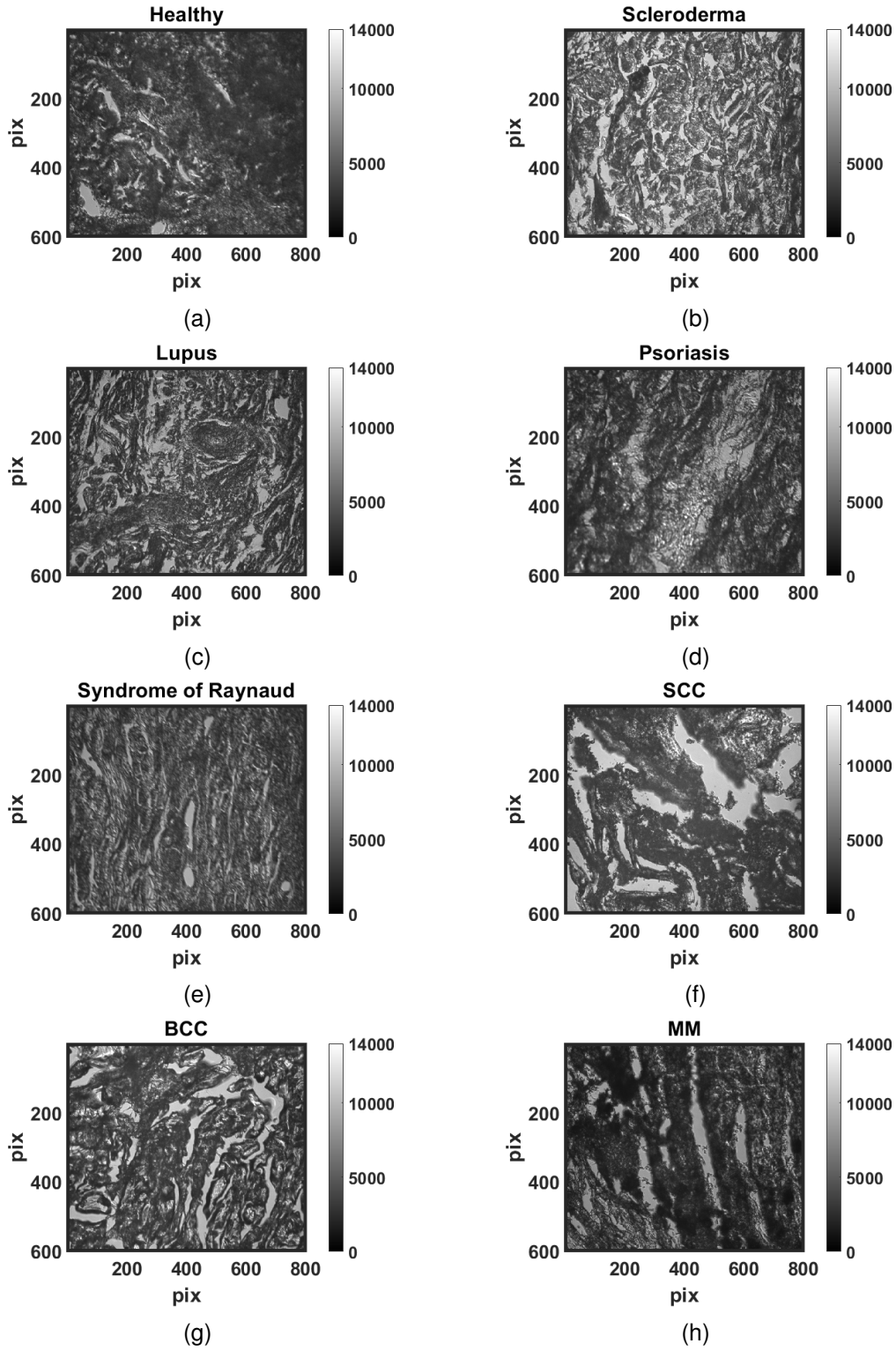


Figure 34: Intensity images of m_{11} in bit depth units: (a) Healthy, (b) Scleroderma, (c) Lupus, (d) Psoriasis, (e) Syndrome of Raynaud, (f) Squamous-cell carcinoma, (g) Basal-cell carcinoma and (h) Melanoma.

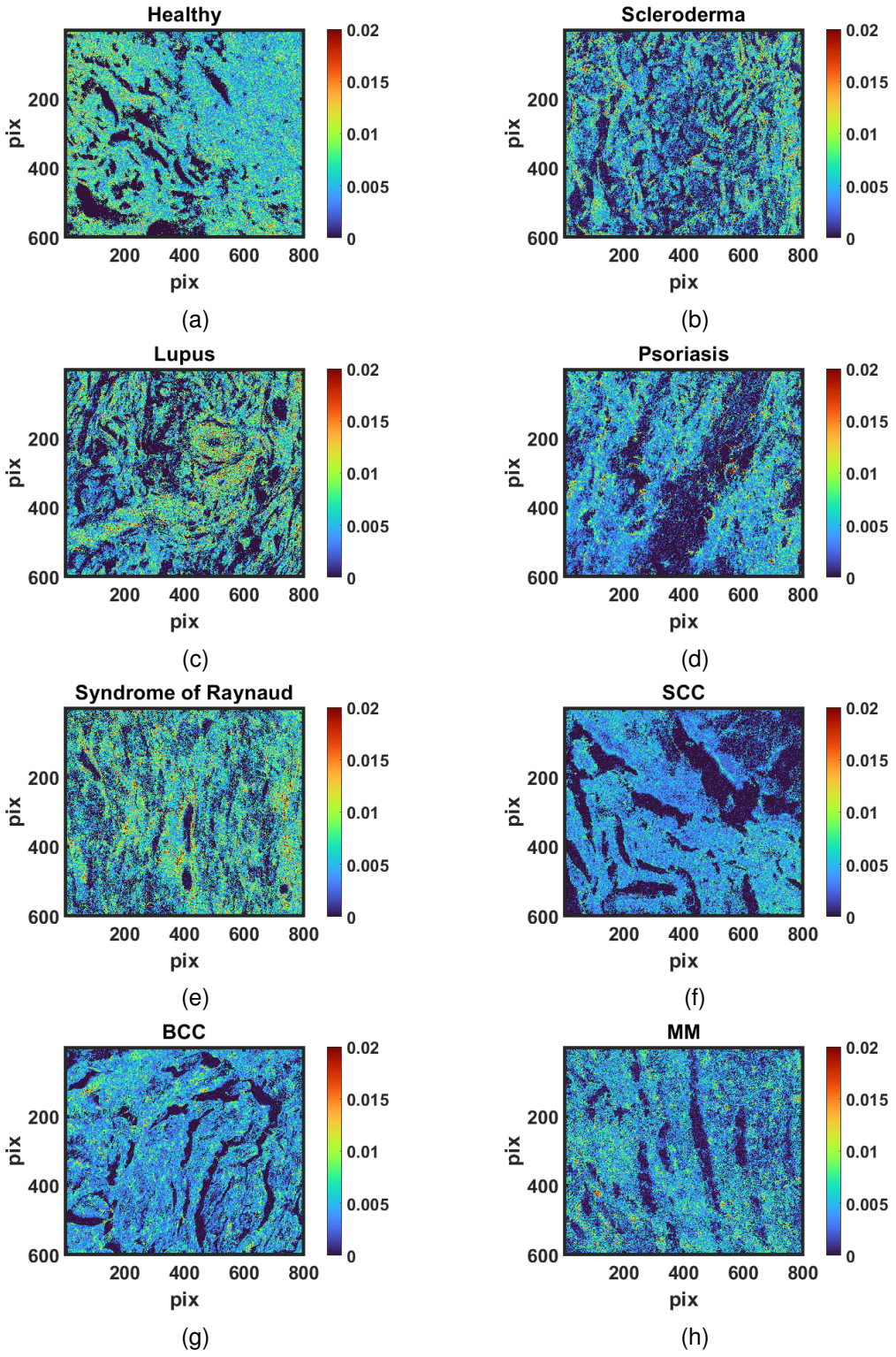


Figure 35: Images of dichroism Δd: (a) Healthy, (b) Scleroderma, (c) Lupus, (d) Psoriasis, (e) Syndrome of Raynaud, (f) Squamous-cell carcinoma, (g) Basal-cell carcinoma and (h) Melanoma.

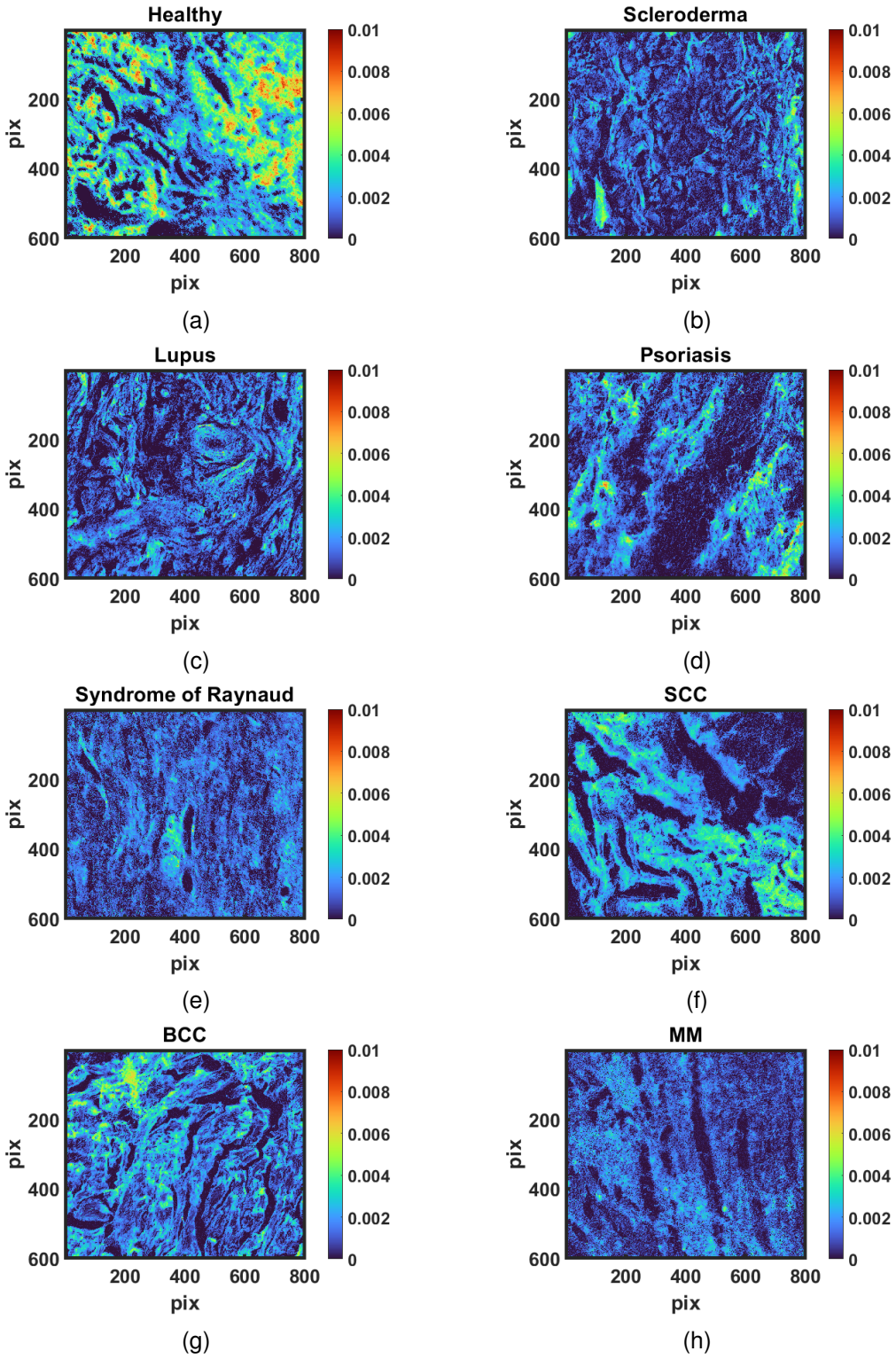


Figure 36: Images of depolarization α_t: (a) Healthy, (b) Scleroderma, (c) Lupus, (d) Psoriasis, (e) Syndrome of Raynaud, (f) Squamous-cell carcinoma, (g) Basal-cell carcinoma and (h) Melanoma.

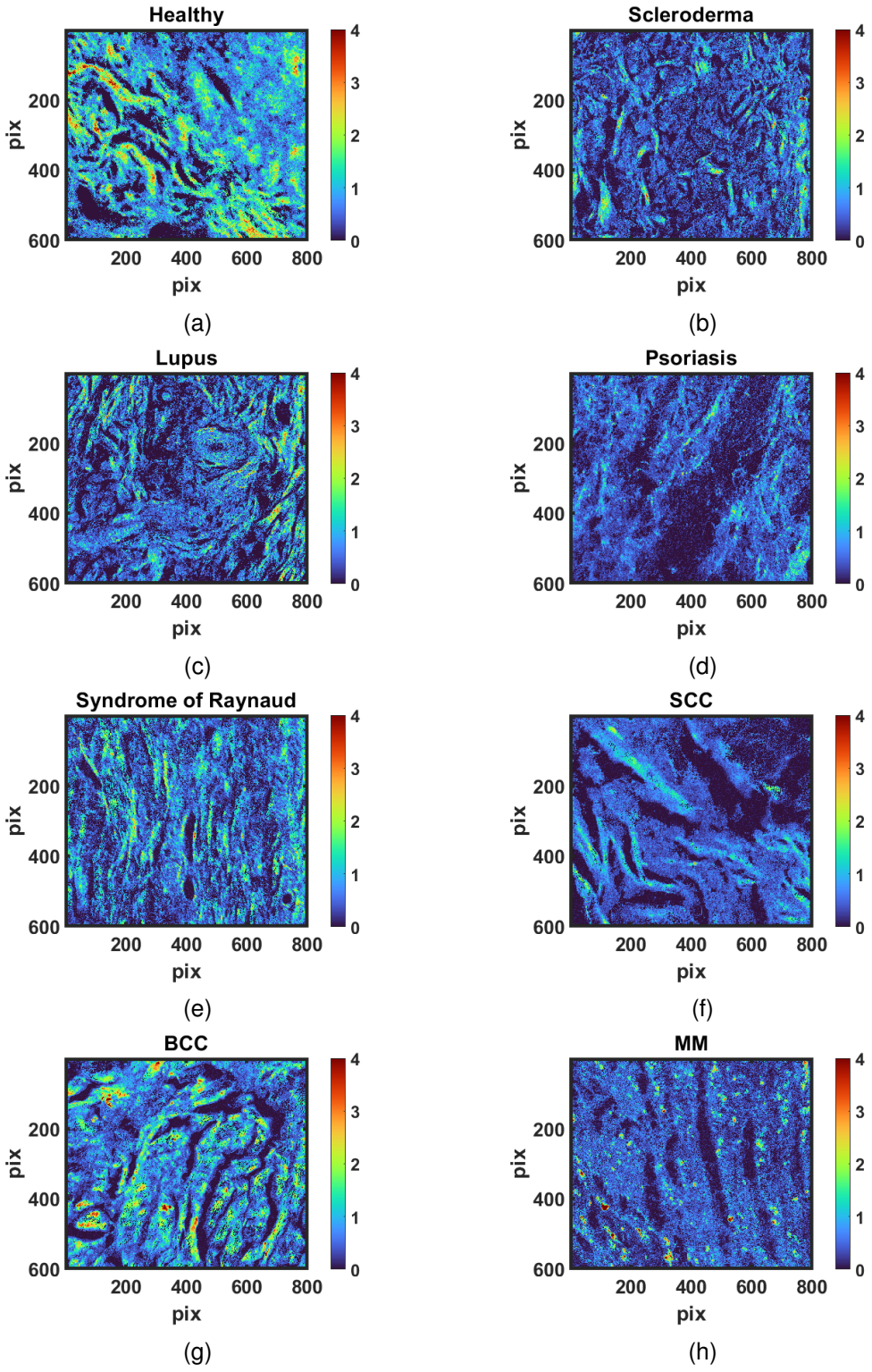


Figure 37: Images of retardance ϕ: (a) Healthy, (b) Scleroderma, (c) Lupus, (d) Psoriasis, (e) Syndrome of Raynaud, (f) Squamous-cell carcinoma, (g) Basal-cell carcinoma and (h) Melanoma.

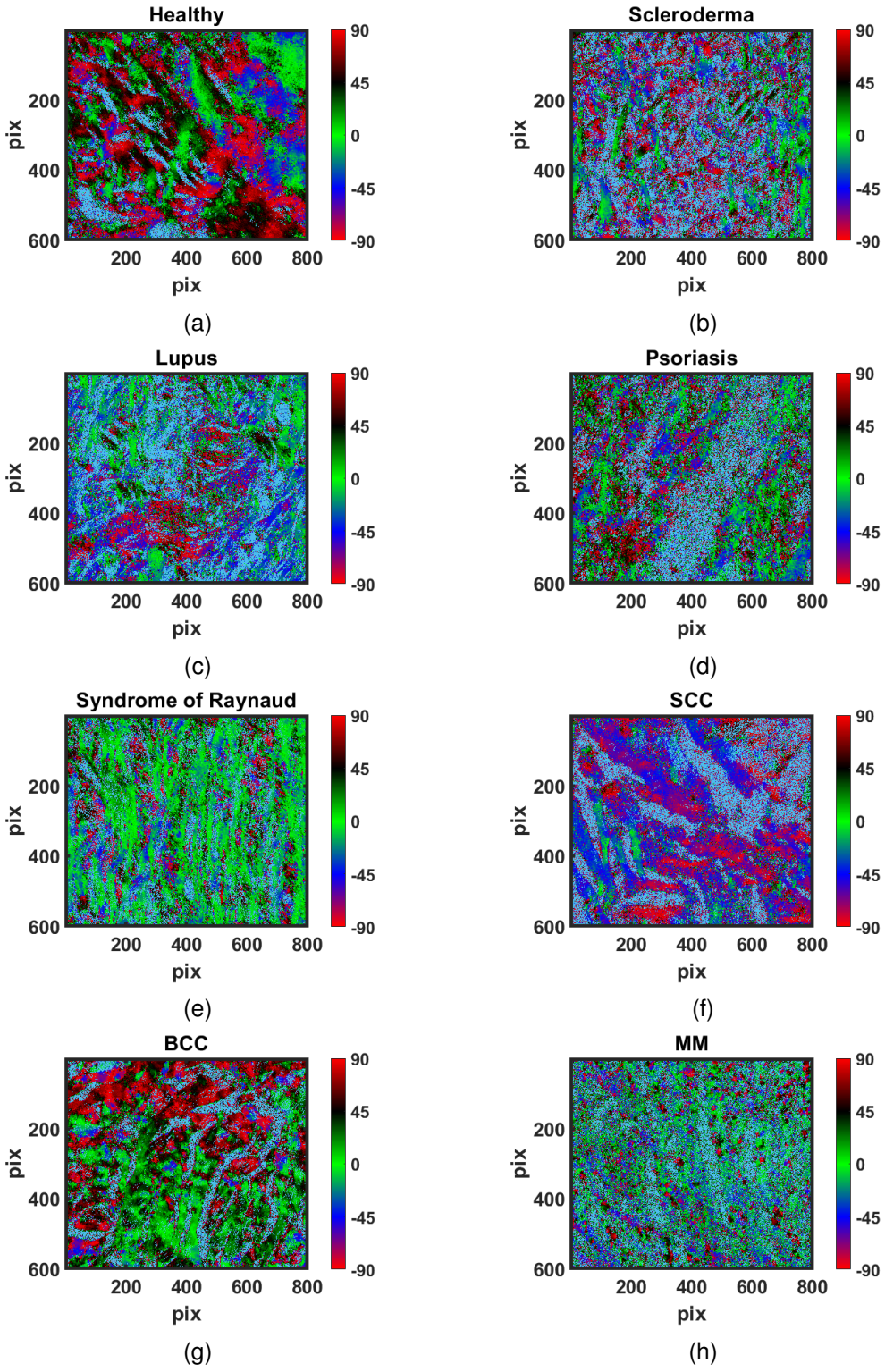


Figure 38: Images of the orientation angle θ of the optic axis: (a) Healthy, (b) Scleroderma, (c) Lupus, (d) Psoriasis, (e) Syndrome of Raynaud, (f) Squamous-cell carcinoma, (g) Basal-cell carcinoma and (h) Melanoma.

As can be seen from the intensity images in Fig.34, there are cavities in the skin tissues, where the glass substrate is visible. These bare glass zones occur with a bit depth around 10^3. Also, some ROIs in the intensity images have significantly low values, whereas the others are far more transparent. This observation holds for all skin lesion images. Analysing the rest of the polarimetric parameters from the differential decomposition, one can observe two trends. Firstly, the algorithm for the glass filtration clears very effectively its contribution. Secondly, the values of the polarimetric parameters for the degenerative and tumorous skin tissue images are decreased with respect to the corresponding values for healthy skin tissue images, used as a reference. This could be attributed to the structural modifications of skin by the pathology, causing morphological alterations in skin tissue, thus, changing its optical properties and affecting its polarimetric response.

Additionally, the orientation angle of the optic axis shows several predominant orientations, in different ROIs of all tissue types, whereas the orientation angle θ is being completely randomized at some zones. This could be related to the destruction of fine ordered fabric of healthy skin tissue upon development of a particular skin lesion. To further amplify and compare the differences between all samples, one must use and compare the measurements from all classes.

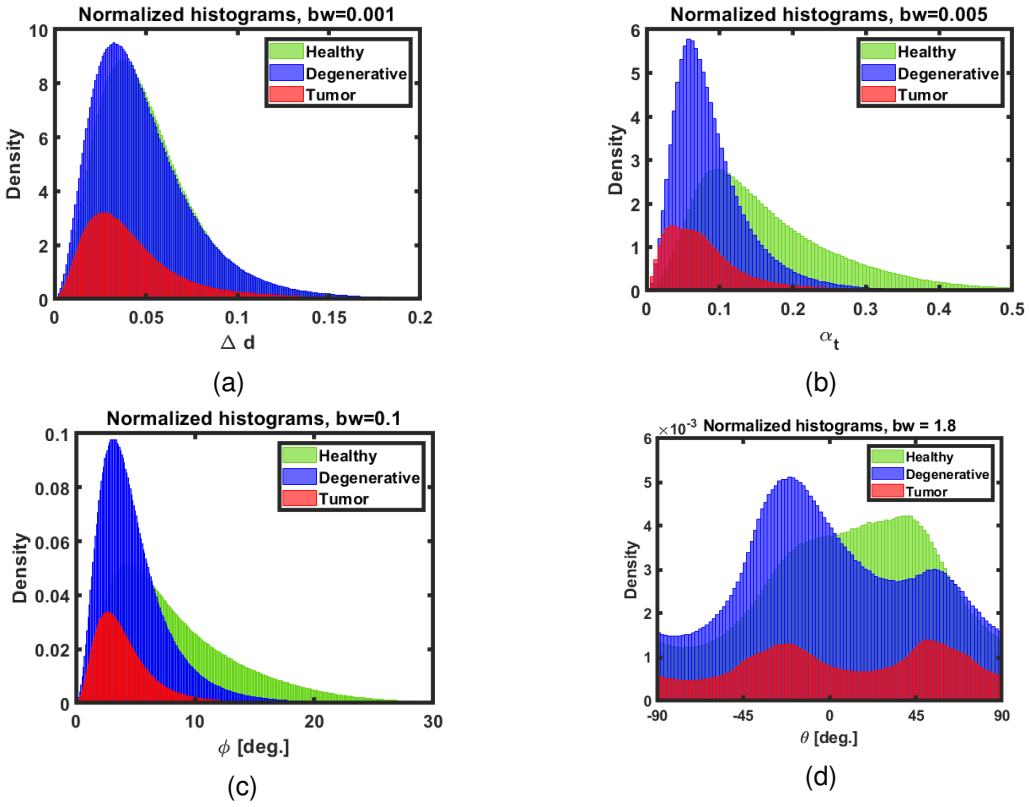


Figure 39: Normalized histograms: (a) dichroism – Δd, (b) depolarization – α_t, (c) retardance – ϕ and (d) orientation angle of the optic axis – θ.

As can be seen from the histograms in Fig39 that include all measurements and samples, all healthy samples demonstrate increased depolarization and retardance, while the distribution of the orientation angle varies significantly in comparison with the other two classes.

The two most sensitive parameters appear to be the depolarization and the retardance, where both parameters show the same trends, namely: they have the highest for the healthy skin samples, lower values for the degenerative skin samples and the lowest values for the tumor samples, respectively. The orientation angle can also be considered as a sensitive parameter, based on its bimodal distributions for the degenerative and tumor skin tissue, contrary to the distribution from the healthy skin tissue measurements. Unfortunately, the distributions of dichroism for both healthy and degenerative skin tissue measurements are indistinguishable, the distribution for the tumor tissue measurements shows slightly lower values of dichroism.

For the healthy skin the extracellular matrix is comprised of the structural proteins and fibers, thus, forming a structure that is highly sensitive to polarized light. All degenerative and malignant formations in skin are characterized by the changes in the extracellular matrix and/or an infiltration of abnormal cells that have usually an increased size and irregular shape. It leads to the loss of tissue anisotropy and changes in morphology, resulting in different polarization response and increase in a contrast between the healthy skin and the pathological samples.

Nevertheless, although the results are quite promising up to this point, in the best case scenario, there should not be an overlap between all distributions, otherwise it may be difficult to differentiate between the three classes with different pathological status. Therefore, to overcome this drawback, the data are further post-processed with the artificial neural network model.

4.3.2 Deep learning

For the purpose of the current dissertation, 15 images per class from all skin lesions were selected, based on the medical doctors' expertise. The classes were separated into three groups: Healthy – 0, Degenerative – 1 and Tumorous – 2. To increase the number of images for the data set, each image was segmented into 12 patches of the size 200$\times$200 pixels. Additionally, once filtered for the glass substrate contribution, the images of all 15 Mueller matrix elements were used as predictors (instead of using polarimetric parameters after decomposing the Mueller matrices), apart from m_{11}, due to its unit value after the normalization. By this way, all predictors were normalized in the range between [-1, 1] and the correlation between the predictors was reduced/avoided. This approach is not only time-consuming and computationally effective, as, generally, the depolarization and/or polarization information is encoded in the Mueller matrix elements of a given sample. Hence, for the purpose of creating a polarimetric model for classification, the decompositions could be avoided. In the end, in total we have three classes for classification, hence 3 x 15 x 15 x 12 or 8100 images. For each image, a mask was created to indicate its class, as mentioned above. This amount of data was found sufficient to further split into the data sets for training, validation and testing with proportions 50:25:25 or 4050:2025:2025 images. By this way, in the validation phase we optimize the model for the training phase and

evaluate its performance with unseen data in the testing phase. Conventionally, we have the images in a single channel, which may not be always enough to make a differentiation between all classes. Instead, we want the neural network to learn more complex features from the images, unnoticeable for the medical doctors, such as edges, curves, lines or shapes on a pixel level. By this way, the diagnostic analysis by the medical doctors may be substantially complemented. For this task a convolutional neural network (CNN) was used. Its architecture is provided in Fig. 40 and Fig. 41. The detailed explanation about the principles of operation behind the current neural network could be found in the Appendix Section 6. In total, our model must process and calculate 3 666 179 parameters, which would require a sufficient amount of computational time. Therefore, to decrease the computational time, a GPU (Graphics Processor Unit) acceleration was also used. For the sake of differentiation between all models used in the dissertation, the deep learning model was given symbolically the name: Skin-HDT700 (deep learning model with skin samples of healthy, degenerative and tumor classes, measured at the wavelength of 700 nm). The use of this particular type of neural network (CNN), will provide us the ability to use the polarimetric images of the corresponding Mueller matrices for the development of a model for computer vision applications. By this way, this deep learning model must be able to recognise shapes, patterns and structures, related to the different skin lesions. When there is a change in the tissue inner structure and reorganization upon degenerative disease or tumor formation, the CNN model should be able to recognise on a pixel level such kind of alterations. Then, not only the medical doctors could be supported with another optical technique for the diagnosis, but this approach may have a potential to diagnose at an early stage the development of a particular type of skin disease.

Firstly, the model learns from the training data and then it tries to minimize the loss function before making predictions on the labeled training data. Then, the model also tries to minimize the loss function of the validation data and to additionally perform classifications on it. Whenever there is a poor performance of the model, the hyperparameters' tuning is applied and the learning process is repeated until reaching satisfactory results. During the training phase, an optimizer "adam" [93, 94] was used, in such a way so that no initial tuning of the learning rate was needed. Instead, this was done iteratively from the build-in Python libraries and the "adam" optimizer. For the loss metric, the cross-entropy was chosen, while 30 epochs with a batch size of 16 were chosen during the training phase.

In Fig.42 the results for the training process are shown, presenting the evolution of the loss function and the model accuracy, with respected to the epochs number. It could be well observed, that the loss functions are optimized until the 15-th epoch, whereafter there is a saturation in loss function's values until the end of the learning process. Analogously, the accuracy is almost equal to unity after the 15-th epoch. The averaged accuracy of the training data set after the classification was found to be 0.98, which is a satisfactory result. Additionally, both curves for the loss function and the model accuracy are almost overlapping, indicating a minimum (or none at all) of overfitting, as initially was targeted. Hence, a good balance was reached between the model complexity with the available data set and the desired accuracy of the classification of skin images with different histological conditions.

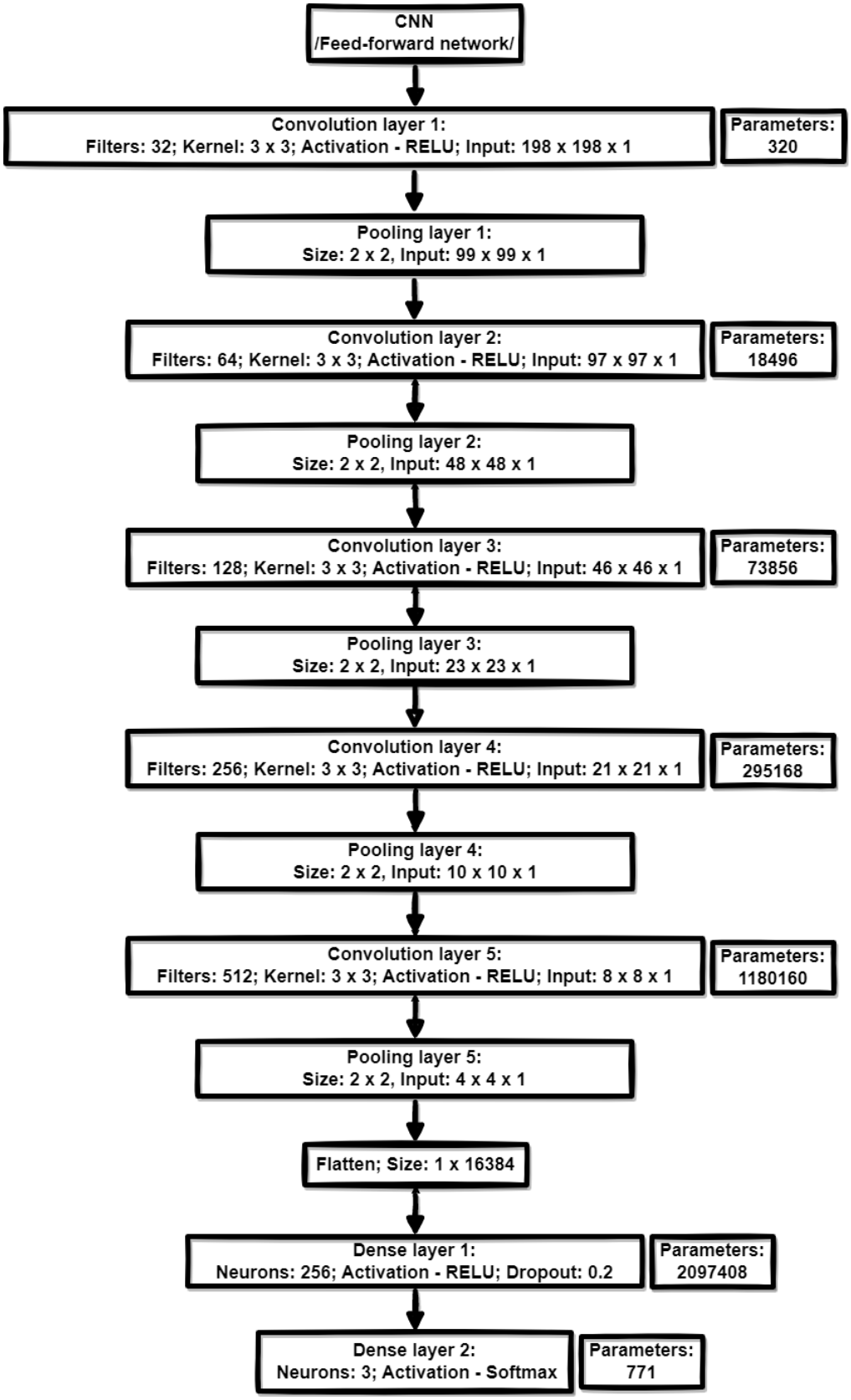


Figure 40: Architecture of the CNN as a flowchart.

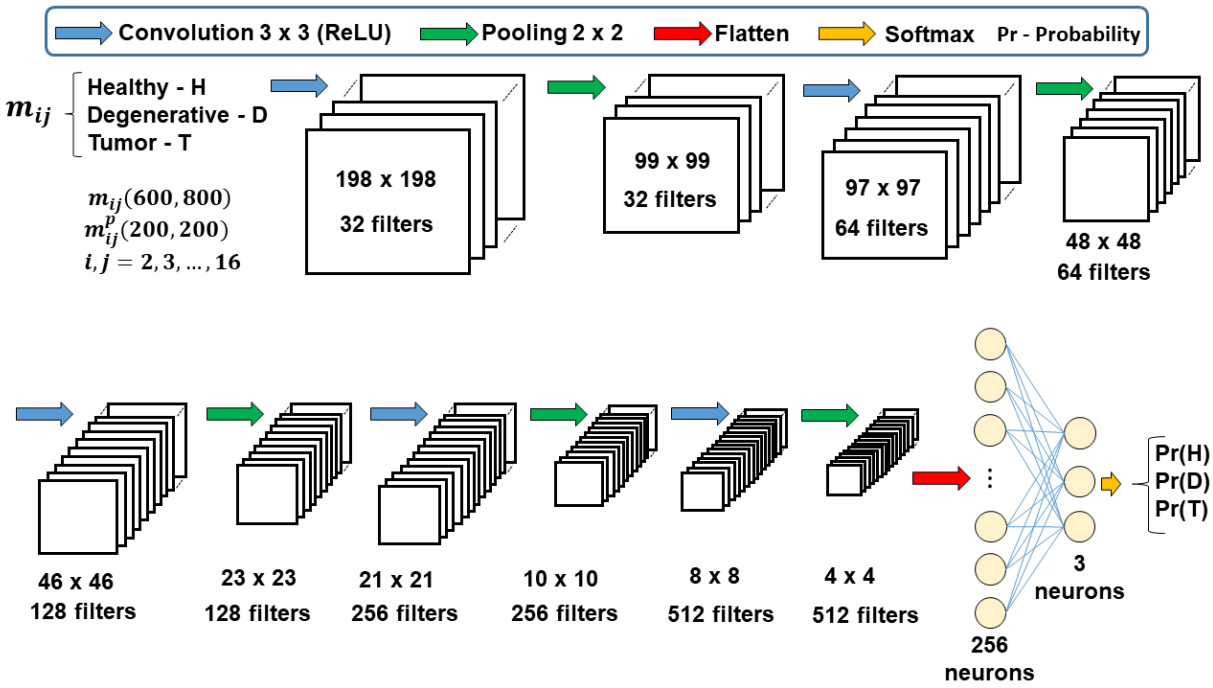
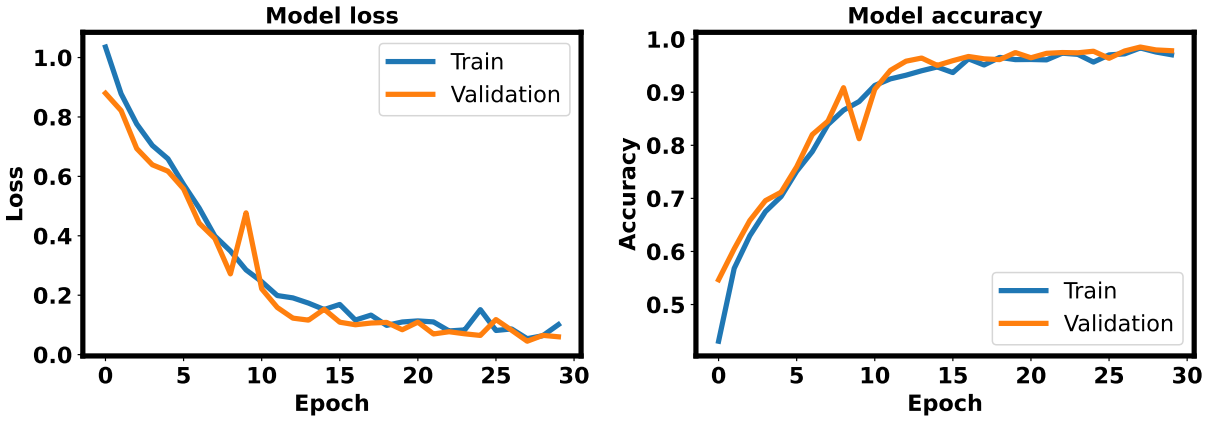


Figure 41: Architecture of the CNN as an illustration.



(a) (b)

Figure 42: CNN evaluation during the training and validation phases: **(a)** model loss and **(b)** model accuracy.

Next, in order to evaluate the real model's performance with unseen data, predictions and classifications were made with the testing data set. From 2025 labelled testing images, only 41 were wrongly classified, which again confirms the excellent performance of the Skin-HDT700 model. For better clarity, the confusion matrix is shown in the Fig.43:

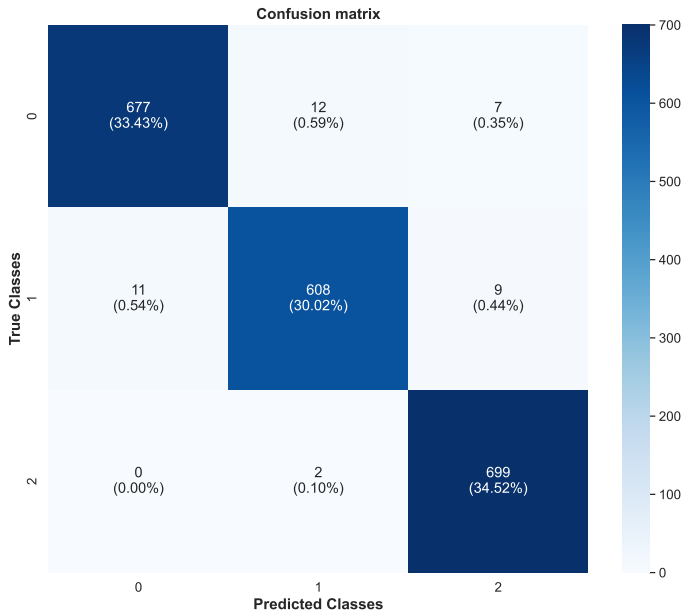
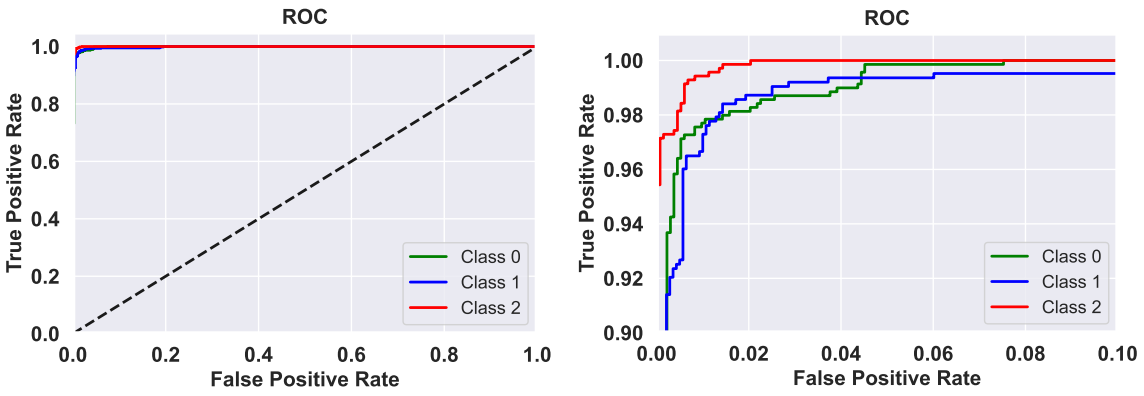


Figure 43: Confusion matrix (CM) of the Skin-HDT700 model. The notations stand for 0 - healthy, 1 - degenerative and 2 - tumor classes.

The main diagonal of the confusion matrix shows the truly predicted classes from the model. On the other hand, the off-diagonal elements represent wrongly classified images. Hence, from the healthy class, the model classifies wrongly 12 images as a degenerative disease and 7 - as a tumor. Analogously, for the degenerative diseases class, the model classifies wrongly 11 images as healthy tissue and 9 - as a tumor. Finally, only 2 images from the tumor class were wrongly classified as a degenerative disease, while none - as healthy tissue.



(a) (b)

Figure 44: ROC curves for the CNN and all classes of the Skin-HDT700 model: **(a)** whole ranges of the false positive and true negative rates, dashed line represents the ROC curve for the completely random classifier and **(b)** reduced ranges for better clarity.

Next, the ROC curves for all three classes are shown in Fig.44, where it could be observed close to perfect square profiles for all three classes. By this way, the model has a good potential to be further developed and fine-tuned with more experimental data and with additional hyperparameter tuning. After its creation, it can be used as a basis of more robust model (a.k.a transfer learning).

4.4 Colon cancer detection: ex-vivo studies

4.4.1 Polarimetric results

This section summarizes the results obtained from a formalin-fixed human, colon specimen. The patient was diagnosed with the colon cancer and underwent the surgery. After, the excision of the diseased part of a colon the biological specimen was cut into two parts - for the histological and optical examination, respectively. The tumor staging was performed by the pathologist during the histology analysis confirming the presence of a tumor zone containing G2-adenocarcinoma (T2, N0, M0). Later on the results of the histological analysis were used as a ground truth for the polarimetric diagnostic measurements. Both healthy and tumor zone thicknesses were about 1 mm; therefore, the polarimetric measurements were done in reflection, using the experimental set-up already described in Chapter 2. To enrich the polarimetric analysis, the polarization properties of a virtual phantom were simulated by means of polarized Monte Carlo algorithm. Also, the measurements of a physical phantom with 2 mm thickness were performed in reflection in order to mimic the polarization response of the real colon tissue specimen in the red to infrared spectral range. A detailed description of the phantom manufacturing can be found in [95, 96]. In short, the main structural difference between all studied turbid media is in their inner structural organization. For example, the virtual phantom is comprised of spherical scatterers of the fixed size and refractive index that are uniformly distributed within optically isotropic host medium, while on the contrary, for all physical samples these assumptions do not hold. Especially, the colon tissue specimen contains scattering particles of various size, most of them are non-spherical and have fluctuating refraction index. For the selected experimental geometry and the described set of samples, the physical realizability filtering, indices of polarimetric purity and the symmetric decomposition algorithms were used (see Chapter 1) in order to extract their polarization response from the recorded/simulated Mueller matrices. All results are presented and discussed below.

Firstly, the general form of all Mueller matrices is very close to that of a diagonal depolarizer. This is especially noticeable for the simulated Mueller matrix of the virtual phantom. However, the off-diagonal elements are non-zero, which in turn is an indication of other polarimetric properties encoded in the Mueller matrices. There is a difference in the angle of incidence and the detection angle and such kind of experimental geometry may be the reason for the observed slight violation from the theoretical model that predicts strictly zero values for the off-diagonal elements of Mueller matrix. Secondly, a rotational symmetry of reflection would yield an equality $m_{22} = -m_{33}$, whereas in Fig. 45 one can easily observe a violation of this criterion for the Mueller matrix of the real phantom. Potentially, this could be due to the presence of non-spherical particles, which are not isotropically oriented. In general, the polarimetric properties of the virtual phantom are closer to that of colon tumor tissue, while the polarimetric properties of a real phantom are closer to the properties of healthy colon tissue. Control of this phenomenon was achieved by simulating an isotropic host medium with ideally spherical scatterers. Setting the scattering coefficient of the real phantom ten times higher in comparison to the virtual one results in an decreased photon mean free path between two scattering events. Consequently, an increase in the number of scattering events leads to an increased randomization of the initial polarization state and higher loss of polarization degree. All real samples are characterized by significant depolarization, with inequality relation of $d_1 > d_2 > d_3$ or higher depolarization for circularly polarized light compared to linearly polarized light - a typical signature of the Rayleigh scattering regime [47].

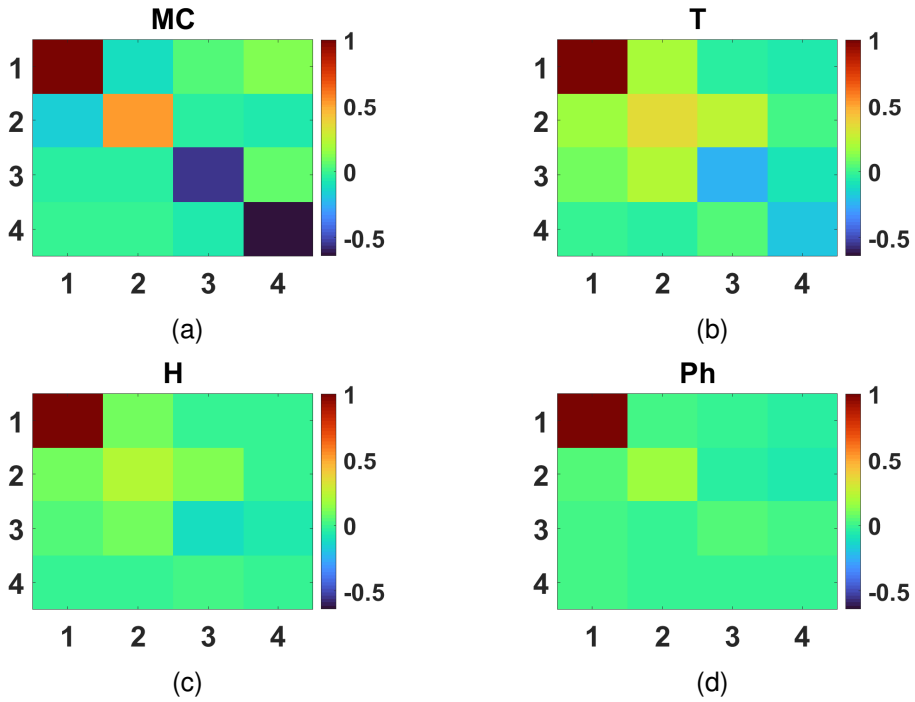


Figure 45: Simulated and measured Mueller matrices: (a) Monte Carlo (MC) (b) Tumor (T), (c) Healthy (H) and (d) Phantom (Ph). Results available from the published material in [63].

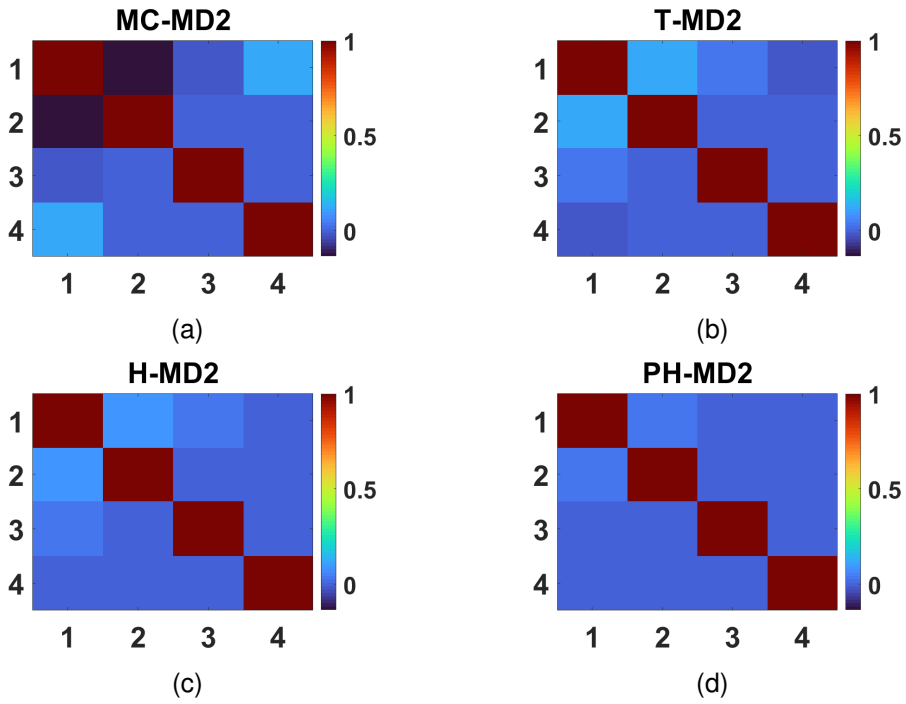


Figure 46: Mueller matrices of the Diattenuator D_2 obtained from the symmetric decomposition for: (a) Monte Carlo (MC) (b) Tumor (T), (c) Healthy (H) and (d) Phantom (Ph). Results available from the published material in [63].

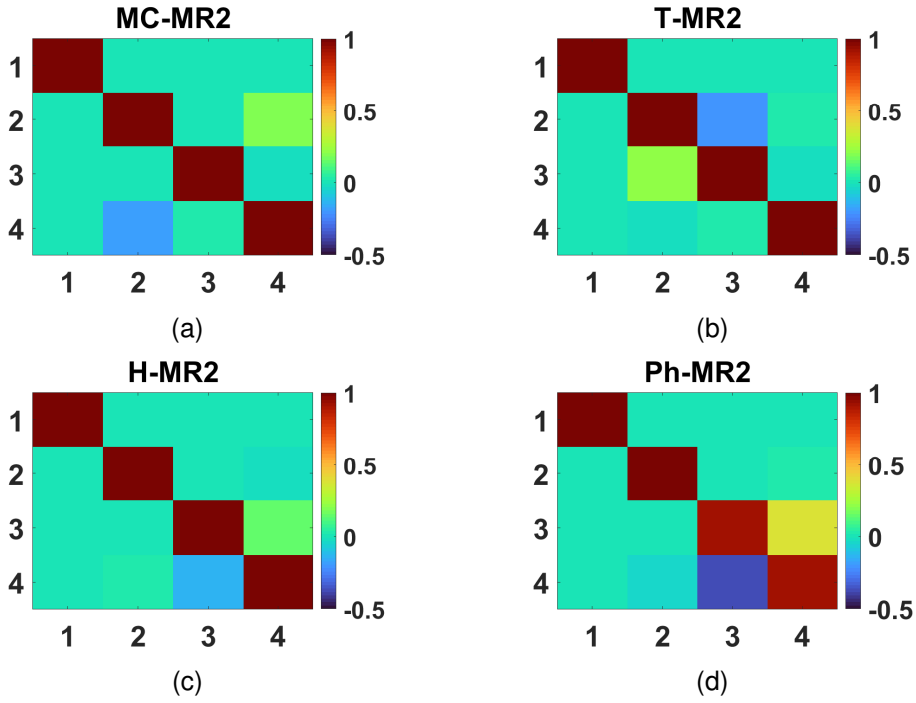


Figure 47: Mueller matrices of the Retarder R_2 obtained from the symmetrical decomposition for: (a) Monte Carlo (MC) (b) Tumor (T), (c) Healthy (H) and (d) Phantom (Ph). Results available from the published material in [63].

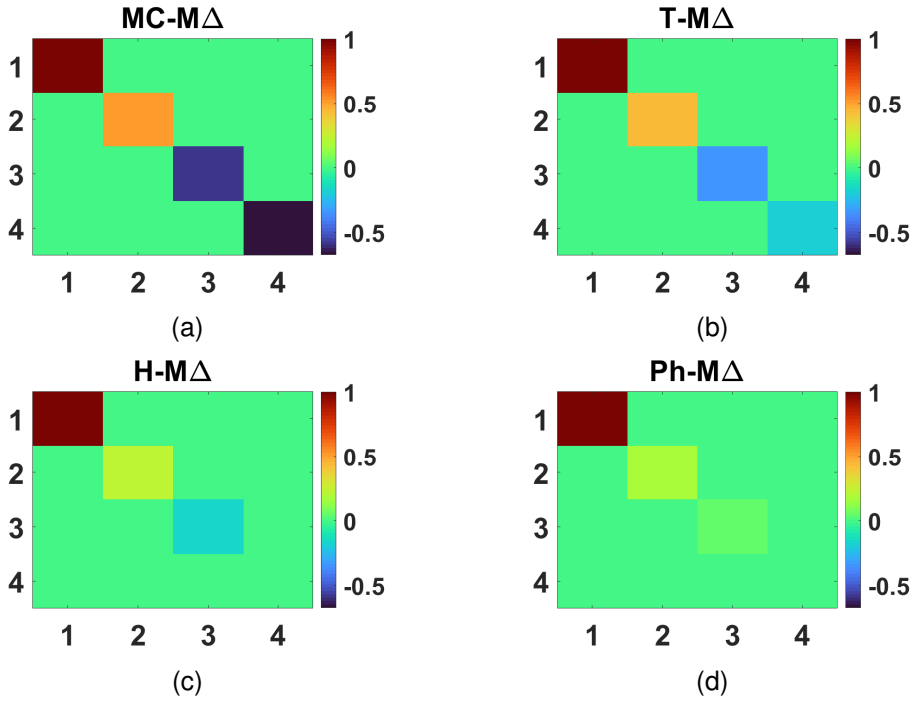


Figure 48: Mueller matrices of the canonical depolarizer M_{Δ} obtained from the symmetrical decomposition for: (a) Monte Carlo (MC) (b) Tumor (T), (c) Healthy (H) and (d) Phantom (Ph). Results available from the published material in [63].

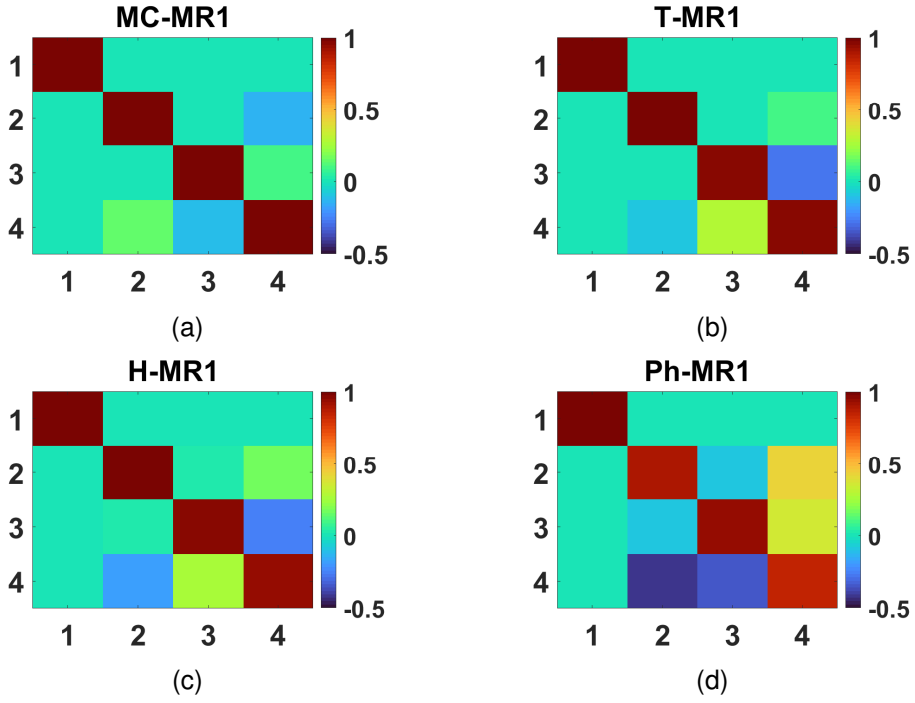


Figure 49: Mueller matrices of the Retarder R_1 obtained from the symmetric decomposition for: (a) Monte Carlo (MC) (b) Tumor (T), (c) Healthy (H) and (d) Phantom (Ph). Results available from the published material in [63].

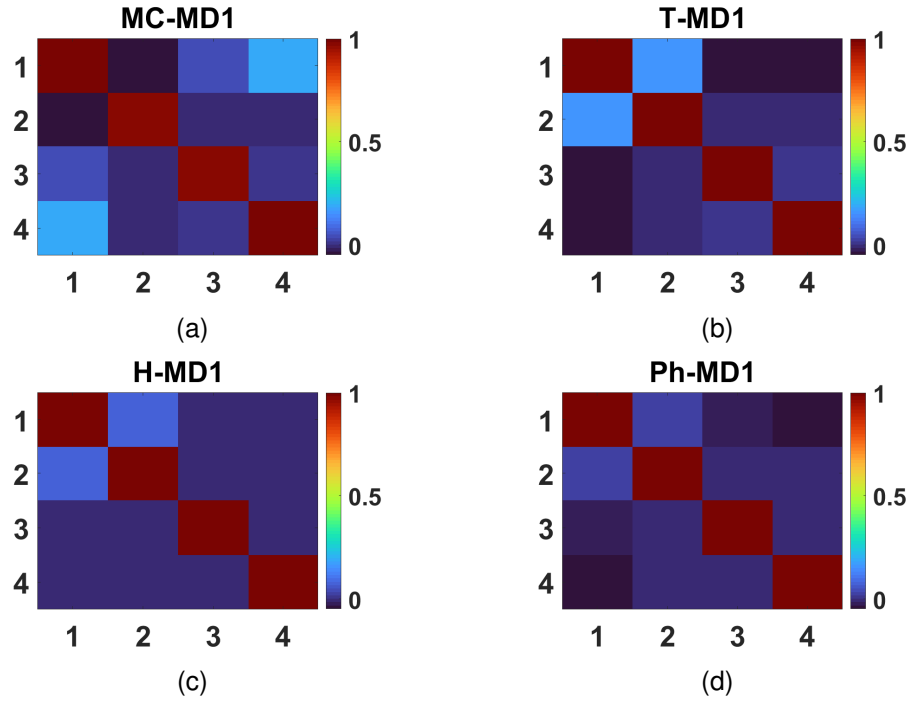


Figure 50: Mueller matrices of the Diattenuator D_1 obtained from the symmetric decomposition for: (a) Monte Carlo (MC) (b) Tumor (T), (c) Healthy (H) and (d) Phantom (Ph). Results available from the published material in [63].

In Table 2 all numerical values of the polarimetric quantities obtained from the symmetric decomposition are summarized.

	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	Δ	$\mathbf{D_1}$	$\mathbf{D_2}$	ϕ_1 **[deg.]**	ϕ_2 **[deg.]**
MC	0.52	0.57	0.69	0.41	0.19	0.18	10.39	10.69
T	0.43	0.34	0.19	0.68	0.17	0.13	17.17	2.47
H	0.24	0.16	0.01	0.87	0.08	0.09	17.81	8.14
PH	0.17	0.05	0.004	0.93	0.05	0.03	19.77	20.73

Table 2: Supplementary table associated with decomposition data. Data available from the published material in [63].

Also, both polarization and depolarization properties of all turbid media can be summarized in Fig. 51, where with **I** the identity matrix was denoted, such as **M*ij*** = δij and with **ID** - the Mueller matrix of an ideal depolarizer correspondingly. The former has only polarization and no depolarization properties, while the latter has only depolarization properties. In Fig.51 all red lines serve as a *guide-to-the-eye* and should not be interpreted as the fit curves, since only categorical variables are presented on the *x*-axes. By this way, it becomes possible to assume a monotonic decrease of all depolarization coefficients d_i and a monotonic increase for the depolarization index Δ. As expected, the simulation predicts less depolarization compared to real samples. Most importantly, the healthy colon tissue excels the depolarization of light in comparison with the malignant colon tissue due to different amount of scattering events within

healthy and tumor zones. As for D_i and ϕ_i values, clearly they are characterized either by decrease for the former and by an increase for the latter, depending on the corresponding turbid media. Nevertheless, several important tissue properties can be elicited. Namely, we observe an increased diattenuation value for the tumor tissue and slightly higher retardance value for the healthy one.

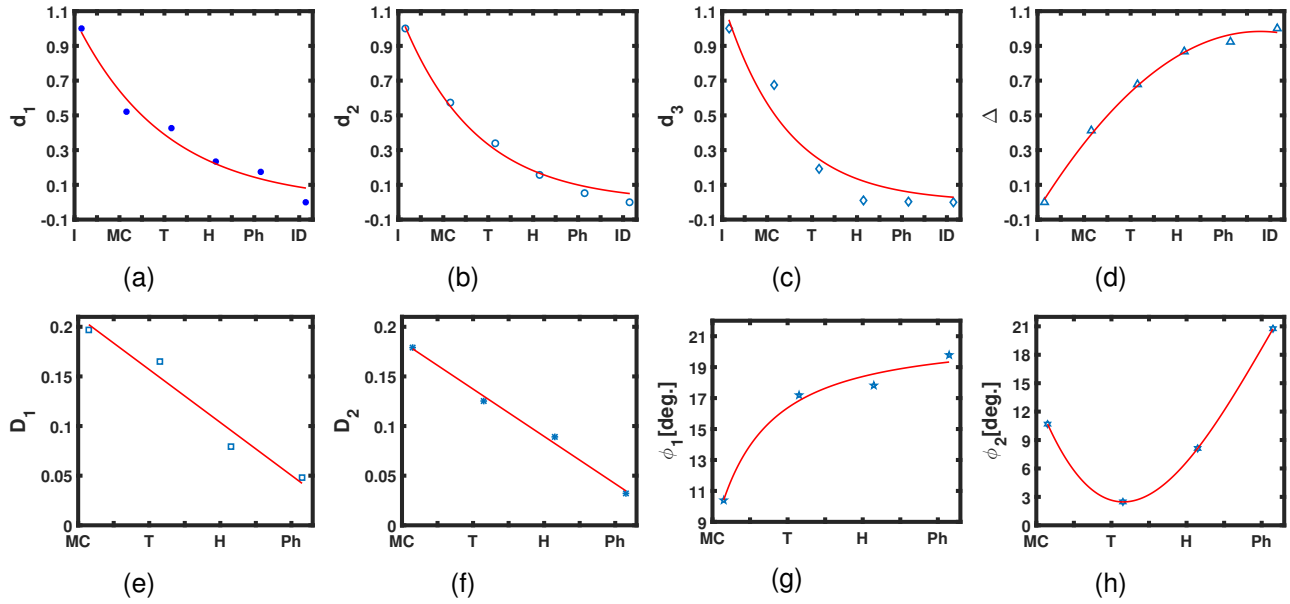


Figure 51: Symmetric decomposition products: a), b), c) depolarization indices – d_i; d) net depolarization index – Δ; e), f) diattenuadtion D_i and g), h) retardance – ϕ_i. Results available from the published material in [63].

So far only single-shot measurements were considered. In order to trace the spatial distribution of polarimetric parameter both healthy and cancer colon zones were scanned independently over a region of interest (ROI) of 1 mm^2 and a step size of 0.2 mm in x-y directions by using a motorized translation stage. For each one of the four input polarization state measurements, the degree of polarization ρ of the measured output Stokes vector was calculated and a normalization procedure was applied (see Eq. (77)). A total of 36 Mueller matrices were obtained for all scanning points within each tissue zone (healthy and cancerous) via Eq. (79). All Mueller matrices were tested and filtered in accordance with the physical realizability criterion. As a result, three matrices per tissue zone were discarded from the data set due to non-compliance. The remaining matrices were decomposed using the symmetric decomposition and the depolarization metric calculations were performed. By visualizing the position of all measurement data points in different 3D depolarization spaces, potentially the support of polarimetry to histopathology analysis may be reinforced. In Fig.52 there are two sets of points, marked by green and red color. They represent the respective values for the cancerous and the healthy tissue zones. Some of the tumor measurements are very closely grouped to the measurements from healthy tissue cluster, making these values diagnostically irrelevant. Fortunately, the rest of the tumor data points are grouped separately from the healthy ones. Hence, these plots enable us to differentiate between the two tissue zones with different pathological status assuming a significant decrease of depolarization presented in the corresponding tumor ROI.

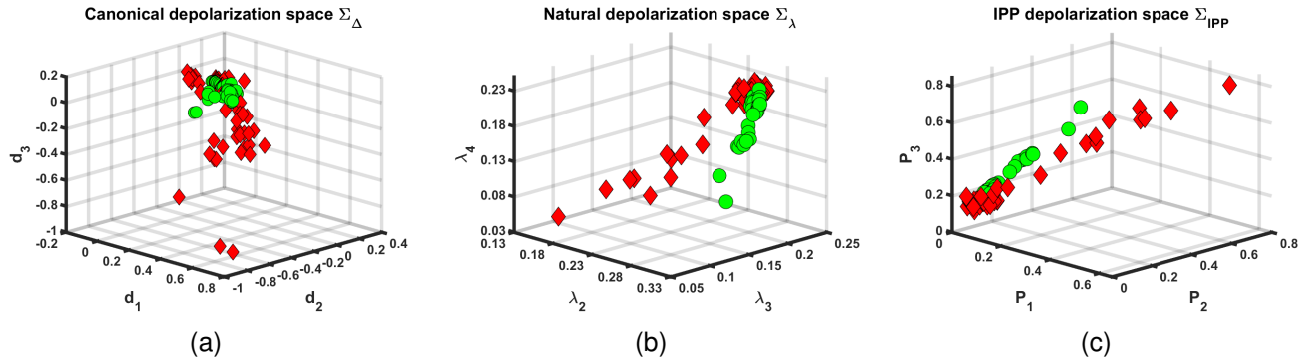


Figure 52: 3D representation of depolarization: (a) Canonical, (b) Natural, (c) Indices of polarimetric purity depolarization spaces, where the symbols represent ● – Healthy, ◆ – Tumor data points. Results available from the published material in [97].

Then, we investigated the ability of the net diattenuation D, the polarizance P, and of the two pairs of polarization parameters (D_i and R_i, $i = 1, 2$) from the symmetric decomposition to discriminate the tumor zone of the colon specimen from the healthy one. As can be seen, both plots in Fig.53(a) and (b) hold great potential to be considered as diagnostically relevant and to support colon histopathology analysis, because they represent two different zones (healthy and cancer) as two distinctive separable clusters. However, the retardance values demonstrate large fluctuations and cannot provide accurate diagnostic information for this set of measurements.

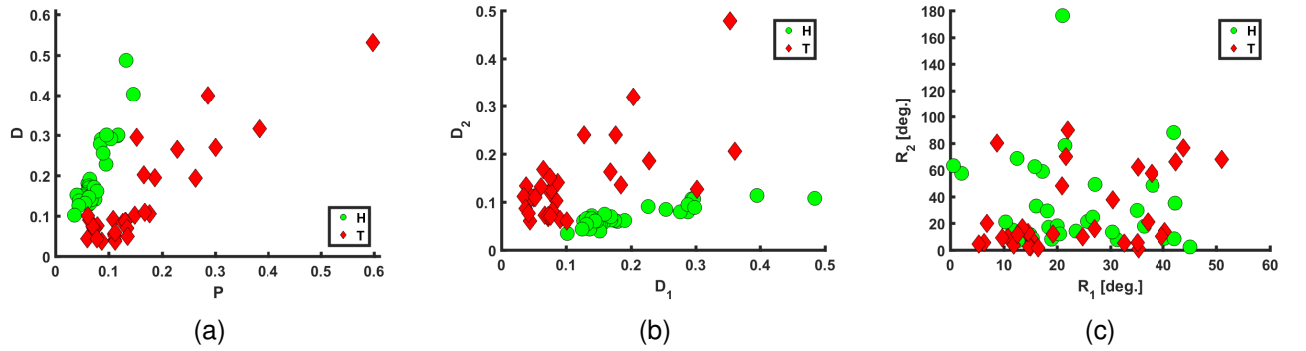


Figure 53: Scatter plots of: (a) D vs P, (b) D_2 vs D_1 and (c) R_2 vs R_1, for both healthy green points) and tumor (red points) zones, where $\phi_{1,2} \equiv R_{1,2}$. Results available from the published material in [97].

When 3-dimensional representation is used, it could be also observed in Fig.54, that both P and D_2 values are very different for the cancerous and the healthy zones of the colon specimen. As the polarimetric measurements were performed at angles of incidence and detection different from the normal to the sample surface, the impact of surface topography is enhanced. The latter may explain the increased diagnostic value of both P and D_2 parameters, which are likely affected by surface and/or volume scattering.

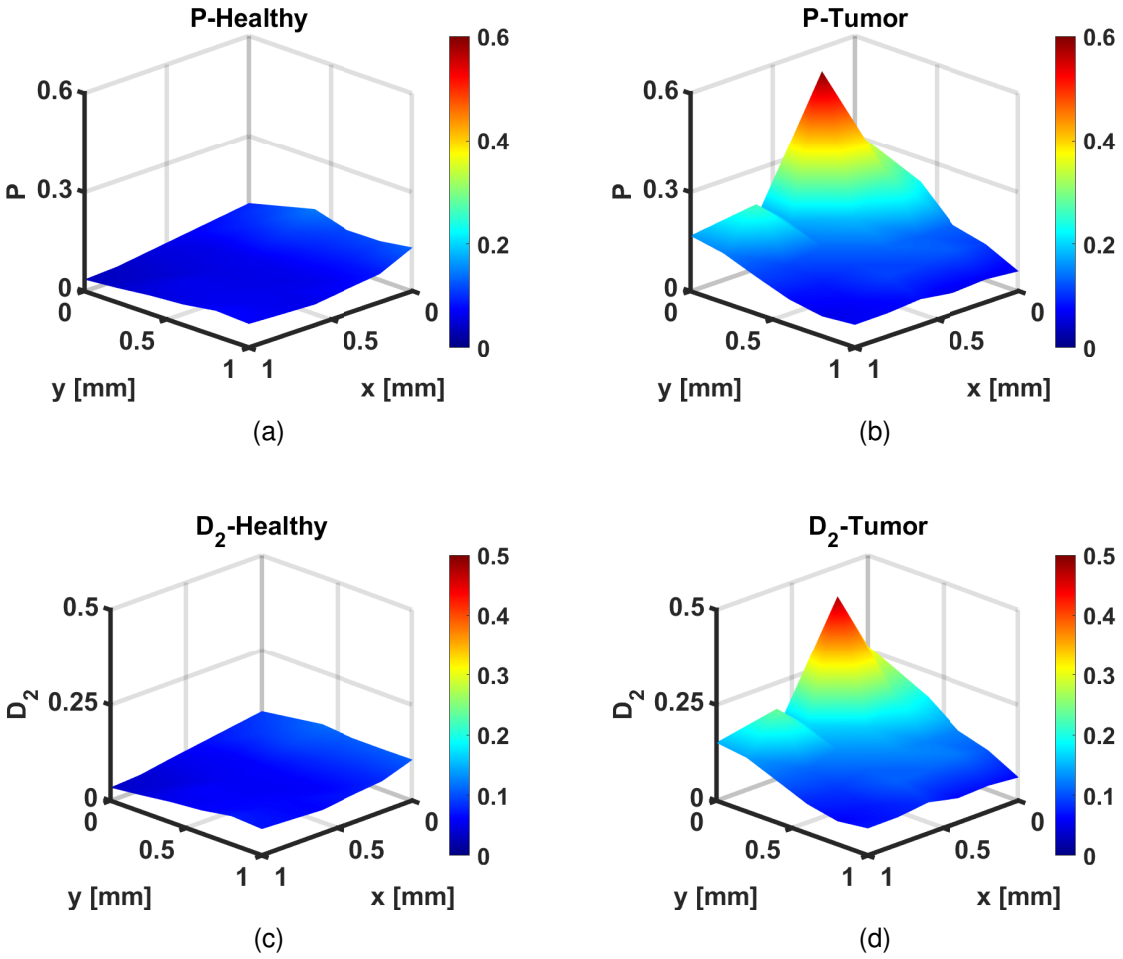


Figure 54: Spatial distribution of the values of polarizance and diattenuation within both tissue zones: (a) P – Healthy, (b) P – Tumor, (c) D_2 – Healthy, (d) D_2 – Tumor. Results available from the published material in [97].

Upon structural modifications in tissues due to malignancy, an alteration of the ECM of collagen is followed by destruction of the collagen cross-links. The effective size of the comprising scattering particles is changed too. Thus, both the polarization and depolarization properties of tissues are inevitably affected when tumors are developing. For this particular reason, an indirect measurement of the polarization entropy S could be regarded as a representative quantity of the tissue spatial heterogeneity. Also, the net depolarization can be characterized with the spatial distribution of Δ (Fig.55a,b) for both tissue sections. Low values of Δ indicate low depolarization of the probing light by the sample and the plots in Fig.55 indicate lower depolarization for the tumor tissue zone suggesting that this parameter could be potentially used as an optical indicator for the tumor spread throughout the scanned tissue zone. Similarly, Fig.55c,d reveals a lower entropy for the cancer zone in comparison to the healthy tissue zone. Specifically, the same spatial location within the cancerous zone is characterized by both lower depolarization Δ from Fig.55a,b and lower entropy S.

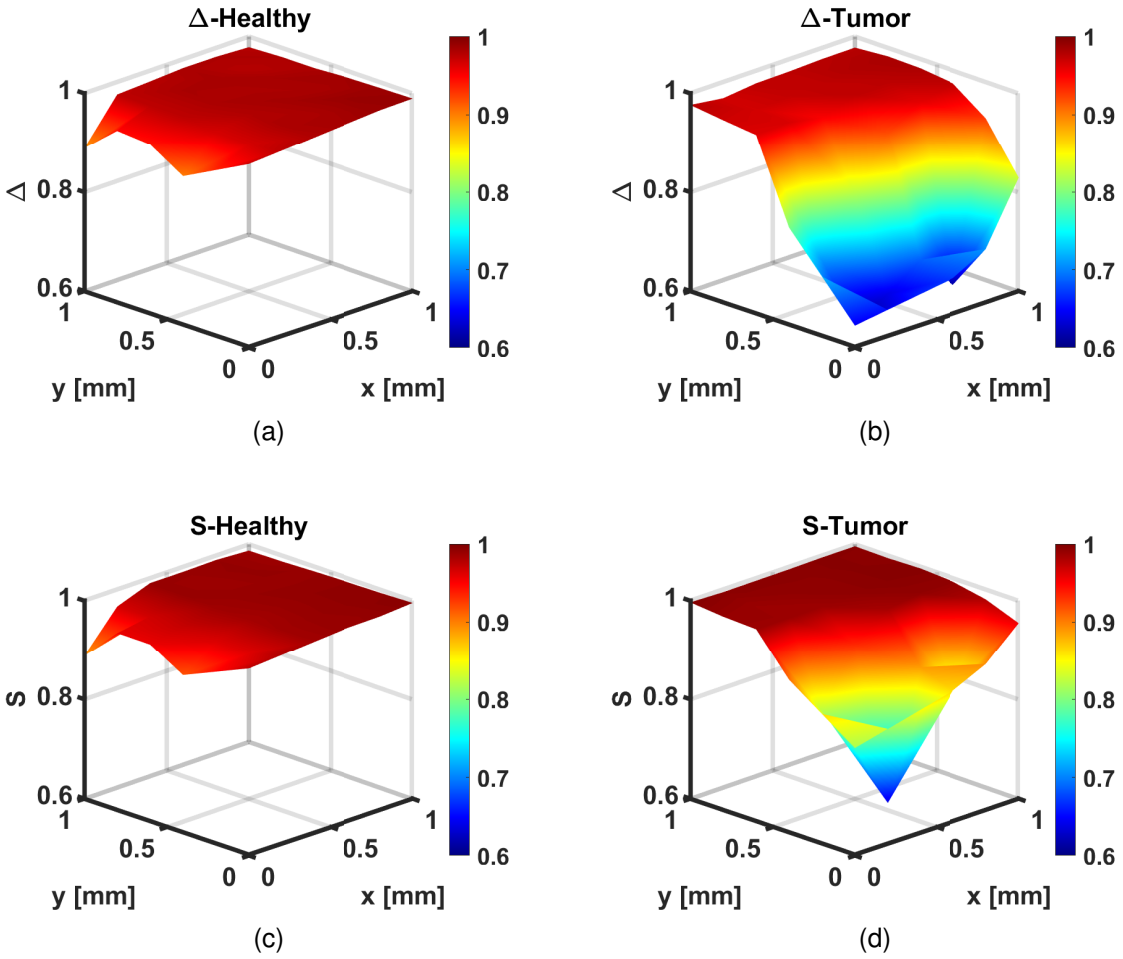


Figure 55: Spatial distribution of the values of depolarization Δ and polarization entropy S within both tissue zones: (a) Δ – Healthy, (b) Δ – Tumor, (c) S – Healthy, (d) S – Tumor. Results available from the published material in [97].

Even more depolarization information can be extracted by analyzing the spatial distributions of PI and P_{Δ}. Fig.56 reveals two important features for both histological conditions – namely a higher polarimetric purity and an increased value of P_{Δ} depolarization index for the tumor tissue zone compared to the healthy one. As a result, the cancerous tissue zone acts as weaker depolarizer compared to the healthy colon tissue, as previously discussed and observed in prior studies [98]. A parallel should be drawn between the depolarization Δ and the depolarization index P_{Δ}. Their interchangeable use should be avoided. As can be seen from Fig.55 and 56, the same scanned area is characterized by opposite magnitude values of Δ and P_{Δ}.

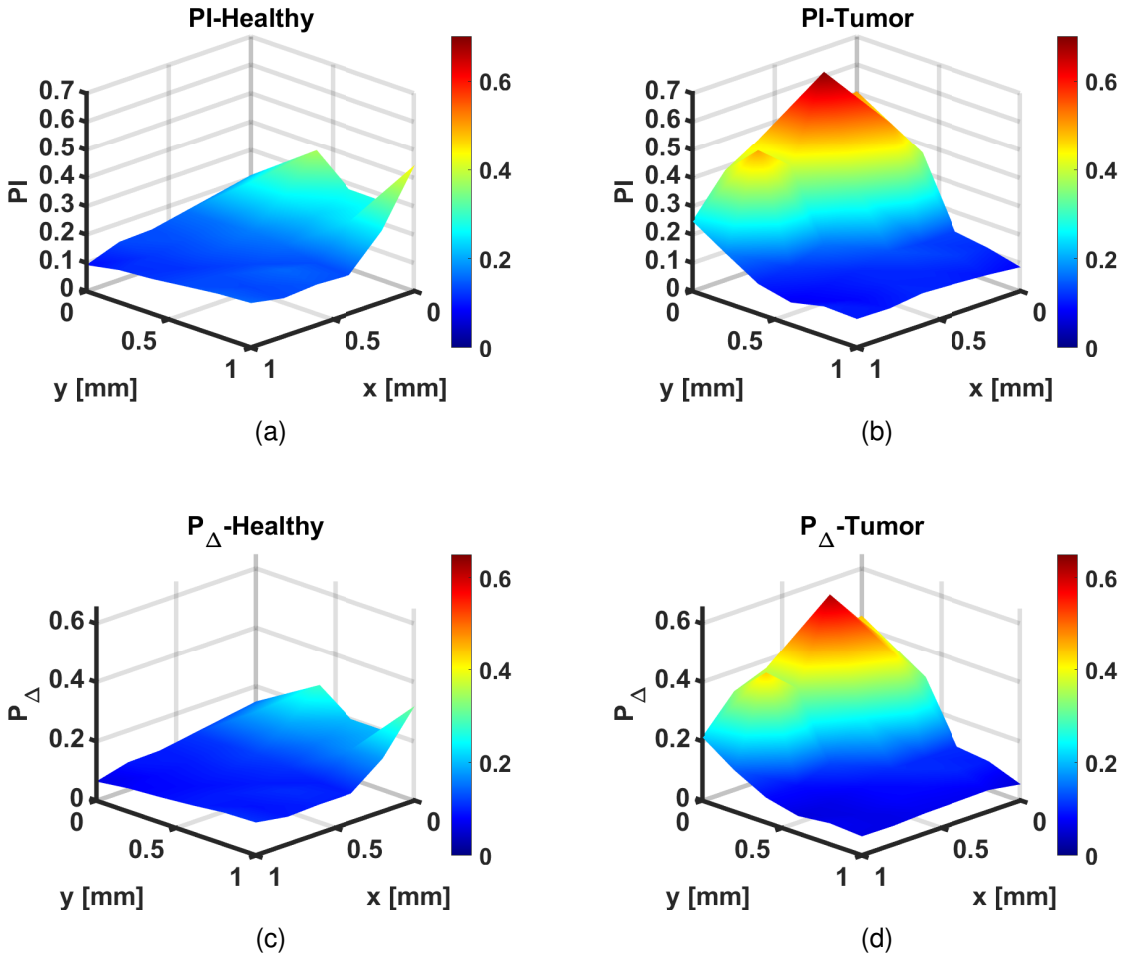


Figure 56: Spatial distribution of the values of the polarimetric purity – *PI* and the Gil-Bernabeu's depolarization index – P_{Δ} within both tissue zones: (a) *PI* – Healthy, (b) *PI* – Tumor, (c) P_{Δ} – Healthy, (d) P_{Δ} – Tumor. Results available from the published material in [97].

4.4.2 Statistical analysis

After including the results from 5 *ex vivo* thick colon samples, in total 670 measurements for both histological conditions were measured. The investigated samples include the colon and gastric adenocarcinomas at different stages: 1) G2-moderately differentiated (intermediate grade) and 3) G3-poorly differentiated (high grade), of varying thicknesses. The samples were provided from multiple patients, thus *a priori* it was not expected to have such a good differentiation as in the case of single-shot measurements and/or scanning of a single sample, because of the inter-patient variability. The following graphical representations in Fig.57 are illustrating this problem:

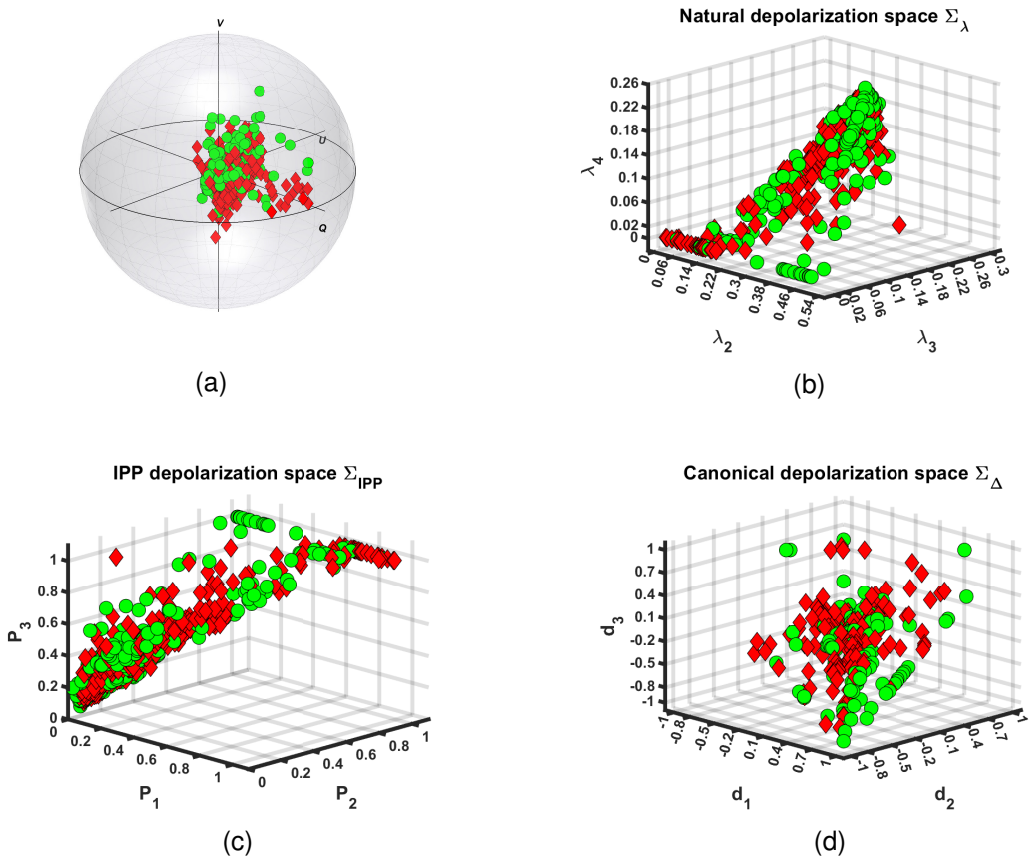


Figure 57: Visualization of polarimetric data sets in different parametric spaces at all spatial locations of the measurements of both healthy and tumor zones of colon tissue ● – healthy and ◆ – tumor via: **(a)** Poincaré sphere for probing (or incident) circular polarization, **(b)** natural, **(c)** IPP and **(d)** canonical depolarization spaces. Results available from the published material in [99].

As can be seen from Fig.57, upon the inclusion of all experimental data from various colon samples with tumors at different stages of development, a superimposing of the majority of data points from both zones of colon tissue is observed. Hence, as expected, the inter-patient variability restricts us to evaluate two separate clusters corresponding to the measurements of healthy and cancerous zones of colon specimens or to find specific trends within either Poincaré sphere or three different depolarization spaces.

After adopting the initial data post-processing sequence, it became possible to extract 20 polarimetric quantities that are to be used as the predictors: $\lambda_{1,2,3,4}$, $P_{1,2,3}$, P_Δ, *PI*, *S, D, P,* $D_{1,2}$, $d_{1,2,3}$, Δ and $\phi_{1,2}$. Initially, the mean values and their standard deviations were calculated. For both data sets from healthy and tumor zones of colon tissue the second statistical moment of the distribution of ϕ_2 was found to be approximately three times higher than the second statistical moment of the distribution of ϕ_1. Thus, we considered the parameter ϕ_2 as an unreliable predictor and, consequently, it was omitted from further analysis. Secondly, the Shapiro-Wilk normality test was computed on a significance level α = 0.05, where the test's results indicated non-Gaussian distribution for all polarimetric quantities. Thus, further on non-parametric statistical tests and machine learning algorithms, which do not require data from the normal distribution were used.

Next, for each of the polarimetric parameters pairs grouped as healthy vs tumor, Mann-Whitney test [67] was computed for the same value of α in order to find out whether the polarimetric pairs were drawn from different or similar distributions. Only for the parameters $\lambda_{1,2}$, P_1, *D, P*,and $D_{1,2}$ this test indicated that these parameters were drawn from the different distributions (*all tests were considered as statistically significant if the computed p-value* $< \alpha$).

After, the data set was reorganized with each column *j* being a polarimetric quantity, where the measurements from both health conditions were concatenated by rows. Then, a factor/categorical variable was added to indicate the health condition as either 0–healthy or 1–tumor. Finally, with the exception of the categorical variable, all other quantities were normalized with the following function $\mathcal{F}_{\mathrm{n}} = (x(j)\text{-}x_{min}(j))\cdot(x_{max}(j)\text{-}x_{min}(j))^{-1}$, in order to restrict them as dimensionless variables that vary within the closed interval [0,1]. This process is also known as the feature selection. From here on, the Principal Component Analysis can be performed, which will be assessed in the next Section. Nevertheless, explicitly the flowchart of the statistical analysis, as well as the feature selection could be summarized in another flowchart, as shown in Fig.58:

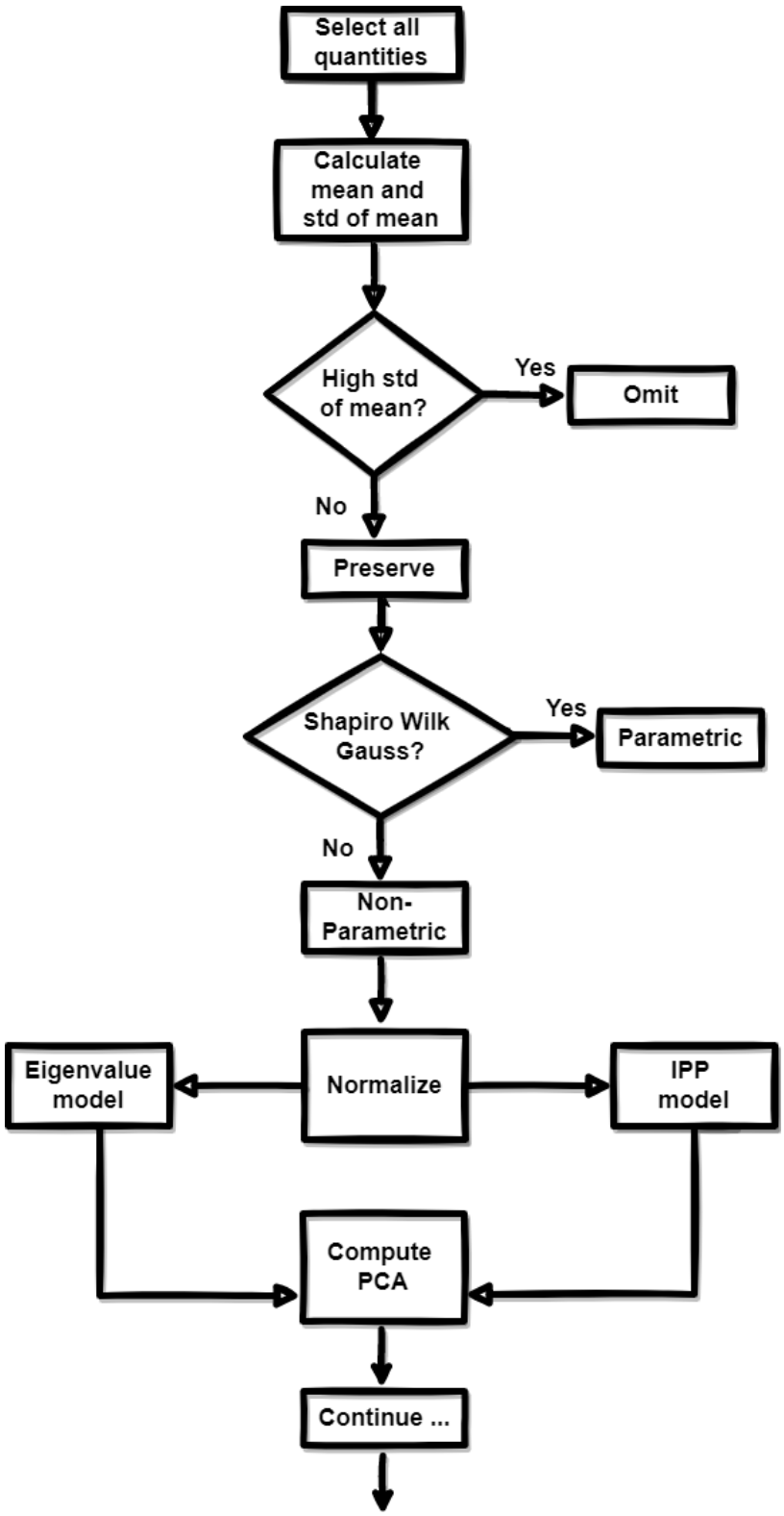


Figure 58: Flowchart of the feature selection.

4.4.3 Principal component analysis

Additional feature selection is necessary, in order to avoid the use of highly correlated predictors and multicollinearity. For instance, Δ, *S*, P_Δ and *PI* data sets were removed from the main model since they are derived from d_i, λ_i and P_i data sets and according to the Mann-Whitney test, their values for both healthy and tumor colon tissue zones are drawn from the same distribution. Moreover, high correlation is also expected for the values of λ_i and P_i as they are obtained from the Hermitian covariance matrix **H**. Therefore, two sub-models were formed: one - omitting all P_i - shortly denoted as – *"eigenvalue model"* and another - omitting all λ_i and referred as – *"IPP model"*. The remaining predictors: *D, P,* $D_{1,2}$, $d_{1,2,3}$ and ϕ_1 were included in both sub-models. In the figures below, the correlation matrices of both sub-models are shown:

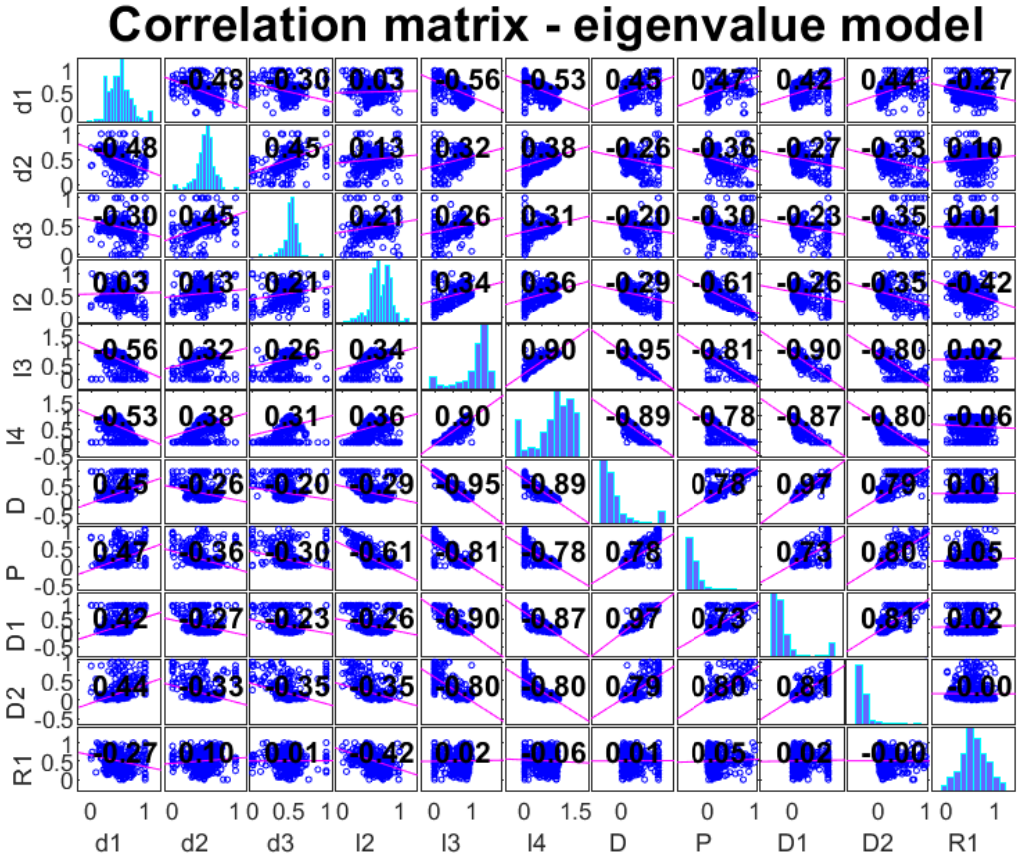


Figure 59: Correlation matrix of the eigenvalue model ($\phi_1 \equiv R_1$ and $\lambda_i \equiv l_i$).

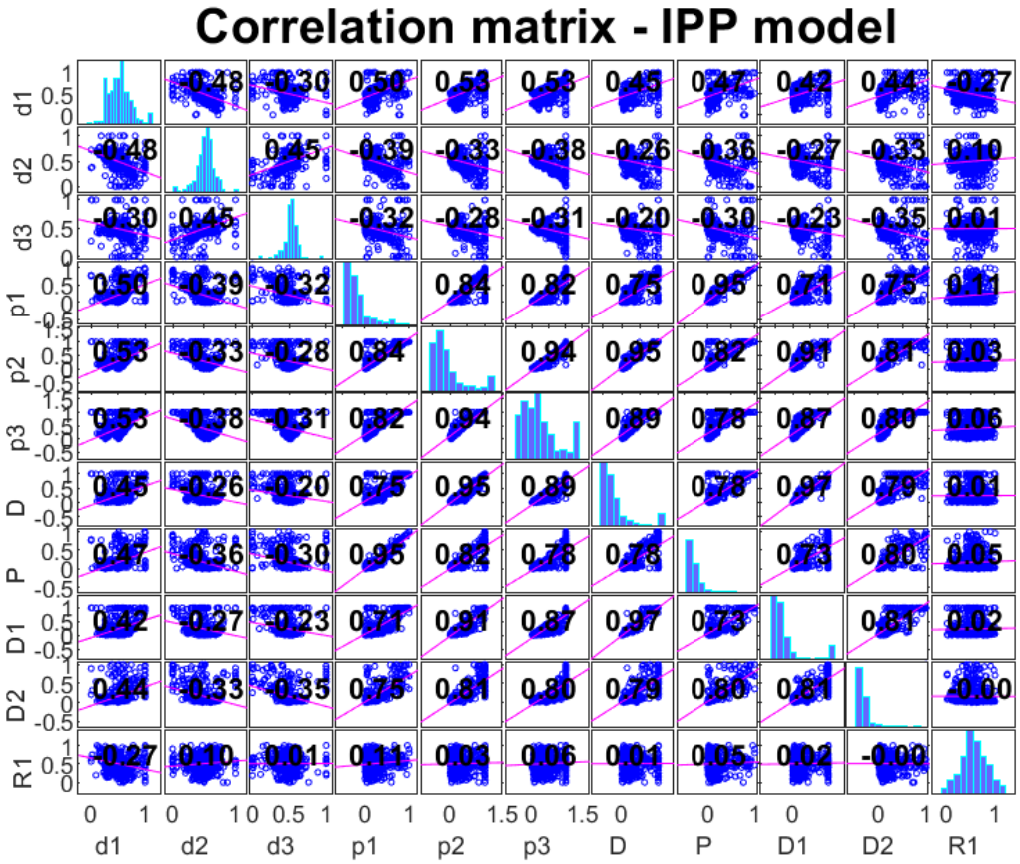


Figure 60: Correlation matrix of the IPP model $\phi_1 \equiv R_1$).

As can be seen in both sub-models there is a significant correlation between the predictors. In other words, in the training process the inclusion of highly correlated predictors may not contribute to better classification. One may plot the combinations of polarimetric quantities like it is shown in Fig.53. However, for *n* predictors, there are *n(n-1)/2* combinations between them. For large number of *n*, such approach would be computationally and analytically ineffective (i.e. 55 plots to be analysed for each of the sub-models). In this case, the principal component analysis (PCA) is an adequate solution. For each of the two sub-models PCA was computed to summarize the available data as shown in Fig. 61. It was shown, that 7 principal components (PCs) retain more than 95 % of the total variance for the eigenvalue model and 6 PCs - for the IPP model. By this way PCA can be combined with the classification ML in order to use the non-redundant features only from both sub-models and any other collinear or highly correlated features can be avoided. To project the experimental data onto the principal component space, one can compute the principal component scores (PCS). As a result, there is no correlation between all PCS in both sub-models, as shown in Fig. 62(a) and (b), whereas 95 % of the total variance is sustained.

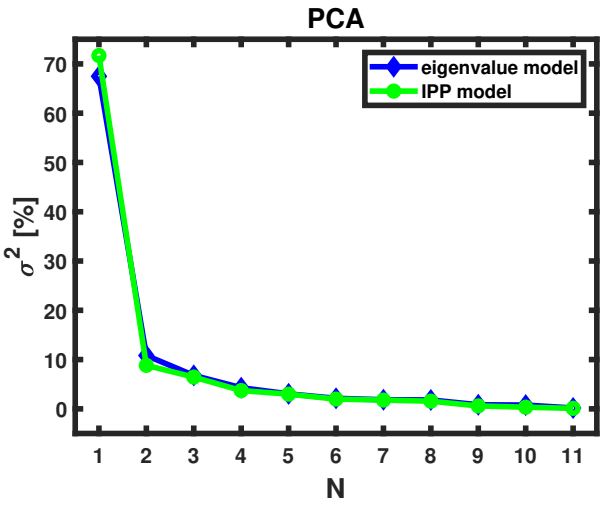
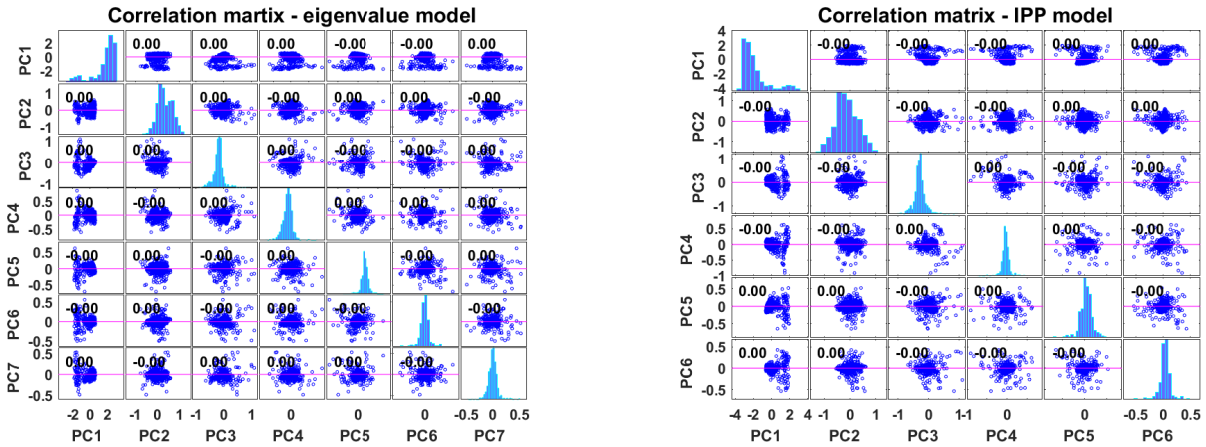


Figure 61: PCA for N number of components, explaining the corresponding percentage of variance σ^2 for both submodels. Results available from the published material in [99].



(a) (b)

Figure 62: **(a)** Correlation matrices for: **(a)** 7 PCs and their scores – eigenvalue model and **(b)** 6 PCs and their scores – IPP model. Results available from the published material in [99].

4.4.4 Machine learning

Firstly, the data sets for both models without the PCs were randomly split to obtain two data subsets for training and testing as follows: 570 samples (85% of the total data) for training and 100 samples (15 % of the total data) for testing. To evaluate the best predictors for tumor detection (see Fig. 63), logistic regression (LR) was trained independently with both models, but without using their PCS. By this way it was found out, that the inclusion of λ_1 is deteriorating for the model performance and this parameter was consequently removed from the analysis. In Fig. 63 the top and bottom axes include 1D distribution of the predictors' normalized data, for both health conditions (0–Healthy, 1–Tumor) respectively. It could be well observed, that d_1, ϕ_1 and λ_2 show excellent detection performance for malignant formations, where the uncertainty intervals (in grey) remain close to the probability values (all blue lines). Although the probabilities for P, D_2 and P_1 parameters are lower and have higher uncertainties compared to the former triplet of polarimetric parameters, each one of the latter triplet could also be identified with sufficient probability values.

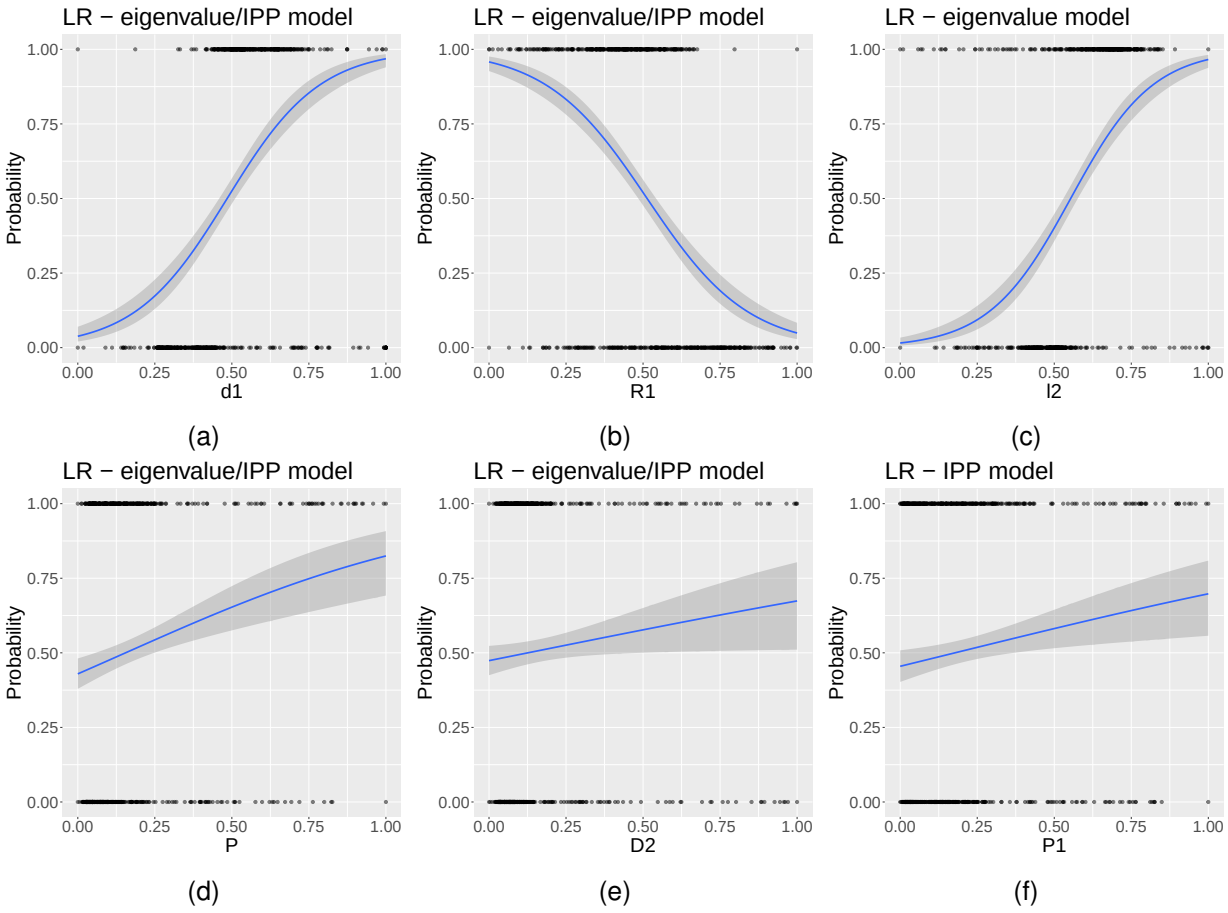


Figure 63: Probability for tumor detection, calculated from LR algorithm: **(a)** d_1, **(b)** R_1, **(c)** λ_2, **(d)** P, **(e)** D_2 and **(f)** D_2. For the panels **a,b,d** and **e** the results are comparable for both sub-models, whereas the panels **c** and **f** show the results computed from the eigenvalue and the IPP sub-models, respectively ($\phi_1 \equiv R_1$ and $\lambda_2 \equiv l_2$). Results available from the published material in [99].

Typically, malignant tumor formations cause morphological alterations in tissues and alter the collagen extracellular matrix, as well as the cellular organelles by modifying their sizes and shapes. This leads to changes in tissue heterogeneity, followed by reduced number of scattering events as ϕ_1 may indicate. Also, Rayleigh-Mie transition of light scattering regime occur, that in turn affects light (de)polarization. Whereas the depolarization parameter d_1 can be considered as a weight coefficient for the Stokes component S_1, higher polarimetric purity would indicate less depolarizing media. By this way, this polarimetric doublet may be considered as noteworthy tumor markers for the angular-resolved measurements with any angles of incidence and detection different from normal.

Next, solely for the classification purpose the logistic regression (LR), random forest (RF) and support vector machines (SVM) algorithms were again trained with the corresponding PCS data subsets for both sub-models, split again randomly with the same proportions. All ML models underwent initial tuning to pick up the best possible hyperparameters. In the case of the RF, a randomly selected fraction of $k=N^{1/2}$ from all predictors was drawn without replications to create an ensemble of decision trees.

For both sub-models, having three predictors per split was found to be the most optimal choice. By setting the number of trees to 30, we reached the same classification accuracy as with 500 trees, while the training time was reduced by an order of magnitude when using smaller number of trees. Without replications, there are 35 possible predictor combinations (3 randomly selected PCs and their scores from total 7) for the eigenvalue sub-model and 20 possible predictor combinations (3 randomly selected PCs and their scores from 6) for the IPP sub-model, calculated from $C_N^k = N!\cdot(k!(N-k)!)^{-1}$. For all decision trees in the ensemble (including replications), the possible number of predictor combinations for training is 4960 for both sub-models, calculated from $K_N^k = (N+k-1)!\cdot(k!(N-1)!)^{-1}$. By this way, RF algorithm could be considered as more reliable MLA for tumor classification, even if the dataset size is small and/or there is a presence of correlated predictors. The out-of-bag (OOB) error was found $\approx$ 5% for the eigenvalue sub-model and $\approx$ 11% for the IPP sub-model. In the case of SVM algorithm, after cross-validation a polynomial kernel of 3$^{\text{rd}}$ degree was found to provide the best classification accuracy with both sub-models. Additional regularisation *C* [36, 37] was necessary to add a penalty for each misclassified data point. Usually, small values of *C* result in smaller margin, which will affect the variance-bias trade-off, resulting in an overfitting model [36, 37]. Conversely, for large *C* values the margins will be increased and the model will allow more data points to be classified from the wrong side of the separating plane, again affecting the variance-bias trade-off [36, 37]. After the cross-validation cycle, the optimal values of *C* were found to be 1 for both sub-models.

After the application of the aforementioned MLA for the data set classification, various other metrics were used to evaluate the classifiers' performances, e. g. their accuracy, sensitivity, specificity, relative risk of misclassification (R_r), receiver operating characteristic (ROC) curve and the corresponding area under the curve (AUC). While the sensitivity represents the portion of the correctly predicted true positive (TP) values (in this study - the correct detection of tumor class), the specificity is related to the amount of the correctly predicted true negative (TN) values (analogously - the correct detection of healthy class). For an ideal classifier the accuracy (a sum

of all true predicted classes normalized to the sum of all true and all false predicted classes), sensitivity and specificity should be 100 %. However, due to the presence of wrongly predicted class values such as false positive (FP – healthy tissue, but detected as tumor) and false negative (FN – tumor tissue, but detected as healthy), the models' detection performance deteriorate. In this regard, the relative risk of misclassifications can be calculated as:

$$R_r = \frac{FP}{FP + TN} \cdot \left[\frac{TP}{TP + FN} \right]^{-1} \tag{88}$$

Ideally, lesser misclassified values will lead to closer proximity of the ROC curve to a step-wise profile. As there is no perfect classifier model, the losses introduced from wrongly predicted class values will always be a considerable factor, which can be simply calculated as 1-AUC. The results from all classification MLAs are presented in Fig. 64 and in Table 3.

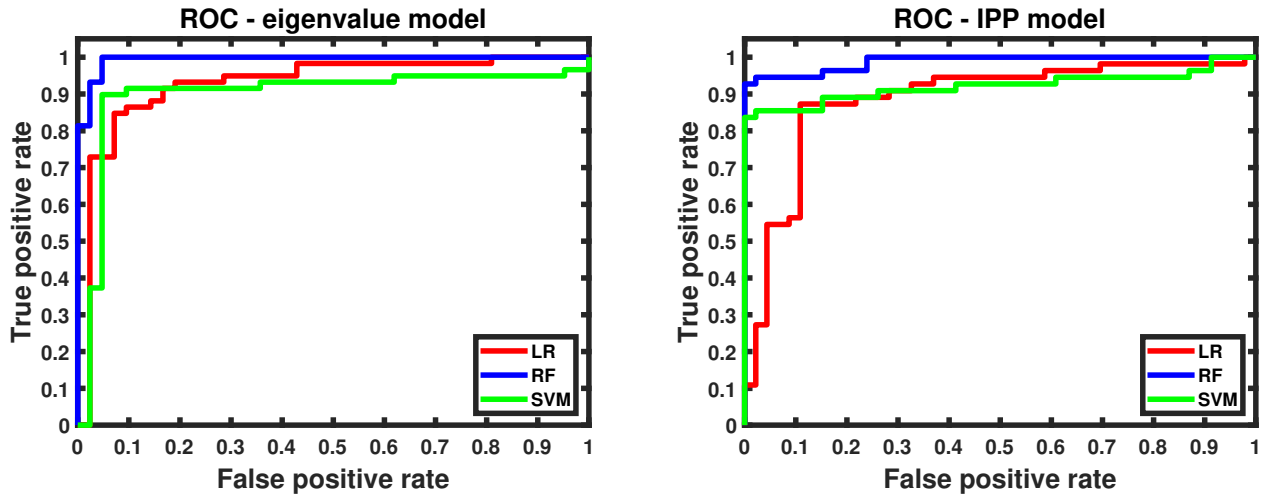


Figure 64: ROC curves for: **(a)** eigenvalue sub-model (trained with 7 PCs and their scores) and **(b)** IPP sub-model (trained with 6 PCs and their scores). Results available from the published material in [99].

Table 3: Supplementary table associated with all classification MLAs performances, where all numerical values are in %. All MLAs were trained with 7 PCs and their scores for the eigenvalue sub-model and 6 PCs and their scores for the IPP sub-model. Data available from the published material in [99].

	Accuracy	Sensitivity	Specificity	AUC	Loss	R_i
LR – (λ_i)	87	85	91	93	7	11
LR – (P_i)	84	80	89	87	13	14
RF – (λ_i)	97	100	93	99	1	7
RF – (P_i)	95	93	98	98	2	3
SVM – (λ_i)	88	92	83	90	10	18
SVM – (P_i)	77	93	59	92	8	45

From the graphical representation in Fig.64 and the values in Table 3, it becomes possible to outline both sub-models performance for tumor tissue classification. To sum up, all ML trained with the corresponding PCS provide reliable accuracy and AUC values close to 1. The eigenvalue sub-model seems to perform better than the IPP sub-model with lower OOB error and higher diagnostic quantities. Whereas LR algorithm is better suited to evaluate the predictor's probability for tumor detection and has higher specificity values than SVM, the latter MLA has higher sensitivity values than LR and is better suited to predict the healthy class. On the other hand, RF algorithm yielded the best results for classification with negligible losses and misclassification risk. However, a parallel should be drawn between RF and SVM. The former can be computed with only two hyperparameters – the number of variables/predictors per each random split and the number of trees. On the other hand, the latter is dependent and highly sensitive to the kernel choice and degree, the regularization parameter(s), choices for support vectors and margins all of which influence the variance-bias trade-off. Additionally, the posterior probabilities for both classes were found to differ at most for RF, whereas for SVM the difference between these values was very small, thus reducing the reliability of SVM for classification for the current study. In the end, another flowchart only for the supervised ML is shown in the Fig.65:

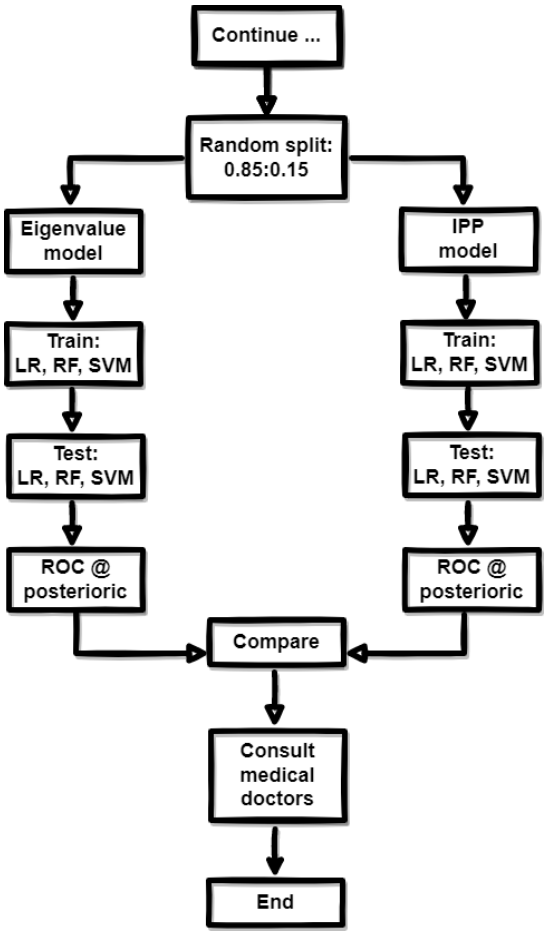


Figure 65: Flowchart of the supervised ML process.

General conclusions

In this Chapter, all experimental results were presented, alongside the corresponding data post-processing parts. Although most of the results originated from the pilot studies, the main milestone was reached. Firstly, our experimental results validate the theory behind all Mueller matrix decomposition algorithms. Secondly, and most importantly, these results are obtained from highly anisotropic and heterogeneous samples that makes is challenging to validate the above-mentioned algorithms. Despite of that, it was empirically shown that the accuracy of our approach for medical diagnosis of tissue suggests the applicability of polarimetry to assist medical doctors with complementary, optical set-up, using light polarization as an optical marker. Thirdly, whenever needed, the polarimetric results were supported by statistical analysis and/or artificial intelligence algorithms. Both data post-processing approaches could be very useful to provide more insights about the available data. For instance, information whether the data come from the normal distribution or not would define the choice of any parametric either parametric or non-parametric tests, as well as machine learning algorithms, accordingly. Also, when multiple distributions are compared and tested whether they are drawn from the same population or not, the error rate from all comparisons will scale up too. Therefore, with the use of *post-hoc* analysis, this drawback could be overcome as well. Of course, this would be on a price of adjusted p-values from the tests for each comparison. Whereas the adjusted p-values are reduced, this would be equivalent of reduced sensitivity of the tests, as smaller changes into the data set cannot be detected. Nevertheless, an initial study of the available data set by means of the statistical analysis is essential prerequisite before further using this data set for machine learning approach.

To begin with, in Section4.1 two decomposition algorithms were used and their results were compared accordingly by using *ex vivo corpus callosum* human, brain thin sections. Both algorithms were able to detect the brain fiber orientations, which has the potential to guide neurosurgeons for more precise localisation of the tumor zone during neurosurgery.

Next, in Section4.2 all presented results showed valuable insights about the impact of different thawing methods on the preservation of the internal micro-sctructure of 3D tissue models compared to the reference unfrozen control tissue model samples. Such an approach allows to better understand the fundamentals of tissue thawing mechanism and has the potential to be extended for reflection geometry with thick samples and become instrumental in the field of regenerative medicine. Also, the statistical analysis with different tests was found complementary and supported the polarimetric results as well.

Then, in Section4.3 the thin sections of human skin at different pathological conditions were selected and measured. The polarimetric images of the selected samples after the decomposition revealed a sensitivity of Mueller polarimetry to such kind of structural changes related either to degenerative or malignant lesions. With the help of image processing and segmentation, the obtained images were used to form a polarimetric data base and to develop a deep learning model for image classification. This approach has the potential to save enormous time to both

physicians and physicists, thus supplementing the final diagnostic conclusions. Nevertheless, this comes on the price of higher computational time and the need to use GPUs to accelerate the computations during the training processes.

The last experimental part included the results obtained with a new polarimetric set-up operating in reflection. This time instead of conventional imaging, the system was used in a scanning mode, where the detector was a single photo-diode. It was found possible by interpolating between the scanning points to obtain non-conventional images of the corresponding ROIs. The symmetric decomposition of the measured Mueller matrices was used and combined with the calculations of polarimetric purity indices and polarization entropy. These frameworks parametric spaces have additionally enriched the polarimetric analysis and were useful for a comparative analysis between the healthy and the tumorous ROIs of the *ex vivo* colon samples. The obtained data were found sufficient to be used for detailed statistical analysis and also to be useful for construction of two polarimetric models, which behaviour for machine learning classification was analysed and compared. By this way, with the help of unsupervised and supervised machine learning, a possible solution to the problem of inter-patient variability was proposed. Namely, with the help of the principal component analysis the redundant information from both the eigenvalue and the IPP sub-models was omitted, while preserving 95 % of the total variance. This was also found suitable to overcome the correlation between all predictors in the sub-models and to avoid overfitting and improve the sub-models' performance. Then, based on the results from the statistical tests, only the algorithms for classification requiring no explicit knowledge of the data distributions were used. While logistic regression provided weaker results after classification in comparison to the random forest and support vector machines algorithms, it was still found feasible to evaluate the predictors probability to detect tumor with the logistic regression algorithm. By this way, medical doctors could be provided with relevant optical markers and this algorithm may have a good future applicability with larger data sets, combined of different polarimetric quantities as well. On the other hand, the reason for the SVM algorithm to underperform in comparison to the RF algorithm is based on the fact, that the number of measurements is higher than the number of features in the data set. On contrary, the RF algorithm is less influenced by the presence of outliers than the rest of the algorithms. The classification results are averaged over all decision trees in the ensemble and by this way the variance-bias trade-off is better balanced. For this particular data set and both sub-models used in our studies, as logistic regression showed, we have many strong predictors. This fact additionally amplifies the significance of RF algorithm, which performance could be worsened if there are less stronger predictors in the data set.

To conclude, with the increasing amount of data, machine and deep learning can come into force to complement the polarimetric results. Nevertheless, there are some key differences between machine learning and deep learning conventions. For instance, the first major difference is related to the human intervention. With machine learning (ML) there are more steps for data pre-processing and feature selection/extraction, in comparison to deep learning (DL) approach. In the framework of the current dissertation, one can recall the following pre-processing steps: i) use of the Mueller matrix decomposition algorithms, ii) data filtration and normalization procedures, iii) the use of principal component analysis with the results from scanning in reflection of the colon

samples before feeding the data to the aforementioned ML algorithms for classification. On the contrary, with DL only the images of the Mueller matrix elements were used, which were initially segmented into the patches for data size increase in order to train the neural network. Potentially, with higher measurements count, the images of full size can be used directly to train the neural network, without the need of image segmentation in patches. Afterwards, the classification of unseen data to the model can be done directly after the physical realizability filtering, without even the need to apply any other decomposition algorithm. Finally, despite all difference between the ML and DL approaches, their application should be done without any misuse and overuse as stated in [100], while the hyperparameter tuning and the correct model choice should be tuned and performed responsibly, thus, assuring the results with low bias and low variance. The very same foundations were followed upon creating the three models presented in this Chapter, namely: eigenvalue, IPP and Skin-HDT700. The former two are applicable for colon samples' binary classification, while the latter one is used for skin images multi-class classification. Due to its reduced accuracy, the IPP model can be disregarded in favor of the eigenvalue sub-model. By this way we can rely on the eigenvalue model for classification of angular-resolved point measurements of *ex vivo* colon samples with single photodiode and to rely on the Skin-HDT700 model for the image classification of *ex vivo* skin samples.

5 Future perspectives

The emerging applications of polarimetry as a diagnostic optical modality for different tissue specimens and tissue types are presented in the review part of the thesis and referenced in the bibliography section. By characterizing the polarization fingerprint of biological tissue specimen polarimetry has the potential to become non-invasive and label-free diagnostic modality for digital histology. The latter can significantly help medical doctors in their diagnostic conclusions. Additionally, significant amount of time could be saved by this way, as sample fixation, sectioning, staining and annotation is extremely time consuming, requiring a lot of work. By measuring the complete Mueller matrix of the given samples, we can generate the images of the corresponding polarimetric parameters derived after the application of the relevant decomposition algorithm. These parameters can be consequently related to the health conditions and may provide better understanding of tissue interactions with polarized light.

The polarimetric studies of the brain tissue histological sections showed encouraging outcome for the visualization of the brain fiber tracts in the images of the azimuth of the optic axis, regardless of the decomposition in use. It is known that brain tumor cells grow chaotically and, contrary to healthy brain white matter, brain tumors do not demonstrate optical anisotropy. Thus, the imaging Mueller polarimetry has the potential to be implemented *in vivo* during neurosurgery in order to visualize the brain fiber tracts and guide the surgeons for more precise excision of brain tumors.

Then, another very important study was dedicated to the assessment of small, but significant internal damages within the 3D tissue models after freezing and consequent defrosting by using different methods. The review of literature and the results from other groups working in this field show that there are internal damages caused by the standard convection rewarming of the cryopreserved biological tissues. It turns out to be very important for the field of regenerative medicine and transplants to be stored at cryogenic temperatures. As was shown, small structural changes in the thawed 3D tissue models can be detected with polarimetry combined with the image segmentation techniques and statistical analysis. In practice, the two latter approaches can be implemented in the clinical practice by providing an user-friendly software to be operational with ease with different types of experimental imaging data sets.

Next, it was shown that polarimetry is also sensitive to multiple types of skin tissue diseases, either degenerative or tumorous. The diagnosis of skin lesions with polarimetric set-up can be done *in vivo* due to the direct access to the lesion site. Again, we demonstrated that when polarimetry is combined with the statistical methods and artificial intelligence (AI) algorithms, it becomes an immensely powerful approach to detect at early stage the development of any particular type of lesion. This was confirmed from the deep learning model Skin-HDT700 and its applicability to predict the lesion type only from the images of Mueller matrix elements. By this way even lesser data processing would be required, by only providing the corresponding images to the classifier. Of course, in order to have a generic model for such type of classification, vast amount of images and data set are required for training, with additional fine tuning of the

model, before deploying it for clinical applications. Nevertheless, the potential of such kind of combination of the polarimetric imaging and AI approach for the detection of the disease-specific signatures on a pixel level is immense and holds great promise for accurate diagnosis.

Finally, the polarimetric scanning experiments with the colon samples in reflection geometry also provided good results in terms of diagnosis of tissue pathological status. To speed up the measurement process, the measurements were done with a different type of polarimeter on the price of no imaging but single-shot measurements. This was found useful to select larger ROIs, explore the impact of varying angles of incidence and detection and to map the corresponding evolution of the polarimetric parameters in the scanned zones. By comparing the results obtained for the healthy and the malignant zones of the colon specimens the significant differences were found in their polarization response, respectively. Inevitably, the results may be affected by the inter-patient variability. Again, with the help of statistical analysis and machine learning (ML), this obstacle was overcome. It was also shown that, when training a ML model with the polarimetric data obtained from the Mueller matrix decompositions, this may results in correlation between the predictors. In order to solve this problem a combination with the unsupervised ML algorithms was found useful for the dimensionality reduction.

To conclude, all results presented in the thesis are from *ex vivo* samples, due to the pilot nature of all projects. We have shown that it is very important to take into account in that case the impact of host medium, if present. For *in vivo* applications far more considerations must be taken into account, i.e. presence of blood, patient movements, etc. Moving towards clinical applications will give rise to many other requirements and considerations, but the initial profound academic studies of tissue polarimetry, combined with the artificial intelligence and statistical analyses are a must for reaching the final goal.

6 Appendix

As a general rule, during the stages of instrument calibration, sample measurements, data processing and post-processing, the final goal is to have the experimental data with low variance and low bias. Ideally, this would contribute to a more precise solution to a real-life problem, e. g. the differentiation between two or more tissue pathologies or the characterization of different thawing mechanisms by means of polarimetric measurements. Having both low variance and bias would also enable more accurate statistical tests or predictions with the available polarimetric model(s). Both requirements were achieved by the meticulous calibration of the experimental set-up, the implementation of the physical realizability filtering and normalization or standardization whenever necessary, as was described in all sections above. Following all these steps, the purpose of this Section is to introduce the basics of the mathematical background underlying all algorithms used for the unsupervised and supervised machine learning.

6.1 Principal component analysis

Let us now denote the data frame with **X** where each column is a polarimetric quantity usable as a predictor, and each row is an independent measurement. Then each column of **X** can be standardized to have zero mean μ and unit variance σ like:

$$\mathbf{X_s}(i) = \frac{\mathbf{X}(i) - \mu(i)}{\sigma(i)}. \tag{89}$$

The covariance matrix **C** of the standardized data can be obtained for the total of N measurements as:

$$\mathbf{C} = \frac{\mathbf{X_s}^{\mathbf{T}}\mathbf{X_s}}{N}. \tag{90}$$

Afterwards, the eigenvector-eigenvalue problem is solved for the matrix **C**, and the eigenvalues with their corresponding vectors are arranged in a descending order. Each eigenvector is normalized to its magnitude, so that the length of all eigenvectors is the same. The diagonal matrix Λ contains all eigenvalues that have a physical meaning of variances, while the eigenvectors represent the directions or the axes in which the original data vary. To quantify the amount of variance contained in each principal component, the quantity **E**(i) is introduced:

$$\mathbf{E}(i) = \frac{\lambda(i)}{tr(\Lambda)}. \tag{91}$$

At this point, one can perform dimensionality reduction by simply selecting those principal components containing, e.g. 0.95 of the total variance, which would be sufficient to describe satisfactorily all patterns and features in the original data. In order to map the original data to the space

of the principal components, the following transformation is applied, by obtaining the principal component scores $\mathbf{Z_{ip}}$ [36–38]:

$$\mathbf{Z_{ip}} = \sum_{k=1}^{p} \mathbf{V_{kp}}^{\mathbf{T}} \mathbf{X_{ik}}, \tag{92}$$

where the matrix $\mathbf{V_{kp}}$ contains all eigenvectors of the covariance matrix $\mathbf{C}$, while $\mathbf{X_{ik}}$ is denoting the standardized data frame of the original data. Explicitly, only for the first principal component, Eq.92 can be expanded as [36–38]:

$$\mathbf{Z_{i1}} = v_{11}x_{i1} + v_{21}x_{i2} + ... + v_{k1}x_{ik}. \tag{93}$$

Moreover, as all eigenvectors are orthogonal to each other, then no correlation is expected between the principal components and their scores, respectively.

To sum up, this algorithm can be used as a way to reduce dimensionality and to overcome correlation between predictors in the data set. By removing the correlated features, the overfitting may be prevented. Additionally, with the dimensionality reduction, the computation time will be decreased too if later on the principal component scores are to be used in combination with other machine learning algorithms. Unfortunately, PCA has its own disadvantages as, firstly, the data must be standardised since otherwise the algorithm will not compute correctly the principal components and their scores. If the choice of the number of principal components is not optimised, this could lead to information loss and the model would not be able to capture all features of the original data. Also, the data interpretation may be more difficult because the principal components and their scores are the linear combinations of the parameters from the original data set. Last but not least, PCA is very sensitive to outliers and special care to these data points must be taken.

6.2 Logistic regression

Let us now consider a data set with target variable whose response falls into a binary classification problem. Instead of trying to create a regression model for the target variable with respect to the other predictors in the data set, we can model the probability of the target to belong to either of the two classes. As the probability will vary within the interval [0, 1], then the logistic function known as sigmoid function $\sigma(x)$, shown in Fig.66, which is also restricted in the interval [0, 1], can be used:

$$\sigma(x) = \frac{1}{1 + e^{-x}}. \tag{94}$$

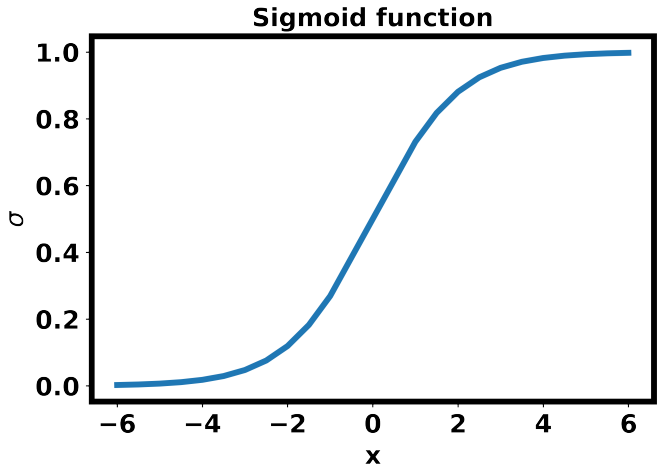


Figure 66: Sigmoid function $\sigma(x)$ of real variable x.

To model the probability $p(x)$ using the logistic function, we can use the following expression [36–38]:

$$p(x) = \frac{e^{\beta_1 x + \beta_0}}{1 + e^{\beta_1 x + \beta_0}}, \tag{95}$$

where β_0 and β_1 are the free term and the slope, respectively. To simplify the expression, one can put $\beta_1 x$ + β_0 = $\alpha(x)$. Then, by multiplying both sides of Eq.95 by 1 + $e^{\alpha(x)}$ and rearranging, we obtain:

$$\frac{p(x)}{1 - p(x)} = e^{\alpha(x)}. \tag{96}$$

The left-hand side of Eq.96 represents the odds, now related exponentially to the regression coefficients in $\alpha(x)$. After taking the logarithm of both sides of Eq.96, we arrive at the point of logarithmic odds, which are now linearly related to the regression coefficients:

$$ln\left[\frac{p(x)}{1 - p(x)}\right] = \alpha(x). \tag{97}$$

Nevertheless, the regression coefficients are unknown and have to be estimated to model all features of the training data and then, consequently, to perform classifications on the testing data. For instance, if we suppose there are two classes (i.e. Healthy: 0 and Tumor: 1), the idea is to estimate the values for the regression coefficients in $\alpha(x)$ so that, when these values are substituted into Eq.95 it would yield a number close to unity for the Tumor class and a number close to zero for the Healthy class. In this case, the modelling can be done by the method of the maximum likelihood [36–38]:

$$L(\beta_0, \beta_1) = \prod_{i:y_i=1} p(x_i) \prod_{i:y_i=0} [1 - p(x_i)]. \tag{98}$$

From here on, what is left is finding the derivative of the logistic function in Eq.98 and determining its maxima and minima for the corresponding regression coefficients, which would be sufficient to model the training data with respect to the target variable.

To sum up, logistic regression is an efficient algorithm to be used when one wants to make predictions on a binary categorical variable, i.e. 0 or 1, "Yes" or "No", "True" or "False", etc. This algorithm is relatively easy to interpret and it makes no assumptions about the distributions of both classes. It performs well with good classification accuracy if the data are linearly separable and, when the data set has fewer dimensions, logistic regression is less prone to overfitting. Also, this algorithm can compute the probability of either of the two classes when making predictions. Conversely, if the number of observations is less than the number of features, this algorithm should be avoided, due to proneness to overfitting. If there is no perfect linear separability in the data set, this will negatively affect the algorithm when making predictions. As a final drawback, it should be emphasized that logistic regression is very sensitive to the predictor correlation and, this may affect the model performance for classification too.

6.3 Random forest

To construct more powerful predictive models, random forest classification algorithm can be utilized. Unfortunately, the use of multiple decision trees suffers from high variance and, if there is a correlation between the predictors, the model can also suffer from multicollinearity. To overcome these problems, random forest draws at random at each node a fraction of the total number of predictors P, usually $\sqrt{P}$. The number of trees T is a hyperparameter as well and can be chosen empirically. Then the model is being trained on each fraction i drawn at random, for all decision trees and the final result of classification $\bar{f}_c$ is then averaged (a.k.a "bagging") [36–38]:

$$\bar{f}_c(i) = \frac{1}{T}\sum_{t=1}^{T} f_t(i). \tag{99}$$

In this way, this algorithm decorrelates the randomly drawn predictors, otherwise the use of all predictors at a node would lead to correlated decision upon classification. Unfortunately, each individual tree has high variance, but low bias however, after averaging, the variance decreases eventually. Additionally, random forest splits the training data into two fractions with proportions 2/3 and 1/3. The former is used to train the model, while the latter is kept for validation to assess the model performance (a.k.a out-of-bag; OOB or out-of-bag error; OOB error).

In summary, random forest is a decision-tree algorithm which does not require additional data scaling and is less influenced by outliers than other ML algorithms. It also does not make any assumptions about the underlying distribution of the data. Moreover, this algorithm can overcome high correlation between the predictors used. Another advantage of algorithm is its ability to provide good balance between the variance-bias trade-off, due to the model principle to average the results across the multiple decision trees. However, this algorithm is recommendable to be used essentially when there are many strong predictors in the data set to perform classifications, otherwise the classification accuracy may decrease. Also, with the increase of the number of

trees and the number of predictors, the computational time would scale accordingly. This factor should be also taken into account when using random forest with very large data sets.

6.4 Support vector machines

Suppose we have a data set with p predictors, originating from two distinct classes. Then we can form a p - dimensional space, in order to separate both classes, by finding a hyperplane with dimension p - 1. However, the question for the choice of the optimal separating hyperplane is generally open; typically, the farthest from both training observations hyperplane is taken. For the sake of simplicity, let us now consider a 2-dimensional case, where the separating hyperplane will be a line. In Fig.67 two scenarios are shown - linearly separable and non-linearly separable data sets.

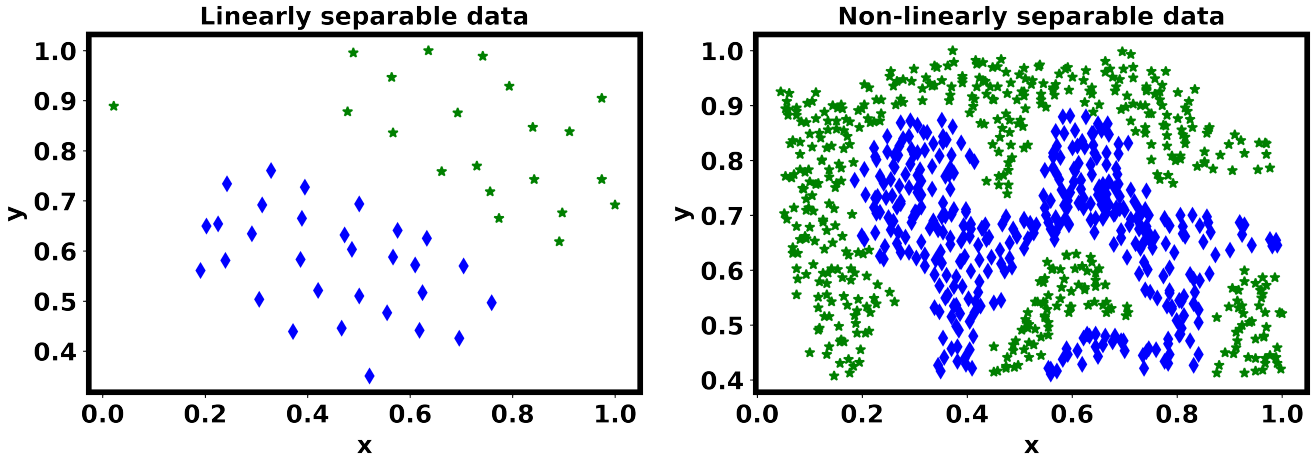


Figure 67: Arbitrary data and provided code solutions available from the course *Machine and Deep Learning, École Polytechnique, INF 554*: **(a)** Linearly separable and **(b)** Non-linearly separable.

By denoting our data frame with D and the two classes by K, and assuming that the data set is linearly separable the hyperplane can be defined as [36–38]:

$$f(x) = D^T\beta + \beta_0 = 0, \tag{100}$$

where $||\beta||$ = 1, so that the hyperplane can be used for classification. If the Eq.100 holds then the data point will be exactly on the hyperplane, while in the case of inequality it will be classified as one of the two available classes i. In practice, often some data points of the two classes may be at close proximity and it will be difficult for the classifier to separate them. For this particular reason, a margin M must be defined, the idea being to maximise the margin M from the hyperplane to the nearest data points from the two classes [36–38]:

$$K_i(D^T\beta + \beta_0) > M. \tag{101}$$

However, a problem in this case may occur. For instance, when we select a classifier relying on a hyperplane which can ideally separate all of the training data points, we may end up with a model which is extremely sensitive to individual observations. Then the margins will be tiny and this would cause a sensitivity to single observations, while we want to keep the margin as large as possible in order to be more certain not to wrongly classify an observation or data point. Explicitly, this would force the model to overfit by studying very well the training data, but failing to perform well on the unseen data. To overcome this problem, we can artificially introduce imperfections in the model by using a hyperplane, which does not separate perfectly the two classes. This technique is implemented by using slack variables ξ_i and a cost variable C. While ξ_i soften the separating constrains, the cost variable makes sure that the sum of all ξ_i is always less than C [36–38]:

$$K_i(D^T\beta + \beta_0) > M(1 - \xi_i). \tag{102}$$

In this way, C becomes a hyperparameter and is tuned via cross-validation. If ξ_i = 0, then the data points are located on the correct side of the margin, but when $\xi_i > 1$ the data points are located on the wrong one. Hence, large values of C suppress large values of the slack variables ξ_i, while small values of C give rise to the slack variables ξ_i. Changes in the position of a single observation would not affect the model and the classifier only if the position is on the correct side of the margin. Those observations which lie directly on the margin, or are located on the wrong side of the margin are defined as support vectors and they do affect the model. In this way, C controls the variance-bias trade-off. When the model is having a higher number of support vectors, lower variance, but high bias are present and conversely, when the model has less support vectors, higher variance is observed, but at lower bias. Therefore, the correct choice of C and the support vectors is critical for the model accuracy. An example can be seen in Fig.68.

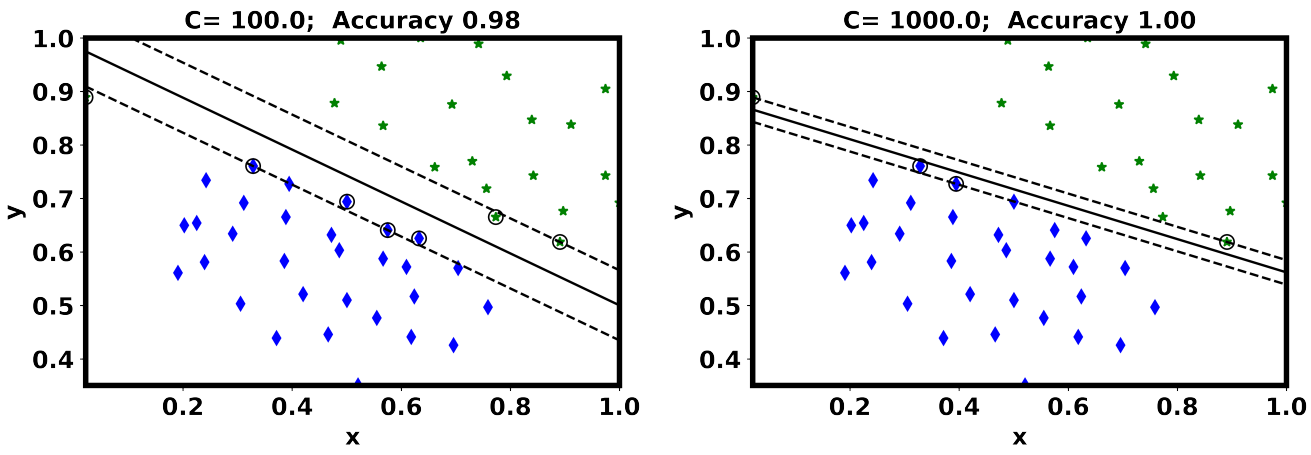


Figure 68: Arbitrary data and provided code solutions available from the course *Machine and Deep Learning, École Polytechnique, INF 554*: **(a)** SVM model with linear kernel, C = 100 and **(b)** SVM model with linear kernel, C = 1000.

In some cases shown in Fig.67b we clearly have a non-linear decision boundary between the

two classes. The solution to the optimisation problem for the most suitable separation boundary involves the inner products of the observations [36–38]:

$$< x_i, x_j^{'} >= \sum_{j=1}^{p} x_{ij} x_{i^{'}j}. \tag{103}$$

For every data point in the training data set there are $n(n-1)/2$ inner product calculations. This is also known as the kernel approach, quantifying the similarities between the training data points. In the linear case, the kernel function has the form of [36–38]:

$$f(x) = \beta_0 + \sum_{i=1}^{n} \alpha_i < x, x_i >, \tag{104}$$

where with α_i are denoted the parameters, which are zero if the inner product is calculated from a non support vector data points pair, and are non-zero in the case of a support vector data points pair. Analogously, one can replace the linear kernel with a non-linear one so that the data points may be better separated. For instance, one can use a Gaussian kernel:

$$K_G(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right), \tag{105}$$

where σ determines how fast the similarity metric will decrease towards 0 if the input data points are further apart, and is defined *a priori*. As σ is a hyperparameter in this case, its optimal value can be determined via cross-validation. Then, the function for the Gaussian kernel equation can be rewritten as [36–38]:

$$f(x) = \beta_0 + \sum_{i=1}^{n} \alpha_i K_G(x_i, x_j). \tag{106}$$

In this case, if an arbitrary data point x_j is far away from the training data point x_i, then the numerator in Eq.105 will be large and the exponential decay will be large too. This will result in little or no contribution to $f(x)$. Conversely, small values in the numerator in Eq.105 would result in higher contribution in the predicted class. In Fig.69 an example is shown with the use of non-linear kernel implemented on a linearly separable data and on a non-linearly separable data. In both cases, an excellent separation between the two classes can be seen. In reality, there is no perfect solution for the optimal kernel choice, the value of the hyperparameters, etc. Therefore, the data must be analysed in detail before making a decision on the model parameters, for support vector machines to yield good classification results, especially with non-linearly separable data set.

To sum up, support vector machines is an efficient algorithm on data sets with multiple features whose number is higher than the number of measurements or data points. Also, it does not require explicit knowledge of the underlying distribution of the data. Very importantly, as SVM uses a subset of training points or only the support vectors in forming the decision boundary function, this makes the algorithm memory efficient. Furthermore, different kernel functions can be chosen if the data are not linearly separable. However, if the number of features is lower

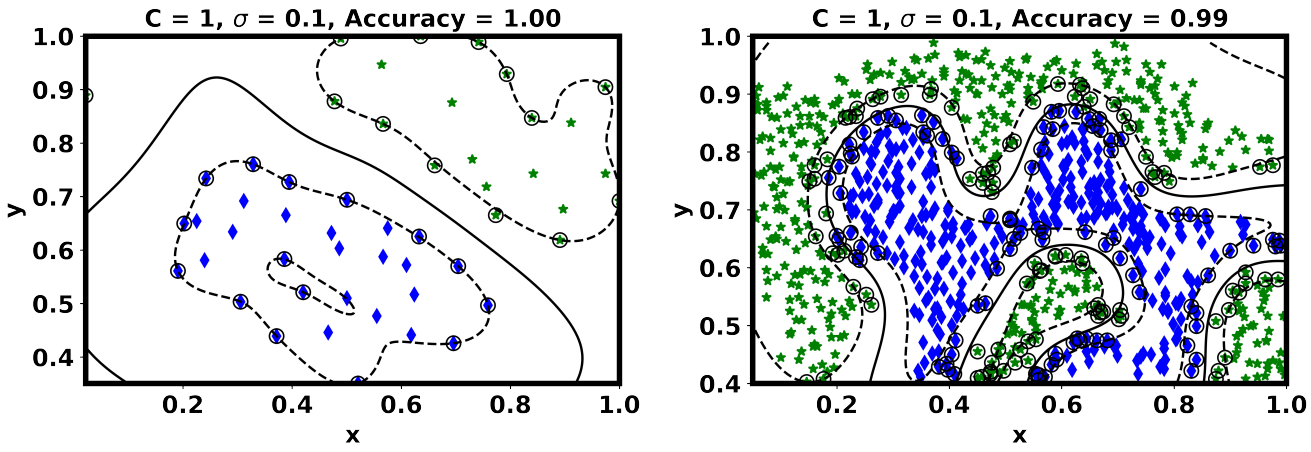


Figure 69: Arbitrary data and provided code solutions available from the course *Machine and Deep Learning, École Polytechnique, INF 554*: **(a)** Linearly separable data and **(b)** Non-linearly separable data. In both cases Gaussian kernel was applied with C = 1 and σ = 0.1.

than the number of data points, the algorithm is not efficient in making predictions, and is also computationally inefficient. In that case, SVM algorithm is prone to overfitting and it becomes crucial to prevent the latter by carefully selecting the kernel functions and the corresponding hyperparameters as well. Finally, SVM does not perform very well when the data set has more noise so that the target classes may be overlapping.

6.5 Deep learning

This class of the algorithms is a subset of machine learning algorithms that is prone to produce impressive results in image classification, computer vision, etc. In deep learning approach, the artificial neural networks are used to mimic the workflow of our brain. Our nervous system is comprised of neurons which communicate with each other along pathways, known as synapses. Analogously, in artificial neural networks, the artificial neurons simulate the biological neurons, thus, simulating the learning process. Usually, fully connected networks are used, where every neuron in a particular layer is connected to all neurons in the sequential layer. An input layer is connected to a hidden layer, which is again connected to an output layer. For every connection between the neurons there is a weight coefficient, calculated during the training phase. The neuron inputs varying and depending on the data type, are multiplied by their weights. Then the sum of the weighted inputs is passed to an activation function, which determines the neuron to activate. During the training process, the loss function is minimized in each epoch and the contribution of each neuron to the overall loss is also assessed. Whenever needed, in each epoch the weights of the neurons are adjusted, so that the loss function is minimized (a.k.a back propagation via gradient descent) [93, 94].

In contrast to machine learning, deep learning has several advantages. For instance, there

is no need for feature engineering which reduces the data pre-processing efforts. Also, deep learning is compliant with and susceptible to GPU parallelization which can reduce significantly the computational time. Moreover, deep learning models can be trained on additional data without the need to start from the scratch. In this way, more features can be learned from a deep learning model and more accurate classification can be reached.

The simplest possible modelling for a neural network would be the following [93, 94]:

$$Y_i = g\Big(b_i + \sum_{i=1}^{n} w_i X_i\Big), \tag{107}$$

where the outputs Y_i of each neuron are updated with their corresponding weights w_i applied to the inputs X_i with some bias b_i and passed to an activation function g. The weights control the signal or the strength of the connection between two neurons. In other words, a given weight will help to decide how much influence the input will have on the output. On the other hand, the biases are an additional inputs into the next layer. Between each layer there must be an activation function in order to extract and link the information between two consecutive layers. Without an activation function a neural network layer would consist of two linear operations: a dot product and an addition, and the model would learn only linear transformations of the input data. In order to learn non-linear features there is a need for non-linearity which can be implemented by the activation function. While the weights control the steepness of the activation function, the biases shift it towards larger or smaller values. Initially, the weights and the biases are assigned random values. In this case, the loss function $\mathcal{L}$ takes high values and classification accuracy is very poor. After each learning step or iteration, the model will try to adjust the weights and the biases in such a way as to minimise the loss function and to improve the classification accuracy. This can be summarised as follows [93, 94]:

$$\theta^{'} = \theta - \eta \nabla_{\theta} \mathcal{L}(\theta), \tag{108}$$

where θ denotes the initial set of parameters, η is the learning rate (which is a hyperparameter to tune) and ∇ denotes the gradient differential operator. Also, the negative sign indicates the direction in which the gradient is updated, which is the opposite to the derivative increase, i.e. towards the minimum value of $\mathcal{L}$. In the best case scenario, the loss function will be convex and will have a global minimum. If one tries to compute the gradient for each training point and average the results, this will be computationally ineffective and would result in a very slow learning of the model. This is also known as gradient descend technique (GD). To overcome this drawback of GD, the data can be randomly split into portions or batches and GD can be applied onto these portions; this is known as stochastic gradient descend (SGD). Then the results can be averaged over the different batches. When a prediction $\hat{y}$ is made of the target y, a cost function C can be used to represent the loss $\mathcal{L}$ of missclassifications in the i-th batch, containing m, the number of data points or inputs [93, 94]:

$$C_i(w_i, b_i) = \frac{1}{m} \sum_{i=1}^{m} \mathcal{L}(y_i, \hat{y}_i), \tag{109}$$

For each batch, the update of the weights and the biases is given by [93, 94]:

$$w_i^{'} = w_i - \frac{\eta}{m} \sum_{i=1}^{m} \frac{\partial C_i}{\partial w_i} \tag{110a}$$

$$b_i^{'} = b_i - \frac{\eta}{m} \sum_{i=1}^{m} \frac{\partial C_i}{\partial b_i} \tag{110b}$$

Since the layers of the network will be executed sequentially, this is known as a feed-forward network. In the input, all available data are received after initial pre-processing, labelling and formatting. Then the convolution, relying on a relationship between the adjacent pixels, takes place. All the features consequently become the input of the next layer. To perform the convolution in each layer, a sliding window or kernel is selected (in this study its size is 3 x 3). Its movement across the image is done in a pixel-by-pixel manner starting from top left to bottom right (a.k.a the stride). Suppose we have an i x j image in a given layer l and we want to perform a convolution via kernel with size n x n. The output will be with size $(i-n+1)$ x $(j-n+1)$, where the convolution with the activation function can be represented as [93, 94, 101]:

$$Y_{ij}^{l} = g\left[b_l + \sum_{k=0}^{n} \sum_{l=0}^{n} w_{kl} X_{(i+k)(j+l)}^{l}\right]. \tag{111}$$

When the complete coverage of the image is done by the kernel, a filter f is produced. Each filter contains different results and information. In each convolution layer, all filters n_f form the so called feature map. The features F_i in the i-th layer can be calculated by [93, 94, 101]:

$$F_i = Px \cdot Py \cdot n_f^i, \tag{112}$$

where Px and Py are the number of pixels in the image of the given layer and n_f^i is the total number of filters for the same layer. On the other hand, the number of parameters per layer P_i can be calculated as [93, 94, 101]:

$$P_i = (Kx \cdot Ky \cdot c_j + 1) \cdot n_f^i, \tag{113}$$

where Kx and Ky are the number of pixels in the convolution kernel and c_j is the color channel of the image, always equal to unity in our case (i.e. gray-level imaging). An activation function is implemented to produce the outputs of the given layer. Typically, a rectified linear unit (RELU) function g is used of the form [93, 94, 101]:

$$g(x) = max(0, x), \tag{114}$$

where any input value less than zero is set to zero.

If the model is too complex it will learn very well the training data and will be able to predict

accurately the training data set, but it will not be able to perform properly on the unseen data. This is known as overfitting; to prevent it, the technique of dropout can be used, typically with a fraction of 0.2 as a regularization method [102]. More precisely, the dropout reduces overfitting by randomly dropping out a part of the neurons during training. This has the effect of preventing the network from relying on any particular set of neurons for making predictions. By forcing the model to use different subsets of neurons for each training sample, the model will learn more robust features that are useful across a variety of inputs. Additionally, overfitting and also computation time can be reduced by adding a pooling layer after each convolutional layer. This downsizing technique examines a square of features and preserves only the one with a maximum value (in this study, the square is 2 x 2). By disregarding the rest of the features, the model becomes more general.

After the final pooling layer, we need to pass all the information to be processed by a layer of neurons. To each neuron and input information, weight and bias coefficients are assigned. Then, we need for the final output of the model an array of 3 probabilities which will help us to decide about the tissue type of classification, according to the model. Therefore, a flattening layer is needed. All the layers before the flattening one will learn the features of the tissue structures of the skin samples. Taking all these features and passing them to a dense layer with 256 neurons (this value was found to be optimal for the current model) results in learning the relationship between these features. Finally, to produce the classifications in the dense layer, 3 neurons are required as we have 3 classes. For this purpose, the "softmax" activation function (A_f) to convert the classification probabilities can be used [93, 94, 101]:

$$A_f(\bar{z})_i = \frac{e^{z_i}}{\sum_j^{N_c} e^{z_j}}. \tag{115}$$

In the above equation $\bar{z}$ is the input vector to the "softmax" function from the previous layer, z_i is the i-th element of the input vector $\bar{z}$, while in the denominator the normalization term ensures that all output values will sum to unity to represent the probability distribution of all classes N_c after the classification. The exponential function ensures that the values in the nominator and the denominator will be positive.

Next, the loss function L_f to be used in our neural network was chosen to be the categorical cross-entropy [93, 94, 101]:

$$\mathcal{L} = -\sum_i^{N_c} T_{c_i} log[A_f(\bar{z}_i)], \tag{116}$$

where T_{c_i} is the true class. $\mathcal{L}$ is subject to minimization by the optimizer, while during the training phase of the neural network an adaptive learning rate of the optimizer was chosen. The loss function is minimised as discussed above by the technique of back propagation and updating the weight and bias values. Depending on the data and the model, different number of epochs will be needed to reach a minimum of the loss function and to maximize the classification accuracy. The epochs are nothing else than the number of times the model processes the entire set of the

training data (for this model, 30 epochs were found to be the optimal number). Overfitting can be also reduced if the training data are split into smaller portions or batches. In this model, a batch size of 16 was selected, which is the number of samples to process at a time during each epoch [103].

7 Bibliography

References

[1] Bertrand N., Drevillon B., Bulkin P., In situ infrared ellipsometry study of the growth of plasma deposited silica thin films, J. Vac. Sci. Technol. A, 16(1), 63-71, (1998).

[2] Gottlieb D., Arteaga O., Mueller matrix imaging with a polarization camera: application to microscopy, Biomed. Opt. Express, 29(21), 34723–34734, 2021.

[3] Manhas S., Swami M.K., Buddhiwant P., Ghosh N., Gupta P.K., Singh K.L, Mueller matrix approach for determination of optical rotation in chiral turbid media in backscattering geometry, Opt. Express, 14378, 190–202, 2006.

[4] Novikova T., Bulkin P., Popov V., Haj Ibrahim B., De Martino A., Mueller polarimetry as a tool for detecting asymmetry in diffraction grating profiles, J. Vac. Sci. Tech. B. 29(5), 051804 (2011).

[5] Perez J. J., Ossikovski R., Polarized Light and the Mueller Matrix Approach (CRC Press, Taylor and Francis, 2016).

[6] Meglinski I., Trifonyk L., Bachinsky V., Vanchulyak O., Bodnar B., Sidor M., Dubolazov O., Ushenko A., Ushenko Y., Soltys I. V., Bykov A., Hogan B., Novikova T., Shedding the Polarized Light on Biological Tissues (Singapore: Spinger Briefs in Applied Science and Technology) (2021).

[7] Mazumder N., Xiang L., Qiu J., Fu-Jen K., In pixel analysis of molecular structure with Stokes vector-resolved second harmonic generation microscopy, Proc. SPIE, 8948, 894822, (2014).

[8] He C., He H., Chang J., Chen B., Ma H., J. Booth M., Polarisation optics for biomedical and clinical applications: A review, Nat. Light: Science & Applications, 10(194), 1–20, 2021.

[9] Sridhar S., Da Silva A., Enhanced contrast and depth resolution in polarization imaging using elliptically polarized light, Journal of Biomed. Opt., 21(7), 071107, 2016.

[10] Kumar N.D., Dey R., Chakraborty S., Panigrahi K.P., Meglinsk I., Ghosh N., Quantitative assessment of submicron scale anisotropy in tissue multifractality by scattering Mueller matrix in the framework of Born approximation, Opt. Commun., 43(15), 172-178, 2018.

[11] Manhas S., Mahesh K. S., Hari S. P., Abha U., Ghosh N., Gupta P. G., Polarized diffuse reflectance measurements on cancerous and noncancerous tissues, 2(10), 581-587, 2009.

[12] Kamal A.M., Pal U.M., Kumar A., Gunabhi R.D., Pandya H. J., Toward the development of portable light emitting diode-based polarization spectroscopy tools for breast cancer diagnosis, 15(3), e202100282, 2012.

[13] Wang J., Zheng W., Lin K., Huang Z., Integrated Mueller-matrix near-infrared imaging and point-wise spectroscopy improves colonic cancer detection, 7(4), 1116-1126, (2016).

[14] Cheong W., Welch A., Prahl S., A review of the optical properties of biological tissues, IEEE J. Quantum Electron., 26, 2166–2185, (1990).

[15] Mazur P., Freezing of living cells: Mechanisms and implications, Am. J. Physiol., 247, C125, (1984).

[16] Karlsson J. O., Toner M., Long-term storage of tissues by cryopreservation: Critical issues, Biomaterials 17, 243, (1996).

[17] Ladanyi C., Mor A., Christianson M. S., Dhillon N., Segars J. H., Recent advances in the field of ovarian tissue cryopreservation and opportunities for research, J. Assist. Reprod. Genet., 34, 709-722, (2017).

[18] Shi Q., Wang Y., Shangwei L., Vitrification versus slow freezing for human ovarian tissue cryopreservation: a systematic review and meta-anlaysis, Nature Sci. Rep. 7, 8538, (2018).

[19] Stanzel B.V., Espanac E.M., Grueterich M., Kawakita T., Parel J.M., Tseng S.C., Binder S., Amniotic membrane maintains the phenotype of rabbit retinal epithelial cells in culture, Exp. Eye Res. 80, 103, (2005).

[20] Manuchehrabadi N., Gao Z., Zhang J., Ring H.L., Shao Q., Liu F., McDermott M., Fok A., Rabin Y., Brockbank K.G.M, Garwood M., Haynes C.L., Bischof J.C., Improved tissue cryopreservation using inductive heating of magnetic nanoparticles, Sci. Transl. Medicine, 9, eaah4586, (2017).

[21] Kobayashi A., Golash N. H., Kirschvink J.L., A first test of the hypothesis of biogenic magnetite-based heterogeneous ice-crystal nucleation in cryopreservation Cryobiology, 72, 216e224, (2016).

[22] Kobayashi A., Horikawa M., Kirschvink J.L., Golash H.N., Magnetic control of heterogeneous ice nucleation with nanophase magnetite: Biophysical and agricultural implications, PNAS, 115(21), 5383–5388, (2018).

[23] Edge S., American Joint Committee on Cancer (AJC12), 6th ed., Springer, Chicago, (2002).

[24] Zaytoun O.M. and Jones J.S., Prostate cancer detection after a negative prostate biopsy: Lessons learnt in the Cleveland Clinic experience, Int. J. Urol. 18, 557–568, (2011).

[25] Day C., Histopathology. Methods and Protocols 1st ed., Humana Press, New Jersey, (2014).

[26] Rodriguez C., Van Eeckhout A., Ferrer L., Garcia-Caurel E., Gonzalez-Arnay E., Campos J., Campos J., Lizana A., Polarimetric data-based model for tissue recognition, Biomed. Opt. Express, 12, 4852–4872, (2021).

[27] Zhu Y., Dong Y., Yao Y., Si L., Liu Y.,Ma H., Probing layered structures by multi-color backscattering polarimetry and machine learning, Biomed. Opt. Express, 12 4324–4339, (2021).

[28] Yousafa M., Ahmad I., Khurshid A., Ikram M., Machine assisted classification of chicken, beef and mutton tissues using optical polarimetry and bagging model, Photodiagn. Photodyn., 31, 101779, (2020).

[29] Queau Y., Leporcq F., Lechervy A., Alfalou A., Learning to classify materials using Mueller imaging polarimetry, Proc. SPIE, 11172, 111720Z, (2019).

[30] Vaughn I., Hoover B., Tyo S., Classification using active polarimetry, Proc. SPIE, 8364, 83640S, (2012).

[31] Panigrahi S., Swarnkar T., Machine learning techniques used for the histopathological image analysis of oral cancer-a review, Open Bioinform. J., 13 106–118, (2020).

[32] 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, (2021).

[33] 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, Photodiagn. Photodyn., 14, 134–141, (2016).

[34] Zhou X., Ma L., Brown W., Little J., Chen A., Myers L., Sumer B. D., Fei B. Automatic detection of head and neck squamous cell carcinoma on pathologic slides using polarized hyperspectral imaging and machine learning, Proc. SPIE, 11603, 16030Q, (2021).

[35] Mukhopadhyay S., Kurmi I., Dey R., Das N., Pradhan S., Pradhan A., et al. Optical diagnosis of colon and cervical cancer by support vector machine., Proc. SPIE, 9887, 98870U, (2016).

[36] James G., Witten D., Hastie T., Tibshirani R, An Introduction to Statistical Learning: with Applications in R, (New York: Springer), (2013).

[37] James G., Witten D., Hastie T., Tibshirani R, Taylor J., An Introduction to Statistical Learning: with Applications in Python, (Springer Nature Switzerland AG: Springer Cham), (2023).

[38] Hastie T., Tibshirani R., Friedman J., The elements of statistical learning: data mining, inference, and prediction, Second edition (New York: Springer), (2009).

[39] Born M., Wolf E., Principles of Optics, 7th edition, Pergamon Press, Cambridge University Press, (1999).

[40] Azzam R.M.A., Bashara N.M., Ellipsometry and polarized light, Amsterdam, North-Holland, (1977).

[41] Goldstein D.H., Polarized Light - Third edition, CRC Press, Taylor and Francis, 2010.

[42] Lihong V. Wang, Hsin Wu, Biomedical optics, principles and imaging, Wiley and Sons, New Jersey, (2007).

[43] Palumbo G., Pratesi R., Lasers and Current Optical Techniques in Biology, (Comprehensive Series in Photochemical & Photobiological Sciences, 2004).

[44] Ghosh N., Vitkin I., Tissue polarimetry: concepts, challenges, applications, and outlook, J. Biomed. Opt., (16), 110801-1-29, (2011).

[45] Ortega-Quijano N., Fanjul-Vélez F., Arce-Diego J.L., Polarimetric study of birefringent turbid media with three-dimensional optic axis orientation, Biomed Opt Express. 5(1), 287 (2014).

[46] Kunnen B., Macdonald C., Doronin A., Jacques S., Eccles M., Meglinski I., Application of circularly polarized light for non-invasive diagnosis of cancerous tissues and turbid tissue-like scattering media, J. Biophotonics, 8, 317-323, (2015).

[47] Macdonald C., Doronin A., F. Pena A., Eccles M., Meglinski I., Diffusing-Wave Polarimetry for tissue diagnostics, Proc. of SPIE., 8940, 894007-1-7, (2014).

[48] Ahmad I., Ahmad M., Khan K., Ashraf S., Ahmad S., Ikram M., Ex vivo characterization of normal and adenocarcinoma colon samples by Mueller matrix polarimetry, J. Biomed. Opt., 20(5), 056012-1-8, (2015).

[49] Jacques S., Roman J.R., Lee K., Imaging superficial tissues with polarized light, Lasers in Surgery and Medicine, 26(2), 119-129, (2000).

[50] Jacques S., Roman J.R., Lee K., Imaging of superficial tissues with polarized light, Annals of Biomedical Engineering, 28(S1), S-39, (2000).

[51] Jacques S., Roman J.R., Lee K., Imaging skin pathology with polarized light, Journal of Biomed. Opt., 7(3), 329-340, (2002).

[52] Roman J.R., Lee K., Prahl S.A., Jacques S., Desing, testing and clinical studies of a handheld polarized light camera, Journal of Biomed. Opt., 9(6), 1305-1310, (2004).

[53] Cloude S.R., Conditions for the Physical Realizability of Matrix Operators in Polarimetry, Proc. SPIE, Polarization Considerations for Optical Systems, 1166, 177-185, (1989).

[54] Azzam R.M.A., Propagation of partially polarized light through anisotropic media with or without depolarization: A differential 4 X 4 matrix calculus, Journal of Opt. Soc. Am., 68(12), 1756 - 1767, (1978).

[55] Ossikovski E., Differential matrix formalism for depolarizing anisotropic media, Opt. Let., 36(12), 2330–2332 (2011).

[56] Lu S.-Y., Chipman R.A., Interpretation of Mueller matrices based on polar decomposition, Journal of Opt. Soc. of Am. A, 13(5), 1106 – 1113 (1996).

[57] Ossikovski R., Analysis of depolarizing Mueller matrices through a symmetric decomposition, J. Opt. Soc. Am. A, 26, 1109–1118 (2009).

[58] Compain E., Poirier S., Drevillon B., General and self-consistent method for the calibration of polarization modulators, polarimeters and Mueller-matrix ellipsometers, Appl. Opt., 38(16), 3490-3502, (1999).

[59] D. Ivanov, L. Si, T. Maragkou, P. Schucht, H. Ma, M.-C. Schanne-Klein, R. Ossikovski, T. Novikova, Impact of corpus callosum fiber tract crossing on polarimetric images of human brain histological sections: ex vivo studies in transmission configuration, J. Biomed. Opt., 28(10), 102908, (2023) .

[60] Tyo J. S., Signal-To-Noise Ratio and Minimization of Systematic Error, Appl. Opt. 41(4), 619–630, (2002).

[61] Zhang H., Xiang H. and Wei Y., Condition numbers for linear systems and Kronecker product linear systems with multiple right-hand sides, Int. J. Comput. Math., 84(12), 1805–1817, (2008).

[62] Twietmeyer K. and Chipman R., Condition Number as a Metric for the Effectiveness of Polarimetric Algorithms, Opt. Eng. 25(16), 2493–2493, (2005).

[63] Ivanov D., Dremin V., Borisova E., Bykov A., Meglinski I., Novikova T. and Ossikovski R., Symmetric decomposition of Mueller matrices reveals a new parametric space for polarimetric assistance in colon cancer histopathology, Proc. SPIE 11646, 1164614, (2021).

[64] Pierangelo A., Manhas S., Benali A., Antonelli M.R., Novikova T., Validire P., Gayet B. and De Martino A., Use of Mueller polarimetric imaging for the staging of human colon cancer, Proc. SPIE 7895, 7895E, (2011).

[65] Nishizawa N., Al-Qadi B. Kuchimaru T., Angular optimization for cancer identification with circularly polarized light, J. Biophoton., 14(3), e202000380, (2020).

[66] Solomon C., Breckon T., Fundamentals with digital image processing, Whiley-Blackwell, John Wiley & Sons, (2011).

[67] Mould, R., Introductory Medical Statistics, 3rd edition, Boca Raton: CRC Press, (1998).

[68] Axer, H., Axer, M., Krings, T. et al., Quantitative estimation of 3-D fiber course in gross histological sections of the human brain using polarized light, J. Neurosci. Methods, 105(2), 121–131, (2001).

[69] Axer, M., Graessel, D., Kleiner, M. et al., High-resolution fiber tract reconstruction in the human brain by means of three-dimensional polarized light imaging, Front. Neuroinformatics 5, 34, (2011).

[70] Sanai N., Berger M. S., Glioma extent of resection and its impact on patient outcome, Neurosurgery 62, 753–764, (2008).

[71] Keles G, Lamborn K., Berger M., Low-grade hemispheric gliomas in adults: A critical review of extent of resection as a factor influencing outcome, J. Neurosurg. 95, 735–745, (2001).

[72] Chaichana, K., Cabrera-Aldana, E., Jusue-Torres, I., When gross total resection of a glioblastoma is possible, how much resection should be achieved?, World Neurosurgery, 82, e257–e265, (2014).

[73] Schucht, P., Beck, J., Seidel, K. et al., Extending resection and preserving function: modern concepts of glioma surgery, Swiss Med. Wkly 145, :w14082, (2015).

[74] Kubben, P. L., Meulen, K. J. T., Schijns, O. E., ter Laak-Poort, M. P., van Overbeeke, J. J., Santbrink, H. V., Intraoperative MRI-guided resection of glioblastoma multiforme: A systematic review, Lancet Oncol., 12(11), 1062–1070, (2011)

[75] Schucht, P. Beck, J. Abu-Isa, J., Andereggen, L., Murek, M., and Kathleen Seidel, K., Stieglitz, L., Raabe, A., Gross total resection rates in contemporary glioblastoma surgery: results of an institutional protocol combining 5-aminolevulinic acid intraoperative fluorescence imaging and brain mapping, Neurosurgery, 75(5), 927-935 (2012)

[76] Schucht, P., Lee, H. R., Mezouar, H. M., Hewer, E., Raabe, A., Murek, M., Zubak, I., Goldberg, J., Kovari, E., Pierangelo, A., Novikova, T., Visualization of white matter fiber tracts of brain tissue sections with wide-field imaging Mueller polarimetry, IEEE Trans. Med. Imaging, 39(12), 4376–4382, (2020)

[77] Novikova, T., Pierangelo, A., Schucht, P., Meglinski, I., Rodriguez-Nunez, O., Lee, H. R., Mueller Polarimetry of Brain Tissues, in Polarized Light in Biomedical Imaging and Sensing, J. C. Ramella-Roman, T. Novikova (Eds), Springer, Cham, 2023

[78] Rodriguez-Nunez, O., Novikova, T., Polarimetric techniques for the structural studies and diagnosis of brain tissue, Adv. Opt. Technol. 11(5-6) 157-171 (2022)

[79] Felger, L., Rodríguez-Núñez, O., Gros, R., Maragkou, T., McKinley, R., Moriconi, S., Murek, M., Zubak, I., Novikova, T., Pierangelo, A., Schucht, P., Robustness of the wide-field imaging Muller polarimetry for brain tissue differentiation and white matter fiber tract identification in surgery-like environment – an ex vivo study, Biomed. Opt. Express, 14(5), 2400-2415 (2023)

[80] Baynes K., Corpus callosum in Encyclopedia of the Human Brain, V.S. Ramachandran (Ed), Academic Press (2002).

[81] Chang T., Zhao G, Ice Inhibition for Cryopreservation: Materials, Strategies and Challenges, Advanced science, 8, 2002425, (2021).

[82] Ito A, Yoshioka K., Masumoto, S., Sato K., Hatae, Y., Nakai, T., Yamazaki, T., Takahashi, M., Tanoue, S., Horie, M., Magnetic heating of nanoparticles as a scalable cryopreservation technology for human induced pluripotent stem cells, Sci. Rep., 10, 13605, (2020).

[83] Zhao G, Ring Hl., Sharma A., Namsrai B., Tran N., Finger E., Garwood M., Haynes C., Bischof J., Preparation of Scalable Silica-Coated Iron Oxide Nanoparticles for Nanowarming, Advanced science, 7(4), 1901624, (2020).

[84] Manuchehrabadi N., Gao, Z., Zhang J., Ring H., Shao Q., Liu F., McDermott M., Fok A., Rabin Y., Brockbank K., Garwood M., Haynes C., Bischof J., Improved tissue cryopreservation using inductive heating of magnetic nanoparticles, Sci. Tansl. Med., 9(379), (2017).

[85] Chiu-Lam A., Staples E., Pepine C., Rinaldi C., Perfusion, cryopreservation, and nanowarming of whole hearts using colloidally stable magnetic cryopreservation agent solutions, Science Advances, 7(2), 1901624, (2021).

[86] Kendall, R., Feghali-Bostwick, C., Fibroblasts in fibrosis: novel roles and mediators, Front Pharmacol., 5(123), 1–13 (2014).

[87] Rajan, N., Habermehl, J., Cote, M., Doillon, C., Mantovani, D., Preparation of ready-to-use, storabale and reconstituted type I collagen from rat tail tendon for tissue engineering applications, Nature 1(6), 2753–2757, (2006).

[88] Yoo S., Ossikovski R. and Garcia-Caurel E., 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(1), 870–877, (2017).

[89] K. Wolff, L. Goldsmith, S. Katz, B. Gilchrest, A. Paller, D. Leffell, Dermatology in general medicine, 7th edition, McGraw-Hill: New York, (2008).

[90] W. James, T. Berger, D. Elston, Andrew's Diseases of the skin: clinical dermathology, 11th edition, Elsevier, (2011).

[91] J. Bolognia, J. Schaffer, K. Duncan, C. Ko, Dermatology Essentials, 1st edition, Elsevier, (2014).

[92] D. Elder, R. Elenitsas, A. Rubin, M. Ioffreda, J. Miller, O. Fred Miller III, Atlas and synopsis of Lever's histopathology of the skin, 3rd edition, LWW, (2013).

[93] Deitel P., Deitel, H., Intro to Python for Computer Science and Data Science, 1st edition, Pearson Education, (2020).

[94] Francois Chollet, Deep Learning with Python, 1st edition, Manning Publications, (2017).

[95] Bykov A., Popov A., Priezzhev A., and Myllyla R., Multilayer tissue phantoms with embedded capillary system for oct and doct imaging, Proc. SPIE, 8091, 80911R, (2011).

[96] Moffitt T. P., Chen Y.-C. and Prahl S. A., Preparation and characterization of polyurethane optical phantoms, J. Biomed. Opt., 11(4), 041103, (2006).

[97] Ivanov, D., Dremin, V., Borisova E., Bykov A., Novikova T., Meglinski I., Ossikovski R., Polarization and depolarization metrics as optical markers in support to histopathology of ex vivo colon tissue, J. Biomed. Opt. Express, 12(7), 4560-4572, (2021).

[98] Pierangelo A., Benali A., Antonelli M., Novikova T., Validire P., Gayet B., De Martino A., Ex-vivo characterization of human colon cancer by Mueller polarimetric imaging, 19(2), 1582-1593, (2011).

[99] Ivanov, D., Dremin, V., Genova T., Bykov A., Novikova T., Ossikovski R., Meglinski I., Polarization-based histopathology classification of ex vivo colon samples supported by machine learning, Front. Phys., 9, 814787, (2022).

[100] Volovici V., Syn N., Ercole A., Zhao J., Liu N., Steps to avoid overuse and misuse of machine learning in clinical research, Nat. Med., 28, 1996–1999, (2022).

[101] Wiatowski T., Boelcskei H., A Mathematical Theory of Deep Convolutional Neural Networks for Feature Extraction, IEEE, 64(3), 1845-1866, (2018).

[102] Phaisangittisagul E., An analysis of the regularization between L2 and dropout in single hidden layer neural network, 7-th International Conference on Intelligent Systems, Modelling and Simulation, 16744225, 1-6, (2016).

[103] Ioffe S., Szegedy C., Batch normalization: Accelerating deep network training by reducing internal covariate shift, Proceedings of the 32-nd International Conference on Machine Learning, PMLR 37, 448-456, (2015).

8 Educational and scientific part

8.1 Courses taken:

1. Machine and Deep Learning, École Polytechnique, Fall semester 2022-2023 - 36 hours;
2. Scientific Python – beginner, University of Bordeaux, Fall semester 2022-2023 - 12 hours;
3. Rendre sa recherche plus transparente, optimiser ses processus: data science en R, University of Bordeaux, Fall semester 2021-2022 - 20 hours;
4. Welcome Day of the Doctoral School, IP Paris, Fall semester 2022-2023 - 3 hours;
5. Welcome Day of the Doctoral School, IP Paris, Fall semester 2021-2022 - 3 hours;
6. Telling a story on your achievements to convince recruiters, Fall semester 2022-2022 - 3.5 hours;
7. Scientific Research Ethics, École Polytechnique, Spring semester 2021-2022 - 6 hours;
8. Conflict management, École Polytechnique, Spring semester 2021-2022 - 21 hours;
9. Identify, prevent and efficiently deal with your stress, École Polytechnique, Spring semester 2021-2022 - 12 hours;
10. French Language Course, A0 level, Science Accueil, Fall semester 2020-2021 - 60 hours;
11. French Language Course, A1 level, Science Accueil, Spring semester 2020-2021 - 60 hours;
12. French Language Course, A2 level, École Polytechnique, Fall semester 2021-2022 - 20 hours;
13. French Language Course, A2 level, École Polytechnique, Spring semester 2021-2022 - 20 hours;
14. French Language Course, B1 level, Télécom Paris, Fall semester 2021-2022 - 19.5 hours;

8.2 Conferences:

1. Oral report, SPIE Photonics West, Polarized light and Optical Angular Momentum for biomedical diagnostics, 6-11 March 2021, San Francisco, CA, USA, title of presentation: *"Symmetric decomposition of Mueller matrices reveals a new parametric space in colon cancer histopathology"*, available at: `https://doi.org/10.1117/12.2578090`;
2. Oral report, SPIE Photonics West, Polarized light and Optical Angular Momentum for biomedical diagnostics, 22-27 January 2022, San Francisco, CA, USA,, title of presentation: *"Polarimentric differentiation of ex vivo colon samples complemented by machine-learning"*, available at: `https://spie.org/photonics-west/presentation/`;
3. Oral report, Journées Imagerie Optique Non Conventionnelle, 17-18 March 2022, Institut Langevin, Paris, France, title of presentation: *"Stokes-Mueller polarimerty as a tool for histopathology assessment of ex vivo colon samples supported with machine learning"*;
4. Poster report, Journées Imagerie Optique Non Conventionnelle, 30-31 March 2023, Institut Langevin, Paris, France, title of presentation: *"Mueller polarimerty assessment of ex vivo skin samples in different histological conditions"*;
5. Oral report, SPIE/OPTICA Europeen Conference on Biomedical Optics (ECBO), 25-29 June 2023, Munich, Germany, title of the presentation: *"Characterization of different thawing mechanisms of fibroblast cell-containing tissue models by Mueller polarimetry and statistical analysis"*.

8.3 Project participation:

1. **Campus France PHC Rila 2022 program, AURORA**
   'Digit**A**l histology of tiss**U**e with Muelle**R** micr**O**scopy and machine lea**R**ning **A**pproach'
   Duration: 2 years
   Partner organization: Institute of Electronics, Bulgarian Academy of Sciences
   Project coordinators: Dr. Tatiana Novikova, LPICM; Dr. Tsanislava Genova.

In the framework of this project, my contribution as a PhD student was to assist in the preparation and submission of the project. Then, after the project's acceptance, to calibrate, operate and perform polarimetric measurements in transmission with the Mueller microscope presented in Section 2.2 with thin histological skin samples. The polarimetric results were presented in Section4.3 after application of the mathematical framework behind the decomposition algorithms in Section1. After, a scientific visit in the Institute of Electronics, "Acad. E. Djakov", Bulgarian Academy of Sciences, Sofia, Bulgaria, was successfully accomplished in which an oral report was made to evaluate the project progress. Additionally, during the visit the Deep learning model was discussed and the results from Polarimetry and Deep learning were included in a manuscript for publication in a peer-review journal (up to the point of the Thesis submission, the manuscript has not been yet sent for publication).

8.4 Publications:

8.4.1 Published in peer-review journals:

1. D. Ivanov, V. Dremin, E. Borisova, A. Bykov, T. Novikova, I. Meglinski, R. Ossikovski, *"Polarization and depolarization metrics as optical markers in support to histopathology of ex vivo colon tissue"*, Biomed. Opt. Expess, 12(7), 4560-4572, 2021, https://www.doi.org/10.1364/BOE.426713;

2. D. Ivanov, V. Dremin, Ts. Genova, A. Bykov, T. Novikova, R. Ossikovski, I. Meglinski, *Polarization-Based Histopathology Classification of Ex Vivo Colon Samples Supported by Machine Learning*, Front. Phys, 9, 814787, 2022, https://doi.org/10.3389/fphy.2021.814787;

3. D. Ivanov, L. Si, T. Maragkou, P. Schucht, H. Ma, M.-C. Schanne-Klein, R. Ossikovski, T. Novikova, *Impact of corpus callosum fiber tract crossing on polarimetric images of human brain histological sections: ex vivo studies in transmission configuration*, J. Biomed. Opt., 28(10), 102908, 2023, https://doi.org/10.1117/1.JBO.28.10.102908;

4. D. Ivanov, A. Hoeppel, J. Schwebler, C. Lotz, S. Dembski, R. Ossikovski, T. Novikova, *Assessment of the Impact of Nanowarming on Microstructure of Cryopreserved Fibroblast-Containing 3D Tissue Models Using Mueller Polarimetry*, Photonics, 10(10), 1129, 2023, https://doi.org/10.3390/photonics10101129.

8.4.2 Published in conference proceedings:

1. D. Ivanov, E. Borisova, T. Novikova, R. Ossikovski, *"Experimental validation of depolarizing Mueller matrix model via ex vivo colon samples"*, Journal of Physics Conference Series, J. Phys.: Conf. Ser., 1859, 012042, 2021, https://www.doi.org/10.1088/1742-6596/1859/1/012042;

2. D. Ivanov, V. Dremin, E. Borisova, A. Bykov, I. Meglinksi, T. Novikova, R. Ossikovski, *"Symmetric decomposition of Mueller matrices reveals a new parametric space in colon cancer histopathology"*, Proc. SPIE, 11646, 1164614, 2021, https://doi.org/10.1117/12.2578090;

3. D. Ivanov, A. Hoeppel, J. Schwebler, C. Lotz, S. Dembski, R. Ossikovski, T. Novikova, *Characterization of different thawing mechanisms of fibroblast cell-containing tissue models by Mueller polarimetry and statistical analysis*, Proc. of SPIE, 12627, 126272G, 2023, https://doi.org/10.1117/12.2669082;

4. T. Genova, D. Ivanov, L. Zaharieva, V. Mircheva, T. Novikova, R. Ossikovski, P. Troyanova, *Differentiation of collagen-related skin diseases through polarimetry and fluorescence*, Proc. of SPIE, 12629, 1262903, 2023,
https://doi.org/10.1117/12.2670922.

8.5 Award:

I4H BME conference grant (ANR France 2030 funding: EUR Bertip, ANR-18-EURE-0002).

8.6 Mentorship:

Master (M2) Student – Myeongseop Kim, April-August 2021, *"Optical measurements of biological tissues with Mueller microscope, combined with Machine Learning"*.

E4H
ENGINEERING FOR HEALTH
INTERDISCIPLINARY CENTER

Certificate

Engineering for Health Annual Forum

5 July 2023 | IP Paris, France

This is to certify that a

BME Conference Fellowship

was awarded to:

Deyan Ivanov

Abdul Barakat
Co-director of the E4H Center

INSTITUT POLYTECHNIQUE DE PARIS

ECOLE
DOCTORALE

Titre: Polarimétrie de Mueller pour l'évaluation de la microstructure tissulaire en histologie numérique et cryoconservation

Mots clés: Polarimétrie du Mueller, Algorithmes de décomposition, Traitement des images, Machine learning, Analyse statistique, Histologie digitale

Résumé: La polarisation de la lumière peut être utilisée comme modalité optique sensible aux échantillons biomédicaux. Un polarimètre imageur dans le visible a été utilisé pour les mesures. A l'aide de celui-ci, les médecins peuvent être assistés dans leur diagnostic par histologie numérique sensible à la polarisation. Pour les échantillons épais, la géométrie en réflexion a été préférée. En transmission, des tranches ex vivo de tissu cervical ont été mesurées pour visualiser l'orientation des fibres cervicales. Des tissus modèles congelés puis décongelés ont été mesurés pour évaluer les dégâts dus à la décongélation. Des échantillons de peau avec des lésions dégénératives et malignes ont été analysés et un modèle deep learning de classification d'images a été élaboré. Un polarimètre non-imageur a été utilisé pour discriminer les réponses d'échantillons de colon ex vivo avec des parties saines et tumorales. Des algorithmes machine learning supervisés et non supervisés ont été aussi employés.

Title: Mueller polarimetry for the assessment of tissue microstructure in digital histology and cryopreservation

Keywords: Mueller polarimetry, Decomposition algorithms, Image processing, Machine learning, Statistical analysis, Digital histology

Abstract: Polarization of light can be used as an optical modality sensitive to biomedical samples. An imaging Mueller polarimeter operating in the visible range was used for the measurements. Thus, medical doctors could be assisted by performing polarization-sensitive digital histology. For thick tissue specimens, reflection geometry and scanning were chosen. In transmission, *ex vivo* brain slides were used to visualize the orientation of brain fiber tracts. Non-frozen, frozen at -80^o C and defrosted with different methods tissue models were measured with the Mueller microscope to assess the damage occurring upon defrosting. Skin samples with degenerative and malignant lesions were analyzed, and a deep learning model was elaborated for image classification. Finally, a polarimetric set-up in reflection was used to scan and classify the polarization response of *ex vivo* colon samples with healthy and tumorous sections. Both supervised and unsupervised machine-learning algorithms were also used.

Institut Polytechnique de Paris
91120 Palaiseau, France