

Aus dem Deutschen Krebsforschungszentrum  
(Vorstand: Prof. Dr. Michael Baumann, Ursula Weyrich)  
Abteilung Intelligente Medizinische Systeme  
(Leiter: Lena Maier-Hein)

# TRANSLATIONAL FUNCTIONAL IMAGING IN SURGERY ENABLED BY DEEP LEARNING

Inauguraldissertation  
zur Erlangung des Doctor scientiarum humanarum (Dr. sc. hum.)  
an der  
Medizinischen Fakultät Heidelberg  
der  
Ruprecht-Karls-Universität

vorgelegt von  
Leonardo Antonio, Ayala Menjivar  
aus  
La Paz, El Salvador

2022



Dekan: Prof. Dr. med. Hans-Georg Kräusslich  
Doktormutter: Lena Maier-Hein



*To the misfits, to the nerds, to those who did not stop believing.*



## Contents

<b>Acronyms</b>	<b>i</b>
<b>List of Figures</b>	<b>iv</b>
<b>List of Tables</b>	<b>xx</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Motivation . . . . .	1
1.2 Background . . . . .	3
1.2.1 Medical fundamentals . . . . .	3
1.2.2 Imaging fundamentals . . . . .	12
1.2.3 Machine Learning fundamentals . . . . .	21
1.3 Related Work . . . . .	28
1.3.1 Band selection in spectral imaging . . . . .	28
1.3.2 Automatic light source estimation . . . . .	32
1.3.3 Perfusion monitoring . . . . .	36
1.4 Objectives and Contributions . . . . .	44
1.5 Outline . . . . .	47
<b>2 Materials and Methodology</b>	<b>49</b>
2.1 Framework for intraoperative functional imaging . . . . .	49
2.2 Spectral imaging . . . . .	50
2.2.1 Real world spectral imaging . . . . .	51
2.2.2 Synthetic spectral imaging . . . . .	55

2.2.3	Task-specific band selection . . . . .	58
2.3	Automatic light source estimation . . . . .	62
2.3.1	Concept overview . . . . .	62
2.3.2	Implementation details . . . . .	63
2.3.3	Hyperparameter settings and baseline . . . . .	65
2.4	Perfusion monitoring . . . . .	67
2.4.1	Perfusion monitoring via regression . . . . .	67
2.4.2	Perfusion monitoring via Out-of-Distribution detection (Clinical study) . . . . .	72
<b>3</b>	<b>Results</b>	<b>79</b>
3.1	Spectral imaging . . . . .	79
3.1.1	Synthetic spectral imaging . . . . .	79
3.1.2	Task-specific band selection . . . . .	82
3.2	Automatic light source estimation . . . . .	85
3.2.1	Light source estimation <i>ex vivo</i> . . . . .	85
3.2.2	Light source estimation <i>in vivo</i> . . . . .	88
3.3	Perfusion monitoring . . . . .	90
3.3.1	Perfusion monitoring via regression . . . . .	90
3.3.2	Perfusion monitoring via Out-of-Distribution detection (Clinical study) . . . . .	100
<b>4</b>	<b>Discussion</b>	<b>117</b>
4.1	Spectral imaging . . . . .	117
4.1.1	Real world spectral imaging . . . . .	118
4.1.2	Synthetic spectral imaging . . . . .	119
4.1.3	Task-specific band selection . . . . .	120
4.2	Automatic light source estimation . . . . .	122
4.3	Perfusion monitoring . . . . .	124
4.3.1	Perfusion monitoring via regression . . . . .	124
4.3.2	Perfusion monitoring via Out-of-Distribution detection (Clinical study) . . . . .	126
<b>5</b>	<b>Summary</b>	<b>135</b>
<b>6</b>	<b>Zusammenfassung</b>	<b>137</b>



<b>A Appendix</b>	<b>141</b>
A.1 Region of interest (ROI) tracking . . . . .	141
A.2 Human kidney tissue Signal-to-Noise Ratio (SNR) . . . . .	141
A.3 Human skin perfusion . . . . .	141
<b>Bibliography</b>	<b>145</b>
<b>Own Contributions and Publications</b>	<b>165</b>
Own share in data acquisition and data analysis . . . . .	165
Own Publications . . . . .	167
<b>Curriculum Vitae</b>	<b>167</b>
<b>Acknowledgement</b>	<b>171</b>
<b>Statutory Declaration</b>	<b>173</b>



## Acronyms

<b>AI</b>	Artificial Intelligence
<b>AHE</b>	Adaptive Histogram Equalization
<b>ASHA</b>	Asynchronous Successive Halving Algorithm
<b>AUC</b>	Area Under the Curve
<b>AU-ROC</b>	Area Under the Receiver Operating Curve
<b>BFS</b>	Best First Search
<b>BMI</b>	Body Mass Index
<b>BoCF</b>	Bag of Color Features
<b>BoFP</b>	Bag-of-Features Pooling
<b>CCD</b>	Charge-Coupled Device
<b>CIFE</b>	Conditional Infomax Feature Extraction
<b>CLAHE</b>	Contrast Limited Adaptive Histogram Equalization
<b>CMIM</b>	Conditional Mutual Information Maximization
<b>CMOS</b>	Complementary Metal-Oxide Semiconductor
<b>CNN</b>	Convolutional Neural Network
<b>CT</b>	Computed Tomography
<b>DCF-CSR</b>	Discriminative Correlation Filter with Channel and Spatial Reliability
<b>DCS</b>	Diffuse Correlation Spectroscopy
<b>DL</b>	Deep Learning

## ACRONYMS

---

<b>DRS</b>	Diffuse Reflectance Spectroscopy
<b>ECoG</b>	Electrocorticography
<b>FLIM</b>	Fluorescence Lifetime Imaging
<b>fMRI</b>	functional Magnetic Resonance Imaging
<b>FWHM</b>	Full Width at Half Maximum
<b>GFP</b>	Green Fluorescence Protein
<b>GPU</b>	Graphics Processing Unit
<b>HSI</b>	Hyperspectral Imaging
<b>HSV</b>	Hue-Saturation-Value
<b>ICAP</b>	Interaction Capping
<b>ICG</b>	Indocyanine Green
<b>INN</b>	Invertible Neural Network
<b>IR</b>	infrared
<b>JMI</b>	Joint Mutual Information
<b>KCl</b>	Potassium Chloride
<b>KDE</b>	Kernel Density Estimation
<b>KL</b>	Kullback-Leibler
<b>KMM</b>	Kernel Mean Matching
<b>LCTF</b>	Liquid Crystal Tunable Filter
<b>LeakyReLU</b>	Leaky Rectified Linear Unit
<b>LED</b>	Light-Emitting Diode
<b>LS</b>	light source
<b>LSI</b>	Laser Speckle Imaging
<b>LUT</b>	look up table
<b>MAE</b>	Mean Absolute Error
<b>MC</b>	Monte Carlo
<b>MCML</b>	Monte Carlo Multi-Layer
<b>MIFS</b>	Mutual Information based Feature Selection
<b>ML</b>	Machine Learning
<b>MLP</b>	Multilayer Perceptron

---

<b>MMD</b>	Maximum Mean Discrepancy
<b>MRI</b>	Magnetic Resonance Imaging
<b>mRMR</b>	Minimum Redundancy Maximum Relevance
<b>MSE</b>	Mean Squared Error
<b>MSI</b>	Multispectral Imaging
<b>NIR</b>	near infrared
<b>OCT</b>	Optical Coherence Tomography
<b>OoD</b>	Out-of-Distribution
<b>OR</b>	Operating Room
<b>PAI</b>	Photoacoustic Imaging
<b>PCA</b>	Principal Component Analysis
<b>PET</b>	Positron Emission Tomography
<b>PHTH</b>	per hundred thousand inhabitants
<b>POS</b>	Polarization Spectroscopy
<b>ReLU</b>	Rectified Linear Unit
<b>RGB</b>	Red-Green-Blue
<b>RKHS</b>	Reproducing Kernel Hilbert Space
<b>RMS</b>	Root-Mean-Square
<b>ROC</b>	Receiver Operating Characteristic
<b>ROI</b>	Region of Interest
<b>SD</b>	Spreading Depolarization
<b>SFDI</b>	Spatial Frequency Domain Imaging
<b>SFS</b>	Sequential Forward Selection
<b>SI</b>	Spectral Imaging
<b>SNR</b>	Signal-to-Noise Ratio
<b>UMAP</b>	Uniform Manifold Approximation and Projection
<b>WAIC</b>	Widely Applicable Information Criterion

## *Acronyms*

---

# List of Figures

- 1.1 **Physiological structure of the brain. (a)** The human brain can be divided into four lobes: frontal, parietal, occipital and temporal, depending on one of their main functions: reasoning, touch sensory, visual and auditory respectively. **(b)** Spreading Depolarizations (SDs) are a phenomenon occurring on the brain cortex that causes massive hyperactivity of neurons followed by inhibition. This phenomenon manifests itself as a wave of depolarized neurons expanding through the cortex. . . . . 5
- 1.2 **The human skin is the largest organ and is structured in a multi-layered fashion.** The human skin is composed of three main layers: epidermis, dermis and subcutaneous layer (*subcutis*). The thickness and precise composition of each varies across the surface of the human body. However, several components can be generally found through the body: hair follicles, nerves, blood vessels, fat, sweat glands, etc. *Image from Madhero88 and M.Komorniczak, CC BY-SA 3.0, via Wikimedia Commons.* . . . . . 8
- 1.3 **Human kidney is perfused through renal artery emanating from the abdominal aorta. (a)** General anatomy of the human kidney. The kidney is covered by the fibrous renal capsule that is composed mainly of fibrous tissue. The adrenal gland sits on top of the kidney and produce a variety of hormones such as adrenaline and cortisol. **(b)** Kidney perfusion through accessory renal arteries branching from the abdominal aorta. Before reaching the renal hilum, arteries usually branch out. One or two accessory arteries are frequently found in humans, while other variations of arterial structures can also be found. . . . . 10

1.4	<b>Assessing tissue perfusion from visual inspection of Red-Green-Blue (RGB) images is challenging.</b> Exemplary laparoscopic images of the kidney surface from different patients under different perfusion conditions. Figures (a), (c) and (e) correspond to perfused kidneys, while figures (b), (d) and (f) represent ischemic kidneys. The induction of ischemia was performed by applying a clamp to the renal artery. . . . .	11
1.5	<b>Hemoglobin is the chromophore with the highest absorption in the visible range for internal organs.</b> (a) Light-tissue interactions in a tissue structure with characteristic layers (e.g. colon). Light is mainly reflected, scattered or absorbed, while luminescence can also occur. Light in the red optical region can penetrate the tissue deeper than light in the blue region. (b) Oxyhemoglobin and deoxyhemoglobin, have similar absorption coefficients while other chromophores like water and lipids absorb much less light in the visible and near infrared (NIR) wavelength range. (c) Spectral images are composed of several bands with narrow light transmission profiles, while Red-Green-Blue (RGB) images are only composed of three bands with wide transmission profiles. Combining the three bands of an RGB image generates a color image that is generally used by medical staff for various procedures. <i>Image taken from Clancy et al. (2020)</i> . . . . .	13
1.6	<b>The presence of fluorescent signal is an indicator of perfusion, while its absence indicates ischemia.</b> Images of a kidney from a female patient after clamping the renal artery and intravenous injection of Indocyanine Green (ICG). The Red-Green-Blue (RGB) image in (a) depicts the presence of an abnormality in the surface of the kidney (Leiomyoma). The absence of fluorescence signal on the kidney (regions without green color in (b)) indicates that the kidney is no longer perfused. . . . .	18
1.7	<b>Spectral Imaging (SI) features a wide range of scanning techniques and detection mechanisms.</b> Scanning can be achieved with a (a) rotating filter wheel, (b) a tunable light source, (c) a spatial line-scan sensor, (d) or a snapshot sensor. These scanning techniques can be coupled with optical parts, depending on the application, such as (e) laparoscopes, (f) flexible endoscopes, (g) microscopes or (h) externally mounted lenses with ring illumination for open surgery. <i>Image taken from Clancy et al. (2020)</i> . . . . .	20



- 
- 1.8 **Invertible Neural Networks (INNs) can learn the *inverse process* in an ill-posed setting.** (a) Standard neural networks require a supervised loss (SL) that causes problems when the mapping between underlying system properties  $\mathbf{x}$  and measurable parameters  $\mathbf{y}$  is not bijective. (b) INNs use a supervised loss only for the well-defined *forward process*  $\mathbf{x} \mapsto \mathbf{y}$ , while an unsupervised loss (USL) is used for the *inverse process*, and to ensure that the latent variables  $\mathbf{z}$  follow a Gaussian distribution. *Figure from Ardizzone et al. (2018).* . . . . . 23
- 1.9 **The core of Invertible Neural Networks (INNs) are *affine coupling layers*.** (a) *Affine coupling layers* split their input into two parts ( $\mathbf{u}_1, \mathbf{u}_2$ ) before being transformed in an alternating manner by functions  $s_i$  and  $t_i$ . (b) The mapping from output to input is done in a similar fashion by rearranging of the computational block. Here,  $\odot$  and  $\oslash$  represent element-wise multiplication and division, respectively. *Figure from (CVL, 2022).* . . . . . 24
- 1.10 **Convolutional operations are in practice applied as matrix multiplications.** Exemplary diagram of a two-dimensional convolutional operation in two channels of an image with two spatial dimensions. . . . . 26
- 1.11 **Oxygenation estimation is possible *in vivo* with random forests.** Small bowel oxygenation estimation (a) before applying a clamp, (b) after applying four clamps, and (c) nine minutes after applying the first clamp. *Image from Wirkert et al. (2016b).* . . . . . 38
- 1.12 **Functional properties are computed from Red-Green-Blue (RGB) images in an iterative process that relies on the Beer-Lambert law.** Spectral data is recovered from RGB images and the spectral power distribution by using a linear square fit, and the concentrations of hemoglobin and oxygen are computed based on a Bayesian estimator. The spectral estimations are then optimized in an iterative process based on such molecular concentrations. *Image from Jones et al. (2017).* . . . . . 39
- 1.13 **Slow sequential spectral recordings can lead to motion artifacts.** Reconstructed Red-Green-Blue (RGB) image and oxygenation estimation of human esophagus. Due to the sequential nature of the recording process (line scan), motion artifacts can be visualized in the image, even under controlled conditions. *Image from (Köhler et al., 2019).* . . . . . 41

<p>1.14 <b>Perfusion estimation in skin is done by pairing a pulse oximeter with an Red-Green-Blue (RGB) camera.</b> This method extracts the perfusion signal from RGB data after performing video stabilization with an optical flow approach. The extracted signal is optimized using the data collected from a pulse oximeter and an outlier weighting estimator. The green box underlines the different steps performed by the estimator. <b>(a)</b> depicts the overall strategy employed to leverage data from an RGB camera and a pulse oximeter, and <b>(b)</b> illustrates the effect of outlier weighting on the perfusion estimation approach. <i>Image from Kumar et al. (2020).</i> . . . . .</p>	42
<p>1.15 <b>This thesis proposes a concept for translational functional imaging in surgery that is validated in a wide range of domains.</b> This concept addresses common roadblocks in the path to translation of functional imaging technologies: a) the lack of labeled <i>in vivo</i> data, b) the slow recording and processing speed of Spectral Imaging (SI) systems, c) the estimation error introduced by the dynamic illumination environment of open procedures, and d) the high inter-patient tissue heterogeneity, for example in the presence of pathologies. . . . .</p>	44
<p>2.1 <b>Our concept for functional imaging during surgical procedures can handle both minimally invasive and open surgery.</b> The pipeline for functional imaging that we propose is capable of optimizing the spectral bands to specific tasks, it can handle open surgical scenarios via a dedicated light source estimation module, and is capable of dealing with data that is out of domain, such as tissue in the presence of pathologies. . . . .</p>	50
<p>2.2 <b>Our multispectral imaging (MSI) system for ischemia monitoring is comparable to current standard equipment in terms of size and weight.</b> <b>(a)</b> Proposed spectral imaging-based laparoscope (left) compared to standard Red, Green, and Blue (RGB) laparoscope (right). <b>(b)</b> Schematic representation of the system developed for spectral tissue analysis in laparoscopic surgery. The dimensions of the laptop and the light source have been scaled down for visualization purposes. . . . .</p>	52

2.3 **Optical properties of our multispectral system.** (a)  $\ell_1$  normalized transmission spectra of the laparoscopes, band-pass filter and C-Mount adapter are shown in the left axis. Extinction coefficients of oxyhemoglobin (HbO<sub>2</sub>) and deoxyhemoglobin (Hb) are shown on the right axis. (b) Representation of the  $4 \times 4$  mosaic pattern of the multispectral camera sensor. Each colored square represents a different filter; these filters form a  $4 \times 4$  pattern that extends over the whole image. (c) The filter responses of the multispectral camera bands. Some bands, such as 5 to 12, show two extra peaks in the spectral response, which are referred to as “second order” peaks. . . . . 54

2.4 **Sequential spectral recordings with glass filters eliminate second order peaks.** Spectral response of the Spectrocam camera filters, no second order peaks are observed. The Spectrocam camera uses a filter wheel, which trades slower recording speed for higher spatial resolution and narrower spectral bands. . . . . 55

2.5 **Labelling multi-layer tissue models by using penetration depth can generate more accurate results.** In contrast to a traditional approach that averages each tissue property up to a depth of  $250\mu m$ , we propose to use a wavelength-dependent penetration depth to assign properties to the entire model. . . . . 58

2.6 **Domain and target-specific simulations can be generated by leveraging unlabeled real spectral data.** Generic simulations are adapted using unlabeled spectral measurements from the target domain. The resulting domain-specific simulations are the basis for the subsequent task-specific band selection. . . . . 59

2.7 **Our method leverages low exposure images to estimate the spectrum of the light source (LS).** Based on low exposure multispectral images, specular highlight masks are computed, which represent regions in the image that are not over- or underexposed and are assumed to contain a dominant specular component. A pre-trained machine learning algorithm tailored to the estimated light source is then applied to high exposure multispectral images to compute functional tissue parameters. . . . . 63

<p>2.8 <b>We use five commonly used types of light sources to validate our method.</b> Reference light sources (LSes). LS 1: Xenon (D-light P 201337 20 endoscopic light source, Karl Storz GmbH, Tuttlingen, Germany); LS 2: Halogen (Halopar 16 GU10 light bulb, Osram<sup>®</sup>, Munich, Germany); LS 3: Fluorescent light (FLS 11W 2700K fluorescent lamp, Paulmann<sup>®</sup>, Springe Völksen, Germany); LS 4: Xenon (Auto LP 5131 endoscopic light source, Richard Wolf GmbH, Knittlingen, Germany); LS 5: Light-Emitting Diode (LED) (Endolight LED 2.2 endoscopic light source, Richard Wolf GmbH, Knittlingen, Germany). <b>(a)</b> Relative irradiance of LS 1-5 <math>\ell_1</math>-normalized. <b>(b)</b> Angular distance matrix between the different LSes. . . . .</p>	<p>64</p>
<p>2.9 <b>Our Convolutional Neural Network (CNN)-based model integrates data preprocessing (e.g. normalization) in the network architecture.</b> <b>(a)</b> The network used for training, whose model parameters are re-used for the live validation network. <b>(b)</b> The live validation network used during the intervention, yielding the functional parameters and Red-Green-Blue (RGB) estimates. The estimated RGB image is only used for visualization. . . . .</p>	<p>71</p>
<p>2.10 <b>We track tissue Regions of Interest (ROIs) with high accuracy in spectral videos.</b> Our tracker analyzes two consecutive images of a spectral video. The RGB images are transformed to Hue-Saturation-Value (HSV) color space and each V channel is normalized using Contrast Limited Adaptive Histogram Equalization (CLAHE), the result is then transformed back to RGB space. The normalized images are fed to a VGG19 and features are extracted from its seventh convolutional layer. These features are then used by the Discriminative Correlation Filter with Channel and Spatial Reliability (DCF-CSR) tracker to generate the new ROI position, which in turn is used to locate the ROI position in subsequent images. <i>Icons by FreePik from flaticon.com</i> . . . . .</p>	<p>74</p>
<p>3.1 <b>The mean penetration depth is smaller than the total penetration depth of each multi-layer tissue model.</b> <b>(a)</b> Two-dimensional representation computed via Principal Component Analysis (PCA) of the complete set of <math>5 \times 10^6</math> simulations of light-tissue interactions. Each dot represents one simulation and its color represents the tissue oxygenation. <b>(b)</b> Distribution of multi-layer tissue model depth and mean light penetration depth (across wavelengths). The mean light penetration depth is usually smaller than the tissue depth. . . . .</p>	<p>80</p>

- 3.2 **Using the penetration depth for multi-layer model labeling assigns properties that are more characteristic of each tissue model.** (a) Relative difference between three-layer tissue models and re-simulated one-layer models for which tissue properties were assigned by averaging each parameter up to 250  $\mu\text{m}$  (orange line) and using the penetration depth for averaging each parameter (green line). (b) Mean Squared Error (MSE) validation loss during training of our CNN model while using the two labeling strategies. Using the penetration depth for model labeling helps during model training by producing a lower validation loss. Model training was done based on the three-layer tissue models. . . . . 81
- 3.3 **Best First Search (BFS) outperforms other band selection algorithms, followed by Sequential Forward Selection (SFS)** (a) Mean Absolute Error (MAE) of various band selection methods applied on the generic training set  $X_g$  and evaluated on the test set  $X_g^{test}$ . *Wrapper* methods (W) outperform *filter* methods (F). The horizontal line represents the result when training and testing on all 101 bands. (b) Ranking stability of different band selection methods across different number of bands. Here, the rank of a method (1: best to 5: worst) is based on the MAE. Each method is color-coded, and the area of each blob at position  $(A_i, \text{rank } j)$  is proportional to the relative frequency  $(A_i)$  each method achieved rank  $j$  for 1000 bootstrap samples. The median rank of each algorithm is indicated by a black cross, and 95% bootstrap confidence intervals across bootstrap samples are indicated by black lines. . . . . 83
- 3.4 **We can adapt generic simulations to a specific domain and use them to estimate oxygenation in real data.** (a) Two dimensional representation, computed via Principal Component Analysis (PCA), of generic Monte Carlo (MC) simulations before and after adapting them to real data from rat head and neck tumor models. (b) Oxygenation map using 20 bands selected with Best First Search (BFS) the tumorous region was identified via Green Fluorescence Protein (GFP) and is indicated by the blue outline. (c) Oxygenation distribution, as estimated by the random forest regressor, for the healthy and tumorous regions identified in (b). The distributions are shown when a regressor is trained on all 101 bands and when training only on 20 selected by BFS. . . . . 84

3.5 **Our method is capable of accurately estimating the light source spectra.** (a) Reference spectra (large circles) and corresponding estimates (crosses) for the light sources (LSes) described in Fig. 2.8. All light source spectra were projected onto the first two principal components of the five reference spectra, determined with Principal Component Analysis (PCA). Crosses of the same color represent different poses of the multispectral camera relative to the same LS. (b) Box plots of the angle between the reference spectrum and the estimated spectrum for each LS. The purple lines represent the mean angular distance for each LS, and the boxes extend between the 25<sup>th</sup> to the 75<sup>th</sup> percentile. . . . . 86

3.6 **Automatic light source estimation can reduce the estimation error of oxygenation *in silico*.** Error in oxygenation estimation when (1) using the reference light source (LS) for training ( $LS_{train} = LS_{test}$ ) (2) using a random LS for training ( $LS_{train} \neq LS_{test}$ ) and (3) using our approach to illuminant estimation to estimate the LS ( $LS_{train} = \hat{LS}_{test}$ ). . . . . 88

3.7 **Our method outperforms all state-of-the-art approaches for different exposure times.** (a) Angular error of baseline methods and our proposed method (Specular highlights) for different exposure times. Box plots of the same color correspond to the same method. (b) Ranking stability of different methods when applied to different exposure times. Here, the rank of a method on each data set (1: best to 5: worst) is based on the angular error. Each method is color-coded, and the area of each blob at position  $(A_i, rank j)$  is proportional to the relative frequency  $(A_i)$  each method achieved rank  $j$  for 1000 bootstrap samples across the different tasks, where one task represents one exposure time. The median rank of each algorithm is indicated by a black cross, and 95% bootstrap confidence intervals across bootstrap samples are indicated by black lines. . . . . 89

3.8 **Automatic light source estimation can reduce the estimation error of oxygenation from spectral images of *in vivo* tissue.** When assuming a constant LS, the estimated blood oxygenation in an Region of Interest (ROI) on the human lips (yellow rectangle) changes when illumination conditions alter (and no longer match the training conditions). Our approach compensates for this by automatic LS calibration. The gap between index 85 and 100 represents the transition phase in which light sources were switched on and off, respectively. . . . . 90

- 
- 3.9 **Generally, lower *learning rates* produced models with the lowest Mean Squared Error (MSE) loss.** (a) Parallel plot of the hyperparameter space explored for hyperparameter tuning. The *learning rate* and the *batch size* are the hyperparameters that mainly influence the training results of our CNN network. (b) Validation loss curves for all trained models, the blue line corresponds to the validation loss of the model with the lowest average MSE loss in the last ten training epochs. . . . . 93
- 3.10 **Our deep learning method estimates oxygenation ( $sO_2$ ) *in silico* with lower estimation error compared to the state-of-the-art random forest.** (a) Oxygenation estimation absolute error for the state-of-the-art random forest and our deep learning approach. (b) Experimental setup for monitoring Spreading Depolarizations (SDs) with a spectral camera, Electrocorticography (ECoG) electrodes were placed at the margins of the craniectomy to monitor SDs. . . . . 94
- 3.11 **Our deep learning method is capable of video-rate visualization of blood oxygenation in the porcine cortex.** (a) Mean oxygenation ( $sO_2$ ) value within each of the defined Regions of Interest (ROIs). The star symbol indicates the spatial location and time point when the Spreading Depolarizations (SDs) were induced by Potassium Chloride (KCl) injection. (b) Electrocorticography (ECoG) electrode measurements typical for SDs. (c) Oxygenation maps of the brain cortex for three different time points (1)-(3), which correspond to the three time points highlighted in a). A deoxygenation wave can be observed traveling from left to right of the brain. 96
- 3.12 **Spectral recording procedure of human skin.** We recorded spectral images as well as spectrometer and oximeter data for each patient of our clinical trial. The spectral and oximeter recordings span across the entire experiment, while the spectrometer measurements were only performed before pressure was applied to the patient’s upper arm and after applying a pressure of 160 mmHg. . . . . 97
- 3.13 **Slight spectral differences can be detected with diffuse reflectance measurements performed with a spectrometer.**  $\ell_2$ -normalized reflectance spectra for all 12 patients of our clinical study, corresponding to two exemplary Regions of Interest (ROIs) annotated in perfused and ischemic tissue (160 mmHg). The solid lines correspond to the median reflectance spectra across all 100 measurements, while the translucent bands around each line correspond to the standard deviation. . . . . 99

3.14	<b>Our deep learning method is capable of monitoring perfusion changes in skin. (a)</b> Representative oxygenation estimations of our deep learning method for patient 2 of our clinical trial. Solid vertical lines represent the time points at which each pressure was applied, while dashed vertical lines represent the moment at which the pressure was released. <b>(b)</b> Perfusion estimations generated with the method proposed by Kumar et al. (2020) for each ROI annotated on patient 2. . . . .	100
3.15	<b>The Monte Carlo (MC) simulations of light transport do not represent pathological tissue in human kidney.</b> Kernel Density Estimation (KDE) of the two-dimensional representation computed with Uniform Manifold Approximation and Projection (UMAP) of the MC simulations and measurements of kidney tissue of patients undergoing partial nephrectomy. . . . .	101
3.16	<b>Multispectral Imaging (MSI) recording procedure in the operating room (OR).</b> We recorded two image sequences (perfused <sub>1</sub> (for training) and perfused <sub>2</sub> (for testing)) before the clamp was applied to the renal artery, and one sequence after applying the clamp (for testing). The laparoscope was removed and re-inserted in the patient’s abdominal cavity before recording the perfused <sub>2</sub> . This ensured equal recording conditions as for the ischemic sequence. . . . .	105
3.17	<b>Our deep learning-based tracker is robust to small variations on the initial Region of Interest (ROI) position initialization. (a)</b> The distance between the original and displaced ROIs remains around 6 px for the entire tracking sequence. <b>(b)</b> ROI position for both original and displaced locations exemplified on patient 1 in the perfused <sub>2</sub> image sequence. The distance between original and displaced ROIs remains approximately constant during the complete image sequence. . . . .	107
3.18	<b>A separation of the normalized spectra can be observed between perfused and ischemic tissue. (a)</b> $\ell_2$ -normalized median diffuse reflectance spectra within each Region of Interest (ROI), different colors correspond to different tissue state, while different line styles represent the different ROIs. <b>(b)</b> Root-Mean-Square (RMS) contrast for each band and each patient. Across almost all patients, the band 11 is the one with the highest contrast, only one ROI of patient 1 deviates from this tendency. The translucent band around each line plot corresponds to the standard deviation across all annotated spectral images. The vertical dashed lines correspond to the camera band with the highest contrast for the majority of patients. . . . .	109



3.19 **High tissue heterogeneity across subjects motivated our personalized approach.** Both the **(a)** principal component analysis (PCA) and **(b)** the mixed model analysis demonstrate the high inter-patient variability of spectral tissue measurements. **(a)** The fact that different tissue states cluster within a subject but do not form a uniform cluster across subjects motivated us to phrase the challenge of ischemia detection as an out-of-distribution (OoD) problem. Solid markers show the cluster centers, transparent markers the raw data points, and the axis labels denote the explained variance of the corresponding principal component (PC). **(b)** Explained variance for the 16 bands of the multispectral imaging (MSI) camera, depicted in Fig. 2.3 . . . 111

3.20 **We treat ischemia monitoring based on multispectral imaging (MSI) as an out-of-distribution (OoD) task which only needs training data from one single patient.** Traditional deep learning (DL) methods (brown) require large amounts of patient data to train a model, while our method (green) only needs data from a single patient. Using an ensemble of invertible neural networks (INNs) as a core component, our algorithm computes the likelihood of ischemia based on a short (several seconds) image sequence acquired at the beginning of each surgery. An important feature of our approach is that the entire training and inference process can be performed during a surgical procedure. . . . . 112

3.21 **Our approach is capable of discriminating between ischemic and perfused kidney states.** **(a)** Examples of the raw Widely Applicable Information Criterion (WAIC) values for both image sequences, each annotated ROI is identified by the translucent squares. The raw WAIC values are further processed to generate our *ischemia index*. **(b)** We calculated the *ischemia index* for every image in each sequences of perfused and ischemic kidney separately for each patient and generated corresponding dot and box plots. The boxes show the interquartile range with the median (solid line) and mean (dashed line) while the whiskers extend up to 1.5 of the interquartile range. Min-max-normalization was performed for clarity of presentation. 113

3.22 **The Beer-Lambert law cannot monitor ischemia in the kidney.** Although Beer-Lambert regression can differentiate ischemic from perfused tissue for some patients (1,2 and 8), a clear separation is no visible for the rest of the patients. The Beer-Lambert law also generates implausible oxygenation values that cannot be interpreted. . . . . 114

3.23	<b>Red-Green-Blue (RGB) data does not contain the information required to classify kidney perfusion state.</b> Although the tissue state (perfused vs ischemic) can be separated for some patients (patients 2 and 3), this is not possible for most of the patients. The so-called ischemia score, evaluated on RGB data, was computed in the same fashion as the <i>ischemia-score</i> . . . .	114
3.24	<b>Pre-training our ensemble of Invertible Neural Networks (INNs) reduced training time while maintaining high accuracy.</b> The tissue state (perfused vs ischemic) can be clearly separated for most patients (except patient 7). Pre-training the ensemble of INN reduces training time and distinguishes better between tissue states (e.g. patient 9). . . . .	115
4.1	<b>Comparison of a representative patient with patient 7 with respect to the algorithm input.</b> (a) Example Red, Green and Blue (RGB) images were selected from the $\text{perfused}_2$ sequence to illustrate the unusual images acquired of the kidney of patient 7. (b) The reflectances of the representative patient (patient 3) differ clearly between perfused and ischemic tissue, while no clear separation can be observed in patient 7. (c) A kernel density estimation (KDE) performed on the data resulting from a patient-specific principal component analysis (PCA) provides further explanation for why our method falsely detected the ischemic kidney data of patient 7 as inlier. The axes labels denote the explained variance of the corresponding principal component (PC). . . . .	131
4.2	<b>Variability across different normalization strategies is insignificant compared to variability across tissue states (perfused vs ischemic).</b> Reflectance spectra of patient 4 after normalizing the data with reference measurements from patients 1 (white 1) and patient 3 (white 2). . . . .	132
A.1	<b>We annotate ROIs in the surface of tissue and ensure high data quality.</b> Region of Interest (ROI) annotation flowchart based on a sequence of spectral images. . . . .	142
A.2	<b>The data within each Region of Interest (ROI) of our clinical trial had a high level of Signal-to-Noise Ratio (SNR)</b> SNR for all patients of our clinical trial on ischemia monitoring. The signal originating from light-tissue interactions is well above noise level, by a factor of $\sim 82$ on average across patients. . . . .	143

A.3 **Slight spectral differences can be detected with diffuse reflectance measurements performed with a spectrometer.**  $\ell_2$ -normalized reflectance spectra for all 12 patients of our clinical study, corresponding to two exemplary Regions of Interest (ROIs) annotated in perfused and ischemic tissue (160 mmHg). The solid lines correspond to the median reflectance spectra across all 100 measurements, while the translucent bands around each line correspond to the standard deviation. . . . . 144

*List of Figures*

---

## List of Tables

- 1.1 Comparison of different functional imaging modalities: Spectral Imaging (SI), Photoacoustic Imaging (PAI), Optical Coherence Tomography (OCT), Diffuse Correlation Spectroscopy (DCS), Fluorescence Lifetime Imaging (FLIM), Polarization Spectroscopy (POS), Laser Speckle Imaging (LSI), Diffuse Reflectance Spectroscopy (DRS), Spatial Frequency Domain Imaging (SFDI), functional Magnetic Resonance Imaging (fMRI). . . 16
- 1.2 Overview of relevant band selection methods. *Selected bands* represent the optimal number of bands reported by the authors. Our proposed method suggests a subset of ten out of 101 bands for the *in-silico* oxygenation experiments, but can be applied to a wide range of tasks. . . . . 30
- 1.3 Comparison of light source spectrum estimation methods. Most of these methods are designed for Red-Green-Blue (RGB) images of natural scenes, only some of them have been extended to Spectral Imaging (SI) and none of them to the medical domain. . . . . 33

1.4	Comparison of functional imaging studies performed over the last decade with a focus on Spectral Imaging (SI). The following imaging modalities are compared Spectral Imaging (SI), Red-Green-Blue (RGB), Photoacoustic Imaging (PAI), Diffuse Reflectance Spectroscopy (DRS), Spatial Frequency Domain Imaging (SFDI), functional Magnetic Resonance Imaging (fMRI), Laser Speckle Imaging (LSI), Polarization Spectroscopy (POS), Polarization Spectroscopy (POS) and Optical Coherence Tomography (OCT). Notice that the population size (biological replicates) was not reported for some studies, and that “human” studies refer to studies performed in humans but where no evidence of a structured clinical trial was found. Clinical studies are highlighted in gray and the methodology proposed in this thesis in this thesis in green. <i>BVF</i> refers to blood volume fraction and <i>sO<sub>2</sub></i> to oxygenation.	37
1.5	Comparison of Region of Interest (ROI) tracking methods. Most ROI tracking algorithms are implemented for natural scenes, which feature a strong contrast between the target and background. . . . .	43
2.1	The simulated ranges of physiological parameters, and their usage in the simulation setup. The properties of each layer of the three-layered tissue models were assigned by uniformly sampling within the ranges specified in this table. Here $v_{\text{HbT}}$ represents the blood volume fraction, $s$ the blood oxygen saturation, $a_{\text{mie}}$ the reduced scattering coefficient at 500 nm, $b_{\text{mie}}$ the scattering power, $g$ the tissue anisotropy, $n$ the refractive index, $d$ the tissue thickness and $\epsilon$ the extinction coefficients of oxyhemoglobin (HbO <sub>2</sub> ) and deoxyhemoglobin (HbO). Notice that all parameters were uniformly sampled within the specified range for each layer independently, which leads to different values across layers. . . . .	56
3.1	Patients recruited in our clinical study. <b>ccRCC</b> : Clear Cell Renal Cell Carcinoma, <b>PRCC</b> : Papillary Renal Cell Carcinoma, <b>ChRCC</b> : Chromophobe Renal Cell Carcinoma . . . . .	102
3.2	Patient comorbidities and data splits for the participants in our clinical study.	103
4.1	Key characteristics of traditional ischemia monitoring with Indocyanine Green (ICG) injection and our proposed non-invasive method. . . . .	129

## Introduction

*The scariest moment is always just before you start.*

— STEPHEN KING

On Writing: A Memoir of the Craft

### 1.1 Motivation

With the increase in population and the discovery of new treatments, the number of medical interventions is increasing worldwide. According to the statistics reported by Eurostat (2022), the median number of surgical procedures performed in Europe per hundred thousand inhabitants (PHTH) has increased by  $\sim 70\%$  between 2009 and 2019. This, according to the World Health Organization (WHO) (WHO, 2022), is accompanied by only a  $\sim 16\%$  increase in the median number of medical doctors in Europe between 2009 (315 doctors PHTH) and 2019 (367 doctors PHTH). This, together with the millions of examinations performed by medical imaging techniques, stresses the need for computer assistance in the medical field.

Imaging technologies are one of the most common tools used in many medical applications to assist and improve clinical decision-making (e.g. oncology, ophthalmology, neurosurgery, etc.). Some of the most widely adopted imaging modalities are Computed Tomography (CT), Magnetic Resonance Imaging (MRI), Ultrasound and Positron Emission Tomography (PET). Despite the success that all those imaging modalities have demonstrated, and the great benefit that their usage brings to patients, their main intended usage is the analysis of tissue morphology. This, unfortunately, ignores a significant amount of physiological information that could be leveraged by medical staff. Functional imaging (Crosson et al., 2010; Zhu et al., 2022; Bozovic et al., 2022; Deffieux et al., 2021), in contrast, focuses on detecting and measuring changes in metabolism, blood flow and chemical composition of biological tissue. This technique has the potential to provide medical staff with rich information that can aid decision-making in applications where functional tissue properties are the main subject of interest, e.g.: organ transplant (perfusion), anesthesiology (water content), partial nephrectomy (perfusion), etc.

Perfusion monitoring is an area of functional imaging that has been extensively explored due to its relevance in medical practice. Assessment of perfusion is a vital part of most medical procedures that involve stopping the blood supply to an organ or part of a

tissue structure. To name a few, perfusion monitoring is linked to medical applications such as kidney partial nephrectomy, myocardial ischemia, brain stroke, brain trauma, spreading depolarizations, mesenteric ischemia, etc. Current clinical practice relies on injection of Indocyanine Green (ICG) into the patient's bloodstream to assess perfusion of different organs or tissue structures. For several decades, ICG has been deemed safe as a contrast agent to be used in various surgical applications such as minimally invasive surgery, although there have been reports of patients undergoing severe complications after ICG injection, such as anaphylactic shock (Chu et al., 2017). Furthermore, ICG suffers from long washout periods because the human body required about 30 min to dispose of the injected ICG. This means that perfusion assessment via ICG can only be done once per surgery. Even more, open surgical procedures might not benefit from ICG injection at all. For example, during open partial nephrectomy, the injection of ICG might be avoided altogether to reduce patient burden.

Within the wide range of functional imaging modalities, Spectral Imaging (SI) plays a special role, as it does not require a contrast agent and can offer direct visual feedback to the surgeon (Red-Green-Blue (RGB) images). The amount of light reflected from biological tissue, which is collected by SI devices, depends on the intrinsic molecular and structural composition of each tissue type, as such, SI records data that is directly linked to functional tissue properties, such as blood content and oxygen concentration. Over the past years, our community has seen several studies in preclinical or proof of concept studies leveraging SI data (Wirkert et al., 2017; Holmer et al., 2018; Dietrich et al., 2021b; Köhler et al., 2019; Köhler et al., 2020). All of these contributions aimed to analyze a functional property in the context of different medical applications (perfusion monitoring, sepsis classification, etc.). Most of these previous approaches have some common characteristics, mainly: **a)** they employ SI devices that are too big or suffer from slow recording and processing times (Holmer et al., 2018; Dietrich et al., 2021b; Köhler et al., 2019; Kumar et al., 2020), **b)** all of them ignore the high complexity that is characteristic of open surgical scenarios in comparison to minimally invasive surgery, **c)** they overlook inter-patient tissue heterogeneity, and **d)** they generally lack validation in clinical trials within the context of surgery.

The usage of big and slow devices could be attributed to the lack of commercially available SI devices that are medically approved. Hence, this limitation leads to the emergence of several in-house devices and methods that might not fulfill clinical translational requirements. For example, most approaches proposed until now ignore challenges arising during open surgical procedures. One of such challenges pertains to light source spectrum estimation. Minimally invasive surgery benefits from a highly controlled environment where the illumination is generally well known. In contrast, open surgical procedures face a highly dynamic environment, where the surgical site can be illuminated by a wide



range of light sources, e.g. ceiling lights, overhead lamps, forehead torches, etc. Given that optical functional imaging techniques rely on accurate knowledge of the light source spectrum, the ambiguity introduced in open procedures hinders their translation to clinics. The lack of medically approved SI devices is also related to the lack of spectral data, which accentuates the lack of clinical validation of many approaches, and drives the need for synthetic data and more sophisticated data analysis. Within this context, all previous approaches have overlooked a critical aspect of translational research. That is, inter-patient tissue heterogeneity. Since most approaches are only validated in studies involving a few patients, tissue differences across participants are generally not addressed. In turn, this hinders the generalizability potential of such approaches.

In this thesis, I aim to fill the gaps in knowledge that I have outlined here. Mainly, I address the issues of slow recording and processing times in SI, I propose a method for automatic light source estimation for open surgery, and explicitly address the issue of high inter-patient tissue heterogeneity with a personalized approach to video-rate and label-free perfusion monitoring. I demonstrate the benefits of these methods in several *in silico*, pre-clinical and clinical trials in a wide range of relevant domains: rat head and neck tumor models, pig brain, human skin and kidney. In the following section, I detail the main objectives and contributions of this thesis, as well as outline its backbone.

## 1.2 Background

This chapter elaborates on the knowledge fundamentals that are relevant for the development of this thesis. More in detail, Sec. 1.2.1 describes the different anatomies and molecular compositions of the different tissue structures that are the subject of study in this thesis. Sec. 1.2.2 describes different biomedical imaging modalities relevant to the topic of functional imaging. Finally, machine and deep learning fundamentals relevant for image analysis throughout this thesis are described in Sec. 1.2.3.

### 1.2.1 Medical fundamentals

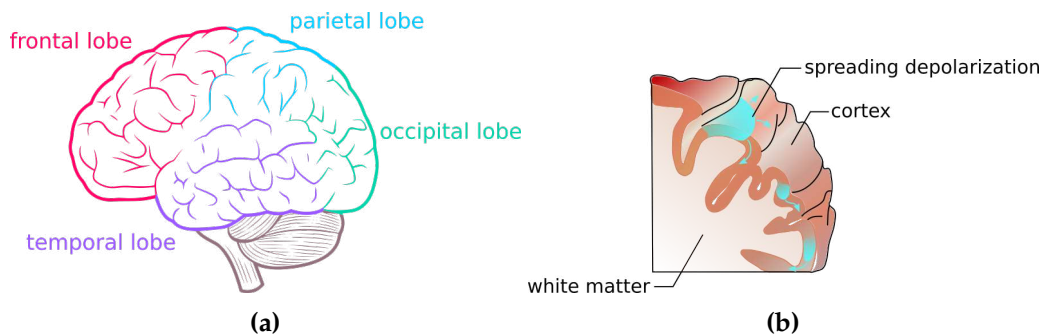
Functional imaging occupies a relevant position in biomedical applications. Mainly because its main subject of study are properties that play a crucial role in many diseases that affect a large portion of the population. Its reach expands from ophthalmology, where Optical Coherence Tomography (OCT) can help detect vessel blockage (Rebolleda et al., 2015), to SI, where diffuse reflectance measurements aid in sepsis classification (Dietrich et al., 2021b). In particular, perfusion assessment has become substantially relevant because it is related to many clinical diseases and conditions, such as myocardial ischemia (Bamberg et al., 2014), Spreading Depolarizations (SDs) (Santos et al., 2014),

and renal hypoxia (Legrand et al., 2008). Tissue perfusion and oxygenation are essential for all metabolic processes in cells and are the primary factors involved in tissue repair (Gottrup, 1994). Poor perfusion (ischemia), leads to medical complications like the ones mentioned previously. Given that the circulatory system is the main responsible for blood and oxygen delivery to different tissue structures in the human body, it becomes relevant to understand the physiology of such tissue structures to understand how perfusion is affected by different factors. In this section, three tissue structures are described (brain, skin and kidney), which cement the background for the medical applications that are presented in this thesis.

### **Brain anatomy and function**

Proper brain perfusion is essential for the survival of any species, including humans. Many brain diseases and conditions are directly tied to changes in perfusion, for example: SDs, stroke, Alzheimer's disease, etc. In particular, a method for monitoring SDs *in vivo* and at video-rate is proposed in this thesis. Therefore, the physiological background relevant for this application is presented in this section.

The brain is considered one of the most complex organs in the human body. It is the center of the nervous system in all vertebrate and almost all invertebrate animals. Most of the complexity associated to the brain does not arise from its physiology, but from the cognitive function that it is responsible for. Humanity has developed many great tools and methodologies to study the function of different organs, however, the study of brain function has remained a challenging task until now. The brain cerebrum is divided in two hemispheres (right and left) which account for  $\sim 85\%$  of the brain mass. Each hemisphere comprises two main components: gray and white matter. The outer part of the cerebrum (cortex) is made up from gray matter, which in itself is mainly composed of billions of neurons and unmyelinated axons arranged in six layers across the section of the cortex. Even though the cortex is only two to four millimeters thick, it accounts for approximately 40% of the brain mass. The inner part of the cerebrum is composed of white matter, which is mainly made from myelinated long axons. The cerebrum can also be divided by function into four lobes: frontal, parietal, occipital and temporal (Fig. 1.1 a)). The frontal lobe is located beneath the frontal bone, and it contains the primary motor cortex and is involved in complex conscious thinking. The parietal lobe is a region of the cerebrum that is located beneath the parietal bone, and it contains the primary sensory cortex and is involved in language acquisition. The occipital lobe is located beneath the occipital bone, and it processes visual information and is related to the understanding of written words. Finally, the temporal lobe is the region of the cerebrum underneath the temporal bone, and it is associated with regions of the brain mainly processing information regarding



**Figure 1.1: Physiological structure of the brain.** (a) The human brain can be divided into four lobes: frontal, parietal, occipital and temporal, depending on one of their main functions: reasoning, touch sensory, visual and auditory respectively. (b) Spreading Depolarizations (SDs) are a phenomenon occurring on the brain cortex that causes massive hyperactivity of neurons followed by inhibition. This phenomenon manifests itself as a wave of depolarized neurons expanding through the cortex.

hearing and equilibrium.

The brain cortex is covered by the *pia mater*, a delicate fibrous layer that allows blood vessels to pass through and nourish the brain. The *arachnoid mater* rests on top of the *pia* and houses cerebrospinal fluid. The *dura mater* is a thick membrane made of dense connective tissue that surrounds the *arachnoid* and spinal cord. It is the outermost layer of the meninges that protects the central nervous system.

There are many medical diseases and conditions associated with the nervous system, and particularly the brain, such as SDs (Santos et al., 2014), Alzheimer's disease, and dementia. SD is a phenomenon occurring in the brain cortex of many species (mammals, reptiles, etc.) that manifests itself as a wave of neuronal hyperactivity followed by inhibition (Dreier et al., 2018; Santos et al., 2014; Ayata and Lauritzen, 2015). The wave of neuronal hyperactivity leads to an accelerated rate of oxygen consumption, thus manifesting itself as a wave of deoxygenated tissue that expands across the cortex (Fig. 1.1 b)). SDs cause arterial dilation accompanied by a red shift (i.e. arterialization) in the color of venous blood, thus indicating changes in oxygenation and perfusion. Although SDs propagate at slow speeds (a few millimeters per minute), they can cause neuronal death, thus leading to deadly consequences. Furthermore, they can worsen existing injuries and detriment patient recovery. The spread of SDs in gray matter does not depend on action potentials, physiological synaptic transmission or long-range axonal connections (Ayata and Lauritzen, 2015). They are, however, dependent on transmembrane ionic and water shifts that locally flood neighboring tissue with depolarizing ions that trigger neuron depolarization in adjacent cells. SDs are triggered when a sufficient stimulus

depolarizes a minimum volume of cortex tissue ( $\sim 1 \text{ mm}^3$ ) in rodents. The stimulus floods the extracellular clearance mechanisms with potassium ions ( $K^+$ ), causing it to surpass a critical threshold concentration of  $\sim 12 \text{ mM}$  (Ayata and Lauritzen, 2015). SDs are generally detected and monitored with Electrocorticography (ECoG) strips, recently more modern imaging modalities have been proposed to monitor the spread of SDs (Santos et al., 2013, 2014). However, these modalities focus on the morphology and spread of SDs, while leaving the analysis of functional properties unaddressed (e.g. oxygenation). The development of a method capable of monitoring functional properties could help the development of new treatments to mitigate the effects of SDs or to prevent them from happening in the first place.

### Skin anatomy and function

The study of skin functional properties, such as perfusion or water content, is relevant for many diseases and conditions where they play a major role: wound healing, psoriasis, sepsis, etc. Therefore, the development of a method to monitor such functional properties in skin could aid in monitoring and treating such diseases and conditions. In the context of this thesis, a method for perfusion monitoring was developed and validated in a clinical trial of human skin. In this section, the physiological background of skin is presented to put the methodology proposed in this work into context.

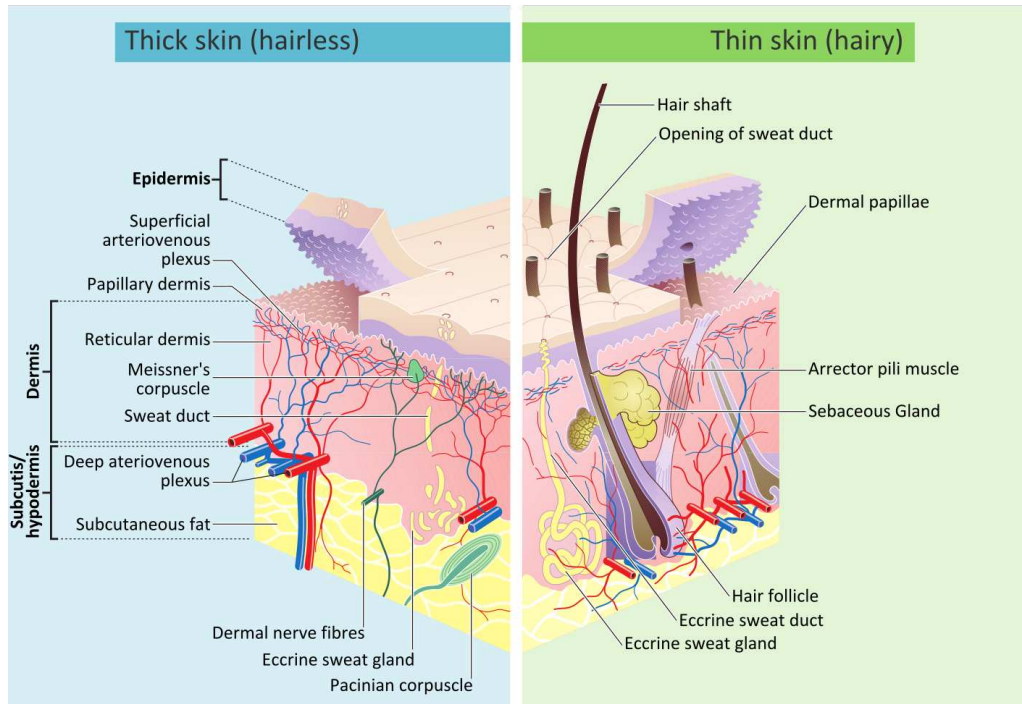
The human skin is the largest organ in the human body and covers it almost in its entirety. The skin is composed of up to seven layers that guard muscles, bones, ligaments, internal organs, etc. Although almost all the skin is covered with hair follicles, it can seem as hairless, there are two main types of skin: hairy and glabrous skin (hairless). Given that the skin is the main organ interfacing with the outer environment, it plays a major immunity role by protecting the body from pathogens and water loss. It also serves as a thermal insulator, sensory interface, vitamin synthesizer, etc.

Human skin can be separated into three main layers, the epidermis (outer layer), the dermis (beneath the epidermis) and the subcutaneous layer (beneath the dermis) (Fig. 1.2). The epidermis is the outermost layer of the skin, which has direct contact with the outer environment. It is made from squamous epithelium (compactly packed cells) that forms a protective barrier against pathogens. Surprisingly, in contrast to the majority of tissue in the human body, the epidermis contains no blood vessels, and cells in this layer are mainly nourished by oxygen diffusion from the surrounding air. The epidermis is mainly composed of Merkel cells, keratinocytes, melanocytes, etc. It can also be subdivided into different regions: *corneum*, *lucidum* (only in palm hands and bottom of feet), *granulosum*, *spinosum* and *basale*. Cells in the epidermis are generated by mitosis, at the *basale* layer, and they slowly move up toward the *corneum* while changing composition and shape.

Eventually, the cells die at the *corneum* as they isolate from the blood source. As cells in the epidermis die, the cytoplasm is released and keratin is inserted as they move towards the top of the skin, where they peel off. Interestingly, the *lucidum* is composed of about five layers of dead and flattened keratinocytes, and it is found only in regions of thick skin (hand palms and bottom of the feet). Under the microscope, the *lucidum* appears translucent, hence the naming.

The dermis is a layer of skin that lies between the epidermis, and contains connective tissue that acts as a stress and strain dampener. It contains many nerve endings that give the sense of touch and heat as well as other structures: hair follicles, sweat glands, blood vessels, etc. The dermis can also be subdivided into two main structural regions: papillary and reticular. The papillary region is composed mainly of loose connective tissue and is connected to the epidermis. On the other hand, the reticular region is composed of dense irregular connective tissue and is usually much thicker than the papillary. It also contains relatively high concentrations of collagenous, elastic and reticular fibers, as well as the roots of the hair and blood vessels. Blood supply to the skin is achieved by two plexuses, one of them is located at the bottom of the reticular region while the other is located in the papillary region. The later one regulates body temperature by increasing or reducing blood flow. Finally, the subcutaneous layer lies below the dermis and is composed mainly of loose connective and adipose tissue. Its main function is the attachment of the dermis to the underlying bone and muscles, as well as blood vessels and nerves. In addition, approximately 50 % of body fat is contained in this layer, and acts as padding and thermal insulation for the body.

The color of the skin is regulated by pigments like melanin, carotene and hemoglobin; melanin being the main one. They can be found in different concentrations and in different locations of the skin. The main purpose of skin pigmentation is the regulation of the amount of ultraviolet radiation that penetrates the skin. Exposure to high amount of radiation can lead to different types of skin cancer such as basal cell carcinoma and melanoma. All of these types of cancer usually feature an increased level of vasculature and blood supply. In particular, melanoma is a type of cancer that originates from melanocytes and is mainly caused by ultraviolet light exposure. The most effective treatment for melanoma is surgical removal if it has not yet spread to other organs. Otherwise, it may require immunotherapy, radiotherapy or chemotherapy, which is why early detection and prevention is so important to improve patient care. Another disease that heavily involves the skin is sepsis. Sepsis is a life-threatening disease that arises when the body overreacts to an infection, causing damage to its own tissue and organs (Hajj et al., 2018; Dietrich et al., 2021b). This overreaction of the body is commonly manifested in the skin by inflammation and is affected by changes in the circulatory system (heart rate and respiration). As such, being able to monitor functional properties in the skin



**Figure 1.2: The human skin is the largest organ and is structured in a multi-layered fashion.** The human skin is composed of three main layers: epidermis, dermis and subcutaneous layer (*subcutis*). The thickness and precise composition of each varies across the surface of the human body. However, several components can be generally found through the body: hair follicles, nerves, blood vessels, fat, sweat glands, etc. *Image from Madhero88 and M.Komorniczak, CC BY-SA 3.0, via Wikimedia Commons.*

(water content, perfusion, etc.) could lead to early detection of sepsis and better patient treatment.

### Kidney anatomy and function

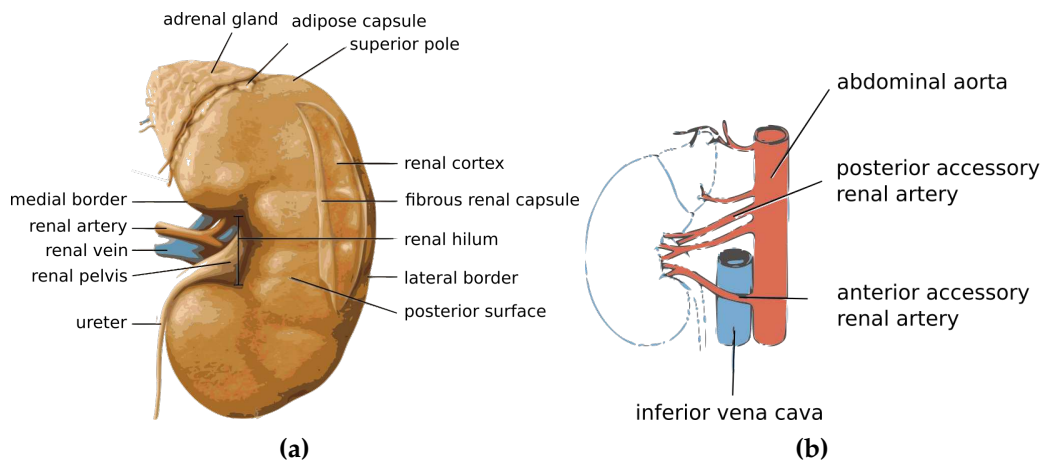
Many surgical procedures require accurate perfusion assessment of an organ, or part of it. Among such procedures are partial nephrectomy, esophagectomy, colorectal resection, hepatectomy, etc. When the blood supply to any biological tissue structure is restricted (ischemia), a shortage of oxygen needed for cellular metabolism is exacerbated over time. Prolonged ischemia can have severe consequences on patients and may even be a cause of death. In this thesis, an approach for label-free ischemia monitoring is proposed and validated in the context of partial nephrectomy. Hence, this section elaborates on the medical background of kidney anatomy, disease and treatment.

The kidneys are organs in the shape of beans that can be found behind the peritoneum. The size of the organ depends on the species, in human adults that is approximately 12 cm. The main function of kidneys is the filtration and removal of wastes, extra water and toxins from the bloodstream in the form of urine. The urine is transported from the kidneys via two thin muscular tubes called ureters, one on each side of the bladder. The kidneys play a crucial role in maintaining a balanced concentration of water, salts and minerals in the bloodstream. Furthermore, kidneys also make hormones that help control blood pressure, make red blood cells.

The functional part of each kidney can be divided into two structures: the renal cortex (Fig. 1.3 a)) and the inner renal medulla. The later one is composed of several cone-shaped renal lobes separated by renal columns. Blood filtering takes place in thousands of structures called Nephrones, which span the medulla and the cortex. The filtering process starts in a portion of the Nephrones located in the renal cortex and called renal corpuscle. A renal tubule connects the cortex with the medullary pyramids (on each renal lobe). The tip (papilla) of each pyramid drains the urine into the renal calyces, which in turn empty into the renal pelvis (Fig. 1.3 a)) and later into the ureter. At the renal hilum, the ureter and renal vein exit the kidney, while the renal artery enters. Generally, fatty tissue surrounds these structures and most of the kidney (adipose capsule) (Manski, 2020).

Each kidney's blood supply emanates from the renal arteries, which branch from the abdominal aorta. Even though the kidneys are relatively small, they receive approximately 20 % of the cardiac output. The renal artery branches into smaller segmental arteries, which also branch into interlobar arteries that penetrate the renal capsule and extend between the renal pyramids. After filtration, the blood exits from the kidney through the renal vein. The supply of blood to the kidneys through the renal artery can have several variants, depending on each person. There might be more than one renal artery that supply each kidney (Listmann and Tubbs, 2020) (Fig. 1.3 b)). Supernumerary renal arteries (two or more) are the most common endovascular anomaly, amounting to 25 % to 40 % of kidneys (Coello-Torà et al., 2020).

Kidneys can suffer from many diseases and conditions such as renal tumors, hydronephrosis, and adrenal tumor. Kidney tumors can be classified in many categories depending on if they are malignant or benign as well as the stage of each. Among the malignant tumors are papillary renal cell carcinoma (PRCC), chromophobe renal cell carcinoma (ChRCC), oncocytoma, clear cell renal cell carcinoma (CCRCC). Many of these tumors are treated by performing partial or total nephrectomy, that is, the partial or complete removal of the kidney. Generally, The RENAL Nephrometry Scoring System (Liu et al., 2020a) can be used to determine the complexity of kidney tumors and decide if partial or radical nephrectomy is more appropriate. The nephrometry score considers the size of the tumor (Radius), how much of the tumor is inside or outside the kidney (Endophyt-

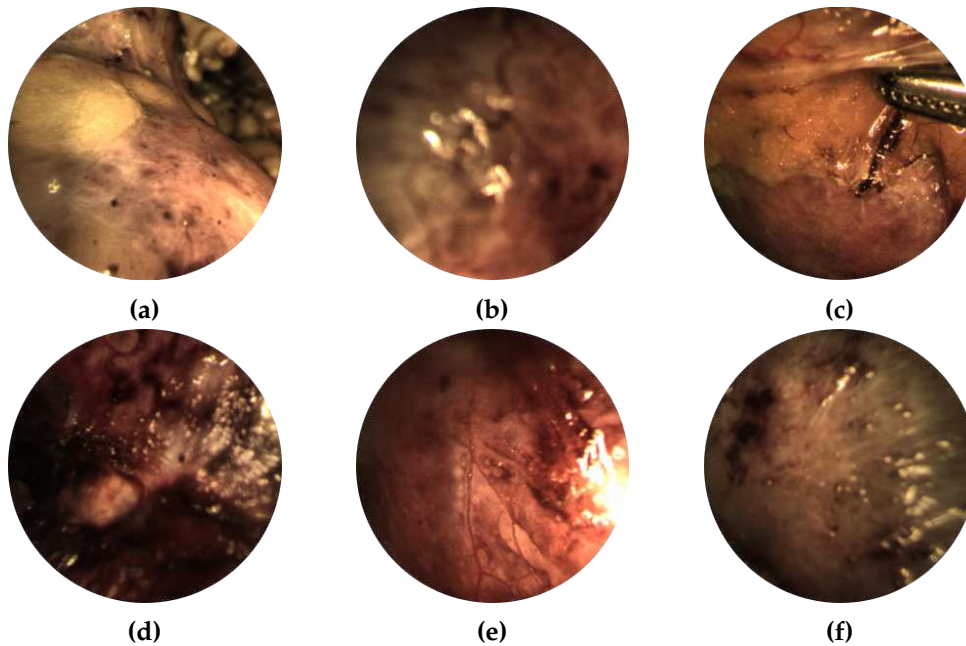


**Figure 1.3: Human kidney is perfused through renal artery emanating from the abdominal aorta. (a)** General anatomy of the human kidney. The kidney is covered by the fibrous renal capsule that is composed mainly of fibrous tissue. The adrenal gland sits on top of the kidney and produce a variety of hormones such as adrenaline and cortisol. **(b)** Kidney perfusion through accessory renal arteries branching from the abdominal aorta. Before reaching the renal hilum, arteries usually branch out. One or two accessory arteries are frequently found in humans, while other variations of arterial structures can also be found.

ic/Exophytic), how close the tumor is to the urinary collecting system (Nearness), whether the mass is on the anterior or posterior surface of the kidney (Anterior), location relative to polar lines (Lines), and whether it is touching the renal artery or vein. In particular, during partial nephrectomy, the blood supply to the kidney is temporarily suppressed by clamping the renal artery before the tumorous region of the kidney is removed. The assessment of kidney perfusion thus becomes crucial to prevent major bleeding during the operation. Unfortunately, assessing perfusion by simple visual inspection has proven to be a challenging task. To illustrate this point, Fig. 1.4 present images of kidneys under different perfusion conditions. A simple visual inspection of these images reveals that it is very difficult to assess the perfusion state of the kidney based only on RGB images. In fact, images in a), c) and e) correspond to perfused kidneys, while images in b), d) and f) represent ischemic kidneys.

Current clinical practice to assess kidney perfusion, thus preventing major bleeding, is performed by ICG injection into the patient's bloodstream (Kaplan-Marans et al., 2019). ICG is a fluorescent molecule that binds to plasma proteins, of which albumin is the main carrier, and accumulates in internal organs (mainly kidney and liver). As such, a fluorescent signal indicates a perfused kidney, while no signal indicates an ischemic





**Figure 1.4: Assessing tissue perfusion from visual inspection of Red-Green-Blue (RGB) images is challenging.** Exemplary laparoscopic images of the kidney surface from different patients under different perfusion conditions. Figures (a), (c) and (e) correspond to perfused kidneys, while figures (b), (d) and (f) represent ischemic kidneys. The induction of ischemia was performed by applying a clamp to the renal artery.

kidney. While this method has been the gold standard for many years, it requires the injection of a contrast agent and can only be performed once per surgery because the human body needs up to 30 minutes to dispose of it. Moreover, deathly complications such as anaphylactic shock has been reported to follow from ICG injection (Chu et al., 2017). Therefore, the development of new imaging techniques for perfusion monitoring that does not require the injection of contrast agents and that can be performed multiple times would improve treatment and patient care.

## 1.2.2 Imaging fundamentals

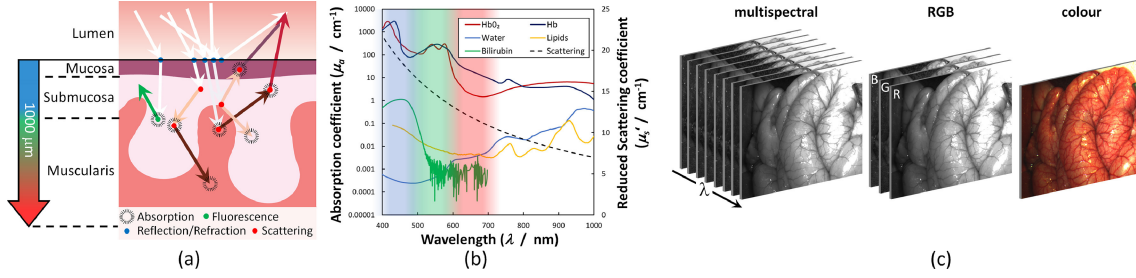
Functional imaging is a technique used mainly in the biomedical domain to monitor and measure changes in metabolism, blood flow, chemical composition, etc. Over the years, it has become the subjects of study of different imaging modalities, where most of them are based on the interaction of light with biological tissue. Among these modalities are: SI, Photoacoustic Imaging (PAI), OCT, Diffuse Correlation Spectroscopy (DCS), Fluorescence Lifetime Imaging (FLIM), Polarization Spectroscopy (POS), Laser Speckle Imaging (LSI), Diffuse Reflectance Spectroscopy (DRS), Spatial Frequency Domain Imaging (SFDI), functional Magnetic Resonance Imaging (fMRI). All of these modalities, except for fMRI, study the interaction of light (typically in the visible or near infrared region) with biological tissue. Therefore, the fundamentals of light-tissue interactions are first presented in Sec. 1.2.2, and fundamentals of functional imaging modalities (with a focus on SI) are presented in Sec. 1.2.2.

### Light-tissue interaction

Among the wide range of imaging modalities for functional imaging, optical imaging is particularly interesting. Optical imaging deals with light-tissue interactions to derive the underlying functional properties, mainly blood flow, blood concentration or oxygenation. The type and strength of interactions between light and biological tissue are governed by the intrinsic optical properties of the tissue (scattering, anisotropy, absorption, etc. Fig. 1.5 a)). At the same time, optical properties of the tissue are defined by its molecular composition, e.g. high concentration of blood leads to large absorption in the visible spectrum ( $400\text{nm} \lesssim \lambda \lesssim 700\text{nm}$ ). During surgery, the main tissue structure of interest are organs. Although the precise molecular composition of different organs can vary substantially, a general trend can be observed. In the human body, molecules like hemoglobin, melanin, water, bilirubin or lipids are commonly found, and among them, hemoglobin features the strongest interaction with light in the visible range (the highest absorption, Fig. 1.5 b)). These molecules are commonly known as *chromophores*.

Light-tissue interactions can be described in a mathematical form using the diffusion theory. However, an analytical solution for the photon diffusion equation is difficult to find for complex tissue and light source geometries. To alleviate these challenges, the interactions between light and tissue can be approximated using a Monte Carlo (MC) method that leverages random sampling.

The light-tissue interactions can be modeled as a sequence of scattering events where the photons are scattered at different tissue locations. After each scattering/event, photons lose part of their energy and travel a certain distance (step size) before a new event may occur. Considering a random variable  $\chi$ . In this case, it represents the step size of a photon



**Figure 1.5: Hemoglobin is the chromophore with the highest absorption in the visible range for internal organs.** (a) Light-tissue interactions in a tissue structure with characteristic layers (e.g. colon). Light is mainly reflected, scattered or absorbed, while luminescence can also occur. Light in the red optical region can penetrate the tissue deeper than light in the blue region. (b) Oxyhemoglobin and deoxyhemoglobin, have similar absorption coefficients while other chromophores like water and lipids absorb much less light in the visible and near infrared (NIR) wavelength range. (c) Spectral images are composed of several bands with narrow light transmission profiles, while Red-Green-Blue (RGB) images are only composed of three bands with wide transmission profiles. Combining the three bands of an RGB image generates a color image that is generally used by medical staff for various procedures. *Image taken from Clancy et al. (2020)*

when it propagates inside a tissue sample, or the scattering angle of photons. A probability distribution  $p : \mathbb{R} \mapsto \mathbb{R}^+$  with support  $[a, b]$  fulfills:

$$\int_a^b p(\chi) d\chi = 1 \quad (1.1)$$

The propagation of photons is simulated by sampling  $\chi$  within the range  $[a, b]$  multiple times to approximate the probability density function. Given a computer-generated pseudo-random variable  $\xi$  that is uniformly distributed in the interval  $(0, 1)$ , the cumulative distribution function  $F : \mathbb{R} \mapsto [0, 1]$  of this uniformly distributed variable is:

$$F_{\xi}(x) = \begin{cases} 0; & \text{if } x \leq 0 \\ x; & \text{if } 0 < x \leq 1 \\ 1; & \text{if } x > 1 \end{cases} \quad (1.2)$$

To sample a non-uniformly distributed function, the assumption is taken that there is a non-decreasing function  $f$  that maps  $\xi \mapsto \chi$  for  $\xi \in (0, 1)$  and  $\chi \in (a, b)$ . That is,  $\chi = f(\xi)$ , therefore both variables have a bijective mapping, which leads to the following (Wang

and Jacques, 1992a):

$$\int_a^{\chi_1} p(\chi) d\chi = \xi_1 \quad \text{s.t.} \quad \xi_1 \in (0, 1); \chi_1 = f(\xi_1) \quad (1.3)$$

Based on this, photon step size and scattering angles can be computed in the following manner.

**Photon step size:** The probability distribution of the photon's free path is sampled  $s \in [0, \infty)$ . The probability of interaction per unit path length in the interval  $(s', s' + ds')$  is:

$$\mu_t = \frac{-dP\{s \geq s'\}}{P\{s \geq s'\} ds'} \quad (1.4)$$

where  $\mu_t = \mu_a + \mu_s$  is the total scattering coefficient

$$d(\ln(P\{s \geq s'\})) = -\mu_t ds' \quad (1.5)$$

This is then integrated in the range  $(0, s_1)$  which leads to an exponential distribution, where  $P\{s \geq 0\} = 1$  is used:

$$P\{s \geq s_1\} = \exp(-\mu_t s_1) \quad (1.6)$$

$$p(s_1) = \frac{dP\{s < s_1\}}{ds_1} = \mu_t \exp(-\mu_t s_1) \quad (1.7)$$

Which can be integrated:

$$\int_0^{s_1} p(s) ds = \xi \quad (1.8)$$

This can then be used to choose a step size:

$$s_1 = \frac{-\ln(\xi)}{\mu_t} \quad (1.9)$$

**Photon scattering angle:** The probability distribution for the cosine of the deflection angle during scattering events is described by the scattering function that Greenstein proposed in 1994:

$$p(\cos \theta) = \frac{1 - g^2}{2(1 + g^2 - 2g \cos \theta)^{3/2}}, \quad (1.10)$$

where  $\theta \in [0, \pi)$  is the scattered angle,  $g = \langle \cos \theta \rangle$  is the anisotropy of the tissue and has a value between -1 and 1.  $g = 0$  means isotropic scattering and  $g = 1$  means forward

scattering. The choice of  $\cos \theta$  can now be interpreted using a random number  $\xi$ :

$$\int_0^{\cos \theta} p(\alpha) d\alpha = \xi \quad \text{s.t.} \quad \xi \in [0, 1] \quad (1.11)$$

$$\cos \theta = \begin{cases} \frac{1}{2g} \left\{ 1 + g^2 - \left[ \frac{1-g^2}{1-g+2g\xi} \right]^2 \right\}; & \text{if } g > 0 \\ 2\xi - 1; & \text{if } g = 0 \end{cases} \quad (1.12)$$

It is worth mentioning that these equations serve as a description of the majority of light-tissue interaction, however, many other interactions are not covered here and depend on the particular application.

### Functional imaging modalities

Imaging modalities are generally developed to meet a wide range of different requirements for different applications. A comparison of some of the most relevant features in the medical domain across imaging modalities is presented in Tab. 1.1. Here, some imaging modalities only feature some characteristics under restrictive conditions or only on some applications (indicated by ✓). Most modalities feature high-resolution sensors, except for DCS (Sutin et al., 2016) and DRS (Kleshnin, 2019; Guo et al., 2022), which are until now point-measurement techniques. On the other hand, fMRI could be considered to have high resolution when imaging large body regions. However, smaller tissue structures usually suffer from low contrast. Moreover, small vessel structures could potentially be visualized via fMRI but only with the use of contrast agents. Recording speed is a characteristic that has particularly eluded the majority of modalities. The imaging modalities that have achieved high recording and processing speed (video-rate) are generally the ones based on sensors that are also employed in photography (Charge-Coupled Device (CCD) and Complementary Metal-Oxide Semiconductor (CMOS)). This could be attributed to the fact that the development of such sensors has benefited from great advances in the last decades, from more sensitive semiconductors to faster read-out techniques. Although medical devices exist for many of these imaging modalities, they suffer from many limitations that hinder their general usability in clinics (except for fMRI). Although fMRI is a widely used preoperative imaging technique, it requires big and expensive equipment that cannot be used to assess functional properties of tissue structures during surgery. Other imaging modalities, such as OCT (Choi et al., 2019; Liu et al., 2019), DCS (Sutin et al., 2016), LSI (Mangraviti et al., 2020) and SFDI (Chen and Durr, 2020) are limited to extracorporeal assessments such as skin perfusion analysis, and devices that can be used during surgery are yet not available. Although there are no medical devices for POS, some proof of concept studies involving a few patients have been published Vasefi et al. (2015). Moreover, OCT

**Table 1.1:** Comparison of different functional imaging modalities: Spectral Imaging (SI), Photoacoustic Imaging (PAI), Optical Coherence Tomography (OCT), Diffuse Correlation Spectroscopy (DCS), Fluorescence Lifetime Imaging (FLIM), Polarization Spectroscopy (POS), Laser Speckle Imaging (LSI), Diffuse Reflectance Spectroscopy (DRS), Spatial Frequency Domain Imaging (SFDI), functional Magnetic Resonance Imaging (fMRI).

Name	High-resolution	High-speed	Wide FOV	Med. device	3D	Intraoperative
SI	✓	✓	✓	✓	✗	✓
PAI	✓	✓	✓	✓	✓	✗
OCT	✓	✗	✓	✓	✓	✓
DCS	✗	✗	✗	✓	✗	✗
FLIM	✓	✗	✓	✗	✓	✗
POS	✓	✓	✓	✗	✗	✓
LSI	✓	✗	✓	✓	✗	✗
DRS	✗	✗	✗	✗	✗	✗
SFDI	✓	✓	✓	✓	✗	✗
fMRI	✓	✗	✓	✓	✓	✗

✓: indicates that the imaging modality meets the corresponding feature (columns).

✓: indicates that the feature is met only in particular applications.

✗: indicates that the imaging modality does not possess the feature.

is a modality that offers a wide range of medically approved devices, although they are limited to the analysis of eyes. Although this technique has the potential to be used during eye surgery, it has not been translated to surgical applications, potentially due to speed requirements and that it is currently limited to big devices.

PAI is a particular imaging modality that has seen many advances recently (Gröhl et al., 2021; Tzoumas et al., 2016; Zhu et al., 2022). This technique can feature high speed and resolution, as well as 3D functional maps of tissue under particular conditions, such as a limited number of recorded spectral regions (bands). Even though medical devices exist that capable of recording photoacoustic images, they are generally big and feature large imaging probes\*. This, paired with the need for a controlled laser light source, limits its applicability in surgical scenarios, which have until now been restricted to a few clinical trials (Gröhl et al., 2021). Only recently, there have been attempts to construct promising small endoscopic PAI probes (Ansari et al., 2020) which could in the future be used for clinical applications. However, these devices are still in the development phase.

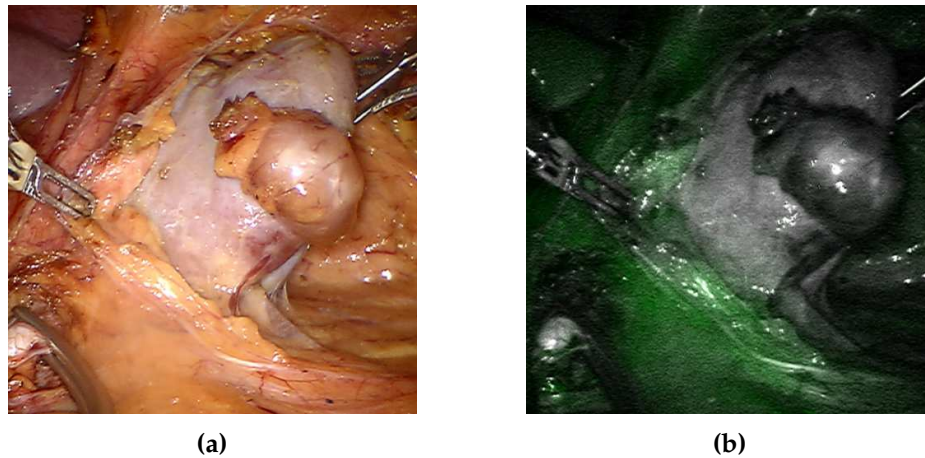
Among these imaging modalities, SI stands out among them, mainly because it leverages fast and high-resolution sensors (both CCD and CMOS) that are commercially available. In addition, medically approved devices have recently become available that are

\*Similar to ultrasound probes.

also capable of being used in surgical applications. Another advantage of SI over other modalities is related to the possibility to record standard RGB images at the same time as spectral data. This reduces the number of devices required during surgery and enhances the surgeon's surgical perception. One of the main disadvantages of SI in comparison to other modalities, such as PAI, is the fact that SI can only image the surface of the tissue. This limitation is introduced by the limited penetration depth of light (usually in the visible range) into biological tissue (a few hundred micrometers). Despite this, SI is one of the imaging modalities that show most promise to be translated into clinical use. For these reasons, this thesis focuses on the translation of SI. Even more, given that perfusion is a functional property with great interest in the medical domain, due to its relevant role in several procedures, and diseases (e.g. organ transplant, myocardial ischemia, brain stroke, etc.), the methods presented in this thesis are validated in perfusion monitoring applications. However, they can be extended to the estimation of other functional properties. Furthermore, since current clinical practice for perfusion assessment of tissue is based on fluorescence imaging, the rest of this section is dedicated to detail fluorescence and spectral imaging in a medical context.

### **Fluorescence imaging**

Fluorescence is a type of interaction between light and matter (e.g. biological tissue), which involves the absorption of light by a molecule and its re-emission at a different wavelength. Certain molecules, called fluorophores, absorb light by changing their internal energy state. This excited molecular state, however, only lasts a few nanoseconds before the molecule attempts to return to its initial state through a series of mechanisms that generate photons at a different wavelength as the incident light (fluorescence). The emission of fluorescence is influenced by many parameters that are mainly related to the environment where the fluorophore resides: polarity, pH, pressure, viscosity, etc. (Dip et al., 2015). Even more, fluorescent molecules have the capacity to bind to specific targets depending on their molecular composition. Some biomolecules contain fluorophores like amino acid tryptophan, however, they are not used because of poor image contrast. In general, external fluorophores are preferred due to their enhanced optical properties: higher brightness, higher quantum efficiency, and enhanced photostability. Among this, organic molecules are preferred due to their small size in comparison to other fluorophores like genetically encoded proteins or nanoparticles (Quantum dots). Ideally, fluorescent molecules should offer a high contrast compared to other intrinsic molecules to facilitate image-guided surgery. For this reason, fluorophores that emit light in the NIR region 650–900 nm are favored. One of the main advantages that working on such optical region offers is that the autofluorescence signal from intrinsic biological molecules is avoided.



**Figure 1.6: The presence of fluorescent signal is an indicator of perfusion, while its absence indicates ischemia.** Images of a kidney from a female patient after clamping the renal artery and intravenous injection of Indocyanine Green (ICG). The Red-Green-Blue (RGB) image in (a) depicts the presence of an abnormality in the surface of the kidney (Leiomyoma). The absence of fluorescence signal on the kidney (regions without green color in (b)) indicates that the kidney is no longer perfused.

In addition, light scattering is reduced in comparison to the visible spectrum. ICG is a fluorophore that fulfills all such requirements, which explains its adoption in many medical applications: renal surgery, cardiac surgery, lymph node biopsy, etc. When injected into patients, ICG binds to plasma and accumulates in internal organs, mainly in liver and kidney. As such, a fluorescent signal is an indicator of perfused organs, while the lack of it refers to ischemic organs. Fig. 1.6 depicts the usage of intravenous ICG during laparoscopic partial nephrectomy to assess the perfusion state of the kidney.

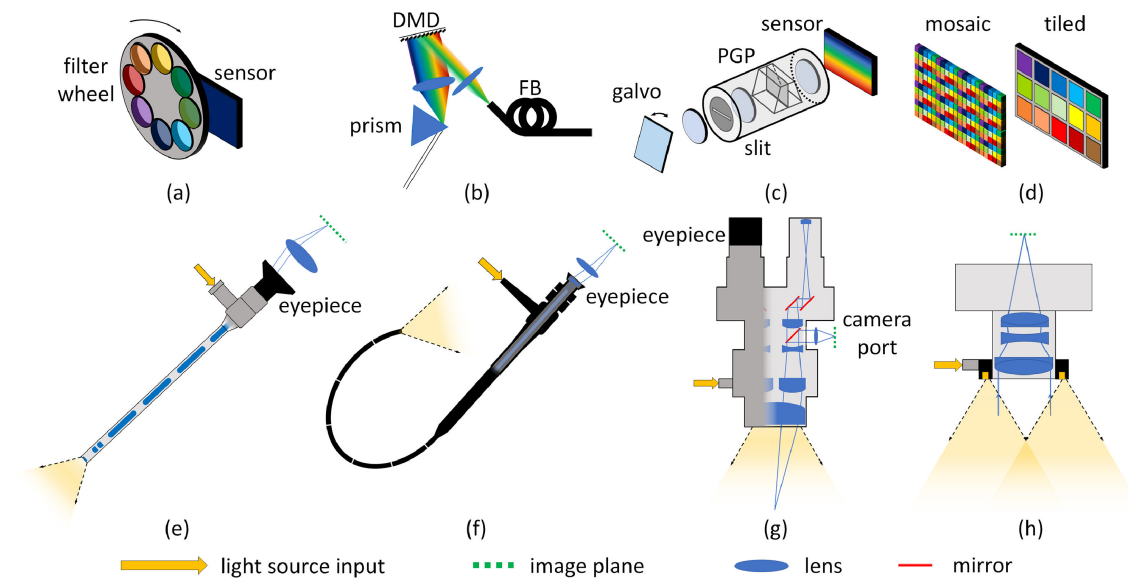
Although ICG has been deemed safe for most patients, there have been reports of deadly complications in some of them, such as anaphylactic shock (Chu et al., 2017). Another drawback of ICG is that it can only be used once per surgery because the human body requires several minutes to dispose of it. Even more, open surgical procedures might not benefit from ICG injection at all. For example, during open partial nephrectomy, the injection of ICG might be avoided altogether to reduce patient burden. Finally, the fluorescent signal emanating from ICG molecules is a relative measure, even low concentrations of ICG could lead to misinterpretation as “perfused” if a gain factor is applied to the signal. Thus, the development of non-invasive and contrast agent-free approaches would remove these limitations, and would also increase usability during open procedures.



## Spectral Imaging (SI)

SI is an imaging modality, which originally designed for aerial imaging, has gained a lot of popularity in the medical domain recently (Clancy et al., 2020; Zhu et al., 2022; Holmer et al., 2018; Dietrich et al., 2021b; Seidlitz et al., 2021). This is mainly due to the ability of SI to acquire images within a wide field of view with high contrast, and to its ability to provide human-interpretable RGB images. SI, in contrast to traditional RGB imaging, records the light reflected from the target object in more and narrower wavelength regions, hereon referred to as “bands” (Fig. 1.5 c)). SI in itself, can be classified in two large categories: a) Multispectral Imaging (MSI), which records images up to tens of bands, and b) Hyperspectral Imaging (HSI), which records images up to hundreds of bands. HSI systems tend to be slower than systems that rely on MSI, mainly because more data needs to be recorded for HSI applications. From here on, the term SI will be used to encompass both MSI and HSI systems, unless particular details of each technique should be highlighted.

SI systems embody a myriad of technical devices, all of them aiming to record multiple spectral bands. Clancy et al. (2020) did an outstanding job at summarizing the existing spectral devices and their features. Among them, rotating filter wheels (Fig. 1.7 a)) have been featured in several applications (Clancy et al., 2015; Wirkert et al., 2016b). Their popularity can be attributed mainly to the fact that they are easy to manufacture and that they feature glass filters that have a narrow transmission profile. One of the main disadvantages of this technique is that the filter wheel introduces mechanical parts that induce motion artifacts during recording. Even more, due to its sequential recording nature, blurring effects can be observed while recording scenes where organ motion is substantial. Another type of technique used in SI involves the usage of tunable light sources (Fig. 1.7 b)), where the scene of interest is illuminated with one wavelength at a time. This can be achieved by using a monochromator or a Liquid Crystal Tunable Filter (LCTF). Although these devices can generate very narrow spectral bands, they can be too big or slow, which hinders their usage in clinical applications. Line scan sensors, in contrast, collect light at all spectral bands simultaneously, but only at specific spatial regions. This is generally achieved by imposing a collimator and an entrance slit between the scene and the sensor, and scanning the spatial scene with a mechanical mirror (Fig. 1.7 c)). Although these devices can be faster than others based on the filter wheels and tunable light sources, they still suffer from blurring effects and could introduce spatially localized bias because different regions of each image are recorded at different time points and may have different focus if the surface of the object of interest is not flat. Finally, one of the promising imaging techniques involves the usage of mosaic sensors (Fig. 1.7 d)). This technique is of particular interest because it can record all spectral bands and spatial



**Figure 1.7: Spectral Imaging (SI) features a wide range of scanning techniques and detection mechanisms.** Scanning can be achieved with a (a) rotating filter wheel, (b) a tunable light source, (c) a spatial line-scan sensor, (d) or a snapshot sensor. These scanning techniques can be coupled with optical parts, depending on the application, such as (e) laparoscopes, (f) flexible endoscopes, (g) microscopes or (h) externally mounted lenses with ring illumination for open surgery. *Image taken from Clancy et al. (2020)*

locations simultaneously. This is achieved by integrating the light filters as a repeating pattern directly into the design of the sensor. In this manner, extremely fast recordings can be achieved while eliminating the risk of motion blur across spectral bands because all of them are recorded at once. This, however, comes at the cost of lower spatial resolution because each different spectral filter occupies a different spatial location within the sensor. Nonetheless, the advances in sensor designs tend to generate smaller pixel sizes, thus increasing spatial resolution with each new camera generation.

Imaging sensors are only half of the story pertaining to SI systems. Such sensors are generally coupled with optical components that aid the image formation process. Depending on the specific application, these optical components can encompass laparoscopes, flexible endoscopes, microscopes or externally mounted lenses (Fig. 1.7 e-h)). On the one hand, laparoscopes and flexible endoscopes are mostly featured in minimally invasive applications such as abdominal surgery or colonoscopy. On the other hand, microscope arrangements and externally mounted lenses are more common for usage in research. Independently of their application, the main objective of such optical components is to both illuminate the scene and to deliver the reflected light to the sensor.

### 1.2.3 Machine Learning fundamentals

The wonder of whether programmable computers could eventually “*think*” on their own has been in the minds of many people for centuries. Even though the word “*think*” carries a lot of philosophical discussion, the question on itself has captivated a lot of interest and has many repercussions in humanity. Nowadays, this is referred to as Artificial Intelligence (AI). Recently, AI has benefited from many successes in the fields of medicine, economics, climate change, etc. More specifically, AI has offered potential improvement in drug discovery and delivery (Jumper et al., 2021; Yang et al., 2022), sepsis classification (Seidlitz et al., 2021), survival prediction (Bello et al., 2019), etc.

AI is an umbrella term that encompasses other fields dedicated to study how programmable computers could “*think*” on their own. In this field, the most complicated challenge resides on how to communicate to a computer the high-level information that we have gathered from the world, e.g. *how do you realize that a friend or partner is angry at you?*. Someone might notice a raised eyebrow, tension on the lips, or just by *gut feeling*. Even more, the way we represent this knowledge in our minds differs from individual to individual. Not surprisingly, hard-coding knowledge into a computer has proven many times to be unsuccessful, and the field of AI has moved on to try to design systems that can acquire such high-level knowledge on its own from raw data, this is now known as Machine Learning (ML). Simple machine learning algorithms, such as logistic regression or naive Bayes, are particularly good at certain tasks like classifying spam e-mail or analyzing market share values. The performance of these type of AI depends considerably on how the data that they exploit is represented. Intuitively, this makes sense, it is much easier to remember that some apples are red than to remember that some apples reflect light mainly in the wavelength region around 650 nm. Logistic regression can correlate data to different possible outcomes (e.g. whether a patient receives treatment A or B). Unfortunately, logistic regression cannot influence how the data is defined in any way. If the data was to be presented to logistic regression in a slightly different way, the performance would likely drop.

Deep Learning (DL) mitigates the representation problem by introducing representations that are expressed in terms of a conglomerate of smaller and simpler representations. *Llamas and alpacas have four legs and fur, but alpacas do not spit*. Examples of DL models are Multilayer Perceptron (MLP), autoencoder, and Convolutional Neural Network (CNN). An MLP is something that translates input values to a different set of output values via the composition of many simple mathematical functions, where each function provides a different representation of the input. One interpretation of DL models is that the representations that are built at different *layers* of the models can represent hierarchical knowledge at different levels, thus making them suitable for learning concepts that are too complex

to describe as a linear set of rules. DL models generally break apart the raw data into a series of nested representations of the same data (edges, colors, background, etc.), and store those representations in different *layers* of the model. The data input that DL models leverage during the learning process is generally known as *input/visible layer* because it is the data that is readily visible to humans. The higher level representation of the input data are stored in *hidden layers*, while the outcome prediction/estimation of a model is stored in the *output layer*. Over the last decade, a myriad of different types of DL models have emerged, many of them designed to perform well on very particular tasks. In this section, the general concept of different types of neural networks is presented together with an ML algorithm called random forests. Both of these types of AI are relevant for the work developed in this thesis.

### Deep Learning regression

There are many neural network architectures with regression capabilities. The implementation of the methods presented in this thesis usually favored accurate and computationally efficient architectures. In this section, the concepts of the recently proposed Invertible Neural Networks (INNs) and the commonly known CNNs are presented. These networks are the basis for the functional parameter estimation methods proposed in this thesis.

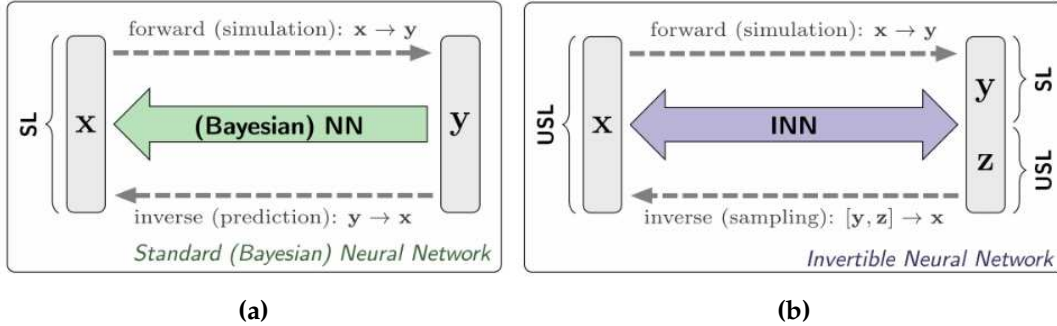
### Invertible Neural Networks (INNs)

The analysis of complex physical environments is usually tied to the restriction that the underlying properties of such systems cannot be directly measured. For example, there is no gold standard to measure blood concentration in biological tissue. To overcome this problem, several models have been proposed to create a relationship that explains how measurable parameters  $\mathbf{y}$  arise from the underlying properties  $\mathbf{x}$ . The mapping  $\mathbf{x} \mapsto \mathbf{y}$  is generally called the *forward process*, while the inference of the underlying properties  $\mathbf{y} \mapsto \mathbf{x}$  is called the *inverse process*. Unfortunately, the *inverse process* is in numerous instances ill-posed, meaning that different measurable parameters  $\{y_i, y_{i+1}, \dots\}$  could map to the same underlying property  $x_i$ . Invertible Neural Networks (INNs) are a category of neural networks that attempt to solve these issues with an architecture that features particular characteristics (Ardizzone et al., 2018):

**Bijjective:** The mapping from inputs to outputs is bijective.

**Efficient:** Both *forward* and *inverse* mappings can be computed efficiently.

**Tractable:** Both *forward* and *inverse* mappings have a tractable Jacobian that allows the explicit computation of posterior probabilities.



**Figure 1.8: Invertible Neural Networks (INNs) can learn the *inverse process* in an ill-posed setting. (a) Standard neural networks require a supervised loss (SL) that causes problems when the mapping between underlying system properties  $x$  and measurable parameters  $y$  is not bijective. (b) INNs use a supervised loss only for the well-defined *forward process*  $x \mapsto y$ , while an unsupervised loss (USL) is used for the *inverse process*, and to ensure that the latent variables  $z$  follow a Gaussian distribution. Figure from Ardizzone et al. (2018).**

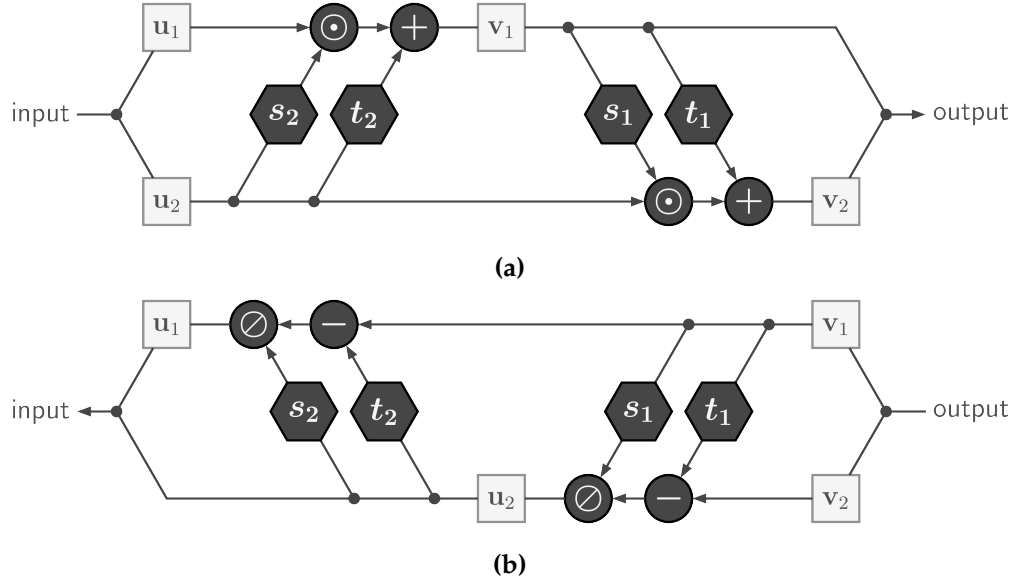
To mitigate the information that is lost during the *forward process*, a latent output variable  $z$  is introduced. This variable  $z$  attempts to capture the information that might be lost during the mapping  $x \mapsto y$ . During training, INNs learn to create a relationship between underlying properties  $x$  and unique pairs of measurable parameters and latent variables  $(y, z) = f(x)$ . This means that INNs also implicitly learn the *inverse process*  $x = f^{-1}(y, z) = g(y, z)$ . During training, it is generally ensured that the density of the latent variable  $z$  is Gaussian distributed. The distribution of  $z$  does not need to be strictly Gaussian, it is, however, required that  $z$  follows a known distribution that can be sampled.

In comparison to standard neural networks (Fig. 1.8 a)), “INNs circumvent a fundamental difficulty of learning inverse problems: Defining a sensible supervised loss for direct posterior learning is problematic since it requires prior knowledge about that posterior’s behavior, constituting a kind of hen-and-egg problem.” –Ardizzone et al. (2018)–.

At the core of INNs are the so-called *affine coupling layers*, which were introduced by Dinh et al. (2016). These layers split their input into two main parts,<sup>†</sup>  $(\mathbf{u}_1, \mathbf{u}_2)$  which are then transformed by functions  $s_i$  and  $t_i$  in an alternating manner (Fig. 1.9 a)).

In the *forward process*, the output is recovered by concatenating the inputs after transformation  $(\mathbf{v}_1, \mathbf{v}_2)$ . The *inverse process* is learned by minor rearrangements, in which case the output is split into two parts and further transformed by the same functions  $s_i$  and  $t_i$  (Fig. 1.9 b)). It is important to notice that the transformation functions  $s_i$  and  $t_i$  do

<sup>†</sup>In Spectral Imaging (SI), the split is done across bands (channels).



**Figure 1.9:** The core of Invertible Neural Networks (INNs) are *affine coupling layers*. **(a)** *Affine coupling layers* split their input into two parts ( $\mathbf{u}_1, \mathbf{u}_2$ ) before being transformed in an alternating manner by functions  $s_i$  and  $t_j$ . **(b)** The mapping from output to input is done in a similar fashion by rearranging of the computational block. Here,  $\odot$  and  $\oslash$  represent element-wise multiplication and division, respectively. Figure from (CVL, 2022).

not necessarily need to be invertible themselves, and can be substituted by any neural network that is trained by backpropagation. The element-wise division introduced in these *affine coupling layers* can lead to numerical problems. This is mitigated by utilizing an exponential function and clipping extreme values after the transformation  $s_i(\cdot)$  (Eq. 1.13). The construction of an INN is then completed by chaining several of these *affine coupling layers*. During INN construction, it is common practice to introduce random channel permutations on each coupling layer to avoid that the first split  $\mathbf{u}_1$  always operates on the second one  $\mathbf{u}_2$  across layers.

$$\begin{aligned}
 \mathbf{v}_1 &= \mathbf{u}_1 \odot \exp(s_2(\mathbf{u}_2)) + t_2(\mathbf{u}_2) \\
 \mathbf{v}_2 &= \mathbf{u}_2 \odot \exp(s_1(\mathbf{v}_1)) + t_1(\mathbf{v}_1) \\
 \mathbf{u}_2 &= (\mathbf{v}_2 - t_1(\mathbf{v}_1)) \odot \exp(-s_1(\mathbf{v}_1)) \\
 \mathbf{u}_1 &= (\mathbf{v}_1 - t_2(\mathbf{u}_2)) \odot \exp(-s_2(\mathbf{u}_2))
 \end{aligned} \tag{1.13}$$

## Convolutional Neural Networks (CNN)

The term Convolutional Neural Network (CNN) (LeCun, 1989) refers to a category of DL models that utilize a type of linear operation called *convolution* in at least one of their layers. Convolutional operations offer several benefits to DL systems: sparse interactions, parameter sharing and equivariant representations (Goodfellow et al., 2016). Other types of neural networks such as MLP process the input data by computing interactions between each element in the input with every element of the output. In contrast, CNNs leverage sparse interactions by computing interactions between local neighborhoods of the input. In general, different layers of a CNN would compute interactions on neighborhoods of different sizes, thus allowing it to analyze the input data at different scales and learn both local and global features. This also implies a reduced number of operations because the number of interactions is only computed locally. Furthermore, the possibility to share parameters from each convolutional operation across the entire DL model allows CNNs reducing dramatically the amount of memory required to store it.

Given two functions  $f$  and  $g$ , the convolution of both, denoted by  $f * g$  is defined as the integral of the product of the two:

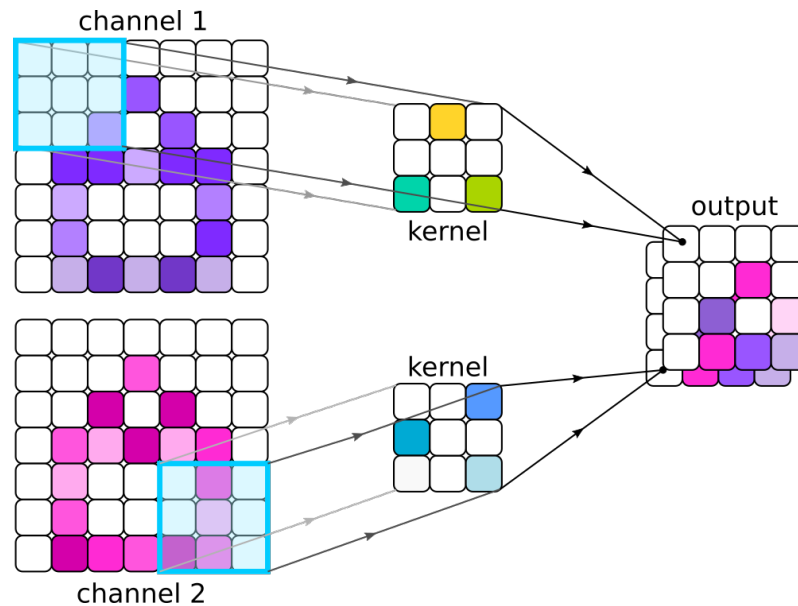
$$(f * g)(t) := \int_{-\infty}^{\infty} f(\tau)g(t - \tau)d\tau, \quad (1.14)$$

where  $g$  is sometimes referred to as the *kernel* function. This operation can be interpreted as the area under the function  $f$  weighted by  $g$ . Theoretically, both  $f$  and  $g$  can be real value functions, however, in practice, sensors may only record data in discrete values. Furthermore, the input  $t$  may also be a multidimensional array representing different types of data, e.g. images. In addition, this convolutional operation can be applied in more than one dimension at the same time. For a two-dimensional image with pixel intensities  $I$ , and a two-dimensional kernel  $K$ , the convolutional operation can be described as:

$$S(i, j) = (I * K)(i, j) = \sum_{a=1}^{N_x} \sum_{b=1}^{N_y} I(a, b)K(i - a, j - b) \quad (1.15)$$

This operation is applied across the entire image similarly to how a sliding window would be used (Fig. 1.10).

The way in which CNNs learn the representation of the input data (training process), equally to other DL models, is through a reward system. Every time that the CNN makes a prediction/estimation, a measure between the output and the expected outcome (labels) is computed and used to update the parameters of the model. Such measure is commonly known as *loss*, performance or reward. The way in which the parameters of the model are updated can be performed in many ways, one of the most common algorithms is



**Figure 1.10: Convolutional operations are in practice applied as matrix multiplications.** Exemplary diagram of a two-dimensional convolutional operation in two channels of an image with two spatial dimensions.

called *backpropagation*. Backpropagation computes the gradient of the loss relative to the input by leveraging the chain rule and updates the parameters of each model layer in the process. In this way, the model adjusts itself after processing each sample and improves its performance over time. For more details in model training and a more rigorous mathematical description of CNNs please refer to the work of Goodfellow et al. (2016).

### Random forest regression

Random forest (Breiman, 2001), also known as random decision forest, represent a group of ensemble learning methods generally used for classification and regression tasks. Each tree in a random forest is built from a sample drawn with replacement from the training data. When splitting each node of a tree, the best split is estimated either by selecting all the input features or by selecting a random subset of features. This injection of randomness in the construction of each tree decreases the variance of the forest estimator and avoids overfitting on the training data.

Given the training sample  $\mathbf{x}_i \in R^n$ , where  $i \in 1, 2, \dots, l$ , and a label vector  $\mathbf{y} \in R^l$ , a decision tree partitions the features in a way that samples with the same or similar labels are grouped together. Considering the data  $Q_m$  at node  $m$  with  $N_m$  samples, each



candidate split  $\kappa = (j, t_m)$  consisting of a feature  $j$  and a threshold  $t_m$  splits the data as:

$$\begin{aligned} Q_m^{left}(\kappa) &= \{(\mathbf{x}, \mathbf{y}) \in Q_m \mid \mathbf{x}_j \leq t_m\} \\ Q_m^{right}(\kappa) &= \{(\mathbf{x}, \mathbf{y}) \in Q_m \mid (\mathbf{x}, \mathbf{y}) \notin Q_m^{left}(\kappa)\} \end{aligned} \quad (1.16)$$

A traditional decision tree leverages all features  $j$  while a random forest tree samples a random subset of features. The quality of each candidate  $\kappa$  is then computed based on a loss function  $H(\cdot)$ . If the target is a continuous value (e.g. oxygenation), the loss can be any of Mean Squared Error (MSE), Mean Absolute Error (MAE), L2 error, etc. Many applications use the MSE, which yields:

$$\begin{aligned} \bar{\mathbf{y}}_m &= \frac{1}{N_m} \sum_{\mathbf{y} \in Q_m} \mathbf{y} \\ H(Q_m) &= \frac{1}{N_m} \sum_{\mathbf{y} \in Q_m} (\mathbf{y} - \bar{\mathbf{y}}_m)^2 \end{aligned} \quad (1.17)$$

The loss  $H(\cdot)$  yields the quality metric  $G(Q_m, \kappa)$ :

$$G(Q_m, \kappa) = \frac{N_m^{left}}{N_m} H(Q_m^{left}(\kappa)) + \frac{N_m^{right}}{n_m} H(Q_m^{right}(\kappa)) \quad (1.18)$$

Finally, the best split candidate is selected such that:

$$\kappa^* = \arg \min_{\kappa} G(Q_m, \kappa) \quad (1.19)$$

This process is repeated until a maximum depth of each tree is reached or the data on the node  $m$  cannot be divided.

Given the input data  $\mathbf{X}$ , a random forest  $\mathcal{F}$  estimates the outcome  $\hat{\mathbf{y}}$  as the average estimation  $\hat{\mathbf{y}}_i$  from each of its decision trees  $i$ . In particular, for the task of oxygenation estimation, the input data  $\mathbf{X}$  can represent either measurements or MC simulations of diffuse reflectances.

## 1.3 Related Work

This chapter elaborates on the related work regarding functional imaging, with a focus on medical applications. Band selection is an important topic in functional imaging, mainly because it allows to speed up recording and processing time by reducing the amount of data that needs to be collected. Hence, a method for band selection is proposed in Sec. 2.2.3 and its related work in Sec. 1.3.1. Processing the collected data in functional imaging is not always straightforward, for example, open surgical scenarios feature a highly dynamic environment where illumination is not known and can even change during surgery. This is a challenge for surgical translation because many methods require a controlled illumination or ground truth information about the light source. This limitation is addressed in this thesis with a method for automatic light source estimation in Sec. 2.3 and the corresponding related work is discussed in Sec. 1.3.2. Finally, estimation of functional properties is particularly challenging under the restriction of scarce data and high inter-patient tissue heterogeneity and lack of gold standard. Methods to address these challenges in the context of ischemia monitoring and perfusion monitoring are presented in Sec. 2.4.2 and Sec. 2.2, its related work is discussed in Sec. 1.3.3.

### 1.3.1 Band selection in spectral imaging

Band selection, or feature selection as it is known in the remote sensing community, generally refers to a wide variety of methods that aim to reduce the number of bands in an image to reduce computational costs while maintaining acceptable performance in a particular task. Even though several biomedical imaging modalities have shown great potential, like SI, their adoption in clinical practice is many times hindered by slow recording and processing speeds. For this reason, band selection related work is presented in this section, and acts as ground base for the method proposed in Sec. 2.2.3. Given that this thesis is mainly concerned with work in the biomedical domain, the methods presented in this section are for the most part related to interventional imaging, as summarized in Tab. 1.2.

Band selection (here on referred to as feature selection) algorithms can be grouped roughly in *filter*, *wrapper* and *embedded* categories (Guyon and Elisseeff, 2003). *Filter* methods determine the best features as a preprocessing step that is independent of the target task (e.g. oxygenation estimation). *Wrapper* methods use the downstream processing pipeline to search for feature combinations that maximize the target performance metric for a given task (e.g. accuracy of tumor detection). Finally, *embedded* methods are incorporated into the training process of a machine learning algorithm, where feature selection becomes part of the model construction. While *filter* methods can rely on unsupervised analysis,

*wrapper* and *embedded* methods usually need labeled data to estimate performance on each set of features.

State-of-the-art methods shown in Tab. 1.2 can be classified into three categories depending on the task being addressed: a) Cancer localization, b) Tissue visualization and c) Functional imaging.

**a) Cancer localization:** The methods in this category focus on selecting bands that maximize the performance of a classifier. In this context, Wood et al. (2008) proposed to select the bands which maximize the Area Under the Curve (AUC) of a Naïve Bayesian classifier’s Receiver Operating Characteristic (ROC). In more detail, they used an algorithm to remove the bands which contribute the least to the AUC while featuring a low AUC if evaluated alone. The algorithm was evaluated on phantoms and stained lung cancer biopsies, and the authors concluded that the use of three wavelengths provides essentially as much information as the use of all sixteen. The contribution of Han et al. (2016) focuses on gastric tumor localization. They selected bands which maximized the symmetric class-conditional Kullback-Leibler (KL) divergence between disease and normal tissue reflectance spectra. A suboptimal, greedy search algorithm (Du et al., 2007) was used for this purpose because the possible number of band subsets grows exponentially with the number of bands. This algorithm iteratively adds bands to a selection set based on the largest incremental increase in the KL measure. The algorithm was evaluated on 12 patients and a total of 21 colorectal tumors, and the authors concluded that five out of 28 bands are useful for outlining tissue disease regions.

**b) Tissue visualization:** Improving the contrast between normal and abnormal tissue has been the target of several approaches. Among these, Gu et al. (2016) selected three bands to discriminate gastric abnormalities from benign tissue. They aimed to replace RGB images with a different set of bands. Their algorithm first selects the band with the highest variance in the visible range from a set of 27 bands. The algorithm then adds subsequent bands iteratively with the criterion of minimizing mutual information between the current set of bands and the previous selection. The set of bands was optimized using 29 images from 12 patients with gastric abnormalities. The authors concluded that the three selected bands increase contrast compared to RGB images. In a similar approach, Nouri et al. (2014, 2016) inspected a number of unsupervised band selection algorithms within the context of hyperspectral ureter surgery. The authors aimed to identify three bands for better visualization (instead of RGB) to present to the surgeon. They evaluated the competing algorithms by several contrast and entropy metrics and assessed how different sets of bands can differentiate structures such as the ureter and adipose tissue. They found that the three best wavelengths to discriminate ureter tissue are situated in the near infrared

**Table 1.2:** Overview of relevant band selection methods. *Selected bands* represent the optimal number of bands reported by the authors. Our proposed method suggests a subset of ten out of 101 bands for the *in-silico* oxygenation experiments, but can be applied to a wide range of tasks.

Reference	Modality	Label-free	Method	Data	Application	Selected bands
<b>Proposed</b>	<b>HSI</b>	<b>yes</b>	<i>wrapper</i>	<i>in vivo</i>	<b>functional imaging</b>	<b>10/101</b>
Preece and Claridge (2004)	MSI	yes	<i>wrapper</i>	<i>in silico</i>	functional imaging	3
Marois et al. (2018)	MSI	yes	<i>wrapper</i>	<i>in silico</i>	functional imaging	6
Asfour et al. (2018)	HSI	yes	<i>wrapper</i>	<i>ex vivo</i>	tissue visualization	4/151
Wood et al. (2008)	MSI	no	<i>wrapper</i>	<i>ex vivo</i>	cancer localization	3/16
Wirkert et al. (2014b)	MSI	yes	<i>filter</i>	<i>in vivo</i>	functional imaging	8/20
Han et al. (2016)	HSI	no	<i>filter</i>	<i>in vivo</i>	cancer localization	5/28
Gu et al. (2016)	MSI	yes	<i>filter</i>	<i>in vivo</i>	tissue visualization	6/16
Nouri et al. (2014, 2016)	HSI	yes	<i>filter</i>	<i>in vivo</i>	tissue visualization	3/141

region. In another work, Guennadi and Andrei (2020) studied the optimal number of bands to increase the contrast of tissue structures. They created a link between image contrast ratio and optical tissue parameters based on the two-flux Kubelka-Munk model. They reported that bands around 550 nm maximize visualization contrast. Furthermore, Asfour et al. (2018) investigated the optimal bands to improve the visualization of atrial ablation lesions. They evaluated the performance of their algorithm in subsets of two, three and four bands based on supersets containing 151 and 31 equidistant bands in the wavelength range between 420 nm and 720 nm. The authors concluded that four bands can be used to enhance the contrast of ablated atrial tissue.

**c) Functional imaging:** In this context, band selection methods aim to select a subset of bands while maintaining good performance in at least one of such tasks (e.g. oxygenation estimation). Wirkert et al. (2014b) proposed to select bands in an unsupervised fashion by maximizing the differential entropy contained in the selected subset of bands. More specifically, the algorithm selects the subset of bands with the highest determinant of the bands' covariance matrix. Assuming an underlying normal distribution, this determinant is proportional to the differential entropy, and thus also to the information contained in the subset of bands. They tested this method in an *in vivo* porcine setting and concluded that a selection of eight bands leads to blood oxygenation values close to the baseline generated with 20 bands. In another work, Marois et al. (2018) selected bands by analyzing the absorption spectra and computing tissue oxygenation with the Beer-Lambert's law while leveraging a wavelength-dependent path length factor. They chose bands that maximize the product of the singular values of the absorption matrix, arguing that this maximizes the orthogonality of the fitted spectra. They evaluated the quality of the chosen bands in an *in silico* setting in the visible and infrared regions and used the root mean squared error of different molecular concentrations as a quality measure. Ultimately, they reported six optimal wavelengths for estimating the concentration of water, lipids, oxyhemoglobin and deoxyhemoglobin. A different idea was proposed by Preece and Claridge (2004). First, they generated Kubelka-Munk light transport theory-based simulations for assessing pigmentation of human skin. A genetic algorithm was then employed to find the best subset for estimating papillary dermis thickness and blood/melanin content. The authors ensured that the bands could invert the parameters uniquely by differential-geometric reasoning. By incorporating the ground truth from simulations allowed them to circumvent problems related to references from real data, as mentioned in Lu et al. (2015). The authors concluded that three bands selected according to their method lead to better results than RGB, but that RGB leads to reasonable results in the investigated context.

In summary, although several methods for band selection have been proposed, several of them require labeled data that can be difficult to acquire (Wood et al., 2008; Han et al., 2016; Nouri et al., 2014). In addition, some methods are based on the maximization of a non-specific measure such as amount of information (Nouri et al., 2016; Gu et al., 2016; Wirkert et al., 2016b), which is unrelated to the specific task at hand (e.g. oxygenation estimation). Finally, some methods (Preece and Claridge, 2004) were designed for *in silico* data because they require knowledge of the tissue molecular composition and can not be adapted to real data.

### 1.3.2 Automatic light source estimation

In the context of optical functional imaging, light source estimation refers to the estimation of the spectrum of a light source within the field of view of a camera. It is important to notice that the term “light source estimation” is strongly related to “color constancy” methods (Khan et al., 2017). Color constancy aims to represent images in a color space that is independent of the type of illumination. Most of the color constancy methods proposed up until now have emanated from the computer vision community, and can be classified in two main groups: machine learning approaches (Sec. 1.3.2) and model-based (Sec. 1.3.2). Tab. 1.3 presents a comparison of the most relevant approaches to light source estimation and color constancy. Although other methods have been proposed in this context, here only the ones with the most potential for clinical translation are discussed. Davi (2011) did an excellent work comparing earlier (less promising) approaches to color constancy.

#### Machine learning light source estimation

Among the machine learning approaches to light source estimation, CNNs are the dominant building block (Hu et al., 2017; Bianco et al., 2015; Hernandez-Juarez et al., 2020; Afifi and Brown, 2019; Laakom et al., 2019). Bianco et al. (2015) proposed a CNN to estimate the light source in a patch-wise manner by training it in natural scenes. Given an RGB image, they sampled non-overlapping patches from it and performed contrast normalization before estimating the light source from each patch. They then aggregated the estimated light sources from each patch to generate a global approximate. Similarly, Hu et al. (2017) paired a convolutional network architecture with a confidence metric that assigns a confidence value for each light source estimation on each analyzed patch of the image. Based on these weights, they generated a global estimate of the illumination to correct each image. In another work, Lo et al. (2021) proposed to learn general feature representations

**Table 1.3:** Comparison of light source spectrum estimation methods. Most of these methods are designed for Red-Green-Blue (RGB) images of natural scenes, only some of them have been extended to Spectral Imaging (SI) and none of them to the medical domain.

Reference	Domain	Modality	Method
<b>Proposed</b>	<b>Human lips, pig liver</b>	<b>SI</b>	<b>Machine learning</b>
Yoon et al. (2020)	Chicken tissue	SI	Model-based
Imai et al. (2011)	Real-world objects	SI	Model-based
Yoo and Kim (2019)	Real-world objects	RGB	Model-based
Kato et al. (2012)	Real-world objects	RGB	Model-based
Koscevic et al. (2021)	Natural scenes	RGB	Machine learning
Afifi et al. (2021)	Natural scenes	RGB	Machine learning
Lo et al. (2021)	Natural scenes	RGB	Machine learning
Laakom et al. (2021)	Natural scenes	RGB	Machine learning
Hernandez-Juarez et al. (2020)	Natural scenes	RGB	Machine learning
Laakom et al. (2020)	Natural scenes	RGB	Machine learning
Xiao et al. (2020)	Natural scenes	RGB	Machine learning
Laakom et al. (2019)	Natural scenes	RGB	Machine learning
Georgoulis et al. (2018)	Natural scenes	RGB	Machine learning
Hu et al. (2017)	Natural scenes	RGB	Machine learning
Bianco et al. (2015)	Natural scenes	RGB	Machine learning
Gu et al. (2014)	Natural scenes	RGB	Machine learning

by comparing similar and dissimilar image pairs (contrastive learning). Given a pair of two random images ( $\mathbf{I}_i, \mathbf{I}_j$ ) from the training set, they created the similar pair by augmenting one of them with a label-preserving augmentation strategy ( $t(\mathbf{I}_i), t'(\mathbf{I}_i)$ ), where  $t(\cdot)$  and  $t'(\cdot)$  represent two random augmentations. The dissimilar pair was generated by augmenting two samples with different scene content ( $t(\mathbf{I}_i), t'(\mathbf{I}_j)$ ). While the authors claim that their method is more robust to the state of the art, evidence of this is not yet clear. In a different approach, Laakom et al. (2020) proposed the Bag of Color Features (BoCF), which is a particular case of Bag-of-Features Pooling (BoFP) (Passalis and Tefas, 2017). In this approach, the authors paired the BoFP with a self-attention mechanism to discard irrelevant spatial information and thus learn only color-related information from regions like reflective surfaces. The idea of using attention mechanisms is very similar to the weighting approach implemented by Hu et al. (2017). Although attention mechanisms could aid the network to focus on regions that contain the most information about the light source, it cannot be guaranteed that the network will always focus on relevant locations. Indeed, Laakom et al. (2020) showed that the network sometimes focused on poorly illuminated regions, thus putting the reliability of this approach into question.

A challenge related to all machine learning algorithms is that their training requires ground truth knowledge on the illuminant, usually obtained by placing a color checker on each scene. This, however, can be seen as a disruptive process for intraoperative procedures. Moreover, the spectral mixing of different light sources can change during the procedure, rendering accurate calibration of the light source based on pre-operative or post-operative data unfeasible.

### Model-based light source estimation

These methods use basic assumptions about the image formation process to obtain the light source spectrum from single images. Khan et al. (2017) identified the four most commonly used methods for light source estimation in RGB images, and discussed how to apply them to multispectral images: **Max-RGB**, **Gray-world**, **Shades-of-gray** and **Grey-edges**. The first three methods work on the assumption that the  $p$ -norm of each band is proportional to the light source spectrum. **Gray-edges** assumes that the average band derivative in any scene is achromatic, therefore containing only information about the light source. The mathematical formulations for these methods are presented in detail in Sec. 2.3.3.

Several widely used methods leverage the *dichromatic reflection model* (Imai et al., 2011) which states that the light reflected from an object can be separated into a specular and a diffuse reflection component:

$$I(i, j, k) = \overbrace{S(i, j, k)}^{\text{diffuse}} L_k + \overbrace{c \cdot L_k}^{\text{specular}}, \quad (1.20)$$

where  $I(i, j, k)$  is the intensity of image band  $k$  at position  $(i, j)$ ,  $S(i, j, k)$  is the surface spectral reflectance,  $c \in [0, 1]$  is a constant and  $L_k$  is the illuminant spectrum in band  $k$ . According to this model, reflections from objects are projected into a two-dimensional space via Principal Component Analysis (PCA) assuming that specular and diffuse reflections form different clusters in such low dimensional space. A linear fit is then applied to each cluster and the light source spectrum derived from it. Although distinctive clusters might form from reflections in materials like metal and plastic, it is not true for biological tissue. Moreover, the assignment of a cluster to either specular or diffuse reflections is ambiguous, thus making it inapplicable in the biomedical domain.

Other model-based approaches are either only applicable to RGB images (Finlayson and Schaefer, 2001) or rely on restrictions not fulfilled by images in the medical domain. For example, Kato et al. (2012) assumes that the object surface is composed of a dielectric material such as plastic or paint. Kaneko et al. (2016) requires materials to be uniform and relies on a statistical daylight model, which is not available for surgical scenarios.



Furthermore, most of the methods in Khan et al. (2017) assume that the average color of a scene is achromatic, which is a strong restriction for surgical scenarios.

In summary, no prior work on illuminant estimation based on SI in surgery has been proposed to date. Methods proposed outside the field of medicine typically suffer either from unrealistic model assumptions (model-based approaches) or the need to acquire labeled training data that is difficult to acquire (machine learning-based approaches).

### 1.3.3 Perfusion monitoring

With the rise of more commercially available SI devices, more studies involving this technology have started to emerge. Clancy et al. (2020) did an astonishing work summarizing and comparing such devices. This review revealed that even though there are nonmedical devices capable of recording spectral data at video-rate, they have not been implemented in clinical settings. The vast majority of devices featured in medical scenarios are big and/or suffer from slow recording times (Dietrich et al., 2021a,b; Köhler et al., 2020; Hu et al., 2020; Takamatsu et al., 2021; Kumar et al., 2020). The only laparoscopic spectral device proposed for clinical use so far (Köhler et al., 2020) takes around five seconds to record one hyperspectral image, which prevents real-time application. This section contains a comparison of different studies performed with a wide range of functional imaging modalities, with a focus on SI. It is important to notice that study types marked as “human” refer to studies performed in humans but where no evidence of a structured clinical trial was found, usually one or two subjects were involved in this proof of concept studies. The analysis of these studies revealed that the majority of them are performed in animals with a limited number of individuals (biological replicates). Even more, the number of clinical studies involving these technologies is scarce. This could be attributed to the general lack of medically approved devices that can be used in large cohort studies involving different institutions.

The related work, presented in Tab. 1.4 that leverages SI and RGB data can be roughly classified into two broad categories based on the underlying method that supports them: a) Machine learning (Sec. 1.3.3) and b) Model-based (Sec. 1.3.3).

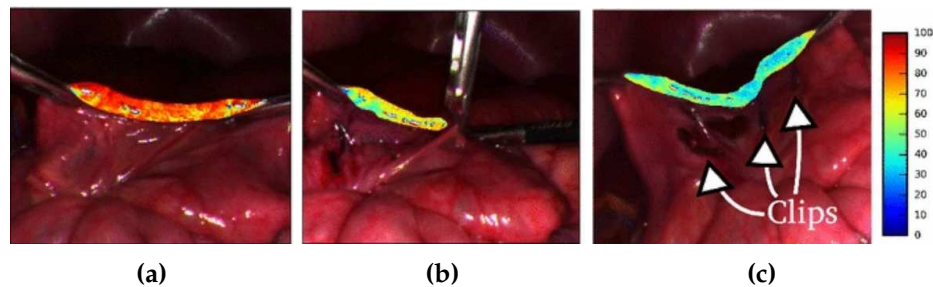
The literature presented in this section is the result of a thorough search involving each imaging modality. Although more literature could exist that involves functional imaging in biomedical settings, the studies presented in this section represents the ones that reported clearly the characteristics depicted in Tab. 1.4. In spite of this, the general trend described here holds true for the rest of the literature: *lack of structured clinical trials and small population sizes*. Furthermore, given that the methods presented in this thesis are mainly concerned with the translation of SI, special focus is placed in this modality. A comparison of different imaging modalities for functional imaging can be found in Sec. 1.2.2.

**Table 1.4:** Comparison of functional imaging studies performed over the last decade with a focus on Spectral Imaging (SI). The following imaging modalities are compared Spectral Imaging (SI), Red-Green-Blue (RGB), Photoacoustic Imaging (PAI), Diffuse Reflectance Spectroscopy (DRS), Spatial Frequency Domain Imaging (SFDI), functional Magnetic Resonance Imaging (fMRI), Laser Speckle Imaging (LSI), Polarization Spectroscopy (POS), Polarization Spectroscopy (POS) and Optical Coherence Tomography (OCT). Notice that the population size (biological replicates) was not reported for some studies, and that “human” studies refer to studies performed in humans but where no evidence of a structured clinical trial was found. Clinical studies are highlighted in gray and the methodology proposed in this thesis in this thesis in green. *BVF* refers to blood volume fraction and *sO<sub>2</sub>* to oxygenation.

Reference	Domain	Modality	Method	Population	Study type	Parameter
Proposed	<i>in silico</i> , rat neck, pig brain, human skin & kidney	SI	Machine learning	$5 \times 10^5$ , 8, 1, 12, 10	clinical	<i>sO<sub>2</sub></i> , <i>BVF</i>
Wirkert et al. (2016a)	Pig bowel	SI	Machine learning	1	animal	<i>sO<sub>2</sub></i> , <i>BVF</i>
Mori et al. (2014)	Human brain	SI	Beer-Lambert	1	human	<i>sO<sub>2</sub></i>
Soares et al. (2021)	Pig bowel	SI	Beer-Lambert	1	animal	<i>sO<sub>2</sub></i>
Clancy et al. (2015)	Pig bowel	SI	Beer-Lambert	1	animal	<i>sO<sub>2</sub></i>
Bouchard et al. (2009)	Rat brain	SI	Beer-Lambert	1	animal	<i>sO<sub>2</sub></i> , <i>BVF</i>
Suh et al. (2006)	Human brain	SI	Beer-Lambert	8	clinical	<i>sO<sub>2</sub></i> , <i>BVF</i>
Holmer et al. (2018)	Human hand	SI	Spectral derivatives	10	clinical	<i>sO<sub>2</sub></i> , <i>BVF</i>
Felli et al. (2020)	Pig liver	SI	Spectral derivatives	6	animal	<i>sO<sub>2</sub></i> , <i>BVF</i>
Bedard and Tkaczyk (2012)	Human lips	SI	Linear unmixing	1	human	<i>sO<sub>2</sub></i>
Jones et al. (2017)	Pig bowel	RGB	Beer-Lambert	1	animal	<i>sO<sub>2</sub></i> , <i>BVF</i>
Kumar et al. (2020)	Human hand	RGB	Oximeter-based	12	clinical	<i>BVF</i>
Gröhl et al. (2021)	Human forearm	PAI	Machine learning	3	clinical	<i>sO<sub>2</sub></i>
Tzoumas et al. (2016)	Pig muscle	PAI	Linear unmixing	8	animal	<i>sO<sub>2</sub></i>
Kleshnin (2019)	Simulations	DRS	Machine learning	$5 \times 10^3$	<i>in silico</i>	<i>sO<sub>2</sub></i>
Guo et al. (2022)	Pig kidney	DRS	Beer-Lambert	1	animal	<i>sO<sub>2</sub></i>
Chen and Durr (2020)	Human feet	SFDI	Beer-Lambert	4	clinical	<i>sO<sub>2</sub></i>
De Vis et al. (2013)	Human brain	fMRI	Pulsed ASL	4	clinical	<i>BVF</i>
Mangraviti et al. (2020)	Rat brain	LSI	Laser speckle	5	animal	<i>BVF</i>
Vasefi et al. (2015)	Human skin	POS	Beer-Lambert	2	human	<i>sO<sub>2</sub></i> , <i>BVF</i>
Choi et al. (2019)	Rat brain	OCT	Doppler shift	5	animal	<i>BVF</i>

### Machine learning perfusion monitoring

One of the most promising approaches to perfusion estimation, and the only one based on machine learning and SI, was proposed by Wirkert et al. (2016b). One of the most challenging aspects of perfusion estimation is the lack of a gold standard method to measure oxygenation *in vivo*. To overcome this, they proposed to use random forests paired with MC simulations of light transport to estimate tissue oxygenation. In this work, they proposed to optimize and train random forests by transforming the MC simulations to the space of an 8-band multispectral camera. They demonstrated in a pig model that *in vivo* oxygenation estimation is possible without the need of labeled data (Fig. 1.11).



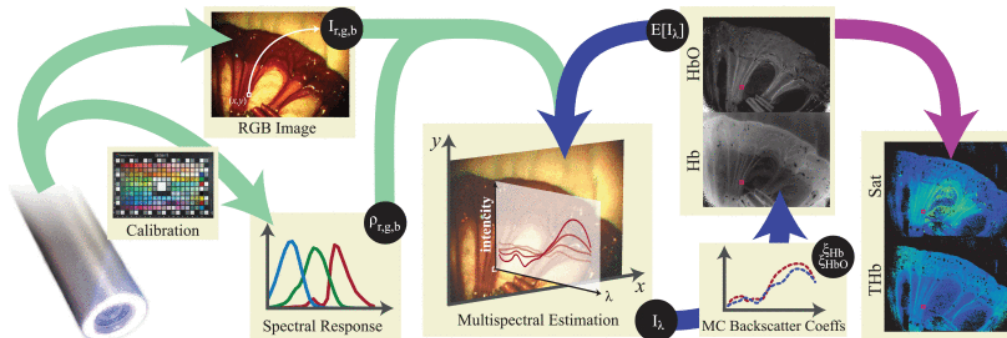
**Figure 1.11: Oxygenation estimation is possible *in vivo* with random forests.** Small bowel oxygenation estimation (a) before applying a clamp, (b) after applying four clamps, and (c) nine minutes after applying the first clamp. Image from Wirkert et al. (2016b).

Although this method removes the need for labeled data, the relatively low estimation speed of random forests paired with the slow sequential recordings of the multispectral camera leads to low translational potential. Furthermore, the data utilized in this approach was only augmented by introducing Gaussian noise during training, this could lead to misrepresentations of real data that contains artifacts such as specular reflections. Despite these limitations, this work paves the road for other methods that aim to leverage synthetic data in functional imaging.

### Model-based perfusion monitoring

Among the methods relying on the Beer-Lambert law, Jones et al. (2017) proposed an interesting approach. They estimated tissue blood content and oxygenation in an iterative process while relying solely on RGB images. They achieved this by reconstructing multispectral data based on RGB images by using a Bayesian *maximum a posteriori* estimate<sup>‡</sup>. In this scenario, the prior is the previous concentration estimations of oxyhemoglobin and

<sup>‡</sup>That is, an estimate of an unknown quantity that equals the mode of the posterior distribution.



**Figure 1.12: Functional properties are computed from Red-Green-Blue (RGB) images in an iterative process that relies on the Beer-Lambert law.** Spectral data is recovered from RGB images and the spectral power distribution by using a linear square fit, and the concentrations of hemoglobin and oxygen are computed based on a Bayesian estimator. The spectral estimations are then optimized in an iterative process based on such molecular concentrations. *Image from Jones et al. (2017).*

deoxyhemoglobin. Based on these multispectral image estimations, the concentrations of hemoglobin are computed via linear least squares fit. The process is then repeated until convergence is reached. With this approach, they were able to compute oxygenation and blood tissue content at a speed of 2.5 Hz. Although this approach leverages RGB cameras that are already part of clinical routine, the relatively slow processing speed hinders its applicability during surgery. As indicated by Jones et al. (2017), the processing speed could be increased to 30 Hz if a look up table (LUT) was precomputed for a specific RGB sensor with 24 bit colors. Even though this increase in speed could allow to image surgical scenes under relatively high organ motion, it requires the computation of a LUT for each camera sensor, thus hindering the translation of this method.

Furthermore, the majority of studies presented in Tab. 1.4 rely on the Beer-Lambert law to estimate functional properties, mainly tissue blood content ( $BVF$ ) and oxygenation ( $sO_2$ ). The Beer-Lambert law (Baker et al., 2014; Jones et al., 2017) states that the intensity of light traveling inside a medium decreases exponentially while increasing the depth of such medium. Even though many studies have developed methods relying in this law (Mori et al., 2014; Soares et al., 2021; Clancy et al., 2015; Bouchard et al., 2009; Suh et al., 2006; Jones et al., 2017), the conditions required to rely on this law are generally not met in clinical applications:

**Absorbers must act independently of each other:** This is a requirement that cannot be guaranteed, especially when fluorescent molecules might be present (e.g. ICG). Fluorescent molecules could emit light that is then partially absorbed by other molecules in the tissue.

**Concentration of absorbers should be homogeneous in the interacting volume:** Different parts of tissue structures can have different concentrations of absorbers. For example, the content of blood can vary within the tissue depending on how it is perfused (e.g. skin).

**The incident light must traverse the same path length inside the medium:** By nature, light can travel different path lengths within the tissue, this can be attributed to tissue heterogeneity (e.g. molecular concentrations, organelles, etc.).

For these reasons, methods that rely on the Beer-Lambert law to compute molecular concentration of biological tissue are not a well-fitted for clinical translation.

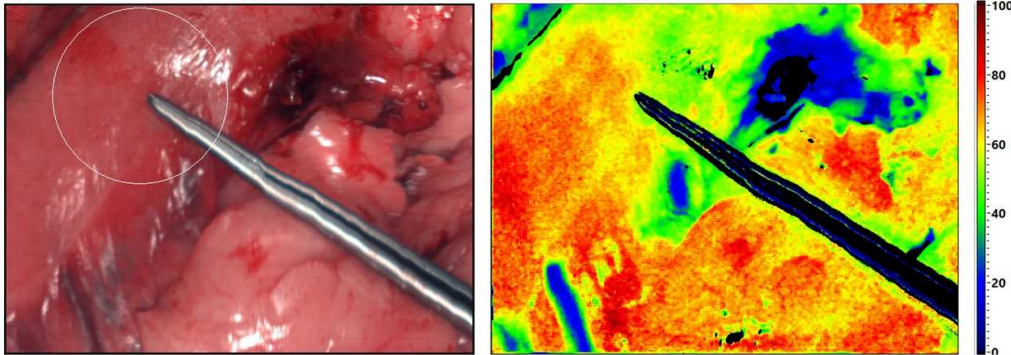
In another recent work, Holmer et al. (2018) proposed using a spectral sequential scanning approach (line scan) coupled with spectral derivatives to estimate functional tissue properties (oxygenation, blood volume fraction and water content). Given that changes in molecular composition of biological tissue would lead to peak intensity changes in the reflectance spectrum, but not in the spectral location of such peaks, they proposed to leverage the first and second derivatives of the reflectance spectrum to quantify molecular concentrations. For this purpose, they computed the first and second derivatives of the reflectance spectrum in wavelength regions that are most affected by changes in concentration of different molecules. For example, hemoglobin features a characteristic peaks between 570 nm and 590 nm, which change intensity depending on the amount of oxygen that it contains. Although this method is simple enough to implement, it requires highly detailed spectral data to reliably compute the first and second spectral derivatives. This, coupled with sequential recordings, leads to slow recording speeds (a few seconds per spectral image). The effect of such slow recordings can be appreciated in Fig. 1.13, where the “wavy” pattern of the surgical instrument is a result of motion artifacts. Such artifacts would only be increased during minimally invasive surgery, where organ motion can be a more influential factor.

Recently, Kumar et al. (2020) proposed to pair an RGB camera with an oximeter to estimate blood perfusion in skin. In this contribution, Kumar et al. (2020) estimated a perfusion-related signal from RGB images by applying a band pass filter between the frequencies of 0.5 Hz and 5 Hz to extract the cardiac-related signal. They then normalized this cardiac-related signal with the same RGB data after applying a filter with cutoff frequency of 0.3 Hz. They computed the final perfusion estimation by pairing the filtered and normalized RGB data with data from the pulse oximeter (Fig. 1.14).

To consider possible outliers, Kumar et al. (2020) employed an M-estimator<sup>§</sup> that assigns lower weights to outlier samples. Similarly to the work by Jones et al. (2017), this method could leverage existing equipment available in the clinics. However, they

---

<sup>§</sup>This can be interpreted as a weighted least square approach.

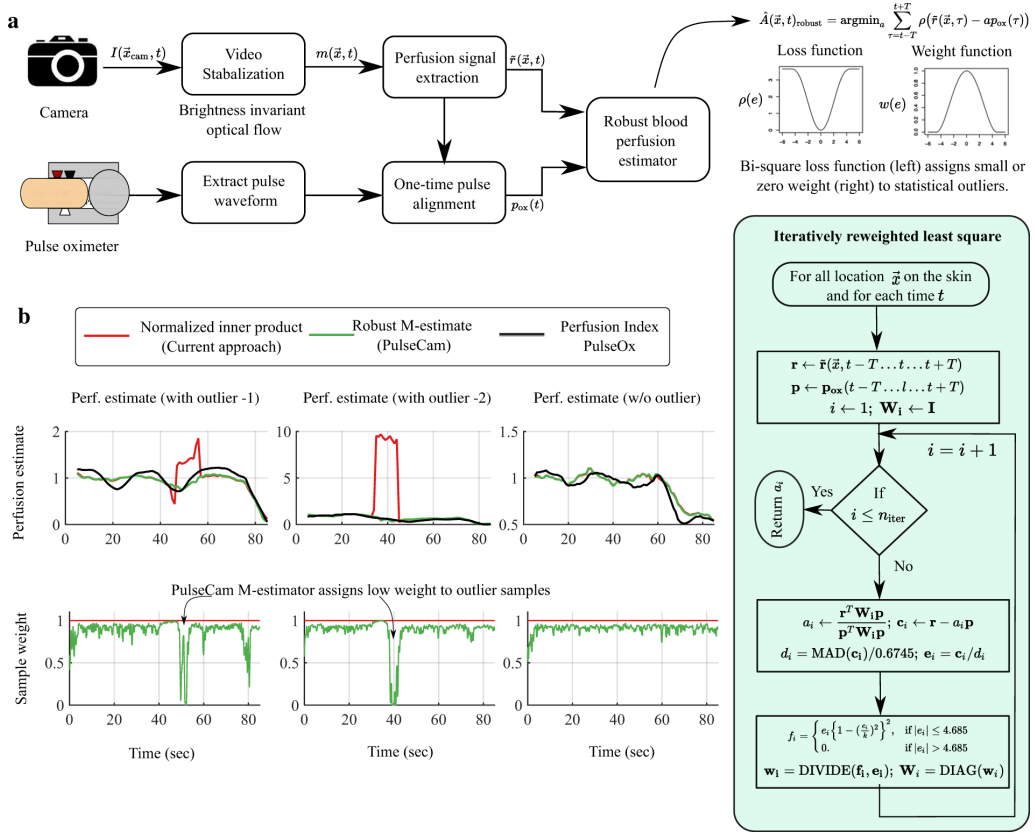


**Figure 1.13: Slow sequential spectral recordings can lead to motion artifacts.** Reconstructed Red-Green-Blue (RGB) image and oxygenation estimation of human esophagus. Due to the sequential nature of the recording process (line scan), motion artifacts can be visualized in the image, even under controlled conditions. *Image from (Köhler et al., 2019).*

require video stabilization across all recorded images with an optical flow approach, which can considerably reduce the perfusion estimation speed. Moreover, the final perfusion estimation for each image is performed with an optimizer in an iterative fashion, which is not guaranteed to converge. Finally, the signals between the pulse oximeter and the RGB camera need to be aligned, which only needs to be done once during each procedure, however, a method to guarantee the success of this step is missing. In turn, these limitations make this method difficult to be adopted in clinical practice.

The computation of functional properties in the medical domain is generally paired with the analysis of Regions of Interest (ROIs) to reduce the effect of unwanted complications, such as organ bleeding, burned tissue, scarring, etc. Moreover, surgical applications are generally focused towards estimating functional properties in specific regions of the surgical scene like organs or parts of them during selective clamping Porpiglia et al. (2018). For applications where recording speed is a limiting factor (Holmer et al., 2018; Köhler et al., 2019), the analysis of individual images is generally preferred. In such cases, tracking ROIs across multiple images is not needed, although slow frame rates hinder translation. On the other hand, in more promising approaches that are video-rate capable, the need for automatic ROIs tracking becomes evident. Unfortunately, until recently, the focus of ROI tracking has been in the domain of natural scenes rather than the medical domain. Tab. 1.5 presents recent work on ROI tracking developed for both natural scenes and medical domains.

Tracking biological ROIs is challenging in many surgical domains, this is due to several factors: a) organs can have very homogeneous surface textures that reduce the available contrast for tracking (e.g. liver), b) the presence of outliers (specular reflections, smoke,



**Figure 1.14: Perfusion estimation in skin is done by pairing a pulse oximeter with an Red-Green-Blue (RGB) camera.** This method extracts the perfusion signal from RGB data after performing video stabilization with an optical flow approach. The extracted signal is optimized using the data collected from a pulse oximeter and an outlier weighting estimator. The green box underlines the different steps performed by the estimator. **(a)** depicts the overall strategy employed to leverage data from an RGB camera and a pulse oximeter, and **(b)** illustrates the effect of outlier weighting on the perfusion estimation approach. *Image from Kumar et al. (2020).*

scaring, etc.) can throw off tracking estimations, and c) the number of datasets with ground truth labels is very scarce in the medical domain. Most of the ROI tracking methods in the medical domain are limited to tracking surgical tools (Speidel et al., 2015; Nwoye et al., 2019; Liu et al., 2005) based on RGB data. Until now, there is still a need for a robust tissue ROI tracker in the medical domain that can be utilized not only in RGB data but on spectral data as well.



**Table 1.5:** Comparison of Region of Interest (ROI) tracking methods. Most ROI tracking algorithms are implemented for natural scenes, which feature a strong contrast between the target and background.

Reference	Domain	Modality	Method
Proposed	Human skin, human kidney	SI	Machine learning
Du et al. (2015)	Human prostate	RGB	Model based
Nwoye et al. (2019)	Surgical tools	RGB	Machine learning
Liu et al. (2005)	Surgical tools	RGB	Edge detection
Speidel et al. (2015)	Surgical tools	RGB	Model based
Boltz et al. (2009)	Natural scenes	RGB	Machine learning
Lukezic et al. (2018)	Natural scenes	RGB	Correlation filters
Zhao et al. (2018)	Natural scenes	RGB	Correlation filters

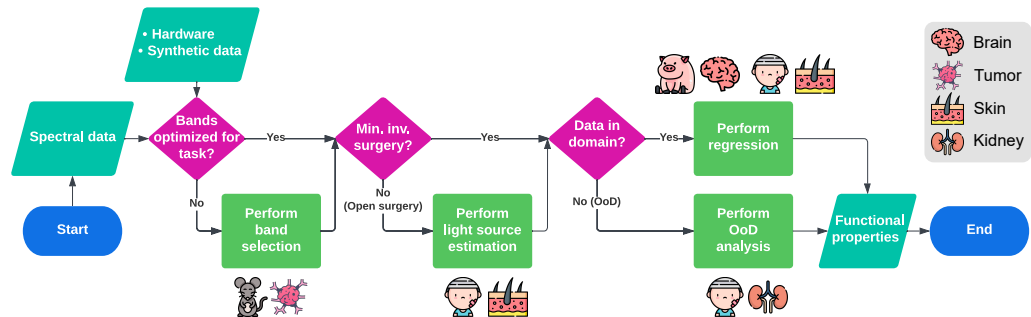
All in all, despite the potential that previously proposed methods have shown, they lack certain characteristics that prevent their translation into clinics:

**Low recording/processing speed:** All the literature presented in this chapter, as well as a result of the review by Clancy et al. (2020), confirm that video-rate perfusion imaging remains an unsolved challenge, mainly due to slow recording devices and slow data processing methods.

**Large devices:** Several of the approaches presented in this chapter require large devices (Holmer et al., 2018; Takamatsu et al., 2021; Dietrich et al., 2021a,b; Kumar et al., 2020) that are difficult to handle for long periods of time during surgery.

**Domain gap:** The overwhelming majority of methods for perfusion estimation validate their approach in specific domains (human skin, pig bowel, etc.), while generalizability to other domains is overlooked.

**High inter-patient variability is ignored:** All studies presented in this section have been validated in small population sizes, thus raising the question of inter-patient variability and the potential generalization of such methods to unseen patients.



**Figure 1.15:** This thesis proposes a concept for translational functional imaging in surgery that is validated in a wide range of domains. This concept addresses common roadblocks in the path to translation of functional imaging technologies: a) the lack of labeled *in vivo* data, b) the slow recording and processing speed of Spectral Imaging (SI) systems, c) the estimation error introduced by the dynamic illumination environment of open procedures, and d) the high inter-patient tissue heterogeneity, for example in the presence of pathologies.

## 1.4 Objectives and Contributions

The main objective of this thesis was to address challenges relevant to the clinical translation of functional imaging. More specifically, the translation of SI in surgical applications. To overcome such challenges, a concept for translational functional imaging in surgery was developed (Fig. 1.15). One of the most common roadblocks in clinical translation of perfusion monitoring methods is the lack of ground truth labels for *in vivo* data. This was addressed in this thesis by leveraging highly accurate synthetic data. The lack of clinical data is also related to devices being too big or too slow to be used in large clinical trials. This was addressed by combining a fast and compact SI device with advanced image processing modules that reduce the amount of data that needs to be recorded, increase accuracy, and speed up data processing via deep learning components. Challenges related to translation to open surgical scenarios were also addressed by introducing a dedicated light source estimation module that reduces potential perfusion estimation errors. In contrast to other related work in the field of functional imaging in surgery, the concept proposed in this thesis was validated in a wide range of domains: *in silico*, rat head and neck, pig brain, and human skin and kidney. Based on these concepts, the following contributions were developed in the context of this thesis.

**Spectral imaging:** This thesis pioneers video-rate perfusion monitoring with SI by modifying existing devices to enable their application during minimally invasive surgery. Furthermore, the first module in the framework proposed in this work aims to speed

up data recording and acquisition by reducing the amount of data required for specific medical tasks and domains. This method increases the speed of SI devices by reducing the amount of spectral bands that they need to record. Even more, it improves accuracy of oxygenation estimation by leveraging large amounts of synthetic data, which resembles real-world conditions, and unlabeled real data to adapt to specific target domains. This approach was validated *in vivo* by studying the oxygenation of normal and tumorous tissue in a rat head and neck model. The results of this analysis were worthy of a publication in the journal *Biomedical Optics Express (BOE)*.

**Automatic light source estimation:** In contrast to minimally invasive surgery, open surgical procedures feature changing illumination conditions that lead to functional parameter estimation errors when the incorrect illumination is assumed. Therefore, an automatic light source estimation module was implemented, which leverages information from specular reflections. Specular reflections are usually undesired and unavoidable artifacts that possess the unique characteristic that the information contained within them originates mainly from the light source that generated them. Given that such reflections appear as over-exposed regions, the light source estimation module leverages low exposure images where such reflections do not appear as over-exposed. By estimating the light source spectrum of the light source, this method is capable of reducing the error produced during functional parameter estimations under changing illumination conditions. The benefits of this approach were demonstrated both *in silico* and *in vivo* human skin, and merited a publication in the *International Journal of Computer Assisted Radiology and Surgery (IJCARS)*.

**Perfusion monitoring:** The computation of functional properties takes place in two separate modules designed to address different challenges (here on referred to as functional modules).

- **Perfusion monitoring via regression:** The first of the functional modules features a deep learning model designed to leverage highly accurate physics-based simulations of light-tissue interactions to regress functional properties (regression module). By leveraging synthetic data, this module addresses the lack of ground truth *in vivo* functional labels. This method was validated in different species (e.g. rats, pigs, and humans) and domains (e.g. head and neck, brain and skin) to test its generalizability. In particular, the potential of this approach was demonstrated by monitoring the harmful SDs (deoxygenation waves) at video-rate on the surface of the brain. Even more, part of this contribution was published in the *Medical Image Computing and Computer Assisted Intervention*

*Society (MICCAI) workshop OR 2.0 and was awarded the **bench to bedside award**. Part of this contribution was also awarded the **young scientist award** by the *international Society for Minimally Invasive Therapy (iSMIT)*.*

- **Perfusion monitoring via Out-of-Distribution (OoD) detection:** The second functional module features a deep learning model that is designed to deal with samples that are not within the domain of the light-tissue simulations that the regression module exploits. In particular, this module rephrases the task of ischemia monitoring as an OoD detection task. By training this OoD module solely on data from individual patients, it is capable of detecting perfusion changes as outliers. This approach addresses the challenge of scarce clinical data by implementing a personalized approach that requires data from individual patients. Furthermore, this module was validated in the first clinical trial employing MSI in minimally invasive surgery (partial nephrectomy) for video-rate ischemia monitoring with deep learning. Part of this contribution was introduced as an oral presentation at the *13th International Conference on Information Processing in Computer Assisted Interventions (IPCAI)*.

In summary, this thesis contributes methods for video-rate perfusion monitoring during surgery that are validated in a wide range of domains: *in silico*, rat head and neck, pig brain, and human skin and kidney. Translation to open procedures of video-rate perfusion monitoring is aided by a dedicated module for automatic light source estimation. Finally, it presents the first clinical trial employing video-rate MSI in laparoscopic partial nephrectomy for ischemia monitoring.

#### **Disclosures to this work:**

The work presented in this thesis was developed in a highly interdisciplinary environment, where many team members contributed with fruitful discussions and analysis. Although I have written this thesis independently, in order to reflect and acknowledge the various team efforts done by the groups that I have collaborated with over the years, the rest of the this thesis is written in first person plural (“We”) instead of first person singular (“I”). For clarity, details about my contributions on each topic presented in this thesis are provided in the “*Own contributions*” chapter at the end of this document.

## 1.5 Outline

In the context of this thesis, the methodology used throughout this thesis is presented in Chap. 2. First, the concept for translational functional imaging is described in Sec. 2.1. Then, the measurement hardware utilized for spectral recording in various medical applications as well as the generation of synthetic data, and domain- and task-specific band selection are presented in Sec. 2.2. Our approach and baseline methods for automatic light source estimation for open surgical applications is presented in Sec. 2.3. Later on, the methodology for perfusion monitoring in various domains, and the first ever clinical trial with SI for ischemia monitoring are presented in Sec. 2.4.2. The experiments and results that back up the claims in this thesis are laid out in Ch. 3. The technical validation of several components of the concept for translational functional imaging (Sec. 2.1), such as the synthetic data generation, and domain and task-specific band selection method are presented in Sec. 3.1.1. The results of the proposed approach for automatic light source estimation in *ex vivo* and *in vivo* experiments are described in Sec. 3.2. The approaches for functional imaging (perfusion monitoring) are then validated in several medically relevant domains, the results of this validation are presented in Sec. 3.3.1, Sec. 3.3.1, and Sec. 3.3.2. Finally, Ch. 4 discussed the methodology and results presented in previous chapters and offers insights into future research directions as well as pinpoints knowledge voids that the biomedical community needs to address to translate functional imaging methodologies into clinical practice. A summary of this thesis is also provided in Ch. 5.



## Materials and Methodology

*You are never wrong to do the right thing.*

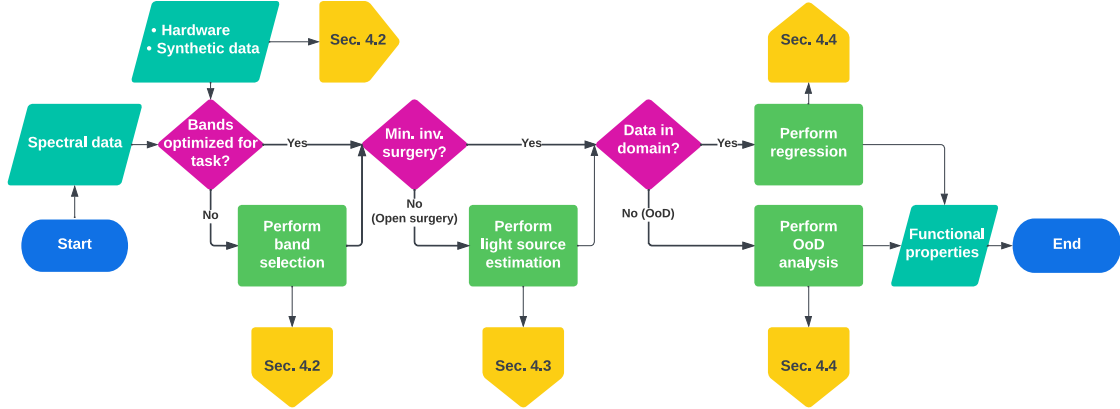
— MARK TWAIN

In this chapter, we present in Sec. 2.1 our general concept for functional imaging in surgical procedures by leveraging several machine and deep learning methodologies. We then elaborate on the hardware used to achieve video-rate perfusion monitoring in surgery and our approach to domain- and task-specific band selection in Sec. 2.2 by exploiting highly accurate Monte Carlo (MC) simulations. We address the challenge of light source estimation in open surgical procedures in Sec. 2.3, and we present our methods for perfusion monitoring in a wide range of domains in Sec. 2.4.

### 2.1 Framework for intraoperative functional imaging

While previous approaches to functional imaging in medicine have focused on specific processing components or medical applications (Clancy et al., 2015; Holmer et al., 2018; Kumar et al., 2020), we introduce a concept that combines several modules to overcome common challenges in translational imaging (Fig. 2.1). The advantage of our methodology is reflected in the success we demonstrate in very different medical applications: partial nephrectomy (Sec. 2.4.2), neurosurgery (Sec. 2.4.1) and skin perfusion monitoring (Sec. 2.4.1). This is also demonstrated through the wide range of biological specimens in which our methodology can be successfully applied: rats (Sec. 2.2.3), pigs (Sec. 2.4.1) and humans (Sec. 2.4.2).

Our concept for translational Spectral Imaging (SI) in surgery comprises four interconnected modules that address the following challenges: a) slow recording and processing speed of SI devices, b) low perfusion monitoring accuracy under changing illumination conditions (open surgery), c) perfusion monitoring under the lack of labeled *in vivo* data, and d) perfusion monitoring in outlier domains (not represented by simulations). One of the biggest challenges related to the translation of optical functional imaging technologies is the low recording and processing speed of imaging devices. For this reason, we propose a method for domain- and task-specific band selection to reduce the amount of data that such devices need to acquire, all of this while maintaining a good performance in specific tasks (Sec. 2.2.3). The second challenge addressed with our concept is related to minimizing the estimation errors that optical imaging modalities might suffer from



**Figure 2.1: Our concept for functional imaging during surgical procedures can handle both minimally invasive and open surgery.** The pipeline for functional imaging that we propose is capable of optimizing the spectral bands to specific tasks, it can handle open surgical scenarios via a dedicated light source estimation module, and is capable of dealing with data that is out of domain, such as tissue in the presence of pathologies.

when the surgical scene illumination changes over time. To overcome this drawback, we propose a dedicated automatic light source estimation module capable of estimating the light source spectrum during open procedures (Sec. 2.3). The third challenge that we address refers to perfusion monitoring during surgery. We approach this challenge with two separate modules. The first of them leverages a deep learning approach and highly accurate physics-based simulations of light-tissue interactions to regress functional properties (e.g. oxygenation and blood content). As such, this method overcomes the lack of *in vivo* labelled data in functional imaging (Sec. 2.4.1). The second perfusion monitoring module addresses the challenge of estimating functional properties in domains that are not well represented by the light-tissue simulations (outlier domains). We propose to achieve this by rephrasing the ischemia monitoring task as an Out-of-Distribution (OoD) detection task. By doing so, our deep learning-based method offers a personalized approach that only requires perfused data from individual patients to detect perfusion changes as outliers (Sec. 2.4.2).

## 2.2 Spectral imaging

The task of functional imaging in an *in vivo* setting is a hard problem, mainly due to the lack of a gold standard method to validate the estimations that any algorithm might produce. Furthermore, the hardware that is required to acquire validation data for such



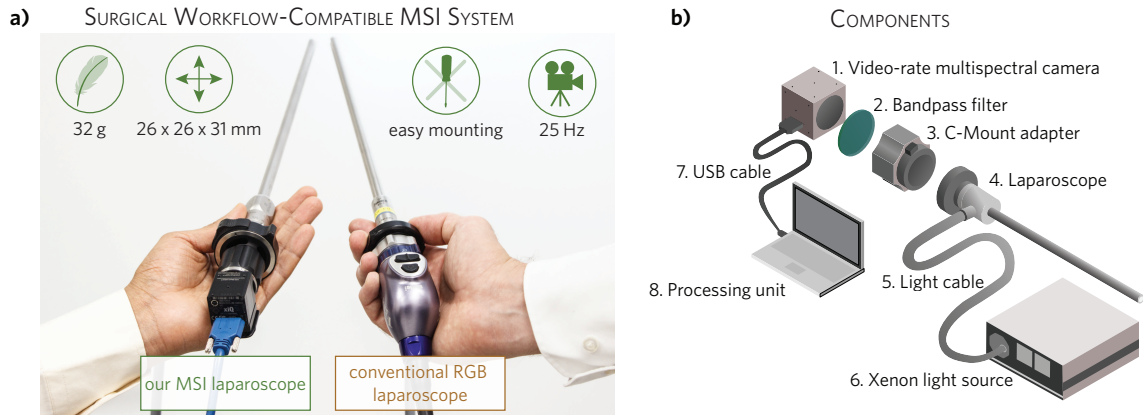
algorithms is big, heavy or suffers from slow recording times (Kumar et al., 2020; Holmer et al., 2018; Clancy et al., 2015). This makes current hardware usage impractical in surgical settings, and in turn renders acquisition of validation data unfeasible. Despite these limitations, a few methods have been proposed over the last decade with the aim of producing reliable oxygenation estimations in biological tissue. In this section, we present a compact imaging system with the aim of recording spectral data during laparoscopic surgery (Sec. 2.2.1). We then present our approach to generate and label multi-layer tissue models with a MC method (Sec. 2.2.2). Finally, we present our approach to domain and task-specific band selection for the task of oxygenation estimation (Sec. 2.2.3).

### 2.2.1 Real world spectral imaging

Over the past decades, our community has seen many advances in the field of spectral imaging, from higher spatial and spectral resolution sensors to smaller camera housings. The most prominent available spectral systems have been listed and described by Clancy et al. (2020). Given that there is no commercially available spectral system capable of recording spectral images at video-rate during surgery, we decided to build our own. More specifically, we designed two spectral systems, a laparoscopic video-rate capable SI system (ideal for minimally invasive surgery), and an open system featuring high spatial resolution and narrower spectral bands (ideal for open static operations). In the following, we describe our customized spectral systems.

**Video-rate multispectral imaging** We propose a multispectral system that enables us to pioneer video-rate Multispectral Imaging (MSI) in laparoscopy. Due to its compact ( $26 \times 26 \times 31$  mm) design, the camera does not add hardware complexity to the Operating Room (OR) (Fig. 2.2 a). In addition, the camera is lightweight (32 g), it enables easy handling over long periods of time, and it operates at a frame rate of 25 Hz. As shown in Fig. 2.2 b), it comprises (1) a snapshot multispectral camera, (2) a 335–610 nm band pass filter, (3) a C-Mount adapter with adjustable focal length, (4) a standard surgical laparoscope, (5) a surgical light cable, (6) a Xenon light source, and (7) a USB cable to connect the MSI camera to the (8) processing unit.

**Multispectral camera** One of the main component of our system is a multispectral camera (MQ022HG-IM-SM4x4-VIS, XIMEA GmbH, Münster, Germany), which is small ( $26 \times 26 \times 31$  mm) and light (32 g). This camera takes advantage of a mosaic snapshot sensor (imec, Leuven, Belgium), which acquires 16 spectral bands at the same time in the wavelength range 450–650 nm. This is achieved by using a  $4 \times 4$  repeating mosaic pattern of filters (see Fig. 2.3 b)); the spectral response of each filter is shown in Fig. 2.3 c).



**Figure 2.2: Our multispectral imaging (MSI) system for ischemia monitoring is comparable to current standard equipment in terms of size and weight. (a)** Proposed spectral imaging-based laparoscope (left) compared to standard Red, Green, and Blue (RGB) laparoscope (right). **(b)** Schematic representation of the system developed for spectral tissue analysis in laparoscopic surgery. The dimensions of the laptop and the light source have been scaled down for visualization purposes.

Multiple bands of the camera show two peaks in the spectral response. These are caused by the Fabry-Pérot cavities (Poirson et al., 1997) that each filter forms, which generates the so-called “second order” peaks. The intensity of such second order peaks depends on the height of the optical cavity of the filters, the refractive index of the sensor material and the cosine of the light incidence angle (Koonen, 2006). The total spatial resolution of the snapshot sensor is  $1088 \times 512$  px, thus yielding a resolution per band of  $272 \times 512$  px.

**Surgery-specific components** Due to intrinsic optical properties of biological tissue, the diffuse reflectance of human tissue in the red region (above  $\approx 620$  nm) is higher than that in the blue region (below  $\approx 490$  nm) \*. To ensure more balanced camera counts across spectral bands, and thus similar signal-to-noise levels across different camera filters, a 335–610 nm band-pass filter (FGB37, Thorlabs Inc., Newton, New Jersey, United States) was placed between the C-Mount adapter and the multispectral camera. The C-Mount adapter (20200043, KARL STORZ SE & Co. KG, Tuttlingen, Germany) has an adjustable focal length, with a maximum of 38 mm. To facilitate the usage of our system during minimally invasive surgery, the camera was connected to standard  $30^\circ$  laparoscopes (26003BA, KARL STORZ SE & Co. KG, Tuttlingen, Germany and Panoview, Richard Wolf, Knittlingen, Germany) via the C-Mount adapter. The  $\ell_1$ -normalized transmission spectra of the band-pass filter, the laparoscope and the C-Mount adapter are shown in Fig. 2.3 a).

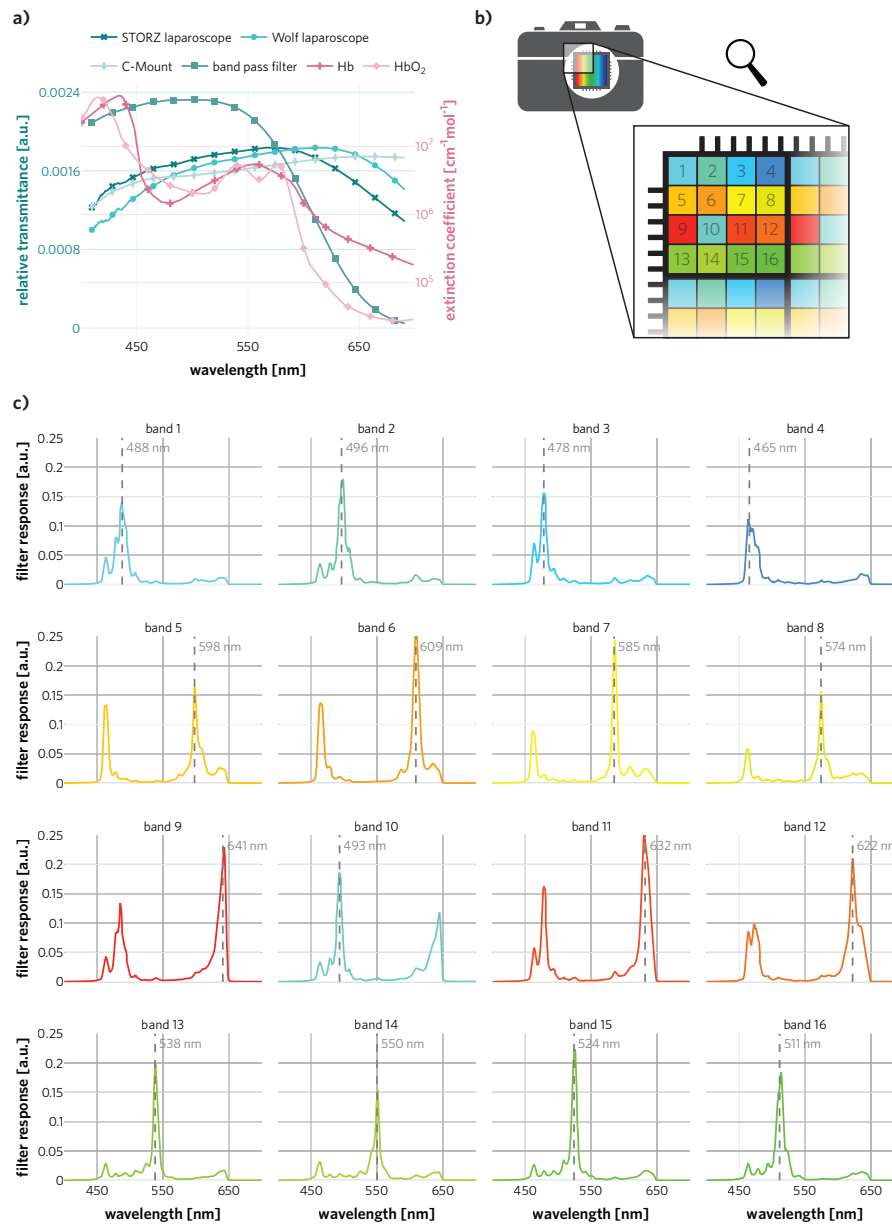
\*Internal organs tend to have a red color.

The extinction coefficients of HbO<sub>2</sub> and Hb are also shown as a reference (Prahl, 1999). We chose a Xenon light source (IP20, Richard Wolf GmbH, Knittlingen, Germany) because, in contrast to Light-Emitting Diodes (LEDs) and halogen lights, it provides more uniformly distributed spectral power across the visible wavelength range. All optical transmission profiles shown in Fig. 2.3 a) were measured with a spectrometer (HR2000+, Ocean Insight, Orlando, USA). 100 transmission measurements of each optical component were averaged and then smoothed across wavelengths by rolling window averaging with a window of  $\approx 19$  nm width. The values in the ranges 400–419 nm and 681–700 nm were ignored after smoothing to avoid undesired border effects.

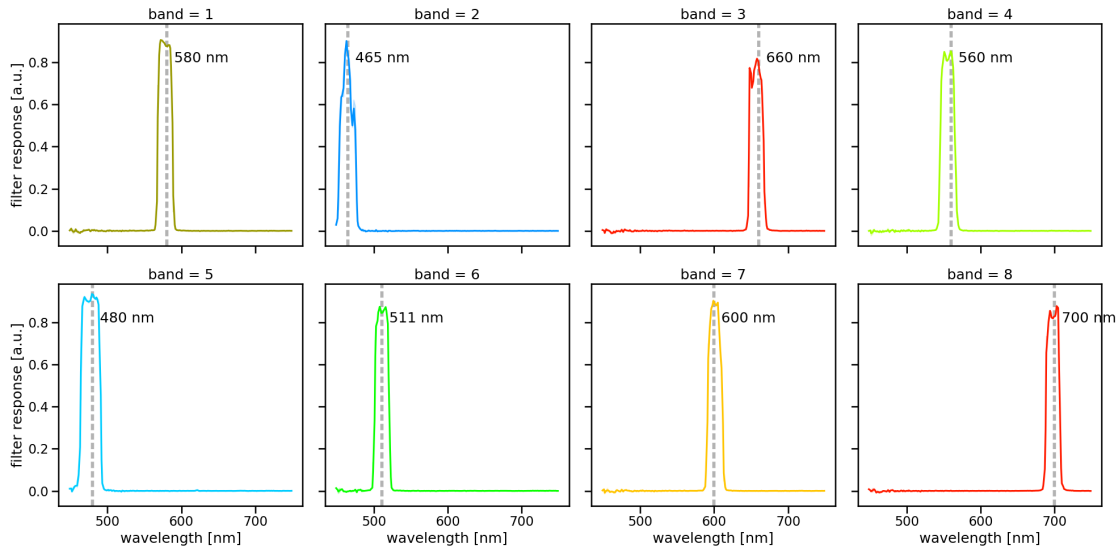
**Image recording and storing** We paired our optical system with a standard laptop for video recording (Msi GE75 Raider 85G, Intel i7, NVIDIA RTX 2080). A custom C++ software (not publicly available) based on the XIMEA Application Program Interface (XiAPI, XIMEA, Münster, Germany) was implemented as an interface to the multispectral camera, and it was used to record the multispectral images.

Even though a spectral system with video-rate recording capabilities has many applications in minimally invasive surgery, other medical fields could benefit from higher spatial resolution at the cost of slower recordings. One of such fields is neurosurgery, where a highly controlled surgical environment reduces organ motion, thus opening the doors to the possibility of a more detailed spatial analysis. With this in mind, we built a system that features higher spatial resolution and narrower spectral bands in comparison to the laparoscopic system (Fig. 2.2), in the following paragraphs we describe such new system.

**High resolution spectral imaging in neurosurgery** To be able to record spectral images at a relatively fast speed at increased spatial resolution we use a 5 Mpx camera (Spectrocam, Pixelteq, Largo, FL, United States) which records multispectral images at a frame rate of approximately 2 Hz. The central wavelengths of the eight filters were determined with the method described by Wirkert et al. (2014a) and set to 465, 480, 511, 560, 580, 600, 660 and 700 nm. This method relies on the endoscopic Sheffield Index (Sheffield, 1985) to select the bands with the most information. All selected bands have a Full Width at Half Maximum (FWHM) of 20 nm except for the 480 nm band which has a FWHM of 25 nm (by camera design). Recording of one multispectral image at resolution  $2058 \times 2456$  px takes approximately 400 ms. For illumination, we use a standard surgical Xenon light source (SCB 300, KARL STORZ SE & Co. KG, Tuttlingen, Germany). The spectral response of each camera filter is presented in Fig. 2.4. In contrast to the filter response of the snapshot camera presented in Fig. 2.3 c), the Spectrocam does not have second order peaks. This



**Figure 2.3: Optical properties of our multispectral system.** (a)  $\ell_1$  normalized transmission spectra of the laparoscopes, band-pass filter and C-Mount adapter are shown in the left axis. Extinction coefficients of oxyhemoglobin (HbO<sub>2</sub>) and deoxyhemoglobin (Hb) are shown on the right axis. (b) Representation of the 4 × 4 mosaic pattern of the multispectral camera sensor. Each colored square represents a different filter; these filters form a 4 × 4 pattern that extends over the whole image. (c) The filter responses of the multispectral camera bands. Some bands, such as 5 to 12, show two extra peaks in the spectral response, which are referred to as “second order” peaks.



**Figure 2.4: Sequential spectral recordings with glass filters eliminate second order peaks.** Spectral response of the Spectrocam camera filters, no second order peaks are observed. The Spectrocam camera uses a filter wheel, which trades slower recording speed for higher spatial resolution and narrower spectral bands.

is because the Spectrocam uses glass filters integrated in a filter wheel instead of the Fabry-Pérot cavities. Using a filter wheel camera trades the speed of a snapshot sensor for higher spatial resolution and narrower spectral bands, which are features more beneficial for neuroimaging applications.

### 2.2.2 Synthetic spectral imaging

Translational functional imaging requires validation in real-world scenarios where real data is collected in specific medical domains. Nonetheless, the development of new methods for functional imaging can benefit greatly with the introduction of digital replicates that imitate the behavior of real devices (*digital twins*). This holds specially true in applications where real data is scarce or difficult to obtain. In this section, we present the physics-based approach that we have adopted to simulate highly accurate light-tissue interactions. In turn, these simulations are part of the foundation for many of the methods presented in this thesis.

**Monte Carlo simulations of light transport** The generation of simulated diffuse reflectance spectra of biological tissue was performed based on previous work (Wirkert et al., 2017). The light transport across three tissue layers was simulated with a MC method

**Table 2.1:** The simulated ranges of physiological parameters, and their usage in the simulation setup. The properties of each layer of the three-layered tissue models were assigned by uniformly sampling within the ranges specified in this table. Here  $v_{\text{HbT}}$  represents the blood volume fraction,  $s$  the blood oxygen saturation,  $a_{\text{mie}}$  the reduced scattering coefficient at 500 nm,  $b_{\text{mie}}$  the scattering power,  $g$  the tissue anisotropy,  $n$  the refractive index,  $d$  the tissue thickness and  $\epsilon$  the extinction coefficients of oxyhemoglobin (HbO<sub>2</sub>) and deoxyhemoglobin (HbO). Notice that all parameters were uniformly sampled within the specified range for each layer independently, which leads to different values across layers.

$v_{\text{HbT}}$ [%]	$s$ [%]	$a_{\text{mie}}$ [cm <sup>-1</sup> ]	$b_{\text{mie}}$ [a.u.]	$g$ [a.u.]	$n$ [a.u.]	$d$ [cm]
0–30	0–100	5–50	0.3–3	0.80–0.95	1.33–1.54	0.002–0.2
simulation framework: GPU-MCML (Alerstam et al., 2010), 10 <sup>6</sup> photons per simulation						
simulated samples: 5.5 × 10 <sup>5</sup>						
sample wavelength range [ $\lambda_{\text{min}}$ – $\lambda_{\text{max}}$ ]: 300–1000 nm, step size 2 nm						

and a model composed of infinitely wide slabs. Each of these slabs was defined by its optical and physiological properties: blood volume fraction  $v_{\text{HbT}}$ , blood oxygen saturation  $s$ , reduced scattering coefficient at 500 nm  $a_{\text{mie}}$ , scattering power  $b_{\text{mie}}$ , anisotropy  $g$ , refractive index  $n$  and layer thickness  $d$ . These parameters were computed based on literature values (Jacques, 2013) that comprise: extinction coefficients of deoxyhemoglobin  $\epsilon_{\text{Hb}}$  and oxyhemoglobin  $\epsilon_{\text{HbO}_2}$ . The absorption  $\mu_a$  and scattering  $\mu_s$  coefficients were computed using Eq. 2.2. The ranges from which each parameter was uniformly sampled as well as general simulation parameters are summarized in Table 2.1. A Graphics Processing Unit (GPU) accelerated version (Alerstam et al., 2010) of the Monte Carlo Multi-Layered (MCML) simulation framework (Wang and Jacques, 1992b) was chosen to generate the diffuse spectral reflectances.

After simulating the diffuse reflectances  $r(\lambda)$  at wavelength  $\lambda$  on the biological tissue, they were transformed into the MSI camera measurement space at band  $k$ , thus yielding the measurement  $r_k$  according to:

$$r_k = \frac{\int_{\lambda_{\text{min}}}^{\lambda_{\text{max}}} \mathcal{T}(\lambda) \cdot I(\lambda) \cdot F_{k,\text{MSI}}(\lambda) \cdot r(\lambda) d\lambda}{\int_{\lambda_{\text{min}}}^{\lambda_{\text{max}}} \mathcal{T}(\lambda) \cdot I(\lambda) \cdot F_{k,\text{MSI}}(\lambda) d\lambda} \quad (2.1)$$

Here,  $\mathcal{T}(\lambda)$  represents the optical transmission profile of the optical components of the hardware setup (e.g. Fig. 2.3 a)),  $I(\lambda)$  is the relative irradiance of the light source, and  $F_{k,\text{MSI}}(\lambda)$  characterizes the  $k^{\text{th}}$  optical filter response of the camera. The transformed simulated spectra  $r_k$  were used to train the models for oxygenation regression (Sec. 2.4.1) and to pre-train the deep learning-based approach to ischemia monitoring (Sec. 2.4.2).

$$\begin{aligned}\mu_a(v_{\text{HbT}}, s, \lambda) &= v_{\text{HbT}}(s \cdot \epsilon_{\text{HbO}_2}(\lambda) + (1 - s) \cdot \epsilon_{\text{Hb}}(\lambda)) \cdot \frac{\ln(10) \cdot 150 \frac{\text{g}}{\text{L}}}{6.45 \times 10^4 \frac{\text{g}}{\text{mol}}} \\ \mu_s(a_{\text{mie}}, b_{\text{mie}}, \lambda, g) &= \frac{a_{\text{mie}}}{1 - g} \left( \frac{\lambda}{500 \text{ nm}} \right)^{-b_{\text{mie}}}\end{aligned}\quad (2.2)$$

**Multi-layer tissue model labeling** Spectral imaging collects the light that is reflected from the surface of the target object. In medical applications, the target object is biological tissue. As such, the total light collected by a spectral imaging system is composed of photons that have traveled inside the tissue, generally up to a few hundred micrometers for light in the visible range. Deeper tissue locations are accompanied by lower light intensities (fewer photons), as depicted in the light intensity profile in Fig. 2.5. This is an effect arising from the fact that the more distance a photon travels inside the tissue, the more likely it is for that photon to be absorbed. The precise penetration depth  $p(\lambda)$  of a photon depends on the photon wavelength, as well as the tissue composition, and is defined as the depth within the tissue at which the intensity of the incident light  $I(\lambda)$  drops to a ratio of  $1/e$  (Fig. 2.5).

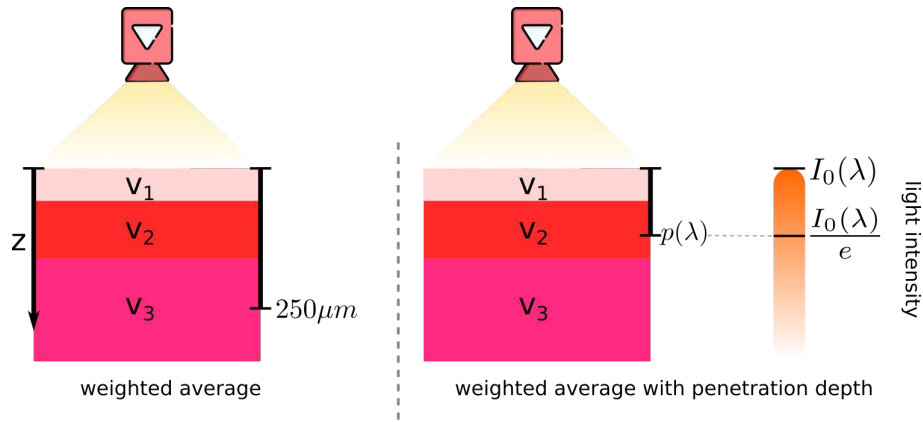
Given that a spectral camera can only assign reflectance values to the tissue surface, any algorithm that aims to leverage this spectral information requires a mechanism to assign tissue properties to the surface of the tissue (e.g. oxygenation). Modelling biological tissue as a multi-layer structure, as opposed to a single layer, benefits from the capability to simulate a wider variety of tissue types. A multi-layer tissue model features stacks of infinite wide slabs, each one with different tissue properties  $v_i$  (Tab. 2.1). Wirkert (2018) proposed to label the entire multi-layer tissue model by assigning one value  $V$  (oxygenation, blood volume fraction, etc.) that is representative across layers. According to Wirkert (2018),  $V$  can be computed with a weighted average to each property up to a depth  $z$  of  $250\mu\text{m}$ :

$$V = \frac{1}{250\mu\text{m}} \int_0^{250\mu\text{m}} v(z) dz \quad (2.3)$$

We hypothesize that a more accurate representation of the tissue parameters of a multi-layer tissue model can be obtained by leveraging the penetration depth of photons at each wavelength  $p(\lambda)$ . As such, the tissue properties  $V$  can be obtained by applying a weighted average across wavelengths in the range  $[\lambda_{\min}, \lambda_{\max}]$ :

$$V = \frac{1}{\lambda_{\max} - \lambda_{\min}} \int_{\lambda_{\min}}^{\lambda_{\max}} \left( \frac{1}{p(\lambda)} \int_0^{p(\lambda)} v(z) dz \right) d\lambda \quad (2.4)$$

Here, the wavelength range  $[\lambda_{\min}, \lambda_{\max}]$  is defined by the range where the imaging



**Figure 2.5: Labelling multi-layer tissue models by using penetration depth can generate more accurate results.** In contrast to a traditional approach that averages each tissue property up to a depth of  $250\mu m$ , we propose to use a wavelength-dependent penetration depth to assign properties to the entire model.

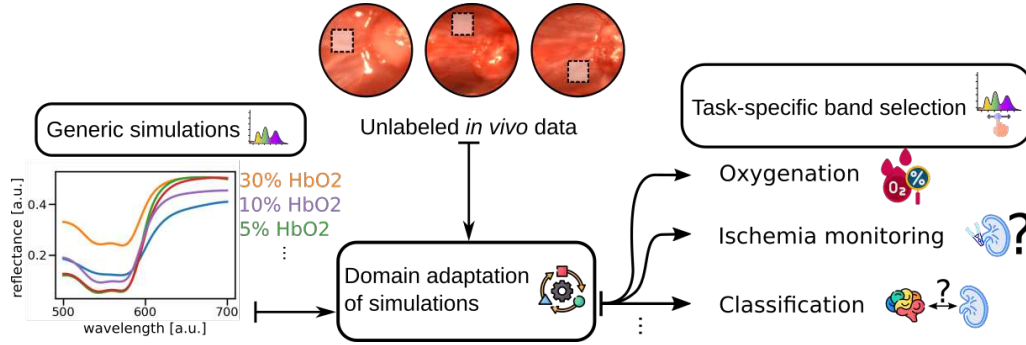
system collects light. This process allows computing each property more precisely by considering how much depth photons travel, in average, inside each tissue model.

### 2.2.3 Task-specific band selection

We propose a method that optimizes a generic model regarding a specific task (e.g. oxygenation monitoring) and domain (e.g. head and neck tumor) without the need for annotated reference measurements. This method features a generic dataset of physics-based simulations, which is leveraged by the proposed domain adaptation technique and thus serves as foundation for the actual band selection algorithms. Fig. 2.6 summarizes our proposed method. In the following paragraphs, we detail the domain adaptation methodology as well as the band selection algorithms that we explored for the task of oxygenation estimation. Note that our concept is not restricted to these specific algorithms, here we exemplify its usability for the task of oxygenation estimation.

**Adaptation to target domain** The simulations presented in the previous paragraph describe a generic tissue, which encompasses a wide variety of tissue types that have hemoglobin as the main light absorber (e.g. brain, kidney, etc.). However, much of the generated data might have little relevance for a specific task, such as cancer localization. This means that bands selected on these generic simulations might be suboptimal for specific domains. Domain adaptation techniques are generally capable of matching simulations with the target domain more closely (Pan and Yang, 2010). We propose





**Figure 2.6: Domain and target-specific simulations can be generated by leveraging unlabeled real spectral data.** Generic simulations are adapted using unlabeled spectral measurements from the target domain. The resulting domain-specific simulations are the basis for the subsequent task-specific band selection.

to use Kernel Mean Matching (KMM) (Huang et al., 2007), as a state-of-the-art domain adaptation technique, to automatically assign a weight to each simulated sample according to how similar they are in comparison to real measurements taken from a target domain. KMM estimates the density ratios between two probability density functions  $\beta(x) = \frac{\rho(x)}{\rho'(x)}$  by minimizing their Maximum Mean Discrepancy (MMD) (Gretton et al., 2009) in a Reproducing Kernel Hilbert Space (RKHS)  $\phi(x) : x \rightarrow \mathfrak{F}$ .

$$\text{MMD}^2(\mathfrak{F}, \beta, \rho, \rho') = \left\| E_{x \sim \rho(x)}[\beta(x)\phi(x)] - E_{x \sim \rho'(x)}[\phi(x)] \right\|^2 \quad (2.5)$$

In the particular application of oxygenation estimation,  $\rho(x)$  represents the distribution of diffuse reflectance simulations  $\mathbf{r}_i$  and  $\rho'(x)$  the distribution of real measurements  $\mathbf{I}_i$ . Thus, the problem of computing the density ratios of these two distributions can be rewritten as follows. For a more complete deduction of this expression, please refer to the work of Miao et al. (2016).

$$\arg \min_{\beta_i} \left\| \frac{1}{n} \sum_{i=1}^n \beta_i \phi(\mathbf{r}_i) - \frac{1}{n'} \sum_{i=1}^{n'} \phi(\mathbf{I}_i) \right\|^2 \quad \text{s.t.} \quad \beta_i \in [0, B] \quad \text{and} \quad \left| \sum_{i=1}^n \beta_i - n \right| \leq n \frac{B}{\sqrt{n}} \quad (2.6)$$

This function correlates the means of simulations  $\mathbf{r}_i$  with real data  $\mathbf{I}_i$  in a RKHS induced by the kernel  $K$ . We set this function to the Gaussian radial basis function kernel:  $\phi(\mathbf{r}_i, \mathbf{r}_j) = e^{-\gamma(\|\mathbf{r}_i - \mathbf{r}_j\|_2)^2}$ . Intuitively, the objective of KMM is to assign weights to Gaussian functions associated with each simulated sample to emulate the distribution of the measurements as closely as possible. The first boundary condition ( $\beta_i \in [0, B]$ ) limits the maximum influence of individual training samples, while the second condition ensures that the term

$\frac{1}{n} \sum_{i=1}^n \beta_i \phi(\mathbf{r}_i)$  is close to a probability distribution (Gretton et al., 2009). Unfortunately, Eq. 2.6 cannot be solved directly due to the possible infinite dimensions of  $\phi$ . To tackle this problem, we use the random kitchen sinks method (Rahimi et al., 2007), which generates a space  $\omega(\mathbf{r}_i)$  that approximates  $\phi(\mathbf{r}_i)$  by sampling from the Fourier transformation of a shift invariant kernel. This method enables using a standard optimizer to solve Eq. 2.6 in its nonkernelized form. The output of this process are the weights  $\beta_i$ , which are used to generate a dataset that resembles the real data  $\mathbf{I}_i$  by sampling with replacement from the original simulated dataset  $\mathbf{r}_i$ .

**Task-specific band selection** The objective of band selection is to select a number of bands required to maintain high performance on a desired task while reducing the amount of data that is needed. We exemplify this procedure in the task of oxygenation estimation. Unless otherwise specified, the metric we employ to assess the performance of different band selection algorithms is Mean Absolute Error (MAE).

For the task of oxygenation estimation, we use the random forest-based method developed by Wirkert et al. (2016b). This method exploits MC simulations to learn the mapping from diffuse reflectances to oxygenation values. Note that for consistency with the feature selection literature, we use the term feature rather than band in this section, but they are equivalent in this context. We studied and compared several popular *filter*- as well as *wrapper*-based approaches, we summarize them in the following paragraphs.

On one hand, *filter* methods generally estimate the amount of information that features provide by computing a specific metric, such as mutual information or conditional mutual information between features and between features and targets. On the other hand, *wrapper* methods select the features that maximize the performance of a model, usually a machine learning model.

For *filter* feature selection, we employ several mutual information-based selection criteria: Minimum Redundancy Maximum Relevance (mRMR) (Peng et al., 2005), Conditional Mutual Information Maximization (CMIM) (Fleuret, 2004), Mutual Information based Feature Selection (MIFS) (Battiti, 1994), Interaction Capping (ICAP) (Jakulin, 2005), Conditional Infomax Feature Extraction (CIFE) (Lin and Tang, 2006) and Joint Mutual Information (JMI) (Yang and Moody, 2000). These methods were originally developed for classification tasks; however, they can be extended to regression by providing a mutual information estimator. For a regression task (oxygenation estimation), such an estimator must compute the conditional mutual information between two feature subsets, even when the target is not categorical. For this purpose, Kraskov's nearest neighbor mutual information estimator (Kraskov et al., 2004) is used to compute the mutual information between subsets. Furthermore, feature subsets are optimized by constructing them in a greedy step-wise manner, as described by Brown et al. (2012). It is worth noting that the

optimal number of features cannot be provided by *filter* feature selection. Instead, these algorithms provide the order with which features should be added to a particular subset. Hence, we report the results for all generated feature set sizes.

As *wrapper* feature selection algorithms, we use Sequential Forward Selection (SFS) and Best First Search (BFS) (Kohavi and John, 1997). Both algorithms generate feature subsets sequentially by optimizing a fitness criterion. SFS iteratively adds single features to a given subset by keeping the feature that most improves the fitness criterion (MAE). Features are added until the criterion has not improved in the last three additions. In contrast, BFS is capable of also removing features from an existing subset and improves the feature search by keeping track of all explored subsets. When BFS is stuck in a local minimum, it returns to a previous feature subset with high performance and resumes the search from that state. We employ the task of oxygenation estimation in all band selection experiments, and leverage the random forest regressor as described by Wirkert et al. (2016b) for this purpose.

$$\arg \min_{s_i \subset S} (l_{\text{MAE}}(y, f(x_{s_i}))) \quad (2.7)$$

Here,  $S$  denotes the complete set of possible features,  $s_i$  are the selected features (a subset of  $S$ ),  $l_{\text{MAE}}(y_i, f(x_{s_i}))$  is the MAE achieved with the subset  $s_i$ .  $f$  represents a regressor that estimates the target variable  $y$  (e.g. oxygenation) based on the features  $s_i$  of sample  $x$ , that is  $x_{s_i}$ .  $l_{\text{MAE}}(y_i, f(x_{s_i}))$  is computed by running a threefold cross-validation on the training data while using only the selected subset of features  $s_i$ . The final subset of features (and thus feature set size) is determined by evaluating the performance of multiple regressors  $f$ , each trained on different feature subsets, on a test dataset and selecting the subset with the minimum MAE.

**Data normalization for band selection** The intensity of the measured diffuse reflectances in real settings is generally affected by multiplicative factors that arise from different measurement settings: e.g. camera angle, camera distance, light source intensity, etc. To mitigate these effects, band normalization is performed. We follow the recommendations from Wirkert et al. (2016b) to normalize each reflectance spectrum by dividing it by its mean, followed by a negative logarithmic transformation ( $-\log$ ) and a further  $\ell_2$  normalization. Given a spectral image  $I \in \mathbb{R}^{N_x \times N_y \times N_s}$  of spatial dimensions  $N_x \times N_y$  and number of bands  $N_s$ , and  $I(i, j, k) \in \mathbb{R}$  represents the intensity value at the pixel position  $(i, j)$  ( $i \in \{1, \dots, N_x\}, j \in \{1, \dots, N_y\}$ ) and on image band  $k$  ( $k \in \{1, \dots, N_s\}$ ). The normalized spectra  $\bar{I}(i, j, k)$  can be computed as follows:

$$\begin{aligned} I^{\log}(i, j, k) &= -\log\left(\frac{I(i, j, k)}{\sum_{k=1}^{N_s} I(i, j, k)/N_s}\right) \\ \bar{I}(i, j, k) &= \frac{I^{\log}(i, j, k)}{\sqrt{\sum_{k=1}^{N_s} (I^{\log}(i, j, k))^2}} \end{aligned} \quad (2.8)$$

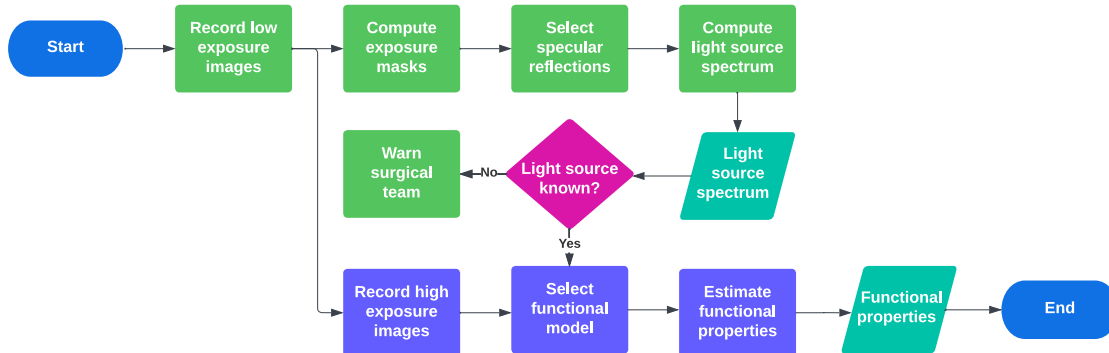
The logarithmic transformation and  $\ell_2$  normalization are not strictly necessary from a theoretical standpoint, but empirically improved results for the task of oxygenation estimation in our band selection experiments. In addition, the data needs to be re-normalized whenever a different set of bands is used because the normalization method works on a per-sample basis and uses all the available bands for normalization. As such, this normalization step theoretically allows our approach to jointly optimize bands and their normalization.

## 2.3 Automatic light source estimation

Estimating the spectrum of the light source (LS) illuminating a surgical scene is one of the key challenges related to translating functional imaging methods to clinical practice. In open surgery, the presence of multiple LSes (e.g. overhead lamps, ceiling lights, etc.) as well as their dynamic nature makes this a challenging problem. In this section, we present our approach to automatic LS estimation during open surgical procedures.

### 2.3.1 Concept overview

Optical functional imaging involves the recording and analysis of relatively high exposure images, which are then processed by an algorithm to estimate a functional property of interest (e.g. oxygenation). Such high exposure images benefit from rich information of light-tissue interactions, which enables the estimation of functional parameters. This type of images also generally features the presence of specular reflections, that is, regions of the tissue that mainly reflect the light originating from the LS without entering the tissue. As such, specular reflections contain mainly information about the spectrum of the LS. Unfortunately, specular reflections are usually present as over saturated regions in high exposure images. We propose to leverage specular reflections, hereon referred to as specular highlights, by recording low exposure images and segmenting specular highlight regions from them (Fig. 2.7).

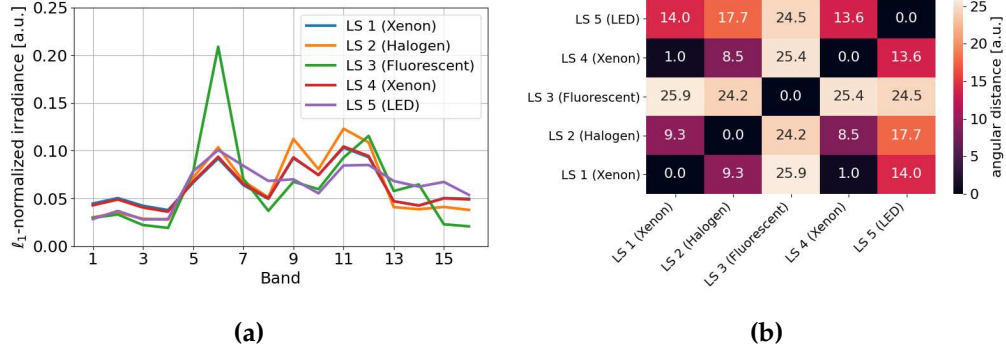


**Figure 2.7:** Our method leverages low exposure images to estimate the spectrum of the light source (LS). Based on low exposure multispectral images, specular highlight masks are computed, which represent regions in the image that are not over- or underexposed and are assumed to contain a dominant specular component. A pre-trained machine learning algorithm tailored to the estimated light source is then applied to high exposure multispectral images to compute functional tissue parameters.

### 2.3.2 Implementation details

Our approach to automatic LS estimation comprises three main steps: a) the recording of low exposure images, b) the segmentation of specular highlight regions based on low exposure recordings and c) the estimation of the LS spectrum based on specular highlight masks.

**Calibration data acquisition** Specular reflections in standard spectral images (i.e. high exposure time) typically saturate the detector, leading to “invalid” pixels. On the other hand, other regions are not well suited to recover the LS spectrum because reflections in those regions are greatly influenced by underlying tissue properties (e.g. oxygenation). We propose to tackle this problem by acquiring dedicated LS calibration images, which are images acquired with a lower exposure time compared to that used for the spectral images that serve as a basis for functional parameter estimation. Even though these images (low exposure) are generally associated with a low Signal-to-Noise Ratio (SNR), they are acquired in a way that “valid” specular highlight pixels (not overexposed and not underexposed) mostly contain the information about the light source. The optimal exposure time for these images is determined based on sequential recordings detailed in Sec. 3.2. To validate our approach, we used five different light sources. The spectra of these are shown in Fig. 2.8 a) and represent some common illumination conditions in the OR. The reference (gold standard) illuminant spectra of all LSes were obtained with a scientific grade spectrometer (HR2000+, Ocean Insight<sup>®</sup>, Largo, USA) over a Spectralon<sup>®</sup>



**Figure 2.8: We use five commonly used types of light sources to validate our method.** Reference light sources (LSes). LS 1: Xenon (D-light P 201337 20 endoscopic light source, Karl Storz GmbH, Tuttlingen, Germany); LS 2: Halogen (Halopar 16 GU10 light bulb, Osram<sup>®</sup>, Munich, Germany); LS 3: Fluorescent light (FLS 11W 2700K fluorescent lamp, Paulmann<sup>®</sup>, Springe Völkse, Germany); LS 4: Xenon (Auto LP 5131 endoscopic light source, Richard Wolf GmbH, Knittlingen, Germany); LS 5: LED (Endolight LED 2.2 endoscopic light source, Richard Wolf GmbH, Knittlingen, Germany). **(a)** Relative irradiance of LS 1-5  $\ell_1$ -normalized. **(b)** Angular distance matrix between the different LSes.

diffuse reflectance standard (SRT-99-050, Labsphere<sup>®</sup>, North Sutton, USA) (Robles-Kelly and Huynh, 2013). We consider the Euclidean angle between two LSes ( $a, b$ ) (or estimated LSes) as a measure of their similarity, as proposed by Khan et al. (2017), which we refer to as *angular error* or *angular distance*:

$$\theta_{ab} = \arccos(\mathbf{L}_a \cdot \mathbf{L}_b), \quad (2.9)$$

where  $\mathbf{L}$  corresponds to the spectra of each light source under study, and  $\theta$  the Euclidean angle between them. The angular distance for the five LSes used in this study ranges from  $1.0^\circ$  (LS1 and LS4; both Xenon) to  $25.9^\circ$  (LS1 and LS3; Xenon and fluorescent) as depicted in Fig. 2.8 b).

**Specular highlight segmentation** Our approach comprises the removal of overexposed and underexposed pixels by selecting pixels with intensities  $I$  in a specific range  $I_{min} < I < I_{max}$ . The minimum intensity  $I_{min}$  is set to the level of dark current for a given exposure time, determined once for each camera. The maximum intensity  $I_{max}$  is set to remove effects arising from the non-linearity in the camera response at high intensities, and is determined according to the camera manufacturer specifications (here:  $I_{max} = 950$ ). Excluding underexposed and overexposed pixels results in a set of pixel indices  $\mathbb{V} \subset \mathbb{N}^2$  corresponding to “valid” pixels. Then, specular highlight pixels are determined by first

computing the *lightness*  $\mathcal{L}(i, j) = \sum_{k=1}^{N_s} \frac{I(i, j, k)}{N_s}$ , where  $N_s$  is the number of camera bands and  $I(i, j, k)$  is the image intensity corresponding to camera band  $k$  at pixel location  $(i, j) \in \mathbb{V}$ . From this “lightness image”, a number  $N_p$  of *highlight pixels* with the highest values of lightness  $\mathcal{L}(i, j)$  are selected. The corresponding indices are represented by  $\mathbb{H} \subset \mathbb{V}$ . Based on an empirical analysis (Sec. 3.2), we set  $N_p = 100$ .

**Light source spectrum estimation** The light source spectrum is computed assuming that the light that is diffusely reflected from the tissue can be neglected in specular highlight pixels. This is a reasonable assumption because specular reflections inherently contain mainly information from the light source. For each  $(i, j) \in \mathbb{H}$ , an estimate of the light source spectrum is computed by  $\ell_1$ -normalizing the intensities across camera bands:

$$L(i, j, k) = \frac{I(i, j, k)}{\sum_{k=1}^{N_s} I(i, j, k)} \quad \text{s.t.} \quad (i, j) \in \mathbb{H}, \quad (2.10)$$

where  $I(i, j)$  is the vector of all camera bands at pixel location  $(i, j)$ , and  $L(i, j, k)$  represents the estimated light source spectrum corresponding to the camera band  $k$ . The final illuminant estimation  $L_k$  in band  $k$  is computed as the mean of all estimations from  $N_p$  individual pixels:

$$L_k = \frac{1}{N_p} \sum_{(i, j)} L(i, j, k) \quad \text{s.t.} \quad (i, j) \in \mathbb{H}. \quad (2.11)$$

### 2.3.3 Hyperparameter settings and baseline

We empirically determined the appropriate values for the two hyperparameters: 1) exposure time  $T_{exp}$  (for the calibration images), and 2) the number of highlight pixels  $N_p$  per image. We performed initial experiments using LS 1-3 (Fig. 2.8). We refer to these LSes as *validation LSes* while we refer to LS 4-5 as *test LSes*. We observed that varying the number of pixels  $N_p$  in the range of 75 – 200 had a negligible influence on performance and thus set  $N_p = 100$ . Furthermore, we found a quality metric  $\mathcal{Q}$ , for which the angular error of illuminant estimations decreases as  $\mathcal{Q}$  increases. This analysis was performed on low exposure images in the validation set (exposure times between 5 ms and 150 ms). We define the quality metric as follows:

$$\mathcal{Q}(T_{exp}) = \text{median}_{(i, j) \in \mathbb{H}(T_{exp})} \left( \frac{\mathcal{L}(i, j, T_{exp})}{\mathcal{L}^d(T_{exp})} \right), \quad (2.12)$$

where  $\overline{\mathcal{L}^d}(T_{exp})$  corresponding to the mean lightness value obtained for dark current images (lights turned off), and  $\mathbb{H}(T_{exp})$  representing the indices of the highlight pixels for exposure time  $T_{exp}$ . Note that  $\mathcal{Q}$  is positive, as  $\overline{\mathcal{L}^d}(T_{exp})$  is small, and does not necessarily increase with exposure time due to overexposed specular highlight pixels. We then propose recording multiple exposure images (5–150 ms, every 5 ms) and selecting the low exposure time  $T_{exp}$  that maximizes  $\mathcal{Q}$ :

$$\hat{T}_{exp} = \arg \max_{T_{exp}} \mathcal{Q}(T_{exp}) \quad (2.13)$$

Note that we also investigated acquiring multiple images of the same  $T_{exp}$  and averaging the corresponding results, but did not find an improvement with this approach.

**Baseline methods** For comparison to the state-of-the-art methods, we compared the performance of our method with four model-based approaches. Model-based methods use basic assumptions on the image formation process to extract the spectrum of the light source directly from (single) images. Khan et al. (2017) identified the four most widely used light source estimation methods on Red-Green-Blue (RGB) images and described how to extend them to spectral images. **M1: Max-RGB** is based on Land’s white patch algorithm (Land, 1977), which states that there is at least one pixel in each image band that produces maximum reflection of the light source. Therefore, combining the maximum reflection from each band, the light source spectrum could be recovered. **M2: Gray-world** assumes that the average value of each band computed over the entire image is achromatic and contains only information about the light source. As such, the light source spectrum can be recovered by combining the average from each band. **M3: Shades-of-gray** is a generalization of **M1** and **M2**, where **M1** is equivalent to using  $\ell_\infty$  normalization on each band, while **M2** is equivalent to using  $\ell_1$  normalization. Given a spectral image, with  $p$  being the order of the  $p$ -norm, the estimated light source spectrum  $L_k$ , at camera band  $k$ , can be derived as follows:

$$\left( \frac{\iint |I_k|^p dx dy}{\iint dx dy} \right)^{1/p} \approx \left( \frac{\sum_{(i,j)}^{N_x, N_y} |I(i, j, k)|^p}{N_x \cdot N_y} \right)^{1/p} \propto L_k, \quad (2.14)$$

where  $I_k$  represents the band  $k$  of a spectral image of spatial dimensions  $N_x \times N_y$ . Here,  $p$  is set to six for **M3**, as suggested by Finlayson and Trezzi (2004). **M4: Gray-edge** is assumed that the average of the reflectance derivative in a scene is achromatic. This can be expressed as follows



$$\left( \frac{\iint |I_k'^\sigma|^p dx dy}{\iint dx dy} \right)^{1/p} \approx \left( \frac{\sum_{(i,j)}^{N_x, N_y} |I^\sigma(i, j, k)|^p}{N_x \cdot N_y} \right)^{1/p} \propto L_k \quad (2.15)$$

$$I_k'^\sigma = \sqrt{(\partial_i I_k^\sigma)^2 + (\partial_j I_k^\sigma)^2}, \quad (2.16)$$

where  $(I_k'^\sigma)$  represents the smoothed derivative of image band  $k$  with a Gaussian filter of standard deviation  $\sigma$ . Following the recommendation of Finlayson and Trezzi (2004),  $p$  is set to six.

## 2.4 Perfusion monitoring

We address the task of perfusion monitoring in two separate modules that deal with different challenges. The first module aims to regress functional properties by exploiting physics based light-tissue interactions generated with a MC method (Sec. 2.4.1). As such, this approach overcomes the lack of labeled *in vivo* data. The second module deals with the task of perfusion monitoring in domains that are not well represented by the MC simulations that the first module requires. This is exacerbated, for example, in domains that feature pathologies such as partial nephrectomy (Sec. 2.4.2).

### 2.4.1 Perfusion monitoring via regression

This section presents the methods corresponding to quantitative oxygenation estimations. First, we present oxygenation estimation via the modified Beer-Lambert law, followed by an approach that relies on oximeter baseline measurements (Kumar et al., 2020). Furthermore, a classical machine learning approach to estimate oxygenation from MC simulations is described, and finally our Convolutional Neural Network (CNN)-based approach to oxygenation estimation is provided.

#### Beer-Lambert regression

The modified Beer-Lambert law (Baker et al., 2014) has been used in many state-of-the-art methods (Clancy et al., 2015; Baker et al., 2014; Jones et al., 2017). In this work, we use it as a baseline to estimate the total blood volume concentration ( $v_{\text{HbT}}$ ) and oxygenation ( $s$ ) from spectral data. This law states that a linear relationship exists between the absorption  $a(\lambda)$  of a media, its attenuation coefficient  $\mu_a(\lambda)$  and the optical path length  $l$  of photons within such media:

$$a(\lambda) \sim -\log(r(\lambda)) = \mu_a(\lambda)l + g \quad (2.17)$$

where  $r$  represents diffuse reflectance, and  $g$  accounts for scattering losses. Considering that blood is the main absorber in internal organs, we can replace the attenuation coefficients by the extinction coefficients of blood

$$a(\lambda) = v_{\text{HbO}_2} \cdot \epsilon_{\text{HbO}_2}(\lambda) \cdot l + v_{\text{Hb}} \cdot \epsilon_{\text{Hb}}(\lambda) \cdot l + g, \quad (2.18)$$

where  $v_{\text{HbO}_2}$  and  $v_{\text{Hb}}$  are the concentrations of oxyhemoglobin (HbO<sub>2</sub>) and deoxyhemoglobin (Hb), and  $\epsilon$  represents their corresponding extinction coefficients. Given measurements at multiple wavelengths, equation (2.18) describes a system of linear equations, which can be solved for  $v_{\text{HbO}_2} \cdot l$  and  $v_{\text{Hb}} \cdot l$  using ordinary least squares regression. Considering multiple measurements  $X$ , this equation can be transformed into a matrix multiplication:

$$\begin{aligned} H &= \begin{bmatrix} \epsilon_{\text{HbO}_2}(\lambda_1) & \epsilon_{\text{Hb}}(\lambda_1) & 1 \\ \epsilon_{\text{HbO}_2}(\lambda_2) & \epsilon_{\text{Hb}}(\lambda_2) & 1 \\ \vdots & \vdots & \vdots \end{bmatrix} \\ X &= \begin{bmatrix} v_{1,\text{HbO}_2} \cdot l & v_{2,\text{HbO}_2} \cdot l & \dots \\ v_{1,\text{Hb}} \cdot l & v_{2,\text{Hb}} \cdot l & \dots \\ g_1 & g_2 & \dots \end{bmatrix} \\ A &= \begin{bmatrix} a_1(\lambda_1) & a_2(\lambda_1) & \dots \\ a_1(\lambda_2) & a_2(\lambda_2) & \dots \\ \vdots & \vdots & \vdots \end{bmatrix} = HX, \end{aligned} \quad (2.19)$$

where  $\{\lambda_1, \lambda_2, \dots\}$  represents the wavelengths of an imaging system (e.g. a spectral camera),  $H$  is the matrix composed of the extinction coefficients,  $X$  is composed of the concentrations  $v_{i,\text{HbO}_2}$  and  $v_{i,\text{Hb}}$  for each sample  $i$ , and  $A$  is the matrix of absorptions for each sample and wavelength. By leveraging linear least square regression, the optimal solution  $\hat{X}$  can be obtained by:

$$\hat{X} = (H^T H)^{-1} H^T A \quad (2.20)$$

$\hat{X}$  thus enables estimating oxygenation  $s$  and total hemoglobin concentration  $v_{\text{HbT}}$  for each sample  $i$  as follows:

$$\begin{aligned} s_i &= \frac{v_{i,\text{HbO}_2} \cdot l}{v_{i,\text{Hb}} \cdot l + v_{i,\text{HbO}_2} \cdot l} = \frac{\hat{x}_{i,1}}{\hat{x}_{i,1} + \hat{x}_{i,2}} \\ v_{i,\text{HbT}} \cdot l &= v_{i,\text{HbO}_2} \cdot l + v_{i,\text{Hb}} \cdot l = \hat{x}_{i,1} + \hat{x}_{i,2} \end{aligned} \quad (2.21)$$

Note that estimations of  $v_{HbT}$  can only be computed up to a constant factor  $l$ , which is the optical path length. Since all extinction coefficients are given at a high spectral resolution, but we use a multispectral sensor with a relatively low number of bands, we adapted the high-resolution extinction coefficients by averaging them over the filter response function of each band. Given that we only have access to camera reflectance measurements  $r_k$ , we approximated  $\epsilon_{Hb}$  and  $\epsilon_{HbO_2}$  at each camera band wavelength by integrating over the wavelength range where the camera captures light  $[\lambda_{min}, \lambda_{max}] = 400\text{--}700\text{ nm}$ , thus yielding:

$$\epsilon_{Hb}(\lambda_k) \sim \frac{1}{\lambda_{max} - \lambda_{min}} \int_{\lambda_{min}}^{\lambda_{max}} F_{k,MSI}(\lambda) \cdot \epsilon_{Hb}(\lambda) d\lambda \quad | \quad k \in \{1, 2, \dots, N_s\} \quad (2.22)$$

$$\epsilon_{HbO_2}(\lambda_k) \sim \frac{1}{\lambda_{max} - \lambda_{min}} \int_{\lambda_{min}}^{\lambda_{max}} F_{k,MSI}(\lambda) \cdot \epsilon_{HbO_2}(\lambda) d\lambda \quad | \quad k \in \{1, 2, \dots, N_s\}, \quad (2.23)$$

where  $F_{k,MSI}$  represents the  $k^{\text{th}}$  filter response of a spectral camera.

### Oximeter-based regression

We use the method proposed by Kumar et al. (2020) for baseline comparison. This method leverages sequences of single band images with high temporal resolution coupled with a pulse oximeter to analyze perfusion dynamic in living tissue. It is worth mentioning that this method was designed to monitor perfusion changes in human skin. Following the notation introduced in Eq. 2.8, the perfusion-related signal  $\mathcal{P}$  at each image location  $(i, j)$  and time point  $t$  is extracted by averaging the sequence of images over a  $(2K + 1) \times (2K + 1)$  spatial region:

$$\mathcal{P}(i, j, t) = \frac{1}{(2K + 1)^2} \sum_{l_2=-K}^K \sum_{l_1=-K}^K I(i + l_1, j + l_2, t) \quad (2.24)$$

Then, the cardiac-related changes in the signal  $I$  are extracted by applying a band-pass filter that eliminates the signal outside the frequency range  $0.5\text{--}5\text{ Hz}^\dagger$ . To make the perfusion signal invariant to the intensity of the incident light, the band-passed signal is normalized with the original signal  $I$  after applying a low-pass filter with a cutoff frequency of  $0.3\text{ Hz}$ :

$$\tilde{\mathcal{P}}(i, j, t) = \frac{\Omega_{0.3-5}[\mathcal{P}(i, j, t)]}{\Omega_{0.3-\infty}[\mathcal{P}(i, j, t)]} \quad (2.25)$$

<sup>†</sup>This is the frequency range of the heart beat.

where  $\tilde{\mathcal{P}}$  is the normalized signal and  $\Omega_{f_1-f_2}$  represents a filtering operation that eliminates signals outside the frequency range between  $f_1$  and  $f_2$ . It is then assumed that the normalized perfusion signal can be separated into two components: a) the ratio of pulsatile perfusion amplitude  $q$  to the averaged surface reflectance  $\iota$  and b) a camera noise model  $n(i, j, t)$ :

$$\tilde{\mathcal{P}}(i, j, t) = Q(i, j, t)\varphi(t) + \tilde{n}(i, j, t); \quad Q(i, j, t) = \frac{q(i, j, t)}{\iota(i, j, t)}, \quad (2.26)$$

where  $\varphi(t)$  corresponds to the pulse signal, and  $\tilde{n}$  is the camera noise after applying the same normalization step presented in Eq.2.25. Based on this model, the term  $Q$  can be computed based on a sequence of images spanning on the time frame  $[t - \delta, t + \delta]$ :

$$Q(i, j, t) = \arg \min_q \sum_{\tau=t-\delta}^{t+\delta} \varrho(\tilde{\mathcal{P}}(i, j, \tau) - q\varphi(\tau)), \quad (2.27)$$

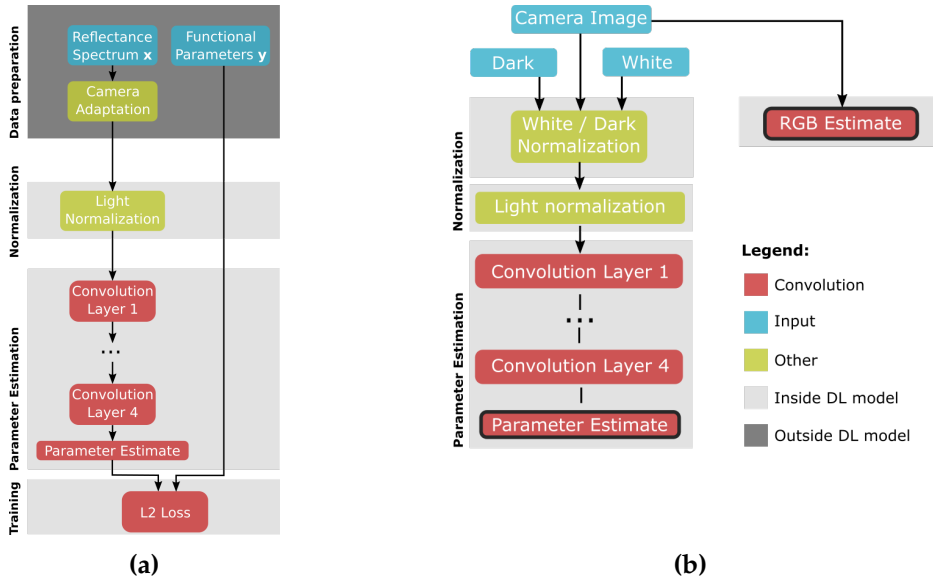
where  $\varrho$  is the Tukey's bi-square loss function with a  $k$ -value of 4.685, which assigns zero weights to outlier samples such as motion artifacts and illumination changes. The pulse signal  $\varphi$  is then obtained by using a pulse oximeter. For more details on the deduction of Eq. 2.27 please refer to the original publication from Kumar et al. (2020), who proposed to use the signal  $Q(i, j, t)$  as a measure to monitor perfusion changes from single-band images.

### Random forest regression

The task of oxygenation estimation via machine learning based on MC simulations was successfully applied by Wirkert et al. (2016b). We adopted this methodology as a baseline and for validation of several of our experiments. As proposed by Wirkert et al. (2016b), we train a random forest regressor on physics-based simulations of light transport in biological tissue (Tab. 2.1). In accordance with Wirkert et al. (2016b), we set the number of trees to ten, the maximum depth per tree to nine, and the minimum number of samples per leaf to ten. In addition, the maximum number of features was set to the number of features of the input samples. As an objective function during training, we use the mean squared error. All random forests presented in our methodology are trained on datasets composed of MC simulations of light transport.

### CNN-based regression

We build upon the work of Wirkert (2018) to generate an end-to-end deep learning pipeline for oxygenation estimation. Our work is based on the hypothesis that (1) deep learning-



**Figure 2.9: Our CNN-based model integrates data preprocessing (e.g. normalization) in the network architecture. (a) The network used for training, whose model parameters are re-used for the live validation network. (b) The live validation network used during the intervention, yielding the functional parameters and Red-Green-Blue (RGB) estimates. The estimated RGB image is only used for visualization.**

based algorithms are better suited than model-based approaches (e.g. Beer-Lambert approaches) for solving the inverse problem of relating diffuse reflectance spectra to the underlying tissue functional parameters, (2) Deep learning approaches can exhibit high accuracy and speed. As shown in Fig. 2.9, the training network learns the functional property estimation based on simulated training data. The weights and biases of the training network are then transferred to the live validation network, which runs in a highly accelerated environment. One advantage of this network design over other state-of-the-art methods is that the necessary data preprocessing steps are integrated into the network architecture and can efficiently be performed on the Graphics Processing Unit (GPU).

We leverage the physics-based simulations presented in Tab. 2.1 to train our network. The diffuse reflectance is an intrinsic property of biological tissue, independent of the imaging system employed during measurements. Nonetheless, recorded reflectances are heavily influenced by imaging parts such as camera filter responses, camera quantum efficiency, and relative irradiance of light source. Thus, transforming the simulated data into a space that resembles the measurement conditions is an important step in our processing pipeline. To achieve this, we use the transformation presented in Eq. 2.1.

In a real scenario, recordings can be affected by changes in illumination, such as illumination intensity and distance of light source to the imaged region. A light normalization step is introduced in the network design to partially reduce this effect (Fig. 2.9). This step is computed on the GPU and comprises normalizing the spectra with its  $\ell_1$  norm, the same normalization is then implemented in the *in vivo* network (Fig. 2.9 b)). To consider possible tissue spatial heterogeneity that can arise in a real scenario, every reflectance training sample is augmented to match a region of  $3 \times 3$  px and Gaussian multiplicative noise at a level of 20% SNR is added to each pixel independently. Both the training and *in vivo* validation network consist of a fully convolutional network architecture. All convolutional layers (1 to 4) are identical, except for the kernel size: the first two convolutional layers have a kernel size of two while the other two have a kernel size of one. The *Parameter Estimate* layer is a sigmoid transformation to scale each physiological parameter to the range  $[0, 1]$  before computing the loss.

For *in vivo* applications (measurements), the adaptation to the optical system is not needed. Instead, images are normalized with measurements of a white reference standard and a dark measurement. Each multispectral recording is then normalized as follows:  $R = \frac{I-D}{W-D}$  where  $I$  represents the raw multispectral image and  $R$  the normalized image. This step is required only once at the beginning of the recordings. Furthermore, the  $\ell_1$  normalization is implemented as part of the deep learning framework. Both the white and dark as well as the  $\ell_1$  normalization steps are integrated in the network architecture to speed up the regression estimation (Fig. 2.9 b)).

#### 2.4.2 Perfusion monitoring via Out-of-Distribution detection (Clinical study)

Laparoscopic surgery presents many challenges regarding ischemia monitoring, mainly high organ motion and tissue heterogeneity. To counteract the effect of organ motion, we present our deep learning-based approach to automatic Region of Interest (ROI) tracking from spectral videos. We then propose a new methodology based on OoD detection for ischemia monitoring in laparoscopic surgery that relies solely on data from individual patients, thus removing the challenges imposed by inter-patient tissue heterogeneity.

##### Region of interest (ROI) tracking

Intraoperative image sequences generally depict highly dynamic scenes with a lot of organ motion. Reliable tracking of ROIs is highly desired in such scenarios. We developed an automated method to track ROIs in spectral image sequences. In this section, we first describe the process of reconstructing RGB images from spectral data, the RGB data then serves as input for our deep learning-based tracker. Finally, we present the post-processing steps we followed to ensure a fair comparison across different ROIs.

**RGB image reconstruction** The reconstruction of RGB images from spectral data was required for (1) automatic ROI tracking, (2) comparison between our approaches that leverage spectral data and the ones using RGB data, and (3) illustration purposes. Such transformation was obtained by computing a transformation matrix  $T$ . A linear regressor was used to compute  $T$  based on the filter response of the multispectral camera (Fig. 2.3), the transmission of each optical component (Fig. 2.3) and the filter responses of an artificial RGB camera. The artificial RGB camera was simulated by three Gaussian filters centered at 460 nm, 550 nm and 640 nm, each with a standard deviation of 42 nm. The spectral response of the multispectral camera was adjusted by multiplying the transmission  $\mathcal{T}$  of each optical component by the matrix  $F_{MSI}$  of filter responses of the multispectral camera.

$$F' = F_{MSI}\mathcal{T} \quad (2.28)$$

By considering the filter response of the artificial camera  $F_{RGB}$  and the spectral response of the multispectral camera  $F'$ , after correcting for the optical components integrated in the system, the transformation matrix can be computed with:

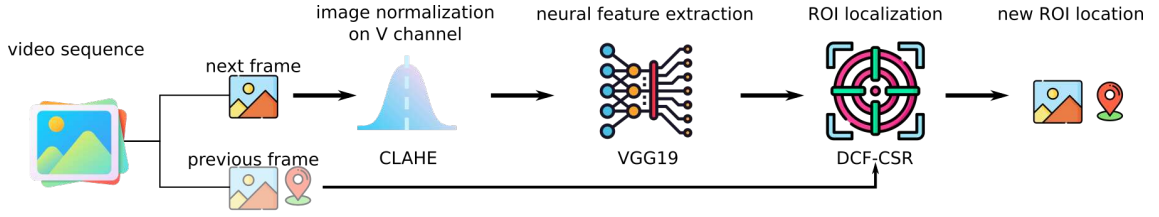
$$T_{\min} = \arg \min_{T'} (\| F_{RGB} - F' T' \|_2) \quad (2.29)$$

$$T = T_{\min} \oslash \mathcal{N}$$

where  $T$  represents the coefficients of the linear regressor,  $\oslash$  corresponds to component-wise division, and  $\mathcal{N}$  is a normalization vector with three elements, which can be computed as:

$$\mathcal{N}_\alpha = \sum_{k=1}^{N_s} (F' T_{\min})_{\alpha,k} \quad s.t. \quad \alpha \in \{1, 2, 3\} \quad (2.30)$$

**Deep learning-based tracker** Our tracker leverages spectral images, which are transformed to RGB space as described in the previous paragraph. To enhance the robustness of the tracker, the RGB images are first transformed to Hue-Saturation-Value (HSV) color space, the V channel is then normalized with Contrast Limited Adaptive Histogram Equalization (CLAHE), and the resulting image is transformed back to RGB. CLAHE enhances the contrast of an image by dividing it into smaller contextual rectangular regions and normalizing the histogram corresponding to each of these regions. Bilinear interpolation is used in the border of each contextual region to avoid visible boundary effects. In contrast to Adaptive Histogram Equalization (AHE), CLAHE clips the histogram at a certain limit and the exceeding pixels are distributed uniformly over the entire histogram. Our tracker sets the kernel size used by CLAHE to 1/8 of the image height by 1/8 of its width, thus



**Figure 2.10: We track tissue Regions of Interest (ROIs) with high accuracy in spectral videos.** Our tracker analyzes two consecutive images of a spectral video. The RGB images are transformed to HSV color space and each V channel is normalized using Contrast Limited Adaptive Histogram Equalization (CLAHE), the result is then transformed back to RGB space. The normalized images are fed to a VGG19 and features are extracted from its seventh convolutional layer. These features are then used by the Discriminative Correlation Filter with Channel and Spatial Reliability (DCF-CSR) tracker to generate the new ROI position, which in turn is used to locate the ROI position in subsequent images. *Icons by FreePik from flaticon.com*

resulting in 64 rectangular contextual regions. The number of bins used for histogram normalization is set to 256 and the clipping limit to 0.01.

The reconstructed and normalized RGB images are then fed into a pre-trained VGG19 neural network (Karen and Andrew (2015)) and deep features are extracted from its seventh convolutional layer. The extracted features are further processed by the tracker Discriminative Correlation Filter with Channel and Spatial Reliability (DCF-CSR) (Lukezic et al. (2018))

This process thus generates a sequence of automatically tracked ROIs across a sequence of spectral images. Fig. 2.10 illustrates the different steps of the proposed tracker.

**ROI data post-processing** The data from each ROI was first normalized with a white ( $W$ ) and dark ( $D$ ) reference recording taken with a highly reflective target (Spectralon<sup>®</sup>, Edmund Optics, Barrington, USA). Given an ROI of dimensions  $N_x \times N_y \times N_s$ , where  $(N_x, N_y)$  are the spatial dimensions and  $N_s$  is the number of spectral bands, the intensity  $I$  of each pixel at spatial location  $(i, j)$  and spectral band  $k$  was normalized according to:

$$I_n(i, j, k) = \frac{I(i, j, k) - D(i, j, k)}{W(i, j, k) - D(i, j, k)} \quad (2.31)$$

In addition, an  $\ell_2$  normalization across different bands was performed to compensate for the influence of light source intensity changes due to changes in the distance of the camera



to the surface of the kidney.

$$\bar{I}(i, j, k) = \frac{I_n(i, j, k)}{\sqrt{\sum_k^{N_s} I_n^2(i, j, k)}} \quad (2.32)$$

The resulting spectra  $\bar{I}(i, j, k)$  can then be compared between different image sequences, and were used to train our deep learning model for ischemia monitoring. The median normalized spectra within each ROI were used for the rest of our analysis.

### Statistical analysis of tissue heterogeneity

For 2D visualization of the high-dimensional spectral data, the data from each ROI was first normalized according to the procedure described in Sec. 2.4.2 yielding a median spectrum per ROI, image, tissue state (perfused *vs.* ischemic) and subject. For each image, we averaged the median spectra across all ROIs, computed the first two principal components based on the data from all subjects and projected the data points onto these new axes (Hotelling, 1933).

We further computed the proportion of explained variability in reflectance by different components using linear mixed models, as suggested by Schreck (2019). Given a patient index  $m \in \{1, 2, \dots, N_p\}$ , and an image index  $n \in \{1, 2, \dots, N_i\}$ , we fitted the following model for each wavelength separately:

$$r_{mn} = \Gamma + \Lambda S_{mn} + \Delta_m + \Psi_{mn} \quad (2.33)$$

where  $r$  represents the averaged median reflectance across tracked ROIs for a given wavelength.  $S_{mn}$  is a variable that indicates the perfusion state of the kidney (1 for ischemic or 0 for perfused). Furthermore,  $\Gamma$  represents an intercept for each linear model,  $\Lambda$  is a fixed state effect,  $\Delta_m \sim \mathcal{N}(0, \sigma_\Delta^2)$  is a random patient effect and  $\Psi_{mn} \sim \mathcal{N}(0, \sigma_\Psi^2)$  are residuals, with respective variances  $\sigma_\Delta^2$  and  $\sigma_\Psi^2$  that are assumed to be independently normally distributed. We computed the proportions of explained variability following the recommendations of Schreck (2019).

### Deep learning-based ischemia monitoring via OoD

We re-phrase the task of ischemia monitoring as an OoD detection problem to leverage data from individual patients. This removes the need for large amounts of annotated training data and avoids the problem of limited robustness, which traditional machine and deep learning approaches suffer from. In particular, our algorithm is trained to compute the likelihood of perfused tissue based on an image sequence acquired at the beginning of

each surgery. In the following paragraphs, we review the principle of Widely Applicable Information Criterion (WAIC) (Watanabe, 2013), explain how to leverage Invertible Neural Networks (INNs) to compute it in real time and present our OoD-based *ischemia index*.

**Widely Applicable Information Criterion (WAIC)** The topic of OoD has benefited from increasing interest in the machine and deep learning community over the last years. However, we are not aware of any previous work in OoD detection in the field of optical imaging. We build our ischemia monitoring method upon the work by Choi et al. (2018) who proposed WAIC as a means to measure the closeness of a new data sample to the distribution of the training data. Watanabe (2013) defined WAIC as:

$$\text{WAIC}(x) = \text{Var}_{\Theta}[\log \nu(x | \Theta)] - \mathbb{E}_{\Theta}[\log \nu(x | \Theta)], \quad (2.34)$$

where  $\Theta$  in general represents the parameters of the model that parametrize  $p$ . The variance and expectation are computed over such parameters  $\Theta \sim \nu(\Theta | X^{\text{tr}})$ , where  $X^{\text{tr}}$  is the training data. Thus,  $\text{WAIC}(x)$  approximates the distance of any sample  $x$  to the training data  $X^{\text{tr}}$ . The variance term  $\text{Var}_{\Theta}[\log \nu(x | \Theta)]$  in Eq. (2.34) implies that the WAIC measure indicates more certainty of a sample  $x$  being in-distribution the closer the sample is to the training distribution. On the other hand, the expectation term  $\mathbb{E}_{\Theta}[\log \nu(x | \Theta)]$  implies that if the sample  $x$  lies in a low density region, then it is expected to be OoD. One of the biggest challenges associated with the implementation of WAIC is the efficient computation of  $\log \nu(x | \Theta)$ . We propose applying INNs to be able to compute  $\log \nu(x | \Theta)$  efficiently and thus leverage  $\text{WAIC}(x)$  in real time during surgical procedures.

**Invertible Neural Networks (INNs) for computing WAIC** We propose to compute  $\log \nu(x | \Theta)$  in two steps: (1) use INNs to convert samples  $x$  into a space that allows the analytical computation of the log-likelihood and then (2) apply the change of variable formula to obtain  $\log \nu(x | \Theta)$ . In more detail, we use INNs (Ardizzone et al., 2018) to transform the spectra (that lies in image space  $X$ ) to a latent space  $Z$ . The latent space  $Z$  is constructed in a way that samples are distributed according to a multivariate standard Gaussian. By denoting the neural network with  $f_{\Theta}: X \subset \mathbb{R}^n \rightarrow Z \subset \mathbb{R}^n$  (with parameters  $\Theta$ ), we can use the change of variable formula to compute the log-likelihood  $\log \nu(x | \Theta)$  for a sample  $x$  as

$$\log \nu(x | \Theta) = -\frac{1}{2} \|f_{\Theta}(x)\|_2^2 - \frac{n}{2} \log(2\pi) + \log |\det Jf_{\Theta}(x)|, \quad (2.35)$$

where  $Jf_{\Theta}$  denotes the network's Jacobian.

To compute the WAIC variance and expectation terms from Eq. (2.34), we generate an

ensemble of networks (default:  $n = 5$ ) with identical architecture but different random seeds. The INNs were implemented using the PyTorch framework and the FrEIA package for the architecture of each INN. Following up on previous work with INNs (Ardizzone et al., 2018; Adler et al., 2019b,a), we applied the *normalizing flow* architecture originally introduced by Dinh et al. (2016) and used the following network default settings: 20 affine coupling blocks (Dinh et al., 2016) with three layer fully connected subnetworks with 256 hidden dimensions, rectified linear unit (ReLU) activations and fixed channel permutations. The INNs were trained by minimizing the Maximum-Likelihood loss  $\mathcal{L}(x) = -\log v(x | \Theta)$  (Eq. 2.35) with the Adam optimizer (Kingma and Ba, 2014b), a learning rate of  $1 \cdot 10^{-4}$  and weight decay of  $1 \cdot 10^{-4}$ . The training data was first z-score normalized, and Gaussian noise was subsequently added with a standard deviation of 0.05.

**Ischemia index** We envision our approach to be applied in clinical practice, which implies that not only the inference but also the training needs to be performed during the actual surgical procedure. Thus, the requirement of fast training time had to be met (seconds rather than minutes or hours). To facilitate fast network convergence based on the patient data, we pre-trained our INN ensemble on simulated data. We leveraged the MC method presented in 2.2.2 to compute high-resolution spectra covering a large space of optical properties that resemble a wide range of physiological states (Tab. 2.1). These high-resolution spectra were then adapted to the imaging system shown in Fig. 2.2. Our proposed approach can leverage spectral information from multiple ROIs. To compute our *ischemia index* from this information, we first aggregate WAIC values belonging to the same ROI in the multispectral image at time point  $t$ ,

$$\text{WAIC}(\text{ROI}(t)) := \text{median WAIC}(\bar{I}(i, j, t)) \quad \text{s.t.} \quad i \in \{1, \dots, N\}; j \in \{1, \dots, M\}, \quad (2.36)$$

where  $\bar{I}(i, j)$  is the image intensity after normalization (Eq. 2.32) at spatial location  $(i, j)$  within the dimensions of the ROI of shape  $N \times M$ . Then, we aggregate all ROIs per image via the mean to obtain the final *ischemia index*:

$$\text{ischemia index}(t) := \frac{1}{N_r} \sum_{k=1}^{N_r} \text{WAIC}(\text{ROI}_k(t)), \quad (2.37)$$

where  $N_r$  represents the number of ROIs. For ease of visualization, we min-max-normalized the *ischemia index* for each patient individually. As this is a strictly monotonic transformation, this has no influence on the Area Under the Receiver Operating Curve (AU-ROC) metric.



*To be yourself in a world that is constantly trying to make you something else is the greatest accomplishment.*

— RALPH WALDO EMERSON

This chapter contains the experiments designed to test the concept of functional imaging in surgery that was presented in Fig. 2.1. This mainly covers the topic of synthetic data generation and task-specific band selection (Sec. 3.1). The experiments and results corresponding to the dedicated automatic light source estimation module are presented in Sec. 3.2. Finally, the task of functional parameter estimation (here perfusion monitoring) is divided into two modules, one dealing with the lack of labeled *in vivo* data (Sec. 3.3.1), and the other with domains that feature a high inter-patient tissue variability and are not well represented by the simulations of light-tissue interactions (Sec. 3.3.2).

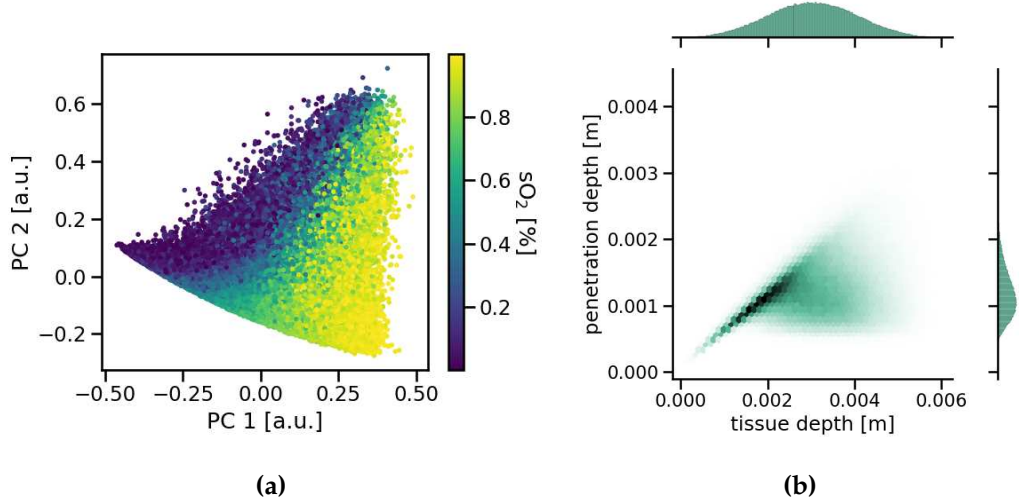
## 3.1 Spectral imaging

The first module in our concept for translational functional imaging (Fig. 2.1) aims to increase the recording and processing speed of optical imaging systems by reducing the amount of data that they need to acquire. We approached this challenge with a domain- and task-specific band selection method capable of leveraging unlabeled *in vivo* data and simulations of light-tissue interactions. In this section, we first present the results of our proposed method for multi-layer tissue model labeling (Sec. 3.1.1) followed by our band selection approach (Sec. 3.1.2).

### 3.1.1 Synthetic spectral imaging

**Research question :** What is the effect of different labeling strategies of a multi-layered tissue model on oxygenation estimation?

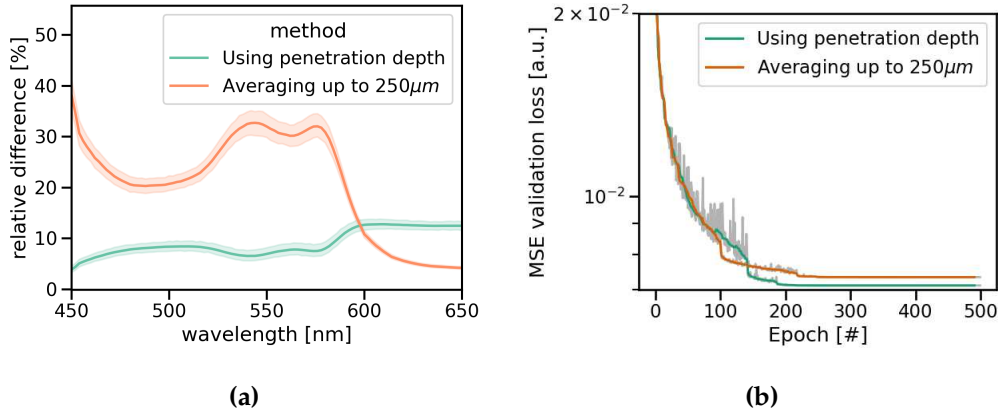
In Eq. 2.3 and Eq. 2.4 we presented two approaches to multi-layer tissue model labeling. The first one assigns labels by averaging the parameter values (e.g. oxygenation) across layers up to a depth of  $250\mu m$ , the later one leverages the penetration depth of photons to average each parameter. We computed the penetration depth at simulation time with a



**Figure 3.1: The mean penetration depth is smaller than the total penetration depth of each multi-layer tissue model. (a)** Two-dimensional representation computed via Principal Component Analysis (PCA) of the complete set of  $5 \times 10^6$  simulations of light-tissue interactions. Each dot represents one simulation and its color represents the tissue oxygenation. **(b)** Distribution of multi-layer tissue model depth and mean light penetration depth (across wavelengths). The mean light penetration depth is usually smaller than the tissue depth.

modified version of the GPU-MCML framework (Alerstam et al., 2010). The penetration depth during Monte Carlo (MC) simulations was obtained by tracking the absorption of photons within each tissue model. Thus, the penetration depth  $p(\lambda)$  was computed as the depth at which the number of photons (not absorbed) dropped to  $1/e$  of its value at the air-tissue interface. The two-dimensional representation, computed via Principal Component Analysis (PCA), of the simulated diffuse reflectances corresponding to  $5 \times 10^6$  multi-layer tissue models is shown in Fig. 3.1 a). The simulated diffuse reflectances were normalized before the computation of the PCA analysis according to Eq. 2.32. Each point represents one multi-layer tissue model and its color corresponds to the tissue oxygenation computed based on Eq. 2.4. Fig. 3.1 b) depicts the distributions of each tissue depth and its corresponding light penetration depth. This indicates that the light penetration depth, for the majority of tissue models, is smaller than its total depth.

To evaluate our approach, we created one-layer tissue models and assigned them the tissue properties extracted from the three-layer models (Tab. 2.1) while using our approach (penetration depth) and the reference method (Eq. 2.3 and Eq. 2.4). The one-layer models were then re-simulated and the resulting diffuse reflectances were compared



**Figure 3.2: Using the penetration depth for multi-layer model labeling assigns properties that are more characteristic of each tissue model. (a)** Relative difference between three-layer tissue models and re-simulated one-layer models for which tissue properties were assigned by averaging each parameter up to 250  $\mu\text{m}$  (orange line) and using the penetration depth for averaging each parameter (green line). **(b)** Mean Squared Error (MSE) validation loss during training of our Convolutional Neural Network (CNN) model while using the two labeling strategies. Using the penetration depth for model labeling helps during model training by producing a lower validation loss. Model training was done based on the three-layer tissue models.

to the ones obtained with the three-layer model. The objective of each labeling strategy is to assign properties that accurately represent the multi-layer tissue models. Therefore, a more appropriate labeling strategy would generate reflectances (from a one-layer model) more similar to the reflectances from the three-layer model. Due to the computational complexity of re-simulating all 500,000 tissue models described in Tab. 2.1, we selected 1,000 random samples to perform this analysis. Fig. 3.2 a) shows the relative difference between the reflectances from the three-layer models and the reflectances from the one-layer models while using each of the labeling strategies. The translucent band around each line corresponds to the standard deviation across all 1,000 re-simulated tissue models. In the wavelength range 450–600 nm we observed that the relative distance is considerably lower while using the penetration depth as labeling strategy. Above 600 nm, the relative distance is slightly higher when using the penetration depth.

Fig. 3.2 b) shows the MSE validation loss during training of our CNN model for oxygenation estimation (Fig. 2.9) while using both labeling strategies. The gray lines correspond to the raw validation losses, and the colored lines correspond to the validation losses after smoothing them with a rolling window of size 20. We observed a slightly lower validation loss for the model trained while using the penetration depth as a labeling

strategy for each tissue model. However, comparison of these two models is challenging because they were trained on different labels: same underlying reflectances but different labeling strategies.

### 3.1.2 Task-specific band selection

**Research question :** How can unlabeled data be leveraged to generate domain-specific MC simulations?

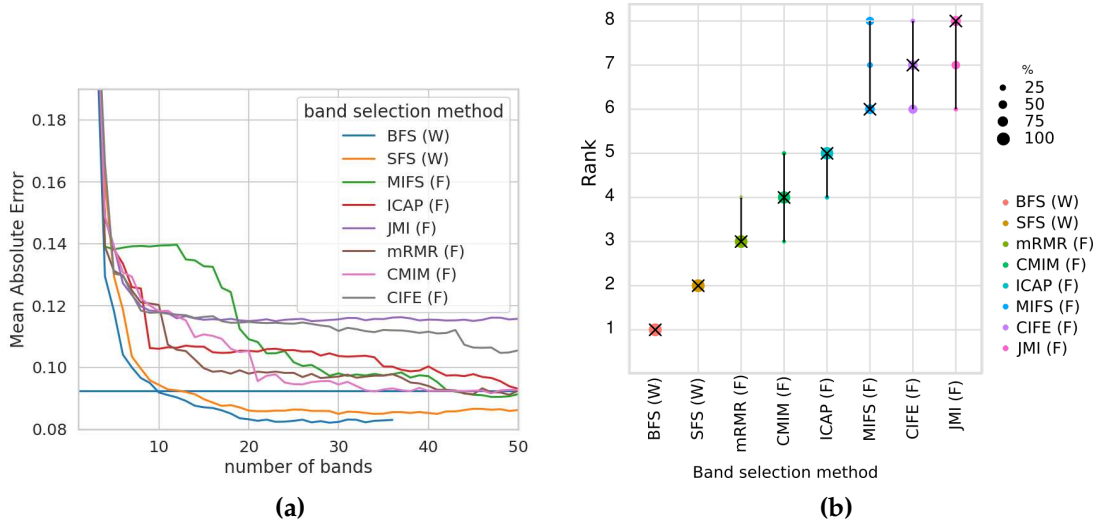
**Research question :** How do different feature selection algorithms perform in the task of oxygenation estimation?

To validate the domain and task-specific band selection method proposed in Sec. 2.2.3, generic MC reflectance simulations  $X_g$  were created from the tissue model described in Tab. 2.1. These generic simulations were randomly separated into a training dataset  $X_g^{train}$  (450,000 simulations) and a test dataset  $X_g^{test}$  (15,000 simulations). Bands were selected on a validation subset of 15,000 simulations, selected randomly from the training dataset. The band selections were evaluated by training the random forest regressor described in Sec. 2.4.1 on the entire training dataset and evaluating its performance on the test set.

The camera used for the *in vivo* evaluation was mimicked by transforming the generic simulations with Eq. 2.1. The camera was simulated with spectral bands every 2 nm from 500 nm to 700 nm, each band represented by a Gaussian transmissions profile with a Full Width at Half Maximum (FWHM) of 20 nm. The quantum efficiency of the spectral camera and the transmission spectra of the optical system were assumed constant within the relatively narrow filter bands. The camera noise was emulated by including multiplicative Gaussian noise at a level of 5%. The wavelength range 500–700 nm was chosen because the simulations and measurements did not match below 500 nm and above 700 nm. Furthermore, a 14 nm shift in the absorption spectrum was detected between the simulations and the measurements. This misalignment was corrected by shifting the simulations. The *in vivo* measurements were further normalized with a dark and white correction, as presented in Eq. 2.31. Measurements were also denoised by blurring them with a Gaussian kernel with a standard deviation of one pixel in the spatial domain.

In the following experiments, we used the random forest regressor described in Sec. 2.4.1. We evaluated the performance of such regressor on several popular *filter* and *wrapper* methods. The Mean Absolute Error (MAE) on the validation set of the generic simulations for all evaluated methods can be visualized in Fig. 3.3 a). A systematic robustness analysis, shown in Fig. 3.3 b), demonstrates that the best method across different

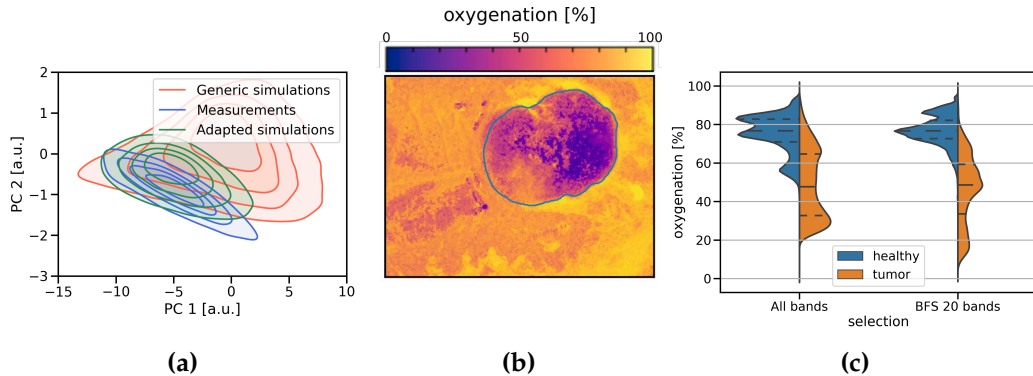




**Figure 3.3: Best First Search (BFS) outperforms other band selection algorithms, followed by Sequential Forward Selection (SFS)** (a) Mean Absolute Error (MAE) of various band selection methods applied on the generic training set  $X_g$  and evaluated on the test set  $X_g^{test}$ . *Wrapper* methods (W) outperform *filter* methods (F). The horizontal line represents the result when training and testing on all 101 bands. (b) Ranking stability of different band selection methods across different number of bands. Here, the rank of a method (1: best to 5: worst) is based on the MAE. Each method is color-coded, and the area of each blob at position  $(A_i, \text{rank } j)$  is proportional to the relative frequency  $(A_i)$  each method achieved rank  $j$  for 1000 bootstrap samples. The median rank of each algorithm is indicated by a black cross, and 95% bootstrap confidence intervals across bootstrap samples are indicated by black lines.

numbers of bands is Best First Search (BFS). The robustness analysis was performed with the *ChallengeR* tool developed by Wiesenfarth et al. (2021). Given that *wrapper* methods performed better than *filter* methods and BFS ranks in first place for all different numbers of bands, it was selected as the standard band selection technique for the following experiments.

We investigated how well a regressor trained on bands selected from a set of domain-specific simulations can estimate oxygenation on *in vivo* data. For this purpose, a dataset composed of eight hyperspectral images of head and neck tumors in a rat model were used (Lu et al., 2015). Each tumor image was recorded by a spectral camera (Maestro, PerkinElmer Inc., Waltham, Massachusetts). This system records hyperspectral images in the wavelength range 450–900 nm with a FWHM of 20 nm. Green Fluorescence Protein (GFP) was used to identify tumorous tissue on each image. More details on the image acquisition process can be found in the original dataset provided by Lu et al. (2015). From



**Figure 3.4:** We can adapt generic simulations to a specific domain and use them to estimate oxygenation in real data. **(a)** Two dimensional representation, computed via Principal Component Analysis (PCA), of generic Monte Carlo (MC) simulations before and after adapting them to real data from rat head and neck tumor models. **(b)** Oxygenation map using 20 bands selected with Best First Search (BFS) the tumorous region was identified via Green Fluorescence Protein (GFP) and is indicated by the blue outline. **(c)** Oxygenation distribution, as estimated by the random forest regressor, for the healthy and tumorous regions identified in (b). The distributions are shown when a regressor is trained on all 101 bands and when training only on 20 selected by BFS.

the eight rat tumor images, five were used for algorithm fine-tuning and three were reserved for the final validation of the band selection results. Our domain adaptation approach (Sec. 2.2.3) was evaluated by adapting the generic simulations (Tab. 2.1) to the hyperspectral *in vivo* rat data. Two subsets (train and test set) of 15,000 simulations (each) were generated by sampling, without repetition, using the weights provided by our method. Here, the Kernel Mean Matching (KMM) parameter  $B$  (see Eq. 2.6) was set to 10, and the kernel function's parameter  $\gamma$  was set to the median pairwise Euclidean sample distance determined on the training data. The two-dimensional representation, computed by PCA, of the generic simulations, rat measurements and domain-adapted simulations is presented in Fig. 3.4 a). This shows that the distribution of the simulations after domain adaptation resembles the *in vivo* rat data more closely.

Bands were then selected with the top-ranking algorithm BFS in the domain-adapted train dataset. The random forest regressor was then trained on the selected bands and evaluated on the domain-adapted test dataset. This analysis revealed that selecting 13 out of the 101 bands was sufficient to reproduce the baseline MAE (using all bands). Furthermore, selecting 20 bands achieved the lowest MAE while still reducing the number of required bands by approximately 80 %. Note that in this setup the band selection algorithm is agnostic to the real measurements as it is presented only domain-specific

generic simulations. This approach allows us to select bands that are specific to some domain without having ground truth oxygenation values for the measurements. The random forest regressor trained on the selected 20 bands was then evaluated on the *in vivo* rat data. We observed that the tumorous regions exhibit lower oxygenation values than the healthy tissue (outside the blue outline). Fig. 3.4 b) exemplifies this observation for one of the eight mice, similar results were observed for the other mice. Fig. 3.4 c) depicts the oxygenation value distribution of the tumorous and the healthy tissue when all 101 bands are used and when only 20 bands are selected with BFS. Based on this, the separation of the healthy from the tumorous tissue is more clear when using only the selected 20 bands.

## 3.2 Automatic light source estimation

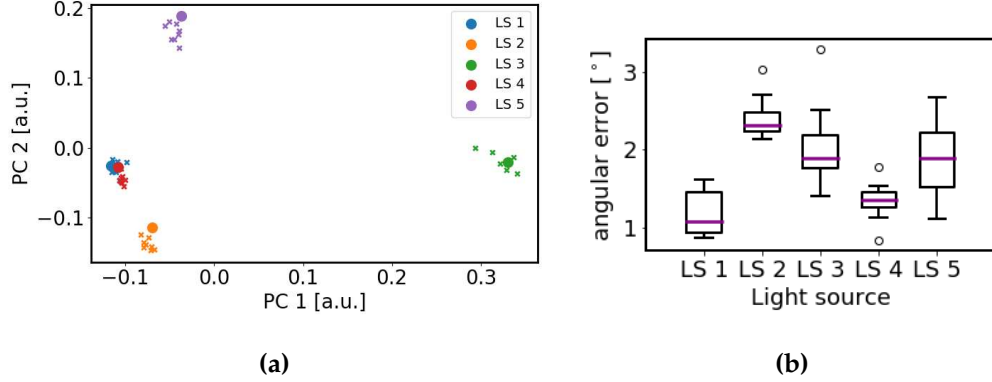
In order to facilitate the translation of optical functional imaging to open surgical scenarios, we designed a method for automatic light source (LS) estimation. We evaluated the performance of our proposed method to LS estimation (Sec. 2.3) in different settings: *in silico*, *ex vivo* and *in vivo*. We studied the performance of our method *in silico* and *ex vivo*, as well as the performance comparison to other state-of-the-art approaches (Sec. 3.2.1). We further analyzed the effect that an automatic approach to LS estimation can have in the task of oxygenation estimation *in vivo* (Sec. 3.2.2).

### 3.2.1 Light source estimation *ex vivo*

**Research question :** How does the proposed approach to light source estimation perform in comparison to state-of-the-art methods?

**Research question :** What is the effect of automatic light source estimations on the task of oxygenation estimation *in silico*?

To address these research questions, we recorded multispectral images of an *ex vivo* pig liver illuminated with the five different LSes described in Fig. 2.8. To study the robustness of our method to LS estimation, we recorded images of eight different camera poses at different exposure times in the range 5–150 ms. One pose corresponds to a fixed camera distance to the tissue surface, and a corresponding angle between the camera and the surface of the tissue. To quantitatively assess the performance of our method for LS estimation, we applied our method to a total of  $5 \times 8 = 40$  images (number of LS  $\times$  number of poses per LS). We then computed descriptive statistics for the angle between the reference spectrum (Fig. 2.8) and the estimated spectrum.



**Figure 3.5: Our method is capable of accurately estimating the light source spectra. (a)** Reference spectra (large circles) and corresponding estimates (crosses) for the light sources (LSes) described in Fig. 2.8. All light source spectra were projected onto the first two principal components of the five reference spectra, determined with Principal Component Analysis (PCA). Crosses of the same color represent different poses of the multispectral camera relative to the same LS. **(b)** Box plots of the angle between the reference spectrum and the estimated spectrum for each LS. The purple lines represent the mean angular distance for each LS, and the boxes extend between the 25<sup>th</sup> to the 75<sup>th</sup> percentile.

The two-dimensional representation of the reference LSes (large blobs) and our estimations (crosses) (Fig. 3.5) shows that the true illuminant is consistently the nearest neighbor to the estimates, except LS 1 and LS 4, which are both Xenon LS from different manufacturers and have an angular error of only 1°. This two-dimensional representation was computed via PCA. The performance of our LS estimation method is summarized in Fig. 3.5 b). It can be seen that the angle between our estimate and the reference is mostly below 3° (only one outlier). The performance for the *test LS* (LS 4-5) is similar to those of the *validation LS* (LS 1-3), which were used to tune the two hyperparameters ( $T_{exp}$  and  $N_p$ ), as described in Sec. 2.3.3.

To quantify the impact of the error in LS estimation on oxygenation estimation, we used the physics-based simulations presented in Tab. 2.1 to simulate a set of 15,000 tissue models, which was divided into a training data set  $X_{train}$  with  $|X_{train}| = 10,000$  and a testing data set  $X_{test}$  with  $|X_{test}| = 5,000$ .  $X_{train}$  was used to generate  $5 + 40 = 45$  *training sets* following the camera space transformation presented in Eq. 2.1, each corresponding to one of the five ground truth LS ( $LS_i \quad \forall i \in \{1, 2, 3, 4, 5\}$ ) or their estimates  $\hat{LS}_i$  ( $n = 5 \times 8$ ; one for each LS and each of the eight camera poses). In total, each of these 45 training sets was composed of 10,000 tuples of tissue properties and corresponding measurements. Note that the training sets for the different illuminants correspond to the same ground truth

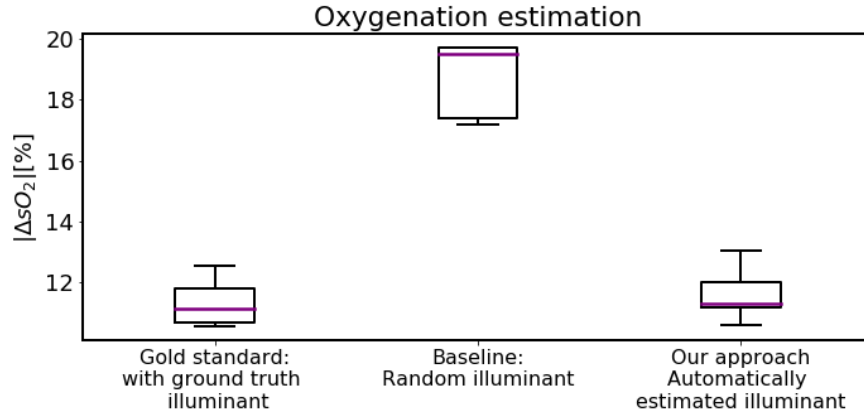
tissue parameters (including oxygenation, which is the parameter we wish to recover). We then trained an oxygenation regressor using the random forest proposed by Wirkert et al. (2016a) and presented in Sec. 2.4.1. To test the performance of all 45 regressors, we used  $X_{test}$  to generate a test set for each of the five reference LSes, following the transformation presented in Eq.2.1. Similarly, as for  $X_{train}$ , each test set comprised of 5,000 tuples of tissue properties and corresponding measurements. We then computed descriptive statistics for the quality of oxygenation estimation in three different scenarios :

1. Using the reference illuminant for training the random forest ( $LS_{train} = LS_{test}; n = 5$ ).
2. Selecting randomly one of the other illuminants for training ( $LS_{train} \neq LS_{test}; n = 20$ ). For each ground truth LS, one of the other four LSes was used for independently training a regressor. This process was repeated for each ground truth LS, thus resulting in 20 different regressors.
3. Using our approach to illuminant estimation to estimate the LS ( $LS_{train} = \hat{L}S_{test}; n = 5 \times 8$ ).

As shown in Fig. 3.6, the mean absolute error in oxygenation estimation when using the ground truth LS for training ranges from 10.5 p.p. (percentage points for LS 1) to 12.5 p.p. (LS 3) with a mean of 11.3 p.p. (averaged over all five LSes). The mean, median and max value of the mean oxygenation error was 11.6 p.p., 11.3 p.p. and 13.1 p.p. when applying our approach ( $n = 5 \times 8$ ). The results were similar for the *validation* LS (mean: 11.3 p.p.) and the *test* LS (mean: 12.0 p.p.). We reduced the mean oxygenation error by an average of 47%, compared to using an arbitrary LS (no calibration performed).

To evaluate the performance of our method in comparison to state-of-the-art approaches, we picked four related methods that fit our requirements: 1) no supervised training is needed and 2) no assumption about homogeneity or composition of the surface is needed. Following the terminology introduced in Sec. 2.3.3, we refer to these methods as **(M1-M4)**. M1-M4 were applied to the *ex vivo* recordings described before such that their performance could be compared with our approach on identical datasets.

Fig. 3.7 a) shows the performance of all methods, quantified by the angular error introduced in Sec. 2.3, for three different exposure times (low: 20 ms, normal: 40 ms, high: 60 ms) and averaged over all poses of all LS (**LS 1 - LS 5**) ( $n = 8$ ). While our method outperformed all the competitors and yielded relatively consistent performance over the three exposure times, the performance of M1-M4 was more sensitive to the exposure time of the recordings. This result is consistently supported by a systematic robustness analysis using the *ChallengeR* tool developed by Wiesenfarth et al. (2021) (Fig. 3.7 b).

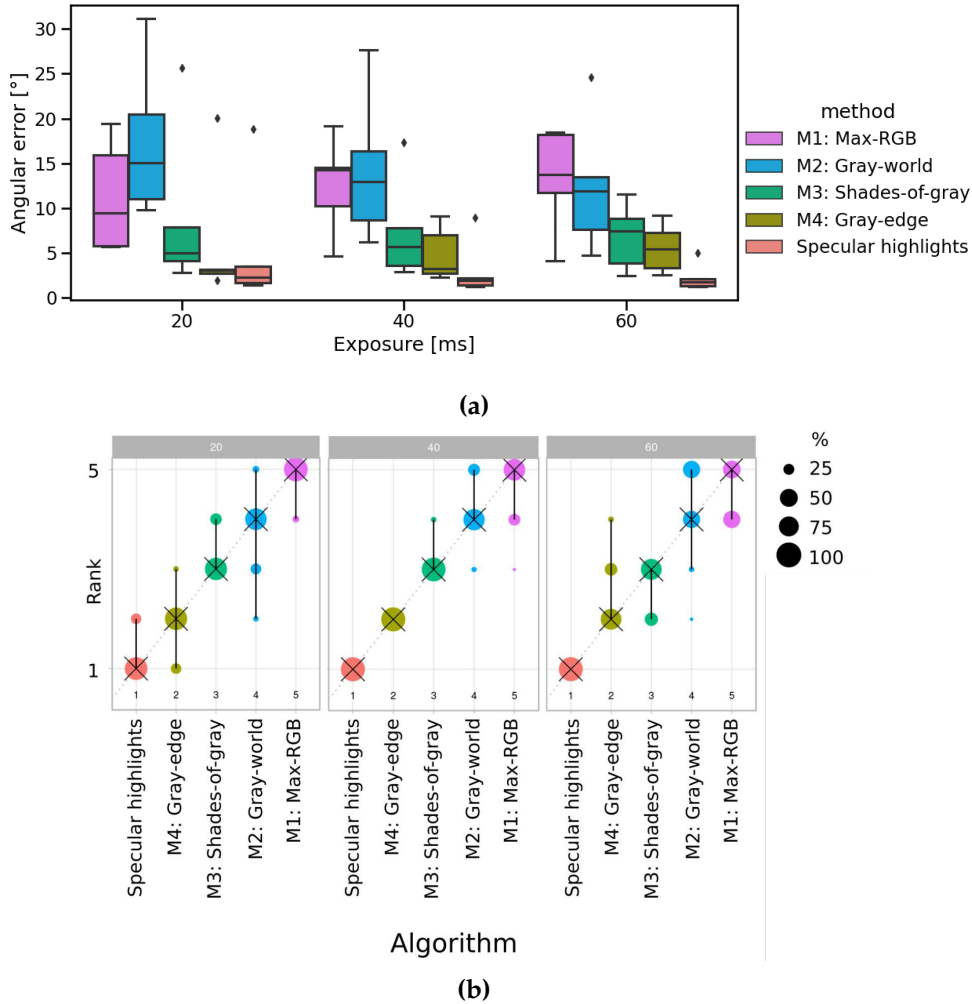


**Figure 3.6: Automatic light source estimation can reduce the estimation error of oxygenation *in silico*.** Error in oxygenation estimation when (1) using the reference light source (LS) for training ( $LS_{train} = LS_{test}$ ) (2) using a random LS for training ( $LS_{train} \neq LS_{test}$ ) and (3) using our approach to illuminant estimation to estimate the LS ( $LS_{train} = \hat{LS}_{test}$ ).

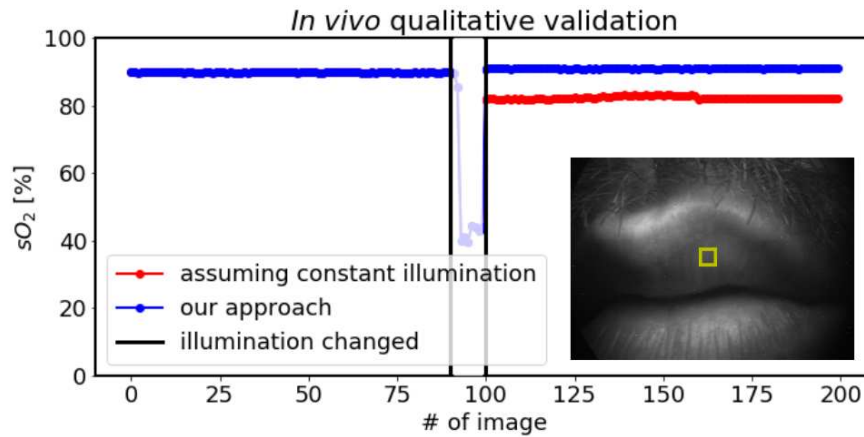
### 3.2.2 Light source estimation *in vivo*

**Research question :** What effect does automatic light source (LS) estimation have on the task of oxygenation estimation *in vivo*?

For *qualitative* validation, we recorded a sequence of multispectral images from the lips of a human subject and switched the LS from LS 1 (Xenon) to LS 5 (Light-Emitting Diode (LED)) in the middle of the recording. These recordings were performed with the compact spectral system presented in Fig. 2.2. We applied our approach to automatic light source calibration to continuously update the regressor to one tailored to the (estimated) light source. As a baseline method, we applied a regressor trained on LS 1 (the first LS used) throughout the whole acquisition process. Qualitative analysis was performed by plotting the mean oxygenation in an Region of Interest (ROI) (Fig. 3.8). The ROI, of size  $30 \times 30$  px, was located in a region of tissue that was not underexposed or overexposed. Given that the recordings lasted only a couple of seconds and the experimental setup was highly controlled, there was no need to automatically track the ROI, as there was no visible tissue motion across the entire image sequence. Given that the recordings only lasted a few seconds and only the LS was modified during recordings, it is reasonable to assume that an ideal regressor would estimate the same oxygenation value for the entire image sequence. Fig. 3.8 shows that the oxygenation estimation error is reduced when



**Figure 3.7: Our method outperforms all state-of-the-art approaches for different exposure times. (a)** Angular error of baseline methods and our proposed method (Specular highlights) for different exposure times. Box plots of the same color correspond to the same method. **(b)** Ranking stability of different methods when applied to different exposure times. Here, the rank of a method on each data set (1: best to 5: worst) is based on the angular error. Each method is color-coded, and the area of each blob at position  $(A_i, rank j)$  is proportional to the relative frequency  $(A_i)$  each method achieved rank  $j$  for 1000 bootstrap samples across the different tasks, where one task represents one exposure time. The median rank of each algorithm is indicated by a black cross, and 95% bootstrap confidence intervals across bootstrap samples are indicated by black lines.



**Figure 3.8: Automatic light source estimation can reduce the estimation error of oxygenation from spectral images of *in vivo* tissue.** When assuming a constant LS, the estimated blood oxygenation in an Region of Interest (ROI) on the human lips (yellow rectangle) changes when illumination conditions alter (and no longer match the training conditions). Our approach compensates for this by automatic LS calibration. The gap between index 85 and 100 represents the transition phase in which light sources were switched on and off, respectively.

utilizing our approach to estimate the LS spectra and therefore use the correct regressor.

### 3.3 Perfusion monitoring

We tackle the challenge of perfusion monitoring in two modules. The first module leverages simulations of light-tissue interactions generated with a MC method to regress functional properties (e.g. oxygenation). We validated this approach in different domains: pig brain and human skin (Sec. 3.3.1). The second module deals with domains that feature a high inter-patient tissue heterogeneity and are not well represented by the MC simulations that the first module exploits. We validated this approach in the first clinical trial involving video-rate spectral imaging in laparoscopy (Sec. 3.3.2).

#### 3.3.1 Perfusion monitoring via regression

We approached the task of regressing functional properties in biological tissue by exploiting large amounts of light-tissue simulations that were generated with a MC method. We employed a deep learning model, which is based on a CNN architecture, and trained it with the physics-based MC simulations (Sec. 3.1) to overcome the lack of *in vivo* labeled



data. Before validating our approach in the domains of pig brain and human skin, we performed an extensive hyperparameter search to optimize our CNN-based method.

**CNN-based perfusion monitoring hyperparameter tuning** It is well known in the deep learning community that hyperparameter selection as well as model architecture search plays a big role in the performance of any given model. We performed a random search for several hyperparameters and network architecture choices to find the parameters that are better suited for the task of oxygenation estimation. Here, we leveraged the physics-based MC simulations presented in Tab. 2.1 for model training. The simulated data was randomly divided into training, validation and test sets; each composed of 450,000, 50,000 and 50,000 samples respectively. Before training, each sample was transformed to the target camera and optical system space with the method introduced in Eq. 2.1. Only the training and validation sets were used for hyperparameter tuning. Each sample (multi-layer tissue model) was assigned a blood volume fraction and oxygenation value based on the weighted average by using our approach to multi-layer tissue model labeling (using penetration depth), as introduced in Eq. 2.4. The MSE loss between network estimations and the labels assigned to each simulated sample was computed during training. Performing an exhaustive hyperparameter search can be computationally expensive. Thus, we performed half-precision training to speed up this process. By doing so, we did not observe substantial differences in the results when training at full-precision. Furthermore, we used the Adam optimizer (Kingma and Ba, 2014a) paired with a *reduce on plateau* learning rate scheduler to aid the convergence of each model. This scheduler reduces the learning rate by a factor of ten, when the validation loss has not decreased in ten consecutive epochs by as much as  $10^{-4}$ .

Traditional data augmentation techniques used in the deep learning community are generally not well suited for the task of oxygenation estimation. For example, random reordering of bands (wavelengths) would get rid of valuable spectral information, such as the characteristic double absorption peaks of hemoglobin (Fig. 3.2). With this in mind, we augmented each sample with a custom strategy. At the beginning of each epoch, the entire training dataset was augmented by: a) random sample shuffling, b) multiplicative Gaussian noise with a Signal-to-Noise Ratio (SNR) of 20, c) brightness variation to randomly selected samples, and d) specular component addition to randomly selected samples. Brightness variation was performed by randomly selecting approximately 10% of the training samples and multiplying them by a normal distribution with mean  $\mu = 1$  and standard deviation of  $\sigma = 0.1$ . Specular reflections are very common in real case scenarios, so we augmented our simulations by introducing specular reflections. To achieve this, we based our augmentation strategy on the dichromatic model, which states that every

reflection can be modeled as the sum of a specular and a diffuse reflection:

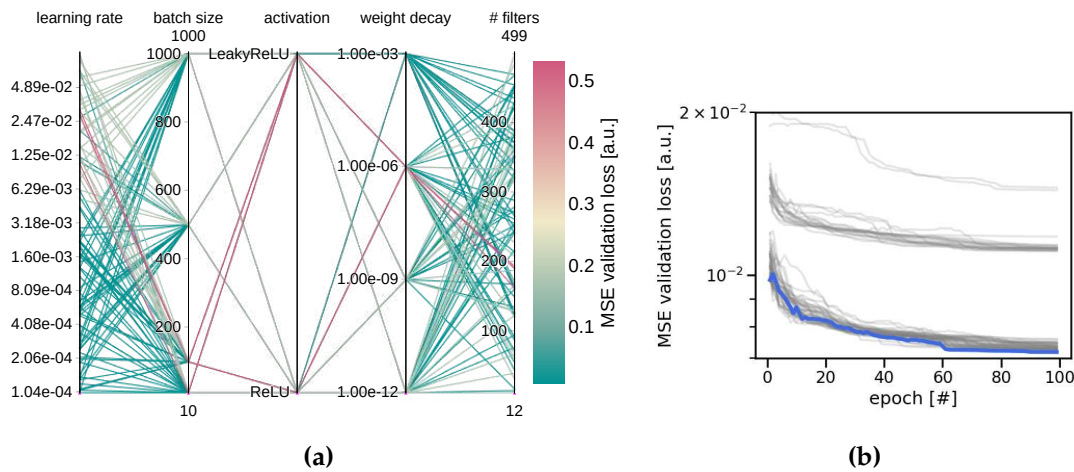
$$\mathbf{X}_s^{aug} = (1 - \kappa)\mathbf{X}_s + \kappa\mathbf{w}, \quad (3.1)$$

where  $\mathbf{X}_s$  represents 10% of the training samples (randomly selected),  $\mathbf{X}_s^{aug}$  are the augmented samples,  $\kappa$  represents the ratio of specular component that is desired, and  $\mathbf{w}$  is the spectra of a light source.

The hyperparameter space explored during this analysis was defined by the following sampling strategy:

- *learning rate*: sampled from a logarithmic uniform distribution  $\mathcal{U}_{\log}(10^{-4}, 10^{-1})$ .
- *batch size*: randomly sampled from  $\{10, 100, 500, 1000\}$ .
- *activation*: randomly selected Rectified Linear Unit (ReLU) or Leaky Rectified Linear Unit (LeakyReLU).
- *weight decay*: randomly sampled from  $\{10^{-12}, 10^{-9}, 10^{-6}, 10^{-3}\}$
- *# filters* (per hidden layer): sampled integers from  $\mathcal{U}(10, 500)$

The hyperparameter space was explored using the Asynchronous Successive Halving Algorithm (ASHA) scheduler (Li et al., 2018) to parallelize the computations. Fig. 3.9 a) shows a parallel plot of the explored hyperparameter space, each line corresponds to one choice of hyperparameters (one model) and the color of each line corresponds to the validation loss on the last epoch during training. The *learning rate* was one of the most influential hyperparameters, lower *learning rate* generally produced models with lower validation loss. In particular, relatively high *learning rates* paired with low *batch size* ( $\leq 200$ ) lead to models that either did not converge or that converged to relatively high validation losses. The rest of the hyperparameters did not seem to considerably influence the final trained model. However, higher *batch sizes* lead to quicker model training, and an intermediate value of *weight decay* generally generated the models with the lowest validation loss. Based on this analysis, we selected the model with the lowest average MSE validation loss on the last ten epochs. The best performing model is highlighted as a blue line in Fig. 3.9 b), the gray lines correspond to all other trained models. The hyperparameters corresponding to this model are the following: *learning rate*:  $1.3 \times 10^{-3}$ , *batch size*: 500, *activation*: LeakyReLU, *weight decay*:  $10^{-9}$ , *# filters*: 353. These hyperparameters were used for the rest of the analysis involving our CNN-based approach to oxygenation estimation.



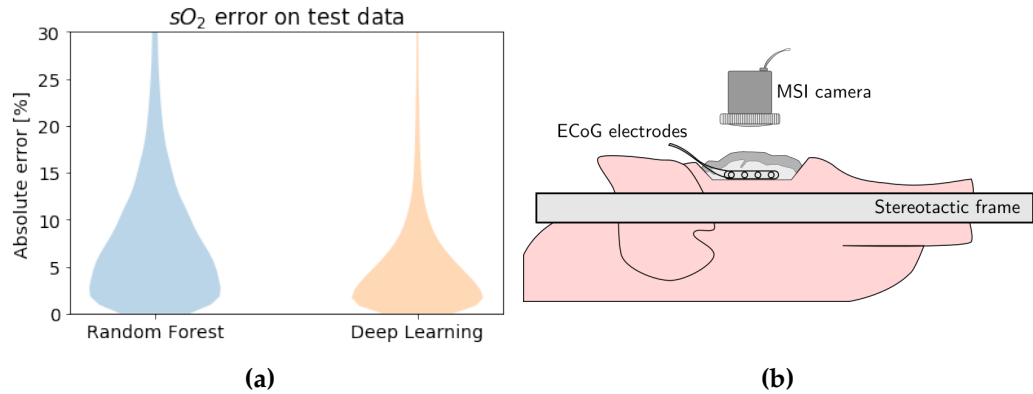
**Figure 3.9: Generally, lower *learning rates* produced models with the lowest Mean Squared Error (MSE) loss. (a) Parallel plot of the hyperparameter space explored for hyperparameter tuning. The *learning rate* and the *batch size* are the hyperparameters that mainly influence the training results of our CNN network. (b) Validation loss curves for all trained models, the blue line corresponds to the validation loss of the model with the lowest average MSE loss in the last ten training epochs.**

### Medical application in brain perfusion

**Research question :** How can perfusion changes be monitored in the brain with a deep learning model?

The validation of our proposed CNN-based approach to oxygenation estimation for neuroimaging is composed of an *in silico* and an *in vivo* study. In the *in silico* study, we compare the performance of our CNN-based oxygenation estimation to the state-of-the-art random forest proposed by Wirkert et al. (2016b) (Sec. 2.4.1). We used the median absolute error of oxygenation ( $sO_2$ ), the interquartile range of  $sO_2$  absolute error and the regression speed as our target metrics. The purpose of the *in vivo* experiments was to investigate if our method is capable of monitoring perfusion changes in porcine brain.

For the *in silico* validation, we compared the random forest baseline to our deep learning (CNN-based) approach. The hyperparameters used for our deep learning model were the ones obtained from our hyperparameter tuning experiment presented in Fig. 3.9. Fig. 3.10 a) shows that the absolute  $sO_2$  estimation error of our proposed deep learning approach is smaller than the random forest. Furthermore, our CNN model is capable of processing  $5.2 \times 10^6$  spectra per second with a median absolute  $sO_2$  estimation error of



**Figure 3.10: Our deep learning method estimates oxygenation ( $sO_2$ ) *in silico* with lower estimation error compared to the state-of-the-art random forest. (a) Oxygenation estimation absolute error for the state-of-the-art random forest and our deep learning approach. (b) Experimental setup for monitoring Spreading Depolarizations (SDs) with a spectral camera, Electrocortigraphy (ECoG) electrodes were placed at the margins of the craniectomy to monitor SDs.**

5.8%. While the random forest only achieves a speed of  $2.7 \times 10^6$  spectra per second with a median absolute  $sO_2$  estimation error of 9.1%.

To investigate if our deep learning model is capable of monitoring perfusion changes in the brain, we chose the monitoring of Spreading Depolarizations (SDs) as the target medical application. SD is a phenomenon in the brain that is usually related to brain trauma and manifests itself as an abrupt depolarization of the neurons in the gray matter. The depolarization of neurons acts as a cascade effect that travels through the gray matter in the form of a wave (Dreier and Reiffurth, 2015). It is well known to be coupled to hemodynamic responses with hypoxic components (Takano et al., 2007), making it a prime target for the qualitative validation of our method for functional Multispectral Imaging (MSI). All experiments were performed in accordance with the relevant guidelines and regulations. Protocols for all experiments were approved by the institutional animal care and use committee in Karlsruhe, Baden-Württemberg, Germany (Protocol No. 35-9185.81/G-174/16). The procedure for inducing and measuring SDs was as follows: As shown in Fig. 3.10 a), a craniectomy exposed the parietal cortex, and spreading depolarizations were induced using 2-5  $\mu$ L Potassium Chloride (KCl) drops placed with a Hamilton syringe in regions selected by visual inspection in the parietal cortex (Santos et al., 2014, 2013).

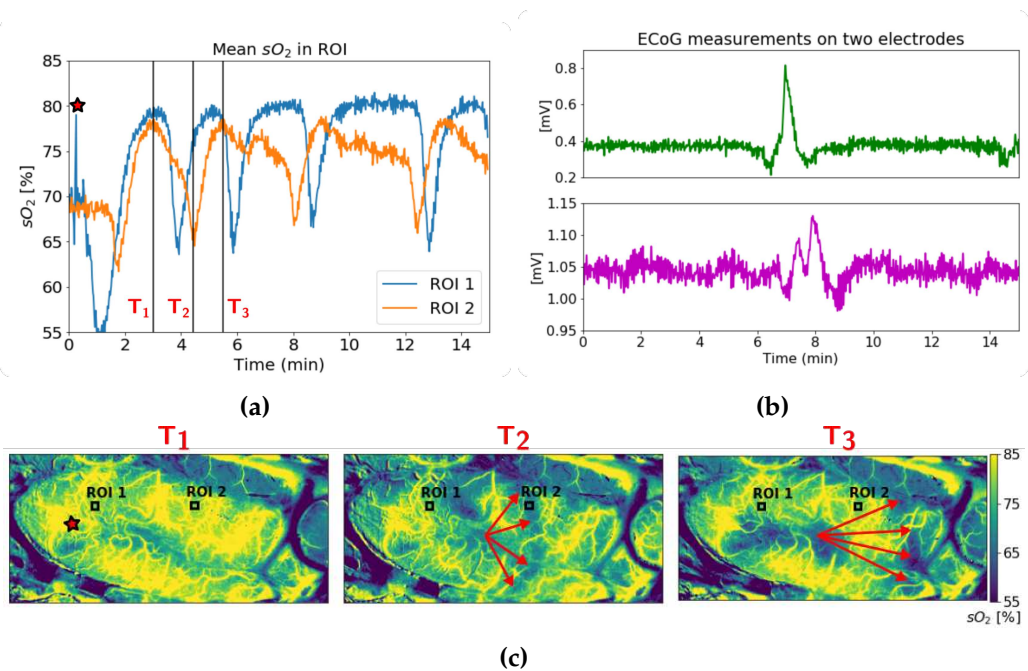
We recorded images of a porcine brain with the high resolution spectral system presented in Sec. 2.2.1. Since the high resolution spectral camera (Spectrocam, Pixelteq, Largo,

FL, United States) records images sequentially through a filter wheel, tissue motion caused by respiration, heart beat or surgical manipulation, for example, could potentially lead to misalignment across the recorded images. To mitigate this problem, we performed elastic registration to compensate for motion across multispectral images. The images corresponding to individual bands of a whole spectral image were registered with the algorithm `bUnwarpJ` (Arganda-Carreras et al., 2006) incorporated in the software FIJI (Rueden et al., 2017), which is based on elastic deformations. The band with the highest contrast (560 nm) was used as reference for the registration algorithm, and registration of one complete multispectral image took  $\sim 12$  s. The implementation of such an elastic transformation to compensate brain movement was inspired by the work of Schöll et al. (2017). This registration was implemented only with the purpose of obtaining a better visualization of hemodynamic changes; it is not a requirement of our proposed network model and motion compensation across images is not necessary when applying a snapshot camera for image acquisition (Fig. 2.2). As demonstrated in Fig. 3.11 c), our method is capable of video-rate visualization of blood oxygenation in the porcine cortex. Two ROIs were defined on the surface of the brain based on the criteria that a) there were no visible vessels within each ROI and b) the ROIs did not cover regions with specular reflection or blood leakage. After SD induction with KCl, SDs can be clearly visualized, as illustrated by the repeated mean oxygenation value drops within each ROI (Fig. 3.11 a)). To ensure that the oxygenation changes that we observed were indeed due to SDs, they were continuously monitored with two ECoG recording strips that were placed on the lateral margins of the craniectomy. Fig. 3.11 b) shows the recordings from both strips, these recordings are typical for SDs.

#### Medical application in human skin perfusion (A clinical trial)

**Research question :** How effective is our deep learning model in monitoring perfusion changes in human skin?

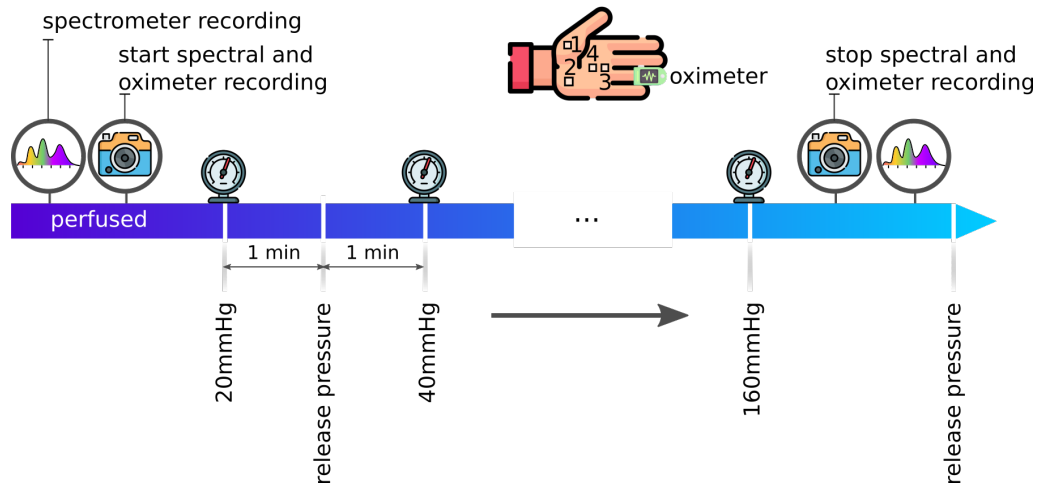
We studied the capabilities of our deep learning approach to oxygenation estimation in the domain of human skin. Building on the work of Kumar et al. (2020), we designed a clinical trial involving 12 healthy patients. All participants were recruited at the German Cancer Research Center (Heidelberg, Germany). The inclusion criteria for all patients were: a) patients were adults ( $\geq 18$  years) and b) patients were healthy (no known conditions). All experiments in this study were performed in accordance with the Declaration of Helsinki and all protocols were approved by the “Ethikkommission der Medizinischen Fakultät Heidelberg” (DE/EKBW03, study reference number: S-530/2020). The study is registered with the German Clinical Trials Register (DRKS00023246).



**Figure 3.11: Our deep learning method is capable of video-rate visualization of blood oxygenation in the porcine cortex.** (a) Mean oxygenation ( $sO_2$ ) value within each of the defined Regions of Interest (ROIs). The star symbol indicates the spatial location and time point when the Spreading Depolarizations (SDs) were induced by Potassium Chloride (KCl) injection. (b) Electrocorticography (ECoG) electrode measurements typical for SDs. (c) Oxygenation maps of the brain cortex for three different time points (1)-(3), which correspond to the three time points highlighted in a). A deoxygenation wave can be observed traveling from left to right of the brain.

Recordings were taken from healthy patients at the German Cancer Research Center. We recorded multispectral images of the surface of the participants' right hand with the spectral system presented in Fig. 2.2. As validation data, we recorded spectrometer diffuse reflectances from four different locations (1)-(4) of the hand, as depicted in Fig. 3.12.

1. **perfused:** Spectral images of the surface of the patients' right-hand palm were recorded for approximately two minutes as a reference measurement at the beginning of each recording session.
2. **ischemic:** After performing the reference measurements (**perfused**), pressure was applied to the upper-right arm with a manual pressure cuff. The pressure was then increased between 20 mmHg and 160 mmHg in steps of 20 mmHg every minute. Spectral recordings for one minute followed the application of each pressure, after



**Figure 3.12: Spectral recording procedure of human skin.** We recorded spectral images as well as spectrometer and oximeter data for each patient of our clinical trial. The spectral and oximeter recordings span across the entire experiment, while the spectrometer measurements were only performed before pressure was applied to the patient’s upper arm and after applying a pressure of 160 mmHg.

which the pressure was released to reduce the stress to the patient’s arm. The spectral recordings were performed continuously during these experiments to study the re-perfusion behavior of each patient. A diagram depicting the measurement process is presented in Fig. 3.12.

3. **spectrometer recordings:** Spectrometer measurements were performed with a scientific grade spectrometer (HR2000+, Ocean Insight, Orlando, USA) and calibrated with white and dark measurements as described in Eq.2.31 using a Spectralon<sup>®</sup> target (SRT-99-050, Labsphere<sup>®</sup>, North Sutton, USA). 100 diffuse reflectance measurements were recorded at each ROI (Fig. 3.12) and at the beginning of each recording session before any pressure was applied to the patients’ arm. Another recording of 100 measurements was done while applying a pressure of 160 mmHg with the manual pressure cuff.
4. **oximeter recordings:** To compare the performance of our method to the baseline approach proposed by Kumar et al. (2020), we recorded the pulsatile signal from an oximeter (MightySat Rx, Masimo, California, USA) placed at the tip of the ring finger of each patient. Continuous oximeter measurements were recorded throughout the entire experiment.

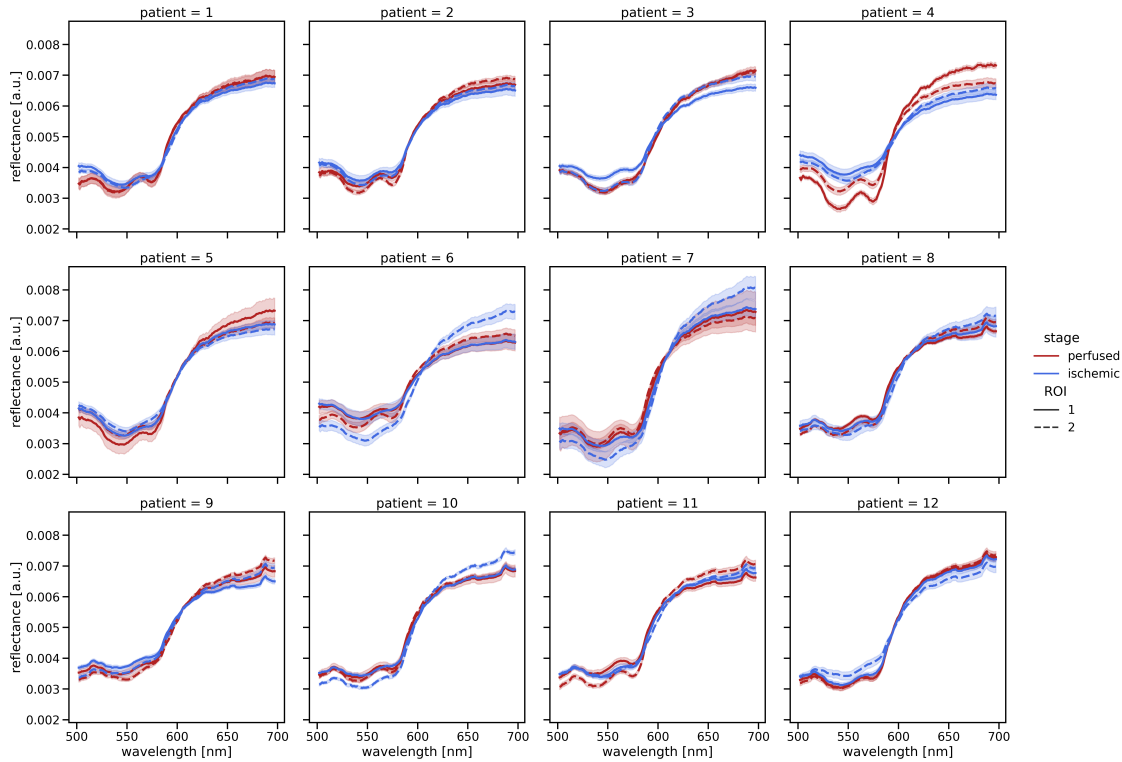
We annotated four ROIs (Fig. 3.12) of size  $30 \times 30$  px on the surface of the palm and

tracked each one across the entire recording sequence using our deep learning-based tracker described in Sec. 2.4.2. We selected each ROI based on the criteria that in each region a) there was no scar, b) there was no visible pigmentation (melanocytic nevi, stains, etc.), and c) pixels were not over saturated or under saturated. Furthermore, we aimed to place each ROI in similar locations of the hand across patients. The data from each ROI was subsequently normalized with a white and dark recording, as well as an  $\ell_2$  norm according to Eq. 2.31 and Eq. 2.32. These normalized ROIs served as input to our deep learning model for oxygenation estimation.

The normalized spectrometer measurements of perfused and ischemic (160 mmHg) skin tissue for all patients is presented in Fig. 3.13. To reduce the amount of noise, each measurement was smoothed with a rolling window of size 5, which corresponds approximately to 2 nm. We only show the data for two of the four annotated ROIs for clarity of visualization, however, a similar behavior was observed in the other ROIs. The spectra for all ROIs is provided in Sec. A.3. Slight differences between the spectra of perfused and ischemic tissue were observed for most patients. In particular, the differences for patients 3, 4 and 6 were considerable compared to the rest of the patients. The perfused measurements depict the characteristic double peak of HbO<sub>2</sub> while the ischemic measurements show less pronounced peaks in accordance to the extinction coefficient of Hb (Fig.2.3). An interesting finding is that for some patients (7 and 10) the reflectance of ischemic tissue is smaller than the one from perfused tissue for  $\lambda \lesssim 600$  nm and the opposite is true for wavelength  $\lambda \gtrsim 600$  nm. The opposite behavior was observed for other patients (4, 5 and 12).

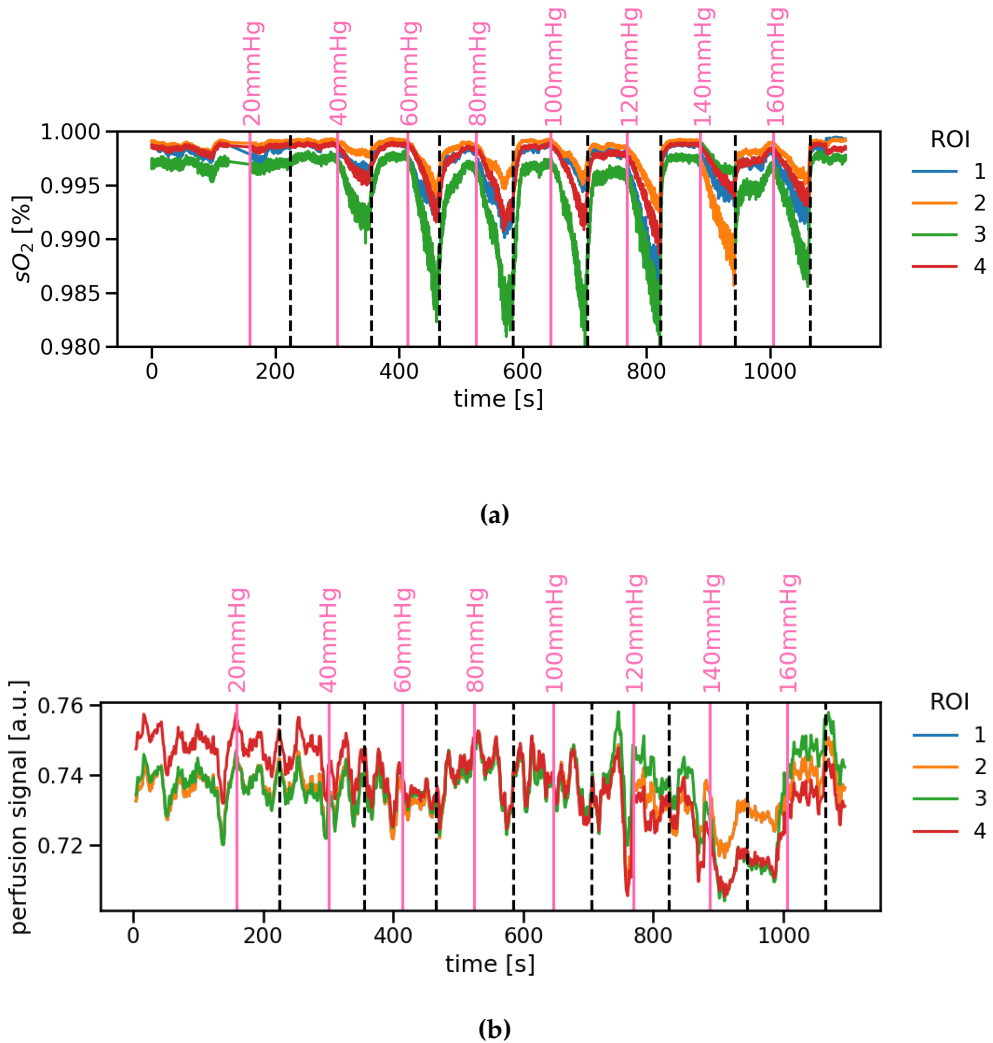
The oxygenation estimations ( $sO_2$ ) of our deep learning method (Fig. 2.9) are presented in Fig. 3.14 a). We observed that our method is capable to detect even small changes in perfusion in all ROIs in comparison to the method proposed by Kumar et al. (2020). For example, our method starts to detect a gradual decrease in oxygenation after applying a pressure of 40 mmHg. After one minute of applying a pressure of 40 mmHg, the oxygenation decreases by approximately 1%. The oxygenation decreased more for higher pressures, by as much as 2%. Interestingly, higher pressures (140 mmHg and 160 mmHg) produced a smaller decrease in oxygenation. A sudden re-oxygenation effect was observed each time the pressure was released (vertical dashed lines). We observed that the re-oxygenation at each ROI occurred more abruptly for higher pressures. Here, we only exemplified the results for one patient; however, similar results were observed for the rest of the patients. It is important to notice that, as expected, the oxygenation estimations differ across different ROIs. Moreover, the relative decrease in oxygenation after applying each pressure to the upper arm differs across ROIs. This occurred even though the reference oxygenation values for each ROI were very similar, as can be seen below 200 s in Fig. 3.14 a).





**Figure 3.13: Slight spectral differences can be detected with diffuse reflectance measurements performed with a spectrometer.**  $\ell_2$ -normalized reflectance spectra for all 12 patients of our clinical study, corresponding to two exemplary Regions of Interest (ROIs) annotated in perfused and ischemic tissue (160 mmHg). The solid lines correspond to the median reflectance spectra across all 100 measurements, while the translucent bands around each line correspond to the standard deviation.

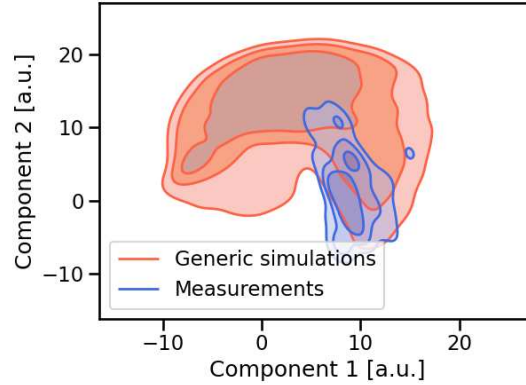
In contrast to our method, perfusion estimations computed with the method proposed by Kumar et al. (2020) (Sec. 2.4.1) do not show a clear change in the perfusion signal after each pressure is applied (Fig. 3.14 b)). However, a slight decrease in the perfusion signal was observed for high pressures. The amount of noise in this perfusion signal is also considerably higher than the one present in our estimations. The perfusion signal also did not show the characteristic re-oxygenation behavior that our method was able to detect. Here, we exemplified the results of the method from Kumar et al. (2020) on patient 2 of our clinical study; however, similar results were observed for the other patients.



**Figure 3.14: Our deep learning method is capable of monitoring perfusion changes in skin.** (a) Representative oxygenation estimations of our deep learning method for patient 2 of our clinical trial. Solid vertical lines represent the time points at which each pressure was applied, while dashed vertical lines represent the moment at which the pressure was released. (b) Perfusion estimations generated with the method proposed by Kumar et al. (2020) for each ROI annotated on patient 2.

### 3.3.2 Perfusion monitoring via Out-of-Distribution detection (Clinical study)

Ischemia monitoring in surgical applications is usually met with a hard challenge: *the presence of pathologies*. This challenge arises from the intrinsic nature of many surgical



**Figure 3.15: The Monte Carlo (MC) simulations of light transport do not represent pathological tissue in human kidney.** Kernel Density Estimation (KDE) of the two-dimensional representation computed with Uniform Manifold Approximation and Projection (UMAP) of the MC simulations and measurements of kidney tissue of patients undergoing partial nephrectomy.

procedures that aim to treat such pathologies: partial nephrectomy, lymph node resection, etc. This implies that any synthetic data (e.g. MC simulations) need to be re-designed to reflect the optical properties of such pathologies. The general lack of knowledge about the optical properties of pathological tissue makes this a hard problem. Even one of the most influential literature sources for optical properties of biological tissue (Jacques, 2013) lacks information about pathologies.

The effect of this lack of information can be visualized in Fig. 3.15, it shows the Kernel Density Estimation (KDE) of the two-dimensional representation computed with Uniform Manifold Approximation and Projection (UMAP) of the MC simulations and the measurements from this clinical trial. Here, the simulations described in Sec. 2.2.2 were transformed to the measurement hardware presented in Sec. 2.2.1 by the method described in Eq. 2.1. Both the simulations and the measurements were  $\ell_1$ -normalized before computing the low dimensionality representation with UMAP. This shows that most of the density from the measurements lies either in the border or outside the distribution of our MC simulations, and it supports the hypothesis that our simulations do not represent pathological tissue in human kidney.

This hindered the application of our proposed method for oxygenation estimation based on MC simulations presented in Sec. 2.2.2, and motivated us to develop a method that does not require knowledge of the optical properties of pathologies. In Sec. 3.3.2 we present the clinical study design used to validate our method for ischemia monitoring. Sec. 3.3.2 presents the Deep Learning (DL)-based method for automatic tissue ROI tracking.

**Table 3.1:** Patients recruited in our clinical study. **ccRCC:** Clear Cell Renal Cell Carcinoma, **PRCC:** Papillary Renal Cell Carcinoma, **ChRCC:** Chromophobe Renal Cell Carcinoma

Patient ID	Age	BMI	Sex	Tumor type	Smoker	Diabetic
1	80	25	male	PRCC	No	No
2	70	22	female	Benign Oncocytoma	No	No
3	49	40	male	ccRCC	No	No
4	56	30	male	Benign Oncocytoma	No	No
5	49	49	female	Leiomyoma	No	Yes
6	82	22	male	ccRCC	No	No
7	67	29	male	ccRCC	Yes	No
8	56	43	male	PRCC	No	No
9	53	31	male	ChRCC	No	No
10	40	24	female	Angiomyolipoma	No	No

Sec. 3.3.2 presents the results of our approach to contrast-free and video-rate ischemia monitoring.

### Study design

The aim of this study was to investigate the feasibility of our approach to contrast agent-free video-rate ischemia monitoring. In this section, we describe the patient recruitment, the in-vivo recording procedure and the study cohorts defined for our validation strategy.

**Ethic declaration** All experiments involving humans were performed in accordance with the Declaration of Helsinki and all protocols were approved by the Landesärztekammer Baden-Württemberg (DE/EKBW01, study reference number: B-F-2019-101). The study is registered with the German Clinical Trials Register (DRKS00020996).

**Patient recruitment** All participants of this clinical trial were recruited in the Städtisches Klinikum Karlsruhe (Karlsruhe, Germany). The inclusion criteria for all subjects were: a) subjects were undergoing partial nephrectomy, and b) subjects were adults ( $\geq 18$  years old). This resulted in a total of 10 subjects with an age range between 40 and 82 years old, seven of which were male and three female. Tab. 3.1 shows the patient ID, age, Body Mass Index (BMI), sex, tumor type, smoker/non-smoker and diabetic/non-diabetic. Tab. 3.2 shows the patient ID, patient comorbidities and the cohort (data split) to which each patient belongs.

**Study cohorts** We designed the study to obtain a reliable reference for our measurements, we only included patients in which the ischemia induction of the kidney was

**Table 3.2:** Patient comorbidities and data splits for the participants in our clinical study.

Patient ID	Comorbidities	Data split
1	Prostate cancer	retrospective
2	Psychiatric illness	retrospective
3		retrospective
4		retrospective
5		retrospective
6	Sigma Carcinoma	prospective
7	Rheumatism	prospective
8		prospective
9		prospective
10	Migraine	prospective

successfully confirmed by Indocyanine Green (ICG) injection as part of the traditional laparoscopic surgical procedure. We also omitted patients for which image recordings were underexposed due to a faulty surgical light guide. Many machine learning-based algorithms suffer from overfitting on the test data (Ball et al., 2020). Even if an algorithm is strictly trained on the training data, a common erroneous procedure involves testing multiple different models or hyperparameter configurations on the test set and then reporting the best model. To avoid the performance overestimation issue that these practices would cause, we finalized the complete algorithm based on existing data sets before conducting the patient study. The only exception was the number of spectral images used for training the networks used to compute the *ischemia index*. To fix this hyperparameter, we separated the recruited patients into a retrospective and prospective split (c.f. Tab. 3.2). The retrospective data split was composed of the first half of the patients (patients 1 to 5) and the prospective split of the second half of the patients (patients 6 to 10). The number of spectral images was determined using only the retrospective data set (Tab. 3.2).

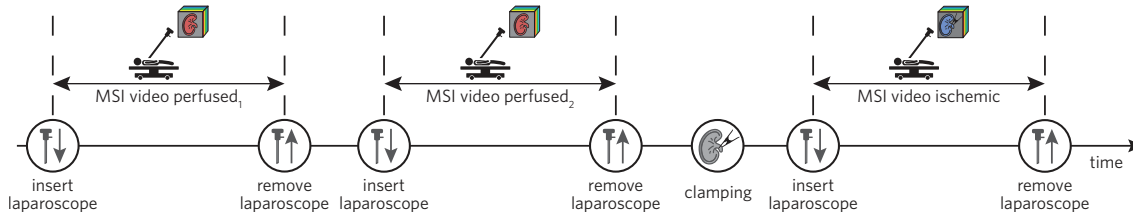
***In vivo* image acquisition** Recordings were taken from subjects undergoing partial nephrectomy at the Städtisches Klinikum Karlsruhe. The standard procedure for partial nephrectomy involves the generation of six surgical ports in the abdominal cavity, three of which were dedicated to the da Vinci<sup>®</sup> robotic arms (Intuitive Surgical Deutschland GmbH, Freiburg, Germany), one was used for the da Vinci<sup>®</sup> RGB surgical camera and the remaining two as assisting ports for other instruments (e.g. scissors, clip appliers, etc.).

We adapted the conventional surgical workflow as follows to record the spectral image sequences that our method leverages. After removing the fatty tissue from the surface of the kidney and locating the renal artery, the laparoscope, attached to our multispectral camera, was inserted through one of the assisting ports and three measurements were

performed to enable training and validation of our ischemia monitoring algorithm. We recorded spectral image sequences for 45 s, which yielded  $\approx 1200$  recorded spectral images (see Fig. 3.16):

1. **First recording of perfused kidney** (training sequence: **perfused<sub>1</sub>**): spectral images of the surface of the kidney were recorded before a clamp was applied to the renal artery. Part of this data was used for training the personalized *ischemia index*, as detailed in Sec. 3.3.2.
2. **Second recording of perfused kidney** (testing sequence: **perfused<sub>2</sub>**): To simulate an unsuccessful clamping, an additional image sequence was recorded before the clamp was applied to the renal artery. For this purpose, the laparoscope attached to our multispectral camera was removed from the surgical port and re-inserted through the same port. This process simulates the actions performed when the renal artery is clamped. Then, spectral images of the surface of the kidney were recorded. The re-insertion of the laparoscope was performed to obtain slightly different acquisition conditions. This is also the case when actual clamping is performed, as the laparoscope needs to be removed before the renal artery can be clamped through the same surgical port (see next paragraph).
3. **Recording of ischemic kidney** (testing sequence: **ischemic**): The laparoscope, attached to our multispectral camera, was removed from the abdominal cavity, such that the surgical team could perform the clamping procedure through the same port. The renal artery was then clamped, and the laparoscope was re-inserted in the abdominal cavity for recording the spectral images. Afterwards, the success of the clamping procedure was confirmed by ICG injection into the subject's bloodstream and subsequent visualization of the fluorescence signal with the Firefly system of the da Vinci<sup>®</sup> robot. The ICG injection was prepared by mixing 50 mg of ICG powder (PULSION Medical Systems SE, Feldkirchen, Germany) with 10 ml of distilled water. This injection was only administered to each patient after all the spectral recordings had finished.

To generate a high-quality dataset, all spectral recordings were performed on parts of the kidney tissue where regions without blood and fatty tissue could be observed. In addition, we ensured that before each recording, a) the camera integration time was set to 40 ms, b) the laparoscope was positioned such that the clean surface of the kidney was visualized in the field of view of the multispectral camera and c) the intensity of the light source was adjusted to minimize the number of specular reflections observed on the surface of the kidney. Furthermore, the light source of the da Vinci<sup>®</sup> Red-Green-Blue (RGB) surgical camera was temporarily turned off during each one of our measurements. The da



**Figure 3.16: Multispectral Imaging (MSI) recording procedure in the operating room (OR).** We recorded two image sequences (perfused<sub>1</sub> (for training) and perfused<sub>2</sub> (for testing)) before the clamp was applied to the renal artery, and one sequence after applying the clamp (for testing). The laparoscope was removed and re-inserted in the patient’s abdominal cavity before recording the perfused<sub>2</sub>. This ensured equal recording conditions as for the ischemic sequence.

Vinci<sup>®</sup> light source was immediately turned on after each of our recordings was completed. While the da Vinci<sup>®</sup> light source was off, the abdominal cavity was still visible thanks to the illumination provided by our Xenon light source. This ensured that the abdominal cavity was visible at all times during each surgery.

### Region of interest (ROI) tracking

**Research question :** How can Regions of Interest (ROIs) of biological tissue be robustly tracked in laparoscopic surgical scenes?

**Research question :** What is the influence of small variations in the initial Region of Interest (ROI) location initialization on the subsequent automatic tracking?

We analyzed the robustness of our tracker by manually initializing two non-overlapping ROIs of size  $30 \times 30$  px on the first image of each sequence (3.3.2). This manual initialization was done while ensuring that within each ROI **a**) there was no adipose tissue, **b**) there was no blood leakage, **c**) the camera counts were not over saturated and not under saturated, and **d**) there was no visible smoke. Each ROI was subsequently tracked across consecutive images with the deep learning-based algorithm presented in Sec. 2.4.2. As the laparoscope, attached to our multispectral camera, needed to be retracted between each recording to introduce surgical instruments through the same port, we were unable to ensure imaging of the same tissue region. Hence, the ROIs of the different sequences do not correspond to the same tissue location.

After automatic tracking, each ROI sequence was inspected to ensure that they **a)** did not disappear from the camera field of view, **b)** at least 95 % of all tracked pixels were not over saturated and not under saturated, and **c)** there was no visible drift of the ROI from the original annotated tissue region. Alg. 1 details the complete ROI annotation process, and a flow chart illustrating this algorithm is provided in Sec. A.1.

The reproducibility of the tracking process was evaluated by tracking each ROI ten times with the same initial location. The location of all ROIs was the same across all ten repetitions. Moreover, our tracking algorithm can run at 30 Hz for images with dimensions of  $272 \times 512$ px.

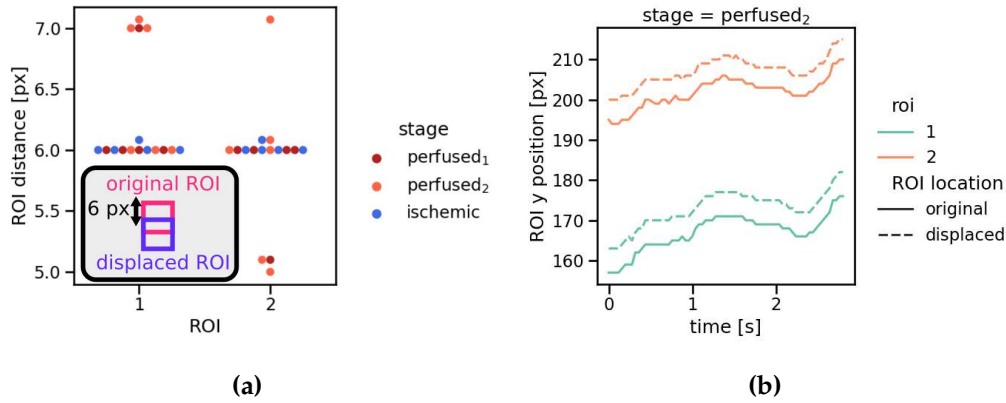
We also evaluated the effect that small variations of the initial ROI location have on the subsequent automatic tracking. We tracked the same manually annotated ROIs while shifting the initial ROI location either vertically or horizontally by 6 px, this is equivalent to 20% of each ROI size. Each tracking (original and displaced) was repeated five times and the distance between original and displaced ROIs was computed by  $d_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$ , where  $(x_i, y_i)$  represents the location of the upper-left corner of ROI  $i$  and each image of the sequence. Fig. 3.17 shows the distribution of the distance between the original and displaced ROIs for all patients and image sequences (perfused<sub>1</sub>, perfused<sub>2</sub>, ischemic) in our clinical study (Sec. 3.3.2). Here, each point corresponds to the median distance between original and displaced ROIs across multiple spectral images for each tuple: (patient, tissue state, ROI). This analysis shows that the distance between original and displaced ROIs remains around its original distance (6 px), thus making our tracking process robust to small variations of the initial manual initialization. We exemplify this robustness in Fig. 3.17. Here, the positions of the original and displaced (by 6 px) ROIs are shown for the perfused<sub>2</sub> image sequence of patient 1. This illustrates that the distance between original and displaced ROIs remains approximately constant, similar results were observed for the rest of the patients and image sequences.

### Ischemia monitoring by Out-of-Distribution (OoD)

**Research question :** How does the inter-patient spectral variability compare to the variability across different tissue states (perfused vs ischemic)?

**Research question :** How can the scarce amount of surgical spectral data be exploited to monitor ischemia during surgery?





**Figure 3.17: Our deep learning-based tracker is robust to small variations on the initial Region of Interest (ROI) position initialization. (a)** The distance between the original and displaced ROIs remains around 6 px for the entire tracking sequence. **(b)** ROI position for both original and displaced locations exemplified on patient 1 in the perfused<sub>2</sub> image sequence. The distance between original and displaced ROIs remains approximately constant during the complete image sequence.

**Research question :** Which are the most effective ischemia monitoring approaches for monitoring perfusion changes during surgery?

The spectral data corresponding to our clinical trial, described in Sec. 3.3.2, was recorded with the system presented in Fig. 2.2. To compensate for tissue and camera motion during image acquisition, we utilized our deep learning-based algorithm (see Sec. 2.4.2), which is capable of tracking ROIs within a spectral video. This tracking algorithm was successfully applied in the spectral data collected from 10 patients undergoing partial nephrectomy, and the results were a prerequisite for further analysis and for computing our deep learning-based *ischemia index*. The results of our analysis are illustrated exemplary in this video. The visualization of the normalized median reflectance spectra within each ROI (Fig. 3.18 a)) reveals that there is a separation between the spectral data of different tissue states (perfused vs ischemic). The Root-Mean-Square (RMS) contrast within each ROI shows that the band with the maximum contrast is band 11 for almost all patients and stages (Fig. 3.18 b)). Only one ROI of the perfused<sub>1</sub> image sequence differs from this observation, in which case band number nine possesses the highest contrast. Furthermore, we ensured that the number of camera counts on each ROI had a SNR of at least 20 to minimize noise effects, the plot of SNR for each camera band within each ROI is supplied in Sec. A.2

**Algorithm 1:** Region of Interest (ROI) tracking. Here  $n$  represents the number of recorded spectral images,  $I_j$  is the  $j^{\text{th}}$  image of the sequence,  $F$  represents the image being visualized to the annotator,  $N_r$  is the number of tracked ROIs on each tracking cycle,  $N = 15$  is the maximum number of times the tracking process is repeated before tracking failure,  $N_f = 100$  is the number of consecutive images where the ROIs are tracked,  $R = 1500$  is the number of images recorded in  $\approx 60$  s and defines the maximum number of times the image requirements are checked before failure.

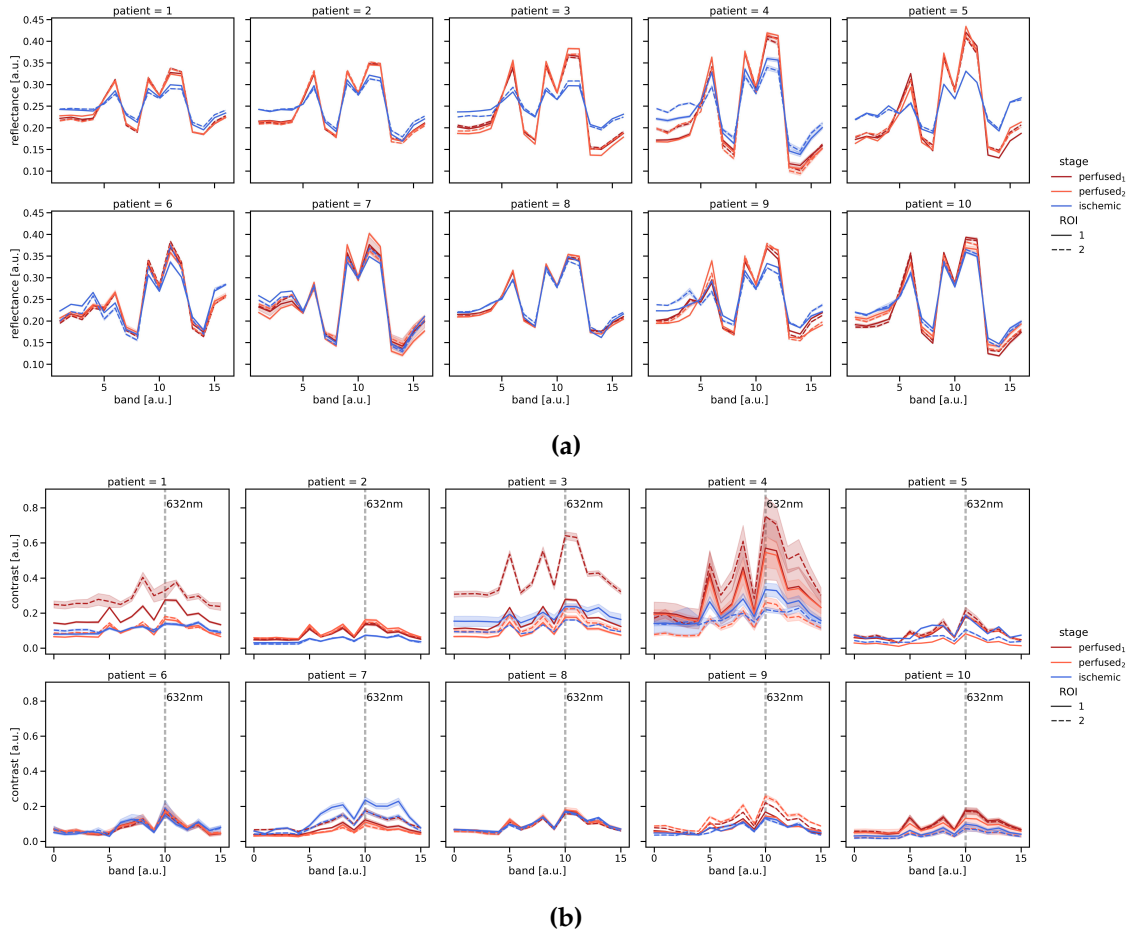
---

```
Requires:  $n \geq 100$ ;  
Input :  $\{I_1, I_2, \dots, I_n\}$  (spectral image sequence)  
 $j \leftarrow 0, F \leftarrow I_0$ ;  
Visualize  $F$  continuously;  
 $i \leftarrow 0, k \leftarrow 0, r \leftarrow 0$ ;  
while  $r < 2 \ \& \ k < R \ \& \ i < N$  do  
  if  $F$  fulfills image requirements then  
    Annotate at least  $N_r$  ROIs;  
    Track ROIs for  $N_f$  consecutive images;  
     $i \leftarrow i + 1$ ;  
    if Tracking fulfills ROI quality requirements then  
       $r \leftarrow r + N_r$ ;  
      Exclude oversaturated and undersaturated pixels;  
    else  
      if There are unannotate regions in first image of current sequence then  
        Exclude previously annotated regions where tracking fails;  
      else  
         $j \leftarrow j + N_f, F \leftarrow I_j$ ;  
      end  
    end  
  else  
     $k \leftarrow k + 1, j \leftarrow j + 1, F \leftarrow I_j$ ;  
  end  
end  
Output:  $r$  tracked ROIs
```

---

**High inter-patient variability for kidney tissue** High inter-patient variability is, in general, an indication of poor generalizability of supervised learning algorithms to unseen patients. In a porcine study of healthy animals, we showed that the biggest source of spectral variability arises from spectral images of different organs rather than the recorded individual or specific acquisition conditions, such as different angles between the camera viewing direction and the tissue surface (Studier-Fischer et al., 2021). However,

### 3.3. Perfusion monitoring



**Figure 3.18: A separation of the normalized spectra can be observed between perfused and ischemic tissue. (a)**  $\ell_2$ -normalized median diffuse reflectance spectra within each Region of Interest (ROI), different colors correspond to different tissue state, while different line styles represent the different ROIs. **(b)** Root-Mean-Square (RMS) contrast for each band and each patient. Across almost all patients, the band 11 is the one with the highest contrast, only one ROI of patient 1 deviates from this tendency. The translucent band around each line plot corresponds to the standard deviation across all annotated spectral images. The vertical dashed lines correspond to the camera band with the highest contrast for the majority of patients.

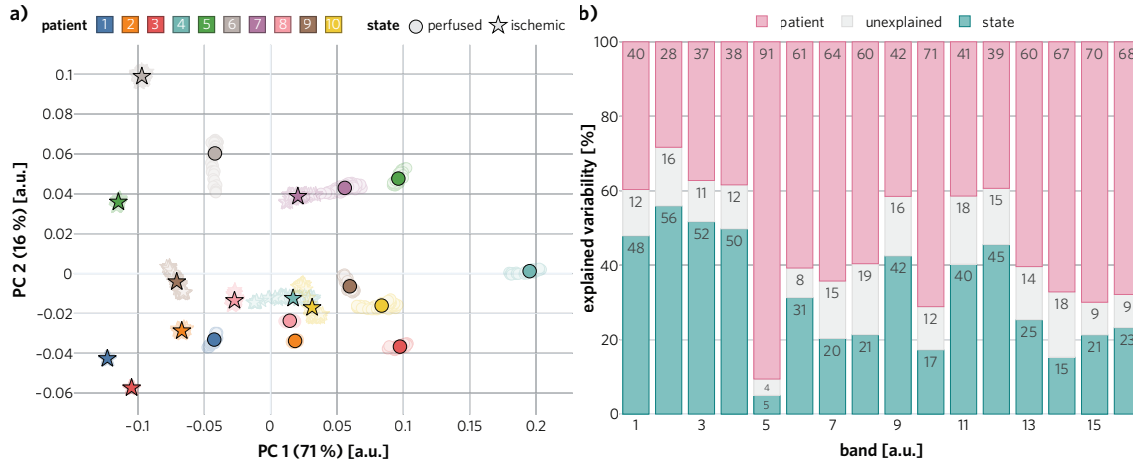
a similar analysis with mixed models applied to the data of our clinical study (Sec. 2.4.2) showed that the spectral data associated with kidney tissue of patients undergoing partial nephrectomy is highly heterogeneous. Fig. 3.19 shows the two-dimensional representation of the median spectra within each ROI (Sec. 3.3.2) via PCA (Hotelling, 1933). It can be seen

that the measurements from different patients gather within clear clusters for each tissue state (indicated by circles (perfused) or a star (ischemic) in Fig. 3.19 a), while clustering of different tissue states across patients cannot be observed (circles and stars do not form clear clusters). Note that the first two principal components (PCs) capture 87% of the variation of the spectral data.

Moreover, we performed a linear mixed model analysis (Sec. 2.4.2) where each patient and each tissue state (perfused and ischemic) represented different components in the model. The data used for this analysis and the two-dimensional representation shown in Fig. 3.19 a) was the median reflectance spectra within each ROI after applying the normalization described in Sec. 2.4.2. This analysis revealed that, in contrast to previous porcine studies, most of the variability in the recorded data can be explained by patients rather than the tissue states (Fig. 3.19 b)). Thus, approaching the challenge of ischemia detection with a supervised algorithm trained on a few patients would suffer from a high risk of poor generalization capabilities to unseen patients. Furthermore, (slight) changes in the acquisition setup, viewing angles etc., might require complete retraining of the method on the entire patient database. This motivated us to develop a personalized approach to ischemia detection, illustrated in Fig. 3.20 where only the data from one individual patient is required.

**Deep learning-based *ischemia index* monitors tissue perfusion** To overcome the limitations that supervised classification approaches suffer in the context of high tissue heterogeneity, we developed an approach to ischemia detection that solely requires spectral data of the individual patient undergoing surgery (see Fig. 3.20).

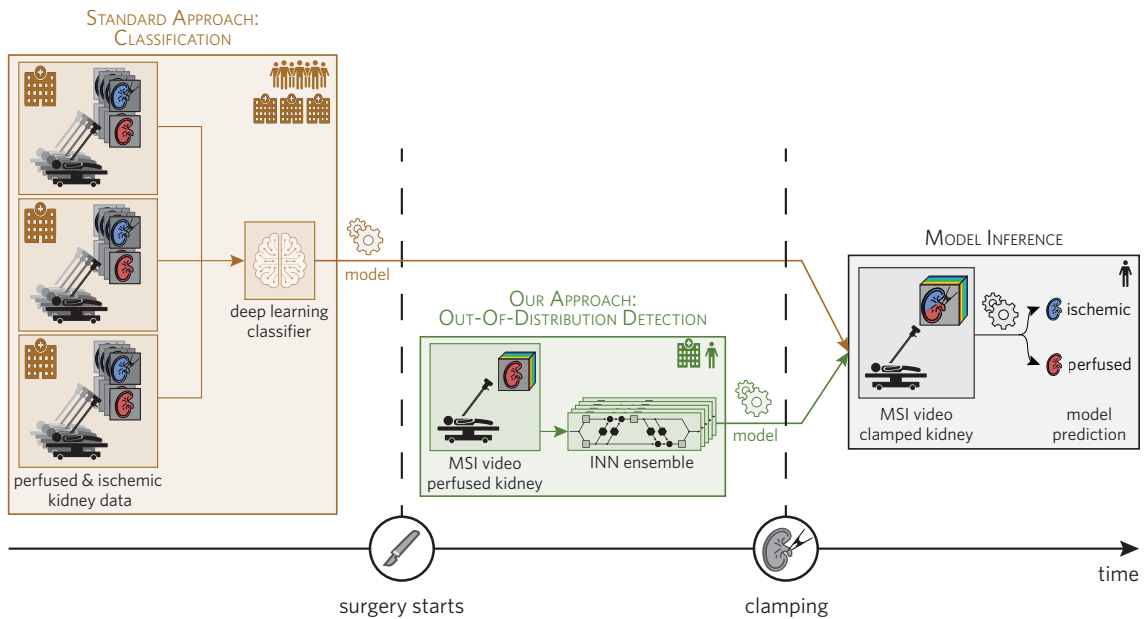
More in detail, we phrase the problem of ischemia detection as an OoD detection problem and determine the perfusion state of biological tissue with an *ischemia index* that relies on an ensemble of Invertible Neural Networks (INNs). As detailed in Sec. 2.4.2, the ensemble of INNs was trained on the normalized ROI data described in Sec. 3.3.2 to approximate the density of the distribution of perfused spectral data. At inference time (after the clamp is applied to the renal artery), the density of the new spectral data is evaluated, and the ensemble predictions are aggregated using the Widely Applicable Information Criterion (WAIC) (Watanabe, 2013). The *ischemia index*, computed for each individual spectral image of the sequence, represents the median WAIC values within each ROI. The data used for training the ensemble of INNs, as described in Sec. 2.4.2, consisted of the two automatically tracked ROIs of size  $30 \times 30$  px (Sec. 3.3.2) located on the perfused tissue (perfused<sub>1</sub>). Training our INNs on different number of spectral images revealed that training on 70 images allowed for faster training times while still providing network convergence. Note that this analysis was only performed on the retrospective patient cohort (patients 1 to 5). In light of this finding, all our analysis was based on the



**Figure 3.19: High tissue heterogeneity across subjects motivated our personalized approach.** Both the (a) principal component analysis (PCA) and (b) the mixed model analysis demonstrate the high inter-patient variability of spectral tissue measurements. (a) The fact that different tissue states cluster within a subject but do not form a uniform cluster across subjects motivated us to phrase the challenge of ischemia detection as an out-of-distribution (OoD) problem. Solid markers show the cluster centers, transparent markers the raw data points, and the axis labels denote the explained variance of the corresponding principal component (PC). (b) Explained variance for the 16 bands of the multispectral imaging (MSI) camera, depicted in Fig. 2.3

first 70 images of each sequence (perfused<sub>1</sub>, perfused<sub>2</sub> and ischemic) where the ROIs were tracked. Fig. 3.21 a) exemplifies the locations where the ROIs were located for patient 2 and the WAIC maps for both testing sequences: perfused<sub>2</sub> and ischemic. Here, higher WAIC values correspond to OoD samples (ischemic) while lower values represent in-distribution samples (perfused). These maps reveal a clear difference between perfused and ischemic tissue.

Furthermore, our patient study, composed of 10 patients undergoing partial nephrectomy, revealed that our *ischemia index* classifies the perfusion state with high accuracy (Fig. 3.21 b)). In fact, in all but one patient, the spectral data corresponding to ischemic tissue could be perfectly separated from those corresponding to perfused tissue. This led to a median Area Under the Receiver Operating Curve (AU-ROC) of 1.0 / 0.9 obtained for the  $n = 10$  patients. To compare our OoD-based approach to the commonly used Beer-Lambert law, we employed the method described in Sec. 2.4.1 to estimate the blood oxygenation in kidney tissue. The spectral data from each ROI was processed in an identical manner as for our *ischemia index* analysis. As depicted in Fig. 3.22, the Beer-Lambert law is incapable of differentiating perfused from ischemic tissue for most of the patients.

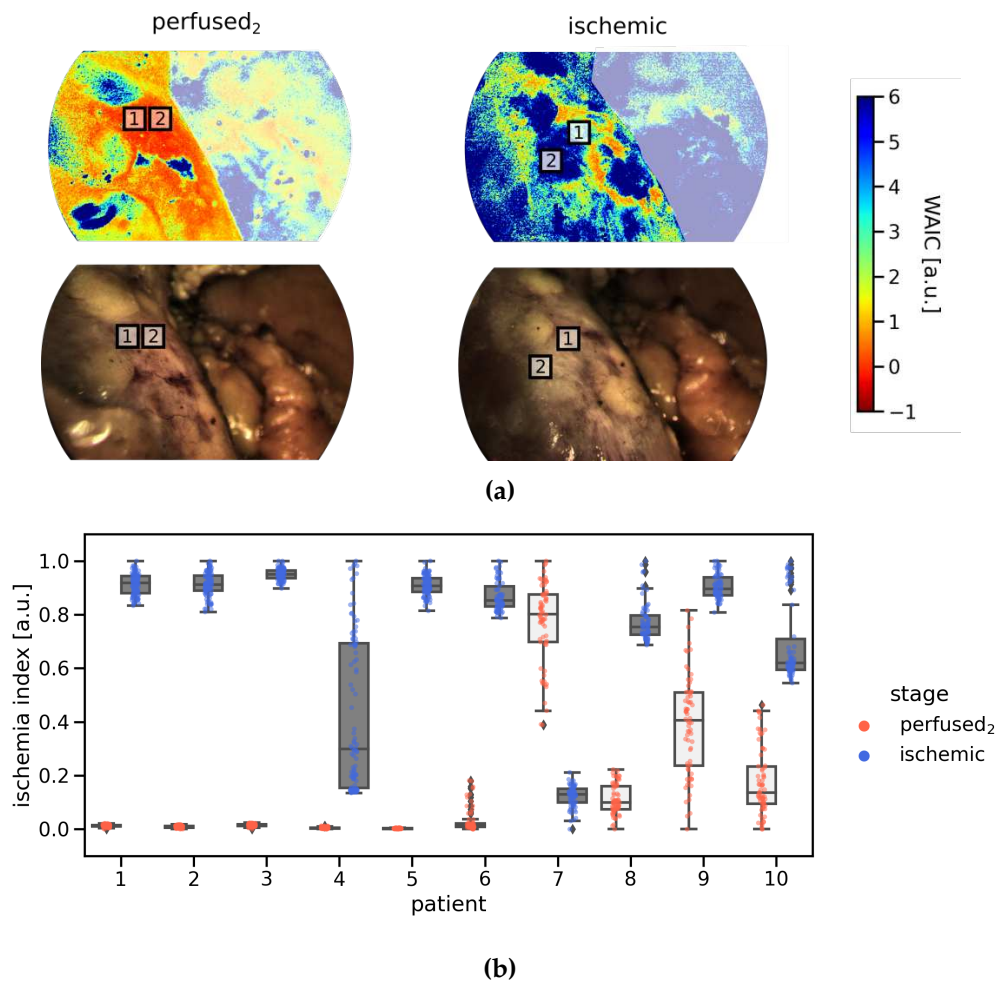


**Figure 3.20:** We treat ischemia monitoring based on multispectral imaging (MSI) as an out-of-distribution (OoD) task which only needs training data from one single patient. Traditional deep learning (DL) methods (brown) require large amounts of patient data to train a model, while our method (green) only needs data from a single patient. Using an ensemble of invertible neural networks (INNs) as a core component, our algorithm computes the likelihood of ischemia based on a short (several seconds) image sequence acquired at the beginning of each surgery. An important feature of our approach is that the entire training and inference process can be performed during a surgical procedure.

Even more, it generates implausible oxygenation values that cannot be interpreted from the physiological point of view.

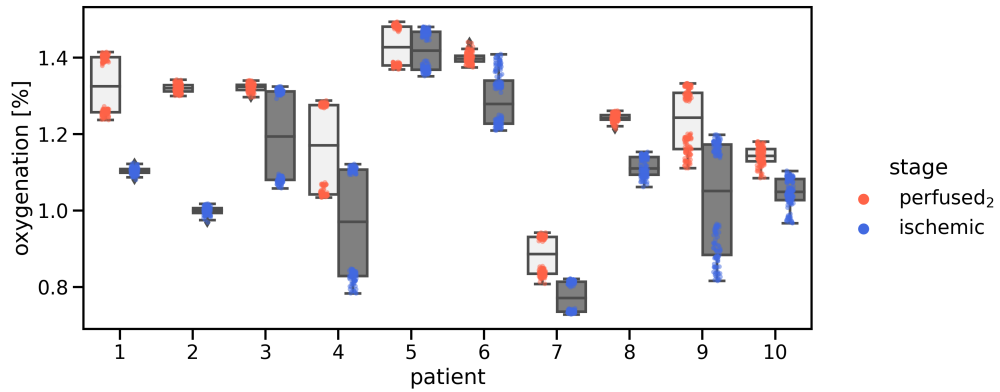
Spectral data is generally scarce in medical applications, so we evaluated its benefit over the more commonly available RGB-data. To this end, we transformed the spectral recordings to RGB images (Sec. 2.4.2) and trained an ensemble of INNs, pre-trained on MC simulations, in the same way as our *ischemia index*. The results of this analysis, presented in Fig. 3.24, show that although the perfusion state could be classified for some patients, it was not possible for most of them.

We studied the hypothesis that accurate MC simulations of light transport on biological tissue, although not representative of pathological tissue in human kidney, could help speed up the training process of our ensemble of INNs. To test this hypothesis, we adapted the MC simulations presented in Sec. 2.2.2 to the imaging system shown in Fig. 2.2 by using the method previously described in Sec. 2.4.2. Each INN was trained on the adapted

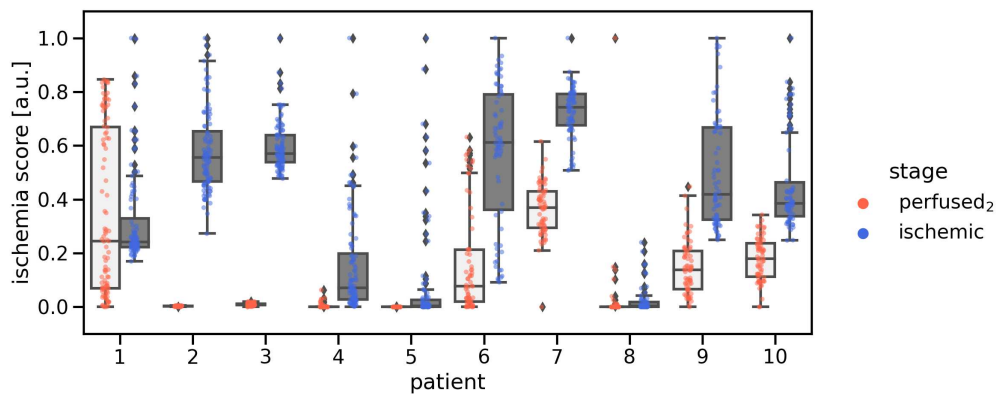


**Figure 3.21: Our approach is capable of discriminating between ischemic and perfused kidney states. (a)** Examples of the raw Widely Applicable Information Criterion (WAIC) values for both image sequences, each annotated ROI is identified by the translucent squares. The raw WAIC values are further processed to generate our *ischemia index*. **(b)** We calculated the *ischemia index* for every image in each sequences of perfused and ischemic kidney separately for each patient and generated corresponding dot and box plots. The boxes show the interquartile range with the median (solid line) and mean (dashed line) while the whiskers extend up to 1.5 of the interquartile range. Min-max-normalization was performed for clarity of presentation.

simulations for 100 epochs using the process described in Sec. 2.4.2, and the training time was recorded. After pre-training, fine-tuning on the patient data was done for only ten epochs while ensuring that the validation loss converged. By doing this pre-training step,



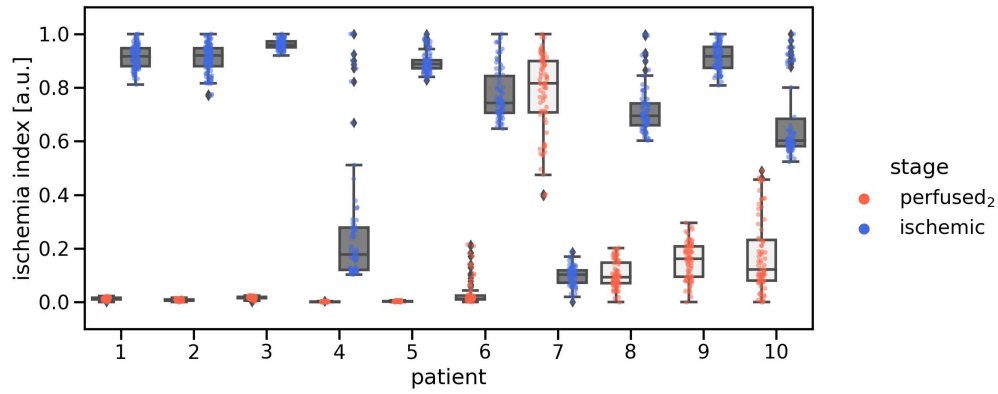
**Figure 3.22: The Beer-Lambert law cannot monitor ischemia in the kidney.** Although Beer-Lambert regression can differentiate ischemic from perfused tissue for some patients (1,2 and 8), a clear separation is no visible for the rest of the patients. The Beer-Lambert law also generates implausible oxygenation values that cannot be interpreted.



**Figure 3.23: Red-Green-Blue (RGB) data does not contain the information required to classify kidney perfusion state.** Although the tissue state (perfused vs ischemic) can be separated for some patients (patients 2 and 3), this is not possible for most of the patients. The so-called ischemia score, evaluated on RGB data, was computed in the same fashion as the *ischemia-score*.

we were able to reduce the total training time by approximately 88% while still being able to detect perfusion changes with high accuracy, as show in Fig. 3.24. While the patient-individual training, after pre-training, took 30 s per network on average, the actual index calculation runs at 120 Hz on the selected ROIs on the kidney, and thus in real time.





**Figure 3.24: Pre-training our ensemble of Invertible Neural Networks (INNs) reduced training time while maintaining high accuracy.** The tissue state (perfused vs ischemic) can be clearly separated for most patients (except patient 7). Pre-training the ensemble of INN reduces training time and distinguishes better between tissue states (e.g. patient 9).



*To leave a whisper of myself in the world, my ghost,  
a magna opera of words.*

— BERNARDINE EVARISTO  
Girl, Woman, Other

We presented a concept for the recording and analysis of spectral images in the field of functional imaging for surgical applications (Fig. 2.1). This concept is capable of exploiting large amounts of physics-based simulations of light transport in biological tissue as well as unlabeled real data, which can be tailored to specific domains and tasks. We proposed to aid translation to open surgical applications by integrating a dedicated automatic light source estimation module. Finally, we proposed to monitor perfusion changes with two approaches. The first approach leverages physics-based simulations of light-tissue interactions to regress functional properties. In this way, we were able to overcome the lack of *in vivo* labeled data. The second approach comprises a personalized approach that deals with domains that are not well represented by the light-tissue simulations. In this chapter, we discuss our approach for domain- and task-specific band selection as well as the Spectral Imaging (SI) hardware for functional imaging (Sec. 4.1). Our method for automatic light source estimation is discussed in Sec. 4.2. Finally, both our methods for perfusion monitoring are discussed in Sec. 4.3.

## 4.1 Spectral imaging

The field of optical functional imaging in the biomedical domain has relied on methods and devices that generally do not fit the requirements for translation to clinical practice. Generally, previous approaches feature large devices that are difficult to handle, they are limited to open procedures, and/or are paired with slow image processing components (Clancy et al., 2015; Baker et al., 2014; Jones et al., 2017; Holmer et al., 2018). We proposed to alleviate this limitations by implementing video-rate SI in surgical applications. In addition, we proposed to increase the speed of spectral devices by reducing the amount of data that they need to acquire. We achieved this by implementing a domain- and task-specific band selection module.

In the following sections, we discuss the spectral hardware setup utilized for validating our approaches (Sec. 4.1.1), our approach to simulations of light-tissue interactions

(Sec. 4.1.2), and the effect of domain and task-specific oxygenation estimations (Sec. 4.1.3).

#### 4.1.1 Real world spectral imaging

Building upon the work of Wirkert (2018), we developed a video-rate Multispectral Imaging (MSI) system for laparoscopic surgery, which to the best of our knowledge is the first video-rate multispectral application in a clinical study. The key strengths of our imaging system are the high acquisition speed (above 25 spectral images per second), the low weight (32 g), the small size (camera cube with edge size of  $26 \times 26 \times 31$  mm) and the fact that it operates with a standard clinical light source. All the changes integrated in our system, in comparison to standard clinical setups, are attached only to the laparoscope, thus remaining always away from the patient and reducing risks for the patient.

Although fast recording speed is a requirement in minimally invasive surgery, other medical applications can benefit from slower systems that feature higher spatial resolution, such as the one presented in our application to brain imaging (Sec. 3.3.1). Given that Spreading Depolarizations (SDs) is a phenomenon that spans over several minutes (potentially hours) and that the brain cortex features a highly complex network of vessels (Fig.3.11), this medical application benefits greatly from a sensor with high spatial resolution at the cost of slower recordings. Slower sequential recordings come at the cost of motion artifacts that can be challenging to overcome. We were able to mitigate the effect of motion artifacts by implementing an inter-band registration process coupled with an elastic registration across multispectral images. It is worth noting that the complex network of vessels in the brain cortex provided many landmarks that the registration algorithm could leverage. However, other tissue structures might feature more uniform surfaces (e.g. liver), thus increasing the challenge of image registration when a sequential recording strategy is used.

Previous approaches in the field of clinical spectral imaging feature big devices with long recording times (several seconds) that are more suited for open surgery (Dietrich et al., 2021a,b; Köhler et al., 2020; Hu et al., 2020; Takamatsu et al., 2021; Kumar et al., 2020; Kester et al., 2011; Stoyanov et al., 2015). Spectral systems in the context of preclinical studies have so far been based on sequential (slow) image recordings (Ayala et al., 2019; Jones et al., 2017; Clancy et al., 2015; Holmer et al., 2018), e.g. using the filter wheel technology (Ayala et al., 2019; Jones et al., 2017; Soares et al., 2021), thus introducing motion artifacts. The snapshot technique applied in our approach has been explored in parallel by other authors (Ebner et al., 2021) in the context of open surgery. Overall, we are not aware of any video-rate spectral laparoscopic system that has so far been applied in a clinical study during surgery. In contrast to the work by Wirkert (2018), several modifications were made to the optical components around the multispectral camera,

all with the aim of increasing data quality and feasibility of our clinical study on partial nephrectomy. The band-pass filter, originally attached to the camera through an extra C-Mount adapter, is now integrated inside the multispectral camera, thus reducing both the size and the number of parts required for our system. Another change in our spectral system is related to the C-Mount adapter, which we upgraded (20200043, KARL STORZ SE & Co. KG, Tuttlingen, Germany) to increase image quality and focal length as well as to ease the application in our clinical trial.

The snapshot technology used by our multispectral camera (Fig. 2.2) enables fast image recordings, as it records all spectral bands at once. However, these benefits come with multiple limitations. Mainly, the snapshot sensor integrated into our multispectral camera (Fig. 2.2) is realized by arrays of Fabry-Pérot filters, which lead to the second order peaks shown in Fig. 2.3. Since these second order peaks appear in the optical wavelength region and blood absorbs light mainly in this region, they need to be considered when the target application is functional imaging. We approached this via the camera transformation introduced in Eq. 2.1. Even though second order peaks should be kept in mind during data analysis, our proposed method for ischemia monitoring in laparoscopic surgery (Sec. 2.4.2) is not influenced by these second order peaks because Invertible Neural Networks (INNs) learn the density of the spectral data, regardless of the influence such peaks have on the reflectance spectra. More specifically, INNs learn to estimate our *ischemia index* based solely on functional changes (e.g. perfusion) and not on hardware configurations, which remains constant across recordings of the same patient.

Currently, most spectral systems targeting the task of perfusion monitoring exploit data in the visible spectrum. However, the spectrum of biological tissue in the infrared (IR) region often features rich information (Trajanovski et al., 2021). As such, a sensor that combines data from the visible spectrum and the IR region could lead to better perfusion estimations. An additional feature required for the application of spectral systems in clinical practice is the capability to provide at the same time: spectral data and reliable Red-Green-Blue (RGB) images, which can be interpreted by the surgical staff. This could be realized, for example, by integrating dedicated RGB filters in the mosaic pattern of a snapshot sensor (Fig. 2.2) while keeping the rest of the pixels for spectral data acquisition.

### 4.1.2 Synthetic spectral imaging

We proposed that more representative tissue properties (e.g. oxygenation) of multi-layer tissue models can be assigned by using the photon penetration depth, which is calculated only once at simulation time. The visualization of the light penetration depth distribution (Fig. 3.1 b) revealed that the majority of photons do not traverse the entire depth of the multi-layer tissue models, thus accentuating the need for a labeling strategy that

considers the penetration depth. Our experiments show that for most wavelengths, in the visible range, the error generated through this new labeling strategy is lower than the one generated by previous work, that is, up to 600 nm (Fig. 3.2 a)). Above 600 nm, our approach yields slightly higher error compared to the baseline approach. This could be because higher wavelengths are related to higher penetration depths. In the limit, where the penetration depth equals the thickness of the tissue, averaging up to 250  $\mu\text{m}$  yields similar results as our approach. Despite these differences above 600 nm, our method was able to reduce the error by 28 p.p. in average on the wavelength range 450–600 nm, while the state of the art performed only 11 p.p. better in the wavelength range 600–650 nm. Furthermore, training our deep learning model for oxygenation estimation based on these two labelling strategies produced similar results (Fig. 3.2 b)), with the model trained with our labeling strategy performing slightly better. It is particularly challenging to interpret these model training results because both models have seen different data labels during training. However, given that our labeling strategy generates more physically accurate results (less error during re-simulation), we decided to use it for the rest of the regression experiments.

Future work in this topic could focus on more accurate estimations of the penetration depth, even more, the absorption profile along the entire 3D model could be leveraged to assign weights to the tissue properties on each layer. Although this approach might yield slight improvements in a model simulated with infinitely wide slabs, we hypothesize that it becomes more relevant when simulating heterogeneous tissue with other tools such as the **SIMPA toolkit** developed by Gröhl et al. (2022). Moreover, the current implementation of the Mone Carlo Multi-Layer (MCML) framework does not consider complex cross-pixel interactions because each pixel is modeled with homogeneous infinitely wide slabs of tissue. Although other frameworks facilitate the computation of diffuse reflectances in heterogeneous media (Gröhl et al., 2022; Yan and Fang, 2020), there is little knowledge of cross-pixel interactions which might become relevant when modeling tissue with high spatial complexity (e.g. vessels, tumor borders, etc.). Addressing this void in knowledge as well as the lack of information of optical properties from various medically relevant tissue structures could prove beneficial to the advance of the generation of realistic synthetic data for biomedical applications.

### 4.1.3 Task-specific band selection

We investigated whether a small subset of bands can provide enough information for specific tasks, mainly tasks related to functional imaging (e.g. oxygenation). This subset of bands can depend on the domain of interest (e.g. type of organ), so we proposed a method to select bands depending on the target domain without the need of labeled *in vivo* data.

Our method leverages physics-based simulations of light-tissue interactions generated with a Monte Carlo (MC) method. We then exemplified the benefits of our approach in the task of oxygenation estimation and the domain of head and neck tumor in a rat model. It is worth mentioning that although we showed the application of our method in a specific task and domain, our methodology can easily be extended to other tasks and domains. One of our main findings was that *wrapper* methods outperform *filter* methods in the task of oxygenation estimation. Here, Best First Search (BFS) was the first ranking band selection method (Fig. 3.3), closely followed by Sequential Forward Selection (SFS). This finding could be explained by the fact that *wrapper* methods leverage the performance of the regressor (here random forests) during band selection, while *filter* methods remain agnostic to it.

The choice of data normalization that we used for this analysis is particularly interesting. Since we normalize each sample by the  $\ell_2$  norm across bands (Eq. 2.8), a re-normalization step is required whenever a new feature is added or removed. We investigated the effect of normalization strategy by implementing the image quotient norm proposed by Styles et al. (2006). By using the image quotient norm, the re-normalization step is unnecessary and therefore eliminates the possibility for the band selection algorithms to jointly optimize the regressor and the normalization. The *in silico* experiments showed higher overall Mean Absolute Error (MAE)s, and the algorithms never surpassed the baseline when training with all available bands. Furthermore, we did not observe any influence of the normalization strategy on the relative performance of the different band selection methods (ranking was the same).

We were further able to reduce the domain gap between the simulations and the *in vivo* data from a head and neck tumor in a rat model (Fig. 3.4 a)). We observed that the domain-selected simulations were in agreement with the real data and that the performance of our regressor surpassed the performance of the baseline (101 bands) when selecting only 20 bands with BFS. This analysis revealed a stronger contrast between the oxygenation distributions of tumorous and healthy tissue when using the subset of 20 bands (Fig. 3.4 c). However, given the lack of a gold standard measure of oxygenation *in vivo*, a quantitative comparison is not possible. Future biomedical research in the field of band selection as well as domain adaptation could prove useful by, for example, reducing the potential bias in many medical datasets. More particularly, domain adaptation methods could help either sample datasets to generate a balanced number of samples across target categories (patients, genders, ethnicities, etc.), or augmenting existing datasets to balance underrepresented categories.

## Conclusions

We addressed the challenge of video-rate SI by implementing a system that can be used in clinical trials. We proposed to increase the speed of spectral systems by reducing the amount of data that they need to acquire. We achieved this by implementing a domain- and task-specific band selection method, which is capable of leveraging unlabeled *in vivo* data to adapt to specific domains. We validated this approach in an *in vivo* rat head and neck tumor model. By reducing the amount of data that spectral devices need to acquire to achieve good performance in specific tasks, we tackled one of the key challenges that translational functional imaging requires, that is, high speed.

## 4.2 Automatic light source estimation

To the best of our knowledge, we are the first to propose an approach to light source (LS) estimation that could easily be applied to spectral imaging in a surgical setting. The methodological innovation that we have presented is mainly related to the LS estimation. We confirmed in our experiments that specular reflections extracted from low exposure spectral images can be processed to recover the LS spectrum with high accuracy. We showed that by leveraging specular reflections to estimate the LS spectrum, higher quality oxygenation estimations can be obtained *in silico* (lower estimation error in Fig. 3.6). This demonstrates that the oxygenation estimations computed with our method are very close to the ground truth. While we performed the optimization of the hyperparameters  $T_{exp}$  (exposure time) and  $N_p$  (number of pixels to aggregate) on a subset of the LS used in our study, we did not observe a decrease in accuracy on the *test LS*. This could indicate that these hyperparameters are independent of the type of LS. Furthermore, our experiments showed that our method outperforms all the competitor state-of-the-art approaches introduced in Sec. 2.3.2 on an identical dataset. Moreover, our method is more robust to changes in the exposure time of the multispectral camera in comparison to the state of the art. This could be explained by the fact that our approach exploits only the information contained in specular reflections, thus making it more robust to changes of the biological tissue (background).

While we have shown the great potential of our methodology for automatic LS estimation, there are several challenges that need to be addressed before a successful clinical translation can be realized. Foremost, our method assumes that the LS illuminates the surgical scene homogeneously in the entire field of view of the camera. This has proven to be a good approximation in initial experiments in a surgical environment; however, more accurate results could be obtained by extending our method such that the LS spectrum could be approximated for different image patches. Depending on the patch size, this ap-



proach could also handle multiple light sources. Secondly, we adapt the machine learning oxygenation regressor by choosing a pre-trained model from a discrete set of regressors, each corresponding to a different LS. A more complete approach would involve adapting a single regressor to dynamically changing illumination conditions. This could be achieved by training the regressor on the MC simulations presented in Tab. 2.1 while conditioning it to the spectrum of known surgical light sources (e.g. Xenon, Light-Emitting Diode (LED), etc.). Furthermore, while we demonstrated the potential of our approach in a wide variety of LSes (Xenon, LED, fluorescent, etc.), clinical validation is required by using a bigger set of light sources commonly found in the Operating Room (OR). This analysis could include LSes of the same type but different vendors as well as new LSes such as overhead lamps and forehead torch lamps.

Our method leverages the information provided by specular reflections in the surgical scene under study. Although specular reflections are very common in optical imaging of internal organs, the presence of specular reflections should be regarded as a key requirement for other types of biological structures where such reflections might be scarce (e.g. skin). A fairly simple quality assurance measure could involve analyzing the camera counts over the entire scene, the presence of specular reflections, in most cases, would be manifested as pixels with high number of counts (either saturated or close to saturated). Finally, our method requires the analysis of low exposure images, which could be regarded as a disruption of the surgical workflow. Although the recording of such low exposure images only requires a few seconds, this could be compensated by implementing a method for illumination change detection, which would trigger the acquisition of low exposure images “on demand”. Given the high accuracy of our method compared to related methods, along with the high processing speed (currently  $\sim 50$  ms), we believe that an occasional acquisition of low exposure images ( $\sim 1$ – $2$  s) is acceptable.

### Conclusions

Open surgery presents many challenges. Its highly dynamic nature is manifested, for example, by the variety of light sources that are used in the OR (e.g. overhead lamps, ceiling lights, forehead torches, etc.). This allows surgeons to use multiple types of light sources, but hinders the translation of functional parameter estimations (e.g. oxygenation). We presented a method for light source spectrum estimation that leverages the information contained in specular reflections by recording low exposure images. We also demonstrated that using this approach can reduce the oxygenation estimation error *in vivo*.

**Clinical applicability:** Our method exploits low exposure spectral recordings, which only require  $\sim 1$ – $2$  s, thus minimizing modifications to clinical routine. Furthermore, our method is robust to different integration times, thus making it more robust than other

related approaches.

**Benefits for the patient:** Reducing functional parameter estimation errors during surgical procedures could improve surgical outcome by providing more accurate information to the surgeon.

**Benefits for the surgeon:** Reducing functional estimation errors by using our approach would provide surgeons with more accurate decision-making information that they can leverage to improve patient treatment.

**Generalization:** Although we have demonstrated the benefits of our method for light source estimation *in vivo*, a study cohort is required to analyze its generalization capabilities to a wide range of patients.

### 4.3 Perfusion monitoring

We approached the challenge of perfusion monitoring with two methods. The first of them overcomes the lack of *in vivo* labeled data by leveraging physics-based simulations of light-tissue interactions (Sec. 4.3.1). However, some surgical domains feature high inter-patient heterogeneity and might not be well represented by the light-tissue simulations. Hence, we proposed a personalized method that utilizes data from individual patients to monitor ischemia at video-rate during minimally invasive surgery (Sec. 4.3.2).

#### 4.3.1 Perfusion monitoring via regression

We presented a deep learning model for perfusion monitoring (exemplified in oxygenation estimation), based on the work by Wirkert (2018), which leverages highly detailed physics-based simulations of light-tissue interactions (Tab. 2.1) generated with a MC method. Compared to previous work, we introduced three key aspects: a) more physically correct multi-layer tissue model labeling strategy, b) extensive neural architecture search and c) custom data augmentation strategies.

While leveraging the MC simulations, our *in silico* experiments revealed a general tendency of our deep learning model, based on a Convolutional Neural Network (CNN) architecture, to perform better in the task of oxygenation estimation for smaller learning rates and larger batch sizes and relatively intermediate values of weight decay. The number of filters per hidden layer as well as the type of activation (Rectified Linear Unit (ReLU) or Leaky Rectified Linear Unit (LeakyReLU)) did not play a major role in the final performance of our models. To train our models with more realistic data, we implemented several data augmentation strategies. Brightness variations of a random subset of samples mimic possible changes in the hardware setup during acquisition of real data (e.g. camera pose or light source intensity). Specular reflections were specifically

introduced through a physics model (*dichromatic model*). Thus, this type of physics-informed neural network learns (during training) about the behavior of specular regions, which is very common in biomedical images.

Although using this types of data augmentation generally led to more stable training and lower validation losses, the lack of a gold standard to estimate oxygenation *in vivo* complicates the analysis of the effect that this strategy might have in a clinical application. We hypothesize that a new generation of physics-informed deep learning models could help bridge the gap between deep learning and clinical applications of oxygenation quantification.

We exemplified the benefits of our deep learning approach compared to the state of the art in our two medical domains (pig brain and human skin). In both applications, our model performed better than state-of-the-art approaches. More in detail, our deep learning model for oxygenation estimation lead to lower oxygenation estimation error *in silico* when compared to the random forest proposed by Wirkert et al. (2016b) (Fig. 3.10). We were able to demonstrate that our model is capable of monitoring perfusion changes *in vivo* through the visualization of SDs in pig brain (Fig. 3.11) and oxygenation changes in an arm-clamping experiment (Fig. 3.14). The success in both of this medical applications shows great promise for future clinical research. On one hand, being able to monitor SDs in the human brain could lead to the development of new drugs that target their spread and even completely counteract their harmful effects. More specifically, future research based on our methodology could lead to automatic detection, tracking and classification of SDs. This type of analysis would also be facilitated by leveraging a high-resolution snapshot sensor instead of sequential band measurements, thus removing the need for registering images across bands. On the other hand, monitoring skin perfusion changes could prove helpful in the field of inflammatory research, where changes of functional properties (e.g. perfusion, water content, etc.) can be associated with many diseases such as sepsis (Dietrich et al., 2021b,a).

Future research could focus on the development of an uncertainty measure that can be assigned to our functional parameter estimations. This could be done, for example, by leveraging our Out-of-Distribution (OoD) approach to compare the MC simulations to real data Sec. 2.4.2. Furthermore, future work in the field of oxygenation estimation for medical applications could benefit from more realistic simulations of light-tissue interactions. In this regard, simulating heterogeneous 3D tissue models instead of multi-layer models would lead to results that are more in agreement with real data. Furthermore, pairing 3D simulations of light transport with a physics-based rendering module would provide a powerful tool in many areas of research. For example, it could provide a way to alleviate the lack of spectral data that is commonly known in the spectral applications of the biomedical community. Moreover, the clinical translation of oxygenation estimation

approaches would benefit from a systematic clinical validation strategy. Despite the success of many regression algorithms, the lack of a gold standard method to estimate oxygenation *in vivo* has held back the progress in this area. Thus, proper clinical validation becomes a requirement for any of these methods to be applied in clinical practice.

### 4.3.2 Perfusion monitoring via Out-of-Distribution detection (Clinical study)

Spectral imaging is a tool with a high potential to tackle challenging problems in the early detection, diagnosis and monitoring of diseases ranging from cancer to cardiovascular and inflammatory diseases. To the best of our knowledge, we have presented the first clinical trial for video-rate ischemia monitoring in laparoscopic surgery enabled by deep learning and MSI. Based on a compact video-rate multispectral laparoscopic imaging system, our approach presents three key advantages over the current state of the art in ischemia monitoring:

1. **No contrast agent:** Our method is completely non-invasive and does not require the administration of Indocyanine Green (ICG) or any other contrast agent. As such, the application of our method has the potential to enhance patient safety by reducing complication risks (e. g. anaphylactic shock).
2. **Multiple uses per surgery possible:** Our method can be used multiple times during surgery. There is no washout period.
3. **Video-rate ischemia monitoring:** Our proposed *ischemia index* can be computed in real time, while the classical ICG approach allows for only a single measurement per surgery.

In the following sections, we discuss the consequences of high tissue heterogeneity, our approach to ischemia detection and a comparison between our proposed approach and current clinical practice.

#### Inter-patient tissue spectral heterogeneity

The work we have presented comprises the first systematic *in vivo* analysis of spectral data from diseased kidney tissue across multiple patients. Our analysis revealed a high inter-patient variability of spectral data. One of the main findings was that the main source of variance of spectral measurements is the patient rather than the perfusion state of the kidney (perfused vs ischemic). This indicates that a supervised machine learning approach could generalize poorly to unseen patients when trained on limited patient data. In addition, the two-dimensional representation of the spectral data computed by Principal

Component Analysis (PCA), and depicted in Fig. 3.19 a) shows that the distance between data from perfused and ischemic tissue varies considerably across patients (distance between circles and stars of the same color). This indicates that the spectral difference between perfused and ischemic tissue is also heterogeneous across patients.

In contrast to recent findings by Studier-Fischer et al. (2021), where the variance arising from different specimens was small compared to other explanatory variables. A possible explanation for this discrepancy is that in the porcine study all organs were healthy, whereas all the human subjects in our clinical trial had kidney cancer (Tab. 3.1). This is an indication that diseased organs might lead to high inter-patient variability.

The lack of a commercially available spectral system for surgical applications limits the number of patients that could be included in a clinical trial from multiple institutions, thus making it hard to determine the mechanism for the change in spectra. In particular, the combination of cancer type and comorbidities is unique for each patient. Furthermore, the highly dynamic environment of any OR offers a significant number of external factors that cannot be controlled. For example, the pose of the laparoscope relative to the kidney changes between each patient due to variables such as port placement, respiratory movements, different anatomy, etc. Since the laparoscope is operated freehand, additional motion induced by the surgeon's hand is involved. Even though we implemented measures to reduce the effect of these factors, such as data normalization, further analysis of these factors is required. In general, the mismatch imbalance between number of patients and the number of possible variable factors makes the association of spectral variability to each factor a challenging problem. Nonetheless, these factors are included as the remaining unexplained variance in Fig. 3.19 b).

Furthermore, we observed that the explained variability of bands 5 and 10 attributed to changes in tissue state (perfused *vs.* ischemic) is especially low in comparison to the variability explained by different patients (Fig. 3.19). This could be due to two main reasons: a) The maxima of the camera filter responses corresponding to these bands are located at wavelengths where the extinction coefficient of Hb and HbO<sub>2</sub> are very similar (Fig. 2.3). Thus, this leads to lower variability between perfused (high HbO<sub>2</sub>) and ischemic (low HbO<sub>2</sub>) tissue. b) Band 5 is affected by a second order peak with a maximum filter response located at a wavelength where the extinction coefficients of Hb and HbO<sub>2</sub> are considerably smaller, thus leading to higher reflectances and a lower influence of different tissue states (perfused *vs.* ischemic).

In the future, our linear mixed model would allow for a more detailed understanding of how different factors influence spectral variability by introducing more factors such as comorbidities. However, with the frequency of partial nephrectomies being limited, and the lack of a commercially available medical spectral system, additional data acquisition is challenging. These results also open the way to an in-depth band selection analysis

with the aim of designing a clinical device that features bands with low inter-patient variability while maintaining high variability across tissue states. Furthermore, it is worth mentioning that we observed differences in the kidney surface structure between surgeries performed on female and male patients. For instance, the fatty tissue was relatively easy to remove from the surface of the kidney from female patients, while it proved to be a more complicated task in male patients. Although we ensured in our clinical trial that we only recorded regions of the kidney with no visible surface fatty tissue, this key difference could be the subject of future research. For example, systematic spectrometer measurements of human fatty tissue could help characterize its optical properties and thus elucidate where the differences between male and female come from\*.

### Ischemia index

We proposed a new approach to ischemia monitoring that removes potential biases from oversimplified models (such as those based on the Beer-Lambert law), and that does not require large amounts of data for algorithm training. Our results demonstrate that by rephrasing the problem of ischemia monitoring as an OoD detection problem, difficulties introduced by the high inter-patient variability can be removed. This led to a personalized approach that only requires data from individual patients, thus avoiding potential confounder effects otherwise introduced by data from other patients (Dietrich et al., 2021b; Zhao et al., 2020). This feature is extremely relevant considering the evidence that many research studies overestimate the performance of deep learning models, mainly due to poor test data selection relative to the training data (Roberts et al., 2021; Shad et al., 2021). Even more, algorithms trained on data from a specific camera may not generalize to slightly different conditions (Dietrich et al., 2021b; Liu et al., 2020b). Given that our method only requires data from perfused tissue for training, and that the training time is approximately 30 s per network (for five models), application in the OR is feasible. As such, the training procedure of our models can easily be performed at the beginning of each surgery before the localization of the renal artery (before clamping), thus avoiding delays in the surgical procedure. Even more, since our model is capable of evaluating spectra at 220 kHz, which translates to an individual Region of Interest (ROI) evaluation rate of 240 Hz, the inference can be done at video-rate.

ICG is the clinical state-of-the-art method to assess ischemia during partial nephrectomy. The main advantages of our method are its non-invasiveness and video-rate capability (Tab. 4.1). One requirement of our approach is the fact that our algorithm needs a clean kidney surface (no fat, no burned tissue, etc.) to perform optimally. Tissue surface effects such as excessive bleeding or scarring may hinder the application of our method, or any

---

\*From an optical perspective.

**Table 4.1:** Key characteristics of traditional ischemia monitoring with Indocyanine Green (ICG) injection and our proposed non-invasive method.

Feature	ICG	Our method
Standard light source	X	✓
Standard camera	X	X
Low per use cost	X	✓
Low investment cost	X	✓
Contrast agent-free	X	✓
Multiple use per surgery	X	✓
Robust to tissue surface imperfections	✓	X
Video-rate	X	✓

other approach relying on spectral data in the visible spectrum. This is because light with wavelength in the visible spectrum has relatively low penetration depth, thus focusing on effects that may occur on the surface of the tissue. However, this disadvantage could be overcome by extending the wavelength range to the infrared range, which is associated with deeper light penetration.

Many state-of-the-art approaches in preclinical studies rely on the Beer-Lambert law to assess perfusion (Baker et al., 2014; Jones et al., 2017). In our application to ischemia monitoring in laparoscopic surgery, this led to non-interpretable functional estimations, such as oxygenation above 100% (Fig. 3.22). Moreover, a clear separation between perfused and ischemic tissue was only possible for 30% of the patients based on oxygenation estimations. We also observed that the effect of multiple ROIs was clearly visible as separate point clusters for each patient (Fig. 3.22), thus hindering the clinical application of methods relying solely on the Beer-Lambert law.

Other model-based methods aiming to assess tissue perfusion, such as the one applied by the TIVITA<sup>®</sup> cameras (Diaspective Vision GmbH, Am Salzhaff-Pepelow, Germany), require the computation of the first and second derivatives of the spectra (Holmer et al., 2018). Unfortunately, this approach is highly sensitive to noise and requires highly detailed spectral data (recordings from many wavelengths). Thus, it has been restricted to hyperspectral systems that are usually associated with long recording times.

In Sec. 2.4.1, we presented a deep learning approach for oxygenation estimation based on spectral data (Ayala et al., 2019; Wirkert et al., 2017). In this method, we leveraged physics-based simulations of light-tissue interactions via a MC method. Analysis of the data distribution of our clinical trial for ischemia monitoring in laparoscopic surgery revealed that our measurements were OoD in comparison to our MC database. This indicated that our oxygenation estimation method could not be used in this application, and that more work is needed for fully realistic spectra generation in the presence of

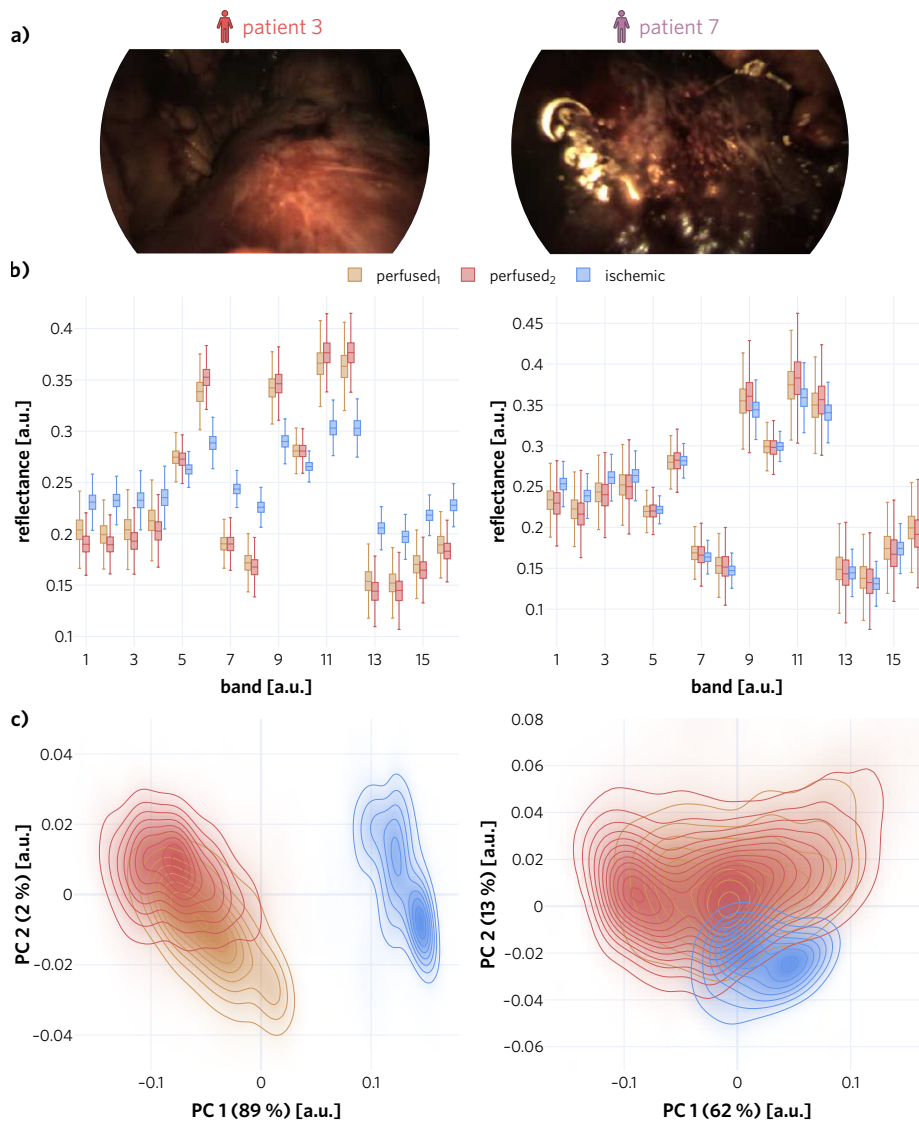
pathologies. To overcome this challenge, we decided to explore the personalized approach presented here.

Even though we were able to monitor ischemia in human kidney, the ROI tracking is a potential confounding factor in our clinical study. Our automatic ROI tracking approach requires the manual localization of each ROI in the surface of the tissue, which could lead to faulty results if the ROI is placed in a region for which our method is not well suited, e.g. fatty tissue. This effect was mitigated by visual inspection of each ROI by trained clinicians, thus ensuring that each ROI was placed in a clean region of the kidney. A translation of our method to clinical practice should take this effect into account, and could be addressed by creating a checklist as the one described in Sec. 2.4.2. Another possible source of bias is related to the reproducibility of the ROI tracking process and from slightly different ROI initialization. Our experiments showed that our ROI tracking can indeed be reproduced when the same ROI location is used over multiple repetitions, and it is robust to small displacements from the original location (Fig. 3.17). Furthermore, future applications of our method could feature the automatic tracking of ROIs with arbitrary shapes to maximize the amount of usable data per spectral frame. In addition, the clinical application of our method would benefit from an uncertainty measure that indicates the probability that changes detected by our ischemia index indeed arise from ischemia induction and not from other confounding factors such as ROI location.

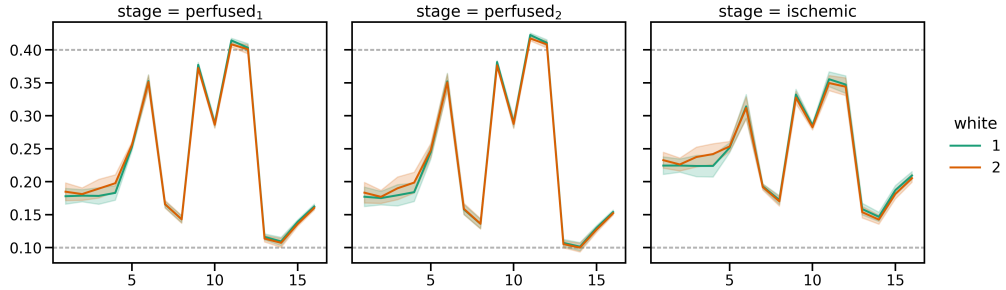
While our method worked perfectly on nine out of ten patients, it failed on patient 7. A more in-depth analysis unveiled that the spectral difference between perfused and ischemic tissue was very low for this patient, as shown in Fig. 4.1. A Kernel Density Estimation (KDE) of the two-dimensional representation of the measurements (computed via PCA) showed that there was a large overlap between the distributions of the perfused and ischemic data. Even more, the variability of the perfused data was considerably high for all bands, compared to the variability of the ischemic data. These two factors might explain why ischemic spectra were erroneously detected as *in domain*. The similarity between the perfused and ischemic spectra could be related to the appearance of burned fatty tissue in the surface of the kidney, which can occur while removing fatty tissue with the da Vinci<sup>®</sup> monopolar scissors. To alleviate problems arising from tissue surface effects (burning, fatty tissue, etc.), we envision increasing the wavelength range of the camera and the light source in the future. It should also be mentioned that the laparoscopic videos of patient 7 looked unusual from a clinical perspective (see Fig. 4.1 a)).

Patient 4 posed a particular challenge in our analysis. Due to data corruption of the white and dark reference measurements, reference measurements from another surgery had to be used for data normalization (Eq. 2.31). Given that hardware, software and Spectralon<sup>®</sup> target were all identical, we did not expect a major influence of this procedure in the analysis of the data from patient 4. In fact, normalizing the data from patient 4 with





**Figure 4.1: Comparison of a representative patient with patient 7 with respect to the algorithm input.** (a) Example Red, Green and Blue (RGB) images were selected from the perfused<sub>2</sub> sequence to illustrate the unusual images acquired of the kidney of patient 7. (b) The reflectances of the representative patient (patient 3) differ clearly between perfused and ischemic tissue, while no clear separation can be observed in patient 7. (c) A kernel density estimation (KDE) performed on the data resulting from a patient-specific principal component analysis (PCA) provides further explanation for why our method falsely detected the ischemic kidney data of patient 7 as inlier. The axes labels denote the explained variance of the corresponding principal component (PC).



**Figure 4.2: Variability across different normalization strategies is insignificant compared to variability across tissue states (perfused *vs* ischemic).** Reflectance spectra of patient 4 after normalizing the data with reference measurements from patients 1 (white 1) and patient 3 (white 2).

reference measurements of multiple other patients showed very low variability across different normalization approaches in comparison to the variability across different tissue states (perfused *vs* ischemic). This effect can be visualized in Fig. 4.2.

Interestingly, our method is capable of adapting to virtually any type of optical system (laparoscopes, filters, etc.), provided that sufficient light reaches the camera sensor, however, the effect of different Signal-to-Noise Ratio (SNR) levels in the training data is unclear. While training our OoD-based ischemia monitoring algorithm, Gaussian noise was added to the data as described in Sec. 2.4.2. Although we found that adding multiplicative noise with a standard deviation of 0.05 led to stable training, further analysis could be performed to study the effect of different SNR levels, as well as other data augmentation strategies, on our final *ischemia-index* estimations. For example, the effect on the performance of our model could be studied while training with a noise level of  $SNR_1$  and evaluating it in data with a noise level of  $SNR_2$ . This analysis would shine light on the generalization capabilities of our methodology.

One of the biggest challenges in this clinical study was the need to remove the laparoscope attached to our multispectral camera (Fig. 2.2) between recordings of the perfused and ischemic tissue. This was because, continuous measurements would require the generation of an additional surgical port in the abdominal cavity of the patient, thus greatly modifying the standard clinical procedure. To reduce stress to the patient, we used the same port that is required for the clamping of the renal artery. This made temporary retraction of our camera necessary, to perform the clamping, thus complicating the ischemia detection because analysis on the same ROI across tissue states could not be guaranteed. In some patients, it was not even manually possible to robustly locate the same tissue region before and after clamping the renal artery due to the lack of physiological landmarks and

the altered pose of the laparoscope. This was a result of the necessary laparoscope removal before clamping the renal artery. We designed our method such that it can leverage data from two non-overlapping ROIs to mitigate the bias that different tissue regions might introduce in our analysis. In addition, we ensured that all our image sequences (perfused<sub>1</sub>, perfused<sub>2</sub> and ischemic) were recorded under the same conditions (after removal and re-insertion of the laparoscope). If our method was to classify the tissue location instead of the perfusion state, it would have failed to correctly classify the perfused<sub>2</sub> sequence. Of note, the challenge of interrupted image acquisition could easily be solved in the future with a clinically certified spectral system, thus enabling spectral recordings and clinical interpretation at the same time.

It is worth mentioning that we investigated the effect of applying our method on reconstructed RGB images (Fig. 3.23). Even though a clear separation between perfused and ischemic tissue could be achieved for two patients (2 and 3), a clear separation could not be observed for the other patients. This analysis also generated considerably wider distributions for each patient compared to the results of the analysis based on spectral data (Fig. 3.24). This result was expected because RGB images lack the rich information across multiple wavelengths that spectral data offers.

## Conclusions

We proposed to monitor perfusion with two approaches. The first approach features a deep learning model that leverages large amounts of simulated data to overcome the lack of *in vivo* labeled data. We validated this approach in different domains: pig brain and human skin (clinical trial). The second approach deals with domains where high inter-patient variability is dominant in comparison to variability across tissue states (perfused *vs.* ischemic). We achieved this by rephrasing the task of ischemia monitoring as an OoD detection task. As such, our personalized approach only requires data from individual patients to detect perfusion changes as outliers. We validated this approach in the first clinical trial involving video-rate ischemia monitoring with SI during minimally invasive surgery.

**Clinical applicability:** Our proposed method could seamlessly be applied in clinical practice, as it leverages existing hardware parts and does not drastically modify current clinical standards. We only integrated the need for a compact multispectral camera that is easy to handle. Furthermore, our deep learning-based ischemia monitoring approach can be trained in under two minutes at the beginning of each surgery, thus making it possible to be applied during surgery.

**Benefits for the patient:** Our method does not require the injection of any contrast agent into the patient's bloodstream, thus reducing burden to the patient.

**Benefits for the surgeon:** The methodology that we have proposed allows for label-free and video-rate ischemia monitoring, thus allowing for multiple repetitions during each surgery and could facilitate surgical decision-making.

**Generalization:** Given that our personalized approach requires only data from individual patients, we mitigated potential biases that might have arisen from other patients.

Many clinical applications currently rely on several imaging modalities such as Positron Emission Tomography (PET), Magnetic Resonance Imaging (MRI), Computed Tomography (CT), etc. All such modalities provide valuable patient data to the clinical staff to aid clinical decision-making and patient care. Despite the undeniable success of such modalities, most of them are limited to preoperative scans and focus on morphology analysis, e.g. tumor segmentation, radiation treatment planning, anomaly detection, etc. Even though the assessment of different functional properties such as perfusion is crucial in many surgical procedures, it remains highly challenging via simple visual inspection. Functional imaging techniques such as Spectral Imaging (SI) link the unique optical properties of different tissue types with metabolism changes, blood flow, chemical composition, etc. As such, SI is capable of providing much richer information that can improve patient treatment and care. In particular, perfusion assessment with functional imaging has become more relevant due to its involvement in the treatment and development of several diseases such as cardiovascular diseases. Current clinical practice relies on Indocyanine Green (ICG) injection to assess perfusion. Unfortunately, this method can only be used once per surgery and has been shown to trigger deadly complications in some patients (e.g. anaphylactic shock).

This thesis addressed common roadblocks in the path to translating optical functional imaging modalities to clinical practice. The main challenges that were tackled are related to a) the slow recording and processing speed that SI devices suffer from, b) the errors introduced in functional parameter estimations under changing illumination conditions, c) the lack of medical data, and d) the high tissue inter-patient heterogeneity that is commonly overlooked. This framework follows a natural path to translation that starts with hardware optimization. To overcome the limitation that the lack of labeled clinical data and current slow SI devices impose, a domain- and task-specific band selection component was introduced. The implementation of such component resulted in a reduction of the amount of data needed to monitor perfusion. Moreover, this method leverages large amounts of synthetic data, which paired with unlabeled *in vivo* data is capable of generating highly accurate simulations of a wide range of domains. This approach was validated *in vivo* in a head and neck rat model, and showed higher oxygenation contrast between normal and cancerous tissue, in comparison to a baseline using all available bands.

The need for translation to open surgical procedures was met by the implementation of an automatic light source estimation component. This method extracts specular reflections from low exposure spectral images, and processes them to obtain an estimate of the light source spectrum that generated such reflections. The benefits of light source estimation were demonstrated *in silico*, *in ex vivo* pig liver, and *in vivo* human lips, where the oxygenation estimation error was reduced when utilizing the correct light source estimated with this method. These experiments also showed that the performance of the approach proposed in this thesis surpasses the performance of other baseline approaches.

Video-rate functional property estimation was achieved by two main components: a regression and an Out-of-Distribution (OoD) component. At the core of both components is a compact SI camera that is paired with state-of-the-art deep learning models to achieve real time functional estimations. The first of such components features a deep learning model based on a Convolutional Neural Network (CNN) architecture that was trained on highly accurate physics-based simulations of light-tissue interactions. By doing this, the challenge of lack of *in vivo* labeled data was overcome. This approach was validated in the task of perfusion monitoring in pig brain and in a clinical study involving human skin. It was shown that this approach is capable of monitoring subtle perfusion changes in human skin in an arm clamping experiment. Even more, this approach was capable of monitoring Spreading Depolarizations (SDs) (deoxygenation waves) in the surface of a pig brain. Even though this method is well suited for perfusion monitoring in domains that are well represented with the physics-based simulations on which it was trained, its performance cannot be guaranteed for outlier domains. To handle outlier domains, the task of ischemia monitoring was rephrased as an OoD detection task. This new functional estimation component comprises an ensemble of Invertible Neural Networks (INNs) that only requires perfused tissue data from individual patients to detect ischemic tissue as outliers. The first ever clinical study involving a video-rate capable SI camera in laparoscopic partial nephrectomy was designed to validate this approach. Such study revealed particularly high inter-patient tissue heterogeneity under the presence of pathologies (cancer). Moreover, it demonstrated that this personalized approach is now capable of monitoring ischemia at video-rate with SI during laparoscopic surgery.

In conclusion, this thesis addressed challenges related to slow image recording and processing during surgery. It also proposed a method for light source estimation to facilitate translation to open surgical procedures. Moreover, the methodology proposed in this thesis was validated in a wide range of domains: *in silico*, rat head and neck, pig liver and brain, and human skin and kidney. In particular, the first clinical trial with spectral imaging in minimally invasive surgery demonstrated that video-rate ischemia monitoring is now possible with deep learning.

## Zusammenfassung

Viele klinische Anwendungen verlassen sich derzeit auf verschiedene Bildgebungsmodalitäten wie Positronen-Emissions-Tomographie (PET), Magnetresonanztomographie (MRT), Computertomographie (CT), usw. Alle diese Modalitäten liefern dem klinischen Personal wertvolle Patientendaten, um die klinische Entscheidungsfindung und die Patientenversorgung zu unterstützen. Trotz des unbestreitbaren Erfolgs solcher Modalitäten beschränken sie sich auf präoperative Scans und konzentrieren sich auf die Analyse der Morphologie, z.B. der Tumorsegmentierung, der Bestrahlungsplanung und der Anomalieerkennung. Auch wenn die Beurteilung verschiedener funktionaler Eigenschaften wie der Perfusion bei vielen chirurgischen Eingriffen von entscheidender Bedeutung ist, bleibt sie bei einer einfachen visuellen Inspektion eine große Herausforderung. Funktionelle Bildgebungstechniken wie Spektrale Bildgebung (SB) verknüpfen Gewebe spezifische optische Eigenschaften mit Stoffwechseleränderungen, Blutfluss, chemischer Zusammensetzung usw. Als solches ist SB in der Lage, viel reichhaltigere Informationen zu liefern, und dadurch die Patientenversorgung zu verbessern. Insbesondere die Perfusionsbeurteilung mit funktionaler Bildgebung hat aufgrund ihres Einflusses auf die Behandlung und auf die Entwicklung verschiedener Erkrankungen, wie Herz-Kreislauf-Erkrankungen und verschiedene ischämischer Zustände, an Bedeutung gewonnen. Die derzeitige klinische Praxis verlässt sich auf die Injektion von Indocyanine Green (ICG), um die Perfusion zu beurteilen. Leider kann diese Methode nur einmal pro Operation angewendet werden und hat nachweislich bei einigen Patienten zu tödlichen Komplikationen geführt (z.B. anaphylaktischer Schock).

Diese Dissertation adressiert allgemeine Herausforderungen auf dem Weg zur Translation optischer funktionaler Bildgebungsmodalitäten. Die Hauptherausforderungen, die adressiert wurden, beziehen sich auf a) die langsame Aufnahme- und Verarbeitungsgeschwindigkeit, unter der SB-Geräte leiden, b) die Fehler, die bei der Schätzung funktionaler Parameter unter wechselnden Beleuchtungsbedingungen entstehen, c) das Fehlen klinischer Daten und d) die hohe Gewebheterogenität zwischen Patienten, die häufig übersehen wird. Dieser Framework folgt einem natürlichen Translationspfad, der mit der Hardwareoptimierung beginnt. Um die Einschränkung zu überwinden, die der Mangel an gelabelten klinischen Daten und die derzeitigen langsamen SB-Geräte mit sich bringen, wurde eine domänen- und aufgabenspezifische Bandauswahlkomponente einge-

führt. Die Implementierung einer solchen Komponente führte zu einer Reduzierung der Datenmenge, die zur Überwachung der Perfusion erforderlich ist. Darüber hinaus nutzt diese Methode große Mengen synthetischer Daten, die gepaart mit nicht gelabelten *in vivo*-Daten in der Lage sind, hochgenaue Simulationen einer Vielzahl von Domänen zu erzeugen. Dieser Ansatz wurde *in vivo* in einem Kopf-Hals-Rattenmodell validiert und zeigte einen höheren Sauerstoffkontrast zwischen normalem und tumorartigem Gewebe im Vergleich zu einer Basislinie, die alle verfügbaren Spektralbänder verwendet.

Die Notwendigkeit einer Translation in offene chirurgische Verfahren wurde durch die Implementierung einer Komponente zur automatischen Schätzung der Lichtquelle erfüllt. Dieses Verfahren extrahiert Totalreflexionen aus Spektralbildern mit geringer Belichtung und verarbeitet sie, um eine Schätzung des Lichtquellenspektrums zu erhalten, das diese Reflexionen erzeugt hat. Die Vorteile der Lichtquellenschätzung wurden *in silico*, in *ex vivo* Schweineleber und *in vivo* menschliche Lippen demonstriert. Dieser Ansatz reduziert die Fehler bei der Schätzung der Sauerstoffsättigung, wenn die richtige Lichtquelle verwendet wurde. Diese Experimente zeigten auch, dass die Leistung des in dieser Arbeit vorgeschlagenen Ansatzes die Performance anderer Basisansätze übertrifft.

Die Schätzung der funktionalen Eigenschaften mit Video-Bildrate wurde durch zwei Hauptkomponenten erreicht: eine Regression und eine Out-of-Distribution (OoD)-Komponente. Das Herzstück beider Komponenten ist eine kompakte SB-Kamera, die mit state-of-the-art deep learning Modellen gekoppelt ist, um funktionale Schätzungen in Echtzeit zu erreichen. Die erste dieser Komponenten verfügt über ein deep learning Modell, das auf einer Convolutional Neural Network (CNN)-Architektur basiert, die auf hochgenauen, physikbasierten Simulationen von Licht-Gewebe-Wechselwirkungen trainiert wurde. Dadurch wurde das Problem des Mangels an *in vivo*-gelabelten Daten überwunden. Dieser Ansatz wurde in einer Perfusionsmonitoring-Studie im Schweinehirn und in einer klinischen Studie mit menschlicher Haut validiert. Es wurde gezeigt, dass dieser Ansatz in der Lage ist, feine Durchblutungsänderungen in der menschlichen Haut am Arm in eine Abbindungsexperiment zu überwachen. Darüber hinaus war dieser Ansatz in der Lage, Spreading Depolarizations (SDs) (Desoxygenierungswellen) an der Oberfläche eines Schweinehirns zu überwachen. Obwohl diese Methode für Perfusionsmonitoring in Domänen gut geeignet ist, die mit den physikbasierten Simulationen, auf denen sie trainiert wurde, gut übereinstimmen, kann ihre Leistung für Ausreißerdomänen nicht garantiert werden. Um diese abweichenden Domänen zu handhaben, wurde die Aufgabe der Ischämieüberwachung in eine OoD-Detektionsaufgabe umformuliert. Diese neue funktionale Schätzungskomponente umfasst ein Ensemble von Invertible Neural Networks (INNs), das nur perfundierte Gewebedaten von einzelnen Patienten benötigt, um ischämisches Gewebe als Ausreißer zu erkennen. Die erste klinische Studie mit einer video-bildfräquenzfähigen SB-Kamera wurde während laparoskopischen partiellen



---

Nephrektomien durchgeführt, um diesen Ansatz zu validieren. Diese Studie ergab eine besonders hohe Gewebheterogenität zwischen verschiedenen Patienten bei Vorliegen von Pathologien (Krebs). Darüber hinaus wurde gezeigt, dass dieser personalisierte Ansatz nun in der Lage ist, Ischämie mit Videorate mit SB während einer laparoskopischen Operation zu überwachen.

Zusammenfassend befasste sich diese Arbeit mit Herausforderungen im Zusammenhang mit der langsamen Bildaufnahme und -verarbeitung während der Operation. Ein Verfahren zur Schätzung der Lichtquelle wird vorgeschlagen, um die Übertragung auf offene chirurgische Eingriffe zu erleichtern. Darüber hinaus wurde die in dieser Arbeit vorgeschlagene Methodik in einer Vielzahl von Bereichen validiert: *in silico*, Rattenkopf und -hals, Schweineleber und -hirn sowie menschliche Haut und Niere. Insbesondere die erste klinische Studie mit spektraler Bildgebung in der minimalinvasiven Chirurgie hat gezeigt, dass mit deep learning jetzt eine Ischämieüberwachung mit Videofrequenz möglich ist.



## A.1 Region of interest (ROI) tracking

We developed an automatic Region of Interest (ROI) tracking algorithm that is enabled by deep learning. The implementation details of such tracked were presented in Sec. 2.4.2. We then leveraged this tracker to annotate ROIs in the surface of human kidney tissue as part of our clinical trial on ischemia monitoring in laparoscopic surgery. In this section, we present the details of our ROI annotation process (Fig. A.1).

## A.2 Human kidney tissue Signal-to-Noise Ratio (SNR)

In Sec. 3.3.2 we presented our clinical trial on ischemia monitoring based on spectral imaging. To demonstrate our Out-of-Distribution (OoD) approach to ischemia monitoring, we annotated and processed several ROIs on the surface of the kidney from each patient. In this section, we present the SNR within each ROI for all patients who took part in our clinical trial. The SNR level across patients was  $\sim 82$ . For this analysis, the mean SNR was computed across camera channels, ROIs and tissue stages (perfused *vs.* ischemic) before computing the mean across different patients. (Fig. A.2). The SNR was computed by dividing the data from each ROI by a dark measurement recorded with the same camera system (Fig. 2.2).

## A.3 Human skin perfusion

The clinical trial involving healthy humans with the aim of assessing the performance of several methods on the task of oxygenation estimation was presented in Sec.3.3.1. For reporting completion, we present the diffuse reflectance spectra of all four locations recorded in the surface of the right-hand palm from each patient (Fig. A.3).

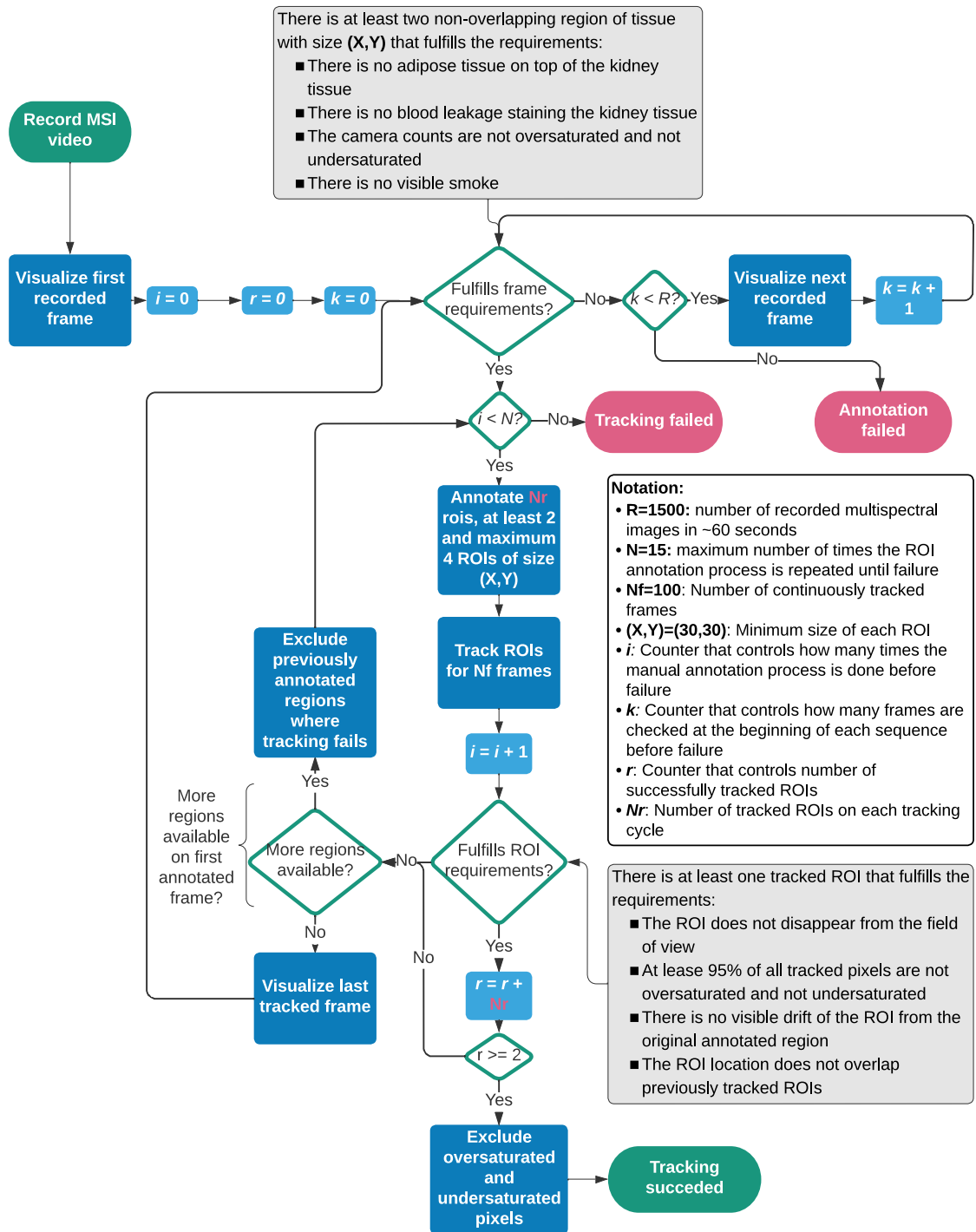
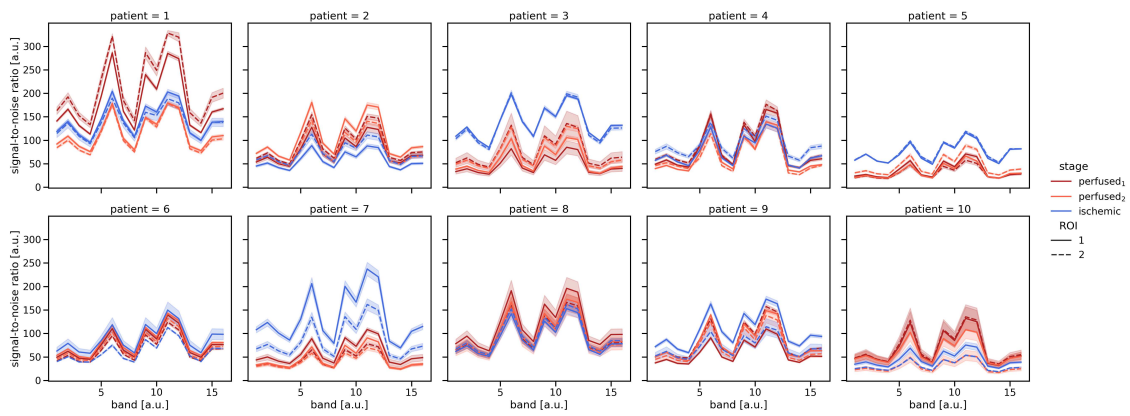
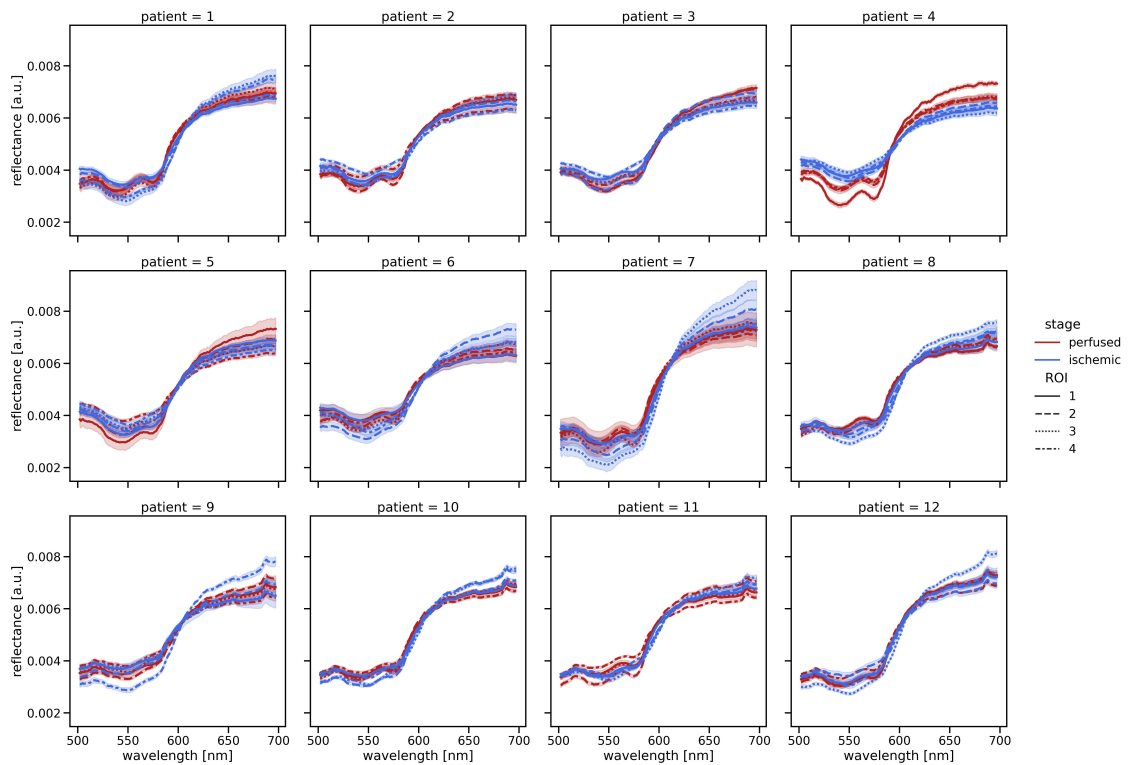


Figure A.1: We annotate ROIs in the surface of tissue and ensure high data quality. Region of Interest (ROI) annotation flowchart based on a sequence of spectral images.



**Figure A.2:** The data within each Region of Interest (ROI) of our clinical trial had a high level of SNR for all patients of our clinical trial on ischemia monitoring. The signal originating from light-tissue interactions is well above noise level, by a factor of  $\sim 82$  on average across patients.



**Figure A.3: Slight spectral differences can be detected with diffuse reflectance measurements performed with a spectrometer.**  $\ell_2$ -normalized reflectance spectra for all 12 patients of our clinical study, corresponding to two exemplary Regions of Interest (ROIs) annotated in perfused and ischemic tissue (160 mmHg). The solid lines correspond to the median reflectance spectra across all 100 measurements, while the translucent bands around each line correspond to the standard deviation.

## Bibliography

- Adler, T. J., Ardizzone, L., Vemuri, A., Ayala, L., Gröhl, J., Kirchner, T., Wirkert, S., Kruse, J., Rother, C., Köthe, U., et al. (2019)a. **Uncertainty-aware performance assessment of optical imaging modalities with invertible neural networks.** *International journal of computer assisted radiology and surgery*, 14(6):997–1007.
- Adler, T., Ayala, L., Ardizzone, L., Kenngott, H., Vemuri, A., Müller-Stich, B., Rother, C., Köthe, U., and Maier-Hein, L. (2019)b. *Out of Distribution Detection for Intra-operative Functional Imaging*, volume 11840 LNCS. Springer. ISBN 9783030326883. doi: 10.1007/978-3-030-32689-0\_8.
- Afifi, M. and Brown, M. (2019). **What Else Can Fool Deep Learning? Addressing Color Constancy Errors on Deep Neural Network Performance.** In *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 243–252. IEEE. ISBN 978-1-7281-4803-8. doi: 10.1109/ICCV.2019.00033. URL <https://ieeexplore.ieee.org/document/9010271/>.
- Afifi, M., Barron, J. T., LeGendre, C., Tsai, Y.-T., and Bleibel, F. (2021). **Cross-Camera Convolutional Color Constancy.** In *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 1961–1970. IEEE. ISBN 978-1-6654-2812-5. doi: 10.1109/ICCV4892.2021.00199. URL <https://ieeexplore.ieee.org/document/9711009/>.
- Alerstam, E., Yip Lo, W. C., Han, T. D., Rose, J., Andersson-Engels, S., and Lilge, L. (2010). **Next-generation acceleration and code optimization for light transport in turbid media using GPUs.** *Biomedical Optics Express*, 1(2):658–675. doi: 10.1364/BOE.1.000658. URL <http://www.opticsinfobase.org/boe/abstract.cfm?URI=boe-1-2-658>.

- Ansari, R., Zhang, E. Z., Desjardins, A. E., and Beard, P. C. (2020). **Miniature all-optical flexible forward-viewing photoacoustic endoscopy probe for surgical guidance.** *Optics Letters*, 45:6238. ISSN 0146-9592. doi: 10.1364/OL.400295.
- Ardizzone, L., Kruse, J., Wirkert, S., Rahner, D., Pellegrini, E. W., Klessen, R. S., Maier-Hein, L., Rother, C., and Köthe, U. (2018). **Analyzing inverse problems with invertible neural networks.** *7th International Conference on Learning Representations, ICLR 2019*. URL <http://arxiv.org/abs/1808.04730>.
- Arganda-Carreras, I., Sorzano, C. O. S., Marabini, R., Carazo, J. M., Ortiz-de Solorzano, C., and Kybic, J. (2006). **Consistent and elastic registration of histological sections using vector-spline regularization.** In Beichel, R. R. and Sonka, M., editors, *Computer Vision Approaches to Medical Image Analysis*, pages 85–95, Berlin, Heidelberg. Springer Berlin Heidelberg.
- Asfour, H., Guan, S., Muselimyan, N., Swift, L., Loew, M., and Sarvazyan, N. (2018). **Optimization of wavelength selection for multispectral image acquisition: a case study of atrial ablation lesions.** *Biomedical Optics Express*, 9(5):2189. ISSN 2156-7085. doi: 10.1364/boe.9.002189. URL <https://www.osapublishing.org/viewmedia.cfm?uri=boe-9-5-2189&seq=0&html=true><https://www.osapublishing.org/abstract.cfm?uri=boe-9-5-2189><https://www.osapublishing.org/boe/abstract.cfm?uri=boe-9-5-2189>.
- Ayala, L. A., Wirkert, S. J., Gröhl, J., Herrera, M. A., Hernandez-Aguilera, A., Vemuri, A., Santos, E., and Maier-Hein, L. (2019). **Live monitoring of haemodynamic changes with multispectral image analysis.** In Zhou, L., Sarikaya, D., Kia, S. M., Speidel, S., Malpani, A., Hashimoto, D., Habes, M., Löfstedt, T., Ritter, K., and Wang, H., editors, *OR 2.0 Context-Aware Operating Theaters and Machine Learning in Clinical Neuroimaging*, pages 38–46. Springer International Publishing. ISBN 978-3-030-32695-1. URL [https://link.springer.com/chapter/10.1007/978-3-030-32695-1\\_5](https://link.springer.com/chapter/10.1007/978-3-030-32695-1_5).
- Ayata, C. and Lauritzen, M. (2015). **Spreading depression, spreading depolarizations, and the cerebral vasculature.** *Physiological Reviews*, 95(3):953–993. ISSN 00319333. doi: 10.1152/physrev.00027.2014.
- Baker, W. B., Parthasarathy, A. B., Busch, D. R., Mesquita, R. C., Greenberg, J. H., and Yodh, A. G. (2014). **Modified beer-lambert law for blood flow.** *Biomedical Optics Express*, 5(11):4053. ISSN 2156-7085. doi: 10.1364/boe.5.004053.
- Ball, T. M., Squeglia, L. M., Tapert, S. F., and Paulus, M. P. (2020). **Double dipping in machine learning: problems and solutions.** *Biological psychiatry. Cognitive neuroscience and neuroimaging*, 5(3):261.



- Bamberg, F., Marcus, R. P., Becker, A., Hildebrandt, K., Bauner, K., Schwarz, F., Greif, M., von Ziegler, F., Bischoff, B., Becker, H.-C., Johnson, T. R., Reiser, M. F., Nikolaou, K., and Theisen, D. (2014). **Dynamic myocardial ct perfusion imaging for evaluation of myocardial ischemia as determined by mr imaging.** *JACC: Cardiovascular Imaging*, 7(3):267–277. doi: 10.1016/j.jcmg.2013.06.008.
- Battiti, R. (1994). **Using mutual information for selecting features in supervised neural net learning.** *IEEE Transactions on neural networks*, 5(4):537–550.
- Bedard, N. and Tkaczyk, T. (2012). **Real-time in vivo visualization of tissue oxygenation and autofluorescence with a snapshot hyperspectral camera for detection of precancerous lesions.** *Biomedical Optics and 3-D Imaging (2012)*, paper BSu3A.76, page BSu3A.76. doi: 10.1364/BIOMED.2012.BSu3A.76. URL <https://www.osapublishing.org/abstract.cfm?uri=BIOMED-2012-BSu3A.76>.
- Bello, G. A., Dawes, T. J. W., Duan, J., Biffi, C., de Marvao, A., Howard, L. S. G. E., Gibbs, J. S. R., Wilkins, M. R., Cook, S. A., Rueckert, D., and O'Regan, D. P. (2019). **Deep-learning cardiac motion analysis for human survival prediction.** *Nature Machine Intelligence*, 1(2):95–104. ISSN 2522-5839. doi: 10.1038/s42256-019-0019-2. URL <https://doi.org/10.1038/s42256-019-0019-2>.
- Bianco, S., Cusano, C., and Schettini, R. (2015). **Color constancy using cnns.** *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, 2015-Octob:81–89. ISSN 21607516. doi: 10.1109/CVPRW.2015.7301275.
- Boltz, S., Debreuve, É., and Barlaud, M. (2009). **High-dimensional statistical measure for region-of-interest tracking.** *IEEE Transactions on Image Processing*, 18(6):1266–1283. ISSN 10577149. doi: 10.1109/TIP.2009.2015158.
- Bouchard, M. B., Chen, B. R., Burgess, S. A., and Hillman, E. M. C. (2009). **Ultra-fast multispectral optical imaging of cortical oxygenation, blood flow, and intracellular calcium dynamics.** *Optics express*, 17(18):15670. doi: 10.1364/OE.17.015670. URL <https://pmc/articles/PMC2760073//pmc/articles/PMC2760073/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC2760073/>.
- Bozovic, G., Schaefer-Prokop, C. M., and Bankier, A. A. (2022). **Pulmonary functional imaging (pfi): A historical review and perspective.** *ACTA RADIOLOGICA*. ISSN 0284-1851. doi: 10.1177/02841851221076324.
- Breiman, L. (2001). **Random forests.** *Machine Learning 2001* 45:1, 45:5–32. ISSN 1573-0565. doi: 10.1023/A:1010933404324. URL <https://link.springer.com/article/10.1023/A:1010933404324>.

- Brown, G., Pocock, A., Zhao, M.-J., and Luján, M. (2012). **Conditional likelihood maximisation: a unifying framework for information theoretic feature selection.** *Journal of machine learning research*, 13(Jan):27–66.
- Chen, M. T. and Durr, N. J. (2020). **Rapid tissue oxygenation mapping from snapshot structured-light images with adversarial deep learning.** *arXiv*, pages 1–18. URL <http://arxiv.org/abs/2007.00760>.
- Choi, H., Jang, E., and Alemi, A. A. (2018). **Waic, but why? generative ensembles for robust anomaly detection.** *arXiv*. URL <https://arxiv.org/abs/1810.01392v4>.
- Choi, W. J., Li, Y., and Wang, R. K. (2019). **Monitoring acute stroke progression: Multiparametric oct imaging of cortical perfusion, flow, and tissue scattering in a mouse model of permanent focal ischemia.** *IEEE Transactions on Medical Imaging*, 38:1427–1437. ISSN 0278-0062. doi: 10.1109/TMI.2019.2895779.
- Chu, W., Chennamsetty, A., Toroussian, R., and Lau, C. (2017). **Anaphylactic shock after intravenous administration of indocyanine green during robotic partial nephrectomy.** *Urology Case Reports*, 12:37. doi: 10.1016/J.EUCR.2017.02.006. URL <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5349462/>.
- Clancy, N. T., Arya, S., Stoyanov, D., Singh, M., Hanna, G. B., and Elson, D. S. (2015). **Intraoperative measurement of bowel oxygen saturation using a multispectral imaging laparoscope.** *Biomedical Optics Express*, 6:4179. ISSN 2156-7085. doi: 10.1364/BOE.6.004179. URL <https://www.osapublishing.org/abstract.cfm?URI=boe-6-10-4179>.
- Clancy, N. T., Jones, G., Maier-Hein, L., Elson, D. S., and Stoyanov, D. (2020). **Surgical spectral imaging.** *Medical Image Analysis*, 63:101699. ISSN 13618415. doi: 10.1016/j.media.2020.101699. URL <https://doi.org/10.1016/j.media.2020.101699>.
- Coello-Torà, I., Segura-Sampedro, J. J., Pérez-Celada, J., Jiménez-Morillas, P., and Morales-Soriano, R. (2020). **[accessory renal artery arising from infrarenal aorta, exposed during lymphadenectomy due to cytoreductive surgery and hipec.].** *Archivos españoles de urología*, 73(1):76–77. ISSN 0004-0614. URL <http://europepmc.org/abstract/MED/31950928>.
- Crosson, B., Ford, A., McGregor, K. M., Meinzer, M., Cheshkov, S., Xiufeng, L., Walker-Batson, D., and Briggs, R. W. (2010). **Functional imaging and related techniques: An introduction for rehabilitation researchers.** *Journal of rehabilitation research and development*, 47:vii. ISSN 07487711. doi: 10.1682/JRRD.2010.02.0017. URL [/pmc/art](http://www.ncbi.nlm.nih.gov/pmc/art)

icles/PMC3225087//pmc/articles/PMC3225087/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC3225087/.

CVL. (2022). **Analyzing inverse problems with invertible neural networks**. Blogpost created by Computer Vision and Learning Lab Heidelberg. Last visited in 06/2022, <https://hci.iwr.uni-heidelberg.de/vislearn/inverse-problems-invertible-neural-networks>.

Davi, F. (2011). **Color constancy**. ISSN 00426989.

De Vis, J. B., Petersen, E. T., Kersbergen, K. J., Alderliesten, T., De Vries, L. S., Van Bel, F., Groenendaal, F., Lemmers, P. M., Hendrikse, J., and Benders, M. J. (2013). **Evaluation of perinatal arterial ischemic stroke using noninvasive arterial spin labeling perfusion mri**. *Pediatric Research* 2013 74:3, 74(3):307–313. ISSN 1530-0447. doi: 10.1038/pr.2013.111. URL <https://www.nature.com/articles/pr2013111>.

Deffieux, T., Demene, C., and Tanter, M. (2021). **Functional ultrasound imaging: A new imaging modality for neuroscience**. *NEUROSCIENCE*, 474(SI):110–121. ISSN 0306-4522. doi: 10.1016/j.neuroscience.2021.03.005.

Dietrich, M., Marx, S., von der Forst, M., Bruckner, T., Schmitt, F. C. F., Fiedler, M. O., Nickel, F., Studier-Fischer, A., Müller-Stich, B. P., Hackert, T., Brenner, T., Weigand, M. A., Uhle, F., and Schmidt, K. (2021)a. **Bedside hyperspectral imaging indicates a microcirculatory sepsis pattern - an observational study**. *Microvascular Research*, 136: 104164. ISSN 0026-2862. doi: <https://doi.org/10.1016/j.mvr.2021.104164>. URL <https://www.sciencedirect.com/science/article/pii/S0026286221000340>.

Dietrich, M., Seidlitz, S., Schreck, N., Wiesenfarth, M., Godau, P., Tizabi, M., Sellner, J., Marx, S., Knödler, S., Allers, M. M., Ayala, L., Schmidt, K., Brenner, T., Studier-Fischer, A., Nickel, F., Müller-Stich, B. P., Kopp-Schneider, A., Weigand, M. A., and Maier-Hein, L. (2021)b. **Machine learning-based analysis of hyperspectral images for automated sepsis diagnosis**. *arXiv*. URL <https://arxiv.org/abs/2106.08445v1>.

Dinh, L., Sohl-Dickstein, J., and Bengio, S. (2016). **Density estimation using real nvp**. *arXiv preprint arXiv:1605.08803*.

Dip, F. D., Ishizawa, T., Kokudo, N., and Rosenthal, R. J., editors. (2015). *Fluorescence Imaging for Surgeons*. Springer International Publishing, Cham. ISBN 978-3-319-15677-4. doi: 10.1007/978-3-319-15678-1. URL <http://link.springer.com/10.1007/978-3-319-15678-1>.

- Dreier, J. P. and Reiffurth, C. (2015). **The stroke-migraine depolarization continuum.** *Neuron*, 86(4):902–922.
- Dreier, J. P., Lemale, C. L., Kola, V., Friedman, A., and Schoknecht, K. (2018). **Spreading depolarization is not an epiphenomenon but the principal mechanism of the cytotoxic edema in various gray matter structures of the brain during stroke.** *Neuropharmacology*, 134:189–207. ISSN 18737064. doi: 10.1016/j.neuropharm.2017.09.027. URL <https://doi.org/10.1016/j.neuropharm.2017.09.027>.
- Du, X., Clancy, N., Arya, S., Hanna, G. B., Kelly, J., Elson, D. S., and Stoyanov, D. (2015). **Robust surface tracking combining features, intensity and illumination compensation.** *International Journal of Computer Assisted Radiology and Surgery*, 10(12):1915–1926. ISSN 18616429. doi: 10.1007/s11548-015-1243-9.
- Du, Z., Jeong, M. K., and Kong, S. G. (2007). **Band selection of hyperspectral images for automatic detection of poultry skin tumors.** *IEEE Transactions on Automation Science and Engineering*, 4(3):332–339. ISSN 1545-5955. doi: 10.1109/TASE.2006.888048. URL <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4266814>.
- Ebner, M., Nabavi, E., Shapey, J., Xie, Y., Liebmann, F., Spirig, J. M., Hoch, A., Farshad, M., Saeed, S. R., Bradford, R., Yardley, I., Ourselin, S., Edwards, A. D., Führnstahl, P., and Vercauteren, T. (2021). **Intraoperative hyperspectral label-free imaging: from system design to first-in-patient translation.** *Journal of Physics D: Applied Physics*, 54(29):294003. ISSN 0022-3727. doi: 10.1088/1361-6463/ABFBF6. URL <https://iopscience.iop.org/article/10.1088/1361-6463/abfbf6><https://iopscience.iop.org/article/10.1088/1361-6463/abfbf6/meta>.
- Eurostat. (2022). **Surgical operations and procedures performed in hospitals.** Last visited in 05/2022, [https://ec.europa.eu/eurostat/databrowser/view/hlth\\_co\\_proc2/default/table?lang=en](https://ec.europa.eu/eurostat/databrowser/view/hlth_co_proc2/default/table?lang=en).
- Felli, E., Taher, M. A., Collins, T., Baiocchini, A., Felli, E., Barberio, M., Ettorre, G. M., Mutter, D., Lindner, V., Hostettler, A., Gioux, S., Schuster, C., and Marescaux, J. (2020). **Hyperspectral evaluation of hepatic oxygenation in a model of total vs . arterial liver ischaemia.** *Scientific Reports*, pages 1–10. ISSN 2045-2322. doi: 10.1038/s41598-020-72915-6. URL <https://doi.org/10.1038/s41598-020-72915-6>.
- Finlayson, G. and Trezzi, E. (2004). **Shades of gray and colour constancy.** pages 37–41. Twelfth Color Imaging Conference: Color Science and Engineering Systems, Technologies, and Applications ; Conference date: 01-11-2004.

- Finlayson, G. D. and Schaefer, G. (2001). **Solving for Colour Constancy using a Constrained Dichromatic Reflection Model**. *International Journal of Computer Vision*, 42(3):127–144. ISSN 1573-1405. doi: 10.1023/A:1011120214885. URL <https://doi.org/10.1023/A:1011120214885>.
- Fleuret, F. (2004). **Fast binary feature selection with conditional mutual information**. *Journal of Machine Learning Research*, 5(Nov):1531–1555.
- Georgoulis, S., Rematas, K., Ritschel, T., Gavves, E., Fritz, M., Gool, L. V., and Tuytelaars, T. (2018). **Reflectance and natural illumination from single-material specular objects using deep learning**. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 40:1932–1947. ISSN 01628828. doi: 10.1109/TPAMI.2017.2742999.
- Goodfellow, I., Bengio, Y., and Courville, A. (2016). *Deep Learning*. MIT Press. <http://www.deeplearningbook.org>.
- Gottrup, F. (1994). **Physiology and measurement of tissue perfusion**. *Annales chirurgiae et gynaecologiae*, 83(3):183–189. ISSN 0355-9521 (Print).
- Gretton, A., Smola, A., Huang, J., Schmittfull, M., Borgwardt, K., and Schölkopf, B. (2009). **Covariate shift by kernel mean matching**. *Dataset shift in machine learning*, 3(4):5. URL [http://www.kyb.mpg.de/publications/attachments/shift-book-for-LeEtAl-webversion\\_5376\[0\].pdf](http://www.kyb.mpg.de/publications/attachments/shift-book-for-LeEtAl-webversion_5376[0].pdf).
- Gröhl, J., Kirchner, T., Adler, T. J., Hacker, L., Holzwarth, N., Hernández-Aguilera, A., Herrera, M. A., Santos, E., Bohndiek, S. E., and Maier-Hein, L. (2021). **Learned spectral decoloring enables photoacoustic oximetry**. *Scientific reports*, 11(1):1–12.
- Gröhl, J., Dreher, K. K., Schellenberg, M., Rix, T., Holzwarth, N., Vieten, P., Ayala, L., Bohndiek, S. E., Seitel, A., and Maier-Hein, L. (2022). **Simpa: an open-source toolkit for simulation and image processing for photonics and acoustics**. *Journal of Biomedical Optics*, 27(8):1 – 21. doi: 10.1117/1.JBO.27.8.083010. URL <https://doi.org/10.1117/1.JBO.27.8.083010>.
- Gu, L., Huynh, C. P., and Robles-Kelly, A. (2014). **Segmentation and estimation of spatially varying illumination**. *IEEE Transactions on Image Processing*, 23:3478–3489. ISSN 10577149. doi: 10.1109/TIP.2014.2330768.
- Gu, X., Han, Z., Yao, L., Zhong, Y., Shi, Q., Fu, Y., Liu, C., Wang, X., and Xie, T. (2016). **Image enhancement based on *in vivo* hyperspectral gastroscopic images: a case study**. *Journal of Biomedical Optics*, 21(10):101412. ISSN 1083-3668. doi:

## Bibliography

---

10.1117/1.JBO.21.10.101412. URL <http://biomedicaloptics.spiedigitallibrary.org/article.aspx?doi=10.1117/1.JBO.21.10.101412>.

Guennadi, S. and Andrei, B. (2020). **Optimization of Band Selection in Multispectral and Narrow-Band Imaging: An Analytical Approach**. *Advances in experimental medicine and biology*, 1232:361–367. ISSN 0065-2598. doi: 10.1007/978-3-030-34461-0\_46. URL <https://pubmed.ncbi.nlm.nih.gov/31893432/>.

Guo, H., Bai, W., Ouyang, W., Liu, Y., Wu, C., Xu, Y., Weng, Y., Zang, H., Liu, Y., Jacobson, L., Hu, Z., Wang, Y., Arafa, H. M., Yang, Q., Lu, D., Li, S., Zhang, L., Xiao, X., Vázquez-Guardado, A., Ciatti, J., Dempsey, E., Ghoreishi-Haack, N., Waters, E. A., Haney, C. R., Westman, A. M., MacEwan, M. R., Pet, M. A., and Rogers, J. A. (2022). **Wireless implantable optical probe for continuous monitoring of oxygen saturation in flaps and organ grafts**. *Nature Communications* 2022 13:1, 13(1):1–12. ISSN 2041-1723. doi: 10.1038/s41467-022-30594-z. URL <https://www.nature.com/articles/s41467-022-30594-z>.

Guyon, I. and Elisseeff, A. (2003). **An introduction to variable and feature selection**. *The Journal of Machine Learning Research*, 3:1157–1182. URL <http://dl.acm.org/citation.cfm?id=944968>.

Hajj, J., Blaine, N., Salavaci, J., and Jacoby, D. (2018). **The “centrality of sepsis”: A review on incidence, mortality, and cost of care**. *Healthcare*, 6(3). ISSN 2227-9032. doi: 10.3390/healthcare6030090. URL <https://www.mdpi.com/2227-9032/6/3/90>.

Han, Z., Zhang, A., Wang, X., Sun, Z., Wang, M. D., and Xie, T. (2016). **In vivo use of hyperspectral imaging to develop a noncontact endoscopic diagnosis support system for malignant colorectal tumors**. *Journal of Biomedical Optics*, 21(1):016001–016001. URL <http://biomedicaloptics.spiedigitallibrary.org/article.aspx?articleid=2481122>.

Hernandez-Juarez, D., Parisot, S., Busam, B., Leonardis, A., Slabaugh, G., and McDonagh, S. (2020). **A Multi-Hypothesis Approach to Color Constancy**. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2267–2277. IEEE. ISBN 978-1-7281-7168-5. doi: 10.1109/CVPR42600.2020.00234. URL <https://ieeexplore.ieee.org/document/9156797/>.

Holmer, A., Marotz, J., Wahl, P., Dau, M., and Kämmerer, P. W. (2018). **Hyperspectral imaging in perfusion and wound diagnostics - methods and algorithms for the determination of tissue parameters**. *Biomedizinische Technik*. ISSN 00135585. doi: 10.1515/bmt-2017-0155.

- Hotelling, H. (1933). **Analysis of a complex of statistical variables into principal components.** *Journal of Educational Psychology*, 24:417–441. ISSN 00220663. doi: 10.1037/H0071325. URL /record/1934-00645-001.
- Hu, Y., Wang, B., and Lin, S. (2017). **Fc 4 : Fully convolutional color constancy with confidence-weighted pooling.** *Cvpr '17*, 1:330–339. doi: 10.1109/CVPR.2017.43.
- Hu, Z., Fang, C., Li, B., Zhang, Z., Cao, C., Cai, M., Su, S., Sun, X., Shi, X., Li, C., Zhou, T., Zhang, Y., Chi, C., He, P., Xia, X., Chen, Y., Gambhir, S. S., Cheng, Z., and Tian, J. (2020). **First-in-human liver-tumour surgery guided by multispectral fluorescence imaging in the visible and near-infrared-i/ii windows.** *Nature Biomedical Engineering*, 4(3):259–271. ISSN 2157-846X. doi: 10.1038/s41551-019-0494-0. URL <https://doi.org/10.1038/s41551-019-0494-0>.
- Huang, J., Smola, A. J., Gretton, A., Borgwardt, K. M., Schölkopf, B., and others. (2007). **Correcting sample selection bias by unlabeled data.** *Advances in neural information processing systems*, 19:601. URL <https://papers.nips.cc/paper/3075-correcting-sample-selection-bias-by-unlabeled-data.pdf>.
- Imai, Y., Kato, Y., Kadoi, H., Horiuchi, T., and Tominaga, S. (2011). **Estimation of multiple illuminants based on specular highlight detection.** *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 6626 LNCS:85–98. ISSN 03029743. doi: 10.1007/978-3-642-20404-3\_7.
- Jacques, S. L. (2013). **Optical properties of biological tissues: a review.** *Physics in Medicine and Biology*, 58(11):R37–R61. ISSN 0031-9155, 1361-6560. doi: 10.1088/0031-9155/58/11/R37. URL <http://stacks.iop.org/0031-9155/58/i=11/a=R37?key=crossref.e58b67eb50f9f6507dd5e75b7744fa07>.
- Jakulin, A. (2005). *Machine learning based on attribute interactions*. PhD thesis, Univerza v Ljubljani.
- Jones, G., Clancy, N. T., Helo, Y., Arridge, S., Elson, D. S., and Stoyanov, D. (2017). **Bayesian estimation of intrinsic tissue oxygenation and perfusion from rgb images.** *IEEE Transactions on Medical Imaging*, 36:1491–1501. ISSN 1558254X. doi: 10.1109/TMI.2017.2665627.
- Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Židek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., Back, T., Petersen, S., Reiman, D., Clancy, E., Zielinski, M., Steinegger, M., Pacholska, M.,

- Berghammer, T., Bodenstein, S., Silver, D., Vinyals, O., Senior, A. W., Kavukcuoglu, K., Kohli, P., and Hassabis, D. (2021). **Highly accurate protein structure prediction with alphafold.** *Nature*, 596(7873):583–589. ISSN 1476-4687. doi: 10.1038/s41586-021-03819-2. URL <https://doi.org/10.1038/s41586-021-03819-2>.
- Kaneko, E., Aoki, H., and Tsukada, M. (2016). **Daylight Spectrum Estimation from Hyper- and Multispectral Image without Area Extraction of Uniform Materials.** *Proceedings - 11th International Conference on Signal-Image Technology and Internet-Based Systems, SITIS 2015*, pages 53–60. doi: 10.1109/SITIS.2015.67.
- Kaplan-Marans, E., Fulla, J., Tomer, N., Bilal, K., and Palese, M. (2019). **Indocyanine green (icg) in urologic surgery.** *Urology*, 132:10–17. ISSN 0090-4295. doi: <https://doi.org/10.1016/j.urology.2019.05.008>. URL <https://www.sciencedirect.com/science/article/pii/S0090429519304467>.
- Karen, S. and Andrew, Z. (2015). **Very deep convolutional networks for large-scale image recognition.** URL <https://arxiv.org/abs/1409.1556>.
- Kato, Y., Horiuchi, T., and Tominaga, S. (2012). **Estimation of multiple light sources from specular highlights.** *International Conference on Pattern Recognition (ICPR)*, pages 2083–2086. ISSN 10514651.
- Kester, R. T., Bedard, N., Gao, L. S., and Tkaczyk, T. S. (2011). **Real-time snapshot hyperspectral imaging endoscope.** *Journal of biomedical optics*, 16(5):056005.
- Khan, H. A., Thomas, J.-B., Hardeberg, J. Y., and Laligant, O. (2017). **Illuminant estimation in multispectral imaging.** *Journal of the Optical Society of America A*, 34:1085. ISSN 1084-7529. doi: 10.1364/JOSAA.34.001085. URL <https://www.osapublishing.org/abstract.cfm?URI=josaa-34-7-1085>.
- Kingma, D. P. and Ba, J. (2014)a. **Adam: A method for stochastic optimization.** URL <https://arxiv.org/abs/1412.6980>.
- Kingma, D. P. and Ba, J. (2014)b. **Adam: A method for stochastic optimization.** *arXiv preprint arXiv:1412.6980*.
- Kleshnin, M. S. (2019). **Deep learning neural network estimation of tissue oxygenation based on diffuse optical spectroscopy.** *Journal of Biomedical Optics*.
- Kohavi, R. and John, G. H. (1997). **Wrappers for feature subset selection.** *Artificial intelligence*, 97(1-2):273–324.



- Köhler, H., Kulcke, A., Maktabi, M., Moulla, Y., Jansen-Winkel, B., Barberio, M., Diana, M., Gockel, I., Neumuth, T., and Chalopin, C. (2020). **Laparoscopic system for simultaneous high-resolution video and rapid hyperspectral imaging in the visible and near-infrared spectral range.** *Journal of Biomedical Optics*, 25(08):086004. ISSN 1083-3668. doi: 10.1117/1.JBO.25.8.086004. URL <https://www.spiedigitallibrary.org/journals/journal-of-biomedical-optics/volume-25/issue-08/086004/Laparoscopic-system-for-simultaneous-high-resolution-video-and-rapid-hyperspectral/10.1117/1.JBO.25.8.086004.full>.
- Koonen, T. (2006). **Fabry-perot interferometer filters.** *Springer Series in Optical Sciences*, 123:271–287. ISSN 03424111. doi: 10.1007/3-540-31770-8\_7. URL [https://link.springer.com/chapter/10.1007/3-540-31770-8\\_7](https://link.springer.com/chapter/10.1007/3-540-31770-8_7).
- Koscevic, K., Subasic, M., and Loncaric, S. (2021). **Iterative convolutional neural network-based illumination estimation.** *IEEE Access*, 9:26755–26765. ISSN 2169-3536. doi: 10.1109/ACCESS.2021.3057072. URL <https://ieeexplore.ieee.org/document/9347438/>.
- Kraskov, A., Stögbauer, H., and Grassberger, P. (2004). **Estimating mutual information.** *Physical review E*, 69(6):066138.
- Kumar, M., Suliburk, J. W., Veeraraghavan, A., and Sabharwal, A. (2020). **Pulsecam: a camera-based, motion-robust and highly sensitive blood perfusion imaging modality.** *Scientific Reports*, 10:1–17. ISSN 20452322. doi: 10.1038/s41598-020-61576-0.
- Köhler, H., Jansen-Winkel, B., Chalopin, C., and Gockel, I. (2019). **Hyperspectral imaging as a new optical method for the measurement of gastric conduit perfusion.** *Diseases of the Esophagus*, 32:6675. ISSN 14422050. doi: 10.1093/dote/doz046.
- Laakom, F., Raitoharju, J., Iosifidis, A., Nikkanen, J., and Gabbouj, M. (2019). **Color Constancy Convolutional Autoencoder.** *arXiv*, pages 1085–1090. doi: 10.1109/SSCI44817.2019.9002684.
- Laakom, F., Passalis, N., Raitoharju, J., Nikkanen, J., Tefas, A., Iosifidis, A., and Gabbouj, M. (2020). **Bag of Color Features for Color Constancy.** *IEEE Transactions on Image Processing*, 29:7722–7734. ISSN 19410042. doi: 10.1109/TIP.2020.3004921.
- Laakom, F., Raitoharju, J., Nikkanen, J., Iosifidis, A., and Gabbouj, M. (2021). **Robust channel-wise illumination estimation.** URL <https://arxiv.org/abs/2111.05681v1>.
- Land, E. (1977). **The retinex theory of color vision.** *Scientific American*.

- LeCun, Y. (1989). **Generalization and network design strategies**. University of Toronto.
- Legrand, M., Mik, E. G., Johannes, T., Payen, D., and Ince, C. (2008). **Renal hypoxia and dysoxia after reperfusion of the ischemic kidney**. *Molecular Medicine*, 14(7):502–516. ISSN 1528-3658. doi: 10.2119/2008-00006.Legrand. URL <https://doi.org/10.2119/2008-00006.Legrand>.
- Li, L., Jamieson, K., Rostamizadeh, A., Gonina, K., Hardt, M., Recht, B., and Talwalkar, A. (2018). **Massively parallel hyperparameter tuning**. URL <https://openreview.net/forum?id=S1Y700lRZ>.
- Lin, D. and Tang, X. (2006). **Conditional infomax learning: an integrated framework for feature extraction and fusion**. *Computer Vision–ECCV 2006*, pages 68–82.
- Listmann, M. and Tubbs, R. S. (2020). **Chapter 19 - the abdominal aorta**. In *Surgical Anatomy of the Lateral Transpsoas Approach to the Lumbar Spine*, pages 185–188. Elsevier. ISBN 978-0-323-67376-1. doi: <https://doi.org/10.1016/B978-0-323-67376-1.00019-7>. URL <https://www.sciencedirect.com/science/article/pii/B9780323673761000197>.
- Liu, B., Sun, M., Liu, Q., Kassam, A., Li, C. C., and Scabassi, R. J. (2005). **Automatic detection of region of interest based on object tracking in neurosurgical video**. *Annual International Conference of the IEEE Engineering in Medicine and Biology - Proceedings*, 7 VOLS:6273–6276. ISSN 05891019. doi: 10.1109/IEMBS.2005.1615931.
- Liu, J., Liu, J., Wang, S., Zhao, H., Tian, C., Shi, B., and Jiang, X. (2020)a. **Three-dimensional nephrometry scoring system: a precise scoring system to evaluate complexity of renal tumors suitable for partial nephrectomy**. *PeerJ*, 8:e8637. ISSN 2167-8359 (Print). doi: 10.7717/peerj.8637.
- Liu, L., Wang, Y., Liu, H. X., and Gao, J. (2019). **Peripapillary region perfusion and retinal nerve fiber layer thickness abnormalities in diabetic retinopathy assessed by oct angiography**. *Translational Vision Science and Technology*, 8:14. ISSN 2164-2591. doi: 10.1167/tvst.8.4.14.
- Liu, Z., Lian, T., Farrell, J., and Wandell, B. A. (2020)b. **Neural network generalization: The impact of camera parameters**. *IEEE Access*, 8:10443–10454. ISSN 21693536. doi: 10.1109/ACCESS.2020.2965089.
- Lo, Y.-C., Chang, C.-C., Chiu, H.-C., Huang, Y.-H., Chen, C.-P., Chang, Y.-L., and Jou, K. (2021). **CLCC: Contrastive Learning for Color Constancy**. In *2021 IEEE/CVF Conference*

- on *Computer Vision and Pattern Recognition (CVPR)*, pages 8049–8059. IEEE. ISBN 978-1-6654-4509-2. doi: 10.1109/CVPR46437.2021.00796. URL <https://ieeexplore.ieee.org/document/9577539/>.
- Lu, G., Wang, D., Qin, X., Halig, L., Muller, S., Zhang, H., Chen, A., Pogue, B. W., Chen, Z. G., and Fei, B. (2015). **Framework for hyperspectral image processing and quantification for cancer detection during animal tumor surgery.** *Journal of Biomedical Optics*, 20(12):126012–126012. URL <http://biomedicaloptics.spiedigitallibrary.org/article.aspx?articleid=2479908>.
- Lukezic, A., Vojir, T., Cehovin Zajc, L., Matas, J., and Kristan, M. (2018). **Discriminative correlation filter tracker with channel and spatial reliability.** *International Journal of Computer Vision*, 126(7):671–688. ISSN 15731405. doi: 10.1007/s11263-017-1061-3. URL <https://doi.org/10.1007/s11263-017-1061-3>.
- Mangraviti, A., Volpin, F., Cha, J., Cunningham, S. I., Raje, K., Brooke, M. J., Brem, H., Olivi, A., Huang, J., Tyler, B. M., and Rege, A. (2020). **Intraoperative laser speckle contrast imaging for real-time visualization of cerebral blood flow in cerebrovascular surgery: Results from pre-clinical studies.** *Scientific Reports*, 10:7614. ISSN 2045-2322. doi: 10.1038/s41598-020-64492-5.
- Manski, D. (2020). *Urologielehrbuch.de*. Dr. med. Dirk Manski.
- Marois, M., Jacques, S. L., and Paulsen, K. D. (2018). **Optimal wavelength selection for optical spectroscopy of hemoglobin and water within a simulated light-scattering tissue.** *Journal of biomedical optics*, 23(4):041202.
- Miao, Y. Q., Farahat, A. K., and Kamel, M. S. (2016). **Ensemble kernel mean matching.** *Proceedings - IEEE International Conference on Data Mining, ICDM, 2016-Janua(February 2016):330–338*. ISSN 15504786. doi: 10.1109/ICDM.2015.127.
- Mori, M., Chiba, T., Nakamizo, A., Kumashiro, R., Murata, M., Akahoshi, T., Tomikawa, M., Kikkawa, Y., Yoshimoto, K., Mizoguchi, M., Sasaki, T., and Hashizume, M. (2014). **Intraoperative visualization of cerebral oxygenation using hyperspectral image data: a two-dimensional mapping method.** *International Journal of Computer Assisted Radiology and Surgery*, 9(6):1059–1072. ISSN 18616429. doi: 10.1007/s11548-014-0989-9.
- Nouri, D., Lucas, Y., and Treuillet, S. (2014). **Efficient tissue discrimination during surgical interventions using hyperspectral imaging.** In *International Conference on Information Processing in Computer-Assisted Interventions*, pages 266–275. Springer. URL [http://link.springer.com/chapter/10.1007/978-3-319-07521-1\\_28](http://link.springer.com/chapter/10.1007/978-3-319-07521-1_28).

- Nouri, D., Lucas, Y., and Treuillet, S. (2016). **Hyperspectral interventional imaging for enhanced tissue visualization and discrimination combining band selection methods.** *International Journal of Computer Assisted Radiology and Surgery*, 11(12): 2185–2197. ISSN 1861-6410, 1861-6429. doi: 10.1007/s11548-016-1449-5. URL <http://link.springer.com/10.1007/s11548-016-1449-5>.
- Nwoye, C. I., Mutter, D., Marescaux, J., and Padoy, N. (2019). **Weakly supervised convolutional lstm approach for tool tracking in laparoscopic videos.** *International Journal of Computer Assisted Radiology and Surgery*, 14(6):1059–1067. ISSN 18616429. doi: 10.1007/S11548-019-01958-6/FIGURES/3. URL <https://link.springer.com/article/10.1007/s11548-019-01958-6>.
- Pan, S. J. and Yang, Q. (2010). **A Survey on Transfer Learning.** *IEEE Transactions on Knowledge and Data Engineering*, 22(10):1345–1359. ISSN 1041-4347. doi: 10.1109/TKDE.2009.191. URL <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5288526>.
- Passalis, N. and Tefas, A. (2017). **Learning Bag-of-Features Pooling for Deep Convolutional Neural Networks.** In *2017 IEEE International Conference on Computer Vision (ICCV)*, pages 5766–5774. IEEE. ISBN 978-1-5386-1032-9. doi: 10.1109/ICCV.2017.614. URL <http://ieeexplore.ieee.org/document/8237876/>.
- Peng, H., Long, F., and Ding, C. (2005). **Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy.** *IEEE Transactions on pattern analysis and machine intelligence*, 27(8):1226–1238.
- Poirson, J., Bretenaker, F., Vallet, M., and Floch, A. L. (1997). **Analytical and experimental study of ringing effects in a fabry-perot cavity. application to the measurement of high finesses.** *J. Opt. Soc. Am. B*, 14(11):2811–2817. doi: 10.1364/JOSAB.14.002811. URL <http://josab.osa.org/abstract.cfm?URI=josab-14-11-2811>.
- Porpiglia, F., Fiori, C., Checcucci, E., Pecoraro, A., Dio, M. D., and Bertolo, R. (2018). **Selective clamping during laparoscopic partial nephrectomy: the use of near infrared fluorescence guidance.** *Minerva Urology and Nephrology*, 70. ISSN 27246051. doi: 10.23736/S0393-2249.17.03046-6.
- Prahl, S. (1999). **Optical absorption of hemoglobin.** <https://omlc.org/spectra/hemoglobin/summary.html>. URL <https://omlc.org/spectra/hemoglobin/summary.html>. Accessed: 2021-12-23.

- Preece, S. J. and Claridge, E. (2004). **Spectral filter optimization for the recovery of parameters which describe human skin.** *IEEE Trans Pattern Anal Mach Intell*, 26(7): 913–922.
- Rahimi, A., Recht, B., and others. (2007). **Random Features for Large-Scale Kernel Machines.** In *NIPS*, volume 3, page 5. URL <https://papers.nips.cc/paper/3182-random-features-for-large-scale-kernel-machines.pdf>.
- Rebolleda, G., Diez-Alvarez, L., Casado, A., Sánchez-Sánchez, C., de Dompablo, E., González-López, J. J., and Muñoz-Negrete, F. J. (2015). **Oct: New perspectives in neuro-ophthalmology.** *Saudi Journal of Ophthalmology*, 29(1):9–25. ISSN 1319-4534. doi: <https://doi.org/10.1016/j.sjopt.2014.09.016>. URL <https://www.sciencedirect.com/science/article/pii/S1319453414001143>.
- Roberts, M., Driggs, D., Thorpe, M., Gilbey, J., Yeung, M., Ursprung, S., Aviles-Rivero, A. I., Etmann, C., McCague, C., Beer, L., Weir-McCall, J. R., Teng, Z., Gkrania-Klotsas, E., Ruggiero, A., Korhonen, A., Jefferson, E., Ako, E., Langs, G., Gozaliasl, G., Yang, G., Prosch, H., Preller, J., Stanczuk, J., Tang, J., Hofmanninger, J., Babar, J., Sánchez, L. E., Thillai, M., Gonzalez, P. M., Teare, P., Zhu, X., Patel, M., Cafolla, C., Azadbakht, H., Jacob, J., Lowe, J., Zhang, K., Bradley, K., Wassin, M., Holzer, M., Ji, K., Ortet, M. D., Ai, T., Walton, N., Lio, P., Stranks, S., Shadbahr, T., Lin, W., Zha, Y., Niu, Z., Rudd, J. H., Sala, E., and Schönlieb, C. B. (2021). **Common pitfalls and recommendations for using machine learning to detect and prognosticate for covid-19 using chest radiographs and ct scans.** *Nature Machine Intelligence* 2021 3:3, 3:199–217. ISSN 2522-5839. doi: [10.1038/s42256-021-00307-0](https://doi.org/10.1038/s42256-021-00307-0). URL <https://www.nature.com/articles/s42256-021-00307-0>.
- Robles-Kelly, A. and Huynh, C. P. (2013). *Spectral Image Formation Process*, pages 17–35. Springer London, London. ISBN 978-1-4471-4652-0. doi: [10.1007/978-1-4471-4652-0\\_3](https://doi.org/10.1007/978-1-4471-4652-0_3). URL [https://doi.org/10.1007/978-1-4471-4652-0\\_3](https://doi.org/10.1007/978-1-4471-4652-0_3).
- Rueden, C. T., Schindelin, J., Hiner, M. C., DeZonia, B. E., Walter, A. E., Arena, E. T., and Eliceiri, K. W. (2017). **Imagej2: Imagej for the next generation of scientific image data.** *BMC Bioinformatics*.
- Santos, E., Schöll, M., et. al., and Sakowitz, O. W. (2013). **Cortical spreading depression dynamics can be studied using intrinsic optical signal imaging in gyrencephalic animal cortex.** *Acta Neurochirurgica, Supplement*.
- Santos, E., Schöll, M., Sánchez-Porrás, R., Dahlem, M. A., Silos, H., Unterberg, A., Dickhaus, H., and Sakowitz, O. W. (2014). **Radial, spiral and reverberating waves of spreading depolarization occur in the gyrencephalic brain.** *NeuroImage*, 99:244–255.

- Schöll, M. J., Santos, E., et. al., and Dickhaus, H. (2017). **Large field-of-view movement-compensated intrinsic optical signal imaging for the characterization of the haemodynamic response to spreading depolarizations in large gyrencephalic brains.** *Journal of Cerebral Blood Flow and Metabolism*.
- Schreck, N. (2019). **Empirical decomposition of the explained variation in the variance components form of the mixed model.** *bioRxiv*.
- Seidlitz, S., Sellner, J., Odenthal, J., Özdemir, B., Studier-Fischer, A., Knödler, S., Ayala, L., Adler, T., Kenngott, H. G., Tizabi, M., et al. (2021). **Robust deep learning-based semantic organ segmentation in hyperspectral images.** *arXiv preprint arXiv:2111.05408*.
- Shad, R., Cunningham, J. P., Ashley, E. A., Langlotz, C. P., and Hiesinger, W. (2021). **Designing clinically translatable artificial intelligence systems for high-dimensional medical imaging.** *Nature Machine Intelligence* 2021 3:11, 3:929–935. ISSN 2522-5839. doi: 10.1038/s42256-021-00399-8. URL <https://www.nature.com/articles/s42256-021-00399-8>.
- Sheffield. (1985). **Selecting band combinations from multispectral data.** *Photogrammetric Engineering and Remote Sensing*, 51(6):681–687.
- Soares, A. S., Stoyanov, D., Lovat, L. B., Chand, M., Clancy, N. T., Bano, S., Lovat, L. B., Lovat, L. B., Chand, M., Chand, M., Stoyanov, D., and Stoyanov, D. (2021). **Intraoperative colon perfusion assessment using multispectral imaging.** *Biomedical Optics Express*, Vol. 12, Issue 12, pp. 7556-7567, 12:7556–7567. ISSN 2156-7085. doi: 10.1364/BOE.435118. URL <https://www.osapublishing.org/viewmedia.cfm?uri=boe-12-12-7556&seq=0&html=true><https://www.osapublishing.org/abstract.cfm?uri=boe-12-12-7556><https://www.osapublishing.org/boe/abstract.cfm?uri=boe-12-12-7556>.
- Speidel, S., Kroehnert, A., Bodenstedt, S., Kenngott, H., Müller-Stich, B., and Dillmann, R. (2015). **Image-based tracking of the suturing needle during laparoscopic interventions.** <https://doi.org/10.1117/12.2081920>, 9415:70–75. ISSN 16057422. doi: 10.1117/12.2081920. URL <https://www.spiedigitallibrary.org/conference-proceedings-of-spie/9415/94150B/Image-based-tracking-of-the-suturing-needle-during-laparoscopic-interventions/10.1117/12.2081920.full><https://www.spiedigitallibrary.org/conference-proceedings-of-spie/9415/94150B/Image-based-tracking-of-the-suturing-needle-during-laparoscopic-interventions/10.1117/12.2081920.short>.
- Stoyanov, D., Elson, D. S., Hanna, G. B., Singh, M., Clancy, N. T., and Arya, S. (2015). **Intraoperative measurement of bowel oxygen saturation using a multispectral imaging**

- laparoscope.** *Biomedical Optics Express*, Vol. 6, Issue 10, pp. 4179-4190, 6(10):4179–4190. ISSN 2156-7085. doi: 10.1364/BOE.6.004179. URL <https://www.osapublishing.org/viewmedia.cfm?uri=boe-6-10-4179&seq=0&html=true><https://www.osapublishing.org/abstract.cfm?uri=boe-6-10-4179><https://www.osapublishing.org/boe/abstract.cfm?uri=boe-6-10-4179>.
- Studier-Fischer, A., Seidlitz, S., Sellner, J., Wiesenfarth, M., Ayala, L., Oezdemir, B., Odenthal, J., Knoedler, S., Kowalewski, K.-F., Haney, C. M., et al. (2021). **Spectral organ fingerprints for intraoperative tissue classification with hyperspectral imaging.** *bioRxiv*.
- Styles, I. B., Calcagni, A., Claridge, E., Orihuela-Espina, F., and Gibson, J. M. (2006). **Quantitative analysis of multi-spectral fundus images.** *Med Image Anal*, 10(4):578–597.
- Suh, M., Bahar, S., Mehta, A. D., and Schwartz, T. H. (2006). **Blood volume and hemoglobin oxygenation response following electrical stimulation of human cortex.** *NeuroImage*, 31(1):66–75. ISSN 10538119. doi: 10.1016/j.neuroimage.2005.11.030.
- Sutin, J., Zimmerman, B., Tyulmankov, D., Tamborini, D., Wu, K. C., Selb, J., Gulinatti, A., Rech, I., Tosi, A., Boas, D. A., and Franceschini, M. A. (2016). **Time-domain diffuse correlation spectroscopy.** *Optica*, 3:1006. ISSN 2334-2536. doi: 10.1364/OPTICA.3.001006. URL <https://www.osapublishing.org/abstract.cfm?URI=optica-3-9-1006>.
- Takamatsu, T., Kitagawa, Y., Akimoto, K., Iwanami, R., Endo, Y., Takashima, K., Okubo, K., Umezawa, M., Kuwata, T., Sato, D., Kadota, T., Mitsui, T., Ikematsu, H., Yokota, H., Soga, K., and Takemura, H. (2021). **Over 1000 nm near-infrared multispectral imaging system for laparoscopic in vivo imaging.** *Sensors*, 21(8). ISSN 1424-8220. doi: 10.3390/s21082649. URL <https://www.mdpi.com/1424-8220/21/8/2649>.
- Takano, T., Tian, G. F., Peng, W., Lou, N., Lovatt, D., Hansen, A. J., Kasischke, K. A., and Nedergaard, M. (2007). **Cortical spreading depression causes and coincides with tissue hypoxia.** *Nature Neuroscience*, 10(6).
- Trajanovski, S., Shan, C., Weijtmans, P. J., Koning, S. G. D., and Ruers, T. J. (2021). **Tongue tumor detection in hyperspectral images using deep learning semantic segmentation.** *IEEE Transactions on Biomedical Engineering*, 68:1330–1340. ISSN 15582531. doi: 10.1109/TBME.2020.3026683.
- Tzoumas, S., Nunes, A., Olefir, I., Stangl, S., Symvoulidis, P., Glasl, S., Bayer, C., Multhoff, G., and Ntziachristos, V. (2016). **Eigenspectra optoacoustic tomography achieves**

- quantitative blood oxygenation imaging deep in tissues.** *Nature Communications* 2016 7:1, 7(1):1–10. ISSN 2041-1723. doi: 10.1038/ncomms12121. URL <https://www.nature.com/articles/ncomms12121>.
- Vasefi, F., MacKinnon, N., Saager, R. B., Durkin, A. J., Chave, R., Lindsley, E. H., and Farkas, D. L. (2015). **Polarization-sensitive hyperspectral imaging in vivo: A multimode dermoscope for skin analysis.** *Scientific Reports*, 4:4924. ISSN 2045-2322. doi: 10.1038/srep04924.
- Wang, L. and Jacques, S. (1992)a. **Monte carlo modeling of light transport in multi-layered tissues in standard c.** <https://omlc.org/software/mc/mcml/MCman.pdf>.
- Wang, L. and Jacques, S. L. (1992)b. **Monte Carlo modeling of light transport in multi-layered tissues in standard C.** *The University of Texas, MD Anderson Cancer Center, Houston.* URL [http://www.researchgate.net/publication/15612297\\_MCM\\_L\T1\textendashMonte\\_Carlo\\_modeling\\_of\\_light\\_transport\\_in\\_multi-layered\\_tissues/file/32bfe50f033cc525b8.pdf](http://www.researchgate.net/publication/15612297_MCM_L\T1\textendashMonte_Carlo_modeling_of_light_transport_in_multi-layered_tissues/file/32bfe50f033cc525b8.pdf).
- Watanabe, S. (2013). **A widely applicable bayesian information criterion.** *Journal of Machine Learning Research*, 14(Mar):867–897. ISSN ISSN 1533-7928.
- WHO. (2022). **Medical doctors (per 10 000 population).** Last visited in 05/2022, [https://www.who.int/data/gho/data/indicators/indicator-details/GHO/medical-doctors-\(per-10-000-population\)](https://www.who.int/data/gho/data/indicators/indicator-details/GHO/medical-doctors-(per-10-000-population)).
- Wiesenfarth, M., Reinke, A., Landman, B. A., Eisenmann, M., Saiz, L. A., Cardoso, M. J., Maier-Hein, L., and Kopp-Schneider, A. (2021). **Methods and open-source toolkit for analyzing and visualizing challenge results.** *Scientific Reports*, 11(1):2369. ISSN 20452322. doi: 10.1038/s41598-021-82017-6. URL <https://doi.org/10.1038/s41598-021-82017-6>.
- Wirkert, S. J., Clancy, N. T., et. al., and Maier-Hein, L. (2014)a. **Endoscopic sheffield index for unsupervised in vivo spectral band selection.** In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, volume 8899. Springer.
- Wirkert, S. J., Clancy, N. T., Stoyanov, D., Arya, S., Hanna, G. B., Schlemmer, H.-P., Sauer, P., Elson, D. S., and Maier-Hein, L. (2014)b. **Endoscopic Sheffield Index for Unsupervised In Vivo Spectral Band Selection.** In Luo, X., Reichl, T., Mirota, D., and Soper, T., editors, *Computer-Assisted and Robotic Endoscopy*, volume 8899, pages 110–120.



- Springer International Publishing, Cham. ISBN 978-3-319-13409-3 978-3-319-13410-9. URL <http://www.springerprofessional.de/011\T1\textemdashendoscopic-sheffield-index-for-unsupervised-in-vivo-spectral-band-selection/5457688.html>.
- Wirkert, S. J., Kenngott, H., Mayer, B., Mietkowski, P., Wagner, M., Sauer, P., Clancy, N. T., Elson, D. S., and Maier-Hein, L. (2016)a. **Robust near real-time estimation of physiological parameters from megapixel multispectral images with inverse monte carlo and random forest regression.** *International Journal of Computer Assisted Radiology and Surgery*, 11:909–917. ISSN 18616429. doi: 10.1007/s11548-016-1376-5.
- Wirkert, S. J., Kenngott, H., Mayer, B., Mietkowski, P., Wagner, M., Sauer, P., Clancy, N. T., Elson, D. S., and Maier-Hein, L. (2016)b. **Robust near real-time estimation of physiological parameters from megapixel multispectral images with inverse Monte Carlo and random forest regression.** *International Journal of Computer Assisted Radiology and Surgery*, 11(6):909–917. ISSN 1861-6410, 1861-6429. doi: 10.1007/s11548-016-1376-5. URL <http://link.springer.com/10.1007/s11548-016-1376-5>.
- Wirkert, S. J., Vemuri, A. S., Kenngott, H. G., Moccia, S., Götz, M., Mayer, B. F. B., Maier-Hein, K. H., Elson, D. S., and Maier-Hein, L. (2017). **Physiological Parameter Estimation from Multispectral Images Unleashed.** In *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2017*, Lecture Notes in Computer Science, pages 134–141. Springer, Cham. ISBN 978-3-319-66178-0 978-3-319-66179-7. doi: 10.1007/978-3-319-66179-7\_16. URL [https://link.springer.com/chapter/10.1007/978-3-319-66179-7\\_16](https://link.springer.com/chapter/10.1007/978-3-319-66179-7_16).
- Wirkert, S. J. (2018). *Multispectral image analysis in laparoscopy – A machine learning approach to live perfusion monitoring.* PhD thesis, Karlsruher Institut für Technologie (KIT).
- Wood, T. C., Thiemjarus, S., Koh, K. R., Elson, D. S., and Yang, G. Z. (2008). **Optimal feature selection applied to multispectral fluorescence imaging.** *Med Image Comput Assist Interv*, 11(Pt 2):222–229.
- Xiao, J., Gu, S., and Zhang, L. (2020). **Multi-Domain Learning for Accurate and Few-Shot Color Constancy.** In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3255–3264. IEEE. ISBN 978-1-7281-7168-5. doi: 10.1109/CVPR42600.2020.00332. URL <https://ieeexplore.ieee.org/document/9157096/>.
- Yan, S. and Fang, Q. (2020). **Hybrid mesh and voxel based monte carlo algorithm for accurate and efficient photon transport modeling in complex bio-tissues.** *Biomed. Opt. Express*, 11(11):6262–6270. doi: 10.1364/BOE.409468. URL <http://opg.optica.org/boe/abstract.cfm?URI=boe-11-11-6262>.

- Yang, H. H. and Moody, J. (2000). **Data visualization and feature selection: New algorithms for nongaussian data.** In *Advances in Neural Information Processing Systems*, pages 687–693.
- Yang, L., Jiang, J., Gao, X., Wang, Q., Dou, Q., and Zhang, L. (2022). **Autonomous environment-adaptive microrobot swarm navigation enabled by deep learning-based real-time distribution planning.** *Nature Machine Intelligence*. ISSN 2522-5839. doi: 10.1038/s42256-022-00482-8. URL <https://doi.org/10.1038/s42256-022-00482-8>.
- Yoo, J.-S. and Kim, J.-O. (2019). **Dichromatic Model Based Temporal Color Constancy for AC Light Sources.** In *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 12321–12330. IEEE. ISBN 978-1-7281-3293-8. doi: 10.1109/CVPR.2019.01261. URL <https://ieeexplore.ieee.org/document/8953584/>.
- Yoon, J., Grigoriou, A., and Id, S. E. B. (2020). **A background correction method to compensate illumination variation in hyperspectral imaging.** *PLOS ONE*, pages 1–21. doi: 10.1371/journal.pone.0229502. URL <http://dx.doi.org/10.1371/journal.pone.0229502>.
- Zhao, D., Fu, H., Xiao, L., Wu, T., and Dai, B. (2018). **Multi-object tracking with correlation filter for autonomous vehicle.** *Sensors (Basel, Switzerland)*, 18(7):2004. ISSN 1424-8220. doi: 10.3390/s18072004. URL <https://pubmed.ncbi.nlm.nih.gov/29932136https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6068606/>.
- Zhao, Q., Adeli, E., and Pohl, K. M. (2020). **Training confounder-free deep learning models for medical applications.** *Nature Communications* 2020 11:1, 11:1–9. ISSN 2041-1723. doi: 10.1038/s41467-020-19784-9. URL <https://www.nature.com/articles/s41467-020-19784-9>.
- Zhu, X., Huang, Q., DiSpirito, A., Vu, T., Rong, Q., Peng, X., Sheng, H., Shen, X., Zhou, Q., Jiang, L., Hoffmann, U., and Yao, J. (2022). **Real-time whole-brain imaging of hemodynamics and oxygenation at micro-vessel resolution with ultrafast wide-field photoacoustic microscopy.** *Light: Science and Applications* 2022 11:1, 11(1):1–15. ISSN 2047-7538. doi: 10.1038/s41377-022-00836-2. URL <https://www.nature.com/articles/s41377-022-00836-2>.

## Own Contributions and Publications

This chapter gives an overview of my contributions in distinction to team efforts.

### Own share in data acquisition and data analysis

This thesis was written in the division of Intelligent Medical Systems (IMSY former CAMI) headed by Prof. Dr. Lena Maier-Hein, who is the primary supervisor for this thesis. Throughout the entire time of my thesis, I was closely collaborating with members of Prof. Dr. Lena Maier-Hein's group, and medical collaborators.

The thesis comprises several interconnected components, mainly: 1) **Oxygenation estimation with spectral imaging**, 2) **Ischemia monitoring by Out-of-Distribution (OoD)** and 3) **Automatic light source estimation**. Due to the translational nature of my thesis, both the data acquisition and analysis involved a team effort from a highly interdisciplinary group. In the following, I attempt to clarify my particular involvement in the different stages of my thesis.

**Own share in Data Acquisition.** The *in vivo* data used for the band selection and domain adaptation experiments was collected by the team of Prof. Baowei Fei. The *in silico* data that was utilized for these experiments, and throughout the rest of this thesis, was generated by myself and Dr. Sebastian Wirkert. The data used in the brain perfusion experiments was collected by myself in conjunction with the team from both Prof. Dr. Lena Maier-Hein and Dr. Edgar Santos. The surgical procedure to expose the brain cortex in the pig model was carried out by the team of Dr. Edgar Santos at Heidelberg University. The data presented in the clinical trial for skin perfusion assessment was collected by myself with the help of M.Sc. Diana Mindroc at the German Cancer Research Center. The

organization and planning of this trial was carried out in conjunction with the team of Prof. Dr. Lena Maier-Hein.

The clinical trial for ischemia monitoring in laparoscopic surgery was organized by myself in conjunction with Dr. Christina Engels, Prof. Dr. Dogu Teber and the team of Prof. Dr. Lena Maier-Hein. All measurements and patient recruitment took place at the Städtisches Klinikum Karlsruhe. The data from this clinical trial was collected by myself together with both the teams from Prof. Dr. Dogu Teber and Prof. Dr. Lena Maier-Hein. In particular, the organization and implementation of this trial was equally carried out by M.Sc. Tim Adler and me.

The *ex vivo* data utilized for the automatic light source experiments was collected by M.Sc. Silvia Seidlitz while I collected the *in vivo* data. The organization of these experiments was carried out by both myself and Silvia Seidlitz.

**Own share in Data Analysis.** I was involved in the development and implementation of all the data analysis presented in this thesis. More in particular, the analysis of the *in vivo* data presented in the band selection and domain adaptation experiments was carried out in conjunction with Dr. Sebastian Wirkert and Dr. Fabian Isensee, while I performed the ranking analysis of the different band selection algorithms and provided help for the rest of the *in silico* analysis.

The analysis of the data collected for the brain perfusion experiments was performed by myself with the help of the team from Prof. Dr. Maier-Hein. I was mainly in charge of the development and implementation of the algorithms presented in this section. Furthermore, the analysis of the data from the skin perfusion clinical trial was performed equally by myself and M.Sc. Diana Mindroc. At the same time, I was responsible for the organization and implementation of this project.

The analysis of the data from the clinical trial on ischemia monitoring during laparoscopic surgery was equally carried out by myself and M.Sc. Tim Adler. I was responsible for the development of the deep learning-based Region of Interest (ROI) tracker, the implementation of the baseline perfusion estimators and the statistical analysis of the data, while M.Sc. Tim Adler was mainly in charge of the development and implementation of the deep learning OoD estimator and statistical analysis. Both, M.Sc. Tim Adler and me, collaborated constantly throughout the analysis of this data, we both provided insightful discussions on one another's work. The mixed models analysis was performed by Dr. Manuel Wiesenfarth, Dr. Nicholas Schreck and the rest of the team of Prof. Dr. Annette Kopp-Schneider. They also provided guidance for the statistical analysis of this trial.

The analysis of the data from the automatic light source estimation experiments was equally performed by myself and M.Sc. Silvia Seidlitz.

## Own publications

In this section, all publications that I was a part of and contributed to during my Ph.D. work are listed. It is subdivided into *First Authorship* and *Co-Authorship*.

### First Authorship - Peer Reviewed International Journal Publications

L. Ayala, F. Isensee, S. Wirkert, A. Vemuri, K. Maier-Hein, B. Fei, L. Maier-Hein. **Biomedical Optics Express**. *Band selection for oxygenation estimation with multispectral/hyperspectral imaging*. doi:10.1364/BOE.441214.

L. Ayala, S. Seidlitz, A. Vemuri, S. Wirkert, T. Kirchner, T. Adler, C. Engels, D. Teber, L. Maier-Hein. **International Journal of Computer Assisted Radiology and Surgery**. *Light source calibration for multispectral imaging in surgery*. doi:10.1007/s11548-020-02195-y.

### First Authorship - Peer Reviewed International Conference Publications

L. Ayala, S. Wirkert, J. Gröhl, M. Herrera, A. Hernandez-Aguilera, A. Vemuri, E. Santos, L. Maier-Hein. **International Workshop on OR 2.0 Context-Aware Operating Theaters (MICCAI 2019)**. *Live monitoring of Haemodynamic changes with Multispectral image analysis*. doi:10.1007/978-3-030-32695-1\_5.

### Co-Authorship

J. Gröhl, K. Dreher, M. Schellenberg, T. Rix, N. Holzwarth, P. Vieten, L. Ayala, S. Bohndiek, A. Seitel, L. Maier-Hein. **Journal of Biomedical Optics**. *SIMPA: an open-source toolkit for simulation and image processing for photonics and acoustics*. doi:10.1117/1.JBO.27.8.083010.

D. Teber, C. Engels, L. Maier-Hein, L. Ayala, S. Onogur, A. Seitel, K. März. **Der Urologe**. *Surgery 4.0-are we ready?*. doi:10.1007/s00120-020-01272-z.


T. Adler, L. Ayala, L. Ardizzone, H. Kenngott, A. Vemuri, B. Müller-Stich, C. Rother, U. Köthe, L. Maier-Hein. **International Workshop on Uncertainty for Safe Utilization of Machine Learning in Medical Imaging (MICCAI 2019)**. *Out of distribution detection for intra-operative functional imaging*. doi:10.1007/978-3-030-32689-0\_8.

# LEONARDO AYALA

Data scientist / Physicist

 Heidelberg, Germany

 leonardo.menjivar@outlook.com

 leonardo-ayala

 +49 1522 6652 484

 leoyala.github.io/whispers

 leoyala

## EDUCATION

### German Cancer Research Center

Heidelberg, Germany

PHD STUDIES

02/2018 - today

- Data science, machine learning and spectral imaging
- Development of physics-based Monte Carlo light transport framework [SIMPA](#)
- Functional imaging in surgery
- Leader of Surgical Spectral Imaging team since 2020

### Instituto Balseiro

Bariloche, Argentina

MSc IN PHYSICS

07/2015 - 12/2016

- Master Thesis: Study of metal/oxide composite nanomaterials for their application as carbon-tolerant solid oxide fuel cell electrodes<sup>1</sup>

### Instituto Balseiro

Bariloche, Argentina

BSc IN PHYSICS

07/2013 - 12/2015

- Bachelor Thesis: Study of metal/oxide composite nanomaterials for their application as carbon-tolerant solid oxide fuel cell electrodes

### Universidad de El Salvador

San Salvador, El Salvador

BSc IN PHYSICS

02/2010 - 07/2013

### Instituto Nacional de Nueva Concepción

Chalatenango, El Salvador

HIGH SCHOOL EDUCATION

01/2008 - 12/2009

## PROFESSIONAL EXPERIENCE

### Optics Laboratory at Universidad de El Salvador

San Salvador, El Salvador

ASSISTING RESEARCHER

01/2017 - 01/2018

- Data analyst for optical measurements

### Young Talent Program (Programa Jovenes Talento)

San Salvador, El Salvador

PHYSICS TEACHER

01/2017 - 12/2017

- Physics instructor
- Teacher for team participating international physics olympiads

### TEMACH High School

San Salvador, El Salvador

SUBSTITUTE TEACHER

04/2017 - 12/2017

- Mathematics teacher

### Young Talent Program (Programa Jovenes Talento)

San Salvador, El Salvador

PHYSICS TEACHER

01/2011 - 07/2013

- Physics instructor
- Teacher for team participating international physics olympiads

## HONORS / AWARDS

2019	<b>Bench to Bedside award</b> , International Conference MICCAI OR 2.0	Germany
2019	<b>Young investigator award</b> , International Conference SMIT	Germany
2015	<b>Three year fellowship</b> , Helmholtz International Graduate School	Germany
2015	<b>Scholarship for Msc education</b> , Instituto Balseiro	Argentina
2013	<b>Scholarship for Bsc education</b> , Instituto Balseiro	Argentina
2009	<b>Scholarship for higher education</b> , High school state proficiency exams	El Salvador
2009	<b>1st Place</b> , High School state proficiency exams	El Salvador

<sup>1</sup>Continuation of BSc. topic



## ABOUT ME

Data scientist, Physicist, Software geek  
"You are never wrong to do the right thing" –Mark Twain

## SKILLS

Python

VTK

PlotlyDASH-RAPIDS

CUDA

Linux

C++

0 +500 h

## LANGUAGES

English • Fluent

German • Conversational

Spanish • Mother tongue

## HOBBIES

Bike riding, reading, soccer, badminton, writing

## NON PROFIT

- [WeltHungerHilfe](#)
- [UNICEF](#)

## KEYNOTES / INVITED TALKS

### Computational Biophotonics - Translating Optical Technologies to Surgery

2019

IEE EMBS INTERNATIONAL STUDENT CONFERENCE

Magdeburg, Germany

## FIRST AUTHOR PUBLICATIONS

### Spectral imaging enables contrast agent-free real-time ischemia monitoring in laparoscopic surgery

2022

L. AYALA, T. ADLER, S. SEIDLITZ, S. WIRKERT, C. ENGELS, A. SEITEL, J. SELLNER, A. AKSENOV, M. BODENBACH, P. BADER, S. BARON, A. VEMURI, M. WIESENFARTH, N. SCHRECK, D. MINDROC, M. TIZABI, S. PIRMAN, B. EVERITT, A. KOPP-SCHNEIDER, D. TEBER, L. MAIER-HEIN. MEDRXIV [doi:10.1101/2022.03.08.22271465](https://doi.org/10.1101/2022.03.08.22271465)

### Band selection for oxygenation estimation with multispectral/hyperspectral imaging

2022

L. AYALA, F. ISENSEE, S. WIRKERT, A. VEMURI, K. MAIER-HEIN, B. FEI, L. MAIER-HEIN. BIOMEDICAL OPTICS EXPRESS. [doi:10.1364/BOE.441214](https://doi.org/10.1364/BOE.441214)

### Video-rate multispectral imaging in laparoscopic surgery: First-in-human application

2021

L. AYALA, S. WIRKERT, A. VEMURI, T. ADLER, S. SEIDLITZ, S. PIRMAN, C. ENGELS, D. TEBER, L. MAIER-HEIN. ARXIV. [arXiv:2105.13901](https://arxiv.org/abs/2105.13901)

### Light source calibration for multispectral imaging in surgery

2020

L. AYALA, S. SEIDLITZ, A. VEMURI, S. WIRKERT, T. KIRCHNER, T. ADLER, C. ENGELS, D. TEBER, L. MAIER-HEIN. INTERNATIONAL JOURNAL OF COMPUTER ASSISTED RADIOLOGY AND SURGERY. [doi:10.1007/s11548-020-02195-y](https://doi.org/10.1007/s11548-020-02195-y)

### Live monitoring of Haemodynamic changes with Multispectral image analysis

2020

L. AYALA, S. WIRKERT, J. GRÖHL, M. HERRERA, A. HERNANDEZ-AGUILERA, A. VEMURI, E. SANTOS, L. MAIER-HEIN. INTERNATIONAL WORKSHOP ON OR 2.0 CONTEXT-AWARE OPERATING THEATERS (MICCAI 2019). [doi:10.1007/978-3-030-32695-1\\_5](https://doi.org/10.1007/978-3-030-32695-1_5)

## CO-AUTHOR PUBLICATIONS

### Robust deep learning-based semantic organ segmentation in hyperspectral images

2021

S. SEIDLITZ, J. SELLNER, J. ODENTHAL, B. ÖZDEMİR, A. STUDIER-FISCHER, S. KNÖDLER, L. AYALA, T. ADLER, H. KENNGOTT, M. TIZABI, M. WAGNER, F. NICKEL, B. MÜLLER-STICH, L. MAIER-HEIN. ARXIV. [arXiv:2111.05408](https://arxiv.org/abs/2111.05408)

### Machine learning-based analysis of hyperspectral images for automated sepsis diagnosis

2021

M. DIETRICH, S. SEIDLITZ, N. SCHRECK, M. WIESENFARTH, P. GODAU, M. TIZABI, J. SELLNER, S. MARX, S. KNÖDLER, M. ALLERS, L. AYALA, K. SCHMIDT, T. BRENNER, A. STUDIER-FISCHER, F. NICKEL, B. MÜLLER-STICH, A. KOPP-SCHNEIDER, M. WEIGAND, L. MAIER-HEIN. ARXIV. [arXiv:2106.08445](https://arxiv.org/abs/2106.08445)

### Spectral organ fingerprints for intraoperative tissue classification with hyperspectral imaging

2021

A. STUDIER-FISCHER, S. SEIDLITZ, J. SELLNER, M. WIESENFARTH, L. AYALA, B. OZDEMIR, J. ODENTHAL, S. KNOEDLER, K. KOWALEWSKI, C. HANEY, I. CAMPLISSON, M. DIETRICH, K. SCHMIDT, G. SALG, H. KENNGOTT, T. ADLER, N. SCHRECK, A. KOPP-SCHNEIDER, K. MAIER-HEIN, L. MAIER-HEIN, B. MÜLLER-STICH, F. NICKEL. BIORXIV. [doi:10.1101/2021.11.24.469943](https://doi.org/10.1101/2021.11.24.469943)

**Optimization of anastomotic technique and gastric conduit perfusion with hyperspectral imaging in an experimental model for minimally invasive esophagectomy**

2021

F. NICKEL, A. STUDIER-FISCHER, B. ÖZDEMİR, J. ODENTHAL, L. MUELLER, S. KNOEDLER, K. KOWALEWSKI, I. CAMPLISSON, M. ALLERS, M. DIETRICH, K. SCHMIDT, G. SALG, H. KENNGOTT, A. BILLETER, I. GOCKEL, C. SAGIV, O. ETZ-HADAR, J. GILDENBLAT, **L. AYALA**, S. SEIDLITZ, L. MAIER-HEIN, B. MÜLLER-STICH. [doi:10.1101/2021.10.03.462901](https://doi.org/10.1101/2021.10.03.462901)

**Surgery 4.0-are we ready?**

2020

D. TEBER, C. ENGELS, L. MAIER-HEIN, L. AYALA, S. ONOGUR, A. SEITEL, K. MÄRZ. DER UROLOGE. [doi:10.1007/s00120-020-01272-z](https://doi.org/10.1007/s00120-020-01272-z)

**Out of distribution detection for intra-operative functional imaging**

2019

T. ADLER, **L. AYALA**, L. ARDIZZONE, H. KENNGOTT, A. VEMURI, B. MÜLLER-STICH, C. ROTHER, U. KÖTHE, L. MAIER-HEIN. INTERNATIONAL WORKSHOP ON UNCERTAINTY FOR SAFE UTILIZATION OF MACHINE LEARNING IN MEDICAL IMAGING (MICCAI 2019). [doi:10.1007/978-3-030-32689-0\\_8](https://doi.org/10.1007/978-3-030-32689-0_8)

**Uncertainty-aware performance assessment of optical imaging modalities with invertible neural networks**

2019

T. ADLER, L. ARDIZZONE, A. VEMURI, **L. AYALA**, J. GRÖHL, T. KIRCHNER, S. WIRKERT, J. KRUSE, C. ROTHER, U. KÖTHE, L. MAIER-HEIN. INTERNATIONAL JOURNAL OF COMPUTER ASSISTED RADIOLOGY AND SURGERY. [doi:10.1007/s11548-019-01939-9](https://doi.org/10.1007/s11548-019-01939-9)

**POSTERS / ABSTRACTS**

**Deep learning for spectral image synthesis**

2022

T. RIX, M. HÜBNER, K. DREHER, J. NÖLKE, **L. AYALA**, M. SCHELLENBERG, J. SELLNER, S. SEIDLITZ, A. STUDIER-FISCHER, B. MÜLLER-STICH, F. NICKEL, A. SEITEL, L. MAIER-HEIN. PHOTONICS WEST, MULTIMODAL BIOMEDICAL IMAGING XVII. [doi:10.1117/12.2608622](https://doi.org/10.1117/12.2608622)



## Acknowledgement

*None of us, including me, ever do great things. But we can all do small things, with great love, and together we can do something wonderful.*

— MOTHER THERESA

To the unnamed that hid behind pages, to the invisible hands that hold mine as I write this, and to those who turned on the light while we traversed this treacherous path together, this is for you. I began this path alone, but as I stand at the end, the end of the beginning that is, I can say with absolute certainty that I am grateful for your company. Over the years, I have picked up people along the path and made them part of my “tribe”; or perhaps they are the ones who picked me. Here, I would like to thank you all (in no particular order), and let you know that I feel a great deal of gratitude to you.

In particular, I feel truly grateful to you, **Lena**, for all the guidance and support that you demonstrated thought the time I was your student. We faced, and hit, many barriers, but I like to think we came out to the other side in pretty good shape. I would also like to thank the entire team of the Division of Intelligent Medical Systems (IMSY former CAMI), both Alumni and current members. Specially thank you, **Sebastian Wirkert, Janek Gröhl and Anant Vemuri**, you were my rock at the beginning of my PhD. Thanks also to you, all the students that I supervised or partially supervised: **Sebastian Pirmann, Marco Hübner, Nina Sautter and Stefanie Martin**, you taught me more than you think. Special thanks also to you **Alexander Seitel** for being my compass when I was lost and could not see the shoreline. Thank you, **Silvia Seidlitz and Tim Adler** for all your collaboration and support during these years, you have been wonderful teammates, I cannot put in words how much you have supported me. Also thank you **Dogu Teber and Christina Engels** for making a large portion of the content of this thesis possible, I can certainly say that it would not have been possible without you! Many thanks to you **Maike Rees**,

**Matthias Eisenmann, Lisa Kausch, Tim Adler, Alexander Seitel, Kris Dreher, Annika Reinke, Marco Hübner and Ahmad Bin Qasim** for helping proofread parts of my thesis. As I once read, *“To write is human, to edit is divine”* — STEPHEN KING, **On Writing: A Memoir of the Craft** —.

My family played a gigantic role on whom I have become, and for that, I will forever be thankful. Thank you for pushing me to pursue science, for all the sacrifices you made, for all the long study sessions, for making me believe in myself; thank you for all of it, and much more. Thank you, also to all my friends for keeping me sane all these years and not letting me fall, I do not know what I will do without you **Ignacio Oliva, Karla Gámez and Andrea Escobar**. Especially you **Lisa**, my personal idol, thank you for all your support, for everything you taught me, for all the remarkable times, for all the challenges we overcame, for walking this path with me, for all I don't know how to put in words and much more; *“I would have written you, myself, if I could put down in words everything I want to say to you. A sea of ink would not be enough. But you built me dreams instead.”* — ERIN MORGENSTERN, **The night circus** —

There have been a great deal of people that helped me be who I am today and that propelled me to reach this point in my life, If I forget to mention someone by name, I can only ask for your forgiveness, but know that I treasure you and what you did for me.

## Eidesstattliche Versicherung

# Statutory Declaration

1. Bei der eingereichten Dissertation zu dem Thema *Translational Functional Imaging in Surgery Enabled by Deep Learning* handelt es sich um meine eigenständig erbrachte Leistung.

*I herewith formally declare that I have written the submitted dissertation *Translational Functional Imaging in Surgery Enabled by Deep Learning* independently.*

2. Ich habe nur die angegebenen Quellen und Hilfsmittel benutzt und mich keiner unzulässigen Hilfe Dritter bedient. Insbesondere habe ich wörtlich oder sinngemäß aus anderen Werken übernommene Inhalte als solche kenntlich gemacht.

*I did not use any third party support except for the quoted literature and other sources mentioned in the text. Content from other work, either literally or in content, has been declared as such.*

3. Die Arbeit oder Teile davon habe ich bislang nicht an einer Hochschule des In- oder Auslands als Bestandteil einer Prüfungs- oder Qualifikationsleistung vorgelegt.

*The thesis has not been submitted to any examination body in this, or similar, form.*

4. Die Richtigkeit der vorstehenden Erklärungen bestätige ich.

*I confirm the correctness of the aforementioned declarations.*

5. Die Bedeutung der eidesstattlichen Versicherung und die strafrechtlichen Folgen einer unrichtigen oder unvollständigen eidesstattlichen Versicherung sind mir bekannt. Ich versichere an Eides statt, dass ich nach bestem Wissen die reine Wahrheit erklärt und nichts verschwiegen habe.

*I am aware of the legal consequences of this declaration. To the best of my knowledge I have told the pure truth and not concealed anything.*

Heidelberg, 31.10.2022

---

Leonardo Antonio Ay-  
ala Menjivar