Physiological Parameter Estimation from Multispectral Images Unleashed

Sebastian J. Wirkert[1(✉)], Anant S. Vemuri[1], Hannes G. Kenngott[2],
Sara Moccia[1,3,4], Michael Götz[5], Benjamin F.B. Mayer[2],
Klaus H. Maier-Hein[5], Daniel S. Elson[6,7], and Lena Maier-Hein[1]

[1] Division of Computer Assisted Medical Interventions, DKFZ, Heidelberg, Germany
s.wirkert@dkfz-heidelberg.de
[2] Department for General, Visceral and Transplantation Surgery,
Heidelberg University Hospital, Heidelberg, Germany
[3] Department of Advanced Robotics, Istituto Italiano di Tecnologia, Genoa, Italy
[4] Department of Electronics, Information and Bioengineering,
Politecnico di Milano, Milan, Italy
[5] Division for Medical and Biological Informatics, DKFZ, Heidelberg, Germany
[6] Hamlyn Centre for Robotic Surgery, Imperial College London, London, UK
[7] Department of Surgery and Cancer, Imperial College London, London, UK

Abstract. Multispectral imaging in laparoscopy can provide tissue reflectance measurements for each point in the image at multiple wavelengths of light. These reflectances encode information on important physiological parameters not visible to the naked eye. Fast decoding of the data during surgery, however, remains challenging. While model-based methods suffer from inaccurate base assumptions, a major bottleneck related to competing machine learning-based solutions is the lack of labelled training data. In this paper, we address this issue with the first transfer learning-based method to physiological parameter estimation from multispectral images. It relies on a highly generic tissue model that aims to capture the full range of optical tissue parameters that can potentially be observed *in vivo*. Adaptation of the model to a specific clinical application based on unlabelled *in vivo* data is achieved using a new concept of domain adaptation that explicitly addresses the high variance often introduced by conventional covariance-shift correction methods. According to comprehensive *in silico* and *in vivo* experiments our approach enables accurate parameter estimation for various tissue types without the need for incorporating specific prior knowledge on optical properties and could thus pave the way for many exciting applications in multispectral laparoscopy.

1 Introduction

Multispectral images (MSI) offer great potential in a large variety of medical procedures. The encoded information about tissue parameters such as oxygenation and blood volume fraction has motivated a considerable body of research related to early cancer detection [1] as well as image-guided therapy involving
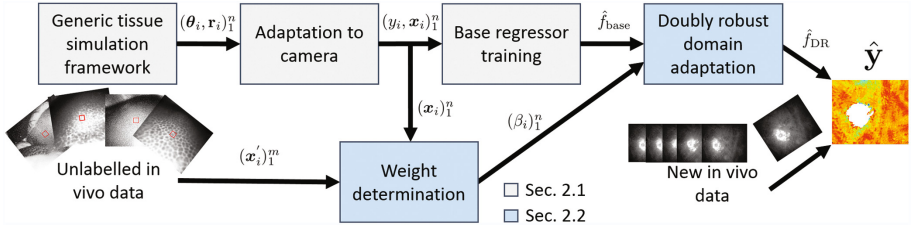
Fig. 1. Overview on our approach. First we create masses of generic, labelled, tissue-reflectance samples $(\boldsymbol{\theta}_i, \mathbf{r}_i)_1^n$. The desired physiological parameter y_i and the reflectances adapted to the camera $\mathbf{x}_i$ are used for training a base regressor $\hat{f}_{\text{base}}$. Weights $(\beta_i)_1^n$ are calculated to fit the simulated data to the measurements. These weights adapt the base regressor to the *in vivo* measurements. Applying the adapted regressor $\hat{f}_{\text{DR}}$ to new images yields their parameter estimates $\hat{\mathbf{y}}$.

bowel anastomosis [2] and transplantation evaluation. However, decoding the reflectance measurements during a medical intervention is not straightforward. Model-based approaches that are sufficiently fast for online execution typically suffer from incorrect base assumptions such as constant scattering or low tissue absorption [2]. Machine learning-based alternatives [3,4] need accurate information about the composition of the underlying tissue for the training phase to account for the lack of annotated real data. This knowledge about optical tissue properties is hard to obtain and is dependent on experimental conditions [5].

To address this bottleneck, we present a novel machine learning-based approach to physiological parameter estimation (Sect. 2) that neither requires real labelled data nor specific prior knowledge on the optical properties of the tissue of interest. The method relies on a broadly applicable model of abdominal tissue that aims to capture a large range of physiological parameters observed *in vivo*. Adaptation of the model to a specific clinical application is achieved by means of domain adaptation (DA) using samples of unlabelled *in vivo* data.

In a comprehensive study (Sect. 3) with seven pigs we show that (1) our model captures a large amount of the variation in real tissue and (2) our transfer learning-based approach enables highly accurate physiological parameter estimation.

2 Methods

Our approach to physiological parameter estimation, which is illustrated in Fig. 1, aims to compensate for the lack of detailed prior knowledge related to optical tissue properties by applying DA. More specifically, we hypothesise that (1) we can use a highly generic tissue model to generate a data set of spectral reflectances that covers the whole range of multispectral measurements that may possibly be observed *in vivo* and that (2) samples of real (unlabelled) measurements can be used to adapt the data to a new target domain. Section 2.1

describes how the generic data set is generated and used for physiological parameter estimation while Sect. 2.2 introduces our approach to DA.

2.1 Generic Approach to Physiological Parameter Estimation in Multispectral Imaging

Our method is built around a single comprehensive data set, which can be used for various camera setups and target structures. Its generation and usage is as described below.

Dataset Generation Using a Generic Tissue Model. A generalization of the layered tissue model developed in [3] was used to create our data set consisting of camera independent tissue-reflectance pairs $(\boldsymbol{\theta}_i, \mathbf{r}_i), i \in \{1...n\}$. The (physiological) parameters are varied within the ranges shown in Table 1 for each of the three layers. Each layer is described by its value for blood volume fraction v_{hb}, scattering coefficient a_{mie}, scattering power b_{mie}, anisotropy g, refractive index n and layer thickness d. Oxygenation s is kept constant across layers [3]. Following the values in [5] and in contrast to [3], b_{mie} is varied, covering all soft, fatty and fibrous tissues. We increase the ranges of v_{hb} by a factor of three to potentially model pathologies. In conjunction with values for haemoglobin extinction coefficients ϵ_{Hb} and ϵ_{HbO2} from the literature, absorption and scattering coefficients μ_a and μ_s can be determined for usage in the Monte Carlo simulation framework. The simulated range of wavelengths λ is large enough for adapting the simulations to cameras operating in the visible and near infrared.

To account for a specific camera setup, $\mathbf{r}_i$ can be transformed to camera reflectances $(c_{i,j})$ at the jth spectral band, using the method described in [4]. Zero mean Gaussian noise w was added to model camera noise.

Table 1. The simulated ranges of physiological parameters, and their usage in the simulation set-up as described in Sect. 2.1. Values used in [3] are denoted if different. Important changes are marked in bold font.

θ	$v_{\mathrm{hb}}[\%]$	$s[\%]$	$a_{\mathrm{mie}}[\frac{1}{\mathrm{cm}}]$	b_{mie}	g	n	$d[\mu\mathrm{m}]$
Layer 1–3:	**0–30**	0–100	5–50	**.3–3**	.8–.95	1.33–1.54	**20–2000**
In [3]:	0–10			1.3		1.36	395–1010

$\mu_a(v_{\mathrm{hb}}, s, \lambda) = v_{\mathrm{hb}}(s\epsilon_{\mathrm{HbO2}}(\lambda) + (1-s)\epsilon_{\mathrm{Hb}}(\lambda))ln(10)150\,\mathrm{g\,L}^{-1}(64,500\,\mathrm{g\,mol}^{-1})^{-1}$

$\mu_s(a_{\mathrm{mie}}, b, \lambda) = \frac{a_{\mathrm{mie}}}{1-g}(\frac{\lambda}{500\,\mathrm{nm}})^{-b_{\mathrm{mie}}}$

Simulation framework: GPU-MCML [6], 10^6 photons per simulation.
Simulated samples: **500** K ([3]: 10 K).
Sample wavelength range: **300–1000** nm ([3]: 450–720 nm), stepsize 2 nm.

Physiological Parameter Regression. To train a regressor for specific physiological parameters $y_i \in \boldsymbol{\theta}_i$, such as oxygenation and blood volume fraction, the corresponding reflectances have to be normalized ($\mathbf{x}_i = \frac{\mathbf{c}_i}{\sum_j c_{i,j}}$) to account for multiplicative factors due to changes in light intensity or camera pose. The combination of normalized camera reflectances and corresponding physiological parameter $(\mathbf{X}, \mathbf{y})$ serve as training data for any machine learning regression method f to obtain the regressor $\hat{f}_{\text{base}}$, which corresponds to a regressor without DA. The physiological parameter estimates during an intervention can be determined using this baseline regressor by evaluating $\hat{f}_{\text{base}}(\mathbf{x}') = y'$ for each recorded multispectral image pixel $\mathbf{x}'$. The next section describes our DA technique to further improve the parameter estimation in a specific clinical context.

2.2 Domain Adaptation

Working with tissue samples from a simulated source domain $p_s(\mathbf{x}, y)$ will inevitably introduce a bias with respect to the target domain $p_t(\mathbf{x}, y)$. We speak of covariate shift if $p_s(y|\mathbf{x}) = p_t(y|\mathbf{x})$ and therefore $\frac{p_t(\mathbf{x},y)}{p_s(\mathbf{x},y)} = \frac{p_t(\mathbf{x})}{p_s(\mathbf{x})} =: \beta(\mathbf{x})$. If p_t is contained in the support of p_s, adding the weights $\boldsymbol{\beta}$ to the loss function of the regressor can correct for the covariate shift [7]. The appeal of this method is, that only recordings $\mathbf{x}'_i$ and no labels y'_i are necessary for adaptation. While the concept of covariate shift has been applied with great success in a number of different medical imaging applications [8,9], major challenges related to transferring it to our problem are estimation of high dimensional $\boldsymbol{\beta}$, and high variance introduced by weighting, both addressed in the next two subsections.

Finding $\boldsymbol{\beta}$ with Kernel Mean Matching and Random Kitchen Sinks. Kernel mean matching (KMM) is a state-of-the-art method for determining $\boldsymbol{\beta}$ [7]. KMM minimizes the mean distance of the samples of the two domains in a reproducing kernel hilbert space $\mathcal{H}$, using a possibly infinite dimensional lifting $\phi : \mathbb{R}^m \to \mathcal{H}$. In its original formulation [7], the kernel trick was used to pose KMM as a quadratic problem. In our problem domain this is not feasible, because calculating the Gram matrix is quadratic in the (high) number of samples. To overcome this bottleneck, we minimize the KMM objective function (see [7]) with an approximate representation of the lifting $\phi(\mathbf{x}_i) \approx z(\mathbf{x}_i)$ determined by the random kitchen sinks method [10]. This enables us to solve the convex KMM objective function in its non-kernelized form using a standard optimizer.

Doubly Robust Covariate Shift Correction. Estimators trained using weighted samples can yield worse result than estimators not accounting for the covariate shift. The reason is that only few samples are effectively "active", providing the risk minimizer with less samples. On the other hand using the unweighted training samples often leads to a reasonable, but biased, estimator [11]. Intuitively the unweighted base regressor $\hat{f}_{\text{base}}$ defined in Sect. 2.1 can be used to obtain an initial estimate. Subsequently, another estimator aims to refine the results with emphasis on the samples with high weight. This is the basic idea

of doubly robust (DR) covariate shift correction [11]. Specifically, we use the residuals $\delta_i = y_i - \hat{f}_{\text{base}}(\mathbf{x}_i)$ weighted by $\boldsymbol{\beta}$ from the last subsection to train an estimator $\hat{f}_{\text{res}}$ on $(\mathbf{X}, \boldsymbol{\delta}, \boldsymbol{\beta})$. The final estimate is $\hat{f}_{\text{dr}}(\mathbf{x}') = \hat{f}_{\text{base}}(\mathbf{x}') + \hat{f}_{\text{res}}(\mathbf{x}')$.

3   Experiments and Results

Based on a comprehensive *in silico* and *in vivo* MSI data set (Sect. 3.1) we validate the quality of our generic tissue model (Sect. 3.2) as well as the performance of the DA based physiological parameter estimation approach (Sect. 3.3).

3.1   Experimental Setup

Images were recorded with a custom-built, multispectral laparoscope, capturing images at eight different wavebands [3] with the 5Mpix Pixelteq (Largo, FL, USA) Spectrocam. We recorded MI data from seven pigs and six organs (liver, spleen, gallbladder, bowel, diaphragm and abdominal wall) in a laparoscopic setting. For all our experiments we used the data set described in Sect. 2.1 for training. We used a random forest regressor with parameters as in [3]. We drew 1000 random directions from the reproducing kernel hilbert space induced by the radial basis function (RBF) kernel with the random kitchen sink method. We set the σ value of the RBF to the approximate median sample distance and the B parameter of the KMM to ten for all experiments.

3.2   Validity of Tissue Model

One of the prerequisites for covariate shift correction is that support of the distribution of true *in vivo* measurements is contained in the support of the simulated reflectances. To investigate this for a range of different tissue types (cf. Sect. 3.1) we collected a total of 57 images, extracted measurements from a 100×100 region of interest (ROI) and corrected them by a flatfield and dark image as in [2]. The first three principal components cover $99\% = 82\% + 13\% + 4\%$ of the simulated variance. For *in vivo* data, $97\% = 89\% + 4\% + 4\%$ of the variance lies on the simulated data's first three principal components.

For a qualitative assessment we projected the *in vivo* measurements on the first two principal components of the simulated data. A selection can be seen in Fig. 2. Apart from gallbladder, all of the *in vivo* data lie on the two dimensional manifold implied by the simulated data. Figure 3 illustrates how changes in oxygenation and perfusion influence the distribution of the measurements.

3.3   Performance of Domain Adaptation

To validate our approach to physiological parameter estimation with reliable reference data, we performed an *in silico* experiment with simulated colon tissue as target domain. For this purpose, we used 15,000 colon tissue samples with corresponding ground truth oxygenation from [3] at a signal-to-noise ratio
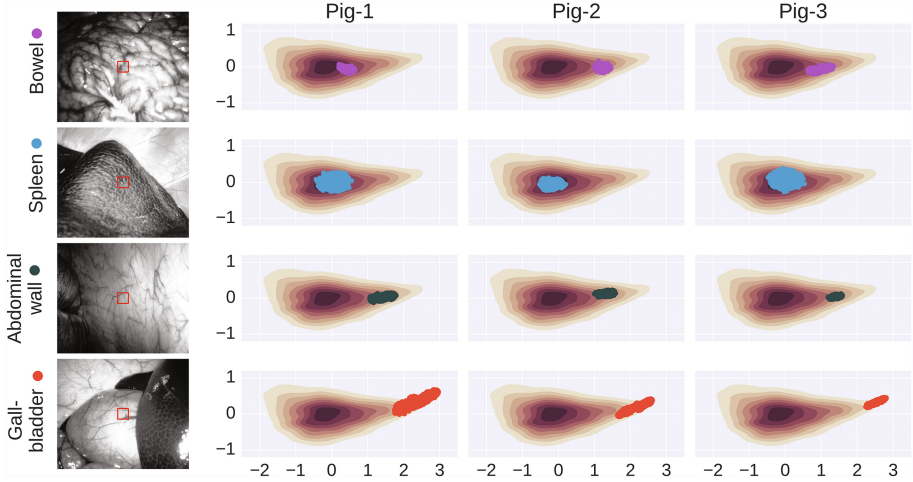
Fig. 2. Four organs from three pigs projected onto the first two principal components of our simulated reflectance data plotted in brown. The images on the left show the 560 nm band recorded for the first pig. The depicted measurements are taken from the red ROI. Except for gallbladder, all organs lie on the non-zero density estimates of the simulated data (See also Sect. 4).
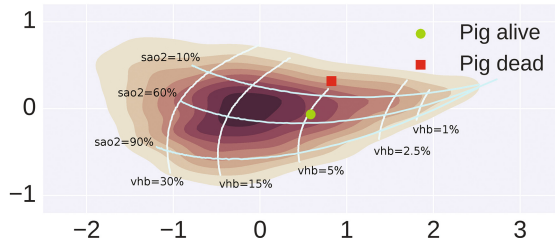

Fig. 3. Liver tissue measurements before and after sacrificing a pig. The grid indicates how varying oxygenation (sao2) and blood volume fraction (vhb) changes the measurements in the space spanned by the first two principal components of the simulations. Note that these lines can not be directly interpretated as sao2 and vhb values for the two points, because other factors such as scattering will cause movement on this simplified manifold.

(SNR $:= \frac{c_{i,j}}{w_{i,j}}$) of 20. 10,000 reflectance samples were selected for DA while the remaining 5,000 samples were used for testing.

We varied the number of training samples between 10^4 and $5 * 10^5$ to investigate how the effective sample size influences results. Our DR DA method reduced the median absolute oxygenation estimation error compared to the base estimator by 25–27% and by 14–25% without the DR correction (Fig. 4a). As expected the difference was smaller for higher effective sample sizes $m_{\text{eff}} = \frac{\|\boldsymbol{\beta}\|_1^2}{\|\boldsymbol{\beta}\|_2^2}$.

(a) *In silico* domain adaptation   (b) *In vivo* domain adaptation
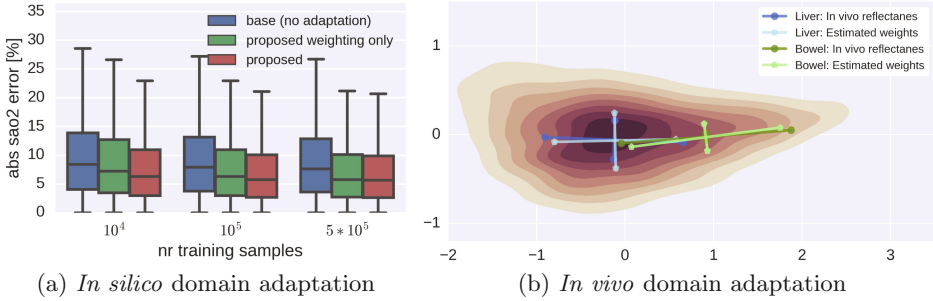
Fig. 4. *In silico* boxplot results (a) and *in vivo* (b) validation results corresponding to the experiments described in Sect. 3.3. (b) Shows the distribution of *in vivo* measurements and adapted *in silico* reflectances in the principal component space of the simulations. For graphical clarity the distributions are visualized as their two principal axes in this space with lengths corresponding to the eigenvalues.

In vivo experiments were performed for each organ by calculating the Euclidean distance of the weighted simulation mean $\frac{1}{n}\sum_i^n \beta_i \mathbf{x}_i$ to the mean of the images. The weighted distance was 36–92% (median 77%) smaller than the unweighted average. See Fig. 4b for a depiction in the principal component space.

4   Discussion

To our knowledge, this paper introduced the first transfer learning-based approach to physiological parameter estimation from multispectral imaging data. As it neither requires real labelled data nor specific prior knowledge on the optical properties of the tissue of interest and is further independent of the camera model and corresponding optics it is potentially broadly applicable to a wide range of clinical applications.

The method is built around a generic data set that can automatically be adapted to a given target anatomy based on samples of unlabelled *in vivo* data. According to porcine experiments with six different target structures, the first three principal components of our simulated data set capture 97% of measured *in vivo* variations. Our hypothesis is that these variations represent the blood volume fraction, oxygenation and scattering. Visual inspection of the first two principal components showed the captured organ data lie within the simulated data, an important prerequisite for the subsequent DA to work. Gallbladder is the exception, most likely due to its distinctive green stain, caused by the bile shining through. Modelling the bile as another chromophore and extending our data set accordingly would be straightforward. In future work we plan to capture an even higher variety of *in vivo* data, involving pathologies such as cancer.

Our experiments further demonstrate the potential performance boost when adapting the generalized model to a specific task using the presented DA technique. An important methodological component in this context was the integration of the recently proposed DR correction method to address the instabilities

when few effective training samples are selected. We also tested this method with another recently proposed DA weighting method [9] with similar results. Both bias and variance of the Euclidean distance of the weighted sample mean to the mean of the *in vivo* images reduced when increasing the number of porcines used for weight determination (not shown). The required training cases for a given application is an interesting future direction of research.

In conclusion, we have addressed the important bottleneck of lack of annotated MSI data, with a novel transfer learning-based method to physiological parameter estimation. Given the highly promising experimental results presented in this manuscript, future work will focus on evaluating the method for a variety of clinical applications including partial nephrectomy and cancer detection.

Conflict of Interest. The authors declare that they have no conflict of interest.

Compliance with Ethical Standards. This article does not contain any studies with human participants. All applicable international, national, and/or institutional guidelines for the care and use of animals were followed.

References

1. Claridge, E., Hidovic-Rowe, D.: Model based inversion for deriving maps of histological parameters characteristic of cancer from ex-vivo multispectral images of the colon. IEEE TMI **33**, 822–835 (2013)
2. Clancy, N.T., et al.: Intraoperative measurement of bowel oxygen saturation using a multispectral imaging laparoscope. Biomed. Opt. Express **6**(10), 4179–4190 (2015)
3. Wirkert, S.J.: others: Robust near real-time estimation of physiological parameters from megapixel multispectral images with inverse Monte Carlo and random forest regression. IJCARS **11**(6), 909–917 (2016)
4. Styles, I.B., et al.: Quantitative analysis of multi-spectral fundus images. Med. Image Anal. **10**(4), 578–597 (2006)
5. Jacques, S.L.: Optical properties of biological tissues: a review. Phys. Med. Biol. **58**(11), R37 (2013)
6. Alerstam, E., et al.: Next-generation acceleration and code optimization for light transport in turbid media using GPUs. Biomed. Opt. Express **1**(2), 658–675 (2010)
7. Huang, J., et al.: Correcting sample selection bias by unlabeled data. In: Advances in Neural Information Processing Systems, vol. 19 (2007)
8. Heimann, T., et al.: Real-time ultrasound transducer localization in fluoroscopy images by transfer learning from synthetic training data. Med. Image Anal. **18**(8), 1320–1328 (2014)
9. Goetz, M., et al.: DALSA: domain adaptation for supervised learning from sparsely annotated MR images. IEEE TMI **35**(1), 184–196 (2016)
10. Rahimi, A., et al.: Random features for large-scale kernel machines. In: NIPS, vol. 3 (2007)
11. Reddi, S.J., et al.: Doubly robust covariate shift correction. In: Twenty-Ninth AAAI Conference on Artificial Intelligence (2015)