



This MICCAI paper is the Open Access version, provided by the MICCAI Society. It is identical to the accepted version, except for the format and this watermark; the final published version is available on SpringerLink.

Deep intra-operative illumination calibration of hyperspectral cameras

Alexander Baumann^{1,2,7}, Leonardo Ayala^{2,7}, Alexander Studier-Fischer^{6,8,9,10},
Jan Sellner^{2,3,4,5}, Berkin Özdemir^{6,9,10}, Karl-Friedrich Kowalewski^{8,9,10},
Slobodan Ilic¹, Silvia Seidlitz^{2,3,4,5,*}, and Lena Maier-Hein^{2,3,4,5,7,*}

¹ Siemens AG, Munich, Germany

² Division of Intelligent Medical Systems, German Cancer Research Center (DKFZ), Heidelberg, Germany

³ Helmholtz Information and Data Science School for Health, Karlsruhe/Heidelberg, Germany

⁴ Faculty of Mathematics and Computer Science, Heidelberg University, Heidelberg, Germany

⁵ National Center for Tumor Diseases (NCT), NCT Heidelberg, a partnership between DKFZ and university medical center Heidelberg

⁶ Department of General, Visceral, and Transplantation Surgery, Heidelberg University Hospital, Heidelberg, Germany

⁷ Medical Faculty, Heidelberg University, Heidelberg, Germany

⁸ Department of Urology and Urosurgery, University Medical Center Mannheim, Medical Faculty of the University of Heidelberg, Mannheim, Germany

⁹ German Cancer Research Center (DKFZ) Heidelberg, Division of Intelligent Systems and Robotics in Urology (ISRU), Heidelberg, Germany

¹⁰ DKFZ Hector Cancer Institute at the University Medical Center Mannheim, Mannheim, Germany

Abstract. Hyperspectral imaging (HSI) is emerging as a promising novel imaging modality with various potential surgical applications. Currently available cameras, however, suffer from poor integration into the clinical workflow because they require the lights to be switched off, or the camera to be manually recalibrated as soon as lighting conditions change. Given this critical bottleneck, the contribution of this paper is threefold: (1) We demonstrate that dynamically changing lighting conditions in the operating room dramatically affect the performance of HSI applications, namely physiological parameter estimation, and surgical scene segmentation. (2) We propose a novel learning-based approach to automatically recalibrating hyperspectral images during surgery and show that it is sufficiently accurate to replace the tedious process of white reference-based recalibration. (3) Based on a total of 742 HSI cubes from a phantom, porcine models, and rats we show that our recalibration method not only outperforms previously proposed methods, but also generalizes across species, lighting conditions, and image processing tasks. Due to its simple workflow integration as well as high accuracy, speed, and generalization capabilities, our method could evolve as a central component in clinical surgical HSI.

* Silvia Seidlitz and Lena Maier-Hein contributed equally to this work.

Keywords: deep learning, hyperspectral imaging, intra-operative imaging, illumination calibration, surgical data science, semantic organ segmentation

1 Introduction

Hyperspectral imaging (HSI) emerges as a promising medical imaging modality that offers distinct advantages over conventional RGB imaging. In particular, HSI captures spectral information across numerous contiguous bands, thereby enriching the representation of the underlying sample. Recent works have demonstrated the resulting enhancement in tissue classification [10, 11, 20, 21, 25], and the capability of estimating physiological tissue parameters [1, 3, 6, 13, 15, 22, 26]. However, in open surgery, spectral data is affected by changes in illumination and must be correctly calibrated whenever lighting conditions vary [8]. The standard approach for experimental surgeries is to switch off all external light sources before acquisition to ensure precise measurements [24]. As this protocol severely disrupts the clinical workflow, it is presumed not to be consistently applied – leading to unreliable data acquisition and severe failures in downstream tasks, as illustrated in Fig. 1. This may be one reason why spectral imaging has not yet found widespread use in clinical practice. Conventional HSI calibration is conducted with physical white reference measurements, capturing the surrounding illumination. However, they pose challenges in terms of time consumption and sterility, rendering them impractical in the operating room (OR) context. The proposed use of white OR rulers as a sterile alternative [4] still presents considerable challenges in the form of additional workload and their small size. A number of automatic calibration algorithms originally devised for RGB imaging, such as Gray-world [5] and Max-RGB [16], recover a global illuminant of the scene based on intensity statistics. An alternative calibration strategy tailored

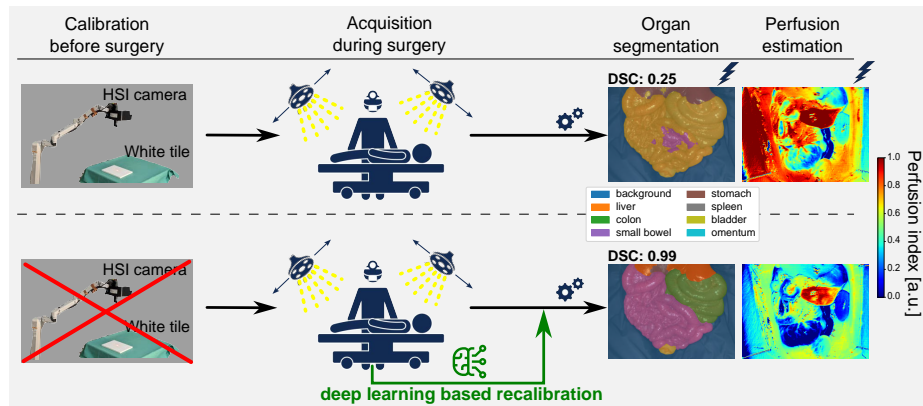


Fig. 1. Motivation: Current hyperspectral cameras, which require known lighting conditions, fail in real-world scenarios with dynamically changing lighting conditions.

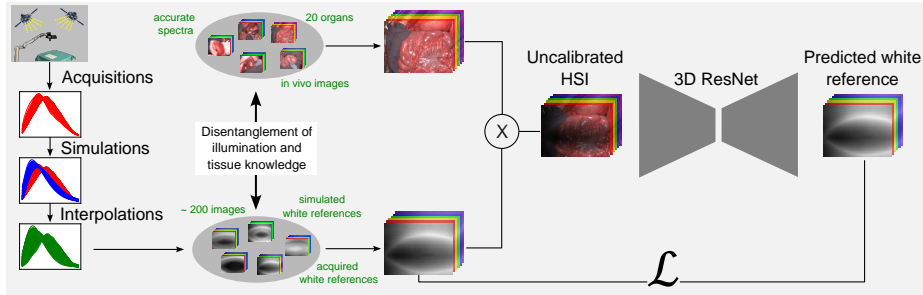


Fig. 2. The proposed approach replaces tedious manual calibration with a dynamic fully-automatic approach. The core of our data-centric method is a 3D-convolutional neural network, trained on in vivo data with artificial light manipulations. At inference time, it takes a raw hyperspectral image as input and generates the corresponding white reference image. The prediction of the white tile image can be used for subsequent calibration of the input image.

specifically to HSI leverages specular highlights for illuminant estimation [2]. However, these methods rely on unrealistic assumptions, such as homogeneous illumination across the entire surgical scene, and were also not tested in in vivo environments during surgery. Multi-illuminant color constancy models devised for RGB imaging have shown potential in overcoming this issue, as they predict pixel-wise illuminants for calibration [7, 9, 14, 19, 23]. Notably, convolutional neural networks (CNNs) have shown superior performance compared to non-learning-based methods [7, 23]. Spectral imaging has so far seen the development of one deep learning approach for multi-illuminant calibration, factorizing reflectance and illumination through an unrolling network [17].

Overall, the methods proposed in the literature either remain untested for surgical HSI and/or are conceptually not suitable for spatially resolved calibration. Given this bottleneck, the mission of our work was to develop a new workflow-optimized calibration approach that enables widespread clinical spectral imaging. Our specific contribution is threefold: (1) We demonstrate that dynamically changing lighting conditions in the OR dramatically affect the performance of in vivo HSI applications, and previously proposed calibration methods fail to restore optimal performance. (2) We present a novel learning-based approach to performing spatially resolved light recalibration of surgical hyperspectral images. Specifically, we propose to replace conventional physical white reference measurements with a data-driven prediction of the corresponding white tile measurement. This enables a seamless and sterile recalibration process during surgery. (3) Based on the downstream tasks of semantic segmentation and physiological parameter estimation, we show that our recalibration method not only outperforms previous methods, but also generalizes across species, lighting conditions, and image processing tasks.

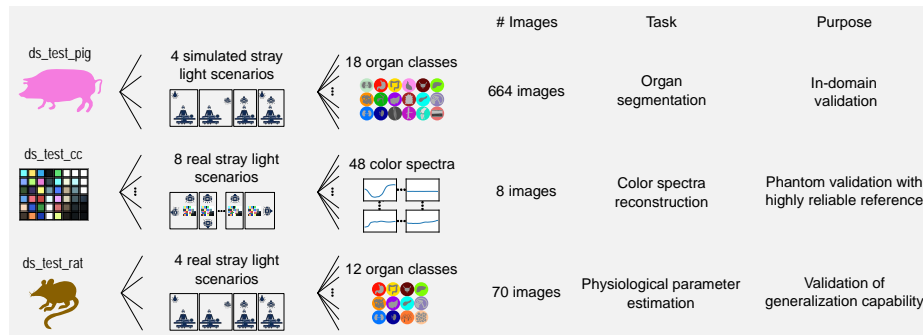


Fig. 3. Testing concept based on data from three species.

2 Materials and Methods

The main issue in data-driven calibration is the generalization to unseen settings. To make our method conceptually robust to domain shifts, we propose estimating the white tile measurement that we would obtain for a given scene rather than directly predicting the recalibrated image. Our approach is based on the hypothesis that capturing a representative set of tissue configurations, including illumination conditions, as a training set is infeasible. We therefore disentangle the space of possible illuminations from the space of possible tissue configurations, as illustrated in Fig. 2. More specifically, we employ a two-dataset training paradigm for the neural network. The first dataset, the illumination dataset, comprises real and simulated white reference images encompassing a wide range of illumination conditions encountered within the OR. The second dataset consists of accurately calibrated HSI cubes of clinically relevant samples. To simulate uncalibrated HSI cubes, each image in the sample dataset is augmented by multiplication with the associated white reference image. Subsequently, the neural network is trained to retrieve the white reference image from the simulated uncalibrated HSI cube. During inference, an uncalibrated HSI cube, possibly acquired with stray light, is fed into the neural network to predict the white reference image needed for illumination calibration.

2.1 Datasets

Training and validation was performed exclusively on porcine data, while testing was performed for unseen stray light scenarios on unseen porcine individuals, a phantom and rats, as summarized in Fig. 3. While the phantom colorchecker board dataset was acquired with the Tivita[®] 2.0 Surgery (Diaspective Vision GmbH, Am Salzhaff, Germany) featuring light-emitting diode (LED) illumination, the others were captured with the halogen-based Tivita[®] Tissue. As these light sources exhibit different behavior when interfering with the main stray light source, namely LED-based OR lights, we validated on both systems.

Measured illumination dataset The purpose of this dataset was to capture a variety of representative OR lighting conditions for algorithm training.

To this end, we acquired white reference images in an OR to capture camera light and additional stray light sources such as surgical lights, ceiling lights, or daylight. The surgical lights used in our study were manufactured by Dr. Mach and are composed of LEDs (Model: LED 8 MC). Diverse stray light scenarios were achieved by varying the angle, distance, and number of surgical lights as well as adjusting blinds or ceiling light, resulting in a wide range of illumination spectra. As the two HSI systems used in this study differ in the light sources, we acquired one illumination dataset with each camera: *ds_dev_ill_led* (LED) and *ds_dev_ill_hal* (halogen).

In vivo porcine development dataset: For model development, we curated a subset of the publicly available HSI dataset HeiPorSPECTRAL [24], *ds_test_pig*, consisting of accurately calibrated hyperspectral images of surgical scenes semantically annotated with 18 organ classes.

Test datasets: Comprehensive validation of our methodology was performed based on the three datasets *ds_test_pig*, *ds_test_cc* and *ds_test_rat* summarized in Fig. 3. In-domain testing of calibration quality was performed with data resembling the training data. To simulate stray light in *ds_test_pig*, we acquired a white reference test set *ds_test_ill_hal* with four stray light scenarios. Colorchecker boards imaged under various lighting conditions were used for the assessment of recalibration performance based on highly reliable reference data (*ds_test_cc*). The effect of the recalibration method on surgical image analysis was assessed by means of the downstream tasks organ segmentation and physiological parameter analysis using in-domain and out of domain in vivo hyperspectral imaging data from porcine models (*ds_test_pig*) and rats (*ds_test_rat*).

2.2 Physics-based illumination simulation

To implement the data-centric recalibration concept, we focused on the model-based generation of plausible white tile data. To overcome the resource-intensive white reference acquisitions, we enhanced *ds_dev_ill_led* and *ds_dev_ill_hal* by synthesizing white tile images based on real L1-normalized white tile images.

LED simulations: Surgical lights are the main source of stray light in the OR. For our LED-based surgical lights, wave interference with an LED-based HSI system is approximately constructive, thus local extrema in the spectrum of the camera light source are preserved. This behavior can be modeled by linear inter- and extrapolations of white reference images. To avoid the generation of duplicates, we first conduct clustering, before combining images of distinct clusters.

Halogen simulations: In contrast to LED spectra, halogen spectra differ in terms of width and the number of local extrema. Consequently, the interference with LED light does not preserve local extrema. To obtain powerful simulations of stray light-affected halogen spectra, we propose to mathematically model the curves. Inspired by Planck’s radiation law, we empirically observed that the following parametric function describes stray light-affected halogen spectra:

$$f_{a,b,c,d}(\lambda) = \frac{(a\lambda - b)^3}{\exp(c\lambda - d) - 1}$$

where f denotes the intensity, λ the wavelength and a, b, c, d the parameters. Least-square optimization to the mean spectra of *ds_dev_ill_hal* yielded parameter ranges for a,b,c,d so that $f_{a,b,c,d}(\lambda)$ adequately models the illumination conditions captured in *ds_dev_ill_hal*. Increasing the upper bounds of these ranges led to halogen spectra with higher levels of stray light. To synthesize a hyperspectral image that features realistic spatial variations of intensity from the simulated light spectrum $f_{a,b,c,d}(\lambda)$, we leveraged the acquired images. More concretely, a hyperspectral image $I(i, j, \lambda)$ is randomly selected from *ds_dev_ill_hal* and divided by its spatially averaged spectrum $\bar{I}(\lambda)$. By multiplication with the simulated spectrum $f_{a,b,c,d}(\lambda)$, we obtain a simulated white reference image $I_s(i, j, \lambda)$ with the simulated spectrum $f_{a,b,c,d}(\lambda)$ as mean spectrum and the spatial variations from I .

$$I_s(i, j, \lambda) = f_{a,b,c,d}(\lambda) \odot I(i, j, \lambda) \oslash \bar{I}(\lambda)$$

To further enhance the coverage of illumination conditions, inter- and extrapolation was performed as for the LED simulations. Overall, this process yielded a set of about 200 different illumination conditions for each light source.

2.3 Neural network implementation details

We feed the HSI cubes into a 3D CNN that employs an autoencoder architecture, utilizing ResNet blocks [12] in both the encoder and decoder. Two design decisions were particularly important for our method’s success: During training, we only optimize the predicted white reference image and not the resulting calibrated sample image. Furthermore, we omit skip connections between the encoder and decoder. Both design choices aim to prevent the model from relying heavily on the content of the sample images, instead focusing on learning the illumination information. As loss function, we employ the MSE-reconstruction loss between the predicted and original white reference image. Further implementation details are provided in Suppl. Tab. 1.

3 Experiments and Results

We investigated the following research questions (RQs):

- (RQ1) How do dynamically changing lighting conditions in the OR affect the performance of hyperspectral image analysis algorithms?
- (RQ2) Are neural networks capable of replacing white tile recalibration of hyperspectral cameras in the OR?
- (RQ3) To what extent can neural network-based recalibration mitigate the performance drop of hyperspectral image analysis algorithms under varying lighting conditions?

For all experiments involving our method, we used the same model trained exclusively on porcine data (*ds_dev_pig*). The generalization to unseen domains

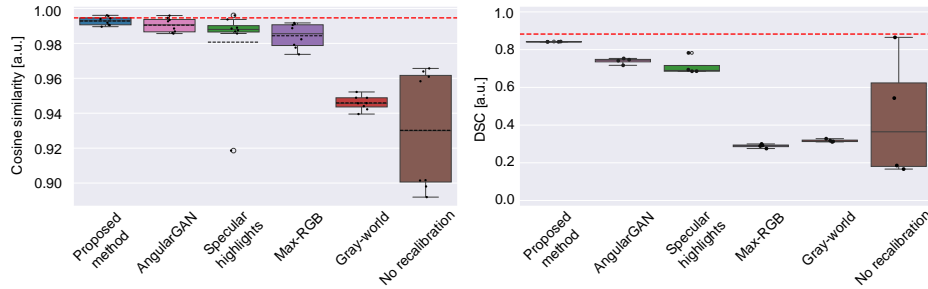


Fig. 4. State-of-the-art methods fail under dynamically changing light conditions. Our approach addresses this issue. (Left) Results on colorchecker dataset *ds_test_cc*. The boxplots show the cosine similarity between recalibrated and reference spectra, averaged across colors. Red line: Gold standard of manual white tile calibration. (Right) Results on semantic segmentation dataset *ds_test_pig*. Red line: Mean DSC in the absence of stray light. Points: Different stray light scenarios.

(here: colorchecker boards and rats) was investigated on untouched test sets.

Experiment RQ1: We used the traditional approaches Gray-world [5], Max-RGB [16], and a method based on specular highlights [2], as baselines. Additionally, we integrated a learning-based method by adapting the RGB-calibration framework AngularGAN [23] to hyperspectral imaging. To this end, we trained AngularGAN on *ds_dev_pig* augmented by the originally acquired white reference images. As downstream tasks, we conducted semantic organ segmentation and physiological parameter estimation on in vivo data.

For the semantic organ segmentation task, we leveraged the accurately calibrated *ds_test_pig* and illumination test set *ds_test_ill_hal* to obtain four stray light-affected versions of each image in *ds_test_pig*. Subsequently, the resulting 664 images were recalibrated by one of the methods, followed by the inference of segmentation masks using a public segmentation model trained on calibrated pig organ images [21]. As segmentation metrics, the Dice similarity coefficient (DSC) and the normalized surface distance (NSD) were used, as recommended by [18]. For physiological parameter estimation, recalibration procedures were applied to *ds_test_rat*, followed by computation of the oxygen saturation, perfusion, hemoglobin, and water index [15]. To gauge calibration performance, mean absolute errors were calculated between the stray light-affected parameters and reference values derived from images devoid of stray light interference. For both downstream tasks, the hierarchical structure of the data was respected during aggregation.

Fig. 4 and Suppl. Fig. 1 show that existing HSI calibration techniques lack adequate accuracy. Even the best performing methods (Specular highlights and AngularGAN) come with a decrease of the DSC of more than 15%.

Similarly (Fig. 5), the tissue parameter maps change substantially under varying lighting conditions, even when manual white tile recalibration was performed (cf. Suppl. Fig. 2).

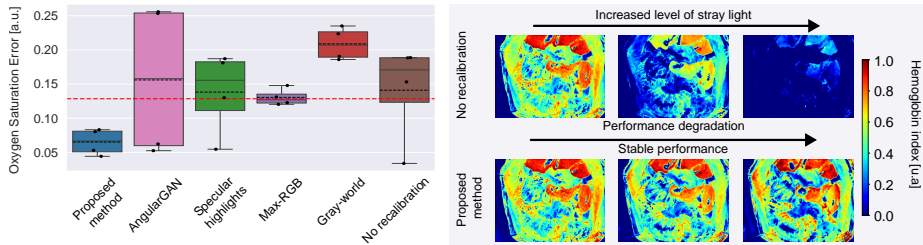


Fig. 5. In contrast to related methods, our approach generalizes across species. (Left) Organ-specific absolute oxygen saturation errors between calibrated rat images without stray light and corresponding stray light images that are recalibrated by one of the methods. Red line: Mean performance of the gold standard (manual white tile calibration). (Right) Our method yields precise hemoglobin index estimates under dynamically changing lighting conditions.

Experiment RQ2 To measure the calibration accuracy in an OOD setting with a highly reliable reference, we applied our model trained on porcine data to recalibrate *ds_test_cc*. As illustrated in Fig. 4, our method demonstrates the highest average cosine similarity. They are almost on par with the gold standard of manually acquiring white tile measurements.

Experiment RQ3 To assess the capability of our method to boost the downstream task performance, we performed Experiment RQ1 on our recalibration approach. As shown in Figs. 4 and 5, our method outperforms previous methods by a large margin. For semantic segmentation, relative improvements of the DSC from 14 % to 191 % were obtained. For oxygen saturation estimation, the error could be reduced by 50 % to 69 %. Similar performance gains were obtained for other tissue parameters (cf. Suppl. Fig. 2). A qualitative assessment of the high fidelity of our recalibrated tissue spectra is available in Suppl. Fig. 3.

4 Discussion

We were the first to provide in vivo evidence that dynamically changing lighting conditions in the OR can cause dramatic failures in HSI downstream analysis such as semantic segmentation. This is a finding of high clinical relevance because the manual white tile-based recalibration of cameras during surgery severely disrupts the clinical workflow and may currently hinder widespread clinical adoption of HSI cameras. The proposed method represents the only calibration model in our analysis that is capable of maintaining high accuracy independently of the downstream task and domain, indicating high applicability for clinical use cases. It also features several major conceptual advantages: White reference measurements are not only impractical as they suffer from sterilization and workflow issues, but are also prone to oversaturation. This explains the suboptimal performance on the rat data (see Fig. 5). Specular highlights and Max-RGB exhibit high calibration accuracy on the colorchecker board, as both methods calibrate

the images with the white color field by design, but fail to generalize to in vivo scenarios. Note that these methods recover a global illumination estimate, which is not sufficient in the case of spatially heterogeneous illumination encountered in the OR. In fact, we also saw drops in performance when reducing the estimations of white tile calibration (classic and data-driven) to a global estimate. Overall, the core strength of our approach is its generalizability. Notably, it outperforms the competing neural network method AngularGAN even when trained on the exact same data as our method. We attribute this to the inherent domain shift of auto-encoded images.

A limitation of our work could be seen in the fact that we did not cover all possible illumination settings that can occur in practice. However, as we focused on the most important light sources (surgical lights and ceiling light) and conducted our validation on highly diverse datasets, we are confident that our conclusions will hold in diverse settings.

In conclusion, our work presents a novel learning-based light calibration method for hyperspectral imaging. The proposed methodology not only outperforms previously proposed approaches in various settings but can be seamlessly incorporated into hyperspectral imaging systems for ORs. Our work could therefore pave the way for clinical workflow-optimized and robust HSI in surgery.

Acknowledgments. This project was supported by the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (NEURAL SPICING, 101002198), the National Center for Tumor Diseases (NCT), Heidelberg’s Surgical Oncology Program, the German Cancer Research Center (DKFZ), and the Helmholtz Association under the joint research school HIDSS4Health (Helmholtz Information and Data Science School for Health). We also acknowledge the support through state funds for the Innovation Campus Health + Life Science Alliance Heidelberg Mannheim from the structured postdoc program for Alexander Studier-Fischer: Artificial Intelligence in Health (AIH).

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Ayala, L., Adler, T.J., Seidlitz, S., Wirkert, S., Engels, C., Seitel, A., Sellner, J., Aksenov, A., Bodenbach, M., Bader, P., et al.: Spectral imaging enables contrast agent-free real-time ischemia monitoring in laparoscopic surgery. *Science advances* **9**(10), eadd6778 (2023)
2. Ayala, L., Seidlitz, S., Vemuri, A., Wirkert, S.J., Kirchner, T., Adler, T.J., Engels, C., Teber, D., Maier-Hein, L.: Light source calibration for multispectral imaging in surgery. *International Journal of Computer Assisted Radiology and Surgery* **15**, 1117–1125 (2020)
3. Ayala, L.A., Wirkert, S.J., Gröhl, J., Herrera, M.A., Hernandez-Aguilera, A., Vemuri, A., Santos, E., Maier-Hein, L.: Live monitoring of haemodynamic changes with multispectral image analysis. In: *OR 2.0 Context-Aware Operating Theaters and Machine Learning in Clinical Neuroimaging: Second International Workshop*,

- OR 2.0 2019, and Second International Workshop, MLCN 2019, Held in Conjunction with MICCAI 2019, Shenzhen, China, October 13 and 17, 2019, Proceedings 2. pp. 38–46. Springer (2019)
4. Bahl, A., Horgan, C.C., Janatka, M., MacCormac, O.J., Noonan, P., Xie, Y., Qiu, J., Cavalcanti, N., Fürnstahl, P., Ebner, M., et al.: Synthetic white balancing for intra-operative hyperspectral imaging. *Journal of Medical Imaging* **10**(4), 046001–046001 (2023)
 5. Buchsbaum, G.: A spatial processor model for object colour perception. *Journal of the Franklin institute* **310**(1), 1–26 (1980)
 6. Clancy, N.T., Jones, G., Maier-Hein, L., Elson, D.S., Stoyanov, D.: Surgical spectral imaging. *Medical image analysis* **63**, 101699 (2020)
 7. Das, P., Liu, Y., Karaoglu, S., Gevers, T.: Generative models for multi-illumination color constancy. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*. pp. 1194–1203 (2021)
 8. Ebner, M., Nabavi, E., Shapey, J., Xie, Y., Liebmann, F., Spirig, J.M., Hoch, A., Farshad, M., Saeed, S.R., Bradford, R., et al.: Intraoperative hyperspectral label-free imaging: from system design to first-in-patient translation. *Journal of Physics D: Applied Physics* **54**(29), 294003 (2021)
 9. Gao, S.B., Ren, Y.Z., Zhang, M., Li, Y.J.: Combining bottom-up and top-down visual mechanisms for color constancy under varying illumination. *IEEE Transactions on Image Processing* **28**(9), 4387–4400 (2019)
 10. Halicek, M., Fabelo, H., Ortega, S., Callico, G.M., Fei, B.: In-vivo and ex-vivo tissue analysis through hyperspectral imaging techniques: revealing the invisible features of cancer. *Cancers* **11**(6), 756 (2019)
 11. Halicek, M., Lu, G., Little, J.V., Wang, X., Patel, M., Griffith, C.C., El-Deiry, M.W., Chen, A.Y., Fei, B.: Deep convolutional neural networks for classifying head and neck cancer using hyperspectral imaging. *Journal of biomedical optics* **22**(6), 060503–060503 (2017)
 12. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. pp. 770–778 (2016)
 13. Holmer, A., Tetschke, F., Marotz, J., Malberg, H., Markgraf, W., Thiele, C., Kulcke, A.: Oxygenation and perfusion monitoring with a hyperspectral camera system for chemical based tissue analysis of skin and organs. *Physiological measurement* **37**(11), 2064 (2016)
 14. Hussain, M.A., Akbari, A.S.: Color constancy algorithm for mixed-illuminant scene images. *IEEE Access* **6**, 8964–8976 (2018)
 15. Kulcke, A., Holmer, A., Wahl, P., Siemers, F., Wild, T., Daeschlein, G.: A compact hyperspectral camera for measurement of perfusion parameters in medicine. *Biomedical Engineering/Biomedizinische Technik* **63**(5), 519–527 (2018)
 16. Land, E.H.: The retinex theory of color vision. *Scientific american* **237**(6), 108–129 (1977)
 17. Li, Y., Fu, Q., Heidrich, W.: Multispectral illumination estimation using deep unrolling network. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*. pp. 2672–2681 (2021)
 18. Maier-Hein, L., Reinke, A., Godau, P., Tizabi, M.D., Buettner, F., Christodoulou, E., Glocker, B., Isensee, F., Kleesiek, J., Kozubek, M., et al.: Metrics reloaded: recommendations for image analysis validation. *Nature methods* pp. 1–18 (2024)
 19. Mutimbu, L., Robles-Kelly, A.: Multiple illuminant color estimation via statistical inference on factor graphs. *IEEE Transactions on Image Processing* **25**(11), 5383–5396 (2016)

20. Seidlitz, S., Sellner, J., Odenthal, J., Özdemir, B., Studier-Fischer, A., Knödler, S., Ayala, L., Adler, T.J., Kenngott, H.G., Tizabi, M., et al.: Robust deep learning-based semantic organ segmentation in hyperspectral images. *Medical Image Analysis* **80**, 102488 (2022)
21. Sellner, J., Seidlitz, S., Studier-Fischer, A., Motta, A., Özdemir, B., Müller-Stich, B.P., Nickel, F., Maier-Hein, L.: Semantic segmentation of surgical hyperspectral images under geometric domain shifts. arXiv preprint arXiv:2303.10972 (2023)
22. Shapey, J., Xie, Y., Nabavi, E., Bradford, R., Saeed, S.R., Ourselin, S., Vercauteren, T.: Intraoperative multispectral and hyperspectral label-free imaging: A systematic review of in vivo clinical studies. *Journal of biophotonics* **12**(9), e201800455 (2019)
23. Sidorov, O.: Conditional gans for multi-illuminant color constancy: Revolution or yet another approach? In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops*. pp. 0–0 (2019)
24. Studier-Fischer, A., Seidlitz, S., Sellner, J., Bressan, M., Özdemir, B., Ayala, L., Odenthal, J., Knoedler, S., Kowalewski, K.F., Haney, C.M., et al.: Heiporspectral-the heidelberg porcine hyperspectral imaging dataset of 20 physiological organs. *Scientific Data* **10**(1), 414 (2023)
25. Trajanovski, S., Shan, C., Weijtmans, P.J., de Koning, S.G.B., Ruers, T.J.: Tongue tumor detection in hyperspectral images using deep learning semantic segmentation. *IEEE transactions on biomedical engineering* **68**(4), 1330–1340 (2020)
26. Wirkert, S.J., Vemuri, A.S., Kenngott, H.G., Moccia, S., Götz, M., Mayer, B.F., Maier-Hein, K.H., Elson, D.S., Maier-Hein, L.: Physiological parameter estimation from multispectral images unleashed. In: *Medical Image Computing and Computer Assisted Intervention- MICCAI 2017: 20th International Conference, Quebec City, QC, Canada, September 11-13, 2017, Proceedings, Part III* 20. pp. 134–141. Springer (2017)