

Hyperspectral Imaging System for Tissue Classification in H&E-Stained Histological Slides

Mateus Martelini Souza

Dept. of Electrical and Computing Eng.

University of São Paulo

São Carlos, Brazil

mateus.martelini.souza@usp.br

Felipe Alvarenga Carvalho

Dept. of Electrical and Computing Eng.

University of São Paulo

São Carlos, Brazil

carvalhofelipe.carvalho@usp.br

Enzo Fabro Vanzela Sverzut

Dept. of Electrical and Computing Eng.

University of São Paulo

São Carlos, Brazil

email address or ORCID

Michelle Barreto Requena

São Carlos Institute of Physics

University of São Paulo

São Carlos, Brazil

requenamichelle@gmail.com

Marlon Rodrigues Garcia

São Carlos Institute of Physics

University of São Paulo

São Carlos, Brazil

marlongarcia@usp.br

Sebastião Pratavieira

São Carlos Institute of Physics

University of São Paulo

São Carlos, Brazil

prata@ifsc.usp.br

Abstract—This paper presents the development of a hyperspectral imaging system for the classification of H&E-stained histological slides. The system was developed to be coupled to a conventional microscope, with software dedicated to control the instrumentation, to show a colorful live image from an RGB camera, and to acquire the hyperspectral imaging using a liquid crystal tunable filter (LCTF). Hyperspectral images of H&E-stained histological slides undergoing photodynamic therapy were classified with four machine learning algorithms to find damaged tissues (crust). The classification results were presented and show that this technique is promising to classify histological tissue regions.

Index Terms—Hyperspectral Imaging, Histological Slides, Photodynamic Therapy, Animal Model, Tissue Classification

I. INTRODUCTION

Hyperspectral imaging (HSI) is a technique to image an object in different wavelengths in order to acquire the spectral information for each pixel. This technique has been extensively used in digital pathology, tissue conditions diagnosis, and for surgery visualization [1–3]. In digital pathology, interesting works have been published in the literature for the detection of neoplastic tissues using HSI, like squamous neoplasia [4], breast cancer [5], colon ovarian cancer [6], and melanoma [7]. Also some studies have also addressed the digital staining of histology slides using hyperspectral imaging [8, 9].

Although there are important works studying skin tissue classification, there is a considerable gap in the published studies for tissue classification in skin histology. This work is dedicated to study the use of different artificial intelligence algorithms in order to classify different tissue regions in H&E stained slides. An interesting tissue area analysed here was the

The authors acknowledge the support provided by Brazilian Funding Agencies: Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001; CNPq (465360/2014-9 and 306919/2019-2) and São Paulo Research Foundation (FAPESP) grants: 2013/07276-1 (CEPOF); 2014/50857-8 (INCT)

(978-0-7381-2418-6/21/\$31.00 ©2021 IEEE)

skin damaged tissue (crust) which is of important relevance in the research area of skin cancer therapies like photodynamic therapy (PDT).

II. MATERIAL AND METHOD

A. Hardware and Software

The hyperspectral system was developed to connect to most of the commercial microscopes available. In this study, the equipment was installed in the trinocular output of an Eclipse Ti-S (Nikon, Japan) but can also be established in the camera output or even in the ocular outcome. The Eclipse Ti-S has a 30 W halogen lamp to illuminate the sample, and the transmitted light is then collimated by the microscope optics and enters the hyperspectral system. The developed hyperspectral system comprises a liquid crystal tunable filter (LCTF, PerkinElmer, USA), which is the spectral component, and two digital cameras, an RGB (DCC-1645C, Thorlabs, USA), to preserve the regular usage of a microscope to find the sample and to focus its objective lenses, and a monochromatic one (DCC-1545M, Thorlabs, USA). Figure 1 presents the system schematic drawing.

The LCTF can filter light in different wavelengths from the visible electromagnetic spectrum to the near-infrared (400 nm to 720 nm) based on light polarization. Changing the equipment's input voltage causes the concatenated liquid crystals to modify their refractive index. These changes can filter light in a specific spectral region, acting as a band-pass filter with a well-defined central wavelength. This filtering defines the spectral range that can pass through LCTF and reaches the monochromatic camera sensor.

The control software was developed and connected to the hyperspectral system to capture the images, control both cameras and change the filter's wavelengths. Both cameras are fully controllable, including the pixel clock, frame rate, exposure time, and the capability of displaying real-time images to the user.

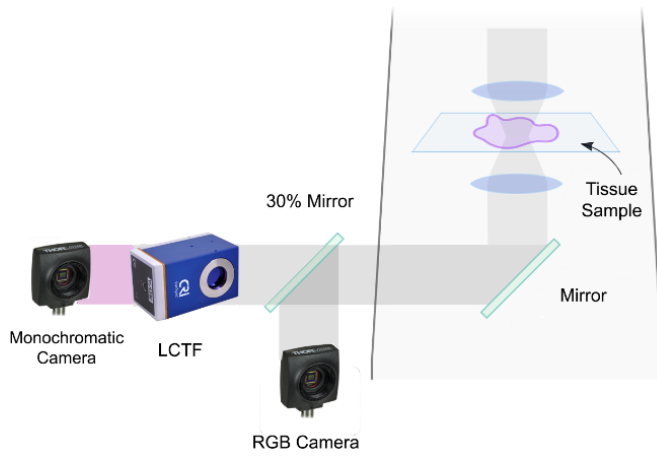


Fig. 1. Schematic drawing of the hyperspectral system (left) and the optical microscope (right). The hyperspectral system is composed by two digital cameras (monochromatic and RGB), the spectral component (LCTF) and a 30% mirror.

While the monochromatic camera function is to capture the hyperspectral images that pass through the filter, the RGB camera is responsible for displaying the colored images in real-time during the process, and being also used to capture images for further processing and comparison with the hyperspectral images.

The control algorithm and the graphical user interface (GUI) were developed on LabVIEW (National Instruments, USA) platform, considering its powerful capabilities to connect different hardware and its fast time-to-development of a complex GUI. The user interface is composed of two picture boxes that display the real-time images from both cameras, in which their parameters can be controlled and changed during the algorithm's execution. Both cameras have the possibility to save images on a chosen folder.

In the automated routine, both LCTF and the monochromatic camera are controlled to acquire specific wavelengths with intervals and steps that can be determined by the user. During this process, the user can choose whether the camera exposure time will be maintained or if it will change the wavelength accordingly to take into account the small LCTF transmittance and the small lamp emission towards smaller wavelengths. All these images and exposure times are recorded.

B. Animal Model

This study had approval from the Animal Use Ethics Committee of the São Carlos Institute of Physics at the University of São Paulo. The experiments were carried out in a male Wistar rat with 300 g on average. Photodynamic therapy (PDT) was performed in the animal dorsum using a commercial system (Lince, MMOptics, Brazil) at an irradiance of 40 mW/cm² until a fluence of 30 J/cm². The irradiance protocol was performed after 2 h of δ -aminolevulinic acid (ALA) incubation in rat skin. The histological slide was

obtained from biopsies collected 72 h after PDT, stained with hematoxylin and eosin (H&E).

C. Image Processing Algorithms

A Python algorithm was developed using the Spyder IDE to process the images and use four different supervised machine learning models. The first model was the K-Nearest Neighbors, which is the simplest and can be used for comparison with the others. The second and third models were the Support Vector Machine (SVM) with linear kernel and RBF kernel, respectively. SVM is a widely used algorithm in hyperspectral image classification and performs very well in this technique [1–3]. The fourth model was Random Forest (RF), which is also commonly used and has a low risk of overfitting [1, 3]. The scikit-learn library was used because it has all the algorithms implemented and it is open source [10]. The code used the classes and functions defined in the library for the algorithms described above, and part of the algorithm was related to image acquisition, preparation, and data and visualization were written by the author.

The monochromatic images ($w \times h$ pixels) acquired by the system described above were loaded and grouped in order to create a hyperspectral image, with the third dimension size equals to the number of acquired wavelengths (d). An algorithm that uses some functions and classes of OpenCV library (Open Source Computer Vision Library) was written for selection of the regions of the image corresponding to each tissue: epidermis, dermis, and damaged tissue (crust), in addition to the glass region of the lamina in which there was no tissue. This procedure was performed because the algorithms are supervised and use this information for the training process and for the calculation of accuracy. Then, the image was normalized by the mean pixels intensity that corresponds to the glass region so that the transmittance of each tissue was obtained, considering the glass of the microscope slide as a reference. This was done to all components of the image, therefore each wavelength has been corrected, preventing from differences in the acquisition parameters from the camera. There was applied another normalization so that the pixels presented values between 0 and 1. This was done because some machine learning models perform better if the input data is in the range between 0 and 1.

First the image were reshaped from $h \times w \times d$ to $h * w \times d$. Thus, each region of interest (ROI) selected before was taken from the $h * w \times d$ array. From the selected data, 90% was applied as a training set and 10% as the test set. After data preparation, the training set was used as an input to the training of the four models. Then, the entire image was used as input for the algorithm to classify each tissue. The array resulted from classification was reshaped again to the dimensions $h \times w \times d$, and there was applied an artificial coloration to the image to identify each classified region.

Each machine learning model from the scikit-learn library has a built-in function that calculates the average accuracy. The test set was used as the input of this function to calculate this score. There was also a post-processing phase which consisted

of applying a median filter with kernel size equals 31. This kind of filter was used in order to increase the homogeneity of the tissue.

III. RESULTS

A. Hardware and Software

The image acquisition and GUI software were developed on LabVIEW, based on Varispec (PerkinElmer, USA) and Thorlab's pre-built functions (SubVI's). A .NET extension file was used to control both cameras, while the LCTF's serial communications were designed using the LabVIEW VISA application.

The software starts with the initialization of both cameras and their serial number loading. A serial identification function is then called to recognize the cameras, allocating their reference ID to its main loop. The two cameras main loops are responsible for displaying the real-time images and, in the case of monochromatic camera, conducting the camera and the LCTF to the hyperspectral imaging loop. Each camera loop is executed in parallel, reading the camera's sensors and displaying the image for the user.

When the user enters the wavelength interval, the incremental step and starts the automated image capture routine, the LCTF is activated, receiving the parameters as an array input and being able to filter light through in the wavelength interval. The monochromatic camera saves the filtered image automatically after every filter change while also updating its exposure time based on an array input.

B. Image Processing Algorithms

After data extraction, the images corresponding to 470 nm, 570 nm, and 630 nm were grouped to form the RGB image (figure 2).

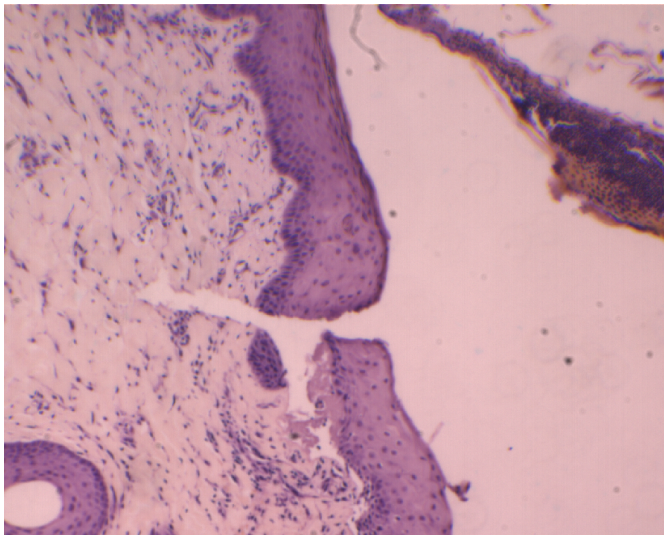


Fig. 2. RGB image of a H&E stained slide of a rat skin submitted to photodynamic therapy (PDT). It is possible to see the epidermis tissue in the middle of the image (purple), the dermis on the left (light purple), and the damaged tissue on the right (darker region). The image was composed with the hyperspectral system using the 470 nm, 570 nm, and 630 nm components.

For the training process, 0.51% of the pixels of the image was used, which corresponded to 6701 pixels. After processing and classifying the hyperspectral image, the Figure 3 was generated, which presents the results of each of the four models. The colors represent the algorithm classifications: dark red for the epidermis, green indicates the dermis classification, blue is the damaged tissue (crust), and yellow the slide glass region.

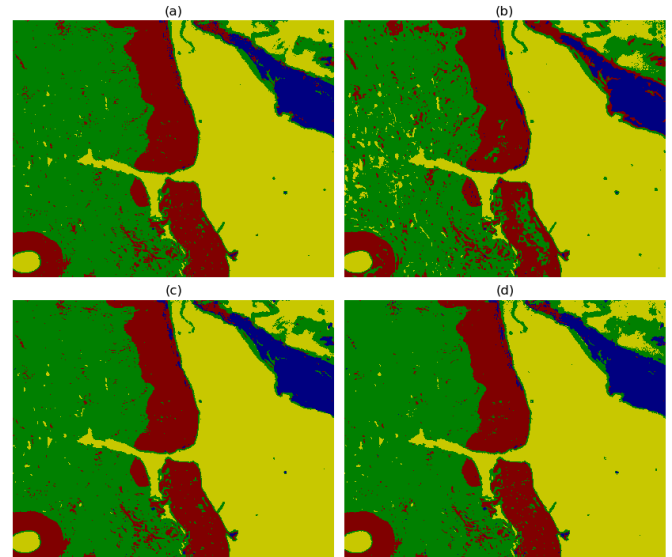


Fig. 3. Resulting image after classification by each model: (a) KNN (b) Linear SVM (c) RBF SVM (d) Random Forest. In the image, each color represents an algorithm classification: green for dermis, dark red for epidermis, yellow for blank region (slide glass), and blue for the damaged region (crust).

Then the test set was used as the input for the function that calculates the average accuracy. The results were generated for each model (Table I).

TABLE I
ACCURACY FOR EACH OF THE MODELS FOR THE TEST SET

Score	KNN	Linear SVM	RBF SVM	RF
Accuracy (%)	96.31	91.79	96.55	96.64

Figure 4 shows the resulting image after applying the median filter, and Table II presents the corresponding score for each model after this post-processing.

TABLE II
ACCURACY FOR EACH OF THE MODELS FOR THE TEST SET AFTER MEDIAN FILTER APPLICATION

Score	KNN	Linear SVM	RBF SVM	RF
Accuracy (%)	98.68	96.37	98.84	98.93

The results indicate that these four models are prominent for classifications of histological slides of H&E-stained tissues. The post-processing increased the accuracy and maybe a good technique to correct the pixels classified incorrectly due to the cell's nuclei.

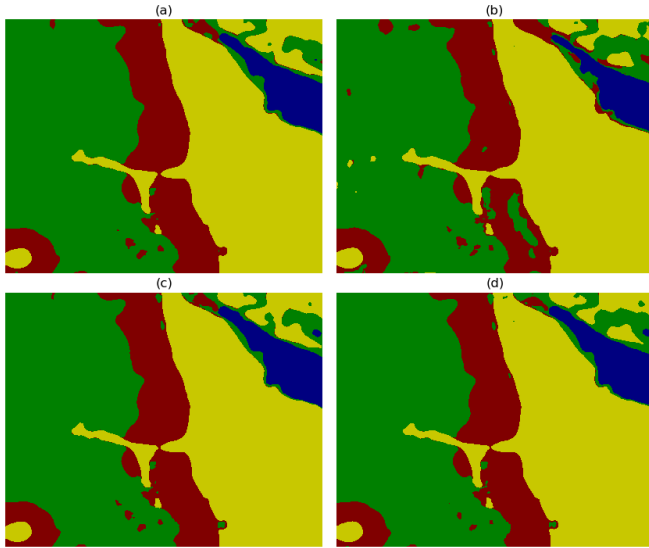


Fig. 4. Resulting classified image after median filter with 31 pixels of kernel size, for each method applied: (a) KNN (b) Linear SVM (c) RBF SVM (d) Random Forest. In the image, each color represents an algorithm classification: green for dermis, dark red for epidermis, yellow for blank region (slide glass), and blue for the damaged region (crust).

IV. CONCLUSIONS

In this study, a hyperspectral system is described and used to image the H&E-stained slide of a rat skin submitted to ALA-mediated photodynamic therapy. Four different algorithms were implemented in this study, which results show that this techniques are promising in the classification of rat skin regions. In particular, this technique is promising for the classification of damaged tissue (crust), which is highly important to quantify the necrotic area (or damaged tissue) in treatment studies like photodynamic therapy.

REFERENCES

- [1] Samuel Ortega, Martin Halicek, Himar Fabelo, Gustavo M Callico, and Baowei Fei. Hyperspectral and multispectral imaging in digital and computational pathology: a systematic review. *Biomedical Optics Express*, 11(6):3195–3233, 2020.
- [2] Guolan Lu and Baowei Fei. Medical hyperspectral imaging: a review. *Journal of biomedical optics*, 19(1): 010901, 2014.
- [3] M Hermes, R Brandstrup Morrish, Laurent Huot, Lichun Meng, Saher Junaid, J Tomko, GR Lloyd, WT Masselink, Peter Tidemand-Lichtenberg, Christian Pedersen, et al. Mid-ir hyperspectral imaging for label-free histopathology and cytology. *Journal of Optics*, 20(2):023002, 2018.
- [4] Guolan Lu, Dongsheng Wang, Xulei Qin, Susan Muller, James V Little, Xu Wang, Amy Y Chen, Georgia Chen, and Baowei Fei. Histopathology feature mining and association with hyperspectral imaging for the detection of squamous neoplasia. *Scientific reports*, 9(1):1–13, 2019.
- [5] Martin Halicek, Samuel Ortega, Himar Fabelo, Carlos Lopez, Marylene Lejeune, Gustavo M Callico, and Baowei Fei. Conditional generative adversarial network for synthesizing hyperspectral images of breast cancer cells from digitized histology. In *Medical Imaging 2020: Digital Pathology*, volume 11320, page 113200U. International Society for Optics and Photonics, 2020.
- [6] Daiki Nakaya, Ayaka Tsutsumiuchi, Shin Satori, Makoto Saegusa, Tsutomu Yoshida, Ako Yokoi, and Masaki Kano. Digital pathology with hyperspectral imaging for colon and ovarian cancer. In *Medical Imaging 2019: Digital Pathology*, volume 10956, page 109560X. International Society for Optics and Photonics, 2019.
- [7] David T Dicker, Jeremy Lerner, Pat Van Belle, DuPont Guerry, 4th, Meenhard Herlyn, David E Elder, and Wafik S El-Deiry. Differentiation of normal skin and melanoma using high resolution hyperspectral imaging. *Cancer biology & therapy*, 5(8):1033–1038, 2006.
- [8] Neslihan Bayramoglu, Mika Kaakinen, Lauri Eklund, and Janne Heikkila. Towards virtual h&e staining of hyperspectral lung histology images using conditional generative adversarial networks. In *Proceedings of the IEEE International Conference on Computer Vision Workshops*, pages 64–71, 2017.
- [9] Pinky A Bautista and Yukako Yagi. Digital staining for histopathology multispectral images by the combined application of spectral enhancement and spectral transformation. In *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 8013–8016. IEEE, 2011.
- [10] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.