

Marianne Maktabi*, Yannis Wichmann, Hannes Köhler, Claire Chalopin, Thomas Neumuth, Boris Jansen-Winkel, Henning Ahle, Dietmar Lorenz, Michael Bange, Susanne Braun, Ines Gockel and René Thieme

Semi-automatic decision-making process in histopathological specimens from Barrett's carcinoma patients using hyperspectral imaging (HSI)

Abstract: Discrimination of malignant and non-malignant cells of histopathologic specimens is a key step in cancer diagnostics. Hyperspectral Imaging (HSI) allows the acquisition of spectra in the visual and near-infrared range (500-1000nm). HSI can support the identification and classification of cancer cells using machine learning algorithms. In this work, we tested four classification methods on histopathological slides of esophageal adenocarcinoma. The best results were achieved with a Multi-Layer Perceptron. Sensitivity and F1-Score values of 90% were obtained.

Keywords: Hyperspectral Imaging, Barrett's carcinoma, histopathological slides, machine learning

<https://doi.org/10.1515/cdbme-2020-3066>

1 Introduction

Discrimination of malignant and non-malignant cells of histopathologic specimens is a key step in cancer diagnostics.

***Corresponding author: Marianne Maktabi:** Innovation Center Computer Assisted Surgery (ICCAS) University Leipzig, Leipzig Germany, Marianne.Maktabi@medizin.uni-leipzig.de

Hannes Köhler, Claire Chalopin, Thomas Neumuth: Innovation Center Computer Assisted Surgery (ICCAS) University Leipzig, Leipzig Germany,

Yannis Wichmann, Boris Jansen-Winkel, , Ines Gockel and René Thieme: University Hospital Leipzig, Leipzig Germany

Henning Ahle: Department of General and Visceral Surgery, Sana Clinic Offenbach GmbH, Offenbach, Germany

Dietmar Lorenz: Department of General, Visceral and Thoracic Surgery, Municipal Hospital of Darmstadt GmbH, Darmstadt, Germany

Michael Bange, Susanne Braun: Institute of Pathology, Sana Clinic Offenbach GmbH, Offenbach, Germany

Hyperspectral imaging (HSI), as recently applied in medicine, is a novel technology combining imaging with spectroscopy. The light reflected from the target object is analysed and provides information about the pathology of this object. HSI allows the acquisition of a spectrum in the visual and near-infrared light range (e.g. 500-1000nm). Based on hypercubes which consists of a three-dimensional dataset of spatial and spectral information it is possible to evaluate the pathology for each pixel of the target object [1]. Moreover, machine learning algorithms can support the identification and classification of cancer cells in the HSI data. This was shown in several in-vivo and ex-vivo studies [2]–[5]. In [6], [7] algorithms to distinguish cancerous and non-cancerous tissue of gastric and pancreatic tumors were applied on stained pathological slides.

In this work we classified four different structures (stroma, background, squamous epithelium and esophageal adenocarcinoma (EAC)) based on four machine-learning methods (e.g. Support Vector Machine (SVM)). Therefore, we

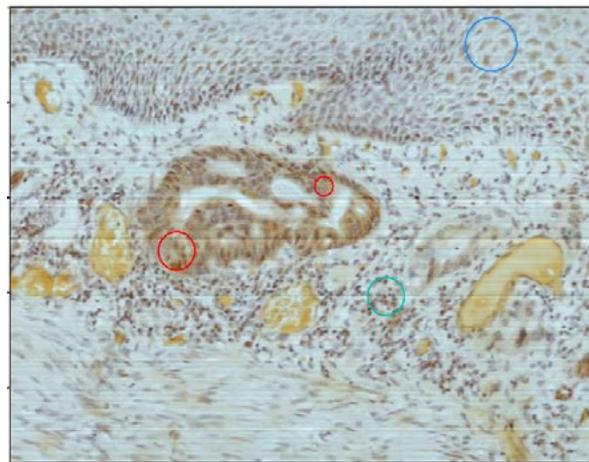


Figure 1: Annotation of the different classes in stained slides done by a medical expert (red circle: EAC, blue circle: stroma, cyan circle: squamous epithelium).

used HSI recordings of histopathological specimens from Barrett's carcinoma patients.

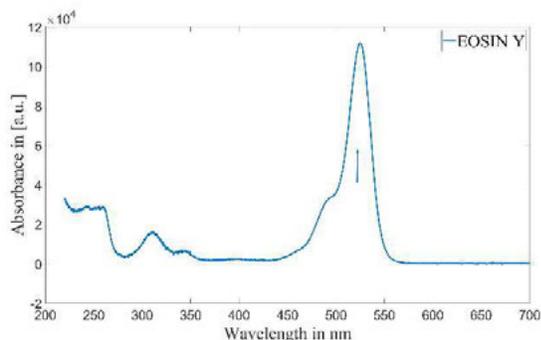


Figure 2: Absorbance spectrum of eosin Y (eosin peak are marked as blue arrow)(www.omlc.org).

2 Methods

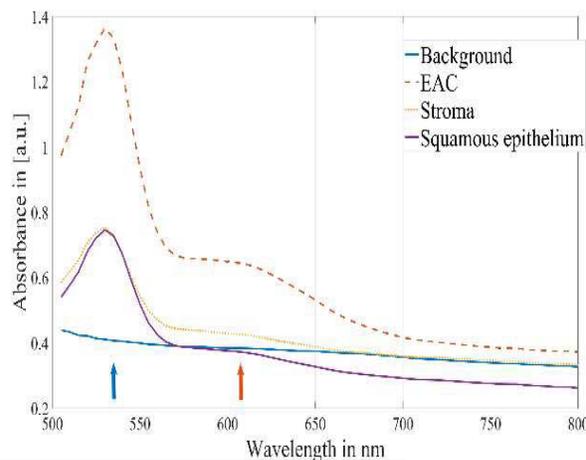
After surgical resection, specimens (n=56) of Barrett's cancer tumor samples were fixed in 4% formaldehyde and sliced (3µm). The slides were stained with hematoxylin and eosin (HE) in a standardized manner. Afterwards, they were recorded with a microscopy HSI camera system. Four different classes were annotated in the obtained HSI data: stroma, blank, squamous epithelium, and EAC. The number of spectra in each class is reported in Table 1. They were used for the training and testing of classification algorithms.

We evaluated in a cross-validation several machine learning algorithms: SVM with poly and linear kernel, Multi-Layer Perceptron (MLP) and Logistic Regression (LR).

Moreover, we balanced the classes by randomly selecting spectra in the different classes of stroma, EAC and blank to minimize the effect of overfitting. In each validation fold, the train and test set contained approximately the same percentage of samples of each class as the complete data set. We tested parameters of the classification methods (e.g. poly and linear kernel for SVM) for each classification algorithm based on a grid search cross-validation.

Table 1: Number of spectra in each class which are used in cross-validation.

Class	Number of Spectra
Stroma	68.215



Class	Number of Spectra
Blank	6.803
Squamous epithelium	64.380
EAC	175.027

3 Results

Figure 3 shows the mean absorbance spectra of the four classes. Differences in the absorbance of squamous epithelium and esophageal adenocarcinoma (EAC) cells are obvious. Moreover the eosin's maximal absorption at 525 nm (blue arrow) is clearly observed. The absorbance spectrum of the eosin is depicted in Figure 2. The hematoxylin's peak at 590 nm (red arrow) is mainly visible on the absorbance spectra of the EAC. However, the intragroup differences for the squamous epithelium and the esophageal adenocarcinoma

Figure 3: Mean absorbance of the four classes stroma, blank, squamous epithelium and EAC (eosin peak are marked as blue arrow and hematoxylin peak are marked as orange arrow).

cells were quite low.

The SVM with a poly kernel showed the worst results with sensitivity with 63% and a F1-Score of 59% against the MLP with the best results (90% sensitivity and F1-Score) for all classes (Table 2). The sensitivity values of the separate classes EAC and background were up to 25% higher than for stroma and squamous epithelium.

Table 2: Classification results of each method.

Classification method	F1-Score in %	Sensitivity in %	Best Parameter
SVM	56 ±0	61±0	Kernel: linear
LR	74±0	74±0	Saga solver and L2 penalty
MLP	90±0	90±0	Activation: tanh, solver: linear-Broyden-Fletcher-Goldfarb-Shanno algorithm

4 Discussion and Conclusion

For the first time in literature, we were able to analyze esophageal adenocarcinoma, tumor stroma and squamous epithelium cells using HSI. Neuronal network with an MLP algorithm provided the best performance.

In gastric cancer specimens a classification was done based on micro-hyperspectral technology using a deep-learning model-based analysis method [7]. The experimental results showed a sensitivity with 96% for cancerous and normal gastric tissue for 30 patients. In contrast to our work, the whole image was annotated in [7], so more spectral data for each patient were available. With a higher number of spectral data, we assume that a deep learning network could improve the results further.

In [8] a support vector machine (SVM) classifier was used to detect pancreatic tumor cell nuclei based on stained spectra using 12 patients with a high number of spectral data (n=656440) in the wavelength range from 420 - 750nm. The best sensitivity (91%) was obtained with the hyperspectral analysis of HE stained pathological slides using high dynamic range images (HDR). In this study, the SVM showed lower results with only 59% sensitivity. Due to the different clinical use cases and the higher number of spectra used in [8], different results were achieved in this work. In future studies, improved results of classifying stroma, squamous epithelium, and EAC by using HDR need to be evaluated.

Furthermore, the histopathological tumor cell identification needs further validation of the training algorithms to foster a semi-automatic decision-making process.

Author Statement

Research funding: This study was funded by the Federal Ministry of Education and Research 13GW0248B.

Conflict of interest: Authors state no conflict of interest. Informed consent: Informed consent has been obtained from all individuals included in this study. Ethical approval: The research related to human use complies with all the relevant national regulations, institutional policies and was performed in accordance with the tenets of the Helsinki Declaration, and has been approved by the authors' institutional review board or equivalent committee.

References

- [1] G. Lu and B. Fei, „Medical hyperspectral imaging: a review“, *J. Biomed. Opt.*, Bd. 19, S. 24, 2014.
- [2] H. Akbari *u. a.*, „Hyperspectral imaging and quantitative analysis for prostate cancer detection“, *J. Biomed. Opt.*, Bd. 17, Nr. 7, S. 0760051, Juli 2012, doi: 10.1117/1.JBO.17.7.076005.
- [3] H. Akbari, K. Uto, Y. Kosugi, K. Kojima, and N. Tanaka, „Cancer detection using infrared hyperspectral imaging“, S. 6, 2011.
- [4] A. Goto *u. a.*, „Use of hyperspectral imaging technology to develop a diagnostic support system for gastric cancer“, *J. Biomed. Opt.*, Bd. 20, Nr. 1, S. 016017, Jan. 2015, doi: 10.1117/1.JBO.20.1.016017.
- [5] H. Ogihara, Y. Hamamoto, Y. Fujita, A. Goto, J. Nishikawa, und I. Sakaida, „Development of a Gastric Cancer Diagnostic Support System with a Pattern Recognition Method Using a Hyperspectral Camera“, *J. Sens.*, Bd. 2016, S. 1–6, 2016, doi: 10.1155/2016/1803501.
- [6] M. Ishigaki *u. a.*, „Diagnosis of early-stage esophageal cancer by Raman spectroscopy and chemometric techniques“, *The Analyst*, Bd. 141, Nr. 3, S. 1027–1033, 2016, doi: 10.1039/C5AN01323B.
- [7] B. Hu, J. Du, Z. Zhang, and Q. Wang, „Tumor tissue classification based on micro-hyperspectral technology and deep learning“, *Biomed. Opt. Express*, Bd. 10, Nr. 12, S. 6370, Dez. 2019, doi: 10.1364/BOE.10.006370.
- [8] M. Ishikawa *u. a.*, „Detection of pancreatic tumor cell nuclei via a hyperspectral analysis of pathological slides based on stain spectra“, *Biomed. Opt. Express*, Bd. 10, Nr. 9, S. 4568, Sep. 2019, doi: 10.1364/BOE.10.004568.