Support vector machine for evaluation of autofluorescence cancer diagnostic parameters

Ts. Genova[1], Al. Zhelyazkova[1], E. Borisova[1,*],L. Zaharieva[1], N. Penkov[2], B. Vladimirov[2]
[1]Institute of Electronics, Bulgarian Academy of Sciences, Bulgaria
[2]University Hospital "Tzaritza Yoanna – ISUL", Bulgaria
*correspondence to :ts.genova@gmail.com

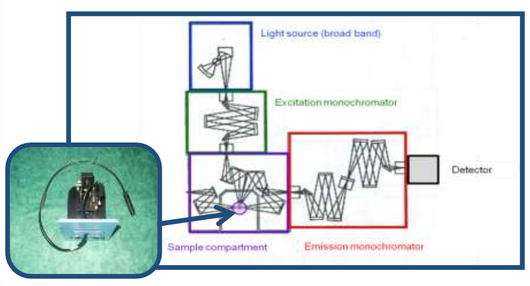
Introduction

Colorectal cancer is the third most common cancer and cause of cancer related deaths worldwide. Early diagnosis and accurate staging of the colorectal cancer are critical for a successful curative treatment. Some optical spectroscopic techniques are investigated for implementing in the clinical diagnostic procedure in order of achieving a greater accuracy of the diagnosis, the most intensively researched one is the fluorescence spectroscopy, mainly because of its sensitivity and potential for revealing tissue cancer related alterations.

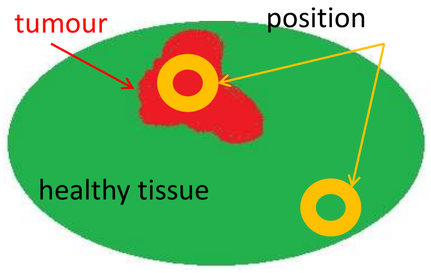
Excitation-emission matrix (EEM) allows presentation of the obtained data as a three-dimensional fluorescence map that provides information about the fluorescence features of the biological tissue samples. The EEM could be used for determining of the range of excitation wavelengths that conduct emission spectra, containing diagnostically significant fluorescence features, for following clinical analysis.

Materials and methods

FluoroLog 3 (HORIBA Jobin Yvon, France) and F-3000 fiber-optic module



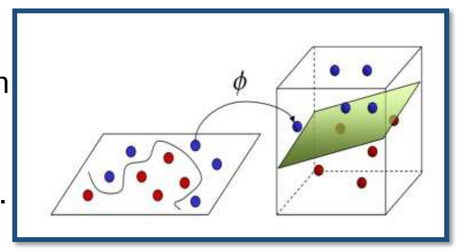
Ex vivo tissue samples scheme:

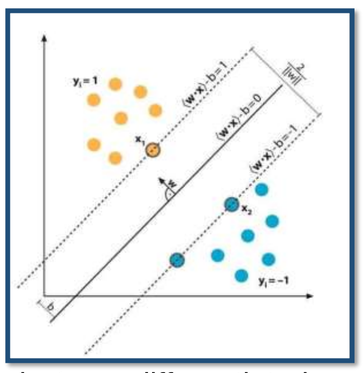


After surgical removal colorectal tissue samples are transported in isothermal conditions and safe-keeping. Point-by-point measurements were taken from the tumour and from the normal mucosa, part of the safety area excised during the tumour removal.
The fluorescence of the samples was evaluated through EEM excitation of 280-440 nm and detection of 300-800 nm.

Support vector machine

Non-linear transformation to a higher dimensional feature space.





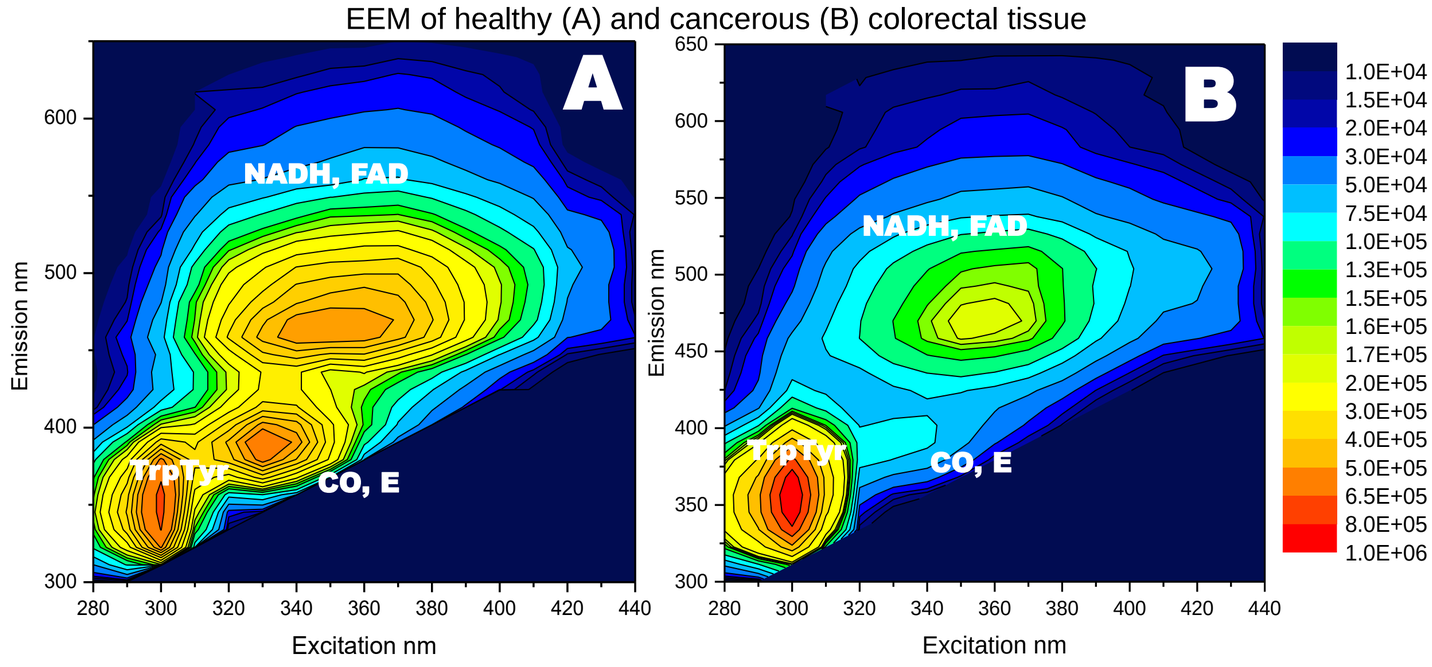
$$Sn = \frac{TP}{(TP+FN)}$$
$$Sp = \frac{TN}{(FP+TN)}$$
$$DA = \frac{(TP+TN}{TP+FP+TN+FN}$$

Two classes yi=1 и yi=-1 are differentiated based on two points (vectors) x1 and x2 with maximum margin hyperplane, defined as the distance between the closest points should be 2/||ω||. Where ω is the normal vector for the hyperplane.

Classification through SVM was performed first with whole fluorescence spectra obtained for one excitation wavelength, to determine the most diagnostically valuable spectra and then for specific parameters, derived from the said spectra. To evaluate the classification the sensitivity (Sn), specificity (Sp) and diagnostic accuracy (DA) were calculated.
Total of 32 measurements were included in the classification – 20 for training and 12 for testing.

Results

EEM of healthy (A) and cancerous (B) colorectal tissue

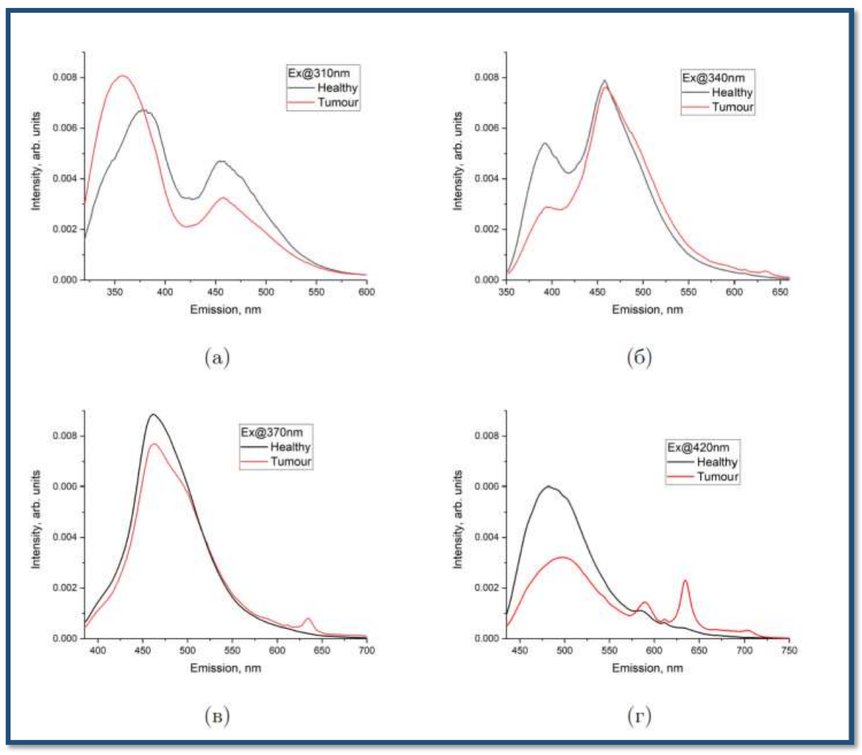


Major recognized autofluorescence sources in the tissues investigated.

Endogenous fluorophore	Tryptophan Tyrosine	Collagen Elastin	NADH	FAD	Porphyrins
Excitation maximum [nm]	280-300	320-360	340-380	450-500	400-450
Emission maximum [nm]	320-400	400,460-500	450-500	450-530	630-690

Statistical evaluation

Exc λ nm	Parameter	Sn %	Sp %
280	60.00	71.43	66.67
290	50.00	66.67	58.33
300	100.00	55.56	66.67
310	100.00	75.00	**91.67**
320	71.43	100.00	83.33
330	80.00	57.14	66.67
340	85.71	100.00	**91.67**
350	75.00	50.00	58.33
360	33.33	100.00	50.00
370	100.00	80.00	**91.67**
380	100.00	55.56	66.67
390	80.00	71.43	75.00
400	100.00	66.67	83.33
410	100.00	71.43	83.33
420	100.00	83.33	**91.67**
430	80.00	71.43	75.00
440	75.00	62.50	66.67



Fluorescence spectra from ex vivo tissue samples for the excitation of 310, 340, 370 and 420 nm.

Exc λ nm	Parameter	Sn %	Sp %	DA %
310	I_{350}/I_{450}	100.00	75.00	83.33
340	I_{390}/I_{460}	80.00	100.00	**91.67**
370	$I_{460}//I_{635}$	60.00	100.00	83.33
420	I_{480}/I_{635}	100.00	66.67	83.33

Conclusions

As a result of the conducted work the excitation wavelength of 340 nm and differentiation based on the parameter I_{390}/I_{460} is evaluated as the most diagnostically valuable, among the investigated. The diagnostic value of this parameter is based mostly on the fluorescence of tryptophan, which is with higher intensity for cancerous tissues and collagen fluorescence, which has higher intensity for healthy tissues. This difference is further enhanced with calculation of the ratio between those two intensities for cancerous and healthy tissues. However some limitations should be pointed out: the investigation is performed on ex vivo tissue samples, hence the strong effect of hemoglobin absorption is strongly reduced and doesn't effect the data; used equipment is highly sensitive and not applicable in clinical practice. All those should be taken into account for our future plans to translate this technique into the clinical practice for endoscopic gastrointestinal cancer diagnosis.

Acknowledgments: The investigations were supported by the Bulgarian National Science Fund under grants #KP06-N28/11/14.12.18 and #KP06-N23/8/2018. Spectrofluorimetric equipment used was purchased under the NSF grant #DO-02-112/2008.