Diagnosis of benign and precancerous skin conditions using bimodal spectroscopy, machine learning and Data Fusion methods

Valentin Kupriyanov[1,2] , Walter Blondel[1] , Christian Daul[1] , Marine Amouroux[1] and Yury Kistenev[2]
[1] The Research Center for Automatic Control (CRAN), University of Lorraine, Nancy, France
[2]Laboratory of Laser Molecular Imaging and Machine Learning, Tomsk State University, Tomsk, Russia
valentin.kupriyanov@univ-lorraine.fr
The research was carried out with the support of a grant under Vernadski international joint PhD program (2021-2024) provided by the French embassy in the Russian Federation
The research was carried out with the support of a grant under the Decree of the Government of the Russian Federation No. 220 of 09 April 2010 (Agreement No. 075-15-2021-615 of 04 June 2021)

Optical methods for skin cancer diagnosis have many advantages over biopsy, which is used traditionally. These advantages are lower cost per test in case of mass screening, non-invasiveness and less time expenditure. One of the most promising fields of use of optical methods for skin cancer diagnosis is multimodal methods, which combine several types of conventional methods. However, the use of multimodal methods requires the application of Data fusion techniques to combine information obtained by different methods. This study presents the results of an analysis of diffuse reflectance spectra and autofluorescence spectra obtained in vivo from the skin of mice with precancerous or benign lesions. The bimodal spectroscopic device used in this study had a spectral range of 317-789 nm, and allowed acquisition of spectra for three values of the distance between the detector and the source - 271, 536 and 834 μm. All spectra were preprocessed using the principal component analysis to solve the dimensionality problem. The classification was carried out by the support vector machine, discriminant analysis, multilayer perceptron and random forest classifier. The results were combined using stacking and voting. This study presents a comparison of the performance these machine learning and Data Fusion methods with more traditional approaches.

Keywords: skin cancer, diffuse reflectance spectroscopy, autofluorescence spectroscopy, multimodal spectroscopic methods, machine learning, data fusion.