Accepted Manuscript

Discrimination of skin cancer cells using Fourier transform infrared spectroscopy

Francisco Peñaranda, Valery Naranjo, Gavin R. Lloyd, Lena Kastl, Björn Kemper, Jürgen Schnekenburger, Jayakrupakar Nallala, Nicholas Stone

PII: S0010-4825(18)30171-9

DOI: 10.1016/j.compbiomed.2018.06.023

Reference: CBM 3003

To appear in: *Computers in Biology and Medicine*

Received Date: 13 April 2018

Revised Date: 22 June 2018

Accepted Date: 23 June 2018

Please cite this article as: F. Peñaranda, V. Naranjo, G.R. Lloyd, L. Kastl, Bjö. Kemper, Jü. Schnekenburger, J. Nallala, N. Stone, Discrimination of skin cancer cells using Fourier transform infrared spectroscopy, *Computers in Biology and Medicine* (2018), doi: 10.1016/j.compbiomed.2018.06.023.

Discrimination of skin cancer cells using Fourier transform infrared spectroscopy

Francisco Peñaranda[a], Valery Naranjo[a,], Gavin R. Lloyd[b,1], Lena Kastl[c], Björn Kemper[c], Jürgen Schnekenburger[c], Jayakrupakar Nallala[d], Nicholas Stone[d]

[a] *Instituto de Investigación e Innovación en Bioingeniería (I3B), Universitat Politècnica de València, Camino de Vera s/n, 46022 Valencia, Spain*
[b] *Biophotonics Research Unit, Gloucestershire Hospitals NHS Foundation Trust, Gloucester, United Kingdom*
[c] *Biomedical Technology Center, University of Münster, Münster, Germany.*
[d] *Biomedical Physics, School of Physics, University of Exeter, Exeter, United Kingdom*

Abstract

Fourier transform infrared (FTIR) spectroscopy is a highly versatile tool for cell and tissue analysis. Modern commercial FTIR microspectroscopes allow the acquisition of good-quality hyperspectral images from cytopathological samples within relatively short times. This study aims at assessing the abilities of FTIR spectra to discriminate different types of cultured skin cell lines by different computer analysis technologies. In particular, 22700 single skin cells, belonging to two non-tumoral and two tumoral cell lines, were analysed. These cells were prepared in three different batches that included each cell type. Different spectral preprocessing and classification strategies were considered, including the current standard approaches to reduce Mie scattering artifacts. Special care was taken for the optimisation, training and evaluation of the learning models in order to avoid possible overfitting. Excellent classification performance (balanced accuracy between 0.85 and 0.95) was achieved when the algorithms were trained and tested with the cells from the same batch. When cells from different batches were used for training and testing the balanced accuracy reached values between 0.35 and 0.6, demonstrating the strong influence of sample preparation on the results and comparability of cell FTIR spectra. A deep study of the most optimistic results was performed in order to identify perturbations that influenced the final classification.

Keywords: Machine learning, Multivariate analysis, Cancer diagnosis, Cytopathology, Fourier transform infrared spectroscopy

1. Introduction

Infrared spectroscopy is a highly promising optical technology in the combined identification and localization of pathophysiological cell and tissue alterations [1]. A limiting factor for the use of Fourier Transform Infrared (FTIR) spectroscopes in biomedical problems has been the sensitivity of the systems, which is related to the relatively high acquisition times needed to obtain measurements of sufficient quality. This problem is of greater concern for single layers and/or individual cells, which contain less biological material than tissues and provide lower signals. Therefore, to obtain spectra from cells with signal-to-noise ratio comparable to tissues, the acquisition time is normally on the order of two or even four times higher for cells than for tissue. Moreover, the measurement of cells has normally been a very tedious task because they are frequently spread out in the sample preparations. These facts have hampered the proper analysis of cell spectra and prevented the systematic assessment of their discriminative power.

Recently, modern FTIR microspectroscopes have increased their acquisition speed mainly thanks to

Email address: vnaranjo@upv.es (Valery Naranjo)
[1] Present address: *Phenome Centre Birmingham, School of Biosciences, University of Birmingham, Birmingham, United Kingdom.*