

AUTOMATIC AND QUANTITATIVE IMAGE ANALYSIS OF MINIMAL SAMPLES WITH APPLICATION IN EARLY LUNG CANCER DIAGNOSIS

B. Ucar, A. Muñoz-Barrutia, T. Pengo, I. Zudaire, C. Ortiz-de-Solórzano
Division of Oncology, Centre for Applied Medical Research (CIMA), Pamplona
University of Navarra
Pío XII 55, 31008 Pamplona, Spain
E-mail: bucar@unav.es

KEY WORDS: Early lung cancer diagnosis, minimal samples, multidimensional image analysis, pseudo-confocal microscopy, blind unmixing, cell segmentation, spot counting.

We propose a method for the automatic analysis of images of Bronchoalveolar Lavage (BAL) samples and evaluate its suitability for the early diagnosis of lung cancer. The samples are stained using a novel cytogenetic technique known as Fluorescence Immunophenotyping and Interphase Cytogenetics as a Tool for the Investigation of Neoplasms (FICTION) [1]. FICTION combines fluorescence immunostaining of a tumor marker (i.e. a protein overexpressed in lung cancer cells) and up to four interesting genes by Multicolor Fluorescence In Situ Hybridization (M-FISH). The combination of these techniques offers higher sensibility than conventional FISH for identifying the presence of cancer cells. BAL samples contain very low numbers -if any- of neoplastic cells. For that reason, the identification of abnormal cells using FISH can be a difficult and time consuming process. FICTION facilitates a rapid selection of rare positive cells by fluorescence immunophenotyping and can simultaneously identify specific genetic aberrations harbored by these cells. Due to its novelty, there is no previous study in the literature on the automation of the quantification of minimal samples stained with FICTION in any type of cancer, not even for the hematological neoplasias for which FICTION was developed several years ago.

Our method takes advantage of a multidimensional and multispectral automated microscopy system to acquire 2D fluorescence immunostaining images of the samples and then 3D stacks of sample areas occupied by candidate cells (immunostaining and FISH).

The images are analyzed as follows: First, we detect immuno-positive and immuno-negative cells -isolated or clustered- on the 2D images using a multispectral quantitative algorithm. Object classification is done using Support Vector Machine classifier previously trained and tested using lavage samples with sprinkled cancer cells from a commercial cell line. Next, immuno-positive cells, candidate of being of tumoral origin, are imaged pseudo-confocally in 3D using sequential acquisition of Z stacks deconvolved using Houghens SVI software. Subsequently, the spectral overlap between the fluorochrome emissions of the FISH channels is eliminated using a blind unmixing algorithm. Finally, clustered cells are detected and segmented to facilitate FISH spot counting on a per cell basis. FISH signals are detected using a 3D modified Top Hat and classified using a Support Vector Machine classifier. We employed a 10-fold cross validation procedure to estimate the classification accuracy of the classifiers.

We have validated the analysis using pure lavage samples of lung cancer. We also plan to present a statistical validation of the automated FICTION analysis to justify its competitive advantage against its conventional counterpart: manually performed FISH and FICTION analysis.

[1] K. Weber-Matthiesen, S. Pressl, B. Schlegelberger, and W. Grote, "Combined immunophenotyping and interphase cytogenetics on cryostat sections by the new FICTION method," *Leukemia*, **7**, 646-649 (1993).