



Proceedings

SY05.06 | Computer Aided Diagnosis

A Workflow For Computer-Aided Cytology In Whole Slide Images: Application In Fine-Needle Aspiration Thyroid Cytology

R. Mormont¹, J.-M. Begon¹, C. Degand², N. D'Haene², I. Salmon², R. Marée^{*1}

¹University of Liège, Systems and Modeling, Liege, Belgium, ²ULB Erasme Hospital, Pathology Department, Brussels, Belgium

Introduction/ Background

Computer-aided cytology has a long history in computer vision research with a large amount of works in cervical cancer screening [Bengtsson et al. 2014; Delga et al. 2014]. However, one still lacks practical systems in many cytology fields due to several technical challenges (incl. the need for significant computational resources and the lack of efficient collaborative tools to efficiently collect and organize realistic and large ground-truth data that would enable large validation of recognition algorithms).

Aims

We present a novel workflow for computer-aided cytology using latest web, databases, and machine learning technologies with the aim to speed up the implementation of such systems. Here, we focus on specializing this framework and applying it for the assessment of thyroid nodules. In practice, physicians have to efficiently stratify patients according to their risk of malignancy in order to identify the best follow-up and therapeutic options. Fine-needle aspiration (FNA) and cytological assessment has become the predominant method used for the primary diagnosis of benign and malignant thyroid nodules, resulting in the categorization of patients as operative or non-operative candidates.

Methods

Data: FNAs were carried out using a 21-gauge needle attached to a 10-mL syringe. The aspirated material was smeared on slides, air-dried and subjected to a Diff-Quick stain. FNAs were scanned (Hamamatsu scanner, 40X, 0.23µm) at the ULB. Digital slides were transferred to a Cytomine (<http://www.cytomine.be/>) [Marée et al., 2016] server at the ULG.

Algorithms: Our framework is generic so that one can specialize each of its component. Here, colour deconvolution [Ruifrok & Johnston 2001] is first applied on original image tiles to detect cells and separate them from background/artifacts. Detected foreground objects are then separated into individual cells or clusters. Localization of individual cells within clusters is then performed using Watershed and the distance transform. All individual cells are then classified using variants of our image classification algorithms. This binary model was trained and optimized using our ground-truth dataset of individual cells where we considered papillary cells with inclusion as positive and all other types of objects (cells with ground glass nuclei, nuclear grooves, normal follicular cells, artifacts, macrophages, polynuclear,...) as negative. Similarly, large clusters are classified using a binary model optimized on the ground-truth dataset to discriminate between normal and proliferative follicular architectural patterns.

Results

Using Cytomine web annotation tools, experts first built an unprecedented ground-truth dataset of various types of normal and abnormal cells and clusters (> 6000 objects from 60 FNA whole-slide images) to train recognition models. Once all cells of new slides are classified by our workflow, predictions are uploaded to the Cytomine-Core server through HTTP requests and can be displayed in the Cytomine-WebUI as sorted galleries



of most suspicious objects. At the conference we will present our qualitative and quantitative evaluation of the different steps of the workflow and discuss limitations and perspectives. This novel Cytomine module will be released as open-source in the near future so that other research groups will be able to train, apply, and extend it on their own data.