Localising Cell Communities in Multi-Gigapixel Digital Pathology Images[1]

The research focus in our Bioimage Analysis (BIA) lab is to examine images of tissue taken from cancer patients with the aim of helping the pathologist come to conclusions that are soundly based on evidence and, therefore, objective. Our group analyses images of cancerous tissues from the University Hospital Coventry and Warwickshire (UHCW), the first hospital in the UK to go completely digital in pathology in recent months. We have a close working relationship with the new Digital Pathology Centre of Excellence at UHCW, publishing joint papers that should improve diagnosis and treatment. Our work on whole slide images (WSI) with size in the range of 100,000×100,000 pixels, using modern digital slide scanners, brings us into contact with GE and Intel, with whom we have ongoing and upcoming joint projects.

A digitised image of a cancerous tissue slide represents the complex tumour microenvironment (TME) consisting of tens of thousands of cells and cell communities of various types and their interactions with each other. Detecting functionally important cell communities within the TME is critical to the analysis of such images but cannot be done without help of expert pathologists. For example, we need to determine the type of the cells we are looking at (there are several hundred human cell types) and whether cells are healthy or diseased. But the computer scientists among us have no medical training. In fact, medical training would be insufficient — the pathologists working with us build on several decades of learning, and on examining hundreds of tissue slides every day.

In order to overcome the above problems, we obtain the *ground truth*, also called the *gold standard* from pathologists who annotate images with their views. This process is subject to human error but is the best information we have. Our biggest problem is not the accur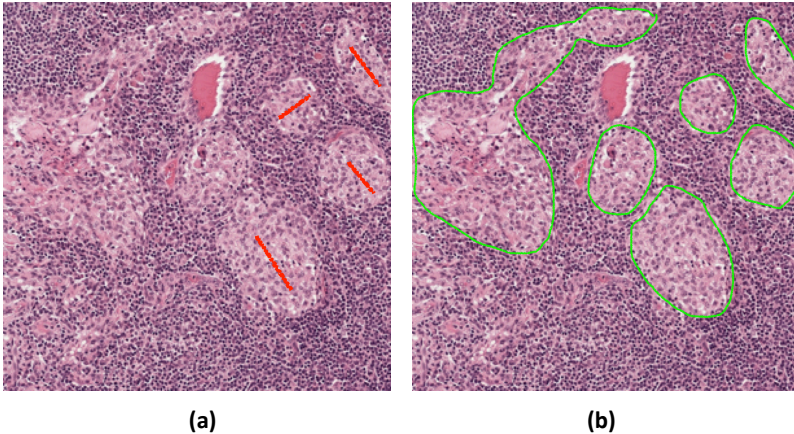acy of the annotation, but the availability of precious pathologist time. They are rightly more concerned with giving good advice to clinicians treating patients with urgent needs, than they are with helping a researcher with a computer program that may or may not in the longer run turn out to be useful. Our approach is to get maximum information from the pathologists with minimal expenditure of their time. Traditionally a pathologist would annotate the histology image by roughly drawing curves surrounding the region of interest. A better solution would be for the pathologist to rapidly mark each area of interest with a single brushstroke of the pen (or stylus or mouse movement). A computer program can then expand the area to an approximation of what a pathologist, with lots of free time and a love of detailed annotation, would have produced. The task in this M.Sc.

(a)   (b)

Figure 1: (a) A snapshot of a small visual field in an image of breast cancer nodal biopsy showing follicles consisting of histiocytes overlaid with "brushstrokes" by expert pathologists, (b) result of the community localisation algorithm.

project is to move as far as possible towards that goal. Of course, the output from such a program would need to be looked at again by a pathologist, with feedback about the success of the program.

The main objectives for the project are:
- Identify salient features of cell communities differentiating them from their surroundings, eg building on our recent work [1] and extending it using [2].
- Based on minimal input from a pathologist, identify and locate the desired cell communities.
- Incorporate the algorithm into a software tool (ideally written in MATLAB) to allow the user to read and draw regions in a whole slide image.

Ph.D. Project: The M.Sc. project sketched above is ambitious and open-ended. If the techniques developed can be applied to help with breast cancer images, it can be extended to inferring cell communities of different kinds in one of the larger cohorts of cancer patients being studied in our group. Once cell communities have been inferred, one can look into characteristics of cell communities for their association with cancer outcome and/or survival. Therefore, this project is highly likely to lead to a suitable Ph.D. project.

References:

[1]. K. Sirinukunwattana, D.R.J. Snead, and N.M. Rajpoot. *A novel texture descriptor for detection of glandular structures in colon histology images.* In SPIE Medical Imaging (pp. 94200S-94200S), 2015.
[2]. M. Tang, L. Gorelick, O. Veksler, Y. Boykov. *GrabCut in One Cut.* IEEE ICCV, 1769-1776, 2013.

---

[1] Please also have a look at the other MSc project "**Visualising Social Networks of Cells in Tumour Microenvironment**"