Reviewer's report

Title: Automated Detection of Regions of Interest for Microarray Experiments: An Image Texture Analysis

Version: 1 Date: 21 November 2006

Reviewer: Donal O'Shea

Reviewer's report:

General
The paper is seems like a nice approach to tackling the problem. The simplicity of using greyscale image segmentation with an estimation of a likelihood ratio test to determine if a segment of a whole section is cancerous seems straightforward. If these ideas were to be implemented it would be good to allow the segmentation algorithm to change (and parameters such as B, P, H). The classification should of course be able to accept new observed data sets.

Some Commentary and Discretionary Revisions

The texture parameters B, P, H: These parameters seem logical to use in describing H&E stained slides. Obviously could use different parameters depending on tissue and stain.

Machine learning algorithm: This is not machine learning in the strict sense. More like pattern recognition. Machine learning usually has some form of feedback loop where the algorithm gets more 'intelligent' as it sees more data. This algorithm compares with a set of known results to make a decision.

I think a workflow diagram might help this paper.

Each of the image blocks must be analysed by the image analysis algorithm (to yield a texture parameter set) and also analysed by a pathologist, and through the machine learning algorithm classified into the cancerous, non-specific, and normal categories.

The advantages of this approach are that no 'difficult' image analysis code needs to be written, characteristics are determined in terms of intensities/colours. Hence the same algorithm could be used for slides with a different tissue type stained with something other than H&E by changing the segmentation algorithm.

Also the machine learning part can be altered by simply getting a pathologist to analyse more images / different types of stain. Hence it would be relatively straightforward to generate a new segmentation algorithm for other stains and a new 'training' dataset for the classification phase, without too much coding overhead.

In the same sense this would be a weakness as the classification depends on how many images are used, the type, the experimental variability, the inter observer variability etc.

2.2.1 Gray scale segmentation

As mentioned in the results section, k-means clustering is biased towards trying to produce clusters of equal size, so it might incorrectly determine a particular quantity for a region. The result also depends on the initial values (the centroid) in each cluster; they chose the low, median, and high
intensity values which guides the clusters in the right direction.

2.2.2 Colour segmentation

I’m not sure why a Weibull distribution was used to model the luminance indices.

k-means clustering using vectors rather than points as a centroid seems like a nice idea. Similar idea as linear regression in that classification would minimize the sum of squares distance from a index to its cluster centroid.

2.3 Texture Parameters of Segmented Histology Image Blocks

The determination of the upper bound for unstained regions could have large variation. It may be different for different types of tissue. However as unstained regions (no cancer) are not of interest they can be ignored (less processing).

2.4 Machine Learning for Detecting Cancer-Specific Tissue Image Regions

As mentioned before the algorithm is not learning/adapting as a result of seeing data, there is no feedback loop. This is more like pattern matching.

The log likelihood ratio test is usually used to test between null and alternative models to fit the data. It seems obvious that the distribution of the normal and cancerous data cannot be described by a parametric distributions (they would be different). The theory for classification seems correct.

The estimation of the ratio of probabilities would hold if both M and N were large, and assuming that the nearest neighbour classifier is an appropriate method to use for texture parameter sets. I assume that is ok to estimate this way.

The N texture profiles from the normal and cancerous reference sets are observed are calculated like any other image, however their status as normal or cancerous is predetermined by a pathologist. This is the pattern matching stage (in the form of NN classifier). Different reference sets could be used for different tissue types, therefore allowing flexibility.

The algorithm outlined for estimation is the same as described in the text.

I’m not sure about support vector regression.

3. Results section

Tables 2 and 3 seem to be the important tables with regards to validation of the method (computer vs pathologist) yet their results are only briefly discussed in the last paragraph of section 3.2.
Discretionary Revisions (which the author can choose to ignore)

What next?: Accept after discretionary revisions

Level of interest: An article whose findings are important to those with closely related research interests

Quality of written English: Acceptable

Statistical review: No

Declaration of competing interests:
I declare that I have no competing interests