Reviewer's report

Title: Large-scale computations on histology images reveal grade-differentiating parameters for breast cancer.

Version: 1 Date: 4 September 2006

Reviewer: Krzysztof OKon

Reviewer's report:

General
I read the paper by Petushi et al. with a very pleasure and interest. However, several problems in the article are to be corrected.

Major Compulsory Revisions (that the author must respond to before a decision on publication can be reached)

Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)
1) "Background", last sentence of 1st paragraph: the description of grading is not coherent, and leave tumors with score 5 or 6 out of it.
2) Methods section: No information is given on how the cases studied were selected nor where were derived from. Are these surgical or biopsy cases? How the cases were selected?
3) Methods section: Where is no mention of the hardware used for image acquisition; the type of microscope and camera should be indicated.
4) Methods section: I understand that the software used was developed explicitly for the study; however this is not directly indicated in the text. If so, I would like to know also that was programming environment, and especially, that libraries and or similar were used.
5) page 6, middle. The term hyperchromatic nuclei is used here in its plain sense for indicating inflammatory cells' nuclei. More descriptive term should be used here, as 'hyperchromatic', 'hyperchromasia' are terms used for describing features of neoplastic cells.
6) In table 1, a column with case numbers, though not adding significance, would make it more readable, because it is not clear without reading comment that individual cases are presented.

Discretionary Revisions (which the author can choose to ignore)
1) "Background", 2nd paragraph, 2nd sentence: should be clearly indicated that though in individual cases histologic grading system may not predict the exact outcome, it gives important prognostic information.
2) "Methods" indicating the class of computer (eg. standard PC or 1024 processor cluster system) used for analysis would show the computation power necessary for created image analysis system.
3) "Classification of section images" - Finding that automatic classification allows to put more cases in low grade & high grade groups (page 12, top) is interesting (as pathologist tend to mislabel cases as the middle group in a 3 groups system), and could be commented in discussion. However I would prefer not to draw a firm conclusion before validation of the obtained results with clinical data is available (i.e. it is shown that C1-C3 have more homogenous behaviour that classical G1-G-3).
4) "Conclusions" - In the text it is said that the presented methods makes possible to classifying breast cancers more exactly; however there are no evidences for it; agreement between two pathologists grading cases is 100%, thus it is difficult to do better... I expect and believe, in future analysis of a larger series of cases, including treatment results should be done. Also, 'clinical significance' should not be mentioned at this moment, as this was not studied so far.
5) Photographs are not perfect; this is possibly due to pdf file format I received, but should be checked by authors.
What next?: Accept after minor essential revisions

Level of interest: An article of importance in its field

Quality of written English: Acceptable

Statistical review: No

Declaration of competing interests:

I declare that I have no competing interests