Reviewer's report

Title: Characterization of Digital Medical Images Utilizing Support Vector Machines

Version: 1 Date: 29 September 2003

Reviewer: Harald Kittler

Reviewer's report:

General

Discretionary Revisions (which the author can choose to ignore)

Minor Compulsory Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)

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

1. In a recent meta-analysis (Rosado et al., Archives of Dermatology, 2002) the following basic quality requirements for experimental studies on computer diagnosis of pigmented skin lesions were suggested
   a. Selection of lesions should be random or consecutive
   b. Inclusion and exclusion criteria should be clearly stated.
   c. All lesions clinically diagnosed as melanocytic lesions should be analyzed
   d. The study setting should be clearly defined
   e. To avoid verification bias, clearly benign lesions which were not excised should be included. For these lesions the diagnostic gold standard could be short term follow-up with digital dermoscopy.
   f. Instrument calibration should be reported.
   g. Repeatability analysis should be carried out (inter- and intra-instrument)
   h. Classification should be carried out on an independent test set
   i. Computer diagnosis should be compared with human diagnosis.

The authors do not provide sufficient data to answer the question whether these quality requirements have been fulfilled or not. For example, the way the lesions were selected is not reported. Selection bias may explain the extremely good results of this study.

2. The authors stated that they used 17 lesions in the training set, which is a very low number. Why only 17 lesions? Did the investigators perform other experiments including more lesions? If yes, what were the results of these experiments. Did the authors only provide data of the experiment with the best results?

3. The authors never tell us how many lesions were included in the test set.

4. It seems that the authors used the test set for optimizing the parameter sigma (see section VGP-DSP comparison). In this case the test set is not independent.

5. Regarding the references, the authors overlooked a large part of the literature which appeared in medical journals.

Advice on publication: Unable to decide on acceptance or rejection until the authors have
responded to the major compulsory revisions

**Level of interest:** A paper whose findings are important to those with closely related research interests

**Quality of written English:** Acceptable

**Declaration of competing interests:**

one