Reviewer's report

Title: Computer-assisted Lip Diagnosis on Traditional Chinese Medicine Using Multi-class Support Vector Machines

Version: 4 Date: 28 June 2012

Reviewer: Yiguang Liu

Reviewer's report:

-Major Compulsory Revisions

The authors added several new and meaningful experimental results, but there are always many problems in this paper.

1. There are also several simple errors in this paper. I suggest the authors read the paper more carefully. For example:
   (1) In abstract, at the end of paragraph 1, there is a redundant space between T and C;
   (2) Also in abstract, page 2, in the paragraph 2, there is a redundant comma in line 4;
   (3) In the last paragraph of page 4, there is a redundant period in line 4;
   (4) In page 18, “lip images classification” # “lip image classification”;
   ...

2. The paper needs serious editing in terms of scientific writing. For example:
   (1) What do you want to convey through the results listed in the abstract?
   (2) Lip diagnosis in TCM is only to distinguish color of lips? Please read related literatures and define your work carefully.
   (3) In abstract, the conclusion is too rough. It haven't summarize your our work fairly.
   (4) In page 4, in the sentence “Automated recognition, if proven accurate …” of first paragraph, for the lip image based diagnosis, what should be recognized from the images?
   ...

3. Figure 1 should be removed. It may be waste of space. If authors can provide different pairs of lip features and syndrome diagnosis, they would be helpful and useful to let readers understand the work clearly.

4. In page 5, first paragraph, “the repeatability of the outcome…” means same symptom may be diagnosed as different syndrome by different TCM doctors?

5. Which work is about PR and DM application in diagnostic standardization in paragraph 2 of page 5?

6. In page 6, which evidence can support your opinion that “In our study, this
approach unfortunately failed”? KNN in your study is completely failed?

7. In page 6, at the end of 1 paragraph, does the sentence “Different from skin color classification, classifying lip images is …” mean that skin color classification is not a multi-class classification problem according to TCM theory?

8. In this paper, your main contribution is that new types of features for lip color classification are introduced. But in the experiments, there are no related comparisons of the different types of features to lip color classification. Are all types of feature introduced in this paper effective? And your work is really too similar with “[8] Li, F.-F., et al. Lip color classification based on support vector machine and histogram. In Proceedings of 3rd International Congress on Image and Signal Processing. 2010. pp:1883-1886”. You should give a comparison experiment between your feature and their features rather than simply give a reference.

9. The framework depicted in figure 2 is not appropriate. The training set and test set are two entities, and image processing and image segmentation are two working procedure, but you use same shapes to depict them. What’s the different between components “classifier” and “classifier/model”? Why does a testing arrow point to classifier component? Test data is also used to train the classifier?

10. I think you should introduce the RFE feature selection method used in your experiments in detail. How to incorporate WSVM, MAPLSC, Naïve Bayes and kNN into RFE?

11. Figure 6 is also useless.

12. I suggest that you should not only tell the readers the basic information about the methods, but also, more importantly, let the readers know how to choose an appropriate classification method for this distinctive work and why do you choose these methods. In this paper, you have included SVM-like and kNN-like method, but “Yiguang Liu, et al. A novel and quick SVM-based multi-class classifier. Pattern Recognition. 2006. 39(11):2258-2264”, “Rong-En Fan, et al. LIBLINEAR: a library for large linear classification, Journal of Machine Learning Research. 2008. 9:1871-1874” and “Yiguang Liu, et al. k-NS: a Classifier by the Distance to the Nearest Subspace. IEEE Transactions on Neural Networks. 2011. 22(8):1256-1268” are more state-of-art and typical. Therefore, you have the duty to tell the researchers that: are they alternative or not?

13. The experimental results should be rewrite to support the authors' opinions and arguments.

14. In Page 20, in the sentence “The proposed classification scheme achieved high accuracy for most classification problems”, what’s the meaning of “for most classification problems”? At the end of the paper, authors said that “These tools are effective not only to find and describe patterns in lip image based on the diagnosis of TCM... It provides the physicians a systematic and objective diagnostic standard...”. It is a little bit exaggerated.

15. Some related references should be cited, such as

A framework and its empirical study of automatic diagnosis of traditional Chinese medicine utilizing raw free-text clinical records, Journal of Biomedical Informatics,

A novel and quick SVM-based multi-class classiﬁer, Pattern Recognition, Vol. 39, No. 11, Nov. 2006, 2258-2264

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

**Quality of written English:** Needs some language corrections before being published

**Statistical review:** No, the manuscript does not need to be seen by a statistician.