Author's response to reviews

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

Authors:

FuFeng Li (li_fufeng@yahoo.com.cn)
Changbo Zhao (changbo0612@126.com)
Zheng Xia (zxia@tmhs.org)
Yiqin Wang (wangyiqin2380@sina.com)
Xiaobo Zhou (xzou@tmhs.org)
Guo-Zheng Li (gzli@tongji.edu.cn)

Version: 5 Date: 22 July 2012

Author's response to reviews: see over
Response letter

Dear editor,

We would like to thank you and the reviewers again for your helpful comments and suggestions. We have included most of them into the new version of our paper. For your convenience, we also have highlighted the major changes with yellow in the new version. Besides, there are many minor changes throughout the text, which we do not point them out. In the following, we summarize the major changes and then present the detailed responses to all of the reviewers.

Response to the reviewers:

Reviewer #1

Comment 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”.
Our response:

Thank you very much for the reviewer’s thorough comments. We have already corrected these errors in our paper.

Comment 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?*
significances. In addition, it also includes the inspection of moistening and aridity, which aims to judge the body fluid and its metabolism. For the fourth comment, we have modified that.

**Comment 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.*

**Our response:**

We have deleted Figure 1. Many thanks for the kindly reminder.

**Comment 4:**

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

**Our response:**

Because diagnosis in TCM is mostly conducted by doctors’ subjective judgments, it always has no object and quantizing indexes. As a result, the diagnosis consequences have weak reproducibility. In reality, for the same patients (or the same syndromes), different TCM doctors may
make different diagnosis about pattern of syndrome owing to that the doctors have different experience. This is called individuation in TCM. For example, the same patient is diagnosed as syndrome of deficiency of Qi by a doctor, and diagnosed as syndrome of Yang asthenia by other doctor; however, the therapy is the same. That is invigorating Qi and warming Yang. And the effect is also the same.

Comment 5:

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

Our response:

We point out PR and DM application in the new version.

Comment 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?

Our response:

Thanks a lot for the reviewer’s reminder. We have to admit the word “failed” maybe too exaggerated to convey our ideas about this algorithm
used in our paper. The accurate idea we want to show is that kNN classification accuracy is worse than other algorithms (SVM, WSVM, MAPLSC, Naïve Bayes) we applied. But kNN can also achieve accuracy more than 70%, it still has the significant research. So we changed the expression “In our study, this approach is slightly worse than other classifiers we applied.” Sincerely thanks again for your thorough comments to improve the scientific expression in this paper.

Comment 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?_

Our response:

According to the background in the introduction, at present, there are just reports about binary classification of skin colors and no works on multi-class classification. But in TCM, there is practical need on multi-class classification about lip colors.

Comment 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_

Our response:
In the process of our experiment, we collected three features including color, haralick and moments from lip images to make classification and comparison and we found that the best effect was achieved when we used these three features together. “The conference paper in 2010 [11]” is about the HSV feature extraction as for a color space. The authors used different features (H,S,V and their mean squares) to perform classification, but they did not make comparison among different classification methods. Our paper is a further extension. In the following, we represent results using color, haralick and moments features respectively using SVM and FLD.

**Total accuracy of classification based on different feature**

<table>
<thead>
<tr>
<th>24 color features:</th>
<th>FLD-78.3%, SVM-79.5%</th>
</tr>
</thead>
<tbody>
<tr>
<td>13 Har features:</td>
<td>FLD-70.8%, SVM-74.7%</td>
</tr>
</tbody>
</table>
47 Moment features: FLD-64.5%, SVM-62.1%
60 Har + Moment features: FLD-77%, SVM-68.1%
84 features (color+Har+Moment): SVM-80.4%
72 features (color+Har+Moment) after F-test: SVM-82.2%

Comment 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?

Our response:

Thanks a lot for the reviewer’s reminder. We checked our framework carefully and earnestly when we received this comment. Indeed, it may have an influence on readers to understand the whole system. So we modified them and made it clear for readers’ convenience. You may read it in the revision.

Comment 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 $k$NN into RFE?

Our response:

Thanks very much for the reviewer’s thorough comments. We have added some detail descriptions in the paper.

Comment 11:

Figure 6 is also useless.

Our response:

Many thanks for your kindly reminder, we have deleted Figure 6.

Comment 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?.

**Our response:**

Thanks very much for the reviewer’s thorough comments. We have collected some detailed information and described them in our paper in the subsection of “2.5 Classification”.

**Comment 13:**

*The experimental results should be rewrite to support the authors’ opinions and arguments.*

**Our response:**

Thanks a lot for the reviewer’s thorough comments. We have rewritten and summarized those scattered descriptions about our experimental results in the paper.

**Comment 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.

Our response:

We have modified that. Many thanks for the kindly reminder.

Comment 15:

Some related references should be cited, such as


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

Our response:

We have read these references carefully and cited them with corresponding explanation in subsection of “2.5 Classification” and “4. CONCLUSIONS”.
Reviewer #3

We are very grateful to the reviewer for checking the revision again.

Thank you very much for your acceptance of our paper.

Best Regards

Sincerely,

Fufeng Li

Email: Li_fufeng@yahoo.com.cn