Author's response to reviews

Title: A practical approach to Sasang constitutional diagnosis using vocal features

Authors:

Jun-Su Jang (junsu.jang@kiom.re.kr)
Boncho Ku (secondmoon@kiom.re.kr)
Young-Su Kim (ys001@kiom.re.kr)
Jiho Nam (jiho0803@kiom.re.kr)
Keun Ho Kim (rkim70@kiom.re.kr)
Jong Yeol Kim (ssmed@kiom.re.kr)

Version: 2 Date: 7 October 2013

Author's response to reviews: see over
Response to Reviewers’ Comments

We are very pleased to have received valuable comments from the reviewer about our manuscript titled, “A practical approach to Sasang constitutional diagnosis using vocal features”. We have revised our manuscript according to the reviewer’s comments. Please check the report to the reviewer’s comments below:

Reviewer 1’s Comments

This paper reports the investigation of a practical diagnosis model that has small number of variables to decrease model complexity in the Sasang constitutional diagnosis using vocal features. Its procedures were generally based on scientific methodologies and conducted with precision, and it was reported in an objective and logically sound manner.

However, there remain some points on parts of the manuscript that should be addressed with attentiveness, and revised:

1. In the material and methods section, it is a little vague and unclear about Sasang constitutional classification. So additional explanations for the decision process or the standard operation procedure in the Sasang constitutional classification should be added for clarity.

   We have added an overview subsection in methods section to give a clear explanation for the Sasang constitutional classification process at the 14th line in page 5. And Figure 1 has added to show the flow chart of the Sasang constitutional classification process.

   “This study followed a typical statistical pattern classification framework. Figure 1 shows the flow chart of the SC classification process. Voice data was separated into two parts, the training set and the test set. In the training phase, vocal features were extracted from the training set and a SC classification model was acquired using the proposed LASSO-based method. Evaluation of the training results was performed to examine the fitting ability of the obtained model to the training set. In the test phase, the same procedure was applied to the test set. Generalization ability of the obtained classifier was evaluated using the test set.”

2. The general presentation of the paper appears to be rather lengthy and redundant. The most effective method of conveying ideas across to the readers is by being concise, succinct, and clear-cut. It is strongly recommended that the authors look over the whole manuscript to rephrase or delete the sentences where possible to make the manuscript more refined and concise (without redundancy), so as to improve the general readability and presentation of the paper.

   We have reduced one page amount of redundant explanations in classification method section. The
whole manuscript has refined for better readability. And the final version has revised by a native English speaker.

**Reviewer 2’s Comments**

Minor Essential Revisions.

This study would be interesting to the readers regarding the medical applications of vocal features in Sasang Constitutional Medicine.

However, authors should revise the manuscript before considering be published, for such issues as follows;

1. **Basically, the contents of study (vocal features and Sasang Constitutional Medicine) would be relatively difficult to understand for general readers.** Then, this manuscript was not written as easy way to follow the study. So, I strongly suggest to add a "flow chart" of study in method session.

   We have added an overview subsection and a figure showing the flow chart of this study in method section. The figure is attached at the end of this document.

   “This study followed a typical statistical pattern classification framework. Figure 1 shows the flow chart of the SC classification process. Voice data was separated into two parts, the training set and the test set. In the training phase, vocal features were extracted from the training set and a SC classification model was acquired using the proposed LASSO-based method. Evaluation of the training results was performed to examine the fitting ability of the obtained model to the training set. In the test phase, the same procedure was applied to the test set. Generalization ability of the obtained classifier was evaluated using the test set.”

2. **The current abstract didn't show the clear summary of study question, method, finding, and conclusion.** For example, method part of abstract omitted important information such as participants. Please make abstract simple and impact, but comprehensive.

   The whole abstract has been rewritten and revised by a native English speaker. Method part now includes the information of participants.

   “Abstract

   Background
Sasang constitutional medicine (SCM) is a type of tailored medicine that divides human beings into four Sasang constitutional (SC) types. Diagnosis of SC types is crucial to proper treatment in SCM. Voice characteristics have been used as an essential clue for diagnosing SC types. In the past, many studies tried to extract quantitative vocal features to make diagnosis models; however, these studies were flawed by limited data collected from one or a few sites, long recording time, and low accuracy. We propose a practical diagnosis model having only a few variables, which decreases model complexity. This in turn, makes our model appropriate for clinical applications.

Methods

A total of 2,341 participants’ voice recordings were used in making a SC classification model and to test the generalization ability of the model. Although the voice data consisted of five vowels and two repeated sentences per participant, we used only the sentence part for our study. A total of 21 features were extracted, and an advanced feature selection method—the least absolute shrinkage and selection operator (LASSO)—was applied to reduce the number of variables for classifier learning. A SC classification model was developed using multinomial logistic regression via LASSO.

Results

We compared the proposed classification model to the previous study, which used both sentences and five vowels from the same patient’s group. The classification accuracies for the test set were 47.9% and 40.4% for male and female, respectively. Our result showed that the proposed method was superior to the previous study in that it required shorter voice recordings, is more applicable to practical use, and had better generalization performance.

Conclusions

We proposed a practical SC classification method and showed that our model having fewer variables outperformed the model having many variables in the generalization test. We attempted to reduce the number of variables in two ways: 1) the initial number of candidate features was decreased by considering shorter voice recording, and 2) LASSO was introduced for reducing model complexity. The proposed method is suitable for an actual clinical environment. Moreover, we expect it to yield more stable results because of the model’s simplicity.”

3. Discussion should be revised, regarding clinical applications as well as its limitations.

We have revised discussion section. Paragraphs regarding clinical applications and limitations have been added in page 13 and 14.

“The proposed method is suitable in an actual clinical environment where patients might not
pronounce a long recording content. Aged people especially have difficulty in uttering vowels for a long period of time. Our approach is also applicable to many voice-based diagnoses such as voice disorder detection, vocal cord dysfunction, and constitutional health diagnosis, expecting more stable results because of the simplicity of the model.

There are, however, some limitations of the proposed method. The classification model is trained by using a predefined sentence, and the same sentence must be pronounced during the test phase. If the recording content slightly changes during the test phase, the corresponding results are not acceptable. This can be problematic when extended to other languages. On the other hand, vowels have the advantage of being applied to other languages since vowels can be pronounced similarly among different languages. A short recording content may not represent patients’ vocal characteristics correctly. Although subjects are asked to pronounce in their ordinary tone without tension, some are nervous during the actual recording resulting in their recordings being different to their ordinary speech. An advanced voice analysis method, capable of dealing with more natural talking, is required for future research.”

**Editor’s Comments**

This manuscript requires some revisions in order to be acceptable. The authors really need to shorten the manuscript and make it read more smoothly. They may need a native English speaking editor to assist with that. In addition, they need to describe the Sasang constitutional classification more clearly. The abstract should be rewritten to include a clear summary of the study question, methods, results and conclusion. The discussion should also include a paragraph on clinical applications.

The whole manuscript has been revised for better readability and redundant sentences are removed. The entire manuscript has been revised by a native English speaker. We have added an overview subsection to give a clear explanation of the Sasang constitutional classification in page 5. The abstract has been completely rewritten and a paragraph on clinical applications has been added at the 20th line in page 13.
Figure 1 - Overview of the SC classification process

Voice data

Training set

Extract vocal features

Learn a classifier

Evaluate training results

Test set

Extract vocal features

Apply the classifier

Evaluate test results

Training phase

Test phase