Reviewer's report

**Title:** Comparing Classification Methods for Diffuse Reflectance Spectra to Improve Tissue Specific Laser Surgery

**Version:** 1  
**Date:** 5 May 2014

**Reviewer:** Diana Porro-Munoz

**Reviewer's report:**

The paper is an extension of a previous work of some of the authors, where it is proposed to use a supervised approach for discriminating among different types of tissues based on diffuse reflectance spectral data. This is important for overcoming some limitations, thus enabling the use of newly developed technology for oral and maxillofacial surgery. The novelty of this paper lies on extending the previous study to large simulated data sets and compare several classification algorithms, in order to determine if there is a better approach than the previous one, for the application.

The paper in general is clear, well written and the goal is well defined. However, I have some major and minor comments that should be revised by the authors.

**Major Compulsory Revisions**

1. My main concern is that, the main contribution of this paper is the comparative study of the different classification algorithms. In order to do such comparison, they should be equally treated such that the comparison really makes sense, or is fare. However, this is not always the case here. In the case of the Classification Trees, authors chose not to do the optimization of the hyperparameters. In fact, it is shown later that the accuracy of this classifier is among the worst and authors recognize that the reason can be the lack of optimization. This optimization should be done in order to make the comparison results creditable.

2. Moreover, the selection of the optimal hyperparameters is based on just one split of what authors define as the tuning data set. In this way, the selection of the optimal value of the hyperparameter can be biased towards that particular partitioning of your data...that particular training set. You cannot know how stable it is or if it is really the optimal value. This is also the reason for using approaches like cross-validation, a number of random splits, etc...for the optimization of the hyperparameters.

This issue also affects the comparison, as you might haven't really found the optimal value for the hyperparameters of some of the classifiers, thus influencing negatively in their performance. Therefore, I believe this is something authors need to work on. You now have data for that!
3. The description of the PDA classifier is not very clear. Of course you don't have to explain the whole method, that's what the reference is for. Still I think you should improve the current description of the method.

4. More on PDA, this method has a smoothing parameter (lambda), which is used to impose a spatial smoothness constraints on the coefficients. You actually mention it in your description of the method as something important. However, there is no mentioning about the optimization of this parameter in the paper. This is related to my initial comments, so it should also be done.

5. In the experimental section, authors mention the term functional data. I guess you mean that spectra have a functional nature, as they are continuous real-valued functions of, in this case, wavelengths. However, this term is never defined before in the paper. This should be done in order for readers to understand your statements.

Minor Essential Revisions

6. Authors should add references to all the classification methods that are used for the paper e.g. QDA, Classification Trees, Random Forests, Neural Networks. Even if most of them are well known in the pattern recognition and machine learning communities, this might not be the case for the medical community. Moreover, I think you are missing a reference when mentioning previous works on a tissue detection algorithm, in the first part of the Existing work section.

7. Authors should introduce the acronyms for the names of the methods before using them in the beginning of Classification methods section. For QDA and PDA for example, the user has to read a lot of things before reaching the description and knowing exactly to which methods authors are referring to.

8. The explanation authors give about the selection of the k number of neighbors at the end of the description of the k-nearest neighbor method, should be moved to the Hyperparameter selection section. As it is now, some information is repeated.

9. Authors use k for denoting a class (section Simulation of new spectra) and also the parameter of the k-nearest neighbor method. This could bring confusion, so the notation should be changed...preferably for the classes , as k is commonly used to denote the nearest neighbors.

10. It would be nice to see the standard deviation of the 10-fold cross-validation of the original data, in order to see how stable is the behavior of the classifiers throughout the different partitioning of the data.

Discretionary Revisions

11. When describing the data sets...you give a name to them...for example, the large one is the Tunning data set. Thus, to be more clear you should stick to this term throughout the paper.
12. Table 9 and Figure 4 are showing pretty much the same information, maybe you could use just one of them.

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

Quality of written English: Acceptable

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

Declaration of competing interests:

'I declare that I have no competing interests